DECISION MAKING WITH THE OWA OPERATOR IN SPORT MANAGEMENT

José M. Merigó [1] **Anna M. Gil-Lafuente** [1]

[1] Department of Business Administration, University of Barcelona
Av. Diagonal 690, 08034 Barcelona, Spain, {jmerigo, amgil}@ub.edu

Abstract

We analyze the use of the OWA operator in the selection of human resources in sport management. We use different business decision making techniques for doing so. We will consider the use of the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. By using the OWA operator, we can parameterize these decision making techniques from the maximum to the minimum result depending on the interests of the decision maker. We will develop an illustrative example about the decision making process to follow in the selection of a football player for a team.

Keywords: Aggregation operators, OWA operator, Decision making, Sport management.

1 INTRODUCTION

The selection of the most appropriate human resources in sports like football, basketball, etc. represents a fundamental problem for its good development. The enterprise needs to analyze how to select the best player according with their interests. In order to solve this problem, the company has to develop a selection process in which they have to compare the different characteristics of each available candidate found in the market with their ideals. Among the great variety of studies existing in selection, this work will focus on the models developed in [5,9] about selection of human resources, the models developed in [6,11-13] about selection of financial products and the models developed in [7-8] about selection of players in sport management. Note that a very useful survey about different decision making methods can be found in [4].

One problem about these selection indexes is that they are neutral against the attitudinal character of the decision maker. Then, when developing the selection process, we cannot manipulate the results according to the interests of the decision maker. This problem becomes important in situations where we want to under estimate or over estimate the decisions in order to be more or less prudent against the uncertain factors affecting the future. One common method for aggregating the information considering the decision attitude of the decision maker is the ordered weighted averaging (OWA) operator introduced in [14]. Since its appearance, the OWA operator has been studied by different authors such as [1-3,10-13,15-16]. Recently, this problem has been considered in other decision making problems such as financial management and human resource management [12-13].

The aim of the paper consists in developing new selection indexes that include the attitudinal character of the decision maker for the selection of human resources in sport management. These new indexes will consist in combining the Hamming distance, the adequacy coefficient and the index of maximum and minimum level, with the OWA operator because then, the neutrality of the old methods will be changed by the OWA operator. We will introduce in the selection of human resources in sport management, the ordered weighted averaging distance (OWAD) operator [12], the ordered weighted averaging adequacy coefficient (OWAAC) [12] and the ordered weighted averaging index of maximum and minimum level (OWAIMAM) [12]. We will also develop an illustrative example in order to see the usefulness of the new approach. Depending on the aggregation operator used, we will see that the results may be different leading to different decisions.

This paper is organized as follows. In Section 2 we briefly review some basic concepts. In Section 3, we analyze the process to follow in the selection of human resources in sport management. In Section 4 we develop an illustrative example of the problem. We focus on the selection of human resources in a football team. Finally, in Section 5 we summarize the main conclusions of the paper.

2 PRELIMINARIES

2.1. OWA OPERATOR

The OWA operator introduced in [14] provides a parameterized family of aggregation operators which have been used in many applications [1-3,10-16]. In the following, we provide a definition of the OWA operator as introduced by [14].

Definition 1. An OWA operator of dimension n is a mapping f: $R^n \rightarrow R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_i \in [0,1]$. Then:

$$f(a_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} w_j b_j \quad (1)$$

where b_j is the jth largest of the a_i.

From a generalized perspective of the reordering step we can distinguish between the descending OWA (DOWA) and the ascending OWA (AOWA) operator. Note that the weights of this two operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the DOWA and w^*_{n-j+1} the jth weight of the AOWA operator.

By choosing a different manifestation of the weighting vector, we are able to obtain different types of aggregation operators such as the maximum, the minimum, the average and the weighted average [14]. Other families of OWA operators can be found in [11,15].

2.2. BUSINESS DECISION MAKING TECHNIQUES

In the literature, we find a wide range of methods for business decision making [1-9,11-16]. In this paper, we will focus on the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. First, we will consider the formulation when all the weights w_j, are equal. We will use two sets A and B.

Definition 2. A normalized Hamming distance (NHD) of dimension n is a mapping d_H: $R^n \times R^n \rightarrow R$ such that:

$$d_H(A,B) = \left(\frac{1}{n} \sum_{i=1}^{n} | a_i - b_i | \right) \quad (2)$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Definition 3. A normalized adequacy coefficient (NAC) of dimension n is a mapping K: $R^n \times R^n \rightarrow R$ such that:

$$K(P_k \rightarrow P) = \frac{1}{n} \sum_{i=1}^{n} [1 \wedge (1 - \mu_i + \mu_i^{(k)}] \quad (3)$$

where μ_i and μ_i^k are the ith arguments of the sets P and P_k respectively.

Definition 4. A normalized index of maximum and minimum level (NIMAM) of dimension n is a mapping K: $R^n \times R^n \rightarrow R$ such that:

$$\eta(P,P_j) = \frac{1}{u+v} \left[\sum_u \left| \mu_i(u) - \mu_i^{(j)}(u) \right| + \sum_v \left(0 \vee (\mu_i(v) - \mu_i^{(j)}(v)) \right) \right] \quad (4)$$

where μ_i and μ_i^k are the ith arguments of the sets P and P_j, j is the jth alternative considered, and u and v are the characteristics to be considered with the Hamming distance and the adequacy coefficient, respectively.

Sometimes, when normalizing these measures, it is better to give different weights to each individual argument. Then, these measures are known as the weighted Hamming distance (WHD), the weighted adequacy coefficient (WAC) and the weighted index of maximum and minimum level (WIMAM).

Definition 5. A WHD of dimension n is a mapping d_{WH}: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_i \in [0,1]$. Then:

$$d_{WH}(A,B) = \left(\sum_{i=1}^{n} w_i | a_i - b_i | \right) \quad (5)$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Definition 6. A WAC of dimension n is a mapping K: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_i \in [0,1]$. Then:

$$K(P_k \rightarrow P) = \sum_{i=1}^{n} w_i [1 \wedge (1 - \mu_i + \mu_i^{(k)}] \quad (6)$$

where μ_i and μ_i^k are the ith arguments of the sets P and P_k respectively.

Definition 7. A WIMAM of dimension n is a mapping K: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_i \in [0,1]$. Then:

$$\eta(P,P_j) = \frac{\sum_u Z_i(u) \times \left| \mu_i(u) - \mu_i^{(j)}(u) \right| +}{\sum_v Z_i(v) \times \left[0 \vee (\mu_i(v) - \mu_i^{(j)}(v)) \right]} \quad (7)$$

where μ_i and μ_i^k are the ith arguments of the sets P and P_j, j is the jth alternative considered, and u and v are the characteristics to be considered with the Hamming distance and the adequacy coefficient, respectively.

3 SELECTION OF HUMAN RESOURCES WITH THE OWA OPERATOR IN SPORT MANAGEMENT

The reason for using the OWA operators in the selection of human resources in sport management appears because the decision maker wants to take the decision with a certain degree of optimism or pessimism rather than with a neutral position. Due to the fact that the traditional methods used in sport management for this purpose [7-8] are neutral against the attitude of the decision maker, the introduction of the OWA operators in these models may change the neutrality and reflect decisions with different degrees of optimism and pessimism. These techniques can be used in a lot of situations but the general ideas about it are the possibility of under estimate or over estimate the problems.

The process to follow in the selection of human resources in sport management with the OWA operator, is similar to the process developed in [5-9] with the difference that the instruments used will include the OWA operator in the selection process. Note that similar models that use the OWA operator have been developed for other selection problems [12]. The 5 steps to follow are:

Step 1: Analysis and determination of the significant characteristics of the available candidates for the team. Theoretically, it will be represented as: $C = \{C_1, C_2, \ldots, C_i, \ldots, C_n\}$, where C_i is the ith characteristic to consider of the candidate.

Step 2: Fixation of the ideal levels of each significant characteristic in order to form the ideal player. That is:

Table 1: Ideal player

	C_1	C_2	…	C_i	…	C_n
P =	μ_1	μ_2	…	μ_i	…	μ_n

where P is the ideal player expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i \in [0,1]$; $i = 1, 2, \ldots, n$, is the valuation between 0 and 1 for the ith characteristic.

Step 3: Fixation of the real level of each characteristic for all the different candidates considered. That is:

Table 2: Available players

	C_1	C_2	…	C_i	…	C_n
P_k =	$\mu_1^{(k)}$	$\mu_2^{(k)}$	…	$\mu_i^{(k)}$	…	$\mu_n^{(k)}$

with $k = 1, 2, \ldots, m$; where P_k is the kth candidate expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i^{(k)} \in [0,1]$; $i = 1, \ldots, n$, is the valuation between 0 and 1 for the ith characteristic of the kth candidate.

Step 4: Comparison between the ideal player and the different candidates considered, and determination of the level of removal using the OWA operator. That is, changing the neutrality of the results to over estimate or under estimate them. In this step, the objective is to express numerically the removal between the ideal player and the different candidates considered. For doing this, it can be used the different available selection indexes such as the Hamming distance, the adequacy coefficient, the index of maximum and minimum level, etc. [5-9,12-13].

Step 5: Adoption of decisions according to the results found in the previous steps. Finally, we should take the decision about which person select. Obviously, our decision will consist in choosing the candidate with the best results according to the index used.

3.1. USING THE OWA OPERATOR AND THE HAMMING DISTANCE

When making the comparison between the ideal player and the available candidates in *Step* 4, it is possible to use different techniques. First, we will consider the combination between the OWA operator and the Hamming distance that it is known as the OWAD operator. It can be defined as follows.

Definition 8. An OWAD operator of dimension n, is a mapping $OWAD: R^n \times R^n \rightarrow R$ that has an associated weighting vector W, with the sum of the weights equal to 1 and $w_j \in [0,1]$ such that:

$$OWAD(P,P_k) = \sum_{j=1}^{n} w_j D_j \quad (8)$$

where D_j represents the jth largest of the $|\mu_i - \mu_i^{(k)}|$, and $k = 1,2,\ldots,m$.

Note that from a generalized perspective of the reordering step it is possible to distinguish between ascending and descending orders. The weights of these operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the

ascending OWAD (AOWAD) and w^*_{n-j+1} the jth weight of the descending OWAD (DOWAD) operator.

As we can see, the OWAD operator is commutative, monotonic, bounded and idempotent. By choosing a different manifestation of the weighting vector, we are able to obtain different types of aggregation operators. For example, the maximum distance is found when $w_1 = 1$ and $w_j = 0$ for all $j \neq 1$. The minimum is found when $w_n = 1$ and $w_j = 0$ for all $j \neq n$. The NHD is obtained when $w_j = 1/n$ for all j. The WHD is found when the ordered position of i is the same than the ordered position of j. Note that in the case of tie in the final result, it could be used in the decision the second best or worst result, and so on.

3.2. USING THE OWA OPERATOR AND THE ADEQUACY COEFFICIENT

In this Section, we introduce the use of the OWA operator in the selection of human resources in sport management with the adequacy coefficient. We will call it the OWAAC. It can be defined as follows.

Definition 9. An OWAAC operator of dimension n, is a mapping $OWAAC$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W, with $w_j \in [0,1]$ and the sum of the weights is equal to 1, such that:

$$OWAAC(P_k \rightarrow P) = \sum_{j=1}^{n} w_j K_j \qquad (9)$$

where K_j represents the jth largest of the $[1 \wedge (1 - \mu_i + \mu_i^{(k)})]$, and $k = 1,2,\ldots,m$.

Note that the final result will be a number between [0,1], being the maximum possible result 1. Note also that from a generalized perspective of the reordering step we can distinguish between descending and ascending orders. The weights of these operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the descending OWAAC (DOWAAC) and w^*_{n-j+1} the jth weight of the ascending OWAAC (AOWAAC) operator. Note also that the OWAAC operator is commutative, monotonic, bounded and idempotent.

Note that it is possible to use different types of aggregation operators such as the maximum, the minimum, the NAC, the WAC, the step-OWAAC, the S-OWAAC, the centered-OWAAC, etc.

Analogously to the OWAAC operator, we can suggest an equivalent removal index that it is a dual of the OWAAC because $Q(P_k \rightarrow P) = 1 - K(P_k \rightarrow P)$. We will call it the ordered weighted averaging dual adequacy coefficient (OWADAC).

Definition 10. An OWADAC operator of dimension n, is a mapping $OWADAC$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W, with $w_j \in [0,1]$ and the sum of the weights is equal to 1, then:

$$OWADAC(P_k \rightarrow P) = \sum_{j=1}^{n} w_j Q_j \qquad (10)$$

where Q_j represents the jth largest of the $[0 \vee (\mu_i - \mu_i^{(k)})]$, and $k = 1,2,\ldots,m$.

The final result will be a number between [0,1]. Note that in this case we usually select the lowest value as the best result. In this case, we can also distinguish between the descending OWADAC (DOWADAC) and the ascending OWADAC (AOWADAC) operator and it is also possible to obtain different families of aggregation operators.

Another interesting issue to consider is the unification point in the selection of human resources in sport management. As it has been explained in [13], the unification point appears when the results obtained in the Hamming distance are the same than the results obtained in the adequacy coefficient. In the new methods suggested in this paper, we also find the unification point when the OWAD and the OWAAC accomplish the theorems explained in [13]. Note that it is possible to find a total unification point or a partial unification point [13]. In the following, we briefly show the main theorem when using the OWA operator.

Theorem 1. Assume $OWAD(P,P_k)$ is the selection of human resources in sport management with the OWAD operator and $OWADAC(P_k \rightarrow P)$ the selection of human resources in sport management with the OWADAC operator. If $\mu_i \geq \mu_i^{(k)}$ for all i, then:

$$OWAD(P,P_k) = OWADAC(P_k \rightarrow P) \qquad (11)$$

Proof. Let

$$OWAD(P,P_k) = \sum_{j=1}^{n} w_j \, | \mu_i - \mu_i^{(k)} | \qquad \text{and}$$

$$OWADAC(P_k \rightarrow P) = \sum_{j=1}^{n} w_j [0 \vee (\mu_i - \mu_i^{(k)})]$$

Since $\mu_i \geq \mu_i^{(k)}$ for all i, $[0 \vee (\mu_i - \mu_i^{(k)})] = (\mu_i - \mu_i^{(k)})$ for all i, then

$$OWADAC(P_k \rightarrow P) = \sum_{j=1}^{n} w_j (\mu_i - \mu_i^{(k)}) = OWAD(P,P_k) \quad \blacksquare$$

Analysing this theorem, we could generalize it for all the players - alternatives considered in the decision problem.

The theorem that explains this generalization is very similar to theorem (1) with the difference that now we consider all the characteristics *i* and all the players *k*.

3.3. USING THE OWA OPERATOR AND THE INDEX OF MAXIMUM AND MINIMUM LEVEL

In this Section, we develop an index for the selection of human resources in sport management that uses the OWA operator in the index of maximum and minimum level. We will call it the OWAIMAM. It is defined as follows.

Definition 11. An OWAIMAM operator of dimension *n*, is a mapping *OWAIMAM*: $R^n \times R^n \to R$ that has an associated weighting vector *W*, with $w_j \in [0,1]$ and the sum of the weights is equal to 1, such that:

$$S(P_k \to P) = \sum_{j=1}^{n} w_j S_j \tag{12}$$

where S_j represents the *j*th largest of all the $|\mu_i - \mu_i^{(k)}|$ and the $[0 \vee (\mu_i - \mu_i^{(k)})]$; with $k = 1,2,\ldots,m$.

Note that from a generalized perspective of the reordering step we could distinguish between descending (DOWAIMAM) and ascending (AOWAIMAM) orders.

In this case, it is also possible to use different types of aggregation operators such as the maximum, the minimum, the NIMAM and the WIMAM.

Analogously to the OWAIMAM operator, we can suggest an equivalent removal index that it is a dual of the OWAIMAM because $R(P_k \to P) = 1 - S(P_k \to P)$. We will call it the ordered weighted averaging dual index of maximum and minimum level (OWADIMAM).

Another interesting issue to consider is the unification point in the selection of human resources in sport management for the index of maximum and minimum level. As it has been explained in [13], in these situations, the index of maximum and minimum level becomes the Hamming distance. Note that it is possible to find a total unification point and a partial unification point [13].

4 ILLUSTRATIVE EXAMPLE

We will develop a decision making problem about selection of players in a football team. We will assume that a football team is looking for a player for a specific position that they need to cover. Obviously, in the market there are a lot of choices but not all of them can be considered as a real choice because his team does not want to sell it. Then, the experts of the team need to consider all the alternatives and develop a first general selection in order to analyze in detail the real alternatives that can be acquired by the team.

Step 1: Analysis and determination of the significant characteristics for the team. Assume that a football team wants to select a player for a forward position and it has 4 candidates P_1, P_2, P_3, P_4, with different characteristics. It is considered for each characteristic a property.

Step 2: Fixation of the ideal level for each significant characteristic. It is defined the ideal player as:

Table 3: Characteristics of the ideal player

	C_1	C_2	C_3	C_4	C_5
P^* =	0.8	0.6	0.9	1	0.8

Step 3: Fixation of the real level of each characteristic for all the different candidates considered. For each of these characteristics, it is found the following information:

Table 4: Available candidates

	C_1	C_2	C_3	C_4	C_5
P_1	0.6	0.7	0.5	0.6	0.8
P_2	0.5	0.7	0.6	0.8	0.5
P_3	0.9	0.5	0.4	0.8	0.7
P_4	0.4	0.6	0.7	0.8	0.8

Step 4: Comparison between the ideal worker and the different candidates considered using the OWA operator. We will consider the NHD, the WHD, the OWAD and the AOWAD operator. In this example, we assume that the company decides to use the following weighting vector: *W* = (0'1, 0'1, 0'2, 0'3, 0'3).

If we elaborate the selection process with the Hamming distance, we will get the results shown in Table 5.

Table 5: Aggregated results with the Hamming distance

	NHD	WHD	OWAD	AOWAD
P_1	0.22	0.23	0.15	0.29
P_2	0.24	0.25	0.21	0.27
P_3	0.2	0.21	0.15	0.25
P_4	0.16	0.14	0.1	0.22

In this case, our decision will consist in selecting the candidate with the smallest distance. Then, we will select P_4 as it gives us the lowest distance in the four cases.

If we develop the selection process with the adequacy coefficient, we will get the following results shown in Table 6. Note that in this case we select the alternative with the highest result being P_4 the optimal player for the team.

Table 6: Aggregated results with the adequacy coefficient

	NAC	WAC	OWAAC	AOWAAC
P_1	0.8	0.78	0.72	0.88
P_2	0.78	0.76	0.74	0.82
P_3	0.82	0.8	0.76	0.88
P_4	0.84	0.86	0.78	0.9

Analogously to this index, we can obtain its equivalent removal index. In an abbreviated form, this index can be obtained by using $Q(P_k \rightarrow P) = 1 - K(P_k \rightarrow P)$. The results are shown in Table 7.

Table 7: Results with the dual adequacy coefficient

	NDAC	WDAC	OWADAC	AOWADAC
P_1	0.2	0.22	0.28	0.12
P_2	0.22	0.24	0.26	0.18
P_3	0.18	0.2	0.24	0.12
P_4	0.16	0.14	0.22	0.1

Finally, if we use the IMAM in the selection process as a combination of the Hamming distance and the adequacy coefficient, we will get the following. Note that in this example, we will assume that the characteristics C_1 and C_2 have to be treated with the adequacy coefficient and the other three characteristics with the Hamming distance. Its resolution is shown in Table 8.

Table 8: Aggregated results with the IMAM

	NIMAM	WIMAM	OWAIM	AOWAIM
P_1	0.22	0.23	0.15	0.29
P_2	0.24	0.25	0.21	0.27
P_3	0.2	0.21	0.15	0.25
P_4	0.16	0.14	0.1	0.22

Then, our decision will consist in select P_4 because it is the candidate with the smallest removal to the ideal. Note that in this example, the results obtained by using the Hamming distance and the IMAM are equal. The reason is that this example fulfils the conditions to enter in a situation of unification point as explained in [13]. Then, in these situations, the Hamming distance and the IMAM have the same formulation.

Analogously to this index, we can obtain its equivalent approximation index. In an abbreviated form, this index can be obtained by using $R(P_k \rightarrow P) = 1 - S(P_k \rightarrow P)$. The results are shown in Table 9.

Table 9: Aggregated results with the dual IMAM

	NDIM	WDIM	OWADIM	AOWADIM
P_1	0.78	0.77	0.85	0.71
P_2	0.76	0.75	0.79	0.73
P_3	0.8	0.79	0.85	0.75
P_4	0.84	0.86	0.9	0.78

As we can see, depending on the particular type of aggregation operator used, the results may lead to different decisions.

5 CONCLUSIONS

We have analysed the selection of human resources in sport management. We have used a new methodology that uses the OWA operator with different measures frequently used in business decision making such as the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. We have seen that the main advantage of using the OWA operator in this type of problems is the possibility of under or over estimate the results. Then, the decision maker is able to take decisions considering its attitudinal character.

In future research, we expect to develop further extensions of these problems by adding new characteristics and applying it to other decision making problems.

Acknowledgements

We would like to thank the anonymous referees for their valuable comments that have improved the quality of the paper.

References

[1] G. Beliakov, A. Pradera, T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer-Verlag, New York, 2007.

[2] H. Bustince, F. Herrera, J. Montero, Fuzzy Sets and their Extensions: Intelligent Systems from Decision Making to Data Mining, Web Intelligence and Computer Vision, Springer-Verlag, 2008.

[3] T. Calvo, G. Mayor, R. Mesiar. *Aggregation Operators: New Trends and Applications*, Physica-Verlag, New York, 2002.

[4] J.R. Figueira, S. Greco, M. Ehrgott, *Multiple criteria decision analysis: State of the art surveys*, Springer, Boston, 2005.

[5] J. Gil-Aluja. *The interactive management of human resources in uncertainty*, Kluwer Academic Publishers, Dordrecht, 1998.

[6] A.M. Gil-Lafuente. *Fuzzy logic in financial analysis*, Springer, Berlin, 2005.

[7] J. Gil-Lafuente. El "índice del máximo y mínimo nivel" en la optimización del fichaje de un deportista, In *X Congreso Internacional de la Asociación Europea de Dirección y Economía de la Empresa* (AEDEM), Reggio Calabria, Italy, pp. 439-443, 2001.

[8] J. Gil-Lafuente, *Algoritmos para la excelencia: Claves para el éxito en la gestión deportiva*, Ed. Milladoiro, Vigo, 2002.

[9] A. Kaufmann, J. Gil-Aluja, *Introducción de la teoría de los subconjuntos borrosos a la gestión de las empresas*, Ed. Milladoiro, Santiago de Compostela, 1986.

[10] N. Karayiannis. Soft Learning Vector Quantization and Clustering Algorithms Based on Ordered Weighted Aggregation Operators, *IEEE Transactions on Neural Networks* 11, Pág. 1093-1105, 2000.

[11] J.M. Merigó. *New extensions to the OWA operators and their application in business decision making*, Unpublished thesis (in Spanish), Department of Business Administration, University of Barcelona, 2007.

[12] J.M. Merigó, A.M. Gil-Lafuente, Using the OWA operators in the selection of financial products, In *Proceedings of the 41st CLADEA Congress*, Montpellier, France, CD-ROM, 2006.

[13] J.M. Merigó, A.M. Gil-Lafuente, Unification point in methods for the selection of financial products, *Fuzzy Economic Review* 13, Pág. 35-50, 2007.

[14] R.R. Yager. On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making, *IEEE Transactions on Systems, Man and Cybernetics B* 18, Pág. 183-190, 1988.

[15] R.R. Yager. Families of OWA operators, *Fuzzy Sets and Systems* 59, Pág. 125-148, 1993.

[16] R.R. Yager, J. Kacprzyk. *The Ordered Weighted Averaging Operators: Theory and Applications*, Kluwer Academic Publishers, Norwell, 1997.

Some Notions on Lever Aggregation Operators

Miguel A. Ballester[1] **Tomasa Calvo**[2]

[1] Universitat Autònoma de Barcelona, MiguelAngel.ballester@uab.es
[2] Universidad de Alcalá, tomasa.calvo@uah.es

Abstract

We make use of the *physical concept* of *lever* to generalize well-known operators based on weights. All these operators can be understood as levers in equilibrium, with weights (masses) located at the points in $[0, 1]$ that opinions represent.

Keywords: Aggregation operators, lever operator and function, weighted aggregation operators.

1 INTRODUCTION

Aggregation of several input values into a single output value is an indispensable tool in many disciplines and applications such as decision making [17], pattern recognition, expert and decision support systems, information retrieval, etc [5]. Aggregation operators are mathematical objets that perform pieces of information coming from different sources (information fusion). The arithmetic mean is probably the most commonly used aggregation operator.

A particular class of aggregation operators are the weighted aggregation operators. Most of these operators are defined with the help of an associated vector of weights. These operators have been used especially in fuzzy decision making [1, 17], where the weights are used to represent the relative importance that the decision maker attaches to different decision criteria. The weighted generalized mean operator has been used in many fields, and it has been studied in the context of fuzzy set aggregation see e.g., [13].

The aim of this work is to capture and generalize the basic idea underlaying all these weighting aggregation operators. This idea simply reproduces the *physical concept* of a *lever*. A lever consists merely of a rigid object that is used with a pivot (or fulcrum). Basically, the mechanical aspects of a lever can be described by saying that the amount of work done is given by *force times distance*. Relevant to us is also the concept of *mechanical equilibrium*, i.e., a system is said to be in mechanical equilibrium when the sum of the forces, and torque, on each particle of the system is zero. In the case of the lever, this definition can be easily interpreted. The forces that exert the masses to the left of the fulcrum, multiplied by their distance to the fulcrum have to equalize the same variable for the masses to the right of the fulcrum. This is basically the idea of mechanical equilibrium in a lever.

Describing weighting averaging operators as *levers* will prove to be useful not only for describing well-known weighting aggregation operators, but also for exploring new families. Some related approaches have been considered by several authors (see [7, 3, 4, 16, 8]) and also applied by [11, 12].

The following section presents basic definitions and notation for the case of weighted arithmetic means. We describe the intuition of our work by analyzing how weighted arithmetic means can be understood as levers, introducing the notions of *lever function* and *lever mean*, key of our analysis. After that, we present the main results on lever operators and we finish with a brief comment on our future work.

2 WEIGHTED LEVER OPERATORS

The main purpose of this paper is to formally study a general class of operators, *weighted lever operators*, that encompasses most of the previously presented aggregation operators. The generalization arises from the consideration of the physical principle that originates most averaging operators using weights, i.e., the principle of moments. Some notation will be useful in subsequent analysis. Given $a \in [0, 1]$, we denote

$L(a) = \{i \in N \text{ such that } x_i < a\}$ and $R(a) = \{i \in N \text{ such that } x_i > a\}$.

Consider the real line in $[0,1]$ to be a rigid object. Take the vector of inputs or opinions, $x \in [0,1]^n$. We can understand the elements of this vector as the set of points over which masses (weights) will be located. Intuitively, an averaging (weighting) n-ary operator should select, for any vector of opinions, the pivot point, i.e., the point for which the lever system is in mechanical equilibrium. The following proposition formally analyzes this property for weighted arithmetic means. [1]

When doing so, we notice that this family is using a very particular distance function over the set $[0,1]$ and we present weighted lever operators as an intuitive generalization.

Proposition 1 For all weighting vector w and for all $x \in [0,1]^n$, the two following equivalent statements hold:

1. $\sum_{i=1}^{n} w_i h(\mathbf{M_w}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}$ where $h(\mathbf{M_w}(\mathbf{x}), \mathbf{x_i}) = \mathbf{x_i} - \mathbf{M_w}(\mathbf{x})$.
2. $\sum_{i \in L} w_i(\mathbf{M_w}(\mathbf{x}) - \mathbf{x_i}) = \sum_{\mathbf{i \in R}} \mathbf{w_i}(\mathbf{x_i} - \mathbf{M_w}(\mathbf{x}))$. [2]

Proof: Consider a weighting vector w and the associated mean $\mathbf{M_w}$. For any point $x \in [0,1]^n$, we have:

$$\sum_{i=1}^{n} w_i \mathbf{M_w}(\mathbf{x}) = (\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{w_i}) \mathbf{M_w}(\mathbf{x}) = \mathbf{M_w}(\mathbf{x}) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{w_i x_i}.$$

Thus clearly $\sum_{i=1}^{n} w_i(x_i - \mathbf{M_w}(\mathbf{x})) = \mathbf{0}$ holds and the first condition is proved.

For 2, simply notice that whenever $x_i = \mathbf{M_w}(\mathbf{x})$, it is $w_i(x_i - \mathbf{M_w}(\mathbf{x})) = \mathbf{0}$ and therefore, the previous equation can be expressed as

$$\sum_{i \in L} w_i(x_i - \mathbf{M_w}(\mathbf{x})) + \sum_{\mathbf{i \in R}} \mathbf{w_i}(\mathbf{x_i} - \mathbf{M_w}(\mathbf{x})) = \mathbf{0},$$

from which second statement directly follows. ∎

These dual characteristics of weighted arithmetic means simply represent the following ideas; aggregation of opinions must be done in the point of $[0,1]$ in which fulcrum is located obtaining a mechanical equilibrium (considering the usual distance notion and a convention about the sign of the exerted work). With this idea in mind, the aim of this paper is to present a very general class of aggregation operators, lever operators, that intuitively work following the principle of levers and mechanical equilibrium. These operators merely generalize the notion of distance.

For the purposes of aggregation operators, any function using the intuitive idea of *the farther, the more work* might be useful, and not only the usual distance over the real numbers as in the case of $\mathbf{M_w}$.

Consider a real two-variable function $h : [0,1]^2 \to (-\infty, +\infty)$. The value $h(a,b)$ is intended to represent the value for which a weight located in b has to be multiplied to represent the work that exerts over the lever with fulcrum at a. Now we present minimal properties over the function h to intuitively represent a useful generalization.

Definition 1. A *lever function* is a mapping $h : [0,1]^2 \to (-\infty, +\infty)$ satisfying:

1. $h(a,a)$ is equal to zero for all $a \in [0,1]$,
2. h is non-decreasing in the second component,
3. h is non-increasing in the first component.

The first condition simply states that masses exert no work over the fulcrum. To interpret condition 2, given the adopted convention, as far as we make the weight to be farther to the right from the fulcrum, the value of the exerted work should be higher. Obviously, if we make the weight to be farther to the left, the value will be higher in absolute terms. This function should be non-increasing in the position of the fulcrum, with a similar reasoning. Lever functions allow us to define the concept of weighted lever operator that extends that of weighted arithmetic mean.

Definition 2. Let $w = (w_1, \ldots, w_n) \in [0,1]^n$ with $\sum_{i=1}^{n} w_i = 1$ and a lever function h. An n-ary operator $\mathbf{L_w^h} : \mathbf{[0,1]^n} \to \mathbf{[0,1]}$ is called a weighted lever operator (for short WLO) associated with w and h if, for all $x \in [0,1]^n$:

$$\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}. \quad (1)$$

In other words, a weighted lever operator should select, given a collection of weights and a lever function, a point of mechanical equilibrium.

[1] We simply represent the classical notion of weighted arithmetic mean based on weights w as $\mathbf{M_w}(\mathbf{x})$.

[2] To simplify the exposition we use notation L and R instead of $L(\mathbf{M_w}(\mathbf{x}))$ and $R(\mathbf{M_w}(\mathbf{x}))$, respectively, when confusion is not possible.

Notice that we have started the analysis with the most general possible view and thus, we have not assumed any symmetry assumption on h. Left and Right might be understood as different concepts. Notice also that, if h represents the usual distance, i.e. $h = d$, over the real numbers, $\mathbf{M_w} = \mathbf{L_w^d}$ for the case of weighted lever operators.

3 FIRST RESULTS ON WLO

We start by discussing a normalization of lever functions. Given the conditions, $h(a, b)$ achieves its maximal value at $(0, 1)$ and its minimal value at $(1, 0)$, allowing to prove the following result.

Proposition 2. Let w be a weighting vector and let h, h' be lever functions, and $h'(a, b) = kh(a, b)$ with $k \in \mathcal{R}$. If an n-ary operator is a weighted lever operator associated to w and h, it is also a weighted lever operator associated with w and h'.

Proof: The proof is straightforward and thus omitted. ∎

From the previous result we can assure that the lever function is unique up to a multiplicative constant. Now it is obvious that we can simply work with lever functions defined in $[-1, 1]$. Given the intuitive monotonicity properties, the maximum and minimum values of h are achieved at $(0, 1)$ and $(1, 0)$ points. We exclude the case in which both of them are zero since it is trivial and assume that at least one of them is different from zero. Consider the largest of the absolute values of h at these two points, and denote it by r. We can normalize the lever function to $\frac{1}{r}h$ that satisfies the extra requirement that either $h(0, 1) = 1$ or $h(1, 0) = -1$. All the intuitions and results of this paper can be extended to any lever function without normalization.

Our first result describes the possibility of constructing n-ary operators from any given lever function and weighting vector.

Theorem 1. Let $h : [0, 1]^2 \to [-1, 1]$ be a lever function, continuous function in the first variable (call it, continuous lever function). Then, for any collection of weights w there exists at least one monotone and idempotent n-ary weighted lever operator associated with w and h.

Proof: Let $h : [0, 1]^2 \to [-1, 1]$ be a continuous lever function and consider a collection of weights w. Fix a point $x \in [0, 1]^n$. Consider the following optimization problem:

$$\min_{a \in [0,1]} |a - x_1| \text{ subject to}$$

$$\sum_{i=1}^{n} w_i h(a, x_i) = 0.$$

We first show that this problem has a solution. Since h is continuous in the first variable the function $g : [0, 1] \to [0, 1]$ defined by $g(a) = \sum_{i=1}^{n} w_i h(a, x_i)$ is also continuous. Given that h is a lever function, h is non-decreasing in the second component and therefore $h(0, x_i) \geq h(0, 0) = 0$ for all $i = 1, \ldots, n$. This implies:

$$g(0) = \sum_{i=1}^{n} w_i h(0, x_i) \geq 0.$$

Similarly $h(1, x_i) \leq h(1, 1) = 0$ for all $i \in N$. This implies:

$$g(1) = \sum_{i=1}^{n} w_i h(1, x_i) \leq 0.$$

This shows that the set $\{a \in [0, 1] \text{ such that } g(a) = 0\}$ is nonempty. Therefore, continuity of the lever function also guarantees that the minimization problem has a unique solution.

Let us define $\mathbf{L_w^h}(\mathbf{x})$ as the point that solves the optimization problem, for all $x \in [0, 1]^n$. It is straightforward to see that this mapping is a weighted lever operator induced by w and h.

To check monotonicity, suppose that $x \leq y$. For all $i \in N$, it is $x_i \leq y_i$ and, being h non-decreasing in its second component, $\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}$ implies $\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{x}), \mathbf{y_i}) \geq \mathbf{0}$. Consider first that $\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{x}), \mathbf{y_i}) > \mathbf{0}$. Clearly, h non-increasing in its first argument implies that $\sum_{i=1}^{n} w_i h(\cdot, y_i)$ is non-increasing. Therefore the solution to the optimization problem has to be $\mathbf{L_w^h}(\mathbf{x}) < \mathbf{L_w^h}(\mathbf{y})$ as desired.

Consider then that $\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{x}), \mathbf{y_i}) = \mathbf{0}$. Suppose, by way of contradiction that $\mathbf{L_w^h}(\mathbf{x}) > \mathbf{L_w^h}(\mathbf{y})$. We prove now that $\mathbf{L_w^h}(\mathbf{x}) > \mathbf{y_1}$. Suppose otherwise. Then, it must be $y_1 \geq \mathbf{L_w^h}(\mathbf{x}) > \mathbf{L_w^h}(\mathbf{y})$. But this contradicts the fact that $\mathbf{L_w^h}(\mathbf{y})$ solves the optimization problem for y. Then, it must be $\mathbf{L_w^h}(\mathbf{x}) > \mathbf{y_1} \geq \mathbf{x_1}$. This implies also that $|\mathbf{L_w^h}(\mathbf{x}) - \mathbf{x_1}| > |\mathbf{L_w^h}(\mathbf{y}) - \mathbf{x_1}|$. As long as $\mathbf{L_w^h}(\mathbf{y})$ does not solve the optimization problem for x, it has to be $\sum_{i=1}^{n} w_i h(\mathbf{L_w^h}(\mathbf{y}), \mathbf{x_i}) \neq \mathbf{0}$. Since

$\sum_{i=1}^{n} w_i h(\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}$ and $\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}) > \mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{y})$, properties of h guarantee that $\sum_{i=1}^{n} w_i h(\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{y}), \mathbf{x_i}) > \mathbf{0}$ and also $\sum_{i=1}^{n} w_i h(\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{y}), \mathbf{y_i}) > \mathbf{0}$ which is a contradiction. Then it cannot be $\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}) > \mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{y})$ and monotonicity is proved.

To check idempotency notice that whenever $x = (t, \ldots, t)$, the fact that $h(t,t) = 0$ guarantees that t belongs to the set $\{a \in [0,1] \text{ such that } g(a) = 0\}$. Given that $x_1 = t$, the solution to the optimization problem is $\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}) = \mathbf{t}$, as desired. ∎

A similar result can be shown when we replaced the lever function for a deviation function (see [3]), but the proof for deviation functions is simpler due to stronger assumptions on continuity and strict monotonicity.

The previous proposition shows how a very large family of functions h can be used to define weighted lever n-ary operators. However, given a vector of weights, the mapping from continuous lever functions to weighted lever operator is not into. There may exist several different operators associated with the same Lever Function. This lack of uniqueness would make the concept of weighted lever less interesting. In the following proposition, we show how a very weak property on lever functions avoids this problem, thus generating an equivalence between lever functions and WLOs (for a given collection of weights).

Theorem 2. Let $h : [0,1]^2 \to [-1,1]$ be a continuous lever function, strictly decreasing in the first component (from now on, SC lever function). Then, for any collection of weights w there exists a unique monotone and idempotent n-ary WLO associated with w and h.

Proof: The previous proposition guarantees existence. For uniqueness, suppose there are two different WLOs, $\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}$ and $\hat{\mathbf{L}}^{\mathbf{h}}_{\mathbf{w}}$ associated with w and h. Without loss of generality, there exists an element $x \in [0,1]^n$ such that $\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}) < \hat{\mathbf{L}}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x})$. By definition of the weighted lever operator, it must be $\sum_{i=1}^{n} w_i h(\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}$. Since h is an SC lever function, it is immediate to see that $\sum_{i=1}^{n} w_i h(\hat{\mathbf{L}}^{\mathbf{h}}_{\mathbf{w}}(\mathbf{x}), \mathbf{x_i}) > \mathbf{0}$ thus contradicting the fact that $\hat{\mathbf{L}}^{\mathbf{h}}_{\mathbf{w}}$ is a weighted lever associated with w and h. ∎

Any weighted lever operator is thus generated by selecting a vector of weights w and a lever function h. This allows us to aggregate opinions in much more general ways that classical methods provide. It is worthwhile to mention however that describing the weighted lever operator implies to solve a functional equation $\sum_{i=1}^{n} w_i h(\mathbf{L}^{\mathbf{h}}_{\mathbf{w}}, \mathbf{x_i}) = \mathbf{0}$. Looking for closed form solutions to this problem might not be easy for several lever functions.

In some cases, imposing additional properties to the lever function might help, thus providing us with well-known classes of n-ary operators. Suppose for instance that h is additive separable, using an order automorphism $\varphi : [0,1] \to [0,1]$, i.e., $h(a,b) = \varphi(b) - \varphi(a)$. In this case, a closed form solution exists and corresponds with a well known class of operators.

Proposition 3. Define, for any order automorphism on $[0,1], \varphi$, the function $h^{\varphi} : [0,1]^2 \to [-1,1]$ by $h^{\varphi}(a,b) = \varphi(b) - \varphi(a)$. Then, for any weighting vector w it is $\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}} = \mathbf{QM}^{\varphi}_{\mathbf{w}}$

Proof: Consider a collection of weights w and an order automorphism on $[0,1], \varphi$. We first note that it is straightforward to see that h^{φ} is a SC lever function. Hence $\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}$ is uniquely defined. Consider any element $x \in [0,1]^n$. The following condition must hold:

$$\sum_{i=1}^{n} w_i h^{\varphi}(\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}.$$

Given the structure of h^{φ}, we can express the previous condition as:

$\sum_{i=1}^{n} w_i(\varphi(x_i) - \varphi(\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x})) = \mathbf{0}$, or equivalently

$\sum_{i=1}^{n} w_i \varphi(x_i) = \sum_{i=1}^{n} w_i \varphi(\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x})) =$

$(\sum_{i=1}^{n} w_i)\varphi(\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x})) = \varphi(\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x}))$.

Since φ is strictly increasing and continuous, φ^{-1} is well defined and thus $\mathbf{L}^{\mathbf{h}^{\varphi}}_{\mathbf{w}}(\mathbf{x}) = \varphi^{-1}(\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{w_i}\varphi(\mathbf{x_i})) = \mathbf{QM}^{\varphi}_{\mathbf{w}}(\mathbf{x})$ as desired. ∎

A kind of invariance might also be interesting. For instance, deviations to the right or the left of the fulcrum should be equally valued in absolute terms, as in principle, the sign could be a matter of convention. Also, the lever function might depend only on the value of the *jump* between the fulcrum and the point in which the mass is located, independently of the position of the fulcrum. Work is exerted equally independently of the place in which the lever is located.

Definition 3. A Lever Function h is said to be an Invariant Lever Function if for all $a, b, c, d \in [0,1]$ the following conditions hold:

1. $h(a,b) = -h(b,a)$

2. $a - b = c - d \Rightarrow h(a, b) = h(c, d)$

The following proposition makes explicit the structure of *invariant lever functions*:

Proposition 4. For any SC-Lever Function h, the following two statements are equivalent:

1. h is an Invariant Lever Function.
2. There exists an order automorphism $\varphi : [0, 1] \longrightarrow [0, 1]$ such that, for all $x, y \in [0, 1]$:

$$h(x, y) = \begin{cases} \varphi(y - x), & \text{if } x \leq y \\ -\varphi(x - y), & \text{otherwise.} \end{cases}$$

Proof: The proof is straightforward and thus omitted. ■

A weighted lever operator associated with an invariant SC lever function (and a weighting vector w) will be denoted by $\mathbf{L}^{\varphi}_{\mathbf{w}}$. This abuse of notation aims to describe the bijective relationship between invariant SC lever functions and order automorphisms. The lever condition in the case of Invariant SC lever functions is:

$$\sum_{i=1}^{n} w_i h(\mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}), \mathbf{x_i}) = \mathbf{0}, \text{ or equivalently,}$$

$$\sum_{i \in L} w_i \varphi(\mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}) - \mathbf{x_i}) = \sum_{\mathbf{i} \in \mathbf{R}} \mathbf{w_i} \varphi(\mathbf{x_i} - \mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x})).$$

A prominent case might be to take the root-power automorphisms,

$$\varphi(t) = t^k, \text{ with } k > 0,$$

and hence the latter expression is now:

$$\sum_{i \in L} w_i (\mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}) - \mathbf{x_i})^{\mathbf{k}} = \sum_{\mathbf{i} \in \mathbf{R}} \mathbf{w_i} (\mathbf{x_i} - \mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}))^{\mathbf{k}}.$$

Let us analyze some of the levers within this family.

- With $k = 1$ we recover the weighted arithmetic means.[3]

[3] In other words, weighted average means are weighted lever operators associated with additive separable invariant SC lever functions.

- With $k = 2$ we have the equation:

$$\sum_{i \in L} w_i (\mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}) - \mathbf{x_i})^{\mathbf{2}} = \sum_{\mathbf{i} \in \mathbf{R}} \mathbf{w_i} (\mathbf{x_i} - \mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}))^{\mathbf{2}}.$$

We show through a very simple algorithm how to calculate the weighted lever operator for this particular case.

Step 1. Take $x \in [0, 1]^n$ and consider $x \nearrow$.

Step 2. For $j = 1, \ldots, n$, evaluate

$$q(i) = \sum_{i \in R(x^{(j)})} w_i (x_i - x^{(j)})^2 - \sum_{i \in L(x^{(j)})} w_i (x^{(j)} - x_i)^2.$$

Step 3. Consider the first value $k \in N \setminus \{n\}$ such that $q(k) \geq 0$ and $q(k + 1) \leq 0$. This value must exist given that weighted lever operators are balanced, and therefore it must be $q(1) \geq 0$ and $q(n) \leq 0$.

Step 4. If $q(k) = 0$ then $x^{(k)}$ is the solution. If $q(k + 1) = 0$ then $x^{(k+1)}$ is the solution. Otherwise, it is $q(k) > 0 > q(k + 1)$ and the solution is in $(x^{(k)}, x^{(k+1)})$.

Step 5. The solution p has to satisfy

$$\sum_{i \in L(x^{(k+1)})} w_i (p - x^{(i)})^2 - \sum_{i \in R(x^{(k)})} w_i (x^{(i)} - p)^2 = 0.$$

Denoting:

1. $K = \sum_{i \in L(x^{(k+1)})} w_i$, representing the sum of weights in the left part,
2. $\Delta = \sum_{i \in L(x^{(k+1)})} w_i x_i - \sum_{i \in R(x^{(k)})} w_i x_i$ and
3. $\Omega = \sum_{i \in L(x^{(k+1)})} w_i (x_i)^2 - \sum_{i \in R(x^{(k)})} w_i (x_i)^2$,

we can represent the lever equation as

$$(2K - 1)p^2 - 2p\Delta + \Omega = 0.$$

Step 6. If $K = \frac{1}{2}$ then $p = \frac{\Omega}{2\Delta}$. If $K \neq \frac{1}{2}$ then p is the value in $(x^{(k)}, x^{(k+1)})$ that solves the second degree equation. This value can be found by standard methods, and its existence is guaranteed by previous results.

- Consider the case in which k tends to zero. Notice that in this case, independently of x_i, $(\mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}) - \mathbf{x_i})^{\mathbf{k}}$ and $(x_i - \mathbf{L}^{\varphi}_{\mathbf{w}}(\mathbf{x}))^{\mathbf{k}}$ tend to 1, and differences between the two sides of the equation just come from the sum of weights. Clearly, the solution of the equation tends to the median element, i.e., the element x such that $\sum_{i\in L} w_i < \frac{1}{2}$ and $\sum_{i\in N\setminus R} w_i > \frac{1}{2}$ whenever it exists.[4]
- In the case in which k tends to infinity, the solution will approach the point $\frac{(x_{(1)}+x_{(n)})}{2}$.

4 Conclusions

Considering the physical concept of a lever we introduce well-known aggregation operators. This approach is related to other ones considered by several authors. In our case, we have introduced the basic notion of a lever function, and the definition of a weighted lever operator associated with that function and a weighting vector. Mainly, we have shown the existence and uniqueness of a weighted lever operator. Our purpose, is to extend these first notions to other well-known weighted aggregation operators, as OWAs or discrete Choquet integrals. Moreover, another extension arises when we consider weights depending on the value of inputs, this approach has been considered in some related topics for several authors (see [3, 8] or [11, 12] and additionally, in an ordinal setting, [6]).

Acknowledgements M.A. Ballester acknowledges the financial support from the Spanish Ministry of Education through the Ramón y Cajal program and projects SEJ2005-01481/ECON, FEDER and the Barcelona Economics Program of CREA. T. Calvo acknowledges the financial support from the Spanish Ministry of Education through the projects MTM2006-08322 and the program PR2007-0193.

References

[1] J. Aczel, On weighted synthesis of judgements, *Aequationes Math.*, **27**, 288-307, 1984.

[2] G. Beliakov, A. Pradera and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer-Verlag, Berlin, Heidelberg, New York, 2007.

[3] P.S. Bullen, *Handbook of Means and Their Inequalities*, Kluwer Academic Publishers, Dordrecht, Boston, London, 2003.

[4] If it does not, there exists a point x such that $\sum_{i\in N\setminus R} w_i = \frac{1}{2}$, and the solution will remain in between this point and the next one in a increasing order.

[4] T. Calvo, R. Mesiar and R.R. Yager, Quantitative Weights and Aggregation, *IEEE Trans. On Fuzzy Systems*, **12 (1)**, 62-69, 2004.

[5] D. Dubois and H. Prade, On the use of aggregation operations in information fusion processes, *Fuzzy Sets and Systems*, **142**, 143-161, 2004.

[6] L. Godo and T. Vicen, Aggregation operators for ordinal qualitative information, *IEEE Trans. on Fuzzy Systems*, **8:2**, 143-154, 2000.

[7] J. Martin, G. Mayor y J. Suñer, Funciones de agregación y el principio de mínima desviación, *Proc. Estylf'2002* , León, 157-163, 2002.

[8] R. Mesiar and J. Spirková, Weighted means and weighted functions, *J. Math. Anal. Appl.*, **42**, 151-160, 2006.

[9] R. Mesiar, Fuzzy set approach to the utility, preference relations, and aggregation operators, *Inf. Sciences*, *176*, 414-422, 2007.

[10] R. Mesiar, J. Špirková and L. Vavriková, Weighted aggregations operators based oon minimization, *Inf. Sciences*, *178*, 1133-1140, 2008.

[11] R.A. Riberio and R.A.M. Pereira, Weights as functions of attribute satisfaction values, *Proc. of the Workshop on Preference Modelling and Applications, EUROFUSE*, Granada, Spain, 937-942, 2001.

[12] R.A. Riberio and R.A.M. Pereira, Aggregation with generalized mixture operators using weighting functions, *Fuzzy Sets and Systems*, **137**, 43-58, 2003.

[13] V. Torra and Y. Narukawa, *Modelling Decisions: Information Fusion and Aggregation Operators*, Springer-Verlag, Berlin, Heidelberg, New York, 2007.

[14] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Trans. on Systems, Man and Cybernetics*, **18**, 183-190, 1988.

[15] R.R. Yager and A. Rybalov, Understanding Median as a Fusion Operator, *Int. Jorn. of General Systems*, **26:3** , 239-263, 1997.

[16] R.R. Yager, Weighted Triangular Norms using generating functions, *Int. J. Intell. System.*, **19**, 217-231, 2004.

[17] H.-J. Zimmermann and P. Zysno, Latent connectives in human decision making, *Fuzzy Sets and Systems*, **4**, 37-51, 1980.

THE FUZZY GENERALIZED HYBRID AVERAGING OPERATOR

José M. Merigó [1] **Montserrat Casanovas** [1]

[1] Department of Business Administration, University of Barcelona
Av. Diagonal 690, 08034 Barcelona, Spain, {jmerigo, mcasanovas}@ub.edu

Abstract

The hybrid averaging is an aggregation operator that uses the weighted average (WA) and the ordered weighted averaging (OWA) operator in the same formulation. We introduce the fuzzy generalized hybrid averaging (FGHA) operator. It is a new aggregation operator that uses generalized means and uncertain information represented by fuzzy numbers in the hybrid averaging operator. The main advantage of this operator is that it generalizes a wide range of aggregation operators that can be used in different applications such as decision making. For example, we could mention the fuzzy hybrid averaging (FHA), the fuzzy hybrid quadratic averaging (FHQA), the fuzzy generalized mean (FGM), the fuzzy generalized OWA (FGOWA), etc.

Keywords: Aggregation operators, Hybrid averaging, Fuzzy numbers, Decision making.

1 INTRODUCTION

The hybrid averaging (HA) operator [17] is an aggregation operator that uses the weighted average (WA) and the ordered weighted averaging (OWA) operator [18-19,21] in the same formulation. Recently [11], it has been suggested an extension of the HA operator for situations where it is not possible to use exact numbers because the information is uncertain. This operator is known as the fuzzy HA (FHA) operator and it represents the uncertain information to be aggregated with the HA operator by using FNs.

Another interesting aggregation operator is the generalized OWA (GOWA) operator [8,20]. It generalizes the OWA operator by using generalized means. Recently [12], a further generalization has been suggested by using FNs in the GOWA operator. This operator is known as the fuzzy GOWA (FGOWA) operator. The main advantage of this operator is that it generalizes a wide range of mean operators such as the fuzzy OWA (FOWA), the fuzzy ordered weighted quadratic averaging (FOWQA), the fuzzy generalized mean (FGM), etc. Note that the FGOWA operator can be further generalized by using quasi-arithmetic means [12]. The result is the Quasi-FOWA operator.

Going a step further, in this paper we introduce the fuzzy generalized hybrid averaging (FGHA) operator. It is an aggregation operator that uses the main characteristics of the GOWA, the FOWA and the HA operator. Then, this operator uses generalized means in the HA operator and in uncertain situations where the available information can not be represented with exact numbers but it is possible to use FNs. By using the HA operator, the FGHA considers the WA and the OWA in the same problem. In decision making problems, this implies that the FGHA operator considers the subjective probability and the attitudinal character of the decision maker in the same formulation. One of the key features of the FGHA operator is that it generalizes a wide range of aggregation operators such as the FGM, the FWGM, the FGOWA, the FWA, the FOWA, the FHA, the fuzzy hybrid quadratic averaging (FHQA), the fuzzy hybrid geometric averaging (FHGA), the fuzzy hybrid harmonic averaging (FHHA), the FOWQA, etc. Then, it is able to provide a more complete view of the decision problem to the decision maker. Obviously, the decision maker will select the particular cases that are in accordance with its interests.

We further generalize the FGHA by using quasi-arithmetic means. The result is the Quasi-FHA operator. We also develop a numerical example of the new approach in a decision making problem about selection of strategies. With this example, we will see that depending on the aggregation operator used, the results may lead to different decisions.

In order to do so, this paper is organized as follows. In Section 2 we briefly describe some basic concepts about the FNs, the FHA and the GOWA operator. In Section 3 we present the FGHA operator. Section 4 briefly analyses some basic families of the FGHA operator. Section 5 introduces the Quasi-FHA operator. In Section 6 we develop a numerical example of the new approach.

Finally, in Section 7 we summarize the main conclusions of the paper.

2 PRELIMINARIES

2.1. FUZZY NUMBERS

The FN was first introduced by [5,22]. Since then, it has been studied and applied by a lot of authors such as [6,9].

A FN is a fuzzy subset of a universe of discourse that is both convex and normal [9]. Note that the FN may be considered as a generalization of the interval number although it is not strictly the same because the interval numbers may have different meanings. In the literature, we find a wide range of FNs [6,9]. For example, a trapezoidal FN (TpFN) A of a universe of discourse R can be characterized by a trapezoidal membership function $A = (\underline{a}, \overline{a})$ such that

$$\begin{aligned} \underline{a}(\alpha) &= a_1 + \alpha(a_2 - a_1), \\ \overline{a}(\alpha) &= a_4 - \alpha(a_4 - a_3). \end{aligned} \tag{1}$$

where $\alpha \in [0, 1]$ and parameterized by (a_1, a_2, a_3, a_4) where $a_1 \le a_2 \le a_3 \le a_4$, are real values. Note that if $a_1 = a_2 = a_3 = a_4$, then, the FN is a crisp value and if $a_2 = a_3$, the FN is represented by a triangular FN (TFN). Note that the TFN can be parameterized by (a_1, a_2, a_4).

In the following, we are going to review the FN arithmetic operations as follows. Let A and B be two TFN, where $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$. Then:

1) $A + B = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$
2) $A - B = (a_1 - b_3, a_2 - b_2, a_3 - b_1)$
3) $A \times k = (k \times a_1, k \times a_2, k \times a_3)$; for $k > 0$.

Note that other operations could be studied [6,9] but in this paper we will focus on these ones.

2.2. FUZZY HYBRID AVERAGING OPERATOR

The FHA operator [11] is an aggregation operator that uses the weighted average (WA) and the OWA operator in the same formulation. Then, it is possible to consider in the same decision making problem, the attitudinal character of the decision maker and its subjective probability. It also deals with uncertain environments where the available information cannot be assessed with exact numbers but it is possible to find approximate results using FNs. Its main advantage is that it provides a more complete view of the aggregation problem because it considers a wide range of particular cases. It can be defined as follows.

Definition 1. Let Ψ be the set of FNs. A FHA operator of dimension n is a mapping FHA: $\Psi^n \to \Psi$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$FHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = \sum_{j=1}^{n} w_j b_j \tag{2}$$

where b_j is the jth largest of the $\hat{a}_i$ ($\hat{a}_i = n\omega_i \tilde{a}_i$, $i = 1,2,\ldots,n$), $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weighting vector of the $\tilde{a}_i$, with $\omega_i \in [0, 1]$ and the sum of the weights is 1, and the $\tilde{a}_i$ are FNs.

Similar to the FOWA operator, it is possible to analyze different properties of the FHA operator [10-11]. Note that in this case, we should also consider the problem of comparing FNs in the reordering process. For simplicity, we recommend to follow the policy explained in [11-12].

2.3. GENERALIZED OWA OPERATOR

The GOWA operator [8,20] is a generalization of the OWA operator by using generalized means. It includes a wide range of mean operators such as the usual OWA, the OWG operator, the ordered weighted quadratic averaging operator (OWQA), etc. It can be defined as follows.

Definition 2. A GOWA operator of dimension n is a mapping $GOWA$: $R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$GOWA(a_1, a_2, \ldots, a_n) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \tag{3}$$

where b_j is the jth largest of the a_i, and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

For further reading on the GOWA operator, see for example, [1-2,4,7-8,10,12,20].

3 THE FUZZY GENERALIZED HYBRID AVERAGING OPERATOR

The fuzzy generalized hybrid averaging operator is an extension of the HA operator that uses generalized means and uncertain information that can be represented by

using FNs. Then, we can represent the problems considering the best and worst result and the possibility that the internal values of the FN will occur. With this generalization, we include in the same formulation a lot of aggregation operators such as the FHA, the fuzzy hybrid quadratic averaging (FHQA), the FGOWA, the FWGM, etc. It can be defined as follows.

Definition 3. Let Ψ be the set of FNs. A FGHA operator of dimension n is a mapping $FGHA: \Psi^n \rightarrow \Psi$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$FGHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \quad (4)$$

where b_j is the jth largest of the $\hat{a}_i$ ($\hat{a}_i = n\omega_i \tilde{a}_i$, $i = 1,2,\ldots,n$), $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weighting vector of the $\tilde{a}_i$, with $\omega_i \in [0, 1]$ and the sum of the weights is 1, the $\tilde{a}_i$ are FNs and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

Note that if $\lambda \leq 0$, we can only use positive numbers R^+, in order to get consistent results. Note also that different types of FNs could be used in the aggregation such as TFNs, TpFNs, L-RFNs, interval-valued FNs, intuitionistic FNs, etc.

In this case, the reordering of the arguments has an additional difficulty because now we are using FNs. Then, in some cases it is not clear which FN is higher, so we need to establish an additional criteria for reordering the FNs. For simplicity, we recommend the following procedure. Select the FN with the highest value in the membership level $\alpha = 1$. Note that if the membership level $\alpha = 1$ is an interval, then, we will calculate the average of the interval.

From a generalized perspective of the reordering step, it is possible to distinguish between the descending FGHA (DFGHA) and the ascending FGHA (AFGHA) operator. The weights of these operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the DFGHA and w^*_{n-j+1} the jth weight of the AFGHA operator.

The FGHA operator is monotonic, commutative and idempotent. It is monotonic because if $\tilde{a}_i \geq \tilde{u}_i$, for all $\tilde{a}_i$, then, $FGHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) \geq FGHA(\tilde{u}_1, \tilde{u}_2 \ldots, \tilde{u}_n)$. It is commutative because any permutation of the arguments has the same evaluation. That is, $FGHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = FGHA(\tilde{u}_1, \tilde{u}_2 \ldots, \tilde{u}_n)$, where $(\tilde{u}_1, \tilde{u}_2 \ldots, \tilde{u}_n)$ is any permutation of the arguments $(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n)$. It is idempotent because if $\tilde{a}_i = \tilde{a}$, for all $\tilde{a}_i$, then, $FGHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = \tilde{a}$. Note that this operator is not bounded by the maximum and the minimum because for some special situations it can be higher and lower, respectively. Mainly, this problem is found when using the hybrid maximum and the hybrid minimum in the aggregation and in similar situations.

A further interesting issue to consider are the measures for characterizing the weighting vector W of the FGHA operator such as the attitudinal character, the entropy of dispersion, the divergence of W and the balance operator. Note that these measures follow the same methodology as the original version developed for the OWA and the GOWA operator [10,18-20].

4 FAMILIES OF FGHA OPERATORS

In this Section, we analyze different families of FGHA operators. We will distinguish between those found in the weighting vector W and those found in the parameter λ.

4.1. ANALYSING THE WEIGHTING VECTOR *W*

By using a different weighting vector in the FGHA operator, we are able to obtain different types of fuzzy aggregation operators. For example, it is possible to obtain the fuzzy hybrid maximum, the fuzzy hybrid minimum, the fuzzy generalized mean (FGM), the fuzzy weighted generalized mean (FWGM) and the FGOWA operator.

Remark 1: The fuzzy hybrid maximum is obtained if $w_1 = 1$ and $w_j = 0$, for all $j \neq 1$. The fuzzy hybrid minimum is obtained if $w_n = 1$ and $w_j = 0$, for all $j \neq n$. More generally, if $w_k = 1$ and $w_j = 0$, for all $j \neq k$, we get for any λ, $FGHA(a_1, a_2 \ldots, a_n) = b_k$, where b_k is the kth largest argument a_i. The FGM is found when $w_j = 1/n$, and $\omega_i = 1/n$, for all a_i. The FWGM is obtained when $w_j = 1/n$, for all a_i. The FGOWA is found when $\omega_i = 1/n$, for all a_i.

Remark 2: Following a similar methodology as it has been developed in [10,19], we could study other particular cases of the FGHA operator. For example, when $w_{j^*} = 1/m$ for $k \leq j^* \leq k + m - 1$ and $w_{j^*} = 0$ for $j^* > k + m$ and $j^* < k$, we are using the window-FGHA operator. Note that k and m must be positive integers such that $k + m - 1 \leq n$.

Remark 3: The olympic-FGHA, is found when $w_1 = w_n = 0$, and for all others $w_{j^*} = 1/(n - 2)$. Note that if $n = 3$ or $n = 4$, the olympic-FGHA is transformed in the median-FGHA and if $m = n - 2$ and $k = 2$, the window-FGHA is transformed in the olympic-FGHA.

Remark 4: Another type of aggregation that could be used is the E-Z FGHA weights. In this case, we should distinguish between two classes. In the first class, we assign $w_{j^*} = (1/q)$ for $j^* = 1$ to q and $w_{j^*} = 0$ for $j^* > q$, and

in the second class, we assign $w_{j^*} = 0$ for $j^* = 1$ to $n - q$ and $w_{j^*} = (1/q)$ for $j^* = n - q + 1$ to n. If $q = 1$ for the first class, the E-Z FGHA becomes the fuzzy hybrid maximum. And if $q = 1$ for the second class, the E-Z FGHA becomes the fuzzy hybrid minimum.

Remark 5: We note that the median can also be used as FGHA operators. If n is odd we assign $w_{(n+1)/2} = 1$ and $w_{j^*} = 0$ for all others. If n is even we assign for example, $w_{n/2} = w_{(n/2)+1} = 0.5$ and $w_{j^*} = 0$ for all others.

Remark 6: For the weighted median-FGHA, we select the argument b_k that has the kth largest argument such that the sum of the weights from 1 to k is equal or higher than 0.5 and the sum of the weights from 1 to $k - 1$ is less than 0.5.

Remark 7: A further interesting family is the S-FGHA operator based on the S-OWA operator [14]. It can be subdivided in three classes: the "orlike", the "andlike" and the generalized S-FGHA operator. The generalized S-FGHA operator is obtained when $w_1 = (1/n)(1 - (\alpha + \beta)) + \alpha$, $w_n = (1/n)(1 - (\alpha + \beta)) + \beta$, and $w_j = (1/n)(1 - (\alpha + \beta))$ for $j = 2$ to $n - 1$ where $\alpha, \beta \in [0, 1]$ and $\alpha + \beta \leq 1$. Note that if $\alpha = 0$, the generalized S-FGHA operator becomes the "andlike" S-FGHA operator and if $\beta = 0$, it becomes the "orlike" S-FGHA operator.

Remark 8: Another family of aggregation operator that could be used is the centered-FGHA operator. We could define a FGHA operator as a centered aggregation operator if it is symmetric, strongly decaying and inclusive. Note that these properties have to be accomplished for the weighting vector w of the FOWA operator but not necessarily for the weighting vector ω of the FWA. It is symmetric if $w_j = w_{j+n-1}$. It is strongly decaying when $i < j \leq (n + 1)/2$ then $w_i < w_j$ and when $i > j \geq (n + 1)/2$ then $w_i < w_j$. It is inclusive if $w_j > 0$. Note that it is possible to consider a softening of the second condition by using $w_i \leq w_j$ instead of $w_i < w_j$. We shall refer to this as softly decaying centered-FGHA operator. Another particular situation of the centered-FGHA operator appears if we remove the third condition. Then, we get the non-inclusive centered-FGHA operator.

Remark 9: Other families of FGHA operators could be studied such as the Gaussian FGHA weights, the nonmonotonic FGHA operator, etc. For more information, see [10,15].

4.2. ANALYSING THE PARAMETER λ

If we analyze different values of the parameter λ, we obtain another group of particular cases such as the usual FHA operator, the fuzzy hybrid geometric averaging (FHGA) operator, the fuzzy hybrid harmonic averaging (FHHA) operator and the fuzzy hybrid quadratic averaging (FHQA) operator.

Remark 10: When $\lambda = 1$, we get the FHA operator.

$$FGHA(a_1, a_2 \ldots, a_n) = \sum_{j=1}^{n} w_j b_j \tag{5}$$

From a generalized perspective of the reordering step we can distinguish between the DFHA operator and the AFHA operator. Note that if $w_j = 1/n$, for all a_i, we get the FWA and if $\omega_j = 1/n$, for all a_i, we get the FOWA operator. If $w_j = 1/n$, and $\omega_j = 1/n$, for all a_i, then, we get the fuzzy average (FA).

Remark 11: When $\lambda = 0$, the FGHA operator becomes the FHGA operator.

$$FGHA(a_1, a_2 \ldots, a_n) = \prod_{j=1}^{n} b_j^{w_j} \tag{6}$$

In this case, it is also possible to distinguish between descending (DFHGA) and ascending (AFHGA) orders. Note that if $w_j = 1/n$, for all a_i, we get the fuzzy weighted geometric average (FWGA) and if $\omega_j = 1/n$, for all a_i, we get the fuzzy OWG (FOWG) operator. If $w_j = 1/n$, and $\omega_j = 1/n$, for all a_i, then, we get the fuzzy GM (FGM).

Remark 12: When $\lambda = -1$, we get the FHHA operator.

$$FGHA(a_1, a_2 \ldots, a_n) = \frac{1}{\sum_{j=1}^{n} \frac{w_j}{b_j}} \tag{7}$$

In this case, we get the descending FHHA (DFHHA) operator and the ascending FHHA (AFHHA) operator. Note that if $w_j = 1/n$, for all a_i, we get the fuzzy weighted harmonic mean (FWHM) and if $\omega_j = 1/n$, for all a_i, we get the fuzzy ordered weighted harmonic averaging (FOWHA) operator. If $w_j = 1/n$, and $\omega_j = 1/n$, for all a_i, then, we get the fuzzy harmonic mean (FHM).

Remark 13: When $\lambda = 2$, we get the FHQA operator.

$$FGHA(a_1, a_2 \ldots, a_n) = \left(\sum_{j=1}^{n} w_j b_j^2 \right)^{1/2} \tag{8}$$

In this case, we also get the descending FHQA (DFHQA) operator and the ascending FHQA (AFHQA) operator. If $w_j = 1/n$, for all a_i, we get the fuzzy weighted quadratic average (FWQA) and if $\omega_j = 1/n$, for all a_i, we get the FOWQA operator. If $w_j = 1/n$, and $\omega_j = 1/n$, for all a_i, then, we get the fuzzy quadratic average (FQA).

Table 1: Fuzzy payoff matrix

	S_1	S_2	S_3	S_4	S_5
A_1	(20,30,40)	(40,50,60)	(30,40,50)	(50,60,70)	(60,70,80)
A_2	(60,70,80)	(30,40,50)	(50,60,70)	(50,60,70)	(20,30,40)
A_3	(40,50,60)	(60,70,80)	(20,30,40)	(50,60,70)	(30,40,50)
A_4	(60,70,80)	(60,70,80)	(50,60,70)	(30,40,50)	(10,20,30)
A_5	(30,40,50)	(60,70,80)	(60,70,80)	(30,40,50)	(30,40,50)

Table 2: Aggregated results

	FA	**FWA**	**FOWA**	**FHA**	**AFHA**
A_1	(40,50,60)	(44,54,64)	(36,46,56)	(36,45,54)	(52,63,74)
A_2	(42,52,62)	(38,48,58)	(38,48,58)	(36,45.5,55)	(40,50.5,61)
A_3	(40,50,60)	(39,49,59)	(36,46,56)	(35,44.5,54)	(43,53.5,64)
A_4	(42,52,62)	(37,47,57)	(37,47,57)	(32.5,43,53)	(41.5,51,61)
A_5	(42,52,62)	(42,52,62)	(39,49,59)	(37.5,47,56.5)	(46.5,57,67.5)

Note that we could analyze other families of FGHA operators by using different values in the parameter λ. Also note that it is possible to study these families individually in a similar way as it has been developed in Section 3 and 4.1.

Note that all the properties and particular cases commented in the FGHA operator, are also included in this generalization. Then, for example, we could mention the problem of reordering the arguments when they are FN. In order to solve this problem, we recommend the method explained in [10-12].

5 THE QUASI-FHA OPERATOR

As it is explained in [1], it is possible to further generalize the GOWA operator by using quasi-arithmetic means. Following the same methodology as in [1,7], we can suggest a similar generalization to the FGHA operator by using quasi-arithmetic means. We will call this generalization the Quasi-FHA operator. It can be defined as follows.

Definition 4. Let Ψ be the set of FN. A Quasi-FHA operator of dimension n is a mapping $QFHA$: $\Psi^n \rightarrow \Psi$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$Quasi\text{-}FHA(\tilde{a}_1, \tilde{a}_2 \ldots, \tilde{a}_n) = g^{-1}\left(\sum_{j=1}^{n} w_j g\left(b_{(j)}\right)\right) \qquad (9)$$

where where b_j is the jth largest of the $\hat{a}_i$ ($\hat{a}_i = n\omega_i\tilde{a}_i$, $i = 1,2,\ldots,n$), $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$ is the weighting vector of the $\tilde{a}_i$, with $\omega_i \in [0, 1]$ and the sum of the weights is 1, the $\tilde{a}_i$ are FNs and $g(b)$ is a strictly continuous monotonic function.

As we can see, we replace b^λ of the FGHA with a general continuous strictly monotone function $g(b)$. In this case, the weights of the ascending and descending versions are also related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the Quasi-DFHA and w^*_{n-j+1} the jth weight of the Quasi-AFHA operator.

6 NUMERICAL EXAMPLE

In the following, we are going to develop a numerical example in order to illustrate the new approach. We will consider a decision making problem where a company is analysing the optimal strategy for the next period. We will use different types of FGHA operators such as the FA, the FWA, the FOWA, the FHA and the AFHA operator.

Assume that a company that it is operating in Europe and North America, is analysing its general policy for the next year and they consider 5 possible strategies to follow.

In order to evaluate these strategies, the group of experts of the company considers that the key factor for the next year is the economic situation. Then, depending on the situation, the expected benefits for the company will be different. The experts have considered 5 possible situations for the next year: S_1 = Very bad, S_2 = Bad, S_3 = Normal, S_4 = Good, S_5 = Very good. The results are shown in Table 1. Note that the results are TFN.

In this example, we assume that the experts use the following weighting vector for all the cases of the WA and the OWA: $W = (0.1, 0.2, 0.2, 0.2, 0.3)$.

With this information, it is possible to aggregate it in order to take a decision. The results are shown in Table 2.

As we can see, depending on the aggregation operator used, the optimal strategy may be different.

A further interesting issue is to establish an ordering of the strategies. This is very useful when we want to consider more than one alternative. The results are shown in Table 3.

Table 3: Ordering of the strategies

	Ordering
FA	$A_2=A_4=A_5 \succ A_1=A_3$
FWA	$A_1 \succ A_5 \succ A_3 \succ A_2 \succ A_4$
FOWA	$A_5 \succ A_2 \succ A_4 \succ A_1=A_3$
FHA	$A_5 \succ A_2 \succ A_1 \succ A_3 \succ A_4$
AFHA	$A_1 \succ A_5 \succ A_3 \succ A_4 \succ A_2$

As we can see, we get different orderings depending on the fuzzy aggregation operator used. Obviously, the decision maker will choose the particular case that is in accordance with its interests.

7 CONCLUSIONS

We have presented the FGHA operator. We have seen that it is an extension of the HA operator that uses generalized means and uncertain information assessed with FNs. Moreover, we have seen that it includes the WA and the OWA in the same formulation because it is an extension of the HA operator. Particularly, this operator includes the FWGM and the FGOWA operator by using the hybrid formulation. It also includes a lot of other types of aggregation operators such as the FHA, the FHGA, the FHQA, etc.

We have further generalized it by using quasi-arithmetic means. As a result, we have obtained the Quasi-FHA operator. We have also developed an application of the FGHA operator in a decision making problem about selection of strategies.

In future research we expect to develop further extensions of the FGHA operator by adding new characteristics in the aggregation and applying it to other decision making problems.

Acknowledgements

We thank the anonymous referees for their comments which have improved the quality of the paper.

References

[1] G. Beliakov. Learning Weights in the Generalized OWA Operators, *Fuzzy Optimization and Decision Making* 4, Pág. 119-130, 2005.

[2] G. Beliakov, A. Pradera, T. Calvo. *Aggregation Functions: A Guide for Practitioners*, Springer-Verlag, Berlin, 2007.

[3] H. Bustince, F. Herrera, J. Montero, Fuzzy Sets and their Extensions: Intelligent Systems from Decision Making to Data Mining, Web Intelligence and Computer Vision, Springer-Verlag, 2008.

[4] T. Calvo, G. Mayor, R. Mesiar. *Aggregation Operators: New Trends and Applications*, Physica-Verlag, New York, 2002.

[5] S.S.L. Chang, L.A. Zadeh. On fuzzy mapping and control, *IEEE Transactions on Systems, Man and Cybernetics* 2, Pág. 30-34, 1972.

[6] D. Dubois, H. Prade. *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.

[7] J. Fodor, J.L. Marichal, M. Roubens. Characterization of the ordered weighted averaging operators, *IEEE Transactions on Fuzzy Systems* 3, Pág. 236-240, 1995.

[8] N. Karayiannis. Soft Learning Vector Quantization and Clustering Algorithms Based on Ordered Weighted Aggregation Operators, *IEEE Transactions on Neural Networks* 11, Pág. 1093-1105, 2000.

[9] A. Kaufmann, M.M. Gupta. *Introduction to fuzzy arithmetic*, Publications Van Nostrand, Rheinhold, 1985.

[10] J.M. Merigó. *New extensions to the OWA operators and their application in business decision making*, Unpublished thesis (in Spanish), Department of Business Administration, University of Barcelona, 2007.

[11] J.M. Merigó, M. Casanovas. Using Fuzzy OWA operators in decision making with Dempster-Shafer belief structure, In *Proceedings of the AEDEM International Conference*, pp. 475-486, Krakow, Poland, 2007.

[12] J.M. Merigó, M. Casanovas. The fuzzy generalized OWA operator, In *Proceedings of the SIGEF Conference*, pp. 504-517, Poiana-Brasov, Romania, 2007.

[13] Z.S. Xu, Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment, *Information Sciences* 168, Pág. 171-184, 2004.

[14] Z.S. Xu, A note on linguistic hybrid arithmetic averaging operator in multiple attribute group decision making with linguistic information, *Group Decision and Negotiation* 15, Pág. 593-604, 2006.

[15] Z.S. Xu, Multi-person multi-attribute decision making models under intuitionistic fuzzy

environment, *Fuzzy Optimization and Decision Making* 6, Pág. 221-236, 2007.

[16] Z.S. Xu, Intuitionistic Fuzzy Aggregation Operators, *IEEE Transactions on Fuzzy Systems* 15, Pág. 1179-1187, 2007.

[17] Z.S. Xu, Q.L. Da. An overview of operators for aggregating the information, *International Journal of Intelligent Systems* 18, Pág. 953-969, 2003.

[18] R.R. Yager. On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making, *IEEE Transactions on Systems, Man and Cybernetics B* 18, Pág. 183-190, 1988.

[19] R.R. Yager. Families of OWA operators, *Fuzzy Sets and Systems* 59, Pág. 125-148, 1993.

[20] R.R. Yager. Generalized OWA Aggregation Operators, *Fuzzy Optimization and Decision Making* 3, 93-107, 2004.

[21] R.R. Yager, J. Kacprzyk. *The Ordered Weighted Averaging Operators: Theory and Applications*, Kluwer Academic Publishers, Norwell, 1997.

[22] L.A. Zadeh. The Concept of a Linguistic Variable and its application to Approximate Reasoning, Part 1, *Information Sciences* 8, Pág. 199-249, 1975; Part 2, *Information Sciences* 8, Pág. 301-357, 1975; Part 3, *Information Sciences* 9, Pág. 43-80, 1975.

TOWARDS A FUZZY ANSWER SET SEMANTICS FOR RESIDUATED LOGIC PROGRAMS

Nicolás Madrid Labrador, Manuel Ojeda-Aciego
Dept. Matemática Aplicada. Universidad de Málaga
{nmadrid,aciego}@ctima.uma.es

Abstract

In this work we introduce the first steps towards the definition of an answer set semantics for residuated logic programs with negation.

Keywords: Fuzzy logic programming, Stable models, Answer Sets.

1 Introduction

Twenty years later of the introduction of the stable model semantics in classical logic programming, we are still seeing a number of different applications of this notion in different logical frameworks.

Answer set semantics is an intuitive and elegant generalisation of the stable model semantics which provides a powerful solution for knowledge representation and non-monotonic reasoning problems. Originally, they were intended to deal with two negations, one strong negation and default negation. The use of these two types of negation is advocated in many contexts of interest, in particular in [13] there use is justified in relation to web rules.

It is convenient to note that stable models were initially aimed at formalizing the use of negation in logic programming as negation-as-failure and, thus, are closely related to reasoning under uncertainty. For instance, the closed world assumption for a given predicate P allows for extracting negative knowledge about P from the absence of positive information about it.

The ideal environment for developing a theory of management of uncertainty is fuzzy logic. Therefore, fuzzy logic programming has become a target theory for a suitable generalization of answer set semantics.

In this paper, we focus on the initial definitions of stable model and answer set in the framework of residuated logic programs [2], as a initial step towards an answer set semantics for multi-adjoint logic programs, which were introduced in [8].

2 Preliminaries

In this section we include the definitions needed to introduce our approach to residuated logic programs with negation. Let us start with the definition of residuated lattice:

Definition 1 *A* residuated lattice *is a triple* $\mathcal{L} = ((L, \leq), *, \leftarrow)$ *such that:*

1. $(L, \leq)$ *is a complete and bounded lattice with largest element 1 and least element 0.*
2. $(L, *, 1)$ *is a commutative monoid unit element 1.*
3. $*$ *and* $\leftarrow$ *form an adjoint pair, i.e:*

$$z \leq (x \leftarrow y) \text{ iff } y * z \leq x \quad \text{for all } x, y, z \in L.$$

If $((L, \leq), *, \leftarrow)$ is a residuated lattice then $*$ is increasing in both components and $\leftarrow$ is increasing in the first argument and decreasing in the second argument (see [7]). Thus, we can interpret the operator $*$ like a conjunction and the operator $\leftarrow$ like an implication.

In the rest of the paper we will consider two different types of negation operators, which will modelize the notions of strong negation $\sim$ and default negation $\neg$ often used in logic programming. As usual, a negation operator, over L, is any decreasing mapping $n : L \to L$ satisfying $n(0) = 1$ and $n(1) = 0$.

The difference between strong and default negation in our context is semantical in essence, and relates to the method we use to infer the truth value of one negated propositional symbol.

The strong negation operator will be used in order to define a measure of consistency of a fuzzy interpretation, whereas the default negation will be used during the construction of a reduct (or program division).

In particular, it is worth to note that the notion of strong negation needs not coincide with the definition of strong negation introduced in [10].

Definition 2 *A* residuated lattice with negation *is an algebra* $(L, *, \leftarrow, \sim, \neg)$*, where* $(L, *, \leftarrow)$ *is a residuated lattice,* $\sim$ *is a strong negation and* $\neg$ *is a default negation.*

In order to introduce our logic programs, we will assume a set Π of propositional symbols. If $p \in \Pi$, then both p and $\sim p$ are called *literals*. We will denote arbitrary literals with the symbol ℓ (possible subscripted), and the set of all literals as Lit.

Definition 3 *Given a residuated lattice with negation* $(L, *, \leftarrow, \sim, \neg)$*, an* extended residuated logic program $\mathbb{P}$ *is a set of weighted rules of the form*

$$\langle \ell \leftarrow \ell_1 * \cdots * \ell_m * \neg \ell_{m+1} * \cdots * \neg \ell_n; \quad \vartheta \rangle$$

where ϑ *is an element of* L *and* $\ell, \ell_1, \ldots, \ell_n$ *are literals.*

Rules will be frequently denoted as $\langle \ell \leftarrow \mathcal{B}; \vartheta \rangle$. As usual, the formula $\mathcal{B}$ is called the *body* of the rule whereas ℓ is called its *head*. We consider *facts* as rules with empty body, which are interpreted as a rule $\langle \ell \leftarrow 1; \quad \vartheta \rangle$.

Definition 4 *A* (fuzzy) interpretation *is a mapping* $I\colon Lit \to L$*; note that the domain of an interpretation can be extended to any rule by homomorphic extension.*

We say that I satisfies *a rule* $\langle \ell \leftarrow \mathcal{B}; \quad \vartheta \rangle$ *if and only if* $I(\mathcal{B}) * \vartheta \leq I(\ell)$ *or, equivalently,* $\vartheta \leq I(\ell \leftarrow \mathcal{B})$.

Finally, I *is a* model *of* $\mathbb{P}$ *if it satisfies all rules (and facts) in* $\mathbb{P}$.

Note that the domain of our interpretations is the whole set of literals Lit, hence we are following an approach which is *not* compositional wrt the strong negation and compositional wrt the default negation.

An extended residuated logic program $\mathbb{P}$ is said to be:

- *definite* if it does not contain negation operators.
- *general* if it does not contain strong negation, but might contain default negation.
- *normal* if it does not contain default negation but it might contain strong negation.

3 Fuzzy stable sets

Our aim in this section is to adapt the approach given in [5] to the general residuated logic programs just defined in the section above.

Let us consider a general residuated logic program $\mathbb{P}$ together with a fuzzy interpretation I. To begin with, we will construct a new general program $\mathbb{P}_I$ by substituting each rule in $\mathbb{P}$ such as

$$\langle p \leftarrow p_1 * \cdots * p_m * \neg p_{m+1} * \cdots * \neg p_n; \quad \vartheta \rangle$$

by the rule[1]

$$\langle p \leftarrow p_1 * \cdots * p_m; \quad \dot{\neg} I(p_{m+1}) * \cdots * \dot{\neg} I(p_n) * \vartheta \rangle$$

Notice that the new program $\mathbb{P}_I$ is definite, that is, does not contain default negation; in fact, the construction closely resembles that of a reduct in the classical case, this is why we introduce the following:

Definition 5 *The program* $\mathbb{P}_I$ *is called the* reduct *of* $\mathbb{P}$ *wrt the interpretation* I.

As a result of the definition, note that given two fuzzy interpretations I and J, then the reducts $\mathbb{P}_I$ and $\mathbb{P}_J$ have the same rules, and might only differ in the values of the weights. By the monotonicity properties of $*$ and $\neg$, we have that if $I \leq J$ then the weight of a rule in $\mathbb{P}_I$ is greater or equal than its weight in $\mathbb{P}_J$.

In the following lemma, we show that every model M of the program $\mathbb{P}$ is a model of the reduct $\mathbb{P}_M$.

Lemma 1 *M is a model of $\mathbb{P}$ iff M is a model of $\mathbb{P}_M$.*

Proof Let us consider a rule in $\mathbb{P}$

$$\langle p \leftarrow p_1 * \cdots * p_m * \neg p_{m+1} * \cdots * \neg p_n; \quad \vartheta \rangle$$

Assume that M is a model of $\mathbb{P}$, then M satisfies the rule above, that is

$$M(p_1 * \cdots * p_m * \neg p_{m+1} * \cdots * \neg p_n) * \vartheta \leq M(p)$$

Now, by the homomorphic extension of interpretations and the adjoint property, the inequality above is equivalent to:

$$\neg M(p_{m+1}) * \cdots * \neg M(p_n) * \vartheta \leq M(p \leftarrow p_1 * \cdots * p_m)$$

that is, M satisfies the rule

$$\langle p \leftarrow p_1 * \cdots * p_m; \quad \neg M(p_{m+1}) * \cdots * \neg M(p_n) * \vartheta \rangle$$

□

Recall that *a fuzzy interpretation* can be interpreted as a L-fuzzy subset. Now, as usual, the notion of reduct allows for defining a *stable set* for a program.

[1] Note the use of a dotted version of the negation symbol to denote the semantical negation operator.

Definition 6 *Let $\mathbb{P}$ be a general residuated logic program and let I be a fuzzy interpretation; I is said to be a* stable set *of $\mathbb{P}$ iff I is a minimal model of $\mathbb{P}_I$.*

Theorem 1 *Any stable set of $\mathbb{P}$ is a minimal model of $\mathbb{P}$.*

Proof Let M be a stable set of $\mathbb{P}$ which, by definition, is a model of $\mathbb{P}_M$. Now, by Lemma 1, M is a model of $\mathbb{P}$. In order to prove minimality, assume that N is another model of $\mathbb{P}$ such that $N \leq M$, and let us show that $N = M$. Let us note the following facts about N:

1. N is a model of $\mathbb{P}_N$, again by Lemma 1.
2. The weights in the rules in $\mathbb{P}_N$ are greater or equal than those in $\mathbb{P}_M$, by the remark to Definition 5.

As a result, since N is a model of $\mathbb{P}_N$, then N is a model of $\mathbb{P}_M$ as well.

Finally, as M is a minimal model of $\mathbb{P}_M$, then $M \subseteq N$. □

Thanks to Theorem 1 we know that every stable set is a model, thus we have a suitable generalization of the concept of stable model in this generalized framework. Hereafter, specially in a semantic context, we will use the term stable model to refer to a stable set.

This approach to reducts and stable models is a conservative extension of the classical approach, as we show below:

Firstly, a classical (bi-valued) extended program $\mathbb{P}$ can be seen as an extended residuated logic program on the lattice $\{0, 1\}$ and such that each rule in $\mathbb{P}$ has weight 1.

Now, let us see that the construction of our (fuzzy) reduct on a classical program provides exactly the classical reduct.

Consider a rule in $\mathbb{P}$, which can be written as follows:

$$\langle p \leftarrow p_1 \wedge \cdots \wedge p_m \wedge \neg p_{m+1} \wedge \cdots \wedge \neg p_n; \quad 1 \rangle$$

Now, given a fuzzy interpretation I, the rule is transformed in the reduct in the rule $p \leftarrow p_1 \wedge \cdots \wedge p_n$ together with a new weight. Obviously, the value of this weight is either 0 or 1, as follows:

- if $p_i \notin M$ for all $i \in \{m+1, ..., n\}$, the assigned value of the rule is

$$1 * \neg p_{m+1} * \cdots * \neg p_n = 1 * \neg 0 * \cdots * \neg 0 = 1$$

- if $p_i \in M$ for some $i \in \{m+1, ..., n\}$; then the formula for constructing the weight in the reduct leads to the value 0.

As a result, the rule is transformed in either

$$\langle p \leftarrow p_1 \wedge \cdots \wedge p_m; \quad 1 \rangle$$

or

$$\langle p \leftarrow p_1 \wedge \cdots \wedge p_m; \quad 0 \rangle$$

Note that the effect of the weight 0 is similar to eliminate the rule, the same as in the classical approach.

Example Let us consider the following program with just one rule $\mathbb{P} = \{p \leftarrow \neg q\}$. In classical logic, for this program there exist exactly four different interpretations, and only one of them is a stable model, namely, $I(p) = 1$ and $I(q) = 0$.

Now, let us consider the extended version of the program above, where the only difference is that we can assign a weight to the rule and that the propositional symbols are evaluated in a residuated lattice with negation

$$\langle p \leftarrow \neg q; \quad \vartheta \rangle$$

Given a fuzzy interpretation $I \colon \Pi \to L$, the reduct $\mathbb{P}_I$ is the rule (actually, the fact) below

$$\langle p; \quad \vartheta * \dot{\neg} I(q) \rangle$$

for which the least model is $M(p) = \vartheta * \dot{\neg} I(q)$, and $M(q) = 0$. As a result, I is a stable model of $\mathbb{P}$ if and only if $I(p) = \vartheta * \dot{\neg} I(0) = \vartheta * 1 = \vartheta$ and $I(q) = 0$. □

4 Fuzzy answer sets

In this section, we concentrate on strong negation and we will consider normal residuated logic programs.

Note that, as our interpretations are defined on the set of literals, every normal program has a least model which can be obtained, for instance, by iterating the immediate consequence operator, see [2].

In the classical case, one has to take into account the interaction between opposite literals in order to reject inconsistent models. The advantage of working in a fuzzy framework is that one can allow that two opposite literals, such as p and $\sim p$, live together ... under some requirements.

Our approach will be based on a generalization of the concept of consistency which we have called *coherence*, to distinguish it from other existing definitions of consistency in a fuzzy setting.

Definition 7 *A fuzzy interpretation I over Lit is* coherent *if for every propositional symbol p the inequality $I(\sim p) \leq \dot{\sim} I(p)$ holds.*

It is easy to check that our notion of coherence coincides with consistency when applied in the framework of classical logic.

The following properties of coherent interpretations will be used in the rest of the section.

Proposition 1 *Let I and J be interpretations such that $I \leq J$ and J is coherent, then I is coherent as well.*

Proof Let p be a propositional symbol, then by the inequality $I \leq J$, the coherence of J and the decreasing property of $\dot{\sim}$ we have

$$I(\sim p) \leq J(\sim p) \leq \dot{\sim} J(p) \leq \dot{\sim} I(p)$$

□

Corollary 1 *If M is a fuzzy coherent model of $\mathbb{P}$ and $T \leq M$ is a another model of $\mathbb{P}$ then T is a coherent model as well.*

As a consequence of the previous corollary we can introduce the following definition:

Definition 8 *Let $\mathbb{P}$ be a normal residuated logic program, we say that $\mathbb{P}$ is* coherent *if its least model is coherent.*

Example Consider the following normal residuated logic program over the unit interval and strong negation $\sim x = 1 - x$:

$$\begin{aligned} \langle p \leftarrow; \quad & 1 \rangle \\ \langle \sim p \leftarrow; \quad & 0'3 \rangle \end{aligned}$$

This program is not coherent because its unique minimal model $M = \{(p, 1), (\sim p, 0'3)\}$ is not a coherent interpretation, since

$$0'3 = M(\sim p) > \sim M(p) = 0$$

□

Once the concept of coherence has been presented, we can introduce the notion of *fuzzy answer set*. Such a set is a fuzzy set of literals, similarly to the classical case, which sometimes will be considered a fuzzy interpretation.

Definition 9 *Let $\mathbb{P}$ be a coherent normal residuated logic program; the* fuzzy answer set *of $\mathbb{P}$ is its least coherent model of $\mathbb{P}$.*

If $\mathbb{P}$ is a positive program, then it is a coherent program and the fuzzy answer set of $\mathbb{P}$ is simply the least fuzzy model of $\mathbb{P}$.

Now, the definition of *fuzzy answer set* for extended residuated logic programs is just a combination of that for normal programs and stable models, via the construction of reducts [6].

Given an *extended* residuated logic program $\mathbb{P}$, it will be transformed into a new *general* logic program $\mathbb{P}^+$ in which we simply "forget" that the rules in the program are constructed from literals, and consider negative literals as new, independent, propositional symbol.

Formally, for any propositional symbol p occurring in $\mathbb{P}$, let p' be a new propositional symbol, which will be called the *positive form* of the negative literal $\sim p$. Every positive literal is, by definition, its own positive form. The positive form of the literal ℓ will be denoted by ℓ^+ and $\mathbb{P}^+$ will stand for the normal program obtained from $\mathbb{P}$ by replacing each rule

$$\ell \leftarrow \ell_1 * \ell_2 * ... * \ell_m * \neg\ell_{m+1} * \cdots * \neg\ell_n$$

by

$$\ell^+ \leftarrow \ell_1^+ * \ell_2^+ * ... * \ell_m^+ * \neg\ell_{m+1}^+ * \cdots * \neg\ell_n^+$$

The transformation above can be easily applied to fuzzy interpretations, in that, an interpretation I such that $I(\sim p) = \vartheta$ is transformed into a fuzzy interpretation I^+ such that $I^+(p') = \vartheta$. As a consequence, we obtain in a straightforward way the following lemma.

Lemma 2 *M is a model of $\mathbb{P}$ if and only if M^+ is a model of $\mathbb{P}^+$.*

It is easy to see that the program $\mathbb{P}^+$ is a general program and, thus, we can consider its stable models. This leads to our definition of fuzzy answer set for an extended program.

Definition 10 *Given an extended residuated logic program $\mathbb{P}$, a* fuzzy answer set *for $\mathbb{P}$ is a stable model for $\mathbb{P}^+$.*

Now, it is convenient to show that the different definitions made for particular classes of programs coincide (whenever it makes sense to establish such a comparison).

To begin with, in the following proposition, we state the relationship between the concepts of fuzzy answer sets for a normal program and stable model for a definite program.

Proposition 2 *Let $\mathbb{P}$ be a normal residuated logic program and I a fuzzy coherent interpretation; then I is a fuzzy answer set of $\mathbb{P}$ if and only if I^+ is a stable model of $\mathbb{P}^+$.*

Proof Follows from the fact that, if $\mathbb{P}$ is normal, then $\mathbb{P}^+$ does not contain default negation; as a result, its only stable model is its least model. □

It is clear, on the other hand, that the concept of fuzzy stable model for a general program coincides with the least model of a definite one; simply, because the latter does not contain negation (neither strong nor default).

At this point it should be clear that there exists another "reasonable" possibility for defining a fuzzy answer set for an extended program; namely, the construction of reduct[2] to be directly applied on an extended program, $\mathbb{P}^I$, and then, compute its least model. If this least model coincides with I. The natural question here is whether I is a fuzzy answer set for $\mathbb{P}$, in the sense of Definition 10.

Lemma 3 *Let $\mathbb{P}$ be an extended residuated logic program, then $(\mathbb{P}^I)^+ = (\mathbb{P}^+)_{I^+}$.*

Proof Let us consider a rule of $\mathbb{P}$

$$\langle \ell \leftarrow \ell_1 * \cdots * \ell_m * \neg\ell_{m+1} * \cdots * \neg\ell_n; \quad \vartheta \rangle$$

when constructing its reduct wrt I we obtain

$$\langle \ell \leftarrow \ell_1 * \cdots * \ell_m; \quad \dot{\neg}(I(\ell_{m+1})) * \cdots * \dot{\neg}(I(\ell_n)) * \vartheta \rangle$$

whose positive form is

$$\langle \ell^+ \leftarrow \ell_1^+ * \cdots * \ell_m^+; \quad \dot{\neg}(I(\ell_{m+1})) * \cdots * \dot{\neg}(I(\ell_n)) * \vartheta \rangle$$

this is the resulting rule under $(\mathbb{P}^I)^+$.

On the other hand, the positivisation of the initial rule is

$$\langle \ell^+ \leftarrow \ell_1^+ * \cdots * \ell_m^+ * \neg\ell_{m+1}^+ * \cdots * \neg\ell_n^+; \quad \vartheta \rangle$$

and when computing its reduct wrt I^+ we obtain

$$\langle \ell^+ \leftarrow \ell_1^+ * \cdots * \ell_m^+; \quad \dot{\neg}(I^+(\ell_{m+1}^+)) * \cdots * \dot{\neg}(I^+(\ell_n^+)) * \vartheta \rangle$$

which is the resulting rule under $(\mathbb{P}^+)_{I^+}$.

Finally, note that, by definition, $\dot{\neg}(I^+(\ell_i)) = \dot{\neg}(I(\ell_i^+))$ for all $i = \{1, ..., n\}$. Therefore the rules are the same under both combinations. □

¿From the above lemma, we can compute a fuzzy answer set for an extended logic program $\mathbb{P}$ by means of, abusing a little bit of terminology, the *extended stable models*.

Formally, we have the following theorem:

Theorem 2 *Let $\mathbb{P}$ a extended residuated logic program. A coherent fuzzy interpretation I is a fuzzy answer set of $\mathbb{P}$ if and only if I is an extended stable of $\mathbb{P}$.*

Proof : Straightforward from Lemma 3. □

[2]Note that we write $\mathbb{P}^I$ to note that $\mathbb{P}$ needs not be a general program; anyway, the construction of the reduct is the same.

5 Other approaches to fuzzy answer set programming

Several approaches have been developed in order to cope with negation in a fuzzy logic programming setting. To the best of our knowledge, one of the first studies of negation in fuzzy logic programming was introduced in [12], which developed a compositional approach to encompass strong negation, in opposition to the well-known but non-compositional approach of [3].

In [4] a foundation for fuzzy logic programming was developed, but his approach was not based on implication rules, but on Horn clauses on which negation is interpreted as default negation. Simultaneously, [1] introduced the so-called antitonic logic programs, for which a well-founded and a stable model semantics was developed; antitonic logic programs are based, as our programs, on a residuated lattice, but their rules must satisfy the condition that their bodies should be either increasing in all variables or decreasing in all variables and, thus, they cannot accomodate rules containing both positive and negative literals in their bodies. Moreover, they consider partial interpretations, in the form of pairs of interpretations in order to keep track of the information about truth and non-falsity.

There are other approaches based on annotated multiple-valued logic [9]. In this paper, the underlying truht-values set is a complete lattice, and the main difference is that interpretations assign an *interval* to literals (in a manner similar to [1]).

More recently, a general introduction to fuzzy answer set programming has been introduced in [11], however their approach is not directly applicable to residuated logic programs, since the programs they considered did not have weights.

6 Conclusions and future work

In this paper we have set the initial definitions in order to start a thorough study of the answer set semantics of extended residuated logic programs. The definitions of fuzzy stable model for general program and of fuzzy answer set for a coherent normal program are used in order to provide the more general definition of fuzzy answer set for an extended fuzzy program. Finally, we conclude with the result that fuzzy answer sets can, equivalently, be computed by means of extended stable models.

A number of issues still have to be studied: for instance, the epistemological implications of the concept of coherence. In this paper, we have only taken into account that the resulting fuzzy answer sets should be validated against some consistency-related notion. Fu-

ture work, should go towards imbricating this notion with threshold computation which turns out to be an important issue for negation-as-failure. For instance, the absence of evidence of p could be interpreted that the value of p is at most a threshold value which cannot be detected by the sensors which provide our information.

Finally, it is important to relate our approach with other existing approaches, and study their possible interactions, advantages and disadvantages.

References

[1] C. V. Damásio and L. M. Pereira. Antitonic logic programs. In *Proc. Logic Programming and Nonmonotonic Reasoning, LPNMR'01*, pages 379–392. Lect. Notes in Artificial Intelligence, 2173, Springer-Verlag, 2001.

[2] C. V. Damásio and L. M. Pereira. Monotonic and residuated logic programs. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'01*, pages 748–759. Lect. Notes in Artificial Intelligence, 2143, 2001.

[3] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D. Gabbay et al, editor, *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 439–513. Clarendon Press, 1994.

[4] R. Ebrahim. Fuzzy logic programming. *Fuzzy Sets and Systems*, 117:215– 230, 2001.

[5] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In *Proc. of ICLP-88*, pages 1070–1080, 1988.

[6] M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9:365–385, 1991.

[7] P. Hájek. *Metamathematics of Fuzzy Logic*. Trends in Logic. Kluwer Academic, 1998.

[8] J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Similarity-based unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146(1):43–62, 2004.

[9] U. Straccia. Annotated answer set programming. Technical Report 2005-TR-51, Istituto di Scienza e Tecnologie dell'Informazione-Consiglio Nazionale delle Ricerche, 2005.

[10] E. Trillas. Sobre funciones de negación en la teoría de conjuntos difusos *Stochastica*, 3(1):47–60, 1979.

[11] D. Van Nieuwenborgh, M. De Cock, and D. Vermeir. An introduction to fuzzy answer set programming. *Ann. Math. Artif. Intell.*, 50(3-4):363–388, 2007.

[12] G. Wagner. Negation in fuzzy and possibilistic logic programs. In T. P. Martin and F. A. Fontana, editors, *Logic programming and soft computing*, chapter 6, pages 113–128. Taylor & Francis, 1998.

[13] G. Wagner. Web rules need two kinds of negation. In *Principles and Practice of Semantic Web Reasoning*, volume 2901 of *Lecture Notes in Computer Science*, pages 33–50, 2003.

The problem of the additive generation of finitely-valued t–conorms

J. Martín
Dept. of Math. and C. S.
Univ. of the Balearic Islands
javier.martin@uib.es

G. Mayor
Dept. of Math. and C. S.
Univ. of the Balearic Islands
gmayor@uib.es

J. Monreal
Dept. of Math. and C. S.
Univ. of the Balearic Islands
jaume.monreal@uib.es

Abstract

This paper deals with the problem of the additive generation of triangular conorms defined on a finite totally ordered set. After defining a new method to construct t–conorms by a nesting procedure, we study the existence of additive generators for t–conorms obtained from nesting appropriate t–conorms in the basic maximum, drastic and Łukasiewicz t–conorms.

Keywords: additive generator, finitely–valued triangular conorm, ordinal sum, nesting of t–conorms.

1 INTRODUCTION

An old problem is whether there exist constructions involving only a one-place real function and the usual addition (or multiplication) which introduce two-place real functions having interesting algebraic properties, in particular, the associativity. More details on this topic can be found in [3], in particular, we point out the following facts:
1) a continuous t–conorm is Archimedean if and only if it has a continuous additive generator;
2) there exist additive generators for the drastic t–conorm (and for other non-continuous t–conorms), but none for the maximum t–conorm.

Fuzzy logic is one of the tools for management of uncertainty; it usually works with a continuous scale, the real interval $[0, 1]$, and the logical connectives are modeled by triangular norms (conjunction) and triangular conorms (disjunction). However, practical applications of fuzzy logic are limited to a finite number of truth values. Thus, technical implementations allow us to work only with a finite (though very large) number of values. On the other hand, when representing vagueness it is usually meaningless to distinguish a high number of truth values; only a small number suffices. In this paper, we deal with triangular conorms defined on a finite ordinal scale.

In full analogy to the representation theorem of continuous t–conorms, there is a characterization of divisible (smooth) finitely-valued t–conorms as ordinal sums of Archimedean finitely-valued t–conorms ([6, 7]). However, some other results concerning discrete t–conorms differ substantially from those obtained for t–conorms on $[0, 1]$. Thus, we know that a t–conorm with non trivial idempotent elements has not an additive generator; this is not true for finitely-valued t–conorms as we can see in this paper.

Sections 2 and 3 contain the main definitions and results which are the basis of the new ones exposed in sections 4 and 5. The proofs in these sections are only shown for the most important non-trivial results.

2 PRELIMINARIES

Consider $L = \{0, 1, 2, \ldots, n\}$, $n \geq 1$, equipped with the usual ordering. We begin recalling basic definitions, examples and properties of finitely-valued t–conorms. A complete exposition of this topic can be found in [7].

Definition 1 *A triangular conorm (briefly t–conorm) on L is a binary operation $S : L \times L \to L$ such that for all $x, y, z \in L$ the following axioms are satisfied:*

1. $S(x, y) = S(y, x)$ *(commutativity)*
2. $S(S(x, y), z) = S(x, S(y, z))$ *(associativity)*
3. $S(x, y) \leq S(x', y')$ *whenever $x \leq x'$, $y \leq y'$ (monotonicity)*
4. $S(x, 0) = x$ *(boundary condition)*

Example 1 We can consider as basic t–conorms: the drastic

$$S_D(x,y) = \begin{cases} x & \text{if } y = 0 \\ y & \text{if } x = 0 \\ n & \text{otherwise} \end{cases}$$

the maximum

$$S_M(x,y) = max(x,y)$$

and the bounded sum or Łukasiewicz t–conorm

$$S_{\mathrm{L}}(x,y) = min(x+y,n)$$

Proposition 1 *Let S be a t–conorm on L. Then we have:*

1. *$S \leq S_D$. Thus, S_D is the largest t–conorm.*
2. *$S(x,y) \geq max(x,y)\ \forall x,y \in L$. Thus S_M is the smallest t–conorm*
3. *$S(x,n) = n\ \forall x \in L$ (n is an annihilator)*
4. *$S(x,x) = x\ \forall x \in L$ if and only if $S = S_M$.*

The property of divisibility can be considered for a t–conorm. This condition is the proper equivalent of the continuity of ordinary t–conorms and it plays a crucial role in our approach.

Definition 2 *A t–conorm S on L is divisible if the following condition holds:*

For all $x,y \in L$ with $x \leq y$ there is $z \in L$ such that $y = S(x,z)$

Observe that the divisibility condition is equivalent to the smoothness condition ([1, 2]): $0 \leq S(x+1,y) - S(x,y) \leq 1$ for all $x,y \in L$, $x < n$.

Given a t–conorm S on L, we say that $x \in L$ is an idempotent element of S if $S(x,x) = x$. Observe that 0 and n are idempotent elements for any t–conorm. A t–conorm S on L is Archimedean if and only if it has as unique idempotent elements the trivial ones 0 and n.

Proposition 2 *S_L is the only divisible Archimedean t–conorm on L.*

Now, we recall a well-known method for constructing a new t–conorm from two given t–conorms

Proposition 3 *Let S_1 be a t–conorm on $L_m = \{0,1,\ldots,m\}$ and S_2 a t–conorm on $L_n = \{0,1,\ldots,n\}$, with $m,n \geq 1$. Consider the binary operation S defined on $L_{m+n} = \{0,1,\ldots,m+n\}$ as follows*

$$S(x,y) = \begin{cases} S_1(x,y)\ if\ (x,y) \in L_m^2 \\ m + S_2(x-m, y-m) \\ \quad if\ (x,y) \in \{m, m+1, \ldots, m+n\}^2 \\ \max(x,y)\ otherwise \end{cases}$$

Then, S is a t–conorm on L_{m+n} that we call the ordinal sum of S_1 and S_2. We will denote $S = \langle S_1, S_2 \rangle$

Next we characterize the class of divisible t–conorms as ordinal sums of Łukasiewicz t–conorms.

Proposition 4 *A t–conorm S on $L = \{0,1,2,\ldots,n\}$ is divisible (smooth) if and only if there exists a set $I = \{0 = a_0 < a_1 < \ldots < a_r < a_{r+1} = n\}$, $0 \leq r \leq n-1$, of elements of L such that*

$$S(x,y) = \begin{cases} \min(a_{i+1}, x+y-a_i) \\ \quad if\ (x,y) \in (a_i, a_{i+1})^2,\ 0 \leq i \leq r \\ \max(x,y)\ otherwise \end{cases}$$

Remark 1

i) We have $S_M \leq S \leq S_{\mathrm{L}}$ for any divisible t–conorm S.

ii) There are exactly 2^{n-1} divisible t–conorms on L.

Example 2

i) There are 2386 t–conorms on L_7, 471 of them are Archimedean. Of the 1915 non–Archimedean t–conorms there are 1021 which are ordinal sums. On the other hand, there are 64 t–conorms that are divisible (only one of them is Archimedean and the other 63 are ordinal sums).

ii) There are 13775 t–conorms on L_8, 2760 of them are Archimedean. Of the 11105 non–Archimedean t–conorms there are 5512 which are ordinal sums. On the other hand, there are 128 divisible t–conorms (only one of them is Archimedean and 127 are ordinal sums).

3 ADDITIVE GENERATORS OF FINITELY–VALUED T–CONORMS

In this section we consider the pseudoinverse of appropriate monotone functions from $L = \{0,1,2,\ldots,n\}$ to $[0,+\infty)$, and we introduce a construction similar to that given in case of ordinary t–conorms.

Definition 3 *An additive generator $f : L \to [0,+\infty)$ of a t–conorm S on L is a strictly increasing function with $f(0) = 0$ such that*

$$S(x,y) = f^{(-1)}(f(x) + f(y))\ \ \forall x,y \in L, \qquad (1)$$

where $f^{(-1)} : [0, +\infty) \to L$ is the pseudoinverse of f, defined by $f^{(-1)}(t) = max\{z \in L; f(z) \leq t\}$.

If S is a t–conorm on L of the form (1) for some f, we say that S is additively generated by f. We often indicate $S = \langle (a_0, a_1, \ldots, a_n) \rangle$ where $a_x = f(x)$, $x \in L$. Of course $0 = a_0 < a_1 < \ldots < a_{n-1} < a_n$.

Remark 2 If $S = \langle (a_0, a_1, \ldots, a_n) \rangle$ then defining $a_x * a_y = \max\{a_z \ ; \ a_z \leq a_x + a_y\}$ we can write $S(x, y) = f^{(-1)}(f(x) + f(y)) = f^{(-1)}(a_x + a_y) = f^{-1}(a_x * a_y)$ for all $x, y \in L$.

Example 3 The basic t–conorms quoted above have additive generator:

- $S_{\mathbf{L}} = \langle (0, 1, \ldots, n-1, n) \rangle$
- $S_M = \langle (0, 1, 3, 7, 2^{n-1} - 1, 2^n - 1) \rangle$
- $S_D = \langle (0, n-1, n, \ldots, 2n-3, 2n-2) \rangle$

A known result about characterization of those t–conorms having additive generator is the following. For more details and results about additive generation of binary operations see [4, 5].

Proposition 5 *A t–conorm S on L has an additive generator if and only if there exists a continuous non–strict Archimedean t–conorm $\hat{S}$ on the real interval $[0, n]$ such that $S(x, y) = \lfloor \hat{S}(x, y) \rfloor$ for all $x, y \in L$, where $\lfloor z \rfloor$ stands for the floor of z (the greatest integer which is less than or equal to z).*

Next we prove that any t–conorm that is an ordinal sum of additively generated t–conorms is also an additively generated t–conorm.

Proposition 6 *Let $f_1 = (a_0, a_1, \ldots, a_m)$ be an additive generator of a t–conorm S_1 on $L_m = \{0, 1, \ldots, m\}$ and $f_2 = (b_0, b_1, \ldots, b_m)$ be an additive generator of a t–conorm S_2 on $L_n = \{0, 1, \ldots, n\}$. Then $f = (a_0, a_1, \ldots, a_m, (2a_m + 1)b_1, (2a_m + 1)b_2, \ldots, (2a_m + 1)b_n)$ is an additive generator of the ordinal sum $\langle S_1, S_2 \rangle$.*

According to previous results we can now establish the following result.

Proposition 7 *Any divisible t–conorm on $L = \{0, 1, \ldots, n\}$ has an additive generator.*

Proposition 8 *Any t–conorm on $L = \{0, 1, \ldots, n\}$ with $n \leq 7$ can be additively generated. There are exactly three t–conorms on $L = \{0, 1, \ldots, 8\}$ without additive generator.*

Proof: An exhaustive computation shows that an additive generator can be obtained for any t–conorm with $n \leq 7$. In case $n = 8$, the three t–conorms without additive generator are:

S_1	0	1	2	3	4	5	6	7	8
0	0	1	2	3	4	5	6	7	8
1	1	4	5	6	6	8	8	8	8
2	2	5	5	7	8	8	8	8	8
3	3	6	7	7	8	8	8	8	8
4	4	6	8	8	8	8	8	8	8
5	5	8	8	8	8	8	8	8	8
6	6	8	8	8	8	8	8	8	8
7	7	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8

S_2	0	1	2	3	4	5	6	7	8
0	0	1	2	3	4	5	6	7	8
1	1	5	5	5	7	8	8	8	8
2	2	5	6	6	7	8	8	8	8
3	3	5	6	8	8	8	8	8	8
4	4	7	7	8	8	8	8	8	8
5	5	8	8	8	8	8	8	8	8
6	6	8	8	8	8	8	8	8	8
7	7	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8

S_3	0	1	2	3	4	5	6	7	8
0	0	1	2	3	4	5	6	7	8
1	1	5	5	6	6	8	8	8	8
2	2	5	5	7	8	8	8	8	8
3	3	6	7	7	8	8	8	8	8
4	4	6	8	8	8	8	8	8	8
5	5	8	8	8	8	8	8	8	8
6	6	8	8	8	8	8	8	8	8
7	7	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8

A direct demonstration of the non–existence of additive generators for these t–conorms can be done. Let us suppose that $\{0, a, b, c, d, e, f, g, h\}$ is an additive generator of S_1. From inequalities $h > 2c$, $f \leq a + c$, $g \leq b + c$, $f > 2b$, $g > a + d$ and $h \leq b + d$, we have $a + b + h > a + b + 2c \geq f + g > a + 2b + d$, thus $h > b + d$ which is a contradiction. The same for the other t–conorms.

In the next section we introduce and study a new method to construct t–conorms.

4 NESTING OF T–CONORMS

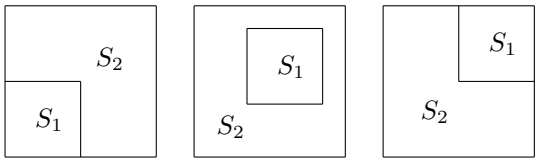
Definition 4 *Given a t–conorm S_2 on $L = \{0, 1, \ldots, r, r+1, \ldots, s, s+1, \ldots, n\}$ with $0 \leq r < s \leq n$*

and a t–conorm S_1 on $\{r, r+1, \ldots, s\}$ we define a binary operation S on L as follows:

$$S(x,y) = \begin{cases} S_1(x,y) & \text{if } r \le x, y \le s \\ S_2(x,y) & \text{otherwise} \end{cases} \quad (2)$$

We say that S is the nesting of S_1 in S_2 (fixed r and s) and we denote $S = [S_1, S_2]$.

Thus, we can distinguish three different situations (see figure below).



$r = 0$ $\quad 0 < r < s < n$ $\quad s = n$

For any t–conorms S_1 and S_2, $S = [S_1, S_2]$ is commutative with 0 and n as neutral and annihilator elements respectively. We are interested in obtaining by this method a new t–conorm.

Proposition 9 *Consider a t–conorm S_2 on $L = \{0, 1, \ldots, r, r+1, \ldots, s, s+1, \ldots, n\}$ with $0 \le r < s \le n$ and a t–conorm S_1 on $\{r, r+1, \ldots, s\}$. The nesting $S = [S_1, S_2]$ is a t–conorm if and only if the following conditions hold, when applicable:*

$$\begin{gathered} S_2(S_1(x,y),z) = S_2(S_2(x,y),z) \\ \forall x, y, z : r \le x, y \le s < z \end{gathered} \quad (3)$$

$$\begin{gathered} S_2(x,y) = \max(x,y) \\ \forall x, y : x < r \le y \le s \end{gathered} \quad (4)$$

Proof: Let us assume that S is a t–conorm. First we observe that the associativity of S and S_2, implies condition (3). From monotonicity of S and boundary conditions of S_1 and S_2 we can write

$$S(0,y) \le S(x,y) \le S(r,y)$$

Thus, for $x < r \le y \le s$

$$y = S_2(0,y) \le S_2(x,y) \le S_1(r,y) = y$$

So, $S_2(x,y) = \max(x,y) = y$ and condition (4) is satisfied.

Reciprocally, from condition (4) we deduce that S is monotone. The associativity of S follows from conditions (3) and (4).

Remark 3

i) In case $r = 0$, a remarkable particular situation is when $S_2(s,x) = \max(s,x)$ for all x. Under this hypothesis the condition (3) is trivially satisfied and $S = [S_1, S_2]$ is just the ordinal sum $\langle S_1, S_2' \rangle$, where S_2' is the t–conorm on $\{0, 1, \ldots, n-s\}$ defined by $S_2'(x,y) = S_2(x+s, y+s) - s$.

On the other hand, any ordinal sum of t–conorms $S = \langle S_1, S_2 \rangle$ is trivially a nesting of this type: $S = [S_1, S]$.

It is also worth observing that if $S = [S_1, S_2]$ is a t–conorm then it is non–Archimedean (s is a non trivial idempotent of S). Reciprocally, if S is a non–Archimedean t–conorm on $\{0, 1, \ldots, n\}$ with s as non–trivial idempotent, then S is the nesting $[S_1, S]$ where S_1 is the restriction of S to $\{0, 1, \ldots, s\}$. Thus the class of non–Archimedean t–conorms on $\{0, 1, \ldots, n\}$ is equal to the class of nestings on the same domain $\{0, 1, \ldots, n\}$ that satisfy condition (3).

ii) If $0 < r < s < n$, condition (4) tell us that the restriction of $S = [S_1, S_2]$ to $\{0, 1, \ldots, s\}^2$, $S|_{\{0,1,\ldots,s\}^2}$, is a t–conorm. So, $S = [S|_{\{0,1,\ldots,s\}^2}, S_2]$ is a nesting of the previous type.

iii) From condition (4), the resultant t–conorms in the case $s = n$ are also ordinal sums, hence they can be interpreted as nestings of the first type.

Next subsections show how we obtain new t–conorms by nesting in the basic t–conorms: maximum, drastic and Łukasiewicz. We will study nestings of the case $r = 0$ because we have seen that the other two nestings can be reduced to it.

4.1 Nesting in the maximum t–conorm

First we note that making nestings in the maximum t–conorm we obtain an ordinal sum of t–conorms.

Proposition 10 *Let S_1 be a t–conorm on $\{0, 1, \ldots, s\}$, and let S_M the maximum t–conorm on $\{0, 1, \ldots, s, \ldots, n\}$ ($s < n$), and consider S_M' the maximum t–conorm on $\{0, 1, \ldots, n-s\}$. Then $[S_1, S_M]$ is a t–conorm satisfying $[S_1, S_M] = \langle S_1, S_M' \rangle$.*

Thus, nesting a t–conorm S_1 in the maximum t–conorm we obtain a new additively generated t–conorm whenever S_1 also is.

4.2 Nesting in the drastic t–conorm

Using the drastic t–conorm, we can state:

Proposition 11 *The nesting $[S_1, S_D]$ of a t–conorm S_1 in the drastic t–conorm S_D is a t–conorm. Moreover, if $(0 = a_0, a_1, \ldots, a_s)$ is an additive generator of S_1 then $(0 = b_0, b_1, \ldots, b_n)$ is an additive generator of $[S_1, S_D]$, where $b_i = (n-s)a_i$ $i = 1, \ldots, s$, $b_{s+1} = 2b_s + 1$, $b_j = b_{j-1} + 1$ $j = s+2, \ldots, n$.*

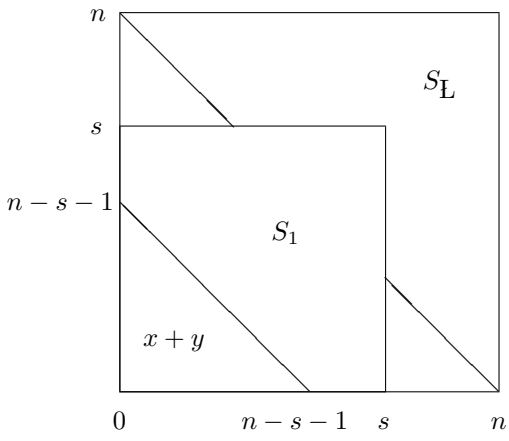
Observe that, except in trivial cases, $[S_1, S_D]$ is a non–Archimedean t–conorm that is not an ordinal sum. We can also observe that this construction can be iterated. Thus, we can consider $[[S_1, S_D], S_D]$, ... , $[\ldots[[S_1, S_D], S_D], \ldots, S_D]$, obtaining new t–conorms additively generated (if S_1 also is).

4.3 Nesting in the Łukasiewicz t–conorm

This is a different case from the two previous ones. Now S_1 needs to satisfy some conditions in order to get a new t–conorm.

Proposition 12 *The nesting $[S_1, S_{\text{Ł}}]$ of a t–conorm S_1 in the Łukasiewicz t–conorm $S_{\text{Ł}}$ is a t–conorm if and only if*

1. $s > \frac{n-2}{2}$
2. $S_1(x, y) = x + y$ *if* $x + y < n - s - 1$
3. $S_1(x, y) \geq n - s - 1$ *if* $x + y \geq n - s - 1$



Proof: First we observe that condition (3) can be written in the form

$$\begin{gathered}\min(S_1(x, y) + z, n) = \min(x + y + z, n) \\ \forall x, y \leq s \quad \forall z > s\end{gathered} \tag{5}$$

Suppose that this condition is satisfied. Taking $x = s$, $y = 1$ and $z = s + 1$ then $\min(S_1(s, 1) + s + 1, n) = \min(2s + 2, n)$. Suppose first $s < \frac{n-2}{2}$, then $S_1(s, 1) + s + 1 = 2s + 2$ which is a contradiction because $S_1(s, 1) = s$. Suppose now $s = \frac{n-2}{2}$, then $S_1(s, 1) + s + 1 \geq n$. Thus $S_1(s, 1) \geq n - s - 1 = s + 1$ which is a contradiction too. Hence $s > \frac{n-2}{2}$.

Now we proof *ii)*. Taking $z = s+1$, if $x+y+s+1 < n$ then from (5) we have $S_1(x, y)+s+1 = x+y+s+1$ and so $S_1(x, y) = x+y$ for all x, y such that $x+y < n-s-1$.

Let us prove now *iii)*. Taking $z = s + 1$ and x, y such that $x + y + s + 1 = n$ condition (5) implies $S_1(x, y) + s + 1 \geq n$, hence $S_1(x, y) \geq n - s - 1$ for all x, y with $x+y = n-s-1$ and finally from monotonicity *iii)* follows.

Reciprocally suppose that *i)*, *ii)* and *iii)* hold. Consider $x, y \leq s$ and $z > s$. If $x + y + z < n$ then $x + y + s + 1 < n$ and $x + y < n - s - 1$, hence from *ii)* $S_1(x, y) = x + y$ and condition (5) is satisfied.
In case $x+y+z \geq n$, we need to prove $S_1(x, y)+z \geq n$. If $x+y < n-s-1$ then $S_1(x, y)+z = x+y+z \geq n$, and if $x+y \geq n-s-1$ then $S_1(x, y)+z \geq n-s-1+s+1 = n$ and condition (5) holds.

The next proposition illustrates how we can get an additive generator for $[S_1, S_{\text{Ł}}]$ from one of S_1.

Proposition 13 *Let S_1 be a t–conorm on $L = \{0, 1, \ldots, s\}$ having $(0 = a_0, a_1, \ldots, a_s)$ as an additive generator. Then $(0 = b_0, b_1, \ldots, b_s, b_{s+1} \ldots, b_n)$ is an additive generator of $[S_1, S_{\text{Ł}}]$, where $b_i = a_i$ $i = 0, \ldots, s$ and $b_{s+1} = 2a_s + 1$, $b_{s+2} = 2a_s + 2 + a_0$, $b_{s+3} = 2a_s + 2 + a_1$, ..., $b_n = 2a_s + 2 + a_{n-s-2}$.*

Proof: Since $b_i = a_i$, $i = 0, \ldots, s$, we only need to show that

$$\begin{gathered}b_i * b_r = b_{\min(i+r,n)} \\ \forall\ i, r:\ 0 \leq i \leq s < r \leq n\end{gathered} \tag{6}$$

where $*$ is the binary operation considered in Remark 3.

Observe also that $b_i * b_j = b_{i+j}$ for all $0 \leq i, j \leq n - s - 2$, because S_1 satisfies conditions described above. This means $b_{i+j} \leq b_i + b_j < b_{i+j+1}$ $\forall i, j \leq n - s - 2$.

We have to study the following three cases:

1) Let's see that $b_i * b_{s+1} = b_{i+s+1}$ whenever $1 \leq i \leq n - s - 2$. We can clearly see that $b_1 * b_{s+1} = b_{s+2}$ because $b_1 + b_{s+1} = b_{s+2}$. And if we suppose $i \geq 2$ then $b_i * b_{s+1} = b_{s+i+1}$ if and only if $b_{s+i+1} \leq b_i * b_{s+1} < b_{s+i+2}$. Since $1 - b_{i-1} \leq b_1 + b_{i-1} \leq b_i$ then $1 + b_{i-1} \leq b_i < b_i + 1$ and the condition (6) holds.

2) In order to see that $b_i * b_{s+p} = b_{\min(s+i+p,n)}$ $p \geq 2$, we only need to observe that $b_{i+p-2} \leq b_i + b_{p-2} < b_{i+p-1}$ (last inequality will be only considered when $i+p-1 \leq n$). From this we obtain $b_{s+i+p} \leq b_i+b_{s+p} < b_{i+s+p+1}$ and condition (6) is satisfied.

3) Finally, we have to see $b_{n-s-1} * b_{s+p} = b_n$ $\forall p \geq 1$. This is true because $b_{n-s-1} + b_{s+p} \geq b_{n-s-1} + b_{s+1} = b_{n-s-1}+2b_s+1 \geq b_{n-s-2}+2b_s+2 = b_n$ and condition (6) holds.

5 Conclusion

This paper deals with the problem of the additive generation of triangular conorms defined on a finite

chain. After introducing a new method to construct t–conorms that extends the ordinal sum procedure, we study the existence of additive generators of t–conorms obtained from nesting appropriate ones in the basic S_M, S_D and $S_{\mathbf{L}}$.

Acknowledgements

The authors acknowledge the support of the Govern Balear grant PCTIB2005GC1-07 and the Spanish DGI grant MTM2006-08322.

References

[1] J. Fodor. Smooth associative operations on finite ordinal scales. *IEEE Transactions on Fuzzy Systems*, vol 8 pp 791-795. 2000.

[2] L. Godo, C. Sierra. A new approach to connective generation in the framework of expert systems using fuzzy logic. In *Proc. 18th IEEE Int. Symposium on Multiple-Valued Logic*, pp. 157-162. 1988.

[3] E.P. Klement, R. Mesiar, E. Pap. *Triangular Norms*. Kluwer. 2000.

[4] J. Martín, G. Mayor, J. Monreal. Additive generation of some classes of finitely–valued t–conorms. Discrete Conjunctive Aggregation Operations (accepted in IMPU 2008).

[5] G. Mayor, J. Monreal. Additive Generators of Discrete Conjunctive Aggregation Operations. *IEEE Transactions on Fuzzy Systems*, vol 15, no. 6, pp 1046-1052. 2007.

[6] G. Mayor, J. Torrens. On a class of operators for expert systems. *Int. J. Intell. Syst.* 8, pp 771-778. 1993.

[7] G. Mayor, J. Torrens. Triangular norms on discrete settings. In *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*. Edited by E.P. Klement and R. Mesiar. Elsevier. 2005.

THE INDUCED MINKOWSKI ORDERED WEIGHTED AVERAGING DISTANCE OPERATOR

José M. Merigó [1] **Montserrat Casanovas** [1]

[1] Department of Business Administration, University of Barcelona,
Av. Diagonal 690, 08034 Barcelona, Spain, {jmerigo, mcasanovas}@ub.edu

Abstract

The Minkowski distance is a distance measure that generalizes a wide range of other distances such as the Hamming and the Euclidean distance. In this paper, we develop a generalization of the Minkowski distance by using the induced ordered weighted averaging (IOWA) operator. We will call it the induced Minkowski OWA distance (IMOWAD). Then, we are able to obtain a wider range of distance measures that includes the Minkowski distance, the Minkowski OWA distance (MOWAD), the induced OWAD (IOWAD), etc. We also present a further generalization that uses quasi-arithmetic means. We will call it the Quasi-IOWAD operator. We end the paper with a numerical example of the new approach.

Keywords: Minkowski distance, Aggregation operators, IOWA operator, Decision making.

1 INTRODUCTION

The distance measures are very useful techniques that have been used in a wide range of applications such as fuzzy set theory, decision making, operational research, etc. The Minkowski distance is one of the main distance measures because it generalizes a wide range of other distances such as the Hamming distance, the Euclidean distance, etc. Often when calculating distances, we want an average result of all the individual distances. We call this the normalization process. In the literature, we find mainly three types of normalized distances. The first one is when we use the arithmetic mean and it is known as the normalized Minkowski distance (NMD). The second one is when we use the weighted average (WA) and it is known as the weighted Minkowski distance (WMD). The third one is when we use the ordered weighted averaging (OWA) operator [1-17] and it is known as the Minkowski ordered weighted averaging distance (MOWAD) operator [5,9]. Note that the MOWAD includes the NMD and the WMD as special cases.

Sometimes, when normalizing the Minkowski distance with the OWA operator, it would be interesting to consider a more general formulation of the attitudinal character. A very useful technique for doing this is the induced ordered weighted averaging (IOWA) operator [14,16]. The IOWA operator provides a parameterized family of aggregation operators such as the maximum, the minimum, the average and the OWA operator.

In this paper, we suggest a new type of distance measure consisting in normalizing the Minkowski distance by using the IOWA operator. Then, the normalization developed will be able to reflect complex attitudinal characters. We will call this generalization as the induced Minkowski OWA distance (IMOWAD) operator. The main advantage of this operator is that it generalizes a wide range of distances such as the NMD, the WMD, the MOWAD [5,9], the induced OWA distance (IOWAD) [7], the induced Euclidean OWA distance (IEOWAD) [8] and a lot of other particular cases. Then, by using this generalization, we get a more complete view of the aggregation process of the individual distances.

We further generalize the IMOWAD operator by using quasi-arithmetic means. As a result, we get the Quasi-IOWAD operator. We also develop an application of the new approach in a decision making problem about selection of investments. We will see that depending on the particular type of IMOWAD operator used, the results may lead to different decisions.

This paper is organized as follows. In Section 2, we briefly review some basic concepts about the Minkowski distance and the IOWA operator. In Section 3 we present the IMOWAD operator. Section 4 analyzes different families of IMOWAD operators. In Section 5 we present the Quasi-IOWAD operator. Section 6 develops a numerical example of the new generalization. Finally, in Section 7 we summarize the main conclusions.

2 PRELIMINARIES

2.1. NORMALIZED MINKOWSKI DISTANCE

The normalized Minkowski distance is a distance measure that generalizes a wide range of distances such as the normalized Hamming distance and the normalized Euclidean distance. In fuzzy set theory, it can be useful, for example, for the calculation of distances between fuzzy sets, interval-valued fuzzy sets, intuitionistic fuzzy sets, etc. It can be formulated for two sets *A* and *B* as follows.

Definition 1. *A normalized Minkowski distance of dimension n is a mapping* d_m*:* $R^n \times R^n \to R$ *such that*:

$$d_m(A,B) = \left(\frac{1}{n} \sum_{i=1}^{n} | a_i - b_i |^{\lambda} \right)^{1/\lambda} \quad (1)$$

where a_i *and* b_i *are the ith arguments of the sets A and B and* λ *is a parameter such that* $\lambda \in (-\infty, \infty)$.

If we give different values to the parameter λ, we can obtain a wide range of special cases. For example, if λ = 1, we obtain the normalized Hamming distance (NHD). If λ = 2, the normalized Euclidean distance (NED).

Sometimes, when normalizing the Minkowski distance, we prefer to give different weights to each individual distance. Then, the distance is known as the weighted Minkowski distance. For two sets *A* and *B*, it can be defined as follows.

Definition 2. *A weighted Minkowski distance of dimension n is a mapping* d_{wm}*:* $R^n \times R^n \to R$ *that has an associated weighting vector W of dimension n such that the sum of the weights is* 1 *and* $w_j \in [0,1]$. *Then*:

$$d_{wm}(A,B) = \left(\sum_{i=1}^{n} w_i \, | a_i - b_i |^{\lambda} \right)^{1/\lambda} \quad (2)$$

where a_i *and* b_i *are the ith arguments of the sets A and B and* λ *is a parameter such that* $\lambda \in (-\infty, \infty)$.

In this case, we can also obtain a wide range of special cases by using different values in the parameter λ. For example, if λ = 1, we obtain the weighted Hamming distance (WHD). If λ = 2, the weighted Euclidean distance (WED). Note that if $\lambda \leq 0$, then, we get similar results as in the generalized mean. For example, if λ = 1, we get the weighted harmonic distance and if λ = 0, we obtain some kind of weighted geometric distance. The problem when using $\lambda \leq 0$, is that sometimes the results are inconsistent. Especially, this problem appears in the weighted geometric distance because if one of the individual distances is 0, then, the whole aggregation is also 0.

2.2. INDUCED OWA OPERATOR

The IOWA operator was introduced by Yager and Filev [16] and it is an extension of the OWA operator. Its main difference is that the reordering step is not developed with the values of the arguments a_i. In this case, the reordering step is developed with order inducing variables. It can be defined as follows.

Definition 3. *An IOWA operator of dimension n is a mapping* f*:* $R^n \to R$ *that has an associated weighting vector W of dimension n such that the sum of the weights is* 1 *and* $w_j \in [0,1]$, *and its argument variable consists of n-tuple of pairs, then*:

$$f(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle \ldots, \langle u_n,a_n\rangle) = \sum_{j=1}^{n} w_j b_j \quad (3)$$

where b_j *is the* a_i *value of the IOWA pair* $\langle u_i,a_i\rangle$ *having the jth largest* u_i, u_i *is the order inducing variable and* a_i *is the argument variable.*

The IOWA operator includes the OWA operator as a particular case and a lot of other situations such as the maximum, the minimum and the average. Note that it is possible to distinguish between the Descending IOWA (DIOWA) operator and the Ascending IOWA (AIOWA) operator.

3 THE INDUCED MINKOWSKI ORDERED WEIGHTED AVERAGING DISTANCE OPERATOR

The IMOWAD operator is a distance measure that uses the IOWA operator in the normalization process of the Minkowski distance. Then, the reordering of the individual distances is developed with order inducing variables. For two sets *X* and *Y*, it can be defined as follows.

Definition 4. *An IMOWAD operator is a mapping* f*:* $R^n \times R^n \to R$ *that has an associated weighting vector W such that* $w_j \in [0, 1]$ *and the sum of the weights is* 1, *then*:

$$f(\langle u_1, x_1, y_1\rangle,\ldots,\langle u_n, x_n, y_n\rangle) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \quad (4)$$

where b_j is the $|x_i - y_i|$ value of the IMOWAD triplet $\langle u_i, x_i, y_i\rangle$ having the jth largest u_i, u_i is the order inducing variable and $|x_i - y_i|$ is the argument variable represented in the form of individual distances and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

Remark 1: The IMOWAD operator is commutative, monotonic, bounded and idempotent. It is commutative because any permutation of the arguments has the same evaluation. That is, $f(\langle u_1, x_1, y_1\rangle, \langle u_2, x_2, y_2\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = f(\langle u_1, s_1, v_1\rangle, \langle u_2, s_2, v_2\rangle, \ldots, \langle u_n, s_n, v_n\rangle)$ where $(\langle u_1, x_1, y_1\rangle, \langle u_2, x_2, y_2\rangle, \ldots, \langle u_n, x_n, y_n\rangle)$ is any permutation of the arguments $(\langle u_1, s_1, v_1\rangle, \langle u_2, s_2, v_2\rangle, \ldots, \langle u_n, s_n, v_n\rangle)$. It is monotonic because if $|x_i - y_i| \geq |s_i - v_i|$, for all i, then, $f(\langle u_1, x_1, y_1\rangle, \langle u_2, x_2, y_2\rangle, \ldots, \langle u_n, x_n, y_n\rangle) \geq f(\langle u_1, s_1, v_1\rangle, \langle u_2, s_2, v_2\rangle, \ldots, \langle u_n, s_n, v_n\rangle)$. It is bounded because the IMOWAD aggregation is delimited by the minimum and the maximum. That is, $\text{Min}\{|x_i - y_i|\} \leq f(\langle u_1, x_1, y_1\rangle, \langle u_2, x_2, y_2\rangle, \ldots, \langle u_n, x_n, y_n\rangle) \leq \text{Max}\{|x_i - y_i|\}$. It is idempotent because if $|x_i - y_i| = |x - y|$, for all i, then, $f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = |x - y|$.

Remark 2: Note that if $x_i = y_i$ for all $i \in [1, n]$, $f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = 0$. Note also that $f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = f(\langle u_1, y_1, x_1\rangle, \ldots, \langle u_n, y_n, x_n\rangle)$.

Remark 3: From a generalized perspective of the reordering step it is possible to distinguish between descending (DIMOWAD) and ascending (AIMOWAD) orders. The weights of these operators are related by $w_j = w^*_{n+1-j}$, where w_j is the jth weight of the DIMOWAD (or IMOWAD) operator and w^*_{n+1-j} the jth weight of the AIMOWAD operator.

Remark 4: If B is a vector corresponding to the ordered arguments b_j^{λ}, we shall call this the ordered argument vector and W^T is the transpose of the weighting vector, then, the IMOWAD operator can be expressed as:

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \left(W^T B\right)^{1/\lambda} \quad (5)$$

Remark 5: Note that if the weighting vector is not normalized, i.e., $W = \sum_{j=1}^{n} w_j \neq 1$, then, the IMOWAD operator can be expressed as:

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \left(\frac{1}{W}\sum_{j=1}^{n} w_j b_j^{\lambda}\right)^{1/\lambda} \quad (6)$$

Remark 6: A further interesting issue is the problem of ties in the reordering process of the order inducing variables. In order to solve this problem, we recommend to follow the policy explained in [16] about replacing the tied arguments by their average. Note that in this case, it would mean that we are replacing the tied arguments by their normalized Minkowski distance.

4 FAMILIES OF IMOWAD OPERATORS

In this Section, we analyze different particular cases of the IMOWAD operator. We distinguish between those families found in the parameter λ and those found in the weighting vector W.

4.1. ANALYSING THE PARAMETER λ

By looking at the parameter λ, we can find a wide range of distance measures such as the IOWAD, the EIOWAD, the induced ordered weighted geometric distance (IOWGD) operator, the induced ordered weighted harmonic averaging distance (IOWHAD) operator and a lot of other cases. Note that in the IMOWAD we also find similar properties than the WMD when $\lambda \leq 0$. Therefore, we may find some inconsistencies, especially in the IOWGD operator.

Remark 7: When $\lambda = 1$, the IMOWAD operator becomes the IOWAD operator.

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \sum_{j=1}^{n} w_j b_j \quad (7)$$

Note that if $w_j = 1/n$, for all a_i, we get the NHD. The WHD is obtained if $u_i > u_{i+1}$, for all i, and the OWAD operator is obtained if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of $|x_i - y_i|$. Note also that it is possible to distinguish between descending (DIOWAD) and ascending (AIOWAD) orders.

Remark 8: When $\lambda = 2$, we get the IEOWAD operator.

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \left(\sum_{j=1}^{n} w_j b_j^2\right)^{1/2} \quad (8)$$

If $w_j = 1/n$, for all a_i, we get the NED. If $u_i > u_{i+1}$, for all i, we get the WED and if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of $|x_i - y_i|^2$, we get the EOWAD operator. In this case, we also get the descending IEOWAD (DIEOWAD) operator and the ascending IEOWAD (AIEOWAD) operator.

Remark 9: When $\lambda = 0$, we get the IOWGD operator.

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \prod_{j=1}^{n} b_j^{w_j} \quad (9)$$

Note that if $w_j = 1/n$, for all a_i, we get the normalized geometric distance and if $u_i > u_{i+1}$, for all i, the weighted geometric distance. If the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of $|x_i - y_i|^{w_j}$, we get the ordered weighted geometric distance operator (OWGD) operator. In this case, it is also possible to distinguish between descending (DIOWGD) and ascending (AIOWGD) orders. Note that the IOWGD can only be used sometimes when all the individual distances are different from 0. That is, when $|x_i - y_i| \neq 0$, for all i.

Remark 10: When $\lambda = -1$, we get the IOWHAD operator.

$$f(\langle u_1, x_1, y_1\rangle, \ldots, \langle u_n, x_n, y_n\rangle) = \frac{1}{\sum_{j=1}^{n} \frac{w_j}{b_j}} \quad (10)$$

Note that if $w_j = 1/n$, for all a_i, we get the normalized harmonic distance. If $u_i > u_{i+1}$, for all i, we get the weighted harmonic distance. If the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of $1/(w_j / |x_i - y_i|)$, we get the ordered weighted harmonic averaging distance operator (OWHAD) operator . In this case, we can also distinguish between the descending IOWHA (DIOWHAD) operator and the ascending IOWHAD (AIOWHAD) operator.

Remark 11: Note that we could analyze other families by using different values in the parameter λ. Also note that it is possible to study these families individually in a similar way as it has been developed in Section 3.

4.2. ANALYSING THE WEIGHTING VECTOR *W*

By choosing a different manifestation of the weighting vector in the IMOWAD operator, we are able to obtain different types of distance aggregation operators. For example, we can obtain the maximum distance, the minimum distance, the NMD, the WMD and the MOWAD operator.

The maximum distance is obtained if $w_p = 1$ and $w_j = 0$, for all $j \neq p$, and $u_p = \text{Max}\{a_i\}$. The minimum distance is obtained if $w_p = 1$ and $w_j = 0$, for all $j \neq p$, and $u_p = \text{Min}\{a_i\}$. More generally, if $w_k = 1$ and $w_j = 0$, for all $j \neq k$, we are using the step-IMOWAD operator. The NMD is found when $w_j = 1/n$, for all a_i. The WMD is obtained if $u_i > u_{i+1}$, for all i, and the MOWAD operator is obtained if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i.

Remark 12: Other families of IMOWAD operators could be used. For more information on the methodology of these families, see for example [6,10,13]. For example, when $w_j = 1/m$ for $k \leq j \leq k + m - 1$ and $w_j = 0$ for $j > k + m$ and $j < k$, we are using the window-IMOWAD operator. Note that k and m must be positive integers such that $k + m - 1 \leq n$.

Remark 13: If $w_1 = w_n = 0$, and for all others $w_j = 1/(n - 2)$, we are using the olympic-IMOWAD. Note that if $n = 3$ or $n = 4$, the olympic-IMOWAD is transformed in the IMOWAD median and if $m = n - 2$ and $k = 2$, the window-IMOWAD is transformed in the olympic-IMOWAD. Also note that the olympic-IMOWAD is transformed in the olympic-MOWAD average if $w_p = w_q = 0$, such that $u_p = \text{Max}_i\{a_i\}$ and $u_q = \text{Min}_i\{a_i\}$, and for all others $w_j = 1/(n - 2)$.

Remark 14: Note that the IMOWAD-median and the weighted IMOWAD-median can also be used as a particular case of the IMOWAD. For the IMOWAD median, if n is odd we assign $w_{(n+1)/2} = 1$ and $w_j = 0$ for all others, and this affects the argument a_i with the $[(n + 1)/2]$th largest u_i. If n is even we assign for example, $w_{n/2} = w_{(n/2)+1} = 0.5$, and this affects the arguments with the $(n/2)$th and $[(n/2) + 1]$th largest u_i. For the weighted IMOWAD median, we select the argument a_i that has the kth largest inducing variable u_i, such that the sum of the weights from 1 to k is equal or higher than 0.5 and the sum of the weights from 1 to $k - 1$ is less than 0.5. Note that if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i, then, we get the MOWAD-median and the weighted MOWAD-median, respectively.

Remark 15: Another interesting family is the S-IMOWAD operator. It can be divided in three classes, the "orlike", the "andlike" and the generalized S-IMOWAD operator. The generalized S-IMOWAD operator is obtained when $w_p = (1/n)(1 - (\alpha + \beta)) + \alpha$, with $u_p = \text{Max}\{a_i\}$; $w_q = (1/n)(1 - (\alpha + \beta)) + \beta$, with $u_q = \text{Min}\{a_i\}$; and $w_j = (1/n)(1 - (\alpha + \beta))$ for all $j \neq p,q$ where $\alpha, \beta \in [0, 1]$ and $\alpha + \beta \leq 1$. Note that if $\alpha = 0$, the generalized S-IMOWAD operator becomes the "andlike" S-IMOWAD operator and if $\beta = 0$, it becomes the "orlike" S-IMOWAD operator.

Remark 16: A further interesting family that could be used is the centered-IMOWAD operator. An IMOWAD operator is a centered aggregation operator if it is symmetric, strongly decaying and inclusive. It is symmetric if $w_j = w_{j+n-1}$. It is strongly decaying when $i < j \leq (n + 1)/2$ then $w_i < w_j$ and when $i > j \geq (n + 1)/2$ then $w_i < w_j$. It is inclusive if $w_j > 0$. Note that it is possible to consider a softening of the second condition by using $w_i \leq w_j$ instead of $w_i < w_j$. We shall refer to this as softly decaying centered-IMOWAD operator. Another particular

situation of the centered-IMOWAD operator appears if we remove the third condition. We will refer to it as a non-inclusive centered-IMOWAD operator.

5 QUASI-IOWAD OPERATOR

The IMOWAD can be generalized by using quasi-arithmetic means in a similar way as it was done in [1-4,6]. We will call it the Quasi-IOWAD operator. It is defined as follows.

Definition 5. *A Quasi-IOWAD operator is a mapping f*: $R^n \times R^n \rightarrow R$ *that has an associated weighting vector W such that* $w_j \in [0, 1]$ *and the sum of the weights is* 1, *then*:

$$f(\langle u_1, x_1, y_1\rangle,\ldots,\langle u_n, x_n, y_n\rangle) = g^{-1}\left(\sum_{j=1}^{n} w_j\, g\left(b_{(j)}\right)\right) \qquad (11)$$

where b_j *is the* $|x_i - y_i|$ *value of the QIOWAD triplet* $\langle u_i, x_i, y_i\rangle$ *having the jth largest* u_i, u_i *is the order inducing variable,* $|x_i - y_i|$ *is the argument variable represented in the form of individual distances, and g is the strictly continuous monotonic function.*

As we can see, when $g(b) = b^{\lambda}$, then, the Quasi-IOWAD becomes the IMOWAD operator. Note that it is also possible to distinguish between descending (Quasi-DIOWAD) and ascending (Quasi-AIOWAD) orders.

Remark 17: Note that all the properties and particular cases commented in the IMOWAD operator are also applicable in the Quasi-IOWAD operator such as the distinction between descending and ascending orders, different families of Quasi-IOWAD operators, the normalization process of the weighting vector when it is different from 1 by using (1 / *W*), etc.

6 NUMERICAL EXAMPLE

In the following, we are going to develop an illustrative example in order to see the results obtained in the aggregation by using different types of IMOWAD operators. We will analyze the selection of investments where an enterprise is looking for the best strategy according to its interests.

Assume that an enterprise that operates in Europe and North America wants to invest some money the next year. In order to do this, the board of directors has established four possible investments S_i that the enterprise could develop in the future.

(1) A_1: Invest in the Asian market.
(2) A_2: Invest in the South American market.
(3) A_3: Invest in the African market.
(4) A_4: Do not develop any investment.

After careful review of the information, the experts have given the following general information. They have summarized the information of the strategies in five main characteristics C_i with the following results.

(1) C_1: Risk of the investment.
(2) C_2: Benefits in the short term.
(3) C_3: Benefits in the long term.
(4) C_4: Difficulty of the investment.
(5) C_5: Other aspects.

Note that the results are valuations between 0 and 1.

Table 1: Available investments

	C_1	C_2	C_3	C_4	C_5
A_1	0.7	0.9	0.8	0.7	0.3
A_2	0.6	0.8	0.7	0.5	0.8
A_3	0.5	0.6	0.8	0.4	0.9
A_4	0.8	0.5	0.6	0.8	0.6

According to the objectives and policies of the enterprise, the experts have established the ideal investment for the company independently of the investments available. They have established the following valuations for it.

Table 2: Characteristics of the ideal investment

	C_1	C_2	C_3	C_4	C_5
I	0.8	0.9	1	0.9	0.9

In order to aggregate the information, the group of experts calculates the attitudinal character of the enterprise. Due to the fact that the attitudinal character depends upon the opinion of several members of the board of directors, it is very complex. Therefore, they to use order inducing variables in the reordering process. The results are shown in Table 3.

Table 3: Order inducing variables

	C_1	C_2	C_3	C_4	C_5
A_1	12	16	20	24	8
A_2	22	18	20	26	28
A_3	14	20	15	18	17
A_4	26	21	19	15	13

With this information, it is possible to develop different methods for selecting an investment. In this example, we will consider the NHD, the NED, the WHD, the WED, the OWAD, the IOWAD, the AIOWAD and the EIOWAD operator. Note that the weighting vector used

is: W = (0.1, 0.2, 0.2, 0.2, 0.3). The aggregated results are shown in Tables 4 and 5.

Table 4: Aggregated results 1

	NHD	WHD	OWAD	IOWAD
A_1	0.22	0.27	0.16	0.26
A_2	0.22	0.21	0.19	0.22
A_3	0.26	0.23	0.21	0.26
A_4	0.24	0.27	0.2	0.27

Table 5: Aggregated results 2

	AIOWAD	EIOWAD	Median	Olympic
A_1	0.18	0.349	0.2	0.1
A_2	0.22	0.249	0.2	0.3
A_3	0.26	0.306	0.3	0.233
A_4	0.21	0.305	0.3	0.3

As we can see, depending on the distance aggregation operator used, the optimal choice is different. Note that the lowest value in each method is the optimal result.

If we establish an ordering of the investments, a typical situation if we want to consider more than one alternative, we will get the following orders shown in Table 6. Note that the first alternative in each ordering is the optimal choice.

Table 6: Ordering of the investments

	Ordering
NHD	$A_3 \succ A_4 \succ A_1 = A_2$
WHD	$A_1 = A_4 \succ A_3 \succ A_2$
OWAD	$A_3 \succ A_4 \succ A_2 \succ A_1$
IOWAD	$A_4 \succ A_1 \succ A_3 \succ A_2$
AIOWAD	$A_3 \succ A_2 \succ A_4 \succ A_1$
EIOWAD	$A_1 \succ A_3 \succ A_4 \succ A_2$
Median-IOWAD	$A_3 = A_4 \succ A_1 = A_2$
Olympic-IOWAD	$A_2 = A_4 \succ A_3 \succ A_1$

As we can see, depending on the particular type of IMOWAD operator used, the results may lead to different decisions.

7 CONCLUSIONS

We have presented the IMOWAD operator. It is a distance measure that uses the IOWA operator in the Minkowski distance. The main advantage of this operator is that it generalizes a wide range of distances such as the NMD, the WMD, the MOWAD, the IOWAD, the EIOWAD, etc. We have studied some of its main properties.

We have further generalized the IMOWAD operator by using quasi-arithmetic means. We have called it the Quasi-IOWAD. We have also developed an application of the new approach in a decision making problem about selection of investments. We have seen that the main advantage of using the IMOWAD is that it gives a more complete view of the decision problem.

In future research, we expect to develop further extensions of the IMOWAD operator by adding new characteristics in the problem and applying it to other fields.

Acknowledgements

We would like to thank the anonymous referees for their insightful comments that have led to an improved version of the paper.

References

[1] G. Beliakov. Learning Weights in the Generalized OWA Operators, *Fuzzy Optimization and Decision Making* 4, Pág. 119-130, 2005.

[2] G. Beliakov, A. Pradera, T. Calvo. *Aggregation Functions: A Guide for Practitioners*, Springer-Verlag, Berlin, 2007.

[3] T. Calvo, G. Mayor, R. Mesiar. *Aggregation Operators: New Trends and Applications*, Physica-Verlag, New York, 2002.

[4] J. Fodor, J.L. Marichal, M. Roubens. Characterization of the ordered weighted averaging operators, *IEEE Transactions on Fuzzy Systems* 3, Pág. 236-240, 1995.

[5] N. Karayiannis. Soft Learning Vector Quantization and Clustering Algorithms Based on Ordered Weighted Aggregation Operators, *IEEE Transactions on Neural Networks* 11, Pág. 1093-1105, 2000.

[6] J.M. Merigó. *New extensions to the OWA operators and their application in business decision making*, Unpublished thesis (in Spanish), Department of Business Administration, University of Barcelona, 2007.

[7] J.M. Merigó, M. Casanovas. Decision making with distance measures and induced aggregation operators, In *Proceedings of the FLINS 2008 Conference*, Madrid, Spain, (accepted).

[8] J.M. Merigó, M. Casanovas. Induced aggregation operators in the Euclidean distance, In *Proceedings of the FUR 2008 Conference*, Barcelona, Spain, (accepted).

[9] J.M. Merigó, A.M. Gil-Lafuente. Using the OWA operator in the Minkowski distance, International *Journal of Computer Science*, (accepted).

[10] Z.S. Xu. An overview of methods for determining OWA weights, *International Journal of Intelligent Systems* 20, Pág. 843-865, 2005.

[11] Z.S. Xu, Q.L. Da. An overview of operators for aggregating the information, *International Journal of Intelligent Systems* 18, Pág. 953-969, 2003.

[12] R.R. Yager. On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making, *IEEE Transactions on Systems, Man and Cybernetics B* 18, Pág. 183-190, 1988.

[13] R.R. Yager. Families of OWA operators, *Fuzzy Sets and Systems* 59, Pág. 125-148, 1993.

[14] R.R. Yager. Induced aggregation operators, *Fuzzy Sets and Systems* 137, Pág. 59-69, 2003.

[15] R.R. Yager. Generalized OWA Aggregation Operators, *Fuzzy Optimization and Decision Making* 3, Pág. 93-107, 2004.

[16] R.R. Yager, D.P. Filev. Induced ordered weighted averaging operators, *IEEE Transactions on Systems, Man and Cybernetics B* 29, Pág. 141-150, 1999.

[17] R.R. Yager, J. Kacprzyk. *The Ordered Weighted Averaging Operators: Theory and Applications*, Kluwer Academic Publishers, Norwell, 1997.

Ajuste Genético de Controladores Difusos para Conducción Autónoma

Enrique Onieva Vicente Milanés Javier Alonso Joshué Pérez Teresa de Pedro

Instituto de Automática Industrial. Consejo Superior de Investigaciones Científicas.
Ctra. Campo Real Km.0,200. 28500 La Poveda, Arganda del Rey. Madrid, Espaa.
{onieva, vmilanes, jalonso, jperez, tere}@iai.csic.es

Resumen

Existe una amplia gama de tecnologías creadas con el fin de dar solución a los problemas de transporte, llamados Sistemas Inteligentes de Transporte; el control automático del volante es uno de los retos más importantes al que se enfrentan investigadores de esta área, pero no es el único. Este trabajo presenta un método para el ajuste automático de controladores difusos capaces de manejar automáticamente el volante de un vehículo convencional, reproduciendo actitudes de un conductor humano en diferentes situaciones. Para hacerlo, se ha recopilado información sobre el estado del coche mientras es conducido por un piloto, gracias a la cual, y por medio de algoritmos genéticos se han obtenido controladores difusos apropiados para el manejo del volante del vehículo. Los dos principales requisitos que debe llevar a cabo el controlador son: ser capaces de manejar el vehículo a través de una ruta prefijada y ejecutar acciones de control suaves, para lograr así, una conducción confortable.

1 INTRODUCCIÓN

Los Sistemas Inteligentes de Transporte (ITS, Intelligent Transportation Systems) usan tecnologías de la información y la comunicación con el fin de conseguir una conducción segura y eficiente. En la industria del automóvil, los sensores se usan principalmente para dar información al conductor y, en algunos casos, están conectados a un ordenador, que ejecutará determinadas acciones de control con el fin de minimizar daos y prevenir colisiones [1].

Actualmente, la implementación del control total del volante de un vehículo es una de las disciplinas de la investigación industrial en vehículos inteligentes que está recibiendo menor atención ahora que en aos anteriores. La razón quizás se encuentre en que queda un muy largo camino hasta que esto llegue al mercado [5], y los fabricantes de vehículos se centran en etapas más maduras, especialmente en la automatización del control de velocidad; algunos ejemplos pueden encontrarse hoy en día en el mercado. Pero no recibe tan poca atención en el área de la investigación académica, tal y como se demuestra en el DARPA Grand Challenge [6], que consiste en una competición de vehículos autónomos financiada por la Agencia de Proyectos de Investigación Avanzada en Defensa (DARPA), la organización central de investigación del Departamento de Defensa de los Estados Unidos. DARPA ha financiado tres competiciones en el área de vehículos autónomos hasta la fecha.

El trabajo llevado a cabo por el Programa AUTOPIA, del Instituto de Automática Industrial, del Consejo Superior de Investigaciones Científicas está centrado en el área de vehículos autónomos. Concretamente, el objetivo del trabajo aquí presentado se centrará en el control automático del volante, también conocido como control lateral, de un vehículo convencional; otras áreas de investigación abordadas por el Programa AUTOPIA son el control longitudinal, el control de la velocidad, y su adaptación a las condiciones de la carretera, mediante la automatización de los pedales de freno y acelerador [3], [4].

El trabajo pretende modelar el comportamiento de un conductor humano mediante el ajuste de un controlador borroso [15] mediante un algoritmo genético [16]. En la sección 2 se presenta, de una manera muy breve, el Programa AUTOPIA, con el objetivo de dar una idea sobre la instrumentación e infraestructura usada para llevar a cabo este trabajo; la sección 3 se presenta el método usado para generar automáticamente controladores difusos, tomando como entrada, una se-

rie de puntos Entrada/Salida y aplicando un método iterativo, consistente en la optimización de la base de datos y de reglas que definen un controlador por separado por medio de algoritmos genéticos. Finalmente, en la sección 4, el método es aplicado para modelar el comportamiento de un conductor en el uso del volante de un vehículo. Una vez generados una serie de controladores difusos, éstos serán probados en una pista de pruebas, con lo que podremos verificar su correcto funcionamiento.

2 EL PROGRAMA AUTOPIA

AUTOPíA pretende transferir las técnicas desarrolladas para el control de robots autónomos al control de vehículos, modificando en la menor medida posible el entorno en que éstos han de evolucionar. Se han realizado inicialmente experimentos simples que muestren los resultados que van obteniéndose y creado la infraestructura que permita llegar a hacer cada vez unas maniobras más sofisticadas.

El Coche y ZOCO: El coche usado en este trabajo es un Citron C3 Pluriel. ZOCO (ZOna de COnducción) es una pista de pruebas de vehículos automáticos que está dedicada exclusivamente a tareas de investigación, es decir, en ella no hay ningún otro tráfico de vehículos, lo que se ha hecho por seguridad. Tiene una forma reticulada, como las manzanas o cuadras de una ciudad, con algunas irregularidades. También posee ZOCO una estación base de posicionamiento global diferencial basado en información geográfica de información vía satélite, DGPS, que puede ser utilizado por sistemas móviles embarcados en los coches, para obtener su posición con una precisión cercana al centímetro. En la figura 1 podemos ver una fotografía del coche utilizado, así como la vista aérea de ZOCO puede verse en la figura 2.

Figura 1: Vehículo Utilizado

Arquitectura de Control: La arquitectura general usa un esquema donde un conjunto de coches independientes se comunican entre sí mediante una estación central, gracias a lo que comparten información con

Figura 2: Vista aérea de ZOCO

el fin de cooperar para llevar a cabo maniobras tal y como son realizadas por humanos. La arquitectura individual instalada en cada vehículo es una arquitectura jerárquica en tres capas: sensorial, de control y de actuación. La capa de control está dividida en tres etapas secuenciales: planificador, copiloto y piloto. En el presente trabajo, nos centraremos en el piloto, ya que es donde se encuentran y ejecutan los controladores difusos capaces de llevar a cabo maniobras simples. El piloto recibe una serie de parámetros de entrada y una maniobra de bajo nivel a realizar (seguir al coche de delante, tomar una curva, parar...) provenientes del copiloto y es capaz de generar una salida aplicable sobre los actuadores del vehículo (pedales y volante).

ORBEX: ORBEX [7], [8] (ORdenador Borroso EXperimental) se usa para la construcción y ejecución de los controladores difusos usados por el piloto. ORBEX trabaja con controladores difusos del tipo TSK [10], con consecuentes Sigleton, lo que permite tomar decisiones de control en un corto período de tiempo y con muy buena precisión; cualidades más que deseables para sistemas de tiempo real, donde el tiempo de inferencia es un aspecto vital, tal y como puede ser la conducción de un coche [9],[11].

3 MÉTODO ITERATIVO PARA EL AJUSTE DE CONTROLADORES BORROSOS

La base de reglas y las funciones de pertenencia que definen un controlador borroso serán modificadas con el fin de adaptarse a las acciones y actitudes mostradas por un conductor humano a la hora de conducir un vehículo. El método se divide en tres fases; primero, se aplica una fase de captura y procesamiento de información; tras lo que, de una forma iterativa, ajustamos las funciones de pertenencia y la base de reglas por medio de dos algoritmos genéticos independientes. Las siguientes cuatro subsecciones describen la fase de captura y procesado de datos, los dos algoritmos genéticos utilizados y el esquema global seguido por el método con el fin de ajustar los controladores difusos.

3.1 CAPTURA Y PROCESADO DE LA INFORMACIÓN

Un conductor humano controlará el vehículo a través de una ruta prefijada, mientras el ordenador de a bordo tomará datos acerca del estado del vehículo y maniobras realizadas. Los datos que pueden ser recogidos por el ordenador pueden ser, por ejemplo: velocidad, posición, giro del volante, presión aplicada sobre los pedales...

Una vez hecho esto, se obtiene un conjunto de datos que será sometido a un proceso de normalización en el intervalo [-1,1], tras lo que se le aplicará una rejilla de dimensiones elegidas por el usuario y, en cada punto, la salida considerada será la media entre los valores más cercanos del conjunto de datos inicial. Por otra parte, existe la posibilidad de aadir nuevos puntos no incluidos en el conjunto de datos inicial, ya que pueden representar zonas del espacio de las que no es posible tener información. Lo que permitirá al usuario introducir información proveniente de su propio conocimiento, tal como *Si el coche esta desviado lo máximo hacia la derecha, gira el volante lo máximo hacia la izquierda.*

3.2 AJUSTE DE LAS FUNCIONES DE PERTENENCIA

En la primera etapa del método, las funciones de pertenencia (FP's) que definen las variables de entrada del controlador difuso serán ajustadas. Una función de pertenencia se definirá mediante cuatro valores reales (a,b,c,d), lo que permitirá trabajar con etiquetas tanto trapezoidales como triangulares (b=c). Con el fin de obtener controladores difusos precisos e interpretables, las funciones de pertenencia deberán cumplir una serie de restricciones semánticas tales como: las funciones de pertenencia deberán estar ordenadas; todos los posibles valores de entrada deben estar cubiertos por una o dos funciones de pertenencia...

Así pues, para representar un conjunto de funciones de pertenencia trapezoidales de forma que pueda ser mejorado por medio de un algoritmo genético, usaremos 4N valores, donde N es el número total de funciones de pertenencia de las variables de entrada. Las variables de salida no serán modificadas. El algoritmo genético estacionario usado para ajustar las funciones de pertenencia usará los siguientes parámetros:

- Inicialización: cada valor v_i de un cromosoma será inicializado a un valor uniformemente distribuido en el intervalo $[m_i \pm \xi]$ con una probabilidad ρ, en otro caso, $v_i = m_i$. Donde m representará el mejor cromosoma encontrado a lo largo de la ejecución. Es decir, la población inicial en cada iteración se creará a partir del mejor individuo encontrados hasta el momento, introduciendo sobre sus valores de las funciones de pertenencia una pequea variación aleatoria.
- Selección: Torneo Binario.
- Operador de Cruce: BLX - α [13].
- Operador de Mutación: Modificación aleatoria bajo una probabilidad de mutación, p_m [11], [12].
- Reemplazamiento: cada nuevo cromosoma reemplazará al peor cromosoma de la población, siempre que obtenga mejor fitness que él.
- Función Objetivo: Para medir la calidad de un determinado controlador, serán tenidos en cuenta dos aspectos: el ajuste a las acciones tomadas por el piloto y tener una superficie de control suave, para no tener cambios de salida bruscos. Para evaluar el primero de ellos, se usará el error cuadrático medio (MSE) entre la salida del controlador y las salidas dadas por el piloto. Para evaluar el segundo, la superficie de control será evaluada bajo una rejilla de 21x21, y se tomará D como la mayor diferencia obtenida entre dos puntos adyacentes de la rejilla. Finalmente, la función objetivo consistirá en una ponderación de las dos medidas comentadas (MSE y D), siendo el valor final de la siguiente forma: F = vMSE + (1-v)D, donde v es elegido por el usuario con el fin de dar más importancia al aspecto deseado.

3.3 AJUSTE DE LA BASE DE REGLAS

La segunda etapa del método consiste en el ajuste de la base de reglas (BR). Para hacerlo, se modificarán los consecuentes asociados a una serie de conjuntos de antecedentes dados. El proceso comienza con la selección, por parte del usuario del número de posibles singleton de Sugeno que podrán ser seleccionados como consecuentes de una determinada regla. Una vez hecho esto, se crea el conjunto de singletons uniformemente distribuido en el rango de la salida y el objetivo del algoritmo genético consistirá en asociar cada posible antecedente con un consecuente, con el fin de obtener la mejor configuración de reglas para el sistema. Independientemente del número de funciones de pertenencia usado para cada una de las variables de entrada, el programa dispone de tres conjuntos de antecedentes prefijados, llamados *Marginal*, *Central* y *Total* respectivamente. El primero de ellos está formado por reglas con antecedente simple, el segundo por reglas compuestas mediante el operador 'Y' y el tercero por la unión de los dos anteriores.

Por ejemplo, si tuviéramos dos entradas, cada una de ellas con 4 etiquetas diferentes, los casos cubiertos por

cada uno de los conjuntos de antecedentes podemos verlo en la tabla 1:

$In_1 \setminus In_2$	$\emptyset$	Et_{21}	Et_{22}	Et_{23}	Et_{24}
$\emptyset$		M	M	M	M
Et_{11}	M	C	C	C	C
Et_{12}	M	C	C	C	C
Et_{13}	M	C	C	C	C
Et_{14}	M	C	C	C	C

Tabla 1: Casos cubiertos por la base de reglas Central (C) y Marginal (M).

De esta forma, el conjunto de antecedentes Marginal almacena todos los antecedentes del tipo:

```
Si (In1 = Et11) Entonces ...
Si (In1 = Et12) Entonces ...
...
Si (In2 = Et23) Entonces ...
Si (In2 = Et24) Entonces ...
```

Mientras que el conjunto de antecedentes Central almacena todos los del tipo:

```
Si (In1 = Et11 ) Y (In2 = Et21) Entonces ...
Si (In1 = Et11 ) Y (In2 = Et22) Entonces ...
...
Si (In1 = Et14 ) Y (In2 = Et23) Entonces ...
Si (In1 = Et14 ) Y (In2 = Et24) Entonces ...
```

Como hemos comentado, el conjunto de antecedentes Total, representará la unión de los conjuntos Marginal y Central. Dicho esto, para representar la base de reglas de forma que pueda ser tratada por un algoritmo genético, usaremos un vector de N enteros, donde N es el número total de reglas de la base seleccionada y en cada posición i del vector se almacenará el índice del consecuente asociado a la regla i. Para el ajuste de la base de reglas, nuevamente usaremos un algoritmo genético estacionario con igual selección, reemplazamiento, y función objetivo al usado para el ajuste de las funciones de pertenencia, pero con distinto operador de cruce; en este caso, usaremos cruce simple [12]. En el caso de la inicialización, nuevamente partiremos del mejor individuo encontrado hasta el momento y dos valores (ξ,ρ); así pues, para generar una población inicial, cada valor m_i del mejor individuo será modificado, con una probabilidad ρ a un valor aleatorio en el intervalo $[m_i \pm \xi]$.

3.4 ESQUEMA DEL MÉTODO

La ejecución del método se hace en dos fases diferenciadas, que serán repetidas iterativamente un cierto número de veces y representarán, respectivamente, los dos algoritmos genéticos comentados con anterioridad; uno de ellos estará encargado de mejorar las funciones de pertenencia, mientras que el otro, de la base de reglas. Gracias a esto conseguimos que, cada vez que obtengamos un buen conjunto de etiquetas, poder encontrar el conjunto de reglas más apropiado para él, al igual que, si encontramos un buen conjunto de reglas, la otra fase del método buscará el conjunto de etiquetas más apropiado. Así, mediante dos *pequeos* algoritmos genéticos independientes, ejecutados iterativamente, logramos la progresiva mejora conjunta de base de reglas y conjunto de etiquetas. Para evaluar la base de reglas se usará el mejor conjunto de funciones de pertenencia encontrado y para evaluar las funciones de pertenencia se usará la mejor base de reglas encontrada. Así pues siempre existirá una copia del mejor individuo encontrado en la población actual. Una vez finalizado, el proceso devolverá la mejor combinación Base de Reglas - Funciones de Pertenencia encontrada.

Inicialmente, se usará una base de reglas aleatoria, con lo que, la primera fase (mejora de etiquetas) podrá evaluarse sobre ella. Una vez aplicado el algoritmo genético encargado de mejorar las etiquetas, el mejor conjunto encontrado será usado para mejorar la base de reglas, y así sucesivamente. El esquema podemos verlo en la figura 3.

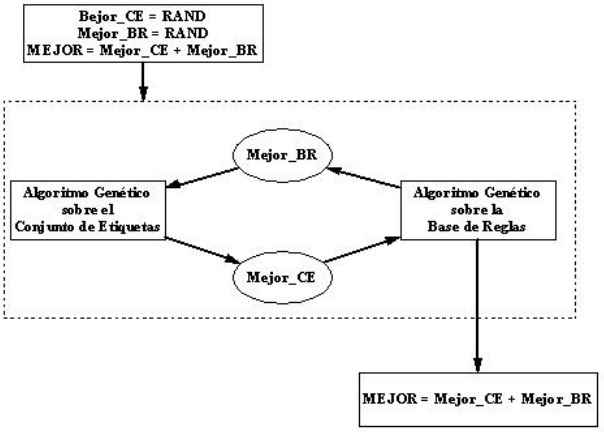


Figura 3: Esquema Usado

Antes de aplicar el método, se realizará la etapa de captura y procesado de información, descrita en la sección 3.1. Los algoritmos genéticos ejecutados iterativamente serán los descritos en las secciones 3.2 y 3.3.

4 EXPERIMENTACIÓN: MODELADO DEL CONTROL LATERAL DE UN VEHÍCULO

El objetivo del experimento consistirá en el modelado del control del volante de un coche (también conocido como control lateral del vehículo). El control automático tanto de la velocidad como de la dirección del vehículo son dos pasos básicos para llegar a conseguir el desarrollo de vehículos inteligentes autónomos. En este experimento, el diseo de la automatización del volante del vehículo gracias al método presentado anteriormente es abordado.

El primer punto del proceso de modelado consistirá en la toma de datos, para ello, un piloto ha seguido la ruta mostrada en la figura 7, mientras el computador embarcado en el vehículo ha ido tomando datos acerca del estado del coche y las acciones tomadas. Para el modelado del control lateral, se usarán dos variables de entrada, obtenidas mediante diferencias entre la ruta GPS de referencia y la posición del vehículo en un momento dado. Las variables de entrada definidas son: el error lateral y el error angular. El error lateral representa la distancia actual del vehículo a línea de referencia, los valores pueden tomar cualquier valor dentro del intervalo (-∞, ∞). Mientras que el error angular es el ángulo existente entre la línea de referencia y el vector de velocidad del vehículo, cuyos valores están restringidos al intervalo (-180,180). La representación gráfica de ambas variables se muestra en la figura 4:

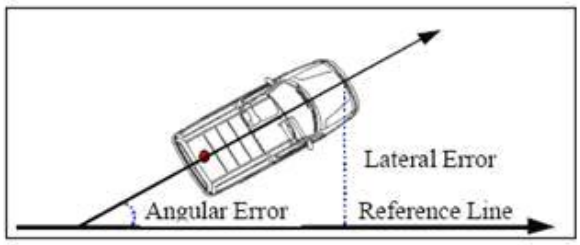


Figura 4: Representación gráfica de las variables de entrada

La salida consistirá en la posición de referencia, que será enviada a un controlador robusto de bajo nivel, capaz de manejar un motor enganchado a la barra de dirección, con el objetivo de de llevar el volante a la posición deseada. A la hora de proceder a la normalización de los valores del conjunto de datos inicial al intervalo [-1,1], se ha asumido que el límite para el error lateral es de ±5 metros y para el error angular (como ya hemos comentado) de ±180 grados. Tras esto, hemos definido una rejilla 21x21 definida en el plano X-Y, y, en cada punto de la rejilla tomamos como salida la media de las salidas más cercanas para el conjunto de datos de entrada. Tal y como podemos ver en la figura 4, existen puntos del espacio de entrada donde no existe salida asociada, así pues, acciones tomadas por el usuario en situaciones extremas, tales como de error angular de ±180 grados,... no son contempladas, ya que no se ha dado el caso durante la toma de datos. Para asegurarnos que casos extremos como estos quedan cubiertos, aadiremos los siguientes puntos al conjunto final:

- (1,1,1): Si el error lateral y angular son máximos a la izquierda, el volante debe girarse lo máximo a la derecha
- (-1,-1,-1): Si el error lateral y angular son máximos a la derecha, el volante debe girarse lo máximo a la izquierda

Una vez terminado el proceso, en la figura 5 podemos ver, tanto el conjunto de datos de partida como el conjunto de datos final, que será el que le pasaremos al programa que realizará el ajuste.

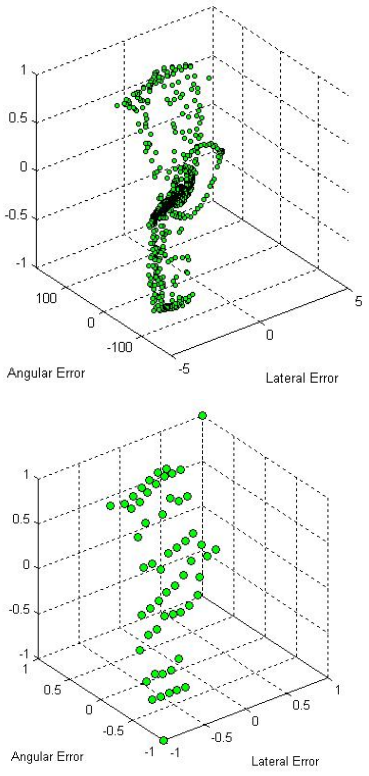


Figura 5: Conjunto de datos inicial (arriba) y final (abajo)

El método se ha ejecutado bajo la siguiente configuración de parámetros. En este caso, aquellos parámetros que son comunes a las fases de ajuste de la

base de reglas y funciones de pertenencia son iguales para ambas fases, excepto ξ. La configuración de parámetros es la siguiente: 100 iteraciones del método, con 15 individuos y 20 generaciones para cada uno de los algoritmos genéticos; probabilidad de mutación de 0.1; α (Cruce BLX) de 0.25; $\xi = 0.25$ (optimización de funciones de pertenencia) y $\xi = 3$ (optimización de base de reglas). Probabilidad de modificación del mejor individuo para la inicialización (ρ) de 0.5. Se usarán 21 posibles consecuentes para las reglas de la base de reglas y 5 funciones de pertenencia para cada una de las variables de entrada. Finalmente, para la función objetivo usaremos v = 0.75, por lo que estaremos dando más importancia a la semejanza con las salidas del piloto que a tener una superficie sin cambios bruscos.

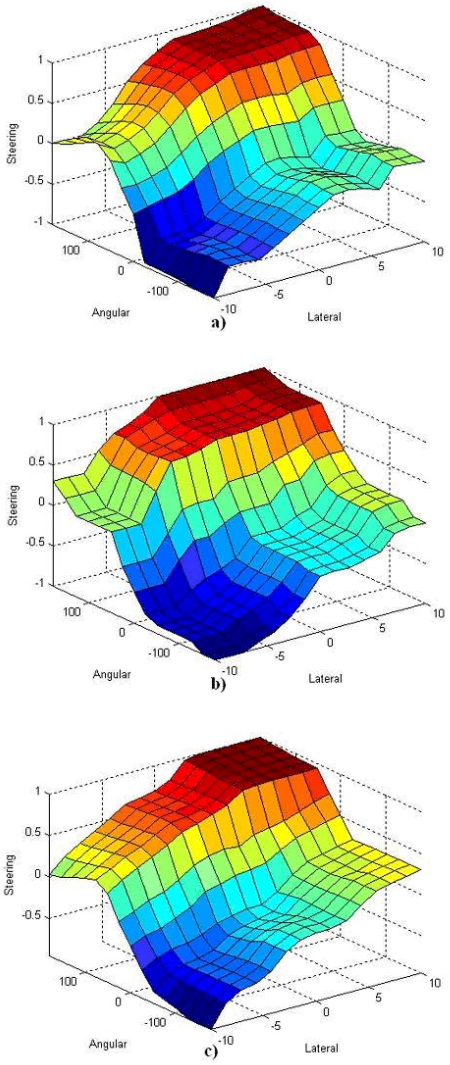


Figura 6: Superficies de control: 5M (a), 5C (b) y 5T (c)

Una vez aplicado el proceso, se han obtenido 3 controladores difusos, que denotaremos *5M* (5 funciones de pertenencia con base de reglas marginal), *5C* (5 funciones de pertenencia con base de reglas central) y *5T* (5 funciones de pertenencia con base de reglas Total) respectivamente. Las superficies de control de cada uno de los controladores se muestran en la figura 6. Tras esto, los controladores han sido puestos a prueba en la Zona de Conducción, para comprobar si son capaces de seguir una ruta prefijada. Todas las rutas han sido realizadas a velocidades entre 15 km/h en curvas y 25 km/h en tramos rectos. El comportamiento seguido por cada uno de los controladores puede apreciarse a continuación.

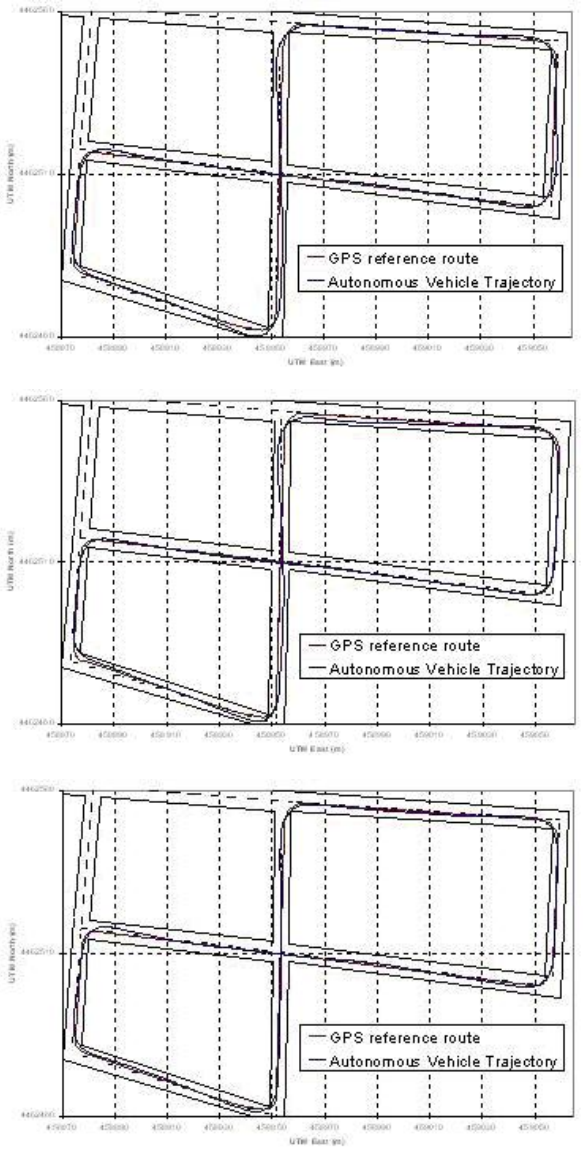


Figura 7: Referencia y Trayectoria seguida por 5M (arriba), 5C (centro) y 5T (abajo).

En las figuras puede apreciarse un relativo buen comportamiento de los controladores; todos ellos son capaces de, con mayor o menor precisión, seguir un tramo de carretera recto pero todos presentan el mismo fallo, tras una curva, se desvían hacia fuera de la línea de referencia; lo que puede ser provocado por el carácter no predictivo del modelo que, unido a la velocidad a la que han sido ejecutadas las pruebas, hacen que, una ven dentro de la curva, no le dé tiempo de reaccionar ante un nuevo punto de referencia GPS. La figura 8 muestra una comparación entre el error lateral y angular normalizado cometido por cada uno de los controladores. El gráfico muestra, debido a que ambas medidas pueden tomar valores positivos y negativos, el valor medio en valor absoluto de cada una de las medidas. El error angular ha sido normalizado a 10.55 grados, que se trata del mayor error angular promedio obtenido, mientras que el error lateral ha sido normalizado a 0.74 metros.

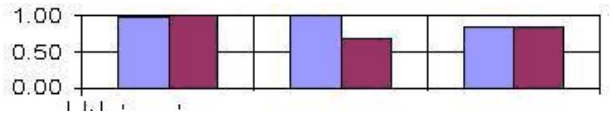


Figura 8: Comparativa del error obtenido. Error Angular (morado) y Error Lateral (magenta)

En la figura puede apreciarse que el peor comportamiento en lo que respecta al error angular ha sido obtenido por el controlador 5C y con respecto al error lateral, por 5M; el error angular obtenido por 5M es muy similar al obtenido con 5C; por tanto, diremos que 5M presenta el peor comportamiento general. Por otra parte, el controlador 5C presenta el menor valor para el error lateral. El valor mínimo para el error angular lo vemos en el controlador 5T, el cual también presenta un buen valor para el error lateral, por lo que diremos que 5T presenta el mejor comportamiento general.

5 CONCLUSIONES

A lo largo del presente trabajo se ha presentado un método para el ajuste automático de controladores difusos para el manejo de un vehículo autónomo, el cual ha sido equipado previamente con la instrumentación necesaria para llevar a cabo una conducción autónoma. Para lograr esto, se ha implementado un algoritmo genético iterativo, el cual es capaz de ajustar separadamente las funciones de pertenencia y la base de reglas del controlador difuso; para lograr un ajuste más preciso e interpretable, se han impuesto una serie de restricciones sobre los controladores, las cuales han ayudado al método, no sólo a ser capaz de obtener controladores apropiados para el manejo del volante del vehículo, sino también de llevar a cabo una conducción suave y confortable. Los controladores obtenidos han sido probados en una pista de pruebas, mostrando, la mayoría de ellos, un buen comportamiento en el seguimiento de tramos de carretera rectos, aunque un tanto peores en lo que a tramos curvos se refiere; esto se debe al carácter no predictivo del modelo usado para generar los controladores, ya que no tienen en cuenta lo que viene por delante de la posición en la que se encuentran, sino únicamente el punto de la ruta en la que se encuentra. Los controladores han mostrado una conducción suave, incluso circulando a velocidades relativamente altas, lo que es un buen punto de partida para posteriores investigaciones en el área de los Sistemas de Transporte Inteligentes.

6 AGRADECIMIENTOS

Este trabajo cuenta con el apoyo de Ministerio de Educación y Ciencia, a través del proyecto SARA, DPI2005-07980-C03-01 y el proyecto Europeo CyberCars2, STREP FP6-028062.

Referencias

[1] Willie D. Jones. "Keeping Cars from, Crashing". IEEE SPECTRUM, September 2001, pp 40-45

[2] J.E. Naranjo, C. González, J. Reviejo, R. García, T. de Pedro, and M.A. Sotelo, "Using Fuzzy Logic in Automated Vehicle Control"IEEE Intelligent Systems. Jan/Feb 2007.

[3] J.E. Naranjo, C. González, J. Reviejo, R. García, and T. de Pedro, "A throttle & brake fuzzy controller: Towards the automatic car", in Lecture Notes on Computer Science, vol. 2809. Berlin, Germany: Springer-Verlag, Jul. 2003, pp. 291-301.

[4] J. E. Naranjo, C. González, J. Reviejo, R. García, and T. de Pedro, "Adaptive fuzzy control for inter-vehicle gap keeping," Special Issue on Adaptive Cruise Control, IEEE Trans. Intell. Transp. Syst., vol. 4, no. 3, pp. 132-142, Sep. 2003.

[5] E. Dickmanns, "The development of machine vision for road vehicles in the last decade," in Proc. IEEE Intelligent Vehicles Symposium., Versailles, France, 2002, pp. 268-281.

[6] Gibbs, W. W. (2006). Innovations from a robot rally. Scientific American, 294(1), 64-71.

[7] R. García Rosa, T. De Pedro. Modeling a fuzzy coprocessor and its programming language. Math-

ware and Soft Computing, vol V, n. 2-3, pp 167-174, 1998.

[8] R. Garcia, T. De Pedro. "First Application of the ORBEX Coprocessor: Control of Unmanned Vehicles". 1999 EUSFLAT-ESTYLF Joint Conference. Mathware and Soft Computing, n. 7, vo12-3, 2000, pp. 265-273.

[9] Huang, S. and W. Ren (1999). Use of neural fuzzy networks with mixed genetic/gradient algorithm in automated vehicle control. IEEE Transactions on Industrial Electronics, 46, 1090-1102

[10] T. Takagi & M Sugeno. Fuzzy Identification of Systems and its Application to Modelling and Control. IEEE Trans. on Syst. Man and Cybernetics, 15:166-132

[11] M. Sugeno, "On stability of fuzzy systems expressed by fuzzy rules with singleton consequents," IEEE Transactions Fuzzy Systems., vol. 7, pp. 201-224, Apr. 1999.

[12] Holland, J.H. (1975).Adaptation in Natural and Artificial Systems. The University of Michigan Press.

[13] Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley, New York.

[14] Eshelman L.J. & Schaffer J.D. (1993). Real Coded Genetic Algorithms and Interval Schemata. Foundation of Genetic Algorithms 2,L. Darrell Whitley (Ed.) (Morgan Kaufmann Publishers, San Mateo), 187-202.

[15] W. Pedrycz, editor. Fuzzy Modelling: Paradigms and Practice. Kluwer Academic Press, 1996.

[16] D. T. Pham and D. Karaboga. Optimum design of fuzzy logic controllers using genetic algorithms. Journal of Systems Engineering, 1 : 114-1 18, 1991.

[17] F. Herrera and J. L. Verdegay, editors. Genetic Algorithms and Soft Computing. Number 8 in Studies in Fuzziness and Soft Computing. Physica-Verlag, 1996.

Una Metodología para la Regulación Adaptativa de Intersecciones Viales mediante Prototipos Deformables Borrosos

Eusebio Angulo, Francisco P. Romero, Ricardo García Ródenas, José A. Olivas

Escuela Superior de Informática de Ciudad Real
Universidad de Castilla – La Mancha
Paseo de la Universidad, 4, 13071, Ciudad Real
Eusebio.Angulo@alu.uclm.es
{FranciscoP.Romero, Ricardo.Garcia, JoseAngel.Olivas}@uclm.es

Resumen

En este trabajo se presenta una aplicación de los Prototipos Deformables Borrosos para el control de la regulación de intersecciones viales mediante semáforos. El objetivo inicial ha sido extraer diferentes patrones de movilidad sobre el flujo de vehículos en una intersección para lo cual se han utilizado técnicas de optimización matemática y de descubrimiento de conocimiento. La representación formal de estos patrones mediante Prototipos Deformables Borrosos proporciona los mecanismos suficientes para establecer un control adaptativo del funcionamiento de los semáforos que regulan la intersección y además dota a la solución de robustez ante la aparición de patrones anómalos de movilidad. A priori, la aplicación real de la metodología ofrece una alternativa de bajo coste sin detrimento de su funcionamiento eficiente.

Palabras Clave: Regulación Adaptiva, Prototipos Deformables Borrosos, Clustering Borroso, Modelos de estimación.

1 INTRODUCCIÓN AL PROBLEMA

El problema del tráfico es un problema grave, presente en la mayoría de las ciudades del mundo, dado que causa tanto problemas económicos como medioambientales. Es una constante preocupación para las administraciones públicas la resolución de estos problemas, para cuya resolución se buscan alternativas como el establecimiento de nuevas infraestructuras viarias (túneles, puentes, etc.), con el coste que suponen y la optimización en la gestión de los elementos que regulan el tráfico dentro de la ciudad (semáforos, rotondas, pasos de peatones, etc.)

Uno de los elementos primordiales dentro de la gestión del tráfico urbano es el funcionamiento de los semáforos que regulan una intersección vial. El elemento fundamental de este problema es el ciclo del semáforo, es decir, la secuencia finita de estados (verde, rojo, ámbar) en la que el semáforo basa su funcionamiento iterativo. Este ciclo del semáforo se debe estudiar desde tres puntos de vista distintos:

i) La longitud en tiempo del ciclo completo del semáforo, es decir, el tiempo desde el comienzo de una fase verde hasta el comienzo de la siguiente fase verde.
ii) El porcentaje del tiempo dedicado para cada una de los estados dentro de un ciclo.
iii) Las transiciones entre los semáforos consecutivos. Un ejemplo típico de esta circunstancia es la presencia de dos semáforos consecutivos en una avenida. En este caso se define el tiempo de transición (offset time) el tiempo desde que un semáforo pasa a estado verde hasta que el siguiente semáforo de la avenida pasa también a verde.

Para resolver este problema se han empleado diversas técnicas matemáticas que van desde la programación matemática con restricciones de equilibrio, útiles para la realización de análisis estratégicos del tráfico, hasta las técnicas de optimización aplicadas a modelos de simulación [8]. Además de estas técnicas matemáticas existen trabajos que han basado su funcionamiento en diferentes técnicas de soft-computing [2]. En el campo del control borroso [3] se han definido diferentes alternativas estableciendo como variables de entrada la densidad de tráfico y la velocidad de los vehículos durante el estado verde del semáforo y como salida el tiempo que debe durar cada uno de los estados dentro del ciclo del semáforo.

Dentro del ámbito de las soluciones evolutivas destaca el trabajo presentado en [7] en el que se presenta una arquitectura basada en algoritmos genéticos que optimiza de forma adaptativa los ciclos de los semáforos. Esta alternativa ha sido implementada en una zona de Santa Cruz de Tenerife y sus resultados han mejorado los

ofrecidos por los semáforos de ciclos de duración pre-establecida. El principal inconveniente que presenta es la tecnología necesaria para su implementación ya que implica la gestión de colas mediante medición de tiempos de permanencia en cola y el número de individuos en esa cola.

La aparición de nuevas tecnologías de monitorización (sensores en pavimento, cámaras, etc.) permite disponer en tiempo real de datos precisos sobre las condiciones del tráfico. Estos datos proporcionan una información valiosa para la definición de los patrones de movilidad que se presentan en la intersección a lo largo del día, es decir la cantidad y la dirección de los movimientos que se realizan en ella en diferentes intervalos de tiempo, los cuales van a ser fundamentales para la regulación eficiente de la intersección mediante semáforos.

En esta comunicación se aborda el problema de la regulación adaptativa de intersecciones planteando una metodología para el control del funcionamiento de los semáforos que la regulan. Esta metodología basa su funcionamiento en la detección automática de patrones de movilidad mediante el análisis del flujo de vehículo en una intersección. El principal problema que se pretende resolver es el tratamiento de patrones de movilidad no estacionaria, es decir, cómo reaccionar ante demandas cambiantes en cada instante del día. Con este fin se realiza una definición formal de los patrones en base a Prototipos Deformables Borrosos [5] aplicados anteriormente a otros problemas como la predicción de incendios [6].

El resto del trabajo se organiza de la siguiente manera: en la sección 2 se detallan los pasos llevados a cabo para la construcción del modelo en base a la extracción de patrones de movilidad a partir de los datos del flujo de vehículos. En la sección 3 se explican las fases de aplicación del modelo de regulación adaptativa en una intersección. En la sección 4 se presenta un experimento computacional con el fin de evaluar la aplicabilidad de la metodología de construcción del modelo. Finalmente, se establecen las conclusiones obtenidas y las líneas de trabajo que permanecen abiertas en este tema.

2 CONSTRUCCIÓN DEL MODELO

El objeto de la metodología desarrollada es regular adaptativamente mediante una serie de semáforos coordinados una intersección como la mostrada en la Figura 1. Se parte de que dicha intersección dispone de sensores que aforan los cuatro carriles, lo cual nos va a permitir disponer de la cantidad de vehículos que entran y salen en cada uno de los sentidos en un intervalo de tiempo dado.

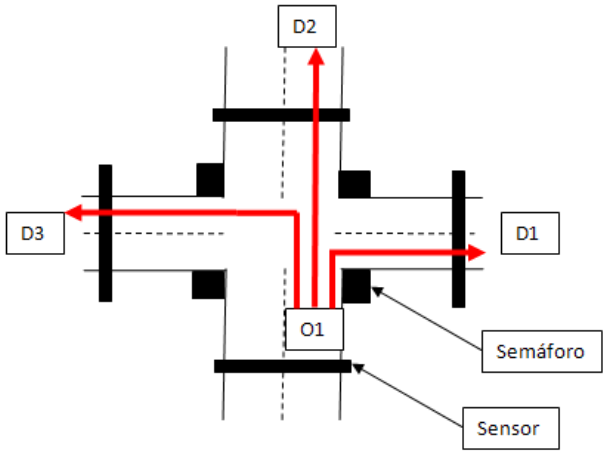


Figura 1: Intersección de 4 carriles con movimientos *U* prohibidos

Con el fin de construir el modelo sobre el que se va a determinar la regulación adaptiva es necesario proceder previamente a la extracción de los patrones de movilidad que se presentan en la intersección. Estos patrones se van a obtener a partir de las observaciones de flujo de vehículos obtenidas por los sensores.

Las diferentes fases a llevar a cabo para la construcción del modelo son las siguientes (Figura 2):

1. Observaciones de flujo entrante/saliente mediante los sensores.
2. Modelo de estimación de matrices O-D de tráfico dinámicas: Este modelo permite estimar los giros en la intersección. Se define matriz O-D como la matriz que contiene en la fila *i-j* el flujo (número de personas por unidad de tiempo) que se incorporan a la intersección en el carril *i* y salen de la misma por el carril *j*. Se asumen los movimientos tipo *U* prohibidos, es decir, consideramos la celdilla *i-i* igual a cero.
3. Extracción de Patrones de movilidad: Establece los patrones de movilidad, en función de las matrices O-D y los representa mediante Prototipos Deformables Borrosos.
4. Modelo de regulación semafórica: Mediante conocimiento experto se determina el comportamiento óptimo de los semáforos para cada uno de los prototipos obtenidos en la fase anterior.

En un determinado periodo del día, se pueden observar diferentes matrices O-D, estas diferencias son debidas al azar, pero todas ellas se considera que corresponden a unas mismas pautas de viaje, nosotros denominamos esto como patrón de movilidad para dicho periodo y a cada una de estas matrices como una realización concreta de dicho patrón de movilidad.

2.1 OBSERVACIONES DE FLUJO

Los sensores en la entrada/salida de cada carril ofrecen los aforos de la intersección en un intervalo de tiempo determinado *t*. De esta forma se obtiene el número de vehículos que pasa por cada sensor, desconociéndose los destinos de cada uno de ellos por mor de los giros que se puedan producir.

Considerada una intersección con *m* entradas, *n* salidas y el período de tiempo de interés dividido en ***N*** intervalos, los datos de entrada del modelo serían:

- Flujos de entrada: $q(t) = [q_1(t), \ldots, q_m(t)]^T$
- Flujos de salida: $y(t) = [y_1(t), \ldots, y_n(t)]^T$

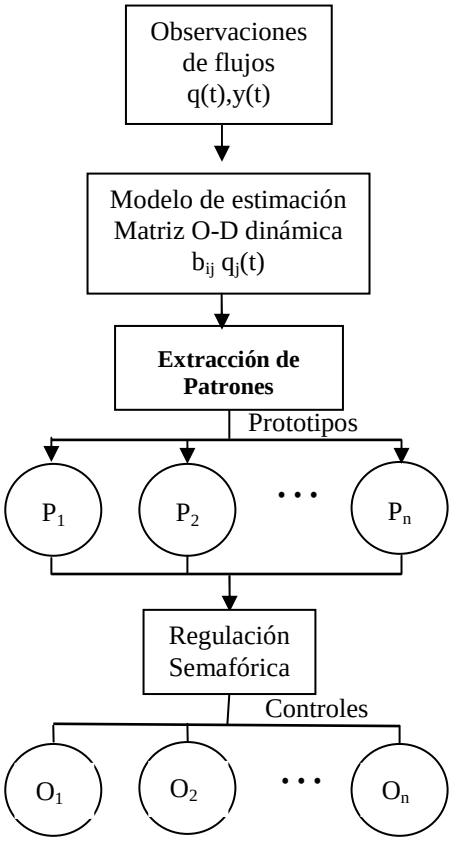


Figura 2: Pasos de la construcción del modelo.

2.2. MODELO DE ESTIMACIÓN: MATRIZ O-D DINÁMICA

Mediante el modelo de estimación que se expone en este apartado, se consigue la matriz O-D con las predicciones completas (incluidos los giros). Esta estimación puede realizarse de forma instantánea a partir de los datos ofrecidos por los sensores de flujo.

Las variables del modelo desarrollado serán las siguientes:

- I_j: Conjunto de valores donde la entrada *i* permite a los usuarios tomar salida j.

$$I_j \ (j=1,\ldots, n).$$

- b_{ij}: Probabilidad de que un vehículo entre por i y tome la salida j.

$$b_{ij} \ (i=1, \ldots, m; j=1, \ldots, n).$$

- b : Vector con probabilidad desde cada entrada *i* hasta la salida j

$$b_j = [b_{ij}] \ \forall i \epsilon I_j \ y \ Q_j=[q_i] \ \forall i \epsilon I_j$$

$$b = [b_1^T,\ldots, b_n^T]^T = [b^{(i)}]$$

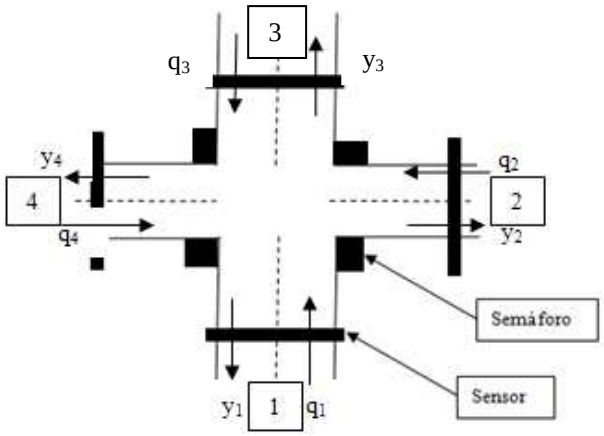


Figura 3: Entradas q_i y salidas y_i en la intersección de 4 carriles con movimientos *U* prohibidos

Para una intersección como la de la Figura 1, donde tenemos *n*=*m*=4, las variables quedarían como sigue:

$$b_1 = [b_{21}, b_{31}, b_{41}]^T \qquad Q_1=[q_2,q_3,q_4]^T$$
$$b_2 = [b_{12}, b_{32}, b_{42}]^T \qquad Q_2=[q_1,q_3,q_4]^T$$
$$b_3=[b_{13}, b_{23}, b_{43}]^T \qquad Q_3=[q_1,q_2,q_4]^T$$
$$b_4=[b_{14}, b_{24}, b_{34}]^T \qquad Q_4=[q_1,q_2,q_3]^T$$

$$b=[\ b_{21}, b_{31}, b_{41}, b_{12}, b_{32}, b_{42}, b_{13}, b_{23}, b_{43}, b_{14}, b_{24}, b_{34}]^T = [\ b^{(1)}, \ldots, b^{(12)}]^T$$

Donde b debe cumplir:

$$b \geq 0, \sum_{j=1}^{n} b_{ij} = 1 \wedge b_{ii} = 0 \quad (1)$$

Por tanto, la suma de probabilidades desde cada entrada *i* hasta una salida *j* debe ser 1 y cada probabilidad tiene que ser mayor que 0. Para los aforos observados *y(t)* y *q(t)* en cada intervalo de tiempo *t*, se resuelve el problema de estimación de b con (formula 2):

$$J(b)=\sum_{s=1}^{t}\sum_{j=1}^{n}\left\{y_j(s)-Q_j(s)b_j\right\}^2$$
$$t=1,\ldots,N \qquad (2)$$

Siendo J_i el conjunto de valores donde la salida j permite a los usuarios entrar en la entrada i. Con (i=1,…, m; j=1,…, n).

El modelo de estimación obtiene la matriz O-D teniendo en cuenta los giros. Estos vectores conforman la entrada de la fase de extracción de patrones de movilidad.

2.3. EXTRACCIÓN DE PATRONES DE MOVILIDAD

Con el objetivo de detectar las relaciones entre los matrices O-D y así poder obtener posteriormente el patrón de movilidad correspondiente se lleva a cabo un proceso de clustering el cual agrupará cada uno de los elementos ofrecidos por similitud. En primer lugar se construirá una matriz de similitud entre cada uno de los elementos de entrada al proceso de extracción de prototipos, en este caso, las matrices resultantes del modelo de estimación. Se utilizará la distancia euclídea entre vectores de la misma dimensión como función para determinar la semejanza entre elementos.

$$d(u,v)=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(v_{ij}-u_{ij}\right)^2} \quad (3)$$

Una vez obtenida esta matriz de similitud se ejecuta un proceso de clustering en dos fases. En primer lugar, con el fin de detectar agrupaciones formadas por intervalos de tiempo sucesivos con valores similares de flujo, se ejecuta un proceso de clustering basado en el particionamiento de grafos[4]. A partir de este primer particionamiento y con el fin de detectar las agrupaciones restantes no contiguas en el tiempo se utiliza un algoritmo jerárquico basado en el tratamiento de grafos que explota la conectividad borrosa entre sus elementos [1]. Los nodos de este grafo serán los clusters detectados durante la primera fase particionante. La utilización de este clustering en dos fases, particionante y jerárquica, permite mantener la complejidad y eficiencia del algoritmo.

Cada uno de las agrupaciones obtenidas corresponde a un patrón de movilidad en la intersección. Cada uno de estos patrones será representado mediante Prototipos Deformables Borrosos que serán representados formalmente por una serie de números borrosos. Estos se obtendrán a partir de la normalización y agregación de las matrices O-D que componen cada uno de los grupos detectados. De esta manera se calcularán los diferentes centros de los números borrosos triangulares, lo cual, junto al tamaño de la base del triángulo (suficientemente grande para poder crear un recubrimiento en el eje X entre 0 y 1) constituyen los únicos valores necesarios para completar su representación. Un ejemplo de representación de cinco prototipos obtenidos de la fase de clustering se puede observar en la Figura 4.

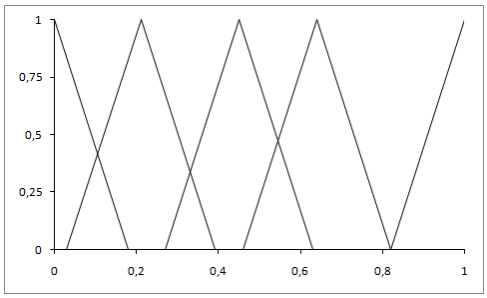


Figura 4: Representación formal de los prototipos.

Con la utilización de una representación formal en base a números borrosos se dispone de un mecanismo sencillo para obtener un grado de pertenencia (entre 0 y 1) de cada una de las situaciones reales que se presenten a cada una de las situaciones prototípicas detectadas.

2.4. REGULACIÓN SEMAFÓRICA

Para cada uno de los patrones de movilidad detectados y definidos mediante prototipos borrosos, se determina, mediante conocimiento experto, la respuesta más adecuada que deben ofrecer estos semáforos ante cada uno de estos patrones. Si se dispone de N prototipos borrosos se establecerán N respuestas que establezcan regulaciones de los semáforos óptimas. Cada una de estas respuestas vendrán representadas mediante marcos o frames con los parámetros que las definen. Estos parámetros, o bien podrán estar expresados numéricamente (tiempo en segundos de duración del periodo verde) o mediante etiquetas lingüísticas. (Tabla 1).

Tabla 1: Descripción Paramétrica de los prototipos.

Prototipo	Nivel de Congestión	Tiempo Verde	Dirección de la demanda
P1	Congestionada	Largo	Ida
P2	*Semicongestionada*	Medio	Ida
P3	*Sin congestión*	Corto	Ida
P4	*Semicongestionada*	Medio	Vuelta
P5	*Sin congestión*	Corto	Vuelta

La utilización de los prototipos deformables borrosos y la definición mediante etiquetas lingüísticas de la regulación semafórica permite diseñar una solución robusta y flexible al problema de la gestión del tráfico en la intersección. Esto permitirá dar una respuesta lógica ante casos extremos como por ejemplo acontecimientos deportivos o culturales que provoquen un incremento o decremento puntual del tráfico en la intersección regulada.

3 APLICACIÓN DEL MODELO

Una vez calculado el modelo éste es aplicable para la gestión diaria de la intersección. Los datos de entrada al sistema de regulación serán las observaciones de flujo en tiempo real, y las salidas expresarán el tipo de regulación a efectuar en cada momento.

En la Figura 5 se pueden observar los diferentes pasos que deben ser ejecutados y que se describen en los siguientes apartados.

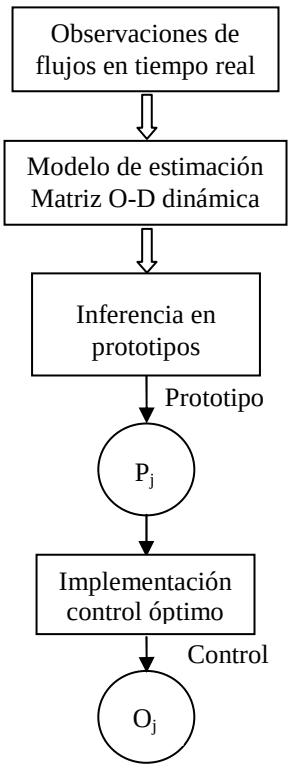


Figura 5: Metodología **on-line**.

3.1. OBSERVACIÓN DE FLUJOS EN TIEMPO REAL

Los sensores colocados en la entrada/salida de cada carril de la intersección ofrecerán datos del flujo de vehículos que se está produciendo en cada instante. Estos datos alimentarán al sistema para determinar cuál es la mejor regulación del tráfico de la intersección.

3.2. MODELO DE ESTIMACIÓN

El modelo de estimación permite obtener la matriz O-D completa (incluyendo giros) a partir de las estimaciones de aforo. Se obtiene exactamente de la misma forma que en la fase de construcción del modelo**.** De esta manera, los elementos y cálculos especificados en el apartado 2.2 serían aplicables también en esta fase.

3.3. INFERENCIA EN PROTOTIPOS

A partir de los valores de entrada de la matriz O-D se determina el patrón de movilidad actual mediante un proceso de inferencia en los prototipos deformables borrosos que forman el modelo. Este proceso se realiza de la siguiente manera:

1. Normalizar los valores del vector O-D de entrada.
2. Agregar los valores normalizados (valor X).
3. Obtener los grados de pertenencia a los prototipos representados mediante los números borrosos. La obtención de un grado de afinidad con los prototipos preestablecidos (μ_i) proporciona información relevante para evaluar la situación presentada (Figura 6).

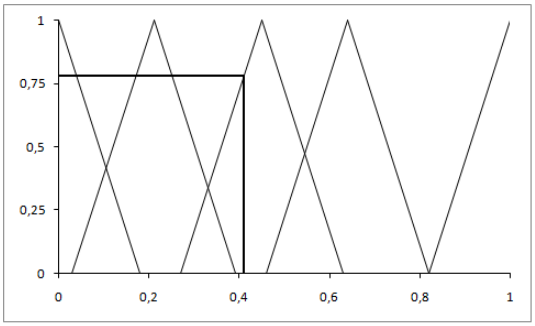


Figura 6: Similitud de un vector a un prototipo (X = 0,41 μ_3=0,78).

Una vez obtenidos los grados de pertenencia de la situación actual a los prototipos del modelo se regresa la definición del prototipo más afín como base de la regulación semafórica óptima a ejecutar.

3.4. IMPLEMENTACIÓN DE CONTROL ÓPTIMO

El prototipo más afín de entre los encontrados en el apartado anterior será la regulación semafórica más adecuada para la situación real del tráfico en dicho intervalo de tiempo.

De esta manera, la salida del sistema contiene todos los parámetros que definirán el comportamiento de los semáforos mientras permanezca el patrón de movilidad detectado. Estos valores serán transmitidos al componente electrónico que se ocupará de transmitir las órdenes a cada uno de los semáforos implicados en la intersección.

4 EXPERIMENTO COMPUTACIONAL

En este apartado de explica un pequeño estudio de simulación que se está desarrollando para evaluar la metodología propuesta. La ventaja de este tipo de análisis frente al estudio de un caso real reside en el conocimiento que se tiene de los mecanismos dinámicos de la demanda, pudiendo así contrastar si la metodología es capaz de detectar los patrones de movilidad correspondientes o por el contrario se ve afectada por las perturbaciones aleatorias que se introducen en los datos. En primer lugar se establece la intensidad de tráfico en cada intervalo de tiempo a partir de un modelo de evolución horaria de la demanda total conocido como el que se muestra en la Figura 7.

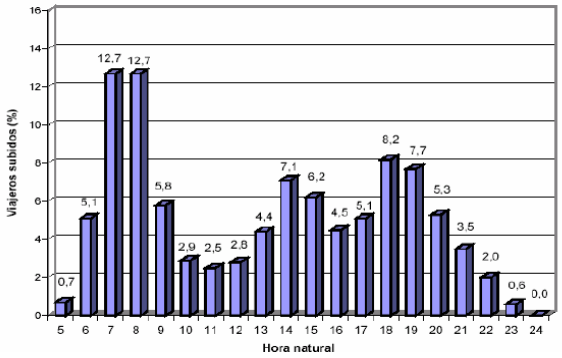


Figura 7: Distribución horaria de la demanda en la red de Cercanías Metropolitana de Madrid.

Se establecen intervalos de tiempo de 5 minutos que van a ir desde las 5:00 de la mañana hasta las 24:00 de la noche, con ello, la matriz O-D observada dispondrá de 228 filas (nº de intervalos de tiempo) y 12 columnas que representan a las diferentes trayectorias a efectuar por los vehículos, dado que las trayectorias en forma de *U* (donde el propio origen es el destino) están prohibidas. Denominando, tal y como se explica en el apartado 2.2, q_i a la demanda de tiempo entrante para cada intervalo de tiempo en cada carril de la intersección e y_i a la demanda saliente. La estimación de las entradas q_i a la intersección se realiza mediante la siguiente expresión:

$$q_i = D * p(t) * u(1-\varepsilon, 1+\varepsilon) \quad (4)$$

donde:

- *D:* demanda total de entrada a la intersección.
- *p(t):* proporción dependiente del instante de tiempo. Este parámetro permite tener en cuenta el sentido de los flujos del tráfico, pues por la mañana sería en un sentido y por las tardes en sentido contrario.
- ε *:* introduce aleatoriedad en la distribución.

La estimación de las salidas y_i en la intersección se realiza mediante la expresión:

$$y_j = \sum_{j \neq i} P_{ij}(t) * q_i$$

Se asume una matriz P_{ij} dependiente del tiempo, la salida y_j se calcula en función de dicha matriz y de las entradas a dicha intersección.

Una vez realizadas las estimaciones de las entradas y las salidas para cada carril, durante todos los intervalos, se establece la matriz observada, donde se almacenan las estimaciones de aforo. El modelo de estimación permite obtener la matriz O-D completa (incluyendo giros) a partir de las estimaciones de aforo y se calcula mediante los elementos y cálculos especificados en el apartado 2.2. Estos datos estimados conforman lo que se denomina matriz predicha**.** En la figura 8 se puede observar como las dos matrices ofrecen valores prácticamente similares.

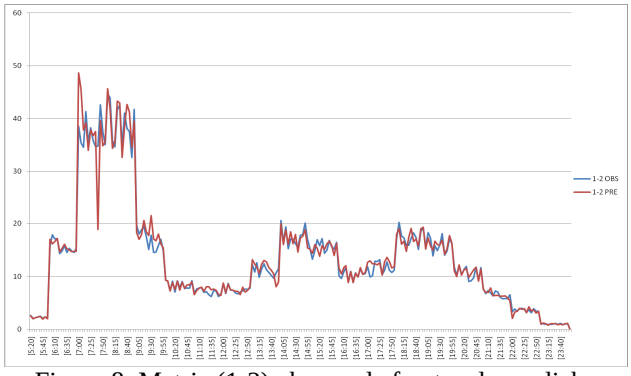


Figura 8: Matriz (1-2) observada frente a la predicha

A partir de estas matrices O-D se realiza el clustering mediante el algoritmo especificado en el apartado 2.3. Los resultados que ofrece este clustering permiten agrupar el resultado en 6 distribuciones distintas (Figura 9), que corresponden a 6 patrones de movilidad distintos que serán representados mediante prototipos deformables borrosos.

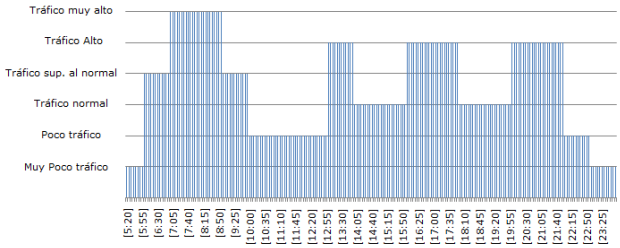


Figura 9: Distribución de las agrupaciones

Los números borrosos resultantes serían los representados en la Figura 10.

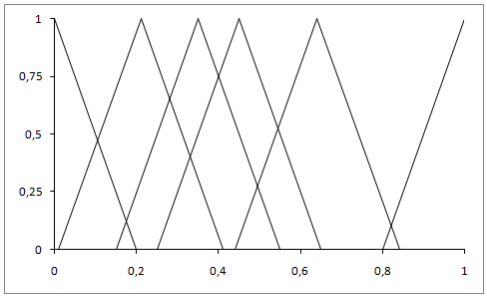


Figura 10: Números Borrosos resultantes.

5 CONCLUSIONES Y TRABAJO FUTURO

Se ha definido en este trabajo una metodología para la definición y extracción de patrones de movilidad mediante técnicas de optimización y lógica borrosa. La utilización de técnicas de descubrimiento de conocimiento prototípico borroso ha permitido crear un modelo más estructurado, significativo y fácilmente actualizable que los existentes actualmente en la literatura.

La metodología aporta un control automático y adaptativo para las intersecciones reguladas semafóricamente, obteniendo una salida del sistema controlada, eliminando así respuestas caóticas del sistema frente a situaciones anómalas. Además, la utilización como entrada de la cantidad de vehículos que fluye por la intersección en un intervalo de tiempo hace que la tecnología a utilizar para su implementación real sea sencilla y su rendimiento muy eficiente.

Se ha realizado un estudio experimental para comprobar la validez de esta metodología mediante la simulación de flujos de vehículos según distribuciones previamente conocidas. Con ello se ha verificado que el planteamiento realizado es realista y puede ser aplicable, a priori, en un entorno real con un funcionamiento satisfactorio.

Como líneas futuras de desarrollo se presenta la experimentación con datos reales de flujo de vehículos sobre intersecciones así como la definición más concreta de los parámetros que determinan cada uno de los patrones de movilidad detectados.

Agradecimientos

La investigación realizada está financiada por el proyecto TIN2007-67494 del Ministerio de Educación y Ciencia y los proyectos PAC06-0059 y PCC08-0081-4388-2 de I+D de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1]Y. Dong, Y. Zhuang, K. Chen, X. Taib, “A hierarchical clustering algorithm based on fuzzy graph connectedness”, Fuzzy Sets and Systems, vol. 157, pp. 1760 – 1774, 2006.

[2] G. Y. Lim, J. J. Kang, and Y. S. Hong, “The optimization of traffic signal light using artificial intelligence,” in Proc. 10th IEEE Int. Conf. Fuzzy Syst., Dec. 2–5, 2001, vol. 3, pp. 1279–1282.

[3] J. López, R. García, A. García Blanco, I. Zúñiga Félix, “Traffic Lights Fuzzy Control Proposals to Improve Vehicular Flow”, Fourth Congress of Electronics, Robotics and Automotive Mechanics, pag. 406-411. 2007.

[4] H. Kawaji, Y. Yamaguchi, H. Matsuda, A. Hashimoto “A Graph-Based Clustering Method for a Large Set of Sequences Using a Graph Partitioning Algorithm”. Genome Informatics Vol. 12: pp 93-102 , 2001.

[5] J. A. Olivas, . “Contribución al Estudio Experimental de la Predicción basada en Categorías Deformables Borrosas”, Tesis Doctoral, Universidad de Castilla La Mancha, España. 2007.

[6] J. A. Olivas J., F. Romero. “FPKD. Fuzzy Prototypical Knowledge Discovery. Application to Forest Fire Prediction” . Proceedings of the SEKE'2000, Knowledge Systems Institute, Chicago, Ill. USA, pp. 47 - 54. 200.

[7] J. Sánchez, M. Galán, E. Rubio, “Applying a Traffic Lights Evolutionary Optimization Technique to a Real Case: “Las Ramblas” Area in Santa Cruz de Tenerife”, IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, pag. 25-40 VOL. 12, NO. 1, 2008

[8] T. Thorpe, “A Physically-Realistic Simulation of Vehicle Traffic Flow”, Technical Report 97-104, Department of Computer Science, Colorado State University, 1997.

PROCESOS DE DESCUBRIMIENTO DE CONOCIMIENTO EN BASES DE DATOS USANDO GRAFOS ASOCIADOS A FILTROS MORFOLÓGICOS

Noé N. Frago, Ramón Fuentes-González, Pedro Burillo
Universidad Pública de Navarra, Dpto. Automática y Computación, 31006 Pamplona, España
{noe.frago, rfuentes, pburillo}@unavarra.es

Resumen

Se definen filtros morfológicos asociados a relaciones difusas y se demuestran propiedades de los mismos. Posteriormente, se prueba la utilidad de la teoría introducida mostrando, con un ejemplo, cómo con el uso de ciertos grafos asociados a esos filtros morfológicos se puede extraer conocimiento relevante en una base de datos con un gran número de ítems.

Palabras Clave: Soft Computing, Filtros Morfológicos, Descubrimiento de Conocimiento, Relaciones Difusas.

1. INTRODUCCIÓN

En el contexto del tratamiento y análisis de imágenes, la *Morfología Matemática* (*MM*) es una teoría que proporciona herramientas útiles para el análisis de estructura en imágenes. Sus fundamentos, introducidos por G. Matheron ([10]) y J. Serra ([11]), se encuentran en la teoría de conjuntos, en la geometría integral y en el álgebra de retículos. En su formulación más amplia, la metodología que aparece en esta teoría para el análisis de imágenes consiste en la consideración de éstas como puntos $A \in L$ de un retículo completo $(L, \leq, \vee, \wedge, 0, 1)$. En este contexto, el problema básico para el tratamiento de imágenes con técnicas de *MM* es el de diseñar operadores no lineales $F: L \to L$ (los *filtros*) de manera que $F(A)$ encapsule alguna información relevante (de tipo geométrico, topológico,...) extraída de la imagen $A \in L$. En el conjunto $\mathcal{L}$ de filtros el orden $\leq$ y las operaciones $\vee$ e $\wedge$ son extensiones de las análogas en L:

$$F_1 \leq F_2 \Leftrightarrow F_1(A) \leq F_2(A) \text{ para toda imagen } A.$$

$$(F_1 \vee F_2)(A) = F_1(A) \vee F_2(A) \text{ y } (F_1 \wedge F_2)(A) = F_1(A) \wedge F_2(A).$$

La mayoría de estos filtros F se construyen a partir de los siguientes operadores en el retículo L:

1. Los operadores básicos del tipo *erosión* $\xi: L \to L$ y del tipo *dilatación* $\delta: L \to L$, que constituyen una *adjunción* (ξ, δ) en L, es decir que están asociados por:

$$\delta(A) \leq B \Leftrightarrow A \leq \xi(B), \quad A \in L \quad B \in L.$$

2. La composición de los operadores anteriores:

Apertura $\gamma = \delta \circ \xi$ (erosión seguida de dilatación) y *cierre* $\zeta = \xi \circ \delta$ (dilatación seguida de erosión).

Utilizando los anteriores se construyen nuevos filtros:

Por composición (en serie): $F_1 = \gamma_1 \circ \gamma_2$, $F_2 = \gamma_1 \circ \gamma_2 \circ \zeta$, o por disyunción (en paralelo): $F_3 = \gamma_1 \vee F_2$, $F_4 = \gamma_1 \vee F_2 \vee \zeta$, etc.

La aplicación inicial de la teoría *MM* fue el tratamiento de imágenes binarias, consideradas como subconjuntos A de puntos $\boldsymbol{x} = (x_1, x_2)$ del plano $U = \mathbb{R}^2$ o del plano $U = \mathbb{Z}^2$. El tratamiento de imágenes binarias, consideradas como puntos $A, B, \ldots$ del retículo $(\wp(U), \subseteq)$, se fundamenta en los operadores *suma* $A \oplus B = \cap \{A + \boldsymbol{x} \,/\, \boldsymbol{x} \in B\}$ y *diferencia* $A \ominus B = \cup \{A + \boldsymbol{x} \,/\, \boldsymbol{x} \in B\}$ *de Minkowski* ([10]). Los filtros que se obtienen son operadores $F: \wp(U) \to \wp(U)$. Más tarde, se extendió la teoría para imágenes en tonos de gris ([11]), imágenes modeladas como aplicaciones $A: \mathbb{R}^2 \to \mathbb{R}^+$ ó $A: \mathbb{Z}^2 \to \mathbb{R}^+$, y operadores filtros $F: (\mathbb{R}^+)^{\mathbb{R}^2} \to (\mathbb{R}^+)^{\mathbb{R}^2}$ ó $F: (\mathbb{R}^+)^{\mathbb{Z}^2} \to (\mathbb{R}^+)^{\mathbb{Z}^2}$. En ambos casos, los operadores básicos *dilatación* δ_R y *erosión* ξ_R ([11]) dependen de una imagen fija R, llamada *elemento o imagen estructurante,* que encapsula el criterio adoptado para lograr el tipo de efecto deseado sobre la imagen inicial A. Por ejemplo, en el caso binario, para una imagen A: $\delta_R(A) = A \oplus R$ y $\xi_R(A) = A \ominus \check{R}$, donde $\check{R} = \{-\boldsymbol{x} \,/\, \boldsymbol{x} \in R\}$ es la imagen simétrica de $R \in \wp(U)$ respecto al origen $\boldsymbol{0} = (0,0) \in U$.

En el campo de la Lógica Difusa se ha desarrollado un *tratamiento morfológico difuso de imágenes* (*FMM*) ([1,2,3]), representando las imágenes por subconjuntos difusos $A \in [0,1]^{\mathbb{R}^2}$ ó $A \in [0,1]^{\mathbb{Z}^2}$. Los filtros son ahora del tipo $F: [0,1]^{\mathbb{R}^2} \to [0,1]^{\mathbb{R}^2}$ ó $F: [0,1]^{\mathbb{Z}^2} \to [0,1]^{\mathbb{Z}^2}$. Los operadores básicos ([1,2]), (*erosiones y dilataciones difusas*), se definen utilizando *t-normas* T, *implicaciones* $\rightsquigarrow$ y *negaciones difusas* $'$ en [0,1].

En un contexto diferente al de tratamiento de imágenes, en ([5,6,7,8]) se *aplican por primera vez, técnicas de Morfología Matemática difusa en problemas de Análisis de Datos* extraídos de una tabla de doble entrada, justificando que la *FMM* es una herramienta útil en Descubrimiento de Conocimiento en Bases de Datos (*KDD*), datos no asociados necesariamente a imágenes.

En este caso, introducimos los operadores básicos erosiones y dilataciones utilizando también un elemento estructurante R que, en lugar de un subconjunto, es aquí *una relación* (difusa o nítida) entre pares de registros:

Si T y $\rightsquigarrow$ son una *t-norma* y una *implicación* en [0,1], si U es un referencial de registros $x, y, \ldots$; si $A \in [0,1]^U$ es un subconjunto difuso de registros, si $R \in [0,1]^{UXU}$ es una relación difusa (*relación estructurante*) y R^{op} su opuesta, entonces, utilizando el operador triángulo $\triangleleft_{\rightsquigarrow}$ ([9]) asociado a $\rightsquigarrow$ y el operador composición $\circ_T$, definimos en ([8,7]) los operadores erosión ξ_R y dilatación δ_R sobre subconjuntos difusos de ítems como:

$$\xi_R(A) = R^{op} \triangleleft_{\rightsquigarrow} A, \quad \delta_R(A) = R \circ_T A, \quad \forall A \in [0,1]^U \; \forall x \in U,$$

es decir:

$$\xi_R(A)(x)= inf\{R^{op}(x,y) \rightsquigarrow A(y) /y\in U\} \ \forall A\in[0,1]^U \ \forall x\in U$$
$$\delta_R(A)(x) = sup\{R(x,y) \mathcal{T} A(y) / y\in U\} \ \forall A\in[0,1]^U \ \forall x\in U .$$

En la línea de lo expuesto en estos párrafos finales, en este trabajo introducimos, por primera vez, nuevos filtros morfológicos básicos sobre relaciones difusas $S\in[0,1]^{U\times U}$ asociadas a una base de datos U. Posteriormente extendemos al mismo contexto otros filtros morfológicos. Finalmente, justificamos su utilidad interpretando ciertos grafos asociados a esos filtros como herramientas para obtener información relevante en la base de datos U.

2. DEFINICIÓN DE EROSIÓN Y DILATACIÓN DE RELACIONES DIFUSAS

Consideramos en $[0,1] \subseteq \mathbb{R}$ la relación de orden usual $\leq$ junto con los siguientes elementos y operaciones:

(i) Una t-norma $\mathcal{T}$ *continua por la izquierda* ([4]), es decir, tal que $\forall a\in [0,1] \ \forall M\subseteq[0, 1]$ verifique:

$$a \mathcal{T} supM = sup(a \mathcal{T} M) .$$

(ii) La *implicación difusa* $\rightsquigarrow_{\mathcal{T}}$ obtenida por *residuación* ([4]) de la t-norma anterior $\mathcal{T}$, es decir:

$$a\rightsquigarrow_{\mathcal{T}} b = sup\{w / a \mathcal{T} w \leq b \} \ \forall(a,b)\in[0, 1]^2 .$$

Y, fijado un referencial $U \neq \emptyset$, en el retículo completo $(\mathcal{F}(U\times U), \leq)$ de las relaciones difusas R, S, Q,... pertenecientes a $[0,1]^{U\times U}$ con el orden $\leq$ extensión usual del de [0,1] las operaciones asociadas a la t-norma $\mathcal{T}$:

(iii) Composición triángulo $\triangleleft_{\rightsquigarrow_{\mathcal{T}}}$ ([9]) y *composición círculo* $o_{\mathcal{T}}$, entre relaciones difusas:

$$(R \triangleleft_{\rightsquigarrow_{\mathcal{T}}} Q)(x,z) = inf\{R(x,y) \rightsquigarrow_{\mathcal{T}} Q(y,z)/y\in U\}\forall(x,z)\in U\times U$$

$$(R \ o_{\mathcal{T}} Q)(x,z) = sup\{R(x,y) \ \mathcal{T} Q(y,z)/y\in U\}\forall(x,z)\in U\times U.$$

(iv) Y la operación $^{op} : [0,1]^{U\times U}\rightarrow[0,1]^{U\times U}$ que hace corresponder a una relación R, su opuesta R^{op}:

$$R^{op}(x,y) = R(y,x) \ \forall(x,y)\in U\times U.$$

Con estos elementos, proponemos las siguientes definiciones de operadores básicos erosión y dilatación en el contexto de relaciones difusas:

Definición 1.- Dada $R\in[0,1]^{U\times U}$, llamada *relación estructurante*, definimos la *erosión* $\xi_R(Q)$ y la *dilatación* $\delta_R(Q)$ por R de la relación difusa $Q\in[0,1]^{U\times U}$ por

$$\xi_R(Q) = R^{op} \triangleleft_{\rightsquigarrow_{\mathcal{T}}} Q,$$
$$\delta_R(Q) = R \ o_{\mathcal{T}} Q.$$

Es decir, son las nuevas relaciones difusas de $[0,1]^{U\times U}$ tales que

$$\begin{aligned}\xi_R(Q)(x,z)&=inf\{R^{op}(x,y) \rightsquigarrow_{\mathcal{T}} Q(y,z)/y\in U\}\forall(x,z)\in U^2\\ \delta_R(Q)(x,z)&= sup\{ R(x,y) \mathcal{T} Q(y,z) /y\in U\} \ \forall(x,z)\in U^2,\end{aligned} \quad (1)$$

Si para $x\in U$ se considera su *R-vecindad*, es decir el subconjunto difuso R_x tal que $R_x(y) = R(x,y) \ \forall y\in U$, las expresiones (1) pueden representarse por:

$$\begin{aligned}\xi_R(Q)(x,z)&= inf\{(R^{op})_x(y)\rightsquigarrow_{\mathcal{T}} Q(y,z) /y\in U\} \ \forall(x,z)\in U^2\\ \delta_R(Q)(x,z) &= sup \ \{ R_x(y) \mathcal{T} Q(y,z) / y\in U\} \ \forall(x,z)\in U^2\end{aligned} \quad (2)$$

Nota. Se verifica: $\xi_{\emptyset}(Q) = U\times U$ y $\delta_{\emptyset}(Q) = \emptyset$ para toda $Q\in[0,1]^{U\times U}$ y $\xi_R(U\times U) = U\times U$, $\delta_R(\emptyset) = \emptyset \ \forall R\in[0,1]^{U\times U}$.

Sean $\leq$, $\wedge$,$\vee$, $'$ y $\mathcal{T}^*$ las extensiones puntuales a $[0,1]^{U\times U}$ del orden $\leq$ y de las leyes $\wedge$,$\vee$, $'$ y $\mathcal{T}^*$ en [0,1].

Proposición 1.- *Para toda relación estructurante R, los operadores erosión y dilatación morfológicos definidos en (1) verifican:*

1) $\xi_R(S \wedge Q) = \xi_R(S) \wedge \xi_R(Q) \ \forall(S,Q)\in[0,1]^{U\times U}\times [0,1]^{U\times U}$.

2) $\delta_R(S \vee Q)=\delta_R(S) \vee \delta_R(Q) \ \forall(S,Q)\in[0,1]^{U\times U}\times [0,1]^{U\times U}$.

3) Si $S \leq Q$, entonces $\xi_R(S) \leq \xi_R(Q) \ \ \& \ \ \delta_R(S) \leq \delta_R(Q)$.

Nota: Si la relación estructurante R es reflexiva, entonces

$$\xi_R(Q) \leq Q \leq \delta_R(Q) \quad \forall Q\in[0,1]^{U\times U}.$$

Un caso particular interesante se obtiene, al considerar la *t-norma* $\mathcal{T}_\ell$ *de Lukasiewicz* en [0,1]:

$$a \ \mathcal{T}_\ell b = max(0, a + b – 1) \quad \forall(a,b)\in[0,1]\times[0,1],$$

cuya implicación asociada $\rightsquigarrow_{\mathcal{T}_\ell}$ viene dada por:

$$a\rightsquigarrow_{\mathcal{T}_\ell} b = min(1, 1 – a + b) \quad \forall(a,b)\in[0,1]\times[0,1].$$

En este caso, las expresiones (1) para los operadores erosión y dilatación son, para cada $(x,z)\in U\times U$:

$$\begin{aligned}\xi_R(Q)(x,z) &= inf\{min(1, 1 - R(y,x) + Q(y,z)) / y\in U\}\\ \delta_R(Q)(x,z) &= sup\{max(0, R(x,y) + Q(y,z) -1) / y\in U\}\end{aligned} \quad (3)$$

El interés de este caso reside en que la t-norma $\mathcal{T}_\ell$ y la implicación $\rightsquigarrow_{\mathcal{T}_\ell}$ están además relacionadas por medio de la negación de Zadeh $a' = 1 – a$ como sigue:

$(a \ \mathcal{T}_\ell b')' = 1–(a \ \mathcal{T}_\ell(1– b)) = a\rightsquigarrow_{\mathcal{T}_\ell} b \ \forall(a,b)\in[0,1]\times[0,1]$,

y como consecuencia, se induce una relación de dualidad entre los operadores erosión y dilatación morfológicos:

$$\xi_R(Q') = (\delta_R^{op}(Q))' \quad \forall Q\in[0,1]^{U\times U} , \quad (4)$$

donde $S'(x,y) = 1–S(x,y) \quad \forall S\in[0,1]^{U\times U} \quad \forall(x,y)\in U\times U.$

Aunque el *elemento estructurante* puede ser cualquier relación difusa $R\in[0,1]^{U\times U}$, hacemos notar que aparece otro caso distinguido cuando son *relaciones nítidas* R, es decir, tales que $R(U\times U) \subseteq \{0,1\}$. Ahora R_x y $(R^{op})_x$ son los subconjuntos $R_x = \{y\in U / xRy\}$, $(R^{op})_x = \{y\in U / yRx\}$.

En este caso, las expresiones (1) de la erosión y dilatación morfológicas no dependen de $\mathcal{T}$ ni de $\rightsquigarrow_{\mathcal{T}}$, ya que son :

$$\begin{aligned}\xi_R(Q)(x,z) &= \begin{cases} 1 \text{ si } R_x^{op} =\emptyset \text{ ó si } (Q(y,z)=1 \ \ \forall y\in R_x^{op}) \\ inf\{Q(y,z) / y\in U \& R(y,x)=1\} \text{ en otro caso}\end{cases}\\ \delta_R(Q)(x,z) &= \begin{cases} 0 \text{ si } R_x =\emptyset \text{ ó (si } Q(y,z)=0 \ \ \forall y\in R_x) \\ sup\{Q(y,z) / y\in U \& R(x,y)=1\} \text{ en otro caso}\end{cases}\end{aligned} \quad (5)$$

3. DEFINICIÓN DE APERTURA Y CIERRE MORFOLÓGICOS DE RELACIONES DIFUSAS

Fijada una t-norma $\mathcal{T}$ continua por la izquierda en [0,1], definimos los operadores *apertura* y *cierre* sobre relaciones $Q \in [0,1]^{U\times U}$ como composición de erosión y dilatación:

Definición 2.- Sea $Q \in [0,1]^{U\times U}$ una relación difusa en U y R una relación estructurante. Los filtros *apertura morfológica* γ_R y *cierre morfológico* ζ_R de la relación Q mediante la relación estructurante R, vienen dados por:

$$\gamma_R(Q) = \delta_R(\xi_R(Q)) \ , \ \zeta_R(Q) = \xi_R(\delta_R(Q)), \qquad (6)$$

es decir :

$$\gamma_R(Q) = R \circ_{\mathcal{T}} (R^{op} \triangleleft_{\mathcal{T}} Q) \ , \ \zeta_R(Q) = R^{op} \triangleleft_{\mathcal{T}} (R \circ_{\mathcal{T}} Q).$$

Proposición 2.- *Sean S y Q dos relaciones difusas de U, y sea R una relación estructurante de U. Entonces los operadores apertura y cierre verifican:*

1) Si $S \le Q$, entonces: $\gamma_R(S) \le \gamma_R(Q)$ & $\zeta_R(S) \le \zeta_R(Q)$.

2) $\gamma_R(Q) \le Q \le \zeta_R(Q) \quad \forall Q \in [0,1]^{U\times U}$.

3) $\gamma_R(\gamma_R(Q)) = \gamma_R(Q)$ & $\zeta_R(\zeta_R(Q)) = \zeta_R(Q) \quad \forall Q \in [0,1]^{U\times U}$.

4) Además, si la relación estructurante R es nítida ó si la t-norma con la que se construye la dilatación es la de Lukasiewicz, se verifica:

$$(\gamma_R^{op}(Q))' = \zeta_R(Q') \quad \forall Q \in [0,1]^{U\times U}.$$

5) Si la relación R es reflexiva y $\le$ *representa la extensión puntual a* $\in [0,1]^{U\times U}$ *del orden en* [0,1]*, entonces*

$$\xi_R(Q) \le \gamma_R(Q) \le Q \le \zeta_R(Q) \le \delta_R(Q) \quad \forall Q \in [0,1]^{U\times U}.$$

Nota. En el retículo $([0,1]^{U\times U}, \le)$, γ_R es un *operador interior* y ζ_R es un *operador clausura*.

4. FILTROS "TOP-HAT" E "HIT-OR-MISS"

En morfología matemática sobre imágenes se consideran otros filtros en cuya construcción interviene el complementario (el "negativo") Q' de la imagen Q. Se pueden utilizar para realzar fronteras o para localizar formas específicas dentro de una imagen. Definiremos dos de esos filtros en el contexto de bases de datos.

Definición 3. Sea $Q \in [0,1]^{U\times U}$ una relación difusa en U y R una relación estructurante. Definimos el *"Top Hat" por la apertura* $TH_{\gamma R}$ y el *"Top Hat" por el cierre* $TH_{\zeta R}$ de Q asociados a la estructurante R y a la t-norma $\mathcal{T}^*$ por:

$$TH_{\gamma R}(Q) = Q \ \boldsymbol{\mathcal{T}}^* (\gamma_R(Q))' \ , \ TH_{\zeta R}(Q) = \zeta_R(Q) \ \boldsymbol{\mathcal{T}}^* Q' \ .$$

(La t-norma $\mathcal{T}^*$ puede ser distinta o coincidir con la $\mathcal{T}$ utilizada para definir la dilatación δ_R y la erosión ξ_R).

Son filtros *diferencia* que determinan relaciones difusas que son fronteras interna y externa de la relación Q.

Definición 4. Sea $Q \in [0,1]^{U\times U}$ una relación difusa en U y $\boldsymbol{R} = (R_1, R_2)$ un par de relaciones estructurantes nítidas tales que $R_1 \cap R_2 = \varnothing$. Definimos el *"Hit or Miss" de Q asociado a* $\boldsymbol{R}$ y a t^* por la $\boldsymbol{\mathcal{T}}^*$-intersección de erosiones:

$$HM_{\boldsymbol{R}}(Q) = HM_{(R_1,R_2)}(Q) = \xi_{R_1}(Q) \ \boldsymbol{\mathcal{T}}^* \xi_{R_2}(Q') \ .$$

Nota. $\forall R$ y $\forall Q \in [0,1]^{U\times U}$ se verifica: $HM_{(R,\varnothing)}(Q) = \xi_R(Q)$.

5. APLICACIÓN A PROCESOS DE KDD: GRAFOS ASOCIADOS A FILTROS MORFOLÓGICOS

Con el fin de utilizar estos filtros como herramientas en procesos de descubrimiento de conocimiento en bases de datos (*KDD*), les asociaremos ciertos grafos como sigue.

Dada la relación difusa $Q \in [0,1]^{U\times U}$, entonces $(Q_\alpha)_{\alpha\in]0,1]}$ y $sop(Q) = \{(x,y) \ / \ Q(x,y) > 0\}$ son respectivamente la familia de sus α-cortes y su soporte. Tanto $sop(Q)$ como cada Q_α son subconjuntos ordinarios de $U\times U$.

Si Q^* es un subconjunto del soporte de Q, $(Q^* \subseteq sop(Q))$, la expresión U_{Q^*} representará el subconjunto de U tal que

$$U_{Q^*} = \{ x \ / \ (x,y) \in Q^* \} \cap \{ y \ / \ (x,y) \in Q^* \}.$$

A cada Q^* le asociamos el *grafo dirigido* $G_{Q^*} = (U_{Q^*}, Q^*)$ con conjunto de vértices U_{Q^*} y conjunto de aristas Q^*.

Con esta consideraciones, tanto la relación difusa inicial $Q \in [0,1]^{U\times U}$, como las $F(Q) \in [0,1]^{U\times U}$ que se obtengan en un proceso de filtrado morfológico, tendrán asociadas familias de grafos dirigidos $(G_{Q^*})_{Q^* \subseteq sop(Q)}$, $(G_{Q^*})_{Q^* \subseteq sopF(Q)}$ construidos tal como se indica en los párrafos anteriores.

Explicación de la Metodología utilizando un ejemplo

En los puntos A1, A2, A3 y A4 de este apartado describimos las distintas fases del proceso de extracción de información de una base de datos mediante operadores morfológicos. Usamos para ello un ejemplo.

Concretamente, extraemos el ejemplo del conjunto V de registros www.isc.uic.edu/~mlearn/MLSummary.html, que constituye una base relacionada con el censo de EE.UU. Tiene 15.000 registros y en ella se recogen datos reales relativos a personas. Son registros anónimos en los que se ha sustituido el nombre por un número identificador. Los campos de la base de datos aparecen en la Tabla 1.

Campos de la Tabla V	
Ca1	Edad
Ca2	Tipo de empleo
Ca3	Sueldo
Ca4	Estudios
Ca5	Años de estudio
Ca6	Estado civil
Ca7	Empleo
Ca8	Estado actual
Ca9	Etnia
Ca10	Sexo
Ca11	Ganancia
Ca12	Pérdida
Ca13	Horas de trabajo a la semana
Ca14	País de nacimiento

Tabla 1

A1. En primer lugar, seleccionar un subconjunto U de V que hará el papel de universo. Puede coincidir con V, ser un subconjunto aleatorio, o un subconjunto propio determinado por algún criterio, etc. En nuestro ejemplo U *está formado por los registros pertenecientes a la base de datos* www.isc.uic.edu/~mlearn/MLSummary.html *tales que* $Ca14$ = "*Mexico*". Su cardinal es $Card(V) = 293$, y de ellos 61, (el 21,82 %), son mujeres.

A2. Selección de una relación difusa $S \in [0,1]^{U \times U}$ entre registros del universo U. Esta relación S será el elemento del que queremos extraer información utilizando filtros morfológicos. Puede ser una relación definida de manera aleatoria o estar definida mediante condiciones sobre los campos de la base de datos. En el ejemplo utilizamos esta segunda opción. Concretamente, definimos S como:

$S(x,y) =$

si $[x.Ca2 = y.Ca2]$ ***entonces*** $1 - \dfrac{|x.Ca1 - y.Ca1|}{57}$, ***en otro caso*** 0.

S es una relación reflexiva y simétrica (tolerancia difusa) cuya descripción es la siguiente: *"En el colectivo U, consideramos que las personas x e y tienen un cierto grado de semejanza S(x,y) si trabajan en empresas u organismos del mismo tipo y tienen una edad parecida"*.

A3. Implementación de los algoritmos que definen las relaciones estructurantes nítidas $R_1, R_2, \ldots$, que se utilizarán para construir los correspondientes filtros $F_{R1}, F_{R2}, \ldots F_{Rn}$ que han de transformar la relación S para conseguir las nuevas relaciones difusas $F_{R1}(S), F_{R2}(S), \ldots F_{Rn}(S)$.

Los operadores erosión necesarios para el ejemplo vienen dados por (5) y se han implementado en una base de datos *ACCESS*, usando las sentencias de *SQL* para la selección de registros.

El filtro F seleccionado para tratar la relación S es el "Hit or Miss" $HM_{\mathbf{R}}$ (definición 4). Hemos utilizado para su implementación $T^* = \wedge$ y las estructurantes nítidas R_1 y R_2:

$R_1(x,y) =$

si $[|x.Ca3 - y.Ca3| < 5000$ **&** $|x.Ca5 - y.Ca5| < 2$ **&** $x.Ca10 = y.Ca10]$ ***entonces*** 1, ***en otro caso*** 0.

$R_2(x,y) =$

si $[|x.Ca3 - y.Ca3| < 5000$ **&** $|x.Ca5 - y.Ca5| < 2$ **&** $x.Ca10 \neq y.Ca10]$ ***entonces*** 1, ***en otro caso*** 0.

Los cardinales del soporte y del 1-corte de la relación $HM_{\mathbf{R}}(S)$ obtenida son:

$$Card(sop(HM_{\mathbf{R}}(S))) = 58655, \quad Card((HM_{\mathbf{R}}(S))_1) = 339.$$

A4. Una vez determinadas las relaciones difusas $F_{R1}(S)$, $F_{R2}(S)$,..., y, utilizando herramientas de representación de grafos, se les asocia familias de grafos dirigidos tal como se ha explicado al comienzo de este apartado 5.

Si es conveniente, se puede reducir la cardinalidad de los conjuntos de vértices y aristas de los grafos que se obtienen utilizando información proporcionada por los campos de la base de datos. Por ejemplo, podemos exigir que alguno o ambos elementos de los pares (x,y) de los α-cortes cumplan condiciones adicionales, condiciones que se expresan como restricciones sobre alguno de los campos $Ca1,\ldots,Ca14$ que aparecen en la Tabla 1.

Finalmente, estos últimos grafos dirigidos serán los elementos de los que extraemos información de la base de datos U.

Los grafos de nuestro ejemplo se han representado con la utilidad de libre acceso para análisis de grafos: *NetDraw* (http://www.analytictech.com/Netdraw/netdraw.htm). En los grafos, la etiqueta de cada uno de los vértices se corresponde con el número de identificación del registro correspondiente en U. Si la cadena *"*Loop*"* aparece en el identificador del vértice x, significa que (x,x) es una arista del grafo. Los vértices circulares representan los registros tales que $Ca10$ = *"mujer"* y los vértices cuadrados para los que $Ca10$ = *"hombre"*. Una línea fina indica arista en un único sentido (el de la flecha). Las líneas más gruesas son aristas en los dos sentidos.

Por ejemplo, en la Figura 1 del ANEXO se muestra el grafo con conjunto de aristas $Q^* = (HM_{\mathbf{R}}(S)_{0.95}_x30)$ de elementos (x,y) tales que $HM_{\mathbf{R}}(S)(x,y) \geq 0.95$ y que cumplen además que la edad de x es mayor o igual que 30 ($x.Ca1 \geq 30$). Este grafo tiene 632 aristas y 158 vértices. El 15,82% son mujeres. En la Figura1, la arista $(x,y) \in (HM_{\mathbf{R}}(S)_{0.95}_x30)$ representa registros x, y tales que:

a1) La persona x tiene 30 años o más.

b1) Toda z que sea del mismo sexo que x y con estudios y sueldo semejantes a x, trabaja en una institución del mismo tipo que y y su diferencia de edad con este y es menor o igual que 2.

c1) Toda persona t de distinto sexo que x con estudios y sueldo semejantes a x, trabaja en una institución de distinto tipo a la del trabajo de y o su diferencia de edad con y es mayor o igual que 55.

En la Figura 2 del ANEXO se muestra el grafo con conjunto de aristas $Q^* = (HM_{\mathbf{R}}(S)_{0.97}_x30)$. Consta de 366 aristas y 132 vértices. La interpretación de $(x,y) \in (HM_{\mathbf{R}}(S)_{0.97}_x30)$ es análoga a la anterior, aunque el número de aristas es menor. El 15,9% son mujeres.

En la Figura 3 del ANEXO se muestra el grafo con conjunto de aristas $Q^* = (HM_{\mathbf{R}}(S)_{1.00}_x30)$. Consta de 143 aristas y 93 vértices. El 11,82% son mujeres.

Aquí $(x,y) \in (HM_{\mathbf{R}}(S)_{1.00}_x30)$ significa que:

a3) Los individuos x e y son de la misma edad y trabajan en instituciones del mismo tipo.

b3) Toda persona z R_1-*semejante* á x (sueldos y estudios similares y del mismo sexo), trabaja en una institución del mismo tipo que x y tiene la misma edad que x.

c3) Toda persona z R_2-*semejante* á x (sueldos y estudios similares pero de distinto sexo, trabaja en una institución de distinto tipo que x o su diferencia de edad con x es de 57 años.

En el grafo de la Figura 3 del ANEXO, cada uno de los subgrafos aislados es tal que los vértices representan personas de la misma edad y que trabajan en el mismo tipo de institución. Por ejemplo, en el subgrafo etiquetado por *"Subg_A"*, los vértices se corresponden con personas que tienen 41 años y trabajan en la empresa privada.

En este grafo, los pares de vértices (x,y) unidos con líneas de mayor grosor, son tales que:
d3) $S(x,y) = 1$
e3) $S(z,x)= 1\ \forall z\in(R_1)_x\cup(R_1)_y$ & $S(z,x)= 0\ \forall z\in(R_2)_x\cup(R_2)_y$

Para los pares de vértices (x,y) unidos con líneas más finas se verifica:
f3) $S(x,y) = 1$
g3) $S(z,x) = 1 \quad \forall z\in(R_1)_x$ & $S(z,x) = 0 \quad \forall z\in(R_2)_x$
h3) $\exists z\in(R_1)_y$ tal que $S(z,x)\neq 1$ ó $\exists z\in(R_2)_y$ tal que $S(z,x)\neq 0$.

Finalmente, señalemos que la complejidad de los algoritmos utilizados en este trabajo para implementar los operadores morfológicos es $O(n^3)$, siendo n el número de registros del referencial U. Si, como hemos expuesto en este apartado, pretendemos obtener una representación de resultados mediante grafos, tendremos que considerar además el costo de la conversión de ficheros de la base de datos al formato utilizado por la herramienta de grafos (en nuestro caso *NetDraw*) y el costo asociado a los análisis realizados con esa herramienta.

5. CONCLUSIONES

Del contenido expuesto en este trabajo podemos concluir:

i) Que la aplicación de filtros morfológicos F a relaciones difusas entre ítems de una base de datos U puede utilizarse para extraer conocimiento no evidente sobre los registros de esa base.

ii) Que los filtros obtenidos por medio de relaciones estructurantes R son filtros de "R-vecindad", característica que introduce una componente "geométrica" en procesos de *KDD* cuando se utiliza la metodología aquí expuesta.

iii) Que el conocimiento obtenido de U mediante los filtros F puede expresarse mediante familias de grafos.

En futuros trabajos extenderemos otros filtros clásicos sobre imágenes al contexto de bases de datos.

REFERENCIAS

[1] B. de Baets, Fuzzy morphology: a logical approach, in Uncertainty Analysis, B. Ayyub and M. Gupta (eds.), 53-67, Kluwer Academic Publishers 1997

[2] Bloch,I. H.Maître. Fuzzy Mathematical Morphologies A comparative study, Pattern Recognition, Vol 28, pp. 1341-1387 (1995).

[3] P. Burillo, R. Fuentes-González, N. Frago, Generation of Fuzzy Mathematical Morphologies. Mathware & Soft Computing Vol VIII, n 1 31-46 (2001).

[4] Erich, P.K., Mesiar R. and Pap E. Triangular Norms. Kluwer Academic Publishers (2000).

[5] Frago, N. y R. Fuentes-González. Aportación al proceso de Descubrimiento de Conocimiento utilizando técnicas de Morfología Matemática. Estylf 2006. Actas del XIII Congreso Español Sobre Tecnologías y Lógica Fuzzy. Ciudad Real (España), Septiembre (2006).

[6] Frago, N. y R. Fuentes-González. Descubrimiento de Conocimiento en Bases de Datos utilizando técnicas de Morfología Matemática Borrosa, Información Tecnológica Vol. 18(6), 39-50 (2007).

[7] Fuentes-González, R., The incorporation of the Mathematical Morphology, the Formal Concept Analysis and the Fuzzy Logic techniques and tools to the data cleaning, aggregating, reducing and mining stages of the Knowledge Discovery process, J.M. Troya and S. Ossowski (Eds.), 147-154. Actas V Jor. Seg. Proy. en tecnología Informática. CEDI 2005. Granada.

[8] Fuentes-González, R., P. Burillo y N. Frago. Técnicas de Morfología Matemática en el tratamiento de Sistemas Relacionales Difusos, Estylf 2002. Actas del XI Congreso Español Sobre Tecnologías y Lógica Fuzzy. León (España), Septiembre (2002).

[9] L.J. Kohout, W. Bandler, Use of fuzzy relations in knowledge representation, acquisition and processing, in: L. Zadeh, J. Kacprzik (eds.), Fuzzy Logic for Management of Uncertainty, 415-435. John Willey & Sons (1992).

[10] Matheron, D. Elements pour une théories des milieux poreux, Masson, Paris (1967).

[11] Serra, J. Image Analysis and Mathematical Morphology. Ac. Press Vol I (1982) -II (1988).

ANEXO (apartado 5). Grafos dirigidos asociados a la relación difusa $HM_R(S)$, obtenidos con la utilidad *NetDraw*.

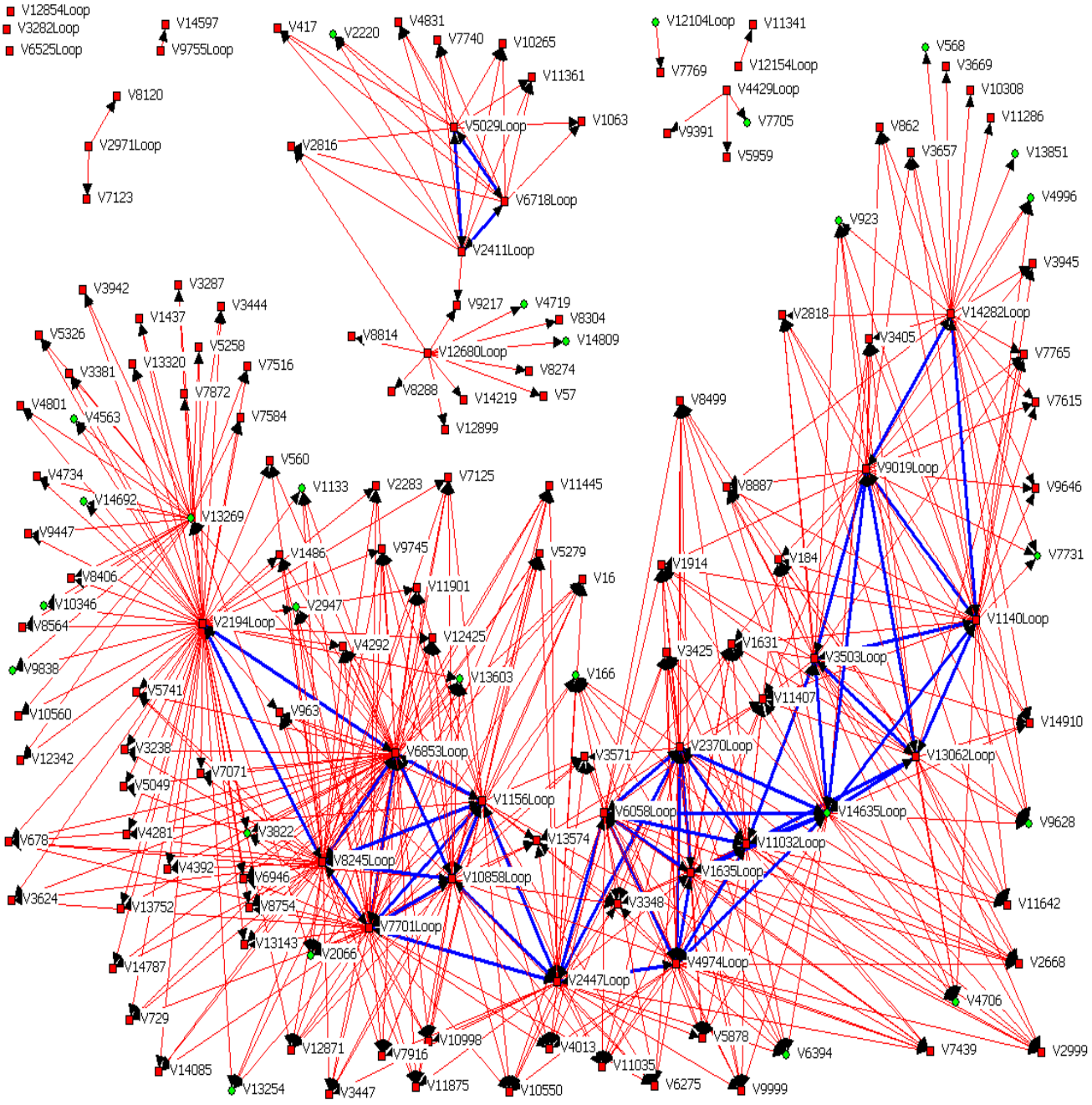


Figura 1: Grafo dirigido asociado a $(HM_{(R_1,R_2)}(S)_{0.95}_x30)$

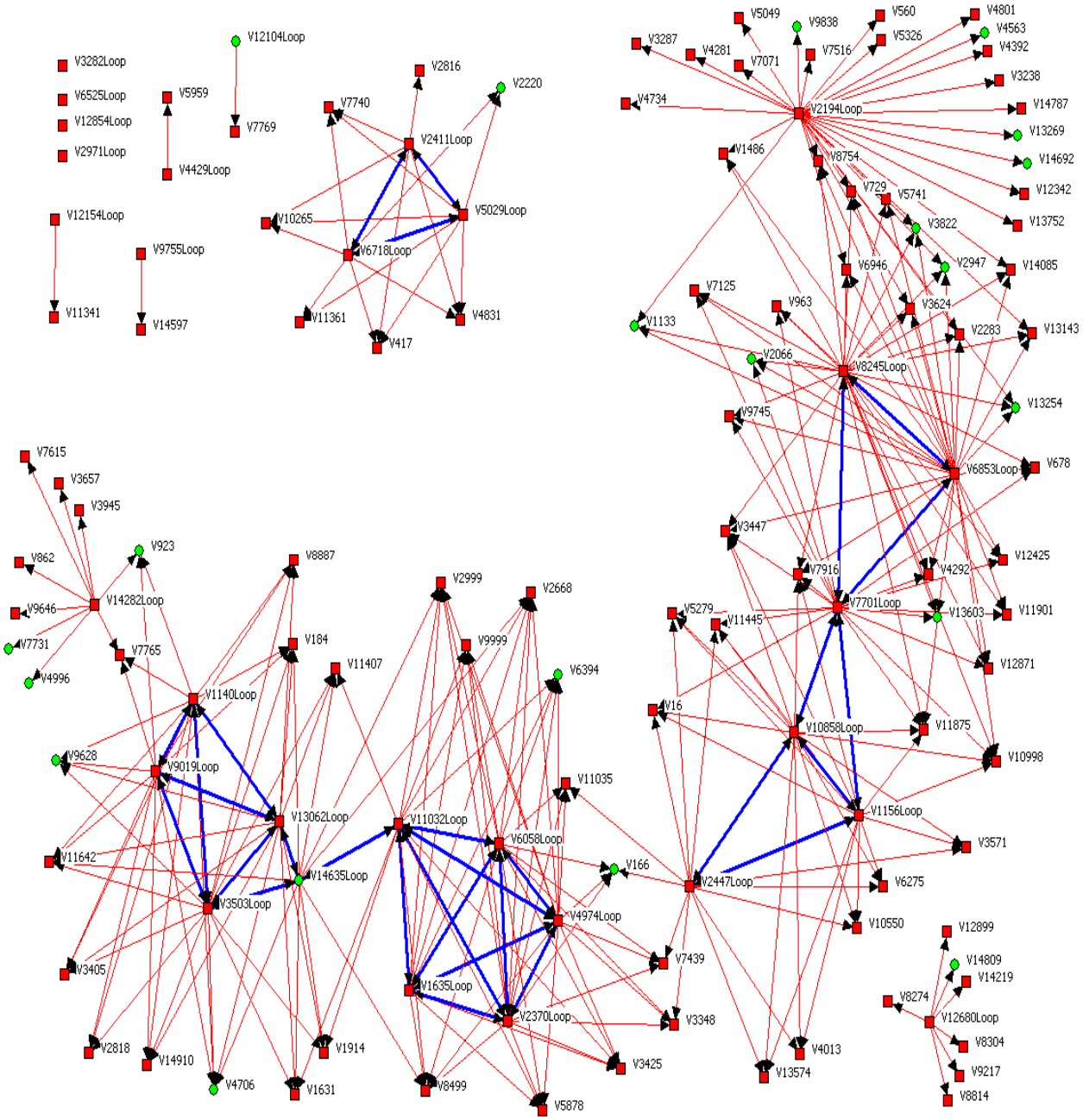


Figura 2: Grafo dirigido asociado a $(HM_{\mathbf{R}}(S)_{0.97}_x30)$

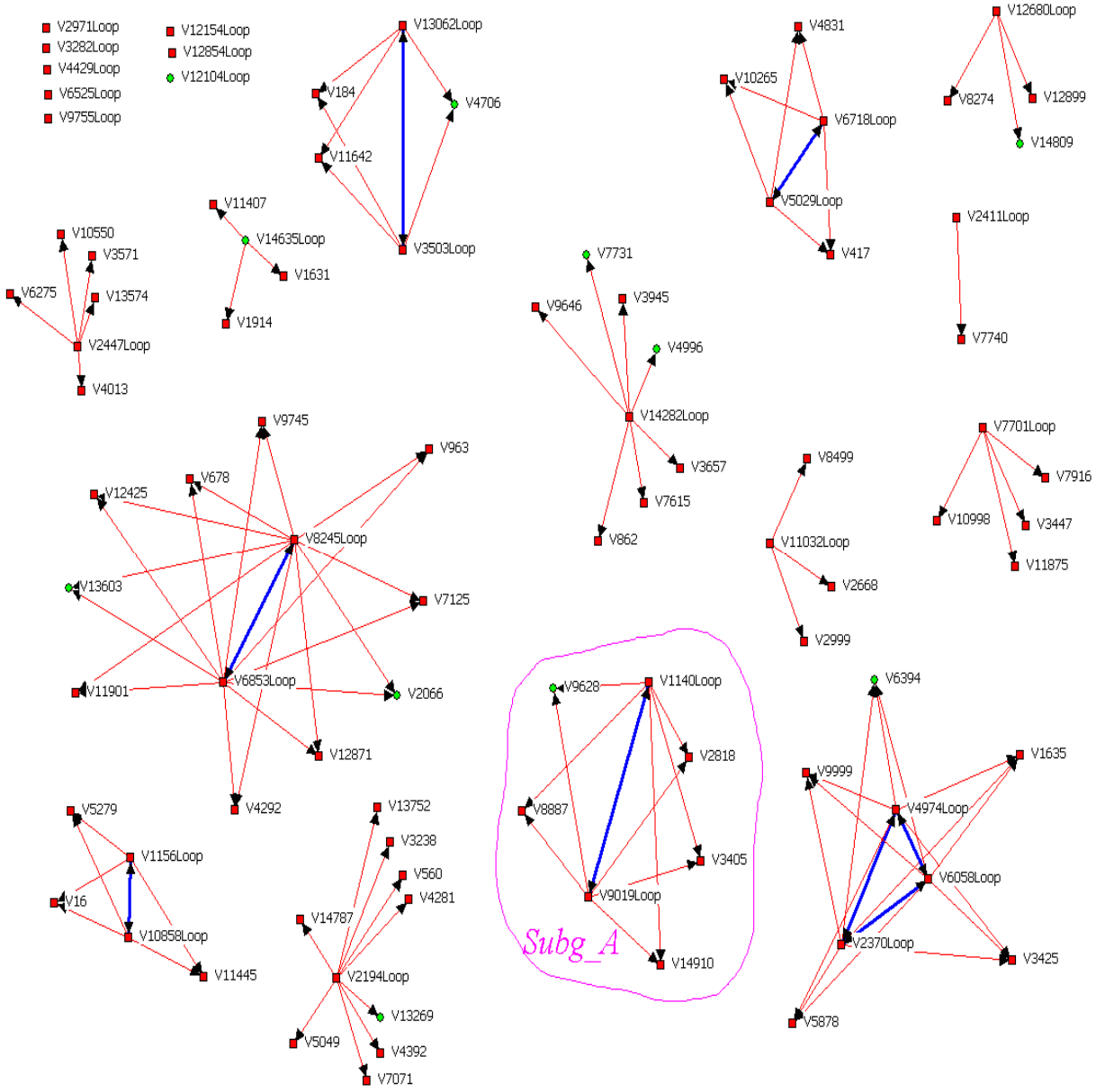


Figura 3: Grafo dirigido asociado a ($HM_{\mathbf{R}}(S)_{1.00}_x30$)

ARQUITECTURA ESCALABLE PARA LA DETECCIÓN DE PATRONES SECUENCIALES DIFUSOS EN MINERÍA DE DATOS CUANTITATIVA

Pablo F. Provasi[1] **Lucio J. Kleisinger**[1] **Francisco R. Villatoro**[2]

[1] Dpto. de Informática, Universidad Nacional del Nordeste - Argentina, pprovasi@exa.unne.edu.ar, lucio@edsistemas.com
[2] Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga - España, villa@lcc.uma.es

Resumen

Se presenta un nuevo enfoque esquemático general, una arquitectura escalable, que permite servir de guía en la construcción de herramientas de minería para la detección de patrones secuenciales difusos en bases de datos de transacciones comerciales. Con el propósito de que estos patrones puedan ser utilizados mediante otras metodologías para decidir políticas de mercadeo, se menciona a modo de ejemplo una base de datos de Supermercado. El trabajo presenta un aspecto teórico del enfoque esquemático, dejando para más adelante la asignación puntual de las diversas herramientas de aplicación a cada etapa del modelo.

Palabras Clave: Arquitectura Guía, Minería, Patrones Secuenciales, Difusos.

1. INTRODUCCIÓN

La minería de datos cuantitativa pretende la extracción de la información que presenta ciertos modelos o patrones a partir de los registros almacenados en grandes bases de datos. Cuando estos patrones involucran cierta secuencia temporal nos referimos a la detección de patrones secuenciales. El problema de la minería de patrones secuenciales en bases de datos de transacciones comerciales fue introducido y resuelto utilizando bases de datos (en adelante DB) con información *crisp* por Agrawal y Strikant [1]. El uso de técnicas *fuzzy* en este contexto fue iniciado por Chen [2, 3]. En este trabajo nos basaremos en el reciente desarrollo de Céline Fiot [4], trabajo que completaremos y extenderemos con énfasis en las bases de datos de transacciones comerciales en supermercados y tiendas virtuales.

Los contenidos del trabajo son los siguientes. En la siguiente sección se introduce un ejemplo que muestra el problema que queremos resolver. En la sección 3 presentaremos las ideas básicas sobre la minería de patrones secuenciales con datos *crisp*. A la minería de datos de patrones secuenciales con datos *fuzzy* se dedica la sección 4. Las ideas fundamentales sobre nuestra nueva arquitectura se presentan en la sección 5. Finalmente, la sección 6 presenta las conclusiones más importantes de este trabajo.

2. EJEMPLO EN UN SUPERMERCADO

El sistema de administración de un supermercado puede almacenar y mantener diversas DB, una de las cuales almacenará los registros de las ventas. Las compras efectuadas por un cliente deben ser consideradas registros de la DB, y cada registro está unívocamente relacionado con un identificador o número de factura. En tiendas virtuales, como las consideradas por Fiot [4], el cliente se puede considerar como un agente en el sistema. Sin embargo, en nuestro caso supondremos que en la DB del supermercado los clientes son anónimos, y sólo se conoce el número de factura o recibo con la lista de productos.

Los datos simples o aislados contenidos en una base de datos ofrecen muy poca o ninguna información que pueda ser de utilidad para el administrador del supermercado. Sin embargo, a medida que estos datos crecen en cantidad (número de registros), se pueden agrupar y relacionar de diversas formas, de manera que se pueda extraer una mayor cantidad de información utilizando técnicas de minería de datos. Surge de lo anterior, que los niveles de información (a) *generados* o (b) *obtenidos* pueden variar significativamente en función de como se organizan: (1) ordenan, (2) agrupan o (3) clasifican los datos. El análisis de estos datos previamente organizados permite (a) generar nueva información mediante la aplicación de reglas de consultas así como (b) obtener información de los mismos.

Ejemplo 1: Sobre la información (b) obtenida de la simple observación de los datos (2) agrupados de una forma determinada, para este caso, en 2 grupos de interés: los que han comprado latas de tomates y los que han comprado latas de pimientos; existen otros grupos que surgen naturalmente y

que pueden o no ser de interés según lo que se desee observar, que son el grupo de los que no pertenecen a ninguno de los 2 grupos de interés y el grupo intersección. De esto podemos inferir mucha información *crisp*: primero, cantidad de clientes que han comprado alguno de los 2 productos, la cantidad que no lo ha hecho y los que han comprado ambos productos. Y si los comparamos entre si de diversas formas, obtendríamos información que nos ayudaría a decidir si nos conviene abastecernos en proporciones similares de estos productos o si nos convendría comprar mayoritariamente un producto en vez de otro. Queda claro que con un simple proceso de agrupación es enorme la cantidad de información que se puede obtener.

Ejemplo 2: Sobre la información (a) generada, siguiendo con el ejemplo anterior. Si a estos grupos además los (3) clasificamos aplicando alguna regla de restricción temporal, podríamos obtener varios grupos más o subgrupos. Surge de esto que preparar grupos con algún significado especial requiere de un proceso de clasificación (aplicación de restricciones a la agrupación).

Bien, hasta aquí parecería que se puede manejar esto con una simple consulta (query) elaborada de una base de datos. En muchos casos esto puede ser suficiente. Sin embargo, el propósito de encontrar patrones secuenciales difusos en una DB es principalmente obtener aquella información que no resulta trivial u obvia de un simple análisis. Para lo que es necesario la aplicación de algunos principios [4]:

I La conversión de los datos cuantitativos en conjuntos difusos y grados de pertenencia.

II El uso de estas titulaciones para extraer las secuencias frecuentes.

Para la etapa (i), convertir datos cuantitativos de tipo *crisp* en conjuntos difusos requiere de un pre procesamiento según se explicó anteriormente. Este pre-procesamiento demanda algún consumo de recursos computacionales que en ocasiones puede ser muy costoso. El propósito de este pre procesamiento consiste en encontrar (agrupar, ordenar, clasificar) *patrones secuenciales* de tipo *crisp*. Un patrón no es otra cosa que un *atributo presente* en un registro de la DB. Los algoritmos de búsquedas de patrones procuran encontrar (o formar) secuencias de atributos presentes, y a esto se le llama minería de patrones secuenciales.

3. MINERÍA DE PATRONES SECUENCIALES (MPS)

Antes de presentar la minería de patrones secuenciales *fuzzy*, consideraremos datos *crisp*. Sea **DB** un conjunto de registros (objetos), donde cada registro **R** consiste de tres elementos de información:

- un identificador de registro **object-id**,
- un registro de tiempo **timestamp**, y
- un conjunto **A** de atributos binarios.

Un atributo **a** puede estar presente '1', o ausente '0'. Si bien los algoritmos intentan encontrar secuencias de atributos presentes, esta claro que su ausencia también brinda significativa información. Un atributo **a** presente es llamado *ítem* **i**, y dentro del contexto de números *crisp*, un *ítem* es una cupla **(atributo, valor)**. Un *conjunto de items* $\{i_1, i_2, \cdots, i_k\}$ es denotado por **I**, donde **I** es un sub conjunto de **A**. Esto es una representación no ordenada. Una *secuencia* **s** es una lista ordenada no vacía de los conjuntos de items **s**, denotado por $\langle s_1, s_2, \cdots, s_p \rangle$. Una **n-secuencia** es una secuencia de **n** items o de tamaño **n**.

Decimos entonces que **R** es un conjunto de registros que contienen los atributo presente (o ítems) **i**. Gráficamente $R_{(I)}$ podría verse así como en la Figura 1.

Object-id	time-stamp	a_1	a_2 ítem-i 2	• • •	a_2 ítem-i *n*
		ausente	presente		presente

Figura 1: Algunos Atributos Presentes o items

3.1. SECUENCIA DE DATOS

Una secuencia de datos $\langle s_1, s_2, \cdots, s_p \rangle$ es una agrupación ordenada de *registros con atributos presentes iguales*, similares o pertenecientes. La agrupación se ordena según los datos de **timestamp**.

3.2. FRECUENCIA DE SECUENCIAS

Los grados de pertenencia se obtienen al evaluar la frecuencia de ocurrencia de las secuencias **freq(s)** con distintos niveles de fuzzyficación. Algunos autores establecen un valor de referencia o parámetro, por ejemplo **minFreq**, con el fin de decidir si una frecuencia es más o menos frecuente. En estos casos se buscan las secuencias de mayor frecuencia, $freq(s) > minFreq$.

4. MINERÍA DE PATRONES SECUENCIALES DIFUSOS (MPSD)

Existen diferentes enfoques para extraer patrones secuenciales difusos, pero en este caso adoptamos una variante basada en Fiot [4], pero todos ellos comienzan por un paso común, donde el universo de cada ítem cuantitativo es particionado en varios conjuntos difusos.

En el contexto de lógica difusa, un *ítem difuso* es la asociación de un *conjunto difuso* **b** a su correspondiente *ítem* **x**, es

decir **[x, b]**. En la DB cada *registro* está asociado a un *ítem difuso* **[x,b]** según su *grado de pertenencia*. Un conjunto de items difusos será denotado por el par **(X,B)**, donde **X** es el conjunto de items, y **B** es un conjunto de conjuntos difusos correspondientes.

4.1. SECUENCIA DE DATOS DIFUSOS

Por otra parte tenemos que, una secuencia **g-k-sequence** $\langle s_1, s_2, \cdots, s_p \rangle$ está formada por g conjuntos de items difusos $s = (X, B)$ agrupados a **k** items difusos **[x,b]**.

4.2. FRECUENCIA DIFUSA DE DATOS DIFUSOS

De acuerdo a las definiciones anteriores podemos decir que la minería de patrones secuenciales difusos consiste en encontrar la frecuencia máxima difusa **g-k-sequence**.

En su trabajo, Fiot et al [4], ofrece una definición general de *frecuencia de una secuencia*, también presenta tres algoritmos para el cálculo de frecuencia de patrones secuenciales difusos que pueden ser de utilidad cuando se construyen herramientas para la minería de patrones secuenciales. Por ejemplo:

- **SpeedyFuzzy:** se basa en la cuenta "admitir / no admitir". Cuenta todos los objetos o items difusos de un conjunto, sin importar el grado, si es mayor que 0 los objetos tienen el mismo peso.
- **MiniFuzzy:** se basa en un umbral de cuenta. Cuenta los objetos o items difusos de un conjunto, pero solo admite aquellos items de la secuencia candidato que tienen un grado de pertenencia mayor a un umbral especificado.
- **TotallyFuzzy:** se basa en una cuenta sigma-contar con umbral. Cuenta cada objeto y cada secuencia. Se toma en cuenta la importancia del conjunto o secuencia de datos, y se considera el mejor grado de pertenencia.

5. ARQUITECTURA GUÍA PARA LA CONSTRUCCIÓN DE HERRAMIENTAS (MPSD)

Este esquema permite oficiar de guía en la producción de herramientas para la minería de patrones secuenciales difusos. Consiste en seguir el camino adecuado según el caso de estudio (ver Figura 2).

El modelo de la arquitectura consta de dos fases: *minería de datos cuantitativos* y *minería de datos difusos*.

En la figura vemos la "Etapa de Clasificación", que consiste en generar grupos o bases de datos de **patrones presentes** luego de aplicar uno de los porcesos alternativos de clasificación. Siguiendo el modelo, se pueden agregar **n** algoritmos con el objeto de probar diferentes tipos de agrupaciones.

Una vez obtenido un conjunto de bases de datos clasificados como salida de la ejecución del proceso elegido, será factible pasar a la etapa siguiente de "Secuenciado", en la cual se podrá obtar por alguno de los algoritmos de ordenamiento conocidos o diseñado con algún criterio especial. Estos algoritmos podrán ser elegidos en base al criterio de ordenamiento más adecuado para las DB generadas en la etapa anterior. Como resultado del secuenciado se optine un nuevo conjunto de bases de datos, agrupadas, clasificadas y ordenadas. Hasta aquí podemos decir que se ha completado la fase correspondiente a *minería de datos cuantitativos*.

La fase de *minería de datos difusos*, presenta dos etapas, "Fuzzyficación" y "Detección y Conteo de freq(s)". La primera consiste en seguir un procedimiento similar a los descritos en las etapas anteriores, seleccionar (o producir) el algoritmo de fuzzyficación adecuado para el tipo de datos contenido en las DB de entrada y efectuar el procesamiento. Como resultado, se obtendrán **n** bases de datos de secuencias difusas.

Ahora bien, en la figura podemos ver la etapa final "Detección y conteo de freq(s)", y vemos que los cuadros que representan procesos, por ejemplo (Algoritmo 1 Detec. y Conteo), pueden ser reemplazados por cualquiera de los tres algoritmos referenciados en la sub sección anterior. También es factible diseñar tantos algoritmos como la necesidad justifique. Por ello la arquitectura de guía pretende aclarar el panorama para quienes se dedican a experimentar en la minería de patrones secuenciales difusos.

6. CONCLUSIONES

En este trabajo nosotros presentamos una introducción teórica de las etapas que implica el desarrollo de herramientas para la minería de patrones secuenciales difusos, y presentamos un esquema general que sirve de guía en la arquitectura de estas herramientas. Los bloques del esquema en cada etapa pueden ser reemplazados por nombres propios de nuevos algoritmos que permitan cumplimentar los pasos requeridos para la MPSD.

Referencias

[1] R. Agrawal, R. Srikant. Mining sequential patterns. *11th International Conference on Data Engineering* (ICDE'95), Taipei, Taiwan, Pág. 3, 1995. Expanded version available as IBM Research Report RJ9910, October 1994.

[2] R.-S. Chen, G.-H. Tzeng, C.C. Chen, Y.-C. Hu. Discovery of Fuzzy Sequential Patterns for Fuzzy Par-

titions in Quantitative Attributes. *ACS/IEEE International Conference on Computer Systems and Applications*. Pág. 144-150, 2001.

[3] R.-S. Chen, Y.-C. Hu. A novel method for discovering fuzzy sequential patterns using the simple fuzzy partition method. *Journal of the American Society for Information Science* 54(7), Pág. 660-670, 2002.

[4] C. Fiot. Fuzzy Sequential Patterns for Quantitative Data Mining. In Galindo, J. (Ed.), *Handbook of Research on Fuzzy Information Processing in Databases*, In Press, 2008.

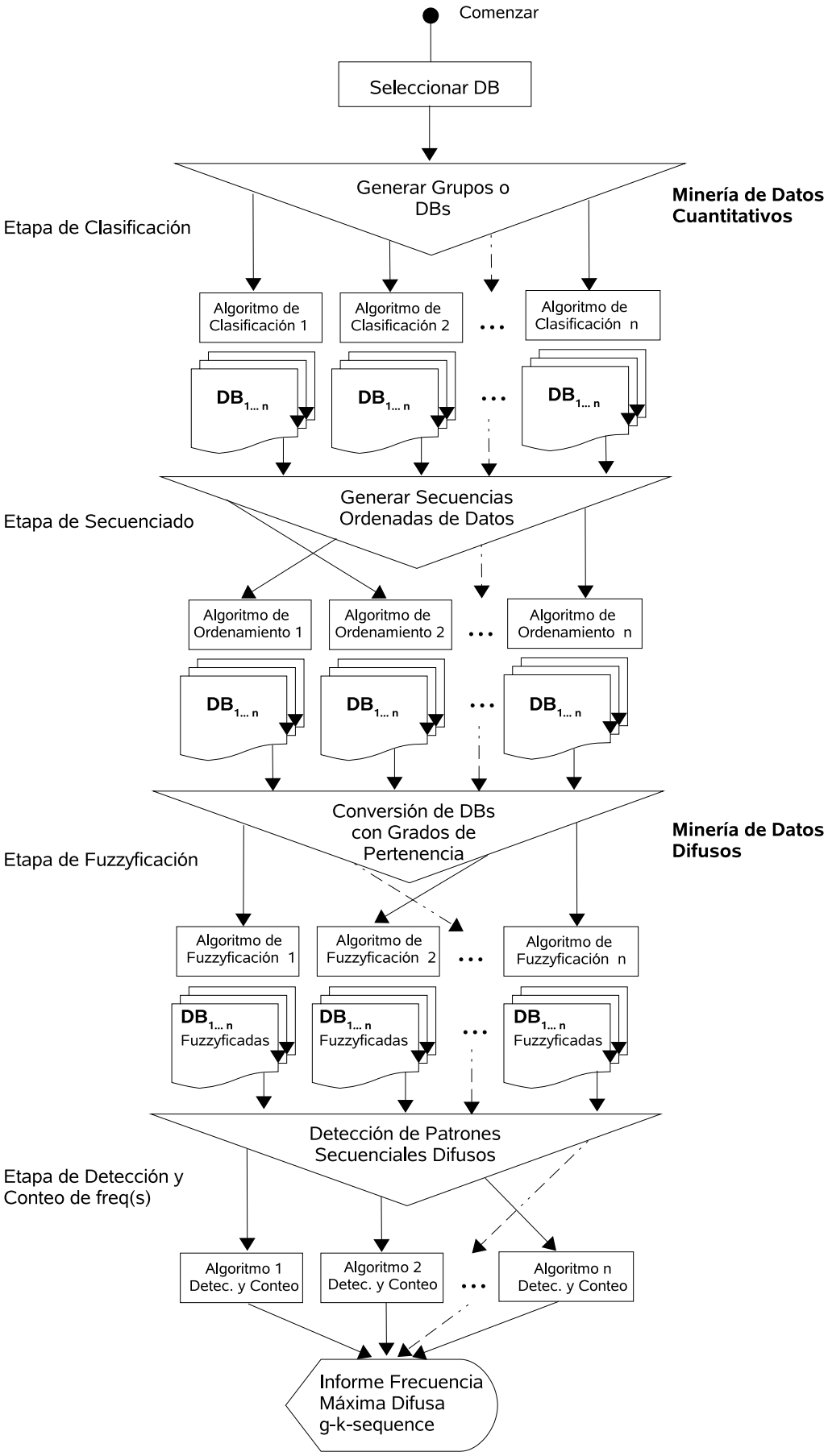


Figura 2: Esquema General para la Arquitectura de Herramientas MPSD

An approach to fuzzy context free languages

Claudio Moraga

European Centre for Soft Computing
33600 Mieres, Asturias, Spain
claudio.moraga@softcomputing.es

Abstract

The paper studies the relationship among fuzzy context free languages and fuzzy push down automata. The degree of membership of the words to the fuzzy language generated by a context free grammar, calculated in a monoid, introduces an additional burden for the accepting push down weighted automaton. Even in the case of languages accepted by fuzzy deterministic push down automata, the relationship with respect to fuzzy context free grammars is complex. A general effective algorithm to deduce one model from the other is still missing. The paper intends to draw attention to some of the open problems in this area.

1 INTRODUCTION

Fuzzy formal languages are fuzzy sets of words on a given alphabet. Therefore, words have a degree of membership to the corresponding language. Fuzzy formal languages are meant to be extensions on "crisp" formal languages and, consequently they preserve their main properties (like, for instance, some forms of closure), however needing special procedures to determine the degree of membership of the possibly resulting new words. There are basically two ways of generating fuzzy formal languages from crisp ones. Since a formal language may be specified by an alphabet and a generating system –(a grammar, in the case of a Chomsky language)- the alphabet (eventually the alphabets) may be replaced by (a) fuzzy set(s) of symbols or the generating system –(the rules of the grammar, in the case of a Chomsky language)- may be *fuzzified*. In the present paper, the second approach will be followed. Similarly, fuzzy automata may be defined considering a fuzzy set of input symbols, fuzzy sets of states, a fuzzy set of transitions or some combination thereof. In the present paper, fuzzy automata work with crisp sets of symbols and states, but with fuzzy transitions. Finite automata of this kind are discussed in the literature under the name of weighted automata [1], [2].

Fuzzy regular languages have been studied (see *e.g.* [3], [4], [6]) and the relationship between a fuzzy finite automaton and the equivalent right-linear and left-linear fuzzy grammars is known [5]. It is also known that given a fuzzy regular grammar, the fuzzy finite automaton accepting that language may be non-deterministic [5], [6]. This shows that different as in the crisp world, fuzzy non deterministic finite automata are more powerful than fuzzy deterministic automata. In the case of fuzzy context free languages, the correspondence between a fuzzy grammar and a fuzzy push down automata accepting the same fuzzy language is not known and will be the subject of study of the present paper. Structural properties of generalized fuzzy context free grammars, however have been studied in [7].

Definition 1:

A fuzzy context free grammar is a 5-tuple (V_T, V_N, S, P, ω), where (V_T, V_N, S, P) is a context free Chomsky grammar and ω: $P \rightarrow [0,1]$ is a weighting function. The set $\{(p, \omega(p))| p \in P\}$ constitutes a fuzzy set of productions.

Definition 2:

A fuzzy context free language is given by the triple $(G, \otimes, \oplus)$, where G is a fuzzy context free grammar –(definition 1), $\otimes$ is chosen such that $([0,1], \otimes)$ is a commutative monoid with identity *e*, and $\oplus$ is its dual based on the standard negation. ($\forall$ x,y $\in$ [0,1], $x \oplus y = 1 - ((1 - x) \otimes (1 - y))$). The operation $\otimes$ applied to the weights of all productions of the grammar needed to derive a word from the initial symbol gives the degree of membership of the word to the fuzzy language generated with the grammar.

It holds:

$$\mu_{L(G)}(w) = \omega(p_i) \otimes \omega(p_j) \otimes \ldots \otimes w(p_z) \qquad (1)$$

where $p_i, p_j, \ldots, p_z \in P$ are (eventually repeated) productions used to derive the word w. (If the grammar is ambiguous, then the possibly different degrees of membership obtained for a same word with different sequences of productions, will be connected with the dual operation $\oplus$ to determine the "best" one.)

Recall that if B is a fuzzy set and B_α denotes the α cut, then $B = \bigcup_{\alpha \in [0,1]} \alpha B_\alpha$. This suggests that a fuzzy language is a union of a set of properly weighted crisp languages. In the case of fuzzy formal languages is interesting to mention, that the component crisp languages do not necessarily belong to the same Chomsky class as the fuzzy language [2], [3], [9]. The following example is adapted from [7].

Let $G = (V_T, V_N, S, P, \omega, \otimes, \oplus)$, such that (V_T, V_N, S, P) is a context free Chomsky grammar, where V_T = {a, b,

c} and V_N = {S, A, B, C, D}. The productions and their respective weights are summarized as follows:

$p_1 : S \rightarrow AB$	$\omega(p_1) = 0.7$	$p_6 : B \rightarrow c$	$\omega(p_6) = 1$
$p_2 : S \rightarrow CD$	$\omega(p_2) = 0.5$	$p_7 : C \rightarrow aC$	$\omega(p_7) = 1$
$p_3 : A \rightarrow aAb$	$\omega(p_3) = 1$	$p_8 : C \rightarrow a$	$\omega(p_8) = 1$
$p_4 : A \rightarrow ab$	$\omega(p_4) = 1$	$p_9 : D \rightarrow bDc$	$\omega(p_9) = 1$
$p_5 : B \rightarrow cB$	$\omega(p_5) = 1$	$p_{10} : D \rightarrow bc$	$\omega(p_{10}) = 1$

Furthermore, $\forall\ x,y \in [0,1]$, let $x \otimes y = x \cdot y$ and $x \oplus y = x + y - x \cdot y$. (The neutral element of $\otimes$ is 0 and the neutral element of $\oplus$ is 1).

The following derivations are possible:

$S \Rightarrow AB \Rightarrow^* a^m b^m c^n$

with $m,n \geq 1$, $m \neq n$ and a global weight of 0.7

$S \Rightarrow CD \Rightarrow^* a^m b^n c^n$

with $m,n \geq 1$, $m \neq n$ and a global weight of 0.5

Additionally it may be shown that words of type $a^n b^n c^n$ with $n \geq 1$ may be obtained by starting both with $S \Rightarrow AB$ and $S \Rightarrow CD$ therefore the global weight is calculated as $0.7 \oplus 0.5 = 0.85$.

If α cuts are generalized to include "<" and "=" (besides ">") as possible relations to α, the following (crisp) languages may be obtained:

$$L(G)_{>0.7} = \{a^n b^n c^n \mid n \geq 1\}$$

$$L(G)_{=0.7} = \{a^m b^m c^n \mid m,n \geq 1,\ m \neq n\}$$

$$L(G)_{<0.7} = \{a^m b^n c^n \mid m,n \geq 1,\ m \neq n\}$$

Notice that even though the grammar G is context free, *all three above languages are context sensitive*. It should also be noticed that if $\otimes$ and $\oplus$ had been chosen as the minimum and maximum, respectively, then $0.7 \oplus 0.5 = 0.7$ and only two crisp α cut languages would be distinguishable.

2 FUZZY PUSH DOWN AUTOMATA

Definition 3:

A fuzzy push down automaton is a 9-tuple $(\Sigma, S, Q, \delta, q_0, s_0, Q_F, \omega, \otimes)$, where Σ denotes the input alphabet and S, the stack alphabet, with s_0 the symbol to identify the bottom of the stack. Q represents the set of states, q_0 the initial state and Q_F, the subset of Q corresponding to the final states. δ and ω will be specified in the next two equations and $\otimes$ is as in definition 2. (Its use will be explained below).

$$\delta \in (Q \times (\Sigma \cup \{\lambda\}) \times S) \times (Q \times S^*) \quad (2)$$

$$\omega: (Q \times (\Sigma \cup \{\lambda\}) \times S \times Q \times S^*) \rightarrow [0,1] \quad (3)$$

The push down automaton is deterministic if and only if

$$|\delta(q,a,s)| + |\delta(q,\lambda,s)| \leq 1 \quad \forall\ (q,a,s) \in Q \times \Sigma \times S \quad (4)$$

(See *e.g.* [10], [12]).

It is simple to see that the set of pairs of transitions obtained with eq. (2) and their corresponding weights calculated with eq. (3), constitutes a fuzzy set. Furthermore, as mentioned earlier, ([0,1], $\otimes$) is a commutative monoid.

Definition 4:

A sequence of transitions moving a deterministic push down automaton from the initial state into a final state (or, equivalently, into any state with an empty stack), will be called a trajectory. The global weight of a trajectory will be obtained by applying $\otimes$ to the weights of the transitions (obtained with eq. (3)) needed to build the trajectory. The sequence of input symbols from Σ that induce a trajectory in the automaton will be called a word and the global weight of the trajectory will be assigned to the word as its membership degree to the containing language.

It becomes apparent that the fuzzy push down automaton generates a fuzzy context free language. The generated language is context free, since it is generated by the "crisp" part of the fuzzy push down automaton. The fuzziness is introduced by associating to the words the weights of the corresponding trajectories.

A fuzzy deterministic push down automaton accepts a fuzzy deterministic context free language iff its crisp part accepts the words (independently of their membership degrees), and for every word accepted in this way, the global weight of the trajectory induced by that word matches the degree of membership of that word to the language.

3 FUZZY PUSH DOWN AUTOMATA AND FUZZY CONTEXT FREE GRAMMARS

In the case of crisp languages, it is known that for every context free grammar there exists a push down automaton that accepts the language generated by the grammar. Furthermore, there are algorithms to generate a push down automata from a given context free grammar (see *e.g.* [10], [11]). These algorithms however require that the grammar is expressed in a Greibach Normal Form and produce a non-deterministic push down automaton even when the language to be accepted is deterministic context free.

When dealing with fuzzy context free languages –(even in the deterministic case)- the situation becomes more complicated, since the number of productions used by the grammar and the number of transitions needed by a possibly accepting push down automaton play a definitive role. Consider, for instance, the grammar G = ({A, S}, {*a,b,c*}, S, {$S \rightarrow aSc$, $S \rightarrow aAc$, $A \rightarrow b$}), which generates the context free language $L(G) = \{a^n b c^n \mid \forall n > 0\}$. For this language it is simple to see by inspection that a push down automaton which first counts the *a*'s by loading them one after the other into the stack, followed by a change of state preserving the stack when reading the symbol *b* and then erasing the top *a* of the stack for each new *c* on the input tape until reaching the bottom of the stack, will accept the words of the language. (Since no word of this language is prefix of another word of the

language, accepting with an empty stack is allowed [12]). The language is deterministic context free (since the accepting automaton is deterministic, even though the generating grammar is not). For a given n > 0, however, the grammar needs applying n-1 times the first rule, once the second one and once the last one, meanwhile an accepting push down automaton needs 2n+1 non-λ transitions (to read and process all symbols from the input tape). If both the grammar and the push down automaton were fuzzy it may be seen that in this case the result of applying ⊗ to the weights of the transitions induced by reading an "*a*" and a "*c*", respectively, should equal the weight of the productions S → *a*S*c* and S → *a*A*c* meanwhile the weight of the production A → *b* should equal the weight of the transition induced by reading the "*b*". (These constraints may be easily satisfied, if at the grammar side, both productions replacing S are assigned the same weight. At the side of the push down automaton if *e.g.* the transition induced by reading an "*a*" is given the weight *e* –(the neutral element of the monoid being used)- then the transition induced by reading a "*c*" should have the same weight as the production S → *a*S*c*. (The same is obtained by giving to "*c*" the unity weight and to "*a*", that of the production.) Under *these* conditions then the fuzzy push down automaton would accept the language generated by the fuzzy context free grammar given above. This, of course, will change from language to language. Notice that equivalent automata with the same state diagram may exist since possible constraints to be satisfied by the automata may have different solutions (recall the two trivial solutions for the above example). A slightly more complex case is the following. Let G = ({*a*,*b*}, {A, S}, S, {S → *b*S*bb* | *b*A*bb*; A → *a*A | *a*}). It is simple to show that L(G) = $\{b^m a^n b^{2m} \mid n > 0, m > 0\}$. This language will be accepted by a deterministic push down automaton with the behaviour illustrated in figure 1.

In the crisp case, the strategy of this automaton is very simple: for each initial *b* that it reads, it loads *bb* into the stack; as soon as it reads one *a* it keeps the stack unchanged until it reads again a *b*. At this point, it will delete one *b* from the stack for each *b* that reads. If the word belongs to the language, the automaton will reach the bottom of the stack. If *m* = 1 only the production rule S → *b*A*bb* is needed once. If however *m* > 1 then the production rule S → *b*S*bb* will be applied *m*-1 times followed once by the production rule S → *b*A*bb*. Notice that if *n* = 1, to produce a single *a* of a word, two productions of the grammar have to be applied: S → *b*A*bb* and A → *a* meanwhile to produce a sequence of *a*'s, notice that the rule S → *b*A*bb*, which closes the prefix string of *b*'s already prepares the beginning of the string of *a*'s, where except for the last *a*, the rule A → *a*A is needed *n*-1 times. Finally the rule A → *a* closes the sequence. The automaton models this peculiarity by changing to the state q_2 when reading the first *a* and staying on state q_2 if additional *a*'s are read and continuing to state q_3 when the first *b* after the middle string of *a*'s is

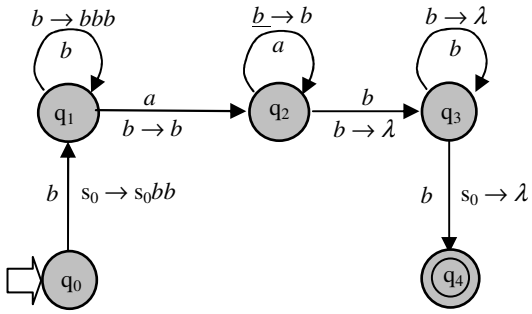


Fig. 1: Crisp push down automaton accepting the language $\{b^m a^n b^{2m} \mid n > 0, m > 0\}$. The edges show the input symbol that has been read and the rewriting action on the top symbol of the stack.

processed. In the fuzzy case, the automaton must however notice when the *last b* of the prefix is read (since ω(S → *b*A*bb*) does not necessarily have to be equal to ω(S → *b*S*bb*)). Similarly it has to notice when the *last a* of the middle string is processed (since ω(A → *a*) does not necessarily have to be equal to ω(A → *a*A)). A possible way to achieve this is to make the automaton notice when the *first* "non-b" following the prefix appears and similarly, when the *first* "non-*a*" after the middle string is read. A correction of the weight cumulated up to these points will be needed to compensate for the change in the production closing the prefix of *b*'s and the sequence of *a*'s. This can be done with the push down automaton shown in figure 1 ***only*** if the weights assigned to productions of the grammar are constrained to satisfy the conditions ω(S → *b*A*bb*) = k_1 ⊗ ω(S → *b*S*bb*) as well as ω(A → *a*) = k_2 ⊗ ω(A → *a*A) for a given operation ⊗ and some $k_1, k_2 \in [0,1]$. (See appendix for further discussion).

If the weights of the grammar rules are as shown in Table 1, where $r_2 = k_1 \otimes r_1$ and $r_4 = k_2 \otimes r_3$, then the fuzzy version of the automaton accepting the above language has to satisfy the constraints shown in table 2.

Table 1. Production weights		
p_1	S → *b*S*bb*	$\omega(p_1) = r_1$
p_2	S → *b*A*bb*	$\omega(p_2) = r_2 = k_1 \otimes r_1$
p_3	A → *a*A	$\omega(p_3) = r_3$
p_4	A → *a*	$\omega(p_4) = r_4 = k_2 \otimes r_3$

For instance consider the word w = *bbbaabbbbbb*. To produce this word, the grammar uses the production sequence p_1, p_1, p_2, p_3, p_4, from where

$$\begin{aligned}\mu_{L(G)}(w) &= r_1 \otimes r_1 \otimes r_2 \otimes r_3 \otimes r_4 \\ &= r_1 \otimes r_1 \otimes k_1 \otimes r_1 \otimes r_3 \otimes k_2 \otimes r_3 \\ &= k_1 \otimes k_2 \otimes r_1 \otimes r_1 \otimes r_1 \otimes r_3 \otimes r_3\end{aligned}$$

since ⊗ is associative and commutative.

A short look at Table 2 allows to see that the weight of the transition processing the last "*b*" of the prefix should have been $k_1 \otimes r_1$. However, since just by reading the first "*a*" the automaton realizes that the former "*b*" was the last

Table 2. Transition weights of the fuzzy deterministic push down automaton				
state / top of stack	input symbol/ next state	grammar rule(s)	transition | transition weight	
q_0 / s_0	b / q_1	S → bSbb	t_1	$\rho_1 = r_1$
q_1 / b	b / q_1	S → bSbb (S → bAbb)	t_2	$\rho_2 = r_1$ ($\rho_2 = k_1 \otimes r_1$)
q_1 / b	a / q_2	A → aA	t_3	$\rho_3 = k_1 \otimes r_3$
q_2 / b	a / q_2	A → aA (A → a)	t_4	$\rho_4 = r_3$ ($\rho_4 = k_2 \otimes r_3$)
q_2 / b	b / q_3	---	t_5	$\rho_5 = k_2$
q_3 / b	b / q_3	---	t_6	$\rho_6 = e$
q_3 / s_0	b / q_4	---	t_7	$\rho_7 = e$

of the prefix and should have been weighted with $k_1 \otimes r_1$, then the "missing" scaling factor k_1 is transferred to the following transition t_3 which corresponds to the transition for the first "a".

To accept this word, the automaton follows the transition sequence t_1, t_2, t_2, t_3, t_4, t_5, t_6, t_6, t_6, t_6, t_7, from where the global weight τ_w of the trajectory is given by

$$\omega(\tau_w) = \rho_1 \otimes \rho_2 \otimes \rho_2 \otimes \rho_3 \otimes \rho_4 \otimes \rho_5 \otimes \rho_6 \otimes \rho_6 \otimes \rho_6 \otimes \rho_6 \otimes \rho_7 = r_1 \otimes r_1 \otimes r_1 \otimes (k_1 \otimes r_3) \otimes r_3 \otimes k_2 \otimes e$$

Since $\otimes$ is associative, commutative and its identity is e, then $\mu_{L(G)}(w) = \omega(\tau_w)$.

Notice that both examples above needed a different treatment to satisfy the constraints leading to fuzzy deterministic automata such that the weights of the trajectories would match the degree of membership of the words as determined by the corresponding fuzzy grammars. No effective transformation algorithm is known to the author to produce a fuzzy deterministic push down automaton from a fuzzy context free grammar (when the 0-cut of the generated language is deterministic context free). Furthermore, if the weights assigned to the productions of the fuzzy context free grammar cannot satisfy the constraints mentioned earlier, a *non*-deterministic fuzzy push down automaton will be required. In the above example, for instance, if the weights given to the productions of the grammar were not constrained to satisfy the stated conditions, then a *non*-deterministic automaton would be required to "correctly guess" when a "b" is the last one of the prefix and an "a" is the last one of the middle sequence.

Fig. 2 shows the diagram of the non-deterministic fuzzy push down automaton to accept the language specified by the grammar of Table 1, if no constraints are imposed on the weights of the productions. Table 3 summarizes the transitions of the non-deterministic fuzzy push down automaton and their corresponding weights as related to the weights of the production rules of the grammar.

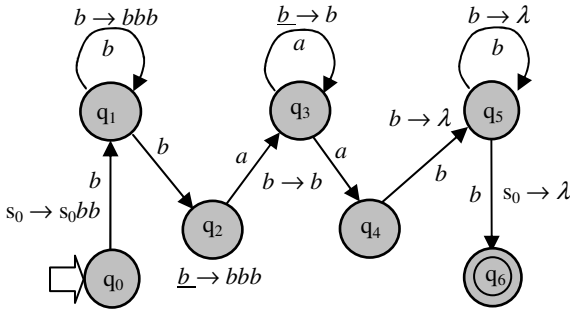


Fig. 2: Non-deterministic fuzzy push down automaton accepting the fuzzy context free language of table 1 if the weights of the productions are not constrained

Table 3. Transition weights of the fuzzy non-deterministic push down automaton						
actual state/ top of stack	input	next statee	associated grammar rules		transitions and weights	
q_0 / s_0	b	q_1	p_1	S → bSbb	t_1	$\rho_1 = r_1$
q_1 / b	b	q_1	p_1	S → bSbb	t_2	$\rho_2 = r_1$
q_1 / b	b	q_2	p_2	S → bAbb	t_3	$\rho_3 = r_2$
q_2 / b	a	q_3	p_3	A → aA	t_4	$\rho_4 = r_3$
q_3 / b	a	q_3	p_3	A → aA	t_5	$\rho_5 = r_3$
q_3 / b	a	q_4	p_4	A → a	t_6	$\rho_6 = r_4$
q_4 / b	b	q_5		(S → bAbb)	t_7	$\rho_7 = e$
q_5 / b	b	q_5		(S → bSbb)	t_8	$\rho_8 = e$
q_5 / s_0	b	q_6		(S → bSbb)	t_9	$\rho_9 = e$

If the same word w = *bbbaabbbbbb* as in the former case is considered, the accepting trajectory τ_w of the automaton follows the transition sequence t_1, t_2, t_3, t_4, t_6, t_7, t_8, t_8, t_8, t_8, t_9 leading to

$$\omega(\tau_w) = \rho_1 \otimes \rho_2 \otimes \rho_3 \otimes \rho_4 \otimes \rho_6 \otimes \rho_7 \otimes \rho_8 \otimes \rho_8 \otimes \rho_8 \otimes \rho_8 \otimes \rho_9 = r_1 \otimes r_1 \otimes r_2 \otimes r_3 \otimes r_4 \otimes e \otimes \ldots \otimes e$$

Since $\otimes$ is associative and commutative, then, as in the former case $\mu_{L(G)}(w) = \omega(\tau_w)$ and the non-deterministic fuzzy automaton accepts the word w.

Now, if the crisp language associated to a fuzzy context free language is non-deterministic context free, then the use of a fuzzy non-deterministic push down automaton to accept the language is unavoidable.

Summarizing it can be seen that if G is a fuzzy context free grammar and L(G) the language generated by the grammar, it is possible that for a given $\otimes$ no deterministic fuzzy push down automaton exists that accepts L(G), meanwhile $L(G)_{>0}$, the crisp context free language associated to L(G) by assigning membership degree 1 to

all words, may well be accepted by a crisp deterministic push down automaton.

Let $\mathcal{DCF}$ and $\mathcal{NDCF}$ denote the set of deterministic context free and non-deterministic crisp context free languages respectively. Similarly, let $\mathcal{FDCF}$ and $\mathcal{FNDCF}$ denote the deterministic and non-deterministic fuzzy context free languages, respectively. Then:

$$\mathcal{DCF} \subset \mathcal{NDCF} \text{ and } \mathcal{FDCF} \subset \mathcal{FNDCF}$$

However, as shown in the example above, depending on the distribution of the weights to the productions and the choice of the monoid, there are fuzzy context free languages $L(G) \in \mathcal{FNDCF} \setminus \mathcal{FDCF}$ such that $L(G)_{>0} \in \mathcal{DCF}$.

4 REMARKS ON INTERPRETATION

A modeling possible application of fuzzy formal languages was mentioned in [4], adding a semantic to words of the language by means of associating a meaning to the weights of the productions. If the weight of a production is interpreted as the *strength* of the production, then it seems reasonable to choose $\otimes$ as a t-norm[1] [13], based on the analogy that the strength of a chain is that of its weakest ring. Consequently, $\oplus$ should be taken as a t-conorm[2] in order to choose the strongest version of a word (analogous to the strongest chain). A quite opposite interpretation is possible, by considering the weight of a production to represent its *cost*. Under this point of view, it seems reasonable to choose $\otimes$ as a t-conorm to determine the "total cost" of a word, meanwhile $\oplus$ should be a t-norm to allow selecting the "most price-convenient" version of a word. The question of semantic interpretation of chains of weights coupled by an arbitrary uninorm is an open problem related to fuzzy languages.

REFERENCES

[1] Schützenberger M.P.: On the definition of a family of automata. *Information and Control* **4**, 245-270, 1961

[2] Droste M., Gastin P.: Weighted automata and weighted logics. *Lecture Notes in Computer Science* **3580**, 513-525, 2005

[3] Mordenson J.N., Fair P.S.: *Fuzzy Mathematics*. Springer, Heidelberg, 2001

[4] Moraga C.: Some properties of fuzzy languages. In: *Computational Intelligence. Theory and Applications*, (B. Reusch, Ed.), 367-374, Springer Verlag, 2006

[5] Moraga C.: Fuzzy formal languages. Research Report European Centre for Soft Computing. FSC-2006-01, 2006

[6] Li Y., Pedrycz W.: Fuzzy finite automata and fuzzy regular expressions with membership values in lattice - ordered monoids. *Fuzzy Sets and Systems* **156**, 68-92, 2005

[7] Asveld P.R.J.: Generalized fuzzy context free grammars. *Theoretical Computer Science* **347**, 167-190, 2005

[8] Yager, R., Rybalov, A.: Uninorm aggregation operators. *Fuzzy Sets and Systems* **80,** 111-120, 1996

[9] Mizumoto M., Toyoda J., Tanaka K.: Examples of formal grammars with weights. *Information Processing Letters* **2**, (4), 311–336, 1973

[10] Unknown author: www.cs.nott.ac.uk/~txa/g51mal.2001/notes/node34.html

[11] Brett A.C.: http://web.uvic.ca/~ling48x/ling484/notes/pda.html

[12] Autebert J.M., Berstel J., Boasson L.: Context-free languages and push-down automata. In: *Handbook of Formal Languages*, 1, 111-174, (Grzegorz Rozenberg and Arto Salomaa, Eds.) Springer, Heidelberg, 1997

[13] Klement E., Mesiar R., Pap E.: *Triangular Norms.* Kluwer, Dordrecht, 2000

[14] Silvert W.: Symmetric summation: A class of operations on fuzzy sets. *IEEE Trans. on Systems, Man, and Cybernetics*, **9**, 659-667, 1979

APPENDIX

On the satisfaction of constrains on the weights of productions.

Consider the structure ([0,1], $\otimes$), where $\otimes$ is chosen to be the uninorm disclosed in [8] and defined as follows for all x, y in [0,1]:

$$x \otimes y = \begin{cases} \max(x, y) \text{ if } \min(x,y) \geq e \\ \min(x, y) \quad \text{otherwise} \end{cases} \tag{A.1}$$

In the example of section 3, let:

$r_{12} = \min(\omega(S \rightarrow bAbb), \omega(S \rightarrow bSbb))$

$r_{34} = \min(\omega(A \rightarrow a), \omega(A \rightarrow aA))$

Without loss of generality let $r_{12} \geq e > r_{34}$. The left hand side of the inequality indicates that $\otimes$ behaves like maximum. Does the following equation have a solution in [0,1]?

$$\omega(S \rightarrow bAbb) = k_1 \otimes \omega(S \rightarrow bSbb) \tag{A.2}$$
$$= \max(k_1, \omega(S \rightarrow bSbb))$$

Eq. (A.2) has a solution if and only if

$\omega(S \rightarrow bAbb) \geq \omega(S \rightarrow bSbb)$,

in which case $k_1 = \omega(S \rightarrow bAbb)$ is the solution. If the inequality does not hold, then the constraint on the weights of the productions cannot be satisfied.

If $e > r_{34}$ then $\otimes$ behaves like minimum. Again the question is, whether the following equation has a solution in [0,1]?

$$\omega(A \rightarrow a) = k_2 \otimes \omega(A \rightarrow aA) \tag{A.3}$$
$$= \min(k_2, \omega(A \rightarrow aA))$$

Eq. (A.3) has a solution if and only if

$\omega(A \rightarrow a) \leq \omega(A \rightarrow aA)$,

in which case $k_2 = \omega(A \rightarrow a)$ is the solution. If the inequality does not hold, then the constraint on the weights of the productions cannot be satisfied.

[1] A t-norm is a uninorm such that $e = 1$

[2] A t-conorm is a uninorm such that $e = 0$

A uninorm is associative, commutative, monotone in both arguments and has a unity e

It becomes fairly obvious that both the distribution of the weights to the productions of a context free grammar and the selection of the uninorm to combine weights of productions into weights of words, are critical to support the realization of an accepting deterministic fuzzy push down automaton. Furthermore the former analysis is valid for all continuous t-norms and t-conorms, since the value of a t-norm is not larger than the minimum of its arguments and the value of a t-conorm is not less than the maximum of its arguments [13]. It is natural to ask whether an operation $\otimes$ exists such that the structure ([0,1], $\otimes$) is strong enough to satisfy the constraints *independently* of their distribution. It is easy to see that if ([0,1], $\otimes$) were an (Abelian) group, the existence of inverses would allow correcting a posteriori a misplaced weight to a production.

Recall the example *bbbaabbbbbb*. (See Table 1). The rule to produce the two first *b*'s has a weight r_1 and the rule to produce the *last b* has a weight r_2. A push down automaton will however realize that it has processed the last *b* when reading the first *a*. This transition therefore would receive the weight $\rho_{first_a} = (r_1)^{-1} \otimes r_2 \otimes r_a$ which corrects the wrong weight (r_1) assigned to the last *b* instead of r_2.

Is there an operation $\otimes$ that builds a group in the squared unit? The author does not know whether such an operation exists. However an "almost everywhere" solution is available.

Let Ξ denote ([0,1] $\times$ [0,1]) \ {$\langle 0,1\rangle$, $\langle 1,0\rangle$}, where $\langle\cdot,\cdot\rangle$ represents a pair of values (to avoid confusion with the open interval (0,1)).

Definition: σ: $\Xi \rightarrow$ [0,1] is a symmetric summation [14]. (See an illustration of the image of σ in figure 3).
For all $x, y \in \Xi$

$$\sigma(x, y) = \frac{xy}{xy + x'y'} \qquad (A.4)$$

where $x' = 1 - x$, represents the basic strong negation in fuzzy sets.

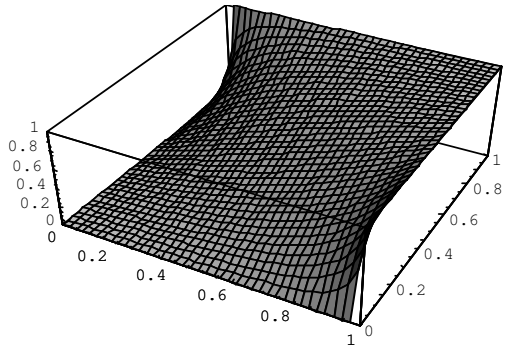


Fig. 3: Graphical representation of $\sigma(x,y)$

Lemma: ((0,1), σ) is an Abelian group.
Proof: If ((0,1), σ) is an Abelian group, its operation has to be associative and commutative; moreover it must have an identity *e* such that for all *y* in (0,1) $\sigma(e,y) = y$ and every element must have an inverse: $\sigma(x, x^{-1}) = e$.

σ is commutative, as can easily be seen in (A.4), and its associativity was shown in [14] .

With respect to the identity:

$\sigma(e,y) = ey / (ey + e'y') = y$

$\Rightarrow \quad e = ey + e'y' = ey + (1 - e)(1 - y)$

$\Rightarrow \quad e(1 - y) = (1 - e)(1 - y)$

$\Rightarrow \quad e = 1 - e$

$\Rightarrow \quad e = ½$

Every element of (0,1) must have an inverse such that $\sigma(x, x^{-1}) = e$. Then from (A.4)

$\sigma(x, x^{-1}) = e = ½$

$\Rightarrow \quad 2xx^{-1} = xx^{-1} + (1 - x)(1 - x^{-1})$
$= xx^{-1} + 1 - x - x^{-1} + xx^{-1}$
$= 2\,xx^{-1} + 1 - x - x^{-1}$

$\Rightarrow \quad x^{-1} = 1 - x = x'$

It follows that ((0,1), σ) is an Abelian group.

Notice that the inverse of an element of ((0,1), σ) corresponds to its fuzzy negation.

By extending to the limit the expression $\sigma(x, x^{-1}) = \sigma(x, x') = e = ½$ the following is obtained:

$\lim_{x\to 0} \sigma(x, x') = \lim_{x\to 0} \sigma(x', x) = \sigma(0,1) = \sigma(1,0)$

and since $\sigma(x, x') = ½$ then it could be assumed that $\sigma(0,1) = \sigma(1,0)$ would also take the value ½. By adding these expressions to the original definition of σ, the domain could be changed from Ξ to [0,1] $\times$ [0,1], but the resulting operation will not be associative:

$\sigma(\sigma(0,1), x)$ would give $\sigma(½, x) = x$, meanwhile $\sigma(0, \sigma(1,x))$ would give $\sigma(0,1) = ½$

The use of the correct definition of σ to combine the weights of productions to determine the membership degree of the generated word to the language, poses however a difficult different problem:

Consider the situation of a sequence of productions that lead to a word and assume that one of the productions has been given the weight 0 and another, the weight 1. Then, the membership degree of the resulting word could not be calculated with σ. It should be noticed, however, that a production with a weight of 0 does not belong to the fuzzy set of productions. Therefore, if a production is needed to generate certain words, then it should not be given the weight 0. On the other hand, a production with a weight 1, certainly belongs to the core of the fuzzy set of productions. The designer must however be aware of the fact that 1 is a dominant element for σ. For all *x* in (0,1], $\sigma(x,1) = 1$. This implies that if only *one* production in a sequence has the weight 1, the generated word would belong to the core of the language, independently of the weights of the other productions. Moreover the only way to secure that a word will belong to the core of the language would be to assign a weight of 1 to *any* of the productions. This is a serious draw back, which in spite of all attractive properties discussed above will probably impair the use of the symmetric summation to aggregate the weights of the productions of a fuzzy grammar to determine the degree of membership of the resulting words to the language.

If σ would however be used, then in the case of non deterministic fuzzy context free languages an additional peculiar problem has to be considered: σ is selfdual. In the case of nondeterministic context free languages, ambiguity may appear and the "dual of σ" will not be effective to elucidate the ambiguity. Depending on the interpretation of the weights, another $\oplus$ function will have to be chosen. For instance, if the weights are interpreted as the *strength* of a rule, then the maximum could be a reasonable choice to elucidate ambiguities. Similarly, minimum would be reasonable to elucidate ambiguities, if the interpretation given to the weight of a production is its *cost*.

ÍNDICES DE SOLAPAMIENTO ENTRE CONJUNTOS DIFUSOS. CONSTRUCCIÓN A PARTIR DE GRADOS DE SOLAPAMIENTO

Humberto Bustince[1] **Javier Fernández**[1] **Javier Montero**[2]

[1] Departamento de Automática y Computación, Universidad Pública de Navarra, Campus Arrosadia s/n, 31006, Pamplona, España, {bustince,fcojavier.fernandez}@unavarra.es

[2] Facultad de Matemáticas, Universidad Complutense, 28040 Madrid, España, monty@mat.ucm.es

Resumen

En este trabajo presentamos las propiedades básicas que, desde nuestro punto de vista, deben cumplir los grados de solapamiento. Estudiamos diferentes métodos de construcción y analizamos las condiciones bajo las cuales las t-normas continuas son grados de solapamiento. Por último, demostramos que si aplicamos operadores de agregación especiales a dichos grados, entonces obtenemos el índice de solapamiento (overlap) entre conjuntos difusos definido por Dubois y el índice de consistencia propuesto por Zadeh.

Palabras Clave: Grado de solapamiento, Indice de overlap, t-norma.

1 GRADO DE SOLAPAMIENTO. DEFINICION Y PROPIEDADES

Sabemos que la teoría de conjuntos difusos [32] ha sido muy útil en la resolución de problemas descritos mediante modelos imprecisos y con gran cantidad de ruido. Por este motivo, esta teoría es una herramienta apropiada para estudiar el problema de la identificación de los objetos que constituyen una imagen (véase [5, 20, 28, 30]).

Para separar el objeto del fondo de una imagen, lo primero que debemos hacer es representar el objeto mediante un conjunto difuso y el fondo mediante otro. El éxito del método de separación radica en la elección correcta de dichos conjuntos difusos, que ni mucho menos tienen que ser disjuntos en el sentido de Ruspini [29] (véase [1, 2, 19]).

Para construir estos conjuntos es necesario conocer la *propiedad* exacta que caracteriza los píxeles que

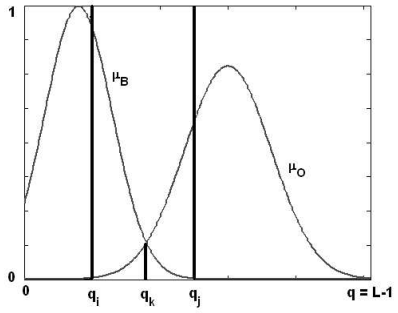


Figura 1: Solapamiento entre dos funciones

pertenecen al objeto (fondo). Esta propiedad determina la expresión de la función de pertenencia asociada al conjunto difuso que va a representar al objeto (fondo) (véase [10, 11, 17]). Normalmente, esta función no es conocida de manera exacta en todos y cada uno de los píxeles de la imagen. Hay píxeles para los cuales el experto está seguro de su pertenencia al fondo o al objeto y para otros duda y es precisamente en éstos, donde el valor de la función de pertenencia no es conocido con exactitud.

Dada una imagen, pedimos a un experto que asigne los siguientes valores a cada píxel de intensidad q:

$\mu_B(q)$: que representa el grado de pertenencia de dicho píxel al fondo.

$\mu_O(q)$: que representa el grado de pertenencia al objeto.

En la Fig.1 mostramos las dos funciones de pertenencia proporcionadas por el experto para representar una imagen dada en la escala de L niveles de gris. De la obsevación de dichas pertenecias deducimos que para intensidades menores que q_i el experto está seguro de que los píxeles pertenecen al fondo. Para intensidades mayores que q_j está seguro de su pertenencia al objeto. Sin embargo, para intensidades entre q_i y q_j el experto no tiene seguridad absoluta de su pertenencia a uno u

otro, siendo la intensidad q_k la que presenta máximo desconocimiento.

De este análisis deducimos que el grado de solapamiento entre las dos funciones lo podemos interpretar como la representación del desconocimiento que tiene el experto en determinar si el píxel en cuestión pertenece al fondo o pertenece al objeto. Por este motivo, definimos el grado de solapamiento entre $\mu_B(q)$ y $\mu_O(q)$ mediante funciones:

$$G_S : [0,1] \times [0,1] \to [0,1]$$

tales que deben cumplir las siguientes propiedades:

(G_S0) G_S depende únicamente de $\mu_B(q)$ y $\mu_O(q)$.

(G_S1) G_S es simétrica. El grado de solapamiento no depende del orden en que consideremos los grados de pertenencia.

(G_S2) El grado de solapamiento, entre los grados de pertenencia asociados a un píxel de intensidad q, es cero si y sólo si por lo menos uno de los grados de pertenencia vale cero. Es decir,

$G_S(\mu_B(q), \mu_O(q))) = 0$ si y sólo si
$\mu_B(q) = 0$ ó $\mu_O(q) = 0$, (i.e. $\min(\mu_B(q), \mu_O(q)) = 0$)

(G_S3) El grado de solapamiento, entre los grados de pertenencia asociados a un píxel de intensidad q, es uno si y sólo si ambos grados de pertenencia valen uno. Es decir,

$G_S(\mu_B(q), \mu_O(q))) = 1$ si y sólo si
$\mu_B(q) = 1$ y $\mu_O(q) = 1$, (i.e. $\mu_B(q) \cdot \mu_O(q) = 1$).

(G_S4) No decreciente; al aumentar los grados de pertenencia aumenta el grado de solapamiento.

(G_S5) Continuidad. El grado de solapamiento entre las dos pertenencias asociadas a un píxel de intensidad q no debe presentar una reacción caótica ante pequeños cambios del grado de pertenencia de dicho píxel al fondo o al objeto.

Definición 1. *Una función $G_S : [0,1]^2 \to [0,1]$ es un grado de solapamiento si satisface las propiedades:* $(G_S0) - (G_S5)$.

2 CONSTRUCCIÓN

Teorema 1. *$G_S : [0,1]^2 \to [0,1]$ es un grado de solapamiento si y sólo si*

$$G_S(x,y) = \frac{f(x,y)}{f(x,y) + h(x,y)}$$

para algún $f, h : [0,1]^2 \to [0,1]$ tal que

1) *f y h son simétricas;*

2) *f es no decreciente y h es no creciente;*

3) *$f(x,y) = 0$ si y sólo si $x \cdot y = 0$;*

4) *$h(x,y) = 0$ si y sólo si $x \cdot y = 1$;*

5) *f y h son continuas;*

Ejemplo 1.

1) Si $f(x,y) = \min(x,y)$ y $h(x,y) = \max(1-x, 1-y)$, entonces $G_S(x,y) = \min(x,y)$ es un grado de solapamiento.

2) Si tomamos $f(x,y) = \sqrt{x.y}$ y $h(x,y) = \max(1 - x, 1-y)$, entonces la construcción propuesta en el Teorema 1 es un grado de solapamiento.

3) Si $f(x,y) = \sqrt{x \cdot y}$ y $h(x,y) = 1 - x \cdot y$, la siguiente expresión es un grado de solapamiento

$$G_S(x,y) = \frac{\sqrt{x \cdot y}}{\sqrt{x \cdot y} + 1 - x \cdot y}$$

Corolario 1. *Bajo las condiciones del Teorema 1 si suponemos que G_S puede ser escrito de dos formas diferentes:*

$$G_S(x,y) = \frac{f_1(x,y)}{f_1(x,y) + h_1(x,y)} = \frac{f_2(x,y)}{f_2(x,y) + h_2(x,y)}$$

para todo $x, y \in [0,1]$ entonces, si definimos

$$F(x,y) = f_1(x,y) - f_2(x,y)$$
$$H(x,y) = h_1(x,y) - h_2(x,y)$$

se cumple:

$$G_S(x,y) = \frac{F(x,y)}{F(x,y) + H(x,y)}.$$

2.1 Caso específico: Uninormas

Sabemos que una uninorma es una función $U : [0,1]^2 \to [0,1]$ que es conmutativa, asociativa, creciente y tiene como elemento neutro $e \in]0,1[$. Por tanto, toda uninorma satisface las propiedades (G_S1) y (G_S4). Sin embargo, en el siguiente teorema demostramos que no satisfacen, por lo general, todas las condiciones demandadas a los grados de solapamiento.

Teorema 2. *No existe ninguna uninorma que sea grado de solapamiento.*

Dem. Sabemos (véase [26]) que no existe ninguna uninorma que sea continua en todo el intervalo unidad, por tanto U nunca satisface la propiedad (G_S5).

2.2 Caso especifico: t-normas

En esta subseccion estudiamos las condiciones bajo las cuales las t-normas cumplen las condiciones demandadas a los grados de solapamiento.

Sabemos que una t-norma es una función $T : [0,1]^2 \to [0,1]$ que es conmutativa, asociativa, creciente y $T(x,1) = x$ para todo $x \in [0,1]$. Por tanto, las t-normas cumplen (G_S1) y (G_S4). También cumplen (G_S3), la demostración es la siguiente: $1 = T(x,y) \leq min(x,y)$, entonces $x = y = 1$. El recíproco es directo teniendo en cuenta que el elemento neutro de T es el uno.

Debemos hacer notar que cuando trabajamos con t-normas, la condición necesaria de la propiedad (G_S2) coincide con la definición de t-norma *positiva*, (véase página 4 en [4]).

Teorema 3. *Sea T una t-norma en $[0,1]$. T satisface (G_S2) si y sólo si T no tiene divisores de cero.*

A continuación vamos a estudiar las condiciones bajo las cuales las t-normas que satisfacen la propiedad arquimediana (es decir, $T(x,x) < x$ para todo $x \in (0,1)$) también satisfacen (G_S2).

Empezamos recordando el teorema de representación de t-normas continuas y arquimedianas dado por Ling, Schweizer y Sklar en [25, 31], (más referencias sobre este teorema pueden encontrarse en [16]).

Teorema 4. *Una t-norma T es continua y arquimediana si y sólo si existe una función $f : [0,1] \to [0,\infty]$ con $f(1) = 0$, continua y estrictamente decreciente tal que*

$$T(x,y) = f^{(-1)}(f(x)+f(y)),$$

donde $f^{(-1)}$ es la pseudoinversa de f definida por

$$f^{(-1)}(x) = \begin{cases} f^{-1}(x) \text{ if } x \leq f(0) \\ 0 \text{ en otro caso} \end{cases}$$

La función f recibe el nombre de generador aditivo de T.

Lema 1. *Sea T una t-norma arquimediana que satisface (G_S5) y sea $f : [0,1] \to [0,\infty]$ un generador aditivo de T. Bajo estas condiciones: T satisface (G_S2) si y sólo si para todo $x, y \in]0,1]$, $f(x) + f(y) < f(0)$ se cumple.*

Dem. Basta tener en cuenta el Teorema 4.

Ejemplo 2. Sea $f(x) = -log(x)$, para todo $x, y \in]0,1]$ se cumple $f(x) + f(y) < f(0)$. Tomando este generador aditivo tenemos: $T(x,y) = xy$.

Teorema 5. *Sea T una t-norma que satisface (G_S5). T es estricta si y sólo si T satisface (G_S2) y es arquimediana.*

Teorema 6. *Si una t-norma T es un grado de solapamiento, entonces T es de uno de los tres tipos siguientes:*

1) $T = \min$;

2) *T es estricta;*

3) *T es la suma ordinal de la familia $\{([a_m, b_m], T_m)\}$ tal que si $T(x,y) = 0$ y $(x,y) \in [0,b_m]^2$, entonces el intervalo $[0,b_m]$ tiene asociada una t-norm T_m continua arquimediana y estricta.*

Dem. Sabemos que todo grado de solapamiento debe satisfacer (G_S5); es decir, debe ser continuo. Por hipótesis T es una t-norma que es grado de solapamiento, entonces es continua. Por la clasificación de t-normas continuas dadas en la página 11 de [16] (véase [27], [25] y [24]), sabemos que para una t-norma continua T tenemos tres posibilidades:

1.- $T = \min$;

2.- T es arquimediana;

3.- existe una familia $\{([a_m, b_m], T_m)\}$ tal que T es la suma ordinal de esta familia en el sentido presentado en [16].

Por hipótesis T es un grado de solapamiento, entonces satisface (G_S5) y (G_S2). Si además T es arquimediana, por el Teorema 5 tenemos que T es estricta.

Supongamos ahora que T es la suma ordinal de la familia $\{([a_m, b_m], T_m)\}$; es decir:

$$T(x,y) = \begin{cases} a_m + (b_m - a_m)T_m(\frac{x-a_m}{b_m-a_m}, \frac{y-a_m}{b_m-a_m}) \\ \qquad \text{if } (x,y) \in [a_m, b_m]^2 \\ \min(x,y) \text{ otherwise} \end{cases}$$

Sabemos que toda t-norm cumple la siguiente propiedad: si $x \cdot y = 0$, entonces $T(x,y) = 0$. Además por hipotesis la t-norm anterior tambien satisface el recíproco, por ser grado de solapamiento. Entonces si $T(x,y) = 0$ pueden ocurrir dos cosas:

a) (x,y) no pertenece a ningun $[a_m, b_m]^2$, en cuyo caso tenemos que $T(x,y) = min(x,y)$ para ese (x,y).

b) El punto (x,y) pertenece a $[a_m, b_m]^2$. Como $T(x,y) = 0 = a_m + (b_m - a_m)T_m(\frac{x-a_m}{b_m-a_m}, \frac{y-a_m}{b_m-a_m})$, entonces $a_m = 0$ y además $b_m \neq 0$ ya que en caso contrario el intervalo considerado sería $[0,0]$ y $x = y = 0$. Sabemos que T cumple (G_S2) y si $T(x,y) = 0$, entonces $T_m(\frac{x}{b_m}, \frac{y}{b_m}) = 0$, luego T_m tambien cumple (G_S2). Por tanto, tenemos que la t-norma continua y arquimediana T_m asociada al intervalo $[0,b_m]$ también cumple (G_S2), entonces por Teorema 5 es estricta.

Ejemplo 3. 1) En la construcción del siguiente grado de solapamiento utilizamos el item 3) del Teorema 6. En estas condiciones tomamos la t-norma producto, (la cual es stricta, continua y arquimediana), para el

correspondiente intervalo $[0, b_m]$.

$$G_S(x,y) = \begin{cases} 2xy \text{ if } (x,y) \in [0,0.5]^2 \\ \min(x,y) \text{ en otro caso} \end{cases}$$

2) En la construcción del siguiente grado de solapamiento tomamos las t-normas producto y la de Lukasiewicz (véase página 84 de [21]). Nótese que en este grado de solapamiento no consideramos ningun intervalo de la forma $[0, b_m]$.

$$G_S(x,y) = \begin{cases} 0.1 + 2.5(x-0.1)(y-0.1) \text{ if } (x,y) \in [0.1,0.5]^2 \\ 0.7 + \max(x+y-1.6, 0) \text{ if } (x,y) \in [0.7,0.9]^2 \\ \min(x,y) \text{ en otro caso} \end{cases}$$

3) La siguiente t-norma no satisface la propiedad $(G_S 2)$. Esto es debido a que en $[0, 0.25]^2$ tomamos la t-norma de Lukasiewicz la cual es continua y arquimediana pero no es stricta.

$$T(x,y) = \begin{cases} \max(x+y-0.25, 0) \text{ if } (x,y) \in [0,0.25]^2 \\ \min(x,y) \text{ en otro caso} \end{cases}$$

3 INDICE DE SOLAPAMIENTO. CONSTRUCCIÓN A PARTIR DE GRADOS DE SOLAPAMIENTO

En esta sección vamos a construir índices de solapamiento (overlap) agregando grados de solapamiento. Por esta razón, empezamos recordando los conceptos de overlap y de consistencia. A continuación justificamos los motivos por los cuales la mayor parte de los índices de overlap que existen en la literatura no satisfacen una de las cuatro condiciones exigidas por Dubois y Prade. Seguidamente mostramos las propiedades que debemos pedir a los operadores de agregación para que al aplicarlos a los grados de solapamiento, estudiados anteriormente, obtengamos índices de solapamiento. Terminamos mostrando la construcción de dichos índices y la forma de obtener las dos expresiones más utilizadas.

Denotamos con $F(U)$ al conjunto de todos los conjuntos difusos sobre el referencial finito no vacío U $(Cardinal(U) = n)$. A los conjuntos difusos sobre U los vamos a representar de la siguiente forma:

$$A = \{(u, \mu_A(u)) | u \in X\}$$

En 1978 Zadeh [33] presentó la extensión natural a la teoría de conjuntos difusos del concepto clásico de índice de overlap (solapamiento) al que llamó *consistencia*:

$$O(A,B) = Sup_{i=1}^{n}(\min(\mu_A(u_i), \mu_B(u_i))).$$

Obviamente, $O(A,B) = 0$ si A y B son completamente disjuntos y $O(A,B) = 1$ si existe un elemento $u_i \in U$ tal que $\mu_A(u_i) = \mu_B(u_i) = 1$.

En 1982 Dubois y Prade [15] presentaron la siguiente axiomatización para el índice de overlap:

Definición 2. *Un índice de overlap es una función $O(A,B)$ de $F(U) \times F(U)$ en el intervalo unidad tal que:*

$(O1)$ $O(A,B) = 0$ *si y sólo si A y B tienen soportes disjuntos;*

$(O2)$ $O(A,B) = 1$, *siempre que ($\mu_A(u_i) = 0$ ó $\mu_B(u_i) = 1$) ó ($\mu_A(u_i) = 1$ ó $\mu_B(u_i) = 0$);*

$(O3)$ $O(A,B) = O(B,A)$;

$(O4)$ *Si $B \leq C$, entonces $O(A,B) \leq O(A,C)$.*

La condición $(O2)$ de esta definición tiene la ventaja de que si A no es difuso, entonces $O(A,A) = 1$. Sin embargo, Dubois, Ostasiewicz y Prade en [14] establecen lo siguiente:

1. Para conjuntos difusos subnormales (es decir, $\mu_A(u_i) < 1$ y $\mu_B(u_i) < 1$ para todo $u_i \in U$), $(O2)$ *debe ignorarse.*

2. El índice *ROC* (véase [15]) no cumple esta propiedad.

También es interesante hacer notar lo siguiente:

Si $A = \{(u_i, \mu_A(u_i) = 0) | u_i \in U\}$, entonces $\min(A,A) = \{(u_i, \mu_{\min(A,A)}(u_i) = 0) | u_i \in U\}$ y por $(O1)$ resulta $O(A,A) = 0$. Si imponemos a la vez $(O2)$, entonces $O(A,A) = 1$. Es decir, tenemos una contradicción.

Estas consideraciones hacen que a los índices de overlap normalmente se les exija sólamente las condiciones: $(O1)$,$(O3)$ y$(O4)$ de la Definición 2.

En el siguiente teorema presentamos un método de construcción de índices de overlap. En este método vamos a agregar funciones de solapamiento. Para llevar a cabo estas construcciones a los operadores de agregación M les vamos a pedir las siguientes propiedades: (véase [9, 12, 23, 18]).

Un operador de agregación de dimensión n [9] es una función:

$$M : [0,1]^n \to [0,1]$$

que satisface las siguientes propiedades (ver también [12] y [13, 3]):

$A1$. $M(x_1, \cdots, x_n) = 0$ si y sólo si $x_i = 0$ para todo $i \in \{1, \cdots, n\}$.

$A2$. $M(x_1, \cdots, x_n) = 1$ si y sólo si $x_i = 1$ para todo $i \in \{1, \cdots, n\}$.

$A3$. Para cualquier par de n-tuplas $(x_1, \cdots, x_n)$ y $(y_1, \cdots, y_n)$ tales que $x_i, y_i \in [0,1]$ para todo $i \in \{1, \cdots, n\}$, si $x_i \leq y_i$ para todo $i \in \{1, \cdots, n\}$, entonces

$$M(x_1, \cdots, x_n) \leq M(y_1, \cdots, y_n);$$

es decir, M es monótona creciente en todos sus argumentos.

Teorema 7. *Sea $M : [0,1]^n \to [0,1]$ un operador de agregación que satisface $A1$, $A3$ y es idempotente ($M(x, \cdots, x) = x$ para todo $x \in [0,1]$)*

Sea $G_S : [0,1]^2 \to [0,1]$ y sea la función:
$O : F(U) \times F(U) \to [0,1]$ *dada por*
$O(A,B) = M_{i=1}^n(G_S(\mu_A(u_i), \mu_B(u_i)))$

En estas condiciones los siguientes ítems se cumplen:

i) O satisface $(O1)$ si y sólo si G_S satisface (G_S2);

ii) O satisface $(O3)$ si y sólo si G_S satisface (G_S1);

iii) Si G_S satisface (G_S4), entonces O satisface $(O4)$.

Ejemplo 4.

1. $O(A,B) = \bigvee_{i=1}^n G_S(\mu_A(u_i), \mu_B(u_i))$
2. $O(A,B) =$
$$\left(\frac{1}{n}\sum_{i=1}^n (G_S(\mu_A(u_i), \mu_B(u_i))^\beta\right)^{\frac{1}{\beta}}, \beta \neq 0$$

Notas al ejemplo

1. Si en la expresión 1. tomamos $G_S(x,y) = \min(x,y)$, entonces tenemos el índice de consistencia dado por Zadeh (véase [33]).
2. Si existe un único u_i tal que $G_S(\mu_A(u_i), \mu_B(u_i)) = 1$, entonces por la expresion 1. tenemos $O(A,B) = 1$. El hecho de que un único elemento u_i haga que $O(A,B) = 1$ nos sugiere utilizar la expresion 2. en lugar de la 1).
3. La expresión 2. cumple la siguiente propiedad:

   $O(A,B) = 1$ if and only if $G_S(\mu_A(u_i), \mu_B(u_i)) = 1$

   En estas condiciones, si $G_S(x,y) = \min(x,y)$ ó $G_S(x,y) = x \cdot y$ ó $G_S(x,y) = \sqrt{x \cdot y}$, entonces:

   $$O(A,B) = 1 \text{ si y sólo si}$$

   $$A = B = \{(u_i, \mu_A(u_i) = \mu_B(u_i) = 1) | u_i \in U\}$$

   por tanto la expresión 2., al igual que la mayor parte de expresiones dadas para el índice de overlap (véase [6, 7, 8, 14, 15]), no satisface la condicion $(O2)$ de la definición 2.

Corolario 2. *Bajo las condiciones del Teorema 7 si imponemos que la funcion $M : [0,1]^2 \to [0,1]$ satisfaga $A2$, entonces la siguiente propiedad se cumple: Si G_S satisface (G_S3), then*

$O(A,B) = 1$ si y sólo si $A = B = \{u_i, \mu_A(u_i) = 1) | u_i \in U\}$;

4 CONCLUSIONES

La relevancia de los problemas de solapamiento en el contexto de la clasificación difusa es evidente, tanto en cuanto no van a existir procedimientos sencillos, como existen en el caso no difuso, para reformular el problema sin solapamientos (véase por ejemplo [1, 2]). En este trabajo hemos propuesto una definición formal para las medidas de solapamiento, relacionándolas con algunas propuestas de índices con fines análogos. Además, hemos caracterizado su aspecto funcional bajo ciertas hipótesis.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por los proyectos de referencia: TIN2006-06190, TIN2007-65981.

Referencias

[1] A. Amo, D. Gómez and J. Montero, Spectral fuzzy classification: a supervised approach, *Mathware and Soft Computing*, **10**, Pág. 141-154, 2003.

[2] A. Amo, J. Montero, G. Biging and V. Cutello, Fuzzy classification systems, *European Journal of Operational Research*, **156**, Pág. 459-507, 2004.

[3] A. Amo, J. Montero, D. Gómez and E. Molina, Representation of consistent recursive rules, *European Journal of Operational Researchg*, **130**, Pág. 29-53, 2001.

[4] M. Baczynski and B. Jayaram, (S,N) - and R-implications: A state-of-the-art survey, *Fuzzy Sets and Systems*, 159, **14**, Pág. 1836-1859, 2007.

[5] J.C. Bezdek, J. Keller, R. Krisnapuram and N. R. Pal, Fuzzy Models and algorithms for pattern recognition and image processing, in *The Handbooks of Fuzzy Sets Series*, Series Editors: D. Dubois and H. Prade, Kluwer Academic Publishers, Boston/London/Dordrecht, 1999.

[6] H. Bustince, M. Pagola, E. Barrenechea, Construction of fuzzy indices from fuzzy DI-subsethood measures: Application to the global comparison of images, *Information Sciences*, 177, **3**, Pág. 906-929, 2007.

[7] H. Bustince, E. Barrenechea, M. Pagola, et al., Weak fuzzy S-subsethood measures. Overlap index, *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 14 **5**, Pág. 537-560, 2006.

[8] H. Bustince, V. Mohedano, E. Barrenechea, et al. Definition and construction of fuzzy DI-subsethood measures, *Information Sciences*, 176, **21**, Pág. 3190-3231, 2006.

[9] H. Bustince H., J. Montero, E. Barrenechea and M. Pagola, Semiautoduality in a restricted family of aggregation operators, *Fuzzy Sets and Systems*, 158, **12**, Pág. 1360-1377, 2007.

[10] H. Bustince, E. Barrenechea, M. Pagola, Restricted equivalence functions, *Fuzzy Sets and Systems*, 157 **17**, Pág. 2333-2346, 2006.

[11] H. Bustince, E. Barrenechea, M. Pagola, Image thresholding using restricted equivalence functions and maximizing the measures of similarity, *Fuzzy Sets and Systems*, 158, **5**, Pág. 496-516, 2007.

[12] T. Calvo, A. Kolesarova, M. Komornikova and R. Mesiar: Aggregation operators: properties, classes and construction methods. In T. Calvo, G. Mayor and R. Mesiar (Eds.): *Aggregation Operators New Trends and Applications* (Physica-Verlag, Heidelberg); Pág. 3-104, 2002.

[13] V. Cutello and J. Montero, Recursive connective rules, *International Journal of Intelligent Systems*, 14, Pág. 3-20, 1999.

[14] D. Dubois and J.L. Koning, Social choice axioms for fuzzy set aggregation. *Fuzzy Sets and Systems* 58, Pág. 339-342, 1991.

[15] D. Dubois, W. Ostasiewicz, H. Prade, Fuzzy Sets: History and Basic Notions, in: D. Dubois, H. Prade (Eds), *Fundamentals of Fuzzy Sets*, Kluwer, Boston, MA, Pág. 21-124, 2000.

[16] J. Fodor and M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support, in: *Theory and Decision Library* (Kluwer Academic Publishers) 1994.

[17] M.G. Forero, Fuzzy thresholding and histogram analysis, in *Fuzzy Filters for Image Processing*, Edited by M. Nachtegael, D. Van der Weken, D. Van de Ville and E.E. Kerre, Springer, Pág. 129-152, 2003.

[18] D. Gómez and J. Montero: A discussion on aggregation operators. *Kybernetika* 40, Pág. 107-120, 2004.

[19] D. Gómez, J. Montero and H. Bustince, On the relevance of some families of fuzzy sets, *Fuzzy Sets and Systems*, 158, Pág. 2429-2442, 2007.

[20] L.K. Huang and M.J. Wang, Image thresholding by minimizing the measure of fuzziness, *Pattern recognition*, 28, **1**, Pág. 41-51, 1995.

[21] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Trends in Logic. *Studia Logica Library*, Kluwer Academic Publishers, Dordrecht, 2000.

[22] E.P. Klement, R. Mesiar, E. Pap, On the relationship of associative compensatory operators to triangular norms and conorms, *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 4, Pág. 129-144, 1996.

[23] G.J. Klir and T.A. Folger: *Fuzzy Sets, Uncertainty and Information.* Prentice Hall, Englewood Cliffs, NJ, 1988.

[24] G. M. Krause, *A strengthened form of Ling's theorem on associative functions*, PhD Thesis, Illinois Institute of Technology, Chicago, 1981.

[25] C.H. Ling, Representation of associative functions, *Publ. Math. Debrecen*, 12, Pág. 189-212, 1965.

[26] R. Mesiar, V. Novák, Open problems from the 2nd International conference of fuzzy sets theory and its applications, *Fuzzy Sets and Systems*, 81, Pág. 185-190, 1996.

[27] P.S. Mostert, A.L. Shields, On the structure of semigroups on a compact manifold with boundary, *Ann. of Math.*, 65, Pág. 117-143, 1957.

[28] N. R. Pal and S. K. Pal, A review of image segmentation techniques, *Pattern recognition*, 26, Pág. 1277-1294, 1993.

[29] E.H. Ruspini, A new approach to clustering, *Information and Control*, 15, Pág. 22-32, 1969.

[30] B. Sankur and M. Sezgin, Survey over image thresholding techniques and quantitative performance evaluation, *J. Electron. Imaging*, 13, **1**, Pág. 146-165, 2004.

[31] B. Schweizer, A. Sklar, *Probabilistic Metric Spaces*, North-Holland, Amsterdam, 1983.

[32] L.A. Zadeh, Fuzzy sets, *Information Control*, 8, Pág. 338-353, 1965.

[33] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets Syst*, 1,Pág. 3–28, 1978.

CONTEXTOS CON CONJUNTOS DE OBJETOS Y ATRIBUTOS IGUALES

Cristina Alcalde[1] **Ana Burusco**[2] **Ramon Fuentes-Gonzalez**[2]

[1] Departamento de Matematica Aplicada. Universidad del Pais Vasco.
Escuela Universitaria Politecnica. Plaza de Europa 1, 20018 San Sebastian, Spain.
c.alcalde@ehu.es
[2] Departamento de Automatica y Computacion. Universidad Publica de Navarra
Campus de Arrosadia, 31006 Pamplona, Spain.
{burusco, rfuentes}@unavarra.es

Resumen

Los contextos L-fuzzy en los que el conjunto de objetos y de atributos coinciden pueden ser interesantes para representar algunas situaciones de la vida real. Por este motivo, realizaremos en este trabajo un estudio de como son los conceptos que se derivan a partir de relaciones de distintos tipos.

En la parte nal del trabajo mostraremos su utilidad cuando trabajemos con relaciones de tolerancia y de similitud.

Palabras Clave: Conceptos formales, conceptos L-Fuzzy, relaciones de tolerancia, relaciones de similitud.

1. INTRODUCTION

En trabajos anteriores ([2],[3]) de nimos los contextos L-Fuzzy (L, X, Y, R), con L ret culo completo, X e Y conjuntos de objetos y atributos respectivamente y $R \in L^{X\times Y}$ una relacion L-Fuzzy entre los objetos y atributos. Esta era una extension al caso Fuzzy de los contextos formales de R. Wille ([5]) cuando se quieren estudiar relaciones entre los objetos y los atributos con valores en un ret culo completo L, en lugar de valores binarios.

Partiendo de estos contextos formales, R. Wille ([5]) obten a informacion mediante los conceptos formales que son pares $(A, B), A \subseteq X, B \subseteq Y$ cumpliendo $A^* = B$ y $B^* = A$, siendo $*$ un operador derivacion que asocia a cada subconjunto de objetos A los atributos relacionados con los elementos de A, y a cada subconjunto de atributos B los objetos relacionados con los elementos de B.

En nuestro caso, para trabajar con estos contextos L-Fuzzy, utilizamos los operadores derivacion 1 y 2 de - nidos mediante las siguientes expresiones:

$\forall A \in L^X, B \in L^Y$

$$A_1(y) = \underset{x\in X}{inf}\{(A(x) \quad \rightarrow R(x, y))\}$$

$$B_2(x) = \underset{y\in Y}{inf}\{(B(y) \quad \rightarrow R(x, y))\}$$

con $\rightarrow$ una implicacion fuzzy de nida en $(L, \ )$, decreciente en el primer argumento y donde A_1 representa de forma difusa a los atributos relacionados con los objetos de A y B_2, a los objetos relacionados con los atributos de B.

La informacion almacenada en el contexto se visualiza mediante los conceptos L-Fuzzy que son pares $(M, M_1) \in (L^X, L^Y)$ con $M \in fix(\varphi)$, conjunto de puntos jos del operador φ, siendo este de nido a partir de los operadores derivacion 1 y 2 antes mencionados como $\varphi(M) = (M_1)_2 = M_{12}$. Estos pares, cuyas primera y segunda componentes son la extension e intension respectivamente, representan de forma vaga a un grupo de objetos que comparten unos atributos.

El conjunto $\mathcal{L} = \{(M, M_1)/M \in fix(\varphi)\}$ con la relacion de orden de nida como: $(M, M_1), (N, N_1) \in \mathcal{L}$, $(M, M_1) \quad (N, N_1)$ si $M \quad N$ (o equiv. $N_1 \quad M_1$) es un ret culo completo al que llamamos ([2],[3]) ret culo de conceptos L-Fuzzy. Un ejemplo de un ret culo de conceptos L-Fuzzy podemos encontrarlo en el apartado 2.1 de este trabajo.

Hasta el momento, los contextos (L, X, Y, R) con los que hemos estado trabajando ten an como objetos y atributos dos conjuntos, en general, distintos. A partir de los mismos se constru a el ret culo de conceptos que extra a informacion de la tabla en la que se pod an observar las relaciones existentes entre los objetos y los atributos.

No existen en la literatura trabajos en profundidad en los que se analicen contextos fuzzy cuyos conjuntos de

objetos X y atributos Y coincidan. Unicamente hemos encontrado en el caso n tido caracterizaciones del ret culo de conceptos para el caso de las relaciones y $\neq$([5]), y algunos estudios que relacionan similitudes y ret culos de conceptos L-Fuzzy([1]). En este trabajo abordaremos esta situacion ya que nos permitira obtener informacion sobre las relaciones entre los elementos que forman dicho conjunto.

En la primera parte, analizaremos distintos tipos de relaciones R para ver como influyen en los conceptos que se obtienen a partir de los contextos (L, X, X, R).

Aprovechando este tipo de conceptos, en la segunda parte del trabajo intentaremos establecer clases de tolerancia y de similitud en X utilizando el ret culo de conceptos.

2. ESTUDIO DE ALGUNOS TIPOS DE RELACIONES EN EL AMBITO DE CONTEXTOS

En lo que resta de trabajo utilizaremos la implicacion de Lukasiewicz $I(a,b) = \text{m n}(1, 1 \quad a + b)$ y $'$ para representar el complementario de un conjunto donde $A'(x) = 1 \quad A(x), \forall x \in X$.

En los siguientes apartados, analizaremos los contextos (L, X, X, R) segun algunas propiedades distinguidas de R.

2.1. R es reflexiva

a) Caso n tido:

Proposicion $\forall (A, B) \in \mathcal{L}$ concepto formal, con $A \neq \emptyset$ y $B \neq \emptyset$, se veri ca que $A \cap B \neq \emptyset$.

Dem Este hecho se comprueba con una simple observacion de la tabla, teniendo en cuenta que los conceptos son los rectangulos maximales en los que, en este caso, siempre intervienen los elementos de la diagonal principal ([5]).

Ejemplo: Tomemos la siguiente relacion

Cuadro 1: Relacion reflexiva

R	x_1	x_2	x_3
x_1	1	0	1
x_2	1	1	0
x_3	0	0	1

El ret culo de conceptos que se obtiene es:

$\{\emptyset, (x_1, x_2, x_3)\}$ $\{(x_1), (x_1, x_3)\}$ $\{(x_2), (x_1, x_2)\}$
$\{(x_3), (x_3)\}$ $\{(x_1, x_2), (x_1)\}$ $\{(x_1, x_3), (x_3)\}$
$\{(x_1, x_2, x_3), \emptyset\}$

Este tipo de relaciones nos permiten establecer modelos para aplicar en situaciones donde, por ejemplo, queremos establecer grupos de acuerdo al grado de simpat a entre personas (no siempre hay una simetr a).

b) Caso fuzzy: En este caso, como se prueba en los siguientes resultados, para cada elemento de X existe un concepto (A, B) donde aparece con grado de pertenencia 1 tanto en el conjunto de objetos A como en el de atributos B. Estos son precisamente los conceptos que se obtienen tomando los elementos de X como conjuntos de partida y calculando el concepto mas proximo a ellos mediante los resultados probados en [3].

Proposicion Si R es reflexiva y tomamos el singleton $A \in L^X$ tal que

$$A(x) = \begin{cases} 1 & \text{si } x = x_i \\ 0 & \text{en otro caso} \end{cases}$$

se demuestra que

$$A(x_i) = A_1(x_i) = A_{12}(x_i) = 1.$$

Dem Si calculamos el conjunto derivado de A

$$A_1(z) = \underset{x \in X}{inf}\{(A(x) \quad \rightarrow R(x,z))\} = R(x_i, z), \forall z \in X.$$

Por tanto, $A_1(x_i) = R(x_i, x_i) = 1$.

$$A_{12}(x) = R(x, x_i) \wedge \underset{z \neq x_i}{inf}\{(R(x_i, z) \quad \rightarrow R(x,z))\}.$$

Luego, $A_{12}(x_i) = R(x_i, x_i) = 1$.

Proposicion Si R es reflexiva, entonces $\forall x_i \in X$, $\exists (A, A_1)$ tal que $A(x_i) = A_1(x_i) = 1$.

Dem Es inmediata teniendo en cuenta que para cada $x_i \in X$ puedo tomar el conjunto $A \in L^X$ tal que $A(x) = 1$, si $x = x_i$, y $A(x) = 0$, en otro caso, y aplicar la propiedad anterior.

Por tanto, el concepto que se obtiene a partir de este conjunto de partida ([3]) cumplira que

$$A(x_i) = A_1(x_i) = 1.$$

Ejemplo: Tomemos la siguiente relacion

Cuadro 2: Relacion fuzzy reflexiva

R	x_1	x_2	x_3
x_1	1	0.2	1
x_2	1	1	0.3
x_3	0.2	0	1

El ret culo de conceptos L-Fuzzy es el siguiente:

Con1:$\{(x_1/0.2, x_2/0.3, x_3/0), (x_1/1, x_2/1, x_3/1)\}$

Con2:$\{(x_1/0.3, x_2/0.3, x_3/0.1), (x_1/1, x_2/0.9, x_3/1)\}$

...

Con11:$\{(x_1/1, x_2/0.3, x_3/1), (x_1/0.2, x_2/0, x_3/1)\}$

...

Con78:$\{(x_1/0.2, x_2/1, x_3/0), (x_1/1, x_2/1, x_3/0.3)\}$

...

Con229:$\{(x_1/1, x_2/0.5, x_3/0.2), (x_1/1, x_2/0.2, x_3/0.8)\}$

...

Con264:$\{(x_1/1, x_2/1, x_3/0.7), (x_1/0.5, x_2/0.2, x_3/0.3)\}$

donde Con229, Con78 y Con11, son los conceptos que se obtienen a partir de x_1, x_2 y x_3 respectivamente.

2.2. R es irreflexiva

a) Caso n tido:

Nunca hay interseccion entre los conjuntos de objetos y de atributos de un mismo concepto ya que, si (A, B) es un concepto y $x \in A \cap B$, entonces $R(x, x) = 1$, y no se cumplir a la irreflexividad.

b) Caso fuzzy:

Proposicion Si R es irreflexiva, entonces

$$\forall A \in L^X, A \geq (A')_2.$$

Dem Si R es irreflexiva, entonces $R \quad I'$. Por tanto, se veri ca que

$$A_1(z) = \inf_{x \in X}\{(A(x) \quad \rightarrow R(x, z))\}$$

$$\inf_{x \in X}\{(A(x) \quad \rightarrow I'(x, z))\} = A'(z).$$

Luego, $A_{12} \geq (A')_2$.

Como ademas en este caso $A = A_{12}$, podemos concluir que $A \geq (A')_2$.

La igualdad no siempre se cumple:

Ejemplo Si tomamos la relacion

Cuadro 3: Relacion fuzzy irreflexiva

R	x_1	x_2	x_3
x_1	0	0.5	0.3
x_2	0.9	0	0.4
x_3	1	0.2	0

y el conjunto $A = \{(x_1/0.3, x_2/0, x_3/1\}$ se tiene que $(A')_2 = \{(x_1/0.3, x_2/0, x_3/0{,}4\}$, luego $A > (A')_2$.

Caso particular: Si $R = I'$ (relacion ser distinto de), tambien aqu los conceptos son de la forma

$$(A, A')/A \in L^X.$$

Dem Si $R = I'$, entonces

$$A_1(z) = \inf_{x \in X}\{(A(x) \quad \rightarrow I'(x, z))\} = A'(z).$$

Luego,

$$A_{12}(x) = \inf_{z \in X}\{(A'(z) \quad \rightarrow R(x, z))\} = A''(x) = A(x).$$

En conclusion, el ret culo de conceptos L-Fuzzy es $\{(A, A')/A \in L^X\}$. Es decir, aparecen pares de conceptos con los conjuntos de objetos y atributos complementarios.

En los siguientes apartados estudiaremos conjuntamente los casos n tido y fuzzy.

2.3. R es simetrica

Proposicion Si R es simetrica, entonces (A, A_1) es un concepto, s y solo s (A_1, A) tambien lo es.

Dem Si R es simetrica, $R = R^{op}$, entonces

$$A_1(z) = \inf_{x \in X}\{(A(x) \quad \rightarrow R(x, z))\} =$$

$$= \inf_{z \in X}\{(A(x) \quad \rightarrow R(z, x))\} = A_2(z).$$

Es decir, $A_1 = A_2$ luego (A, A_1) es un concepto, s y solo s (A_1, A) tambien lo es.

2.4. R es una relacion de tolerancia (reflexiva y simetrica)

Si calculamos el ret culo de conceptos de un contexto donde la relacion es de tolerancia, observamos que en algunos conceptos (A, B) los conjuntos de objetos A y atributos B son iguales. Creemos que estos conceptos son los que establecen los v nculos mas fuertes entre grupos de elementos y pueden servir para resolver preguntas del tipo: elige un grupo de personas de una empresa para formar un equipo. En el apartado 3 de este trabajo veremos un ejemplo.

En general, aunque la relacion no sea de tolerancia, para resolver estas cuestiones sera interesante elegir los conceptos que tengan identicos conjuntos de objetos y de atributos (esto equivale a buscar un tipo especial de rectangulos maximales que son los cuadrados maximales en el caso de los conceptos formales).

2.5. R es una relacion de similitud (reflexiva, simetrica y transitiva)

Para ver como son los conceptos que surgen de ciertos conjuntos fuzzy especiales, necesitaremos del siguiente resultado.

Lema Si la relacion R es transitiva, se veri ca que

$$R(x,y) \quad R(z,x) \quad \rightarrow R(z,y), \forall x,y,z \in X.$$

Dem Como $R \circ R \quad R$, entonces $R \circ_T R \quad R$, siendo T la t-norma asociada a la implicacion con la que estamos trabajando, es decir,

$$R(z,x) T R(x,y) \quad R(z,y), \forall x,y,z \in X$$

lo que es equivalente a

$$R(x,y) \quad R(z,x) \quad \rightarrow R(z,y), \forall x,y,z \in X.$$

Proposicion Si R es de similitud y tomamos el singleton $A \in L^X$ tal que

$$A(x) = \begin{cases} 1 & \text{si } x = x_i \\ 0 & \text{en otro caso} \end{cases}$$

se demuestra que

$$A_1(x) = A_{12}(x) = R(x_i,x) = R(x,x_i), \forall x \in X.$$

Dem Calculando el conjunto derivado de A tenemos

$$A_1(z) = \underset{x \in X}{inf}\{(A(x) \quad \rightarrow R(x,z))\} = R(x_i,z), \forall z \in X.$$

Calculando de nuevo el derivado

$$A_{12}(x) = \underset{z \in X}{inf}\{(A_1(z) \quad \rightarrow R(x,z))\} =$$

$$= \underset{z \in X}{inf}\{(R(x_i,z) \quad \rightarrow R(x,z))\}$$

Teniendo en cuenta la simetr a de la relacion R y el lema anterior tenemos que

$$\underset{z \in X}{inf}\{(R(x_i,z) \quad \rightarrow R(x,z))\} = R(x_i,x)$$

Luego, $A_1(x) = A_{12}(x) = R(x,x_i) = R(x_i,x), \forall x \in X.$

3. TOLERANCIAS Y CLASIFICACIONES EN EL CONJUNTO DE OBJETOS

Numerosas situaciones de la vida real nos obligan a establecer normas que permitan clasi car o agrupar los elementos de un conjunto. Pensemos, por ejemplo, en establecer grupos de trabajo entre los empleados de una empresa.

En algunas ocasiones, son las propias relaciones entre los miembros del grupo las que los establecen: podemos preguntarle a la persona responsable de la seccion correspondiente como se coordinan entre s esos trabajadores. El resultado ser a una relacion reflexiva y simetrica. Sin embargo, si los propios miembros del grupo opinan entre s tendr amos, en la mayor parte de las situaciones, unicamente una relacion reflexiva.

En el primer caso, si tomamos una relacion R de tolerancia (reflexiva y simetrica) y construimos el ret culo de conceptos del contexto (L,X,X,R) quedandonos con aquellos conceptos que tienen el mismo conjunto de objetos y de atributos, estos nos pueden servir para establecer agrupamientos en X que representar an grupos de consenso.

Veamos un ejemplo:

Ejemplo: Tomemos la siguiente relacion.

Cuadro 4: Relacion fuzzy reflexiva y simetrica

R	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1	1	0	0.8	0.9	0
x_2	1	1	0.8	0.9	0.5	0
x_3	0	0.8	1	0	0	0.8
x_4	0.8	0.9	0	1	1	0
x_5	0.9	0.5	0	1	1	0
x_6	0	0	0.8	0	0	1

Trabajaremos con el contexto (L,X,X,R) donde $L = \{0, 0.1, \ldots 1\}$ y construiremos los conceptos L-Fuzzy que se derivan ([3]) de tomar como conjuntos de partida todos los posibles conjuntos n tidos, es decir, $A \in L^X$, con $A(x) = 0$ o 1. Estos conceptos son elementos destacados del ret culo de conceptos que resumen la informacion que del mismo se desprende.

Como nuestro interes es el de formar grupos de consenso entre los elementos de X, nos quedaremos con aquellos conceptos que tengan iguales los grados de pertenencia de los objetos y los atributos (ser an cuadrados maximales en la teor a de conceptos). En este caso son los siguientes (solamente aparece el conjunto de objetos, porque coincide con el de atributos):

$\{x_1/0, x_2/0, x_3/0.8, x_4/0, x_5/0, x_6/1\}$

$\{x_1/1, x_2/1, x_3/0, x_4/0.8, x_5/0.5, x_6/0\}$

$\{x_1/0.8, x_2/0.5, x_3/0, x_4/1, x_5/1, x_6/0\}$

que pueden representar, respectivamente, los grupos de consenso de la Figura 1.

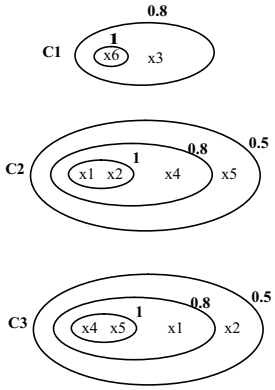


Figura 1: Grupos de consenso

Como puede comprobarse, estos grupos estan muy relacionados con las clases de tolerancia que se obtienen a partir de la misma relacion mediante los procesos habituales (Klir[4]).

Para el ejemplo que nos ocupa, la interpretacion de estas clases puede ser la de distintos agrupamientos que se pueden hacer en los elementos de X de cara a formar grupos de trabajo.

- Como se ve en C1, los elementos x_3 y x_6 pueden formar pareja con un consenso alto (0.8).

- En C2 tenemos que claramente (grado 1) x_1 y x_2 estan estrechamente ligados, algo menos con x_4 (grado 0.8) y en menor medida con x_5.

- Finalmente, en C3 vemos que tambien hay una estrecha relacion entre x_4 y x_5 a los que podr amos añadir x_1 y en menor medida x_2.

4. CLASES DE SIMILITUD

Si la relacion con la que trabajamos es de similitud (reflexiva, simetrica y transitiva) se deduce de forma sencilla, teniendo en cuenta la proposicion del apartado 2.5, que los conceptos que se obtienen de los conjuntos de partida $A(x) = 1$, si $x = x_i$, y $A(x) = 0$, en otro caso, tienen iguales los conjuntos de objetos y atributos. Ademas, estos conjuntos de objetos y atributos representan exactamente las clases de similitud de las que habla Klir([4]).

Ejemplo: Tomemos la relacion representada en el Cuadro 5.

En este caso, si construimos los conceptos que se derivan de los conjuntos de partida $x_1, x_2, x_3, x_4, x_5, x_6$ obtenemos los siguientes resultados:

$\{x_1/1, x_2/0.8, x_3/0, x_4/0.8, x_5/0.5, x_6/0\}$
$\{x_1/0.8, x_2/1, x_3/0, x_4/1, x_5/0.5, x_6/0\}$

Cuadro 5: Relacion fuzzy reflexiva, simetrica y transitiva

R	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1	0.8	0	0.8	0.5	0
x_2	0.8	1	0	1	0.5	0
x_3	0	0	1	0	0	0.8
x_4	0.8	1	0	1	0.5	0
x_5	0.5	0.5	0	0.5	1	0
x_6	0	0	0.8	0	0	1

$\{x_1/0, x_2/0, x_3/1, x_4/0, x_5/0, x_6/0.8\}$
$\{x_1/0.8, x_2/1, x_3/0, x_4/1, x_5/0.5, x_6/0\}$
$\{x_1/0.5, x_2/0.5, x_3/0, x_4/0.5, x_5/1, x_6/0\}$
$\{x_1/0, x_2/0, x_3/0.8, x_4/0, x_5/0, x_6/1\}$

que coinciden con las las de la matriz.

5. CONCLUSIONES Y TRABAJO FUTURO

Tras analizar contextos L-Fuzzy donde coinciden los conjuntos de objetos y atributos y donde la relacion R cumple diversas propiedades, hemos encontrado una conexion entre la Teor a de Conceptos y las clases de tolerancia y similitud asociadas a la relacion R.

Este es solo un primer analisis de estos tipos de contextos en funcion de las propiedades de R. Nuestro interes es el de profundizar en otras propiedades aqu no estudiadas (ordenes, distintos tipos de preferencias, etc.) y sus posibles aplicaciones a distintos campos como las ciencias sociales y la econom a.

Referencias

[1] R.Bělohlavek. Fuzzy relational Systems. IFSR International Series on Systems Science and Engineering 20 (2002).

[2] A.Burusco, R.Fuentes-Gonzalez. The Study of the L-Fuzzy Concept Lattice. *Mathware and Computing* 1 No.3 (1994), pp.209-218.

[3] A.Burusco, R.Fuentes-Gonzalez. Construction of the L-Fuzzy Concept Lattice. *Fuzzy Sets and Systems* 97 No.1 (1998), pp.109-114.

[4] G.J.Klir, B.Yuan. Fuzzy Sets and Fuzzy Logic, Prentice Hall,(1995).

[5] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. *Ordered Sets*, Reidel, Dordrecht-Boston, (1982), pp. 445-470.

THE CUTWORTHY APPROACH FOR FUZZY INTERVAL ORDERS

Susana Díaz[1] **Susana Montes**[2]

[1] Dept. of Statistics and O.R., University of Oviedo, Calvo Sotelo s/n, 33071 Oviedo, Spain
diazsusana@uniovi.es
[2] Dept. of Statistics and O.R., University of Oviedo, Edificio Polivalente de Viesques, 33203 Gijón, Spain
montes@uniovi.es

Abstract

Interval orders admit five equivalent ways to be expressed in the case of dealing with crisp sets. Each one of those five definitions can be translated to fuzzy sets. However, the counterparts for fuzzy relations of those five properties are not equivalent. In this contribution we consider each one of those five notions and we study the connection among them and their α-cuts. More precisely, we study for each property whether the fuzzy sets that satisfy that condition can be characterized by the corresponding crisp property of their α-cuts. The results obtained prove that only the Ferrers property propagates among a fuzzy relation and its α-cuts. In the other cases the implication does not hold in any one of the directions.

Keywords: interval order, Ferrers property, biorder, α-cut.

1 Introduction

Preference modeling theory deals with the judgments of decision makers over a set of alternatives. Those judgments are formalized by preference structures. A preference structure is a triplet of binary relations. Those relations cover the three possible situations when comparing two alternatives: preference, indifference and incomparability. This way every preference structure shows the opinion of a decision maker over a set of alternatives. The decision maker is asked to compare all the alternatives by pairs. His/her answer over each pair is formalized by one of the three relations that make up a preference structure. In classical or crisp preference modeling, from every preference structure can be obtained a relation, called large preference relation. It completely characterizes the preference structure from which it is built.

The original binary relations considered in preference modeling are crisp. They can only take the values 0 and 1 that express absence and presence of relation among the alternatives. In real life, specially when dealing with human decisions, comparison of alternatives involve nuances that crisp relations cannot reflect. Fuzzy sets and in particular fuzzy relations try to solve that problem. They can take any value in the interval $[0, 1]$. This ample range of values allows to formalize the intensity of the answers of the decision maker. Fuzzy preference relations model in a better way real life decisions. So they are a topic of interest since several years ago (see, for example, [3, 4, 5, 7, 9, 11, 12]).

Regardless of the context (crisp or fuzzy sets theory), preferences must be coherent. Several properties are considered to provide coherence to preference structures. Interval orders are one of the most important types of "coherent" preference structures. They do not classify any pair of alternatives as incomparable. Moreover, they model coherent decisions, since the strict preference relation must be transitive and they are consistent with real world, since its indifference relation may be intransitive. The formal definition of crisp interval order will be provided in the following section. In addition to the original notion, there are four well-known characterizations for the concept. They involve the Ferrers property of both the large and strict preference relations as well as the notion of biorder. Those five concepts are not equivalent when handling fuzzy relations (see [8]). This opens the problem of which one expresses in a better way the original idea of interval order. Since many properties of fuzzy structures are investigated using the transferability to cut sets, in this contribution we compare the five possible definitions by studying the propagation of the properties of their α-cuts.

The contribution is structured in five sections. After the introduction we recall basic notions about crisp and fuzzy preference structures in Sections 2 and 3, respectively. In Section 4 we study the α-cuts of the five (equivalent in the crisp case) definitions of interval order. Finally, in Section 5 we briefly discuss the results obtained.

2 Crisp preference structures

Suppose a decision maker that compares two alternatives in a set of alternatives A. He/she always takes a decision among the following three options: (i) he/she clearly prefers one to the other; (ii) the two alternatives are indifferent to him/her; (iii) he/she is unable to compare the two alternatives. Accordingly, three (binary) relations are defined on A: the *strict preference relation* P, the *indifference relation* I and the *incomparability relation* J.

Given the set of alternatives A, We will denote A^2 the set of ordered pairs of alternatives, $A^2 = \{(a,b) \mid a,b \in A\}$, where the pairs (a,b) and (b,a) are different elements of A^2. Every binary relation defined on A can be seen as a subset of A^2. The binary relation Q connects two alternatives a and b if and only if the pair (a,b) is in the subset Q. Using this notions of subsets, for a relation Q on A, its converse or transpose is defined as $Q^t = \{(b,a) \mid (a,b) \in Q\}$, its complement as $Q^c = \{(a,b) \mid (a,b) \notin Q\}$ and its dual as $Q^d = (Q^t)^c$.

Definition 2.1 [17] *A preference structure on A is a triplet (P,I,J) of relations on A that satisfy:*

(i) P and J are irreflexive and I is reflexive;

(ii) P is asymmetrical, I and J are symmetrical;

(iii) $P \cap I = \emptyset$, $P \cap J = \emptyset$ and $I \cap J = \emptyset$;

(iv) $P \cup P^t \cup I \cup J = A^2$.

In the crisp sets context the quadruplet (P, P^t, I, J) establishes a particular partition of A^2. Since the relation P^t can be obtained from P, the triplet (P,I,J) contains all the information about the answers of the decision maker.

Every preference structure can be identified with a unique reflexive relation called *large preference relation*. It is denoted by R and defined as $R = P \cup I$. This relation contains all the information of the preference structure since we can reobtain the triplet from R:

$$(P,I,J) = (R \cap R^d, R \cap R^t, R^c \cap R^d). \qquad (1)$$

Observe that the relation R is the union of the strict preference and the indifference. The usual interpretation of this relation is the following one:

$$aRb \text{ if and only if } a \text{ is at least as good as } b.$$

We say that a relation Q is *complete* if for any two alternatives in the set A, it holds at least aQb or bQa. In many cases the large preference relation R associated to a preference structure is forced to be complete. Intuitively, this means that the decision maker is able to compare any two alternatives. In [16] it was proven that the large preference relation is complete if and only if the associated preference structure does not connect any pair of alternatives (a,b) by the incomparability relation, that is, if $J = \emptyset$.

$$R \text{ complete} \quad \Leftrightarrow \quad J = \emptyset.$$

Those preference structures whose incomparability relation $J = \emptyset$ are usually denoted as (P,I).

A relation Q satisfies the *Ferrers property* (see among others [16, 17]) if

$$(aQb \quad \wedge \quad cQd) \quad \Rightarrow \quad (aQd \quad \vee \quad cQb),$$

for all a,b,c,d in A.

A relation Q is a *biorder* if

$$(aQb \quad \wedge \quad cQb \quad \wedge \quad cQd) \quad \Rightarrow \quad aQd.$$

Let us denote the composition of two binary relations Q_1 and Q_2 by $Q_1 \circ Q_2$. That is, $a(Q_1 \circ Q_2)b$ if and only if there exists c such that $aQ_1c \wedge cQ_2b$. Then the equivalent compositional definition of biorder is $Q \circ Q^d \circ Q \subseteq Q$.

Definition 2.2 *An* interval order *is a preference structure (P,I,J) such that $J = \emptyset$ and*

$$P \circ I \circ P \subseteq P.$$

In this context of crisp relations it is well-known ([16, 17]) that for any binary relation Q,

$$Q \text{ is Ferrers} \quad \Leftrightarrow \quad Q \text{ is a biorder.}$$

The equivalence holds in particular for R and P. It is easy to prove that any reflexive relation satisfying the Ferrers property is complete. As a consequence, for any preference structure, if R is Ferrers, then R is complete. We have already commented that the relation R is complete if and only if the associated preference relation does not admit incomparability relation

(P, I). In this setting it is also well-known (see again [16, 17]) that

$$P \text{ is Ferrers} \quad \Leftrightarrow \quad R \text{ is Ferrers.}$$

More precisely, for interval orders, Monjardet [16] showed the equivalence among the following notions:

Proposition 2.1 [16] *Let (P, I) be an AFPS without incomparability and R its large preference relation. The following conditions are equivalent.*

a) (P, I) is an interval order.

b) P is Ferrers.

c) P is a biorder.

d) R is Ferrers.

e) R is a biorder.

Therefore, for crisp relations we have five different (equivalent) ways of defining an interval order.

3 Fuzzy preference structures

In this section we define for fuzzy relations the notions recalled in the previous section for crisp relations.

3.1 Triangular norms

As it is well-known, triangular norms are the classical operators employed for defining the intersection of fuzzy relations. We recall that the binary operator $T : [0,1] \times [0,1] \to [0,1]$ is a *triangular norm* or *t-norm* for short if it is commutative, associative, monotone and has 1 as neutral element. The most important t-norms are the *Minimum*: $T_{\mathbf{M}}(x, y) = \min(x, y)$, the *Product* $T_{\mathbf{P}}(x, y) = xy$, *Lukasiewicz*: $T_{\mathbf{L}}(x, y) = \max(x + y - 1, 0)$ and the *Drastic product*:

$$T_{\mathbf{D}}(x, y) = \begin{cases} \min(x, y) & \text{, if } \max(x, y) = 1, \\ 0 & \text{, otherwise.} \end{cases}$$

The drastic product is the smallest t-norm and the minimum t-norm is the greatest one. That is, for any t-norm T it holds that $T_{\mathbf{D}} \leq T \leq T_{\mathbf{M}}$.

An important family of t-norms are the rotation invariant t-norms [13]. A *rotation invariant t-norm* is a t-norm T that verifies

$$T(x, y) > z \quad \Leftrightarrow \quad T(x, 1 - z) > 1 - y,$$

for any $x, y, z \in [0, 1]$.
These operators satisfy in particular that $T(x, y) > 0 \Leftrightarrow x + y > 1$.

The largest rotation invariant t-norm is the *nilpotent minimum* $T_{\mathbf{nM}}$ defined as

$$T_{\mathbf{nM}}(x, y) = \begin{cases} \min(x, y) & \text{, if } x + y > 1, \\ 0 & \text{, otherwise.} \end{cases}$$

Conversely, the usual operators used for defining the union of fuzzy relations are the t-conorms. A *triangular conorm* or *t-conorm* is a binary relation $S : [0,1] \times [0,1] \to [0,1]$ commutative, associative, monotone and that has 0 as neutral element.

It is easy to prove that t-norms and t-conorms are closely related. For any t-norm T it holds that the relation S defined as follows

$$S(x, y) = 1 - T(1 - x, 1 - y), \forall (x, y) \in [0, 1]^2,$$

is a t-conorm. It is called *dual t-conorm* of T. Conversely, for each t-conorm S the relation T defined as $T(x, y) = 1 - S(1 - x, 1 - y)$ is a t-norm.

A wide study of t-norms and t-conorms can be found in [14].

3.2 Fuzzy binary relations

Crisp relations lack of flexibility. They only take values in {0,1}. This becomes a drawback for modeling human behaviour and in general, real life. Fuzzy binary relations solve the problem of flexibility. They take a value in the interval $[0, 1]$ that expresses the intensity of the connection by the relation considered between two elements [19].

A fuzzy binary relation is defined as a function Q from the cartesian product $A \times A$ into the interval $[0, 1]$. For every pair $(a, b) \in A \times A$, the value $Q(a, b)$ shows the degree of truth of the fact that aQb in the crisp sense. The complementary, transpose (or converse) and dual of the binary relation Q are defined for any $(a, b) \in A \times A$ as follows:

$$\begin{aligned} Q^c(a, b) &= 1 - Q(a, b), \\ Q^t(a, b) &= Q(b, a), \\ Q^d(a, b) &= 1 - Q(b, a). \end{aligned}$$

For every fuzzy binary relation Q and every $\alpha \in [0, 1]$, Q_α will represent the α-cut of Q, i.e., the crisp binary relation defined by $Q_\alpha = \{(a, b) \in A | Q(a, b) \geq \alpha\}$.

Thus, for any $(a, b) \in A$,

$$Q(a, b) = \sup\{\alpha \in [0, 1] | (a, b) \in Q_\alpha\}$$

and therefore, any fuzzy binary relation is completely characterized by their α-cuts. Then, we could study the properties of a fuzzy binary relation by means of

the properties of its α-cuts, which are crisp binary relations. In general, this is known as "cutworthy" approach (for the notion of cutworthiness see e.g., the book by Klir and Yuan [15]).

In the fuzzy sets context the notion of completeness admits different generalizations. The two most important notions are the following ones. Given a relation Q we say that it is

- *Strongly complete:* if $Q(a,b) = 1$ or $Q(b,a) = 1, \forall a, b \in A$.
- *weakly complete:* if $Q(a,b) + Q(b,a) \geq 1, \forall a, b \in A$.

Let us consider a t-norm T, a t-conorm S and two fuzzy binary relations Q_1 and Q_2 defined on the same set of alternatives A. The T-intersection of Q_1 and Q_2 is denoted $Q_1 \cup_T Q_2$, the S-union is denoted $Q_1 \cap_S Q_2$. They are defined for any two alternatives a, b in A as follows:

$$Q_1 \cup_T Q_2(a,b) = T(Q_1(a,b), Q_2(a,b)),$$
$$Q_1 \cap_S Q_2(a,b) = S(Q_1(a,b), Q_2(a,b)).$$

Let us consider a t-norm T and two fuzzy binary relations Q_1 and Q_2 defined on the set A. The T-composition of Q_1 and Q_2 is denoted by $Q_1 \circ_T Q_2$. It is defined by

$$Q_1 \circ_T Q_2(a,b) = sup\{T(Q_1(a,c), Q_2(c,b)) \mid c \in A\}.$$

3.3 Fuzzy preference structures

The notion of preference structure in the setting of fuzzy relations was a topic of debate for several years (see [3]) until Van de Walle et al. [18] introduced the notion of additive fuzzy preference structure.

Definition 3.1 *An* additive fuzzy preference structure *on the set of alternatives A is a triplet of fuzzy binary relations (P, I, J) satisfying*

- *P and J are irreflexive and I is reflexive.*
- *P is $T_{\mathbf{L}}$-asymmetric ($P \cap_{T_{\mathbf{L}}} P^t = \emptyset$), I and J are symmetric.*
- *$P \cap_{T_{\mathbf{L}}} I = P \cap_{T_{\mathbf{L}}} J = I \cap_{T_{\mathbf{L}}} J = \emptyset$.*
- *$(P \cup_{S_{\mathbf{L}}} I)^c = P^t \cup_{S_{\mathbf{L}}} J$.*

Every preference structure has associated a *fuzzy large preference relation* R obtained as the Łukasiewicz union of P and I, $R = P + I$. The value $R(a,b)$ expresses the degree of truth of the assertion "a is at least as good as b". Given that relation R, we can build many different AFPS by means of generators [2].

Definition 3.2 *A* generator i *is a symmetric function* $i : [0,1]^2 \to [0,1]$, *such that*

$$T_L \leq i \leq T_M.$$

Observe that although it is bounded by two t-norms, it is not necessarily a t-norm. Given a reflexive relation R and a generator i, we can obtain an additive fuzzy preference structure as follows:

$$P = R - i(R, R^t),$$
$$I = i(R, R^t),$$
$$J = I - (R + R^t - 1).$$

Additive fuzzy preference structures verify the following important property:

$$P(a,b) + P(b,a) + I(a,b) + J(a,b) = 1.$$

For crisp relations (any of the definitions of) completeness of the large preference relation R is equivalent to the condition $J = \emptyset$. In this context the following equivalence holds:

Proposition 3.1 [6] *Let R be a reflexive fuzzy relation R and J the incomparability relation of the preference structure obtained from R by a generator i. It holds that*

$$J = \emptyset \Leftrightarrow \begin{cases} R \text{ is weakly complete} \\ i|_S = T_L \end{cases}$$

where $S = \{(x,y) \in [0,1]^2 : \exists (a,b) \in A^2$ *with* $R(a,b) = x, R(b,a) = y\}$.

In this case the strict preference and the indifference relations are obtained from R as follows:

$$I = R + R^t - 1; \quad P = R^d.$$

3.4 Fuzzy interval orders

We introduce some appropriate generalizations of the concepts of Ferrers, biorder and interval order for fuzzy relations and additive fuzzy preference structures. We will stay within the usual setting of t-norms.

Definition 3.3 [12] *Given a t-norm T and a t-conorm S, a fuzzy binary relation Q on A is* (T,S)-Ferrers *if the following inequality holds:*

$$T(Q(a,b), Q(c,d)) \leq S(Q(a,d), Q(c,b))$$

for any $a, b, c, d \in A$.

When S is the dual t-conorm of T, the (T,S)-Ferrers property is just called *T-Ferrers*. In this work, we will always assume that S is the dual t-conorm of T. Thus, we will just talk about T-Ferrers relations.

Definition 3.4 [11] *Given a fuzzy binary relation Q on A, Q is a* T-biorder *if it holds that*

$$Q \circ_T Q^d \circ_T Q \subseteq Q\,,$$

or, equivalently,

$$T(T(Q(a,b), 1 - Q(c,b)), Q(c,d)) \leq Q(a,d)$$

for any $a, b, c, d \in A$.

Definition 3.5 [4] *An additive fuzzy preference structure (P, I, J) on a set of alternatives A is called a* T-interval order *if it holds that*

$$J = \emptyset \quad \text{and} \quad P \circ_T I \circ_T P \subseteq P\,,$$

or, equivalently,

$$J(a,b) = 0 \text{ and } T(T(P(a,b), I(b,c)), P(c,d)) \leq P(a,d)$$

for any $a, b, c, d \in A$.

Clearly, if $T_1 \leq T_2$, every T_2-Ferrers relation is also T_1-Ferrers, every T_2-biorder is also a T_1-biorder and every T_2-interval order is also a T_1-interval order.

Let us consider an additive fuzzy preference structure (P, I, J) and its associated large preference relation R. Given a t-norm T, we can provide a "fuzzy version" of each one of the five properties involved in Proposition 2.1:

a) (P, I, J) is a T-interval order,

b) P is T-Ferrers,

c) P is a T-biorder,

d) R is T-Ferrers,

e) R is a T-biorder.

In general, we cannot establish an equivalence between any two of the five points above for any t-norm T (see [8]). Many authors have characterized fuzzy structures by means of their α-cuts. Thus, in the following section we study the behaviour of the five properties with respect to their α-cuts.

4 Interval orders and α-cuts

In [10] the following result was proven.

Proposition 4.1 *Let Q be a binary fuzzy relation on a finite set of alternatives A. Then*

$$Q\ T_{\mathbf{M}}\text{-Ferrers} \Leftrightarrow Q_\alpha \text{ Ferrers } \forall \alpha \in [0,1]$$

As a consequence we get the following result.

Corollary 4.2 *Let R be a binary fuzzy relation defined on a finite set of alternatives A and let P be the asymmetric relation associated to R: $P = R^d$. Then*

$$R\ T_{\mathbf{M}}\text{-Ferrers} \Leftrightarrow R_\alpha \text{ Ferrers } \forall \alpha \in [0,1]$$
$$P\ T_{\mathbf{M}}\text{-Ferrers} \Leftrightarrow P_\alpha \text{ Ferrers } \forall \alpha \in [0,1]$$

Since a crisp relation is Ferrers if and only if it is a biorder, it also follows from Proposition 4.1 the following result.

Corollary 4.3 *Let R be a binary reflexive fuzzy relation defined on a finite set of alternatives A and let P be the asymmetric relation associated to R: $P = R^d$. Then*

$$R\ T_{\mathbf{M}}\text{-Ferrers} \Leftrightarrow R_\alpha \text{ biorder } \forall \alpha \in [0,1]$$
$$P\ T_{\mathbf{M}}\text{-Ferrers} \Leftrightarrow P_\alpha \text{ biorder } \forall \alpha \in [0,1]$$

This result cannot be translated to fuzzy biorders.

Proposition 4.4
- *Not every reflexive fuzzy $T_{\mathbf{M}}$-biorder R satisfies that all its α-cuts are biorders.*
- *The converse implication neither holds.*

Let us recall that in [6] we proved the following result.

Proposition 4.5
- *$T_{\mathbf{nM}}$ is the largest t-norm T such that any $T_{\mathbf{M}}$-Ferrers relation is a T-biorder.*
- *$T_{\mathbf{nM}}$ is the largest t-norm T such that any $T_{\mathbf{M}}$-biorder is a T-Ferrers relation.*

It follows from here that neither every fuzzy $T_{\mathbf{M}}$-Ferrers relation is a $T_{\mathbf{M}}$-biorder, nor every $T_{\mathbf{M}}$-biorder is a $T_{\mathbf{M}}$-Ferrers fuzzy relation. Thus,

Corollary 4.6 *Let Q be a fuzzy reflexive relation defined on a finite set of alternatives A. Then*

$$Q\ T_{\mathbf{M}}\text{-biorder} \begin{array}{c} \not\Leftarrow \\ \not\Rightarrow \end{array} Q_\alpha \text{ biorder } \forall \alpha \in [0,1]$$

However, this does not imply that for the particular case of reflexive or asymmetric relations like R or P, some (or both) of the implications hold. Bufardi [1] proved what follows:

$$P \quad T_{\mathbf{M}}\text{-Ferrers} \not\Rightarrow P \quad T_{\mathbf{M}}\text{-biorder.}$$

As a consequence,

$$P \text{ is } T_{\mathbf{M}}\text{-biorder} \not\Leftarrow P_\alpha \text{ is a biorder } \forall \alpha \in [0,1]$$

The converse implications neither holds. This follows from the following result, where we assure that Proposition 4.5 cannot be improved, even in the case the fuzzy binary relation is asymmetric.

Proposition 4.7 *Let P be an asymmetric fuzzy relation. It holds that*

$$P\ T_{\mathbf{M}}\text{-biorder} \Rightarrow P\ T_{\mathbf{nM}}\text{-Ferrers}$$

and this is the strongest implication possible.

Therefore, not every asymmetric relation P that is a $T_{\mathbf{M}}$-biorder is $T_{\mathbf{M}}$-Ferrers and from here,

Corollary 4.8 *Let P be an asymmetric fuzzy relation. Then*

$$P\ T_{\mathbf{M}}\text{-biorder} \not\Rightarrow P_\alpha \text{ biorder } \forall \alpha \in [0,1]$$

However, the implication does hold for some particular α-cuts.

Theorem 4.9 *Let Q be a fuzzy binary relation. It holds that*

$$Q\ T_{\mathbf{M}}\text{-biorder} \Rightarrow Q_\alpha \text{ biorder } \forall \alpha \in [0, 0.5]$$

In particular, the implication holds for any asymmetric relation P. It is clear that the converse implication does not hold. Recall that it does not hold even when considering every α-cut with $\alpha \in [0,1]$.

From Theorem 4.9 another question arises. The t-norm $T_{\mathbf{nM}}$ is the largest one for which the equivalence between T-biorder and T-Ferrers holds. Then, it makes sense to study if a $T_{\mathbf{nM}}$-biorder can be characterized by those α-cuts such that $\alpha \leq 0.5$. However, the answer is negative. No one of the implications hold.

Proposition 4.10 *Let P be an asymmetric fuzzy relation.*

(i) If for all $\alpha \leq 0.5$ P_α is a biorder, then not necessarily P is a $T_{\mathbf{nM}}$-biorder.

(ii) If P is a $T_{\mathbf{nM}}$-biorder, then not necessarily P_α is a biorder. And this holds for any $\alpha \leq 0.5$.

Let us recall that the implication (ii) means that there exists no α such that we can assure that P_α is a biorder when P is a $T_{\mathbf{nM}}$-biorder.

Next we focus on reflexive relations. We attain to similar conclusions as for the strict preference relations.

For reflexive relations it holds the following.

Proposition 4.11 *Let R be a fuzzy reflexive relation.*

- *It holds that*

$$R\ T_{\mathbf{M}}\text{-biorder} \Rightarrow R\ T_{\mathbf{nM}}\text{-Ferrers}$$

and this is the strongest possible implication.

- *It holds that*

$$R\ T_{\mathbf{M}}\text{-Ferrers} \Rightarrow R\ T_{\mathbf{nM}}\text{-biorder}$$

and this is the strongest possible implication.

It follows from here that:

$$R\ T_{\mathbf{M}}\text{-biorder} \not\Leftarrow R_\alpha \text{ biorder } \forall \alpha \in [0,1]$$

The converse implication neither holds imposing reflexivity.

$$R\ T_{\mathbf{M}}\text{-biorder} \not\Rightarrow R_\alpha \text{ biorder } \forall \alpha \in [0,1]$$

However, the implication does hold for some particular α's as we proved in Theorem 4.9. The result holds particularly for reflexive relations R.

As we discussed for P, the converse implication does not hold since it does not hold for all α in $[0,1]$.

As for strict preference relations, we next study whether a reflexive $T_{\mathbf{nM}}$-biorder can be characterized by its α-cuts for $\alpha \leq 0.5$. The answer once more time is no.

Proposition 4.12 *Let R be a fuzzy reflexive relation.*

- *If for all $\alpha \leq 0.5$ R_α is a biorder, then not necessarily R is a $T_{\mathbf{nM}}$-biorder.*
- *If R is a $T_{\mathbf{nM}}$-biorder, then not necessarily R_α is a biorder. And this holds for every $\alpha \leq 0.5$.*

However, when we consider fuzzy interval orders we obtain better results. Thus,

Proposition 4.13 *Let (P, I, J) be an additive fuzzy preference structure. (P, I, J) is a $T_{\mathbf{M}}$-interval order if and only if*

$$J_\alpha = \emptyset \quad \text{and} \quad P_\alpha \circ I_\alpha \circ P_\alpha \subseteq P_\alpha, \forall \alpha \in [0,1].$$

The triplet $(P, I, \emptyset)$ is a $T_{\mathbf{M}}$-interval order if and only if P is a $T_{\mathbf{M}}$-biorder. Therefore, we cannot characterize a $T_{\mathbf{M}}$-interval order by the Ferrers property of the α-cuts of R and P as we discussed in this section.

After Proposition 4.12, the reader could think that this result contradicts Theorem 1 given in [6]. We would have a contradiction if $(P_\alpha, I_\alpha, \emptyset)$ were the (crisp) preference structure associated to the crisp reflexive relation R_α. However this does not hold. The (crisp) preference structure associated to the α-cut of a (fuzzy) reflexive relation is not the triplet of the α-cuts of the relations in the preference structure associated to the fuzzy reflexive relation. This is, given a fuzzy reflexive relation, R and the associated preference structure (P, I, J), the (crisp) preference structure obtained from R_α is not the triplet $(P_\alpha, I_\alpha, J_\alpha)$. This does not hold even although the incomparability relation J is empty.

5 Conclusion

This work deals with the definition of fuzzy interval order. The definition given for a crisp interval order admits five equivalent ways to be expressed as we recall in Proposition 2.1. In the fuzzy sets context those equivalences do not hold in general. Therefore, we have a priori five different options to define a fuzzy interval order. Since the five options are not equivalent, it is necessary to determine which is the best one to formalize what a fuzzy interval order is. We understand that the more classical properties a definition verifies, the better it is. Some properties of fuzzy relations are characterized by the corresponding crisp property of the α-cuts of the relation. In this contribution we have considered the five definitions that appear at the end of Subsection 3.4. We have studied the connection of each of them with the corresponding crisp property of its α-cuts. The results obtained lead us to consider the Ferrers property as a definition better than the other ones. Only the Ferrers property of a fuzzy relation can be characterized by the Ferrers property of their α-cuts.

Acknowledgements

The research in this paper is partly supported by the Spanish Ministry of Education and Science grant MTM2007-61193. This grant partially participates of ERDF.

References

[1] A. Bufardi. An alternative definition for fuzzy interval orders. *Fuzzy Sets and Systems*, 133:249–259, 2003.

[2] B. De Baets and J. Fodor. Generator triplets of additive fuzzy preference structures. In *Proc. RelMiCS6*.

[3] B. De Baets and J. Fodor. Twenty years of fuzzy preference structures (1978-1997). *JORBEL*, 37:61–82, 1997.

[4] B. De Baets and B. Van de Walle. Weak and strong fuzzy interval orders. *Fuzzy Sets and Systems*, 79:213–225, 1996.

[5] B. De Baets, B. Van de Walle, and E. Kerre. Fuzzy preference structures without incomparability. *Fuzzy Sets and Systems*, 76:333–348, 1995.

[6] S. Díaz, B. De Baets, and S. Montes. T-ferrers relations versus T-biorders. *Lecture Notes in Artificial Intelligence*, 2715:269–276, 2003.

[7] S. Díaz, B. De Baets, and S. Montes. Additive decomposition of fuzzy pre-orders. *Fuzzy Sets and Systems*, 158:830–842, 2007.

[8] S. Díaz, S. Montes, and B. De Baets. A comparative study of T-interval orders, T-Ferrers and T-biorders. In *Proc. IPMU 2008*.

[9] S. Díaz, S. Montes, and B. De Baets. Transitivity bounds in additive fuzzy preference structures. *IEEE Transactions on fuzzy systems*, 15(2):275–286, 2007.

[10] J. P. Doignon, B. Monjardet, M. Roubens, and Ph. Vincke. Biorder families, valued relations, and preference modelling. *J. Math. Psych.*, 30:435–480, 1986.

[11] J. Fodor. Traces of binary fuzzy relations. *Fuzzy Sets and Systems*, 50:331–341, 1992.

[12] J. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, 1994.

[13] S. Jenei. New family of triangular norms via contrapositive symmetrization of residuated implications. *Fuzzy Sets and Systems*, 110:157–174, 2000.

[14] E. P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[15] G. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic*. Prentice-Hall, New Jersey, 1995.

[16] B. Monjardet. Axiomatiques et propriétés des quasi-ordres. *Math. Sci. Hum.*, 63:51–82, 1978.

[17] M. Roubens and Ph. Vincke. Preference modelling. In *Lecture Notes in Economics and Mathematical Systems*, volume 250. Springer, Berlin, 1985.

[18] B. Van De Walle, B. De Baets, and E. Kerre. Characterizable fuzzy preference structures. *Annals of Operational Research*, 80:105–136, 1998.

[19] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

SOBRE MEDIDAS DE $\mathcal{T}$-INCOMPATIBILIDAD PARA CONJUNTOS BORROSOS DE ATANASSOV

Elena E. Castiñeira[1] **Susana Cubillo**[1] **Wilmer Montilla**[2]

[1] Dpto. Matemática Aplicada, Facultad de Informática, Universidad Politécnica de Madrid (UPM)
28660, Boadilla del Monte, Madrid. {ecastineira, scubillo}@fi.upm.es
[2] Área de Matemática, Universidad Nacional Abierta, 5201 Venezuela, wilmer-montilla@hotmail.com

Resumen

Continuando con el estudio sobre la incompatibilidad entre conjuntos borrosos de Atanassov iniciado en [5], en este artículo se presentan distintas formas de medir la incompatibilidad, bien sea considerando la familia de t-normas intuicionistas t-representables mediante t-normas y t-conormas conjugadas, respectivamente, con la t-norma y t-conorma de Łukasiewicz, o bien tomando otra familia de t-normas intuicionistas no representables. Por último, se muestran varios resultados en los que se evidencia la relación entre las distintas medidas de incompatibilidad propuestas.

Palabras Clave: Conjuntos borrosos intuicionistas de Atanassov, t-normas intuicionistas, t-normas intuicionistas representables, medidas de incompatibilidad.

1 PRELIMINARES

Un conjunto borroso intuicionista de Atanassov (AIFS) A sobre un universo E [2, 3] es un conjunto borroso en el sentido de Goguen (véase [10]) cuya función de $\mathbb{L}$-pertenencia es $\chi^A : E \to \mathbb{L}$ dada por $\chi^A(x) = (\mu_A(x), \nu_A(x))$, con $(\mathbb{L}, \leq_{\mathbb{L}})$ tal que $\mathbb{L} = \{(a_1, a_2) \in [0,1]^2 \mid a_1 + a_2 \leq 1\}$ y el orden $\leq_{\mathbb{L}}$ definido como:

$$\overline{a} = (a_1 a_2) \leq_{\mathbb{L}} (b_1, b_2) = \overline{b} \Leftrightarrow \begin{cases} a_1 \leq b_1, \\ a_2 \geq b_2, \end{cases}$$

donde $0_{\mathbb{L}} = (0,1)$ y $1_{\mathbb{L}} = (1,0)$ son, respectivamente, los elementos mínimo y máximo de $(\mathbb{L}, \leq_{\mathbb{L}})$. El conjunto de todos los AIFSs es isomorfo al de sus funciones de $\mathbb{L}$-pertenencia, que se representa por $\mathbb{L}^E$.

Para extender la noción de incompatibilidad a los AIFSs es necesario, al igual que en el caso borroso, contar con instrumentos que modelen la intersección. Así, en [6, 8] se define el concepto de t-norma intuicionista de la siguiente manera.

Definición 1.1 *Una t-norma intuicionista es una operación binaria $\mathcal{T} : \mathbb{L}^2 \to \mathbb{L}$ que es conmutativa, asociativa, monótona y con $1_{\mathbb{L}}$ como elemento neutro, es decir, $\mathcal{T}(\overline{a}, 1_{\mathbb{L}}) = \overline{a}$ para todo $\overline{a} \in L$.*

En [6], Cornelis y Deschrijver observan que se pueden construir t-normas intuicionistas usando convenientemente t-normas y t-conormas (recordemos que son operaciones binarias en $[0,1]$ conmutativas, asociativas, monótonas y con elemento neutro 1 y 0, respectivamente -véase, por ejemplo, [1, 11]-). Desafortunadamente, el recíproco no siempre es cierto, por lo que para distinguir entre estos dos tipos de t-normas intuicionistas, Deschrijver *et al.* en [8] y [9] proponen la noción de t-representabilidad.

Definición 1.2 ([8, 9]) *Una t-norma intuicionista $\mathcal{T}$ se dice t-representable si existen una t-norma T y una t-conorma S tales que $\mathcal{T}(\overline{a}, \overline{b}) = (T(a_1, b_1), S(a_2, b_2))$, $\forall \overline{a} = (a_1, a_2), \overline{b} = (b_1, b_2) \in \mathbb{L}$.*

A las t-normas intuicionistas que no cumplen la condición anterior se les llama no representables. Ahora estamos en condiciones de introducir la definición de incompatibilidad entre AIFSs:

Definición 1.3 ([4]) *Sea $\mathcal{T} : \mathbb{L}^2 \to \mathbb{L}$ una t-norma intuicionista, diremos que dos conjuntos borrosos de Atanassov A y B, o alternativamente sus funciones de $\mathbb{L}$-pertenencia $\chi^A = (\mu_A, \nu_A), \chi^B = (\mu_B, \nu_B) \in \mathbb{L}^E$, son $\mathcal{T}$-incompatibles si $\mathcal{T}(\chi^A(x), \chi^B(x)) = 0_{\mathbb{L}}$ para todo $x \in E$.*

De forma análoga al caso borroso, no es suficiente con señalar que dos conjuntos de Atanassov son $\mathcal{T}$-incompatibles, o no lo son; por ello, a continuación se introducen los axiomas mínimos que una función debería cumplir para medir el grado de $\mathcal{T}$-incompatibilidad entre dos conjuntos de Atanassov.

Definición 1.4 ([5]) *Sea $\mathcal{T} : \mathbb{L}^2 \to \mathbb{L}$ una t-norma intuicionista. Se dice que $I_{\mathcal{T}} : \mathbb{L}^E \times \mathbb{L}^E \to [0,1]$ es una medida de $\mathcal{T}$-incompatibilidad si verifica las siguientes condiciones:*

(i) $I_{\mathcal{T}}(\chi^{0_{\mathbb{L}}}, \chi^{0_{\mathbb{L}}}) = 1$, *siendo* $\chi^{0_{\mathbb{L}}}(x) = 0_{\mathbb{L}}$, $\forall x \in E$.

(ii) $I_{\mathcal{T}}(\chi^A, \chi^B) = 0$, *si existe* $x \in E$ *tal que* $\mathcal{T}(\chi^A(x), \chi^B(x)) \neq 0_{\mathbb{L}}$.

(iii) Simetría: $I_{\mathcal{T}}(\chi^A, \chi^B) = I_{\mathcal{T}}(\chi^B, \chi^A), \forall \chi^A, \chi^B \in \mathbb{L}^E$.

(iv) Antimonotonía: Dados $\chi^A, \chi^B \in \mathbb{L}^E$ *con* $\chi^A \leq_{\mathbb{L}} \chi^B$ *se tiene que* $I_{\mathcal{T}}(\chi^B, \chi^C) \leq I_{\mathcal{T}}(\chi^A, \chi^C)$ *para todo* $\chi^C \in \mathbb{L}^E$.

2 MEDIDAS DE $\mathcal{T}_{W_\varphi W_\varphi^*}$-INCOMPATIBILIDAD

En esta sección estudiaremos algunas funciones que miden la incompatibilidad respecto a t-normas intuicionistas t-representables asociadas a conectivos conjugados de los de Łukasiewicz. Es decir, si para cualesquiera $a, b \in [0,1]$, $W(a,b) = \mathrm{Max}(0, a+b-1)$ y $W^*(a,b) = \mathrm{Min}(a+b, 1)$ definen la t-norma y la t-conorma de Łukasiewicz, respectivamente, y φ es un automorfismo de orden del intervalo unidad ($\varphi : [0,1] \to [0,1]$ es una función biyectiva y creciente), consideraremos la t-norma t-representable mediante los conectivos conjugados con W y W^* mediante φ, o *φ-conjugados*, $W_\varphi = \varphi^{-1} \circ W \circ (\varphi \times \varphi)$ y $W_\varphi^* = \varphi^{-1} \circ W^* \circ (\varphi \times \varphi)$, i.e. $\mathcal{T}_{W_\varphi W_\varphi^*}(\overline{a}, \overline{b}) = \big((\varphi^{-1} \circ W)(\varphi(a_1), \varphi(b_1)), (\varphi^{-1} \circ W^*)(\varphi(a_2), \varphi(b_2))\big)$ para cualesquiera $\overline{a} = (a_1, a_2), \overline{b} = (b_1, b_2) \in \mathbb{L}$.

Obsérvese que dos AIFSs $\chi^A = (\mu_A, \nu_A), \chi^B = (\mu_B, \nu_B) \in \mathbb{L}^X$ son $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibles si y sólo si para todo $x \in E$ se verifica que $W_\varphi(\mu_A(x), \mu_B(x)) = 0$ y $W_\varphi^*(\nu_A(x), \nu_B(x)) = 1$, lo que equivale a que

$$\forall x \in E \begin{cases} \varphi(\mu_A(x)) + \varphi(\mu_B(x)) \leq 1 \\ \varphi(\nu_A(x)) + \varphi(\nu_B(x)) \geq 1 \end{cases} \quad (1)$$

Por lo que si consideramos las regiones en $[0,1]^2$:

$$\begin{aligned} R_\varphi &= \{(a,b) \in [0,1]^2 \mid \varphi(a) + \varphi(b) \leq 1\} \\ R_\varphi^* &= \{(a,b) \in [0,1]^2 \mid \varphi(a) + \varphi(b) \geq 1\}, \end{aligned} \quad (2)$$

se tiene que χ^A y χ^B son $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibles si y sólo si, para todo $x \in E$, se verifica que $(\mu_A(x), \mu_B(x)) \in R_\varphi$ y $(\nu_A(x), \nu_B(x)) \in R_\varphi^*$; por ello tiene sentido denominar a R_φ y R_φ^* *regiones de $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad*. Esta observaciones sugieren que una posible forma de cuantificar la incompatibilidad sería medir, con la distancia euclídea, cuánto falta para que (μ_A, μ_B) ó (ν_A, ν_B) alcancen la frontera entre las regiones de incompatibilidad (véase Fig. 1).

Así, se tiene la familia de funciones, dependiente de la familia de automorfismos de orden en $[0,1]$, dada en la siguiente proposición e introducida en [5], que es una familia de medidas de incompatibilidad como a continuación probaremos.

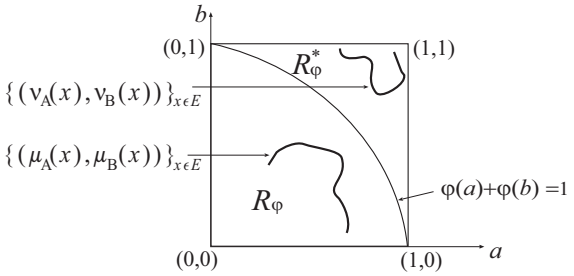


Figura 1: Regiones de $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad y AIFSs χ^A y χ^B $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibles

Proposición 2.1 *Sean φ un automorfismo de orden de $[0,1]$ y W_φ, W_φ^* la t-norma y la t-conorma φ-conjugadas con las de Łukasiewicz. Considérense la t-norma t-representable asociada a W_φ y W_φ^*, $\mathcal{T}_{W_\varphi W_\varphi^*}$, y la función $I_{W_\varphi W_\varphi^*} : \mathbb{L}^E \times \mathbb{L}^E \to [0,1]$ definida, para cada $\chi^A = (\mu_A, \nu_A), \chi^B = (\mu_B, \nu_B) \in L^E$, por*

$$I_{W_\varphi W_\varphi^*}(\chi^A, \chi^B) = \mathrm{Min}\left(\frac{\inf\limits_{x \in E} d\big((\mu_A(x), \mu_B(x)), R_\varphi^*\big)}{d((0,0), R_\varphi^*)}, \frac{\inf\limits_{x \in E} d((\nu_A(x), \nu_B(x)), R_\varphi)}{d((1,1), R_\varphi)} \right),$$

donde d es la distancia euclídea, y R_φ y R_φ^* son las regiones de $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad definidas en (2). Se verifica que $I_{W_\varphi W_\varphi^*}$ es una medida de $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad.

Demostración. (i) $I_{W_\varphi W_\varphi^*}(\chi^{0_{\mathbb{L}}}, \chi^{0_{\mathbb{L}}}) = 1$ se comprueba trivialmente.

(ii) Sean $\chi^A, \chi^B \in \mathbb{L}^E$, y supongamos que existe $x \in E$ tal que $\mathcal{T}_{W_\varphi W_\varphi^*}(\chi^A(x), \chi^B(x)) \neq 0_{\mathbb{L}}$, esto es, $W_\varphi(\mu_A(x), \mu_B(x)) > 0$ ó $W_\varphi^*(\nu_A(x), \nu_B(x)) < 1$, o lo que es lo mismo $\varphi(\mu_A(x)) + \varphi(\mu_B(x)) > 1$ ó $\varphi(\nu_A(x)) + \varphi(\nu_B(x)) < 1$, es decir $(\mu_A(x), \mu_B(x)) \in R_\varphi^*$ ó $(\nu_A(x), \nu_B(x)) \in R_\varphi$. Así, $I_{W_\varphi W_\varphi^*}(\chi^A, \chi^B) = 0$.

(iii) Sean $\chi^A, \chi^B \in \mathbb{L}$. Dado que $(a,b) \in R_\varphi$ si y sólo si $(b,a) \in R_\varphi$ (i. e. R_φ es simétrica respecto a $a+b=1$), y lo mismo sucede con R_φ^*, entonces para todo $x \in E$ se verifica que

$$\begin{aligned} d\left((\mu_A(x), \mu_B(x)), R_\varphi^*\right) &= \inf_{(a,b) \in R_\varphi^*} d\left((\mu_A(x), \mu_B(x)), (a,b)\right) \\ &= \inf_{(b,a) \in R_\varphi^*} d\left((\mu_B(x), \mu_A(x)), (b,a)\right) \\ &= d\left((\mu_B(x), \mu_A(x)), R_\varphi^*\right) \end{aligned}$$

De la misma manera, se obtiene que $d((\nu_A(x),\nu_B(x)),R_\varphi) = d((\nu_B(x),\nu_A(x)),R_\varphi)$. Por tanto, se concluye que $I_{W_\varphi W_\varphi^*}$ es simétrica.

(iv) Sean $\chi^A,\chi^B,\chi^C \in \mathbb{L}^E$ tales que $\chi^A \leq_{\mathbb{L}} \chi^B$. Para probar la antimonotonía de $I_{W_\varphi W_\varphi^*}$ es suficiente probar que para todo $x \in E$ se satisfacen las desigualdades $d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right) \geq d\left((\mu_B(x),\mu_C(x)),R_\varphi^*\right)$ y $d((\nu_A(x),\nu_C(x)),R_\varphi) \geq d((\nu_B(x),\nu_C(x)),R_\varphi)$. Veamos que así es.

Si $d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right) = 0$, se tiene que $(\mu_A(x),\mu_C(x)) \in R_\varphi^*$, por ser R_φ^* cerrado, luego $\varphi(\mu_A(x))+\varphi(\mu_C(x)) \geq 1$. Además, de $\mu_A(x) \leq \mu_B(x)$, para todo $x \in E$, se sigue que $\varphi(\mu_B(x))+\varphi(\mu_C(x)) \geq 1$, esto es, $(\mu_B(x),\mu_C(x)) \in R_\varphi^*$, y por lo tanto, $d\left((\mu_B(x),\mu_C(x)),R_\varphi^*\right) = 0$.

Si $d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right) > 0$, entonces, dado que R_φ^* es un conjunto compacto, la distancia se alcanza en la frontera de R_φ^*, que denotaremos por $B(R_\varphi^*)$. Sea $(a^*,b^*) \in B(R_\varphi^*)$ el punto tal que

$$d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right) = d((\mu_A(x),\mu_C(x)),(a^*,b^*))$$

y sean los tres segmentos siguientes contenidos en el segmento $[\mu_A(x),1]\times\{\mu_C(x)\}$ (véase Fig. 2):

$$I_1 = [\mu_A(x),a^*]\times\{\mu_C(x)\},$$

$$I_2 = [a^*,\varphi^{-1}(1-\varphi(\mu_C(x))]\times\{\mu_C(x)\},$$

$$I_3 = [\mu_A(x),1]\times\{\mu_C(x)\}\cap R_\varphi^*.$$

Dado que $\mu_A(x) \leq \mu_B(x)$ para todo $x \in E$, se tiene que $(\mu_B(x),\mu_C(x)) \in [\mu_A(x),1]\times\{\mu_C(x)\} = I_1\cup I_2\cup I_3$

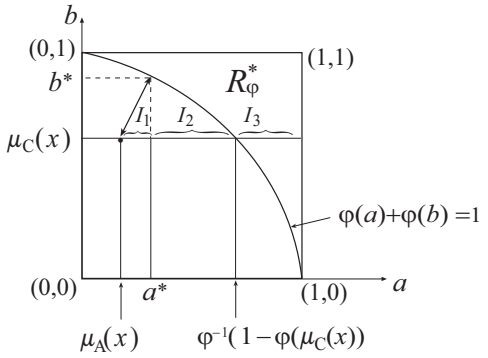


Figura 2: Segmentos I_1, I_2 e I_3 y $d((\mu_A(x),\mu_C(x)),R_\varphi^*)$

Así hay tres posibilidades:

a) Si $(\mu_B(x),\mu_C(x)) \in I_1$ entonces $\mu_A(x) \leq \mu_B(x) \leq a^*$ y así $0 \leq a^*-\mu_B(x) \leq a^*-\mu_A(x)$, por lo que $(a^*-\mu_B(x))^2 \leq (a^*-\mu_A(x))^2$ y

$d\left((\mu_B(x),\mu_C(x)),R_\varphi^*\right) \leq$

$$\begin{aligned}
&\leq d((\mu_B(x),\mu_C(x)),(a^*,b^*))\\
&= \sqrt{(a^*-\mu_B(x))^2+(b^*-\mu_C(x))^2}\\
&\leq \sqrt{(a^*-\mu_A(x))^2+(b^*-\mu_C(x))^2}\\
&= d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right).
\end{aligned}$$

b) Si $(\mu_B(x),\mu_C(x)) \in I_2$ entonces

$d\left((\mu_B(x),\mu_C(x)),R_\varphi^*\right) \leq$

$$\begin{aligned}
&\leq d\left((\mu_B(x),\mu_C(x)),(\mu_B(x),\varphi^{-1}(1-\varphi(\mu_B(x))))\right)\\
&= \varphi^{-1}(1-\varphi(\mu_B(x)))-\mu_C(x).
\end{aligned}$$

Además,
$a^* \leq \mu_B(x)$ implica que $b^* = \varphi^{-1}(1-\varphi(a^*)) \geq \varphi^{-1}(1-\varphi(\mu_B(x)))$ y así $\varphi^{-1}(1-\varphi(\mu_B(x)))-\mu_C(x) \leq b^*-\mu_C(x) \leq \sqrt{(a^*-\mu_A(x))^2+(b^*-\mu_C(x))^2} = d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right)$.

c) Si $(\mu_B(x),\mu_C(x)) \in I_3$ entonces
$d\left((\mu_B(x),\mu_C(x)),R_\varphi^*\right) = 0 \leq d\left((\mu_A(x),\mu_C(x)),R_\varphi^*\right)$.

De la misma forma, se prueba que $d((\nu_A(x),\nu_C(x)),R_\varphi) \geq d((\nu_B(x),\nu_C(x)),R_\varphi)$.

Luego, ha quedado demostrado que $I_{W_\varphi W_\varphi^*}$ es una medida de $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad. □

El resto de esta sección se dedica a construir otras funciones para medir la $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibilidad. En primer lugar, continuaremos con medidas definidas a través de la distancia euclídea y las regiones de incompatibilidad; y, finalmente, construiremos otra función aplicando otro punto de vista.

Observación 2.2 Sea φ un *automorfismo de orden recíproco* del intervalo unidad, esto es, un automorfismo de orden que cumpla con la condición $\varphi(1-t) = 1-\varphi(t)$ para todo $t \in [0,1]$. Si $\chi^A = (\mu_A,\nu_A), \chi^B = (\mu_B,\nu_B) \in \mathbb{L}^E$ son $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibles, entonces la segunda desigualdad de (1) es equivalente a $1 \geq \varphi(1-\nu_A(x))+\varphi(1-\nu_B(x))$ para todo $x \in E$. Por lo tanto, χ^A y χ^B son $\mathcal{T}_{W_\varphi W_\varphi^*}$-incompatibles, si y sólo si, para todo $x \in E$, se verifica que $(\mu_A(x),\mu_B(x)) \in R_\varphi$ y $(1-\nu_A(x),1-\nu_B(x)) \in R_\varphi$.

Esta observación sugiere construir otra familia de medidas de incompatibilidad asociada ahora a la familia de automorfismos recíprocos del intervalo unidad, considerando sólo la región R_φ^*, como indica la siguiente proposición, cuya demostración es similar a la anterior.

Proposición 2.3 *Sean φ un automorfismo de orden recíproco del intervalo unidad y $\mathcal{T}_{W_\varphi W_\varphi^*}$ la t-norma t-representable asociada a W_φ y W_φ^*, siendo W_φ, W_φ^* la t-norma y la t-conorma φ-conjugadas con las de Łukasiewicz. Considérese $I^2_{W_\varphi W_\varphi^*} : \mathbb{L}^E\times\mathbb{L}^E \to [0,1]$*

tal que para cada $\chi^A=(\mu_A,\nu_A),\chi^B=(\mu_B,\nu_B)\in\mathbb{L}^E$,

$$I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)=\mathrm{Min}\left(\frac{\inf\limits_{x\in E} d\left((\mu_A(x),\mu_B(x)),R^*_\varphi\right)}{d((0,0),R^*_\varphi)},\right.$$

$$\left.\frac{\inf\limits_{x\in E} d\left((1-\nu_A(x),1-\nu_B(x)),R^*_\varphi\right)}{d((0,0),R^*_\varphi)}\right),$$

donde d es la distancia euclídea y R^*_φ *la región definida (2). Se verifica que* $I^2_{W_\varphi W^*_\varphi}$ *es una medida de* $\mathcal{T}_{W_\varphi W^*_\varphi}$*-incompatibilidad.*

En la proposición anterior se ha limitado el resultado a automorfismos recíprocos, ya que de lo contrario no se cumpliría la condición (ii) de la definición de medidas de $\mathcal{T}$-incompatibilidad. Por ejemplo, si para todo $x\in E$ es $\chi^A(x)=(\frac{1}{4},\frac{1}{2})$ y $\chi^B(x)=(\frac{1}{8},\frac{1}{2})$, y considerando el automorfismo $\varphi(a)=a^2$, se tiene que χ^A y χ^B no son $\mathcal{T}_{W_\varphi W^*_\varphi}$-incompatibles, pero $I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$ es mayor que cero. Por ello, para salvar este inconveniente se propone la siguiente familia de medidas de $\mathcal{T}_{W_\varphi W^*_\varphi}$-incompatibilidad.

Proposición 2.4 *Sean* φ *un automorfismo de orden de* $[0,1]$ *y* W_φ,W^*_φ *las t-norma y t-conorma* φ*-conjugadas con las de Łukasiewicz. Si* $\mathcal{T}_{W_\varphi W^*_\varphi}$ *es la t-norma t-representable asociada a* W_φ *y* W^*_φ*, considérese* $I^3_{W_\varphi W^*_\varphi}:\mathbb{L}^E\times\mathbb{L}^E\to[0,1]$ *tal que para cada* $\chi^A=(\mu_A,\nu_A),\chi^B=(\mu_B,\nu_B)\in L^E$,

$$I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)=\mathrm{Min}\left(\frac{\inf\limits_{x\in E} d\left((\mu_A(x),\mu_B(x)),R^*_\varphi\right)}{d((0,0),R^*_\varphi)},\right.$$

$$\left.\frac{\inf\limits_{x\in E} d\left((\varphi^{-1}(1-\varphi(\nu_A(x))),\varphi^{-1}(1-\varphi(\nu_B(x)))),R^*_\varphi\right)}{d((0,0),R^*_\varphi)}\right),$$

donde d es la distancia euclídea y R^*_φ *la región definida en (2). Entonces,* $I^3_{W_\varphi W^*_\varphi}$ *es una medida de* $\mathcal{T}_{W_\varphi W^*_\varphi}$*-incompatibilidad.*

Nótese que esta función mide la mínima distancia a la región R^*_φ de los conjuntos imagen $Im(\mu_A\times\mu_B)$ e $Im(N_\varphi\circ\nu_A\times N_\varphi\circ\nu_B)$, siendo $N_\varphi(a)=\varphi^{-1}(1-\varphi(a))$, que es una negación fuerte (isomorfismo de $[0,1]$ decreciente e involutivo).

Observación 2.5 Si φ es una automorfismo de orden recíproco de $[0,1]$, entonces las medidas $I^2_{W_\varphi W^*_\varphi}$ y $I^3_{W_\varphi W^*_\varphi}$, construidas en las proposiciones anteriores, coinciden.

Finalizamos esta sección con otra forma de medir la $\mathcal{T}_{W_\varphi W^*_\varphi}$-incompatibilidad motivada de manera diferente a las anteriores. Sabemos que $\chi^A=(\mu_A,\nu_A),\chi^B=(\nu_B,\nu_B)\in\mathbb{L}^E$ son $\mathcal{T}_{W_\varphi W^*_\varphi}$-incompatibles si y sólo si se verifican las inecuaciones de (1), pero esto es equivalente a que

$$\sup_{x\in E}\left(\varphi(\mu_A(x))+\varphi(\mu_B(x))\right)\le 1$$
$$\inf_{x\in E}\left(\varphi(\nu_A(x))+\varphi(\nu_B(x))\right)\ge 1$$

Por ello, tiene sentido cuantificar la incompatibilidad calculando cuánto falta para que una de las anteriores desigualdades no se cumpla. Así, se tiene el siguiente resultado, cuya prueba sigue técnicas parecidas a las de la proposición 3.2

Proposición 2.6 *Sean* φ *un automorfismo del intervalo unidad y* W_φ,W^*_φ *las t-norma y t-conorma conjugadas a las de Łukasiewicz mediante* φ*. Si* $\mathcal{T}_{W_\varphi W^*_\varphi}$ *es la t-norma t-representable asociada con* W_φ *y* W^*_φ*, considérese* $I^4_{W_\varphi W^*_\varphi}:\mathbb{L}^E\times\mathbb{L}^E\to[0,1]$ *tal que para* $\chi^A=(\mu_A,\nu_A),\chi^B=(\mu_B,\nu_B)\in\mathbb{L}^E$*, es* $I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)=$

$$\mathrm{Max}\Bigg(0,\mathrm{Min}\left(1-\sup_{x\in E}(\varphi(\mu_A(x))+\varphi(\mu_B(x))),\right.$$
$$\left.\inf_{x\in E}(\varphi(\nu_A(x))+\varphi(\nu_B(x))-1)\right)\Bigg).$$

Se verifica que $I^4_{W_\varphi W^*_\varphi}$ *es una medida de* $\mathcal{T}_{W_\varphi W^*_\varphi}$*-incompatibilidad.*

3 MEDIDAS DE $\mathcal{T}_{W_\varphi}$-INCOMPATIBILIDAD

Sea φ un automorfismo de orden del intervalo unidad y consideremos la t-norma intuicionista no representable [4, 8] definida para $(\overline{a},\overline{b})=(a_1,a_2,b_1,b_2)\in\mathbb{L}^2$ por

$$\mathcal{T}_{W_\varphi}(\overline{a},\overline{b})=\begin{cases}\overline{a} & \text{si } \overline{b}=1_{\mathbb{L}}\\ \overline{b} & \text{si } \overline{a}=1_{\mathbb{L}}\\ \left(W_\varphi(a_1,b_1),W^*_\varphi(N_\varphi(a_1),N_\varphi(b_1))\right) & \text{e.o.c}\end{cases}\quad(3)$$

Al igual que en la sección anterior, para cada automorfismo de orden de $[0,1]$ construiremos dos medidas de $\mathcal{T}_{W_\varphi}$-incompatibilidad según las dos distintas vías anteriores.

Observemos que $\chi^A=(\mu_A,\nu_A),\chi^B=(\mu_B,\nu_B)\in\mathbb{L}^E$ son $\mathcal{T}_{W_\varphi}$-incompatibles si y sólo si alguno de ellos es $\chi^{0_{\mathbb{L}}}$ o si para todo $x\in E$, $(W_\varphi(\mu_A(x),\mu_B(x)),W^*_\varphi(N_\varphi(\mu_A(x)),N_\varphi(\mu_B(x))))=0_{\mathbb{L}}$, lo que equivale a que $\varphi(\mu_A(x))+\varphi(\mu_B(x))\le 1$. Por tanto, χ^A,χ^B son $\mathcal{T}_{W_\varphi}$-incompatibles si y sólo si, para todo $x\in E$, $(\mu_A(x),\mu_B(x))\in R_\varphi$, siendo R_φ la región definida en (1), por lo que R_φ también será la *región de* $\mathcal{T}_{W_\varphi}$*-incompatibilidad*. Esta observación sugiere construir una familia de medidas de incompatibilidad asociada a la familia de automorfismos del intervalo unidad, considerando cuánto falta para que (μ_A,μ_B) alcance la frontera de la región de $\mathcal{T}_{W_\varphi}$-incompatibilidad con R^*_φ. Veamos:

Proposición 3.1 *Sea $\mathcal{T}_{W_\varphi}$ la t-norma definida en (3) y consideremos la función $I_{W_\varphi} : \mathbb{L}^E \times \mathbb{L}^E \to [0,1]$ definida para $\chi^A{=}(\mu_A,\nu_A), \chi^B{=}(\mu_B,\nu_B) \in \mathbb{L}^E$ por*

$$I_{W_\varphi}(\chi^A,\chi^B){=}\frac{\underset{x\in E}{\text{Inf}}\, d\left((\mu_A(x),\mu_B(x)),R^*_\varphi\right)}{d((0,0),R^*_\varphi)},$$

donde d es la distancia euclídea y R^*_φ es la región definida en (2). Entonces, I_{W_φ} es una medida de $\mathcal{T}_{W_\varphi}$-incompatibilidad.

Proposición 3.2 *Sea $\mathcal{T}_{W_\varphi}$ la t-norma definida en (3). La función $I^2_{W_\varphi} : \mathbb{L}^E \times \mathbb{L}^E \to [0,1]$ definida para cada $\chi^A{=}(\mu_A,\nu_A), \chi^B{=}(\mu_B,\nu_B) \in \mathbb{L}^E$ por*

$$I^2_{W_\varphi}(\chi^A,\chi^B){=}\text{Max}\Big(0, 1 - \underset{x\in E}{\text{Sup}}(\varphi(\mu_A(x)) + \varphi(\mu_B(x)))\Big),$$

es una medida de $\mathcal{T}_{W_\varphi}$-incompatibilidad.

Demostración. Es trivial comprobar que $I^2_{W_\varphi}(\chi^{0_\mathbb{L}},\chi^{0_\mathbb{L}}){=}1$.

(ii) Sean $\chi^A, \chi^B \in \mathbb{L}^E$, y supongamos que existe $x \in E$ tal que $\mathcal{T}_{W_\varphi}(\chi^A(x), \chi^B(x)) \neq 0_\mathbb{L}$, se verifica:

a) Si $\chi^A(x) = 1_\mathbb{L}$ ó $\chi^B(x) = 1_\mathbb{L}$, entonces $\mu_A(x) = 1$ ó $\mu_B(x) = 1$, y así $\underset{x\in E}{\text{Sup}}\,(\varphi(\mu_A(x)) + \varphi(\mu_B(x))) \geq 1$.

b) Si $\chi^A(x) \neq 1_\mathbb{L}$ y $\chi^B(x) \neq 1_\mathbb{L}$, $\mathcal{T}_{W_\varphi}(\chi^A(x), \chi^B(x)) = (W_\varphi(\mu_A(x),\mu_B(x)), W^*_\varphi(N_\varphi(\mu_A(x)), N_\varphi(\mu_B(x)))) \neq 0_\mathbb{L}$, lo que implica que $\varphi(\mu_A(x)) + \varphi(\mu_B(x)) > 1$ ó $\varphi(\varphi^{-1}(1 - \varphi(\mu_A(x)))) + \varphi(\varphi^{-1}(1 - \varphi(\mu_B(x)))) < 1$, por lo que $\underset{x\in E}{\text{Sup}}\,(\varphi(\mu_A(x))+\varphi(\mu_B(x))) > 1$.

Por consiguiente, tanto si se da a) como b), se obtiene que $I^2_{W_\varphi}(\chi^A,\chi^B) = 0$.

(iii) Que $I^2_{W_\varphi}$ es simétrica es inmediato por definición.

(iv) Sean $\chi^A, \chi^B, \chi^C \in \mathbb{L}^E$ tales que $\chi^A \leq_\mathbb{L} \chi^B$, entonces $\mu_A(x) \leq \mu_B(x)$ para todo $x \in E$, y por ser φ creciente $\varphi(\mu_A(x)) \leq \varphi(\mu_B(x))$ para todo $x \in E$, de donde se tiene que

$$\text{Max}\Big(0, 1 - \underset{x\in E}{\text{Sup}}(\varphi(\mu_A(x)) + \varphi(\mu_C(x)))\Big)$$
$$\geq \text{Max}\Big(0, 1 - \underset{x\in E}{\text{Sup}}(\varphi(\mu_B(x)) + \varphi(\mu_C(x)))\Big).$$

Esto es, $I^2_{W_\varphi}(\chi^A,\chi^C) \geq I^2_{W_\varphi}(\chi^B,\chi^C)$.

Luego, $I^2_{W_\varphi}$ es una medida de $\mathcal{T}_{W_\varphi}$-incompatibilidad. □

4 COMPARACIÓN ENTRE LAS MEDIDAS DE INCOMPATIBILIDAD

Sea $\mathcal{T}_{W_\varphi W^*_\varphi}$ la t-norma intuicionista t-representable asociada con W_φ y W^*_φ y consideremos las cuatro medidas de $\mathcal{T}_{W_\varphi W^*_\varphi}$-incompatibilidad definidas en la Sección 2, entonces se cumple la siguiente relación:

Proposición 4.1 *Sea el automorfismo de orden $\varphi = id_{[0,1]}$. Entonces para todo $\chi^A = (\mu_A,\nu_A), \chi^B = (\mu_B,\nu_B) \in \mathbb{L}^E$ es cierto que:*

$$I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}I^i_{W_\varphi W^*_\varphi}(\chi^A,\chi^B),\ \textit{para } i = 2,3,4.$$

Demostración. En efecto, dado que $\varphi = id_{[0,1]}$ es obvio que $I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$; además

$$I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}\text{Max}\Big(0, \text{Min}(1{-}\underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x)),$$
$$\underset{x\in E}{\text{Inf}}(\nu_A(x) + \nu_B(x) - 1)\Big).$$

Luego, hay tres posibilidades para $I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$:

a) Si $I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = 0$, entonces $1 - \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x)) \leq 0$ ó $\underset{x\in E}{\text{Inf}}(\nu_A(x) + \nu_B(x) - 1) \leq 0$.

Supongamos en primer lugar, que $1 - \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x)) \leq 0$. Si además fuese $1 < \underset{x\in E}{\text{Sup}}(\mu_A(x){+}\mu_B(x))$, entonces existiría $x_0 \in E$ tal que $(\mu_A(x_0),\mu_B(x_0)) \in R^*_\varphi$, y así $I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = 0$. Si por el contrario, fuese $1 = \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x))$, entonces para cada $n \in \mathbb{N}$ se puede encontrar $x_n \in E$ tal que $d\left((\mu_A(x_n),\mu_B(x_n)),R^*_\varphi\right) < 1/n$, con lo que $\underset{x\in E}{\text{Inf}}\, d\left((\mu_A(x),\mu_B(x)),R^*_\varphi\right) = 0$, y así $I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}0$.

Supongamos ahora que $\underset{x\in E}{\text{Inf}}(\nu_A(x) + \nu_B(x) - 1) \leq 0$. Por una parte, como antes, se llega a que $\underset{x\in E}{\text{Inf}}\, d\left((\nu_A(x),\nu_B(x)),R_\varphi\right) = 0$, obteniéndose que $I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}0$. Por otra parte, como $0 \leq 1 - \underset{x\in E}{\text{Inf}}(\nu_A(x){+}\nu_B(x)){=}\underset{x\in E}{\text{Sup}}(1 - \nu_A(x) - \nu_B(x))$, entonces $\underset{x\in E}{\text{Sup}}(1{-}\nu_A(x)){+}(1{-}\nu_B(x)) \geq 1$, y de nuevo, como antes, $\underset{x\in E}{\text{Inf}}\, d\left((1 - \nu_A(x), 1 - \nu_B(x)),R^*_\varphi\right) = 0$, con lo que $I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B){=}0$.

b) Si $0 < I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = 1 - \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x))$, entonces de la primer desigualdad se tiene que $\mu_B(x) \leq 1 - \mu_A(x)$ para todo $x \in E$, y de la segunda, que $1 - \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x)) \leq \underset{x\in E}{\text{Inf}}(\nu_A(x) + \nu_B(x) - 1) = 1{-}\underset{x\in E}{\text{Sup}}((1{-}\nu_A(x)){+}(1{-}\nu_B(x)))$, por lo que $\underset{x\in E}{\text{Sup}}((1{-}\nu_A(x)) + (1 - \nu_B(x))) \leq \underset{x\in E}{\text{Sup}}(\mu_A(x) + \mu_B(x)) \leq 1$; y así (véase Fig. 3):

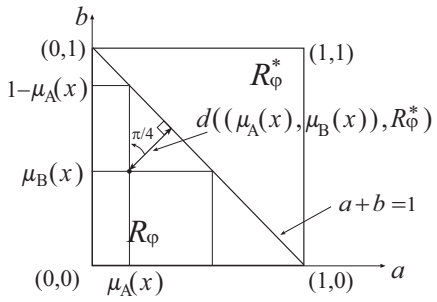


Figura 3: Distancia de $(\mu_A(x), \mu_B(x))$ a R^*_φ

$I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = \frac{\inf\limits_{x\in E} d((\mu_A(x),\mu_B(x)),R^*_\varphi)}{d((0,0),R^*_\varphi)} = \frac{\inf\limits_{x\in E}(1-\mu_A(x)-\mu_B(x))\frac{\sqrt{2}}{2}}{\frac{\sqrt{2}}{2}} = I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$.

c) Si $0 < I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = \inf\limits_{x\in E}(\nu_A(x)+\nu_B(x)-1)$, entonces $\nu_B(x) \geq 1-\nu_A(x)$, para todo $x \in E$, y $1 - \sup\limits_{x\in E}(\mu_A(x)+\mu_B(x)) \geq \inf\limits_{x\in E}(\nu_A(x)+\nu_B(x)-1)$, por lo que (véase Fig. 4):

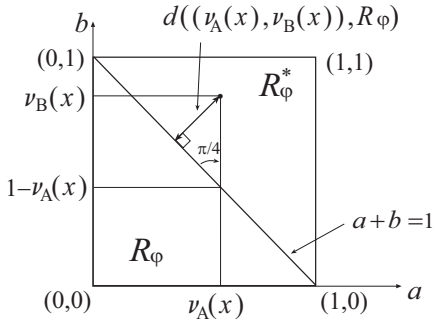


Figura 4: Distancia de $(\nu_A(x), \nu_B(x))$ a R_φ

$I_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = \frac{\inf\limits_{x\in E} d((\nu_A(x),\nu_B(x)),R_\varphi)}{d((1,1),R_\varphi)} = \frac{\inf\limits_{x\in E}(\nu_B(x)-(1-\nu_A(x)))\frac{\sqrt{2}}{2}}{\frac{\sqrt{2}}{2}} = I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$.

Y además, $\nu_A(x) \geq 1-\nu_B(x)$ para todo $x \in E$, por lo que para cada $x \in E$ $I^2_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = I^3_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = \frac{\inf\limits_{x\in E} d((1-\nu_A(x),1-\nu_B(x)),R^*_\varphi)}{d((0,0),R^*_\varphi)} = \frac{\inf\limits_{x\in E}(\nu_A(x)-(1-\nu_B(x)))\frac{\sqrt{2}}{2}}{\frac{\sqrt{2}}{2}} = I^4_{W_\varphi W^*_\varphi}(\chi^A,\chi^B)$. □

Sea $\mathcal{T}_{W_\varphi}$ la t-norma intuicionista no representable definida previamente y consideremos las dos medidas de $\mathcal{T}_{W_\varphi}$-incompatibilidad dadas en la Sección 3, entonces se cumple la siguiente relación:

Proposición 4.2 *Sea el automorfismo de orden* $\varphi = id_{[0,1]}$. *Entonces, para todos* $\chi^A = (\mu_A,\nu_A), \chi^B = (\mu_B,\nu_B) \in \mathbb{L}^E$ *es cierto* $I_{W_\varphi}(\chi^A,\chi^B) = I^2_{W_\varphi}(\chi^A,\chi^B)$.

Por último, es inmediato:

Proposición 4.3 *Dados el automorfismo de orden* $\varphi = id_{[0,1]}$ *y* $\chi^A = (\mu_A,\nu_A), \chi^B = (\mu_B,\nu_B) \in \mathbb{L}^E$. *Se tiene que si* $0 < 1 - \sup\limits_{x\in E}(\mu_A(x)+\mu_B(x)) \leq \inf\limits_{x\in E}(\nu_A(x)+\nu_B(x)-1)$, *entonces es cierto que:*

$$I^i_{W_\varphi W^*_\varphi}(\chi^A,\chi^B) = I^j_{W_\varphi}(\chi^A,\chi^B), \quad i = 1,..,4,\ j = 1,2$$

5 CONCLUSIONES

En este trabajo se han presentado, en primer lugar, diversas funciones que miden el grado de incompatibilidad entre dos conjuntos borrosos de Atanassov, considerando para ello la familia de t-normas intuicionistas t-representables $\mathcal{T}_{W_\varphi W^*_\varphi}$ asociada con W_φ y W^*_φ. En segundo lugar, se definen diferentes medidas de incompatibilidad considerando ahora la familia de t-normas intuicionistas no representables $\mathcal{T}_{W_\varphi}$. Y finalmente, se ha estudiado la relación entre las distintas medidas de incompatibilidad propuestas.

Como inmediatas y futuras líneas de investigación que continuarían este trabajo, entre otras, señalamos las siguientes: ampliación del modelo teórico considerando axiomas que modelicen la continuidad de las medidas de compatibilidad; búsqueda de mecanismos para agregar estas medidas; relación con otras medidas borrosas, en particular, con las medidas de contradicción; investigación de su uso para evaluar los resultados en los procesos de inferencia.

Referencias

[1] C. Alsina, M. Frank, B. Schweizer. *Associative Functions: Triangular Norms and Copulas*, World Scientific (Singapore, Thailand), 2006.

[2] K. T. Atanassov. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, **20**, Pág. 87-96, 1986.

[3] K. T. Atanassov, *Intuitionistic Fuzzy Sets*, Physica-Verlag (New York, USA), 1999.

[4] E. Castiñeira, S. Cubillo. Algunas propiedades algebraicas de los conjuntos borrosos intuicionistas. *Actas del XII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)*, Pág. 297-302, Jaén, 2004.

[5] E. Castiñeira, S. Cubillo, W. Montilla. On the incompatibility between two AIFS. Aparecerá en *Proceedings of the 8th International FLINS Conference on Computational Intelligence in Decision and Control*, Madrid, 2008.

[6] C. Cornelis, G. Deschrijver. The compositional rule of inference in an intuitionitic fuzzy logic framework. *Proceedings of 13th European Summer School in Logic, Language and Information (ESSLLI)*, Pág. 83-94, Helsinki, 2001.

[7] S. Cubillo, C. Torres, E. Castiñeira. Self-contradiction and contradiction between two intuitionistic fuzzy sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **16** (3), Pág. 283-300.

[8] G. Deschrijver, C. Cornelis, E. Kerre. Intuitionistic fuzzy connectives revisited. *Proceedings of IX Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Annecy (France), Pág. 1839-1844, 2002.

[9] G. Deschrijver, C. Cornelis, E. Kerre, On the Representation of Intuitionistic Fuzzy T-Norms and T-Conorms, *IEEE Transactions on Fuzzy Systems*, **12** (1), Pág. 45-61, 2004.

[10] J. A. Goguen. L-fuzzy sets. *Journal of Mathematical Analysis and Applications*, **18**, Pág. 145-174, 1967.

[11] E.P. Klement, R. Mesiar, E. Pap. *Triangular Norms*, Kluwer Academic Publishers (Dordrecht, Netherlands), 2000.

A RESTRICTION LEVEL APPROACH TO PROBABILITY AND STATISTICS

Daniel Sánchez, Miguel Delgado, and María-Amparo Vila

Depto. Ciencias de la Computación e I.A., Universidad de Granada
E.T.S.I.I.T., C/Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
{daniel,vila}@decsai.ugr.es, mdelgado@ugr.es

Abstract

In this paper we propose to employ RL-numbers, a representation of imprecise quantities, in order to represent probabilities of imprecise events. The resulting probabilities are called RL-probabilities. We extend several operations on probabilities to the imprecise case, and we show that the usual properties of crisp probabilities are preserved. As additional advantages, the extension of crisp operations to the imprecise case and their implementation is unique and straightforward. Finally, we describe the application of the approach in order to perform statistical tests on imprecise data.

Keywords: RL-numbers, Gradual Numbers, Probabilities of Imprecise Events, RL-probability, Fuzzy Statistics.

1 INTRODUCTION

Different approaches have been proposed in order to represent and operate with imprecise probabilities and probabilities of fuzzy events (see [1] for a review). One of the main drawbacks of the existing approaches is how to extend the usual arithmetic operations on crisp probabilities so that the ordinary properties of the probability theory and statistics are preserved. As an example, probabilities represented by fuzzy numbers have to deal with the well-known drawbacks of fuzzy arithmetic, like the increment in the imprecision of the results after several operations and the loss of intuitive properties like $p(A) + p(\bar{A}) = 1$.

In this paper we propose RL-numbers to represent the probability of imprecise events. RL-numbers were proposed in [4] as a representation of imprecise quantities that preserve the ordinary properties of arithmetic operations. At the same time, the imprecision of the results is not necessarily greater than the imprecision of the arguments (it can be even the case that the result of an operation between RL-numbers is a crisp number). As additional advantages, the proposal is simple to implement and understand, and the extension of crisp operations to the imprecise case is unique and straightforward. We call probabilities represented by RL-numbers *RL-probabilities*.

We also propose an extension of statistical tests to the case of imprecise data based on RL-probability and operations on them. The result is a measure of the evidence in favor or against rejecting the null hypothesis. The rationale behind this result is that, since there is imprecision in the data, there is imprecision in the result.

The paper is organized as follows: in section 2 we give the definition of RL-numbers as introduced in [4] on the basis of the RL-representation of imprecise properties introduced in [3]. We introduce probabilities based on RL-numbers in section 3, and we discuss about the properties of this proposal. Section 4 introduces our approach to statistical tests based on RL-probabilities. Finally, section 5 is devoted to our conclusions and future research.

2 RL-REPRESENTATION AND RL-NUMBERS

RL-numbers are based on the RL-representation of imprecise properties proposed in [3]. Starting from a slightly different approach, they were previously proposed in [2], where they are called *gradual numbers*. We summarize RL-representation as proposed in [3] in section 2.1; then we explain the definition of RL-numbers we proposed in [4] in section 2.2.

2.1 RL-REPRESENTATION

The RL-representation of an imprecise property is a collection of crisp sets, each crisp set corresponding to a crisp realization of the property under a *restriction* rule. We distinguish between *atomic* and *derived* properties. Atomic properties are those that cannot be defined in terms of other properties in our problem. Derived properties are defined by logical operations on other properties.

In [3] we consider that atomic imprecise properties are represented by fuzzy sets, and hence restrictions are of the form *degree* $\geq \alpha$ with $\alpha \in (0,1]$, and restriction levels are associated to values $\alpha \in (0,1]$. In the same case, the crisp realization of an atomic imprecise property represented by a fuzzy set A in the restriction level α corresponds to the α-cut A_α.

For every property we assume that there is a finite set of restriction levels $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verifying that $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$, $m \geq 1$. We call such sets *RL-sets*. The consideration that a RL-set is finite is not a practical limitation since humans are able to distinguish a limited number of restriction or precision levels and, in practice, the limit in precision and storage of computers allows us to work with a finite number of degrees (and consequently, of levels) only. In practice, the RL-set for an atomic property represented by a fuzzy set A on an universe X is

$$\Lambda_A = \{A(x) \mid x \in \mathit{support}(A)\} \cup \{1\} \tag{1}$$

The RL-set employed to represent a derived property is obtained as the union of the RL-sets of the atomic properties in terms of which the property is defined. Finally, the RL-representation of an imprecise property on X is defined in [3] as follows:

Definition 2.1 *A RL-representation is a pair* (Λ, ρ) *where* Λ *is a RL-set and* ρ *is a function*

$$\rho : \Lambda \to \mathcal{P}(X) \tag{2}$$

The function ρ indicates the crisp realization that represents the imprecise property for each restriction level. As an example, the RL-representation for an atomic imprecise property defined by a fuzzy set A is the pair (Λ_A, ρ_A), where Λ_A is obtained using equation (1), and $\rho_A(\alpha) = A_\alpha \ \forall \alpha \in \Lambda_A$.

Given an imprecise property P represented by (Λ_P, ρ_P), we define the set of crisp representatives of P, Ω_P, as

$$\Omega_P = \{\rho_P(\alpha) \mid \alpha \in \Lambda_P\} \tag{3}$$

For an atomic property A, the set of crisp representatives Ω_A is the set of significant α-cuts of A, as we have seen. However, notice that in definition 2.1 there is no restriction about the possible crisp representatives for non-atomic properties. In particular, as a consequence of operations, they don't need to be nested, so the final RL-representation of a derived property is not always equivalent to the α-cut representation of fuzzy sets.

In order to define properties by operations, it is convenient to extend the function ρ to any RL $\alpha \in (0,1]$. Let (Λ, ρ) be a RL-representation with $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verifying $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$. Let $\alpha \in (0,1]$ and $\alpha_i, \alpha_{i+1} \in \Lambda$ such that $\alpha_i \geq \alpha > \alpha_{i+1}$. Then

$$\rho(\alpha) = \rho(\alpha_i) \tag{4}$$

A RL-representation (Λ, ρ) on X is a crisp set $A \subseteq X$ iff $\forall \alpha \in \Lambda$, $\rho(\alpha) = A$.

Given a RL-representation (Λ_A, ρ_A) for an atomic property A, the values of Λ_A can be interpreted as values of possibility of a possibility measure defined $\forall \rho_A(\alpha_i) \in \Omega_A$ as $Pos(\rho_A(\alpha_i)) = \alpha_i$. Following this interpretation we define a basic probability assignment for each $Y \in \Omega_A$ and a fuzzy set μ_A representing an object-centered summary of the imprecise property A as

$$m(Y) = \sum_{\alpha_i \mid Y = \rho(\alpha_i)} \alpha_i - \alpha_{i+1} \tag{5}$$

$$\mu_A(x) = \sum_{Y \in \Omega_A \mid x \in Y} m_A(Y) \tag{6}$$

This is useful since it is easier to understand for those familiar with fuzzy sets, though information about how the representatives are distributed on restriction levels (and hence, crucial information) is lost. In other words, both representations are not equivalent in general. However, as a particular case, for any atomic property A and $\forall x \in X$

$$\mu_A(x) = A(x). \tag{7}$$

Operations on RL-representations are based on the following ideas:

1. There is a single membership scale, i.e., let F and G be fuzzy subsets of X, then $F(x) = G(x')$ means that x is F just like x' is G. Hence, it makes sense to consider the representation of a knowledge base in a certain restriction level as a knowledge base consisting of the representation in that RL of all the properties and concepts that form the KB.

2. Crisp operations are extended to imprecise properties by operating in each restriction level. In

general, let $f : \mathcal{P}(X)^n \to \mathcal{P}(X)$ be a crisp operation. Then, f is extended to RL-representations as follows: let $(P_1, \ldots, P_n)$ be imprecise properties defined on X with P_i represented by a RL-representation $(\Lambda_{P_i}, \rho_{P_i})$. Then, $f(P_1, \ldots, P_n)$ is represented by $(\Lambda_{f(P_1,\ldots,P_n)}, \rho_{f(P_1,\ldots,P_n)})$ where

$$\Lambda_{f(P_1,\ldots,P_n)} = \bigcup_{1 \leq i \leq n} \Lambda_{P_i} \tag{8}$$

and, $\forall \alpha \in \Lambda_{f(P_1,\ldots,P_n)}$,

$$\rho_{f(P_1,\ldots,P_n)}(\alpha) = f(\rho_{P_1}(\alpha), \ldots, \rho_{P_n}(\alpha)) \tag{9}$$

The representation in a certain RL of the extension of a certain crisp operation is the result of that operation on the representatives in the same RL of the arguments involved.

Let us remark that operations extended in this way verify the usual properties of their crisp counterparts [3]. As particular cases, $\rho_{\neg A}(\alpha) = \overline{\rho_A(\alpha)}$ $\forall \alpha \in \Lambda_A$, and $\rho_{A \wedge B}(\alpha) = \rho_A(\alpha) \cap \rho_B(\alpha)$, $\rho_{A \vee B}(\alpha) = \rho_A(\alpha) \cup \rho_B(\alpha)$, $\forall \alpha \in \Lambda_A \cup \Lambda_B$.

2.2 RL-NUMBERS

On the basis of RL-representations and operations, we introduced in [4] RL-numbers as a representation of imprecise quantities. This approach offers two main advantages:

- RL-numbers are representations of imprecise quantities that can be easily obtained by extending usual crisp measurements to fuzzy sets.
- Arithmetic and logical operations on RL-numbers are straightforward and unique extensions of the operations on crisp numbers, verifying the following:
  - They verify all the usual properties of crisp arithmetic and logical operations.
  - The imprecision does not necessarily increase through operations, and can even diminish. The maximum imprecision is related to the number of restriction levels employed.

RL-natural numbers are obtained as the possible cardinalities of RL-representations, the latter being the measurement of cardinality of the (crisp) set representative of the concept in each level. Then, and following the ideas about operations in section 2.1, arithmetic operations are extended by operating on each restriction level, and subsequently RL-integers, RL-rational numbers and RL-real numbers are defined. The following definitions and properties are from [4]:

Definition 2.2 *A RL-real number is a pair* $(\Lambda, \mathcal{R})$ *where* Λ *is a RL-set and* $\mathcal{R} : (0, 1] \to \mathbb{R}$.

We shall note $\mathbb{R}_{RL}$ the set of RL-real numbers. The RL-real number R_x is the representation of a (precise) real number x iff $\forall \alpha \in \Lambda_{R_x}$, $\mathcal{R}_{R_x}(\alpha) = x$. We shall denote such RL-real number as R_x or, equivalently, x, since in the crisp case, the set Λ_{R_x} is unimportant. Operations are extended as follows:

Definition 2.3 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *and let* $R_1 \ldots R_n$ *be RL-real numbers. Then* $f(R_1, \ldots, R_n)$ *is a RL-real number with*

$$\Lambda_{f(R_1,\ldots,R_n)} = \bigcup_{1 \leq i \leq n} \Lambda_{R_i} \tag{10}$$

and, $\forall \alpha \in \Lambda_{f(R_1,\ldots,R_n)}$

$$\mathcal{R}_{f(R_1,\ldots,R_n)}(\alpha) = f(\mathcal{R}_{R_1}(\alpha), \ldots, \mathcal{R}_{R_n}(\alpha)) \tag{11}$$

It is obvious that operations defined in this way are consistent extensions of crisp operations. It is important to remark that operations not defined for certain combinations of real values are not defined for RL-numbers that verify that combination in at least one restriction level. This is the case of division by 0 for example, i.e., R/R' is defined iff $0 \notin \Omega_{R'}$.

In order to define a ranking of RL-numbers, the notion or RL-ranking is introduced as follows:

Definition 2.4 *A RL-ranking is a pair* (Λ, δ) *where* Λ *is a RL-set and*

$$\delta : (0, 1] \to \{=, <, >\}$$

Notice that RL-rankings are imprecise rankings based on the possible rankings of the representatives of RL-numbers. This definition is based on the same ideas that the definitions of RL-numbers, and allows us to extend procedures based on rankings to restriction levels. On this basis, the notion of ranking of RL-numbers is as follows:

Definition 2.5 *The ranking of two real numbers* R_1 *and* R_2 *is a RL-ranking* $(\Lambda_{R_1,R_2}, \delta_{R_1,R_2})$ *with* $\Lambda_{R_1,R_2} = \Lambda_{R_1} \cup \Lambda_{R_2}$ *and* $\forall \alpha \in \Lambda_{R_1,R_2}$,

$$\delta_{R_1,R_2}(\alpha) = \begin{cases} = & \text{if } \mathcal{R}_{R_1}(\alpha) = \mathcal{R}_{R_2}(\alpha) \\ < & \text{if } \mathcal{R}_{R_1}(\alpha) < \mathcal{R}_{R_2}(\alpha) \\ > & \text{if } \mathcal{R}_{R_1}(\alpha) > \mathcal{R}_{R_2}(\alpha) \end{cases} \tag{12}$$

We say that $R_1 = R_2$ (resp. $R_1 < R_2$, $R_1 > R_2$) iff $\forall \alpha \in \Lambda_{R_1,R_2}$, $\mathcal{R}_{R_1}(\alpha) = \mathcal{R}_{R_2}(\alpha)$ (resp. $\mathcal{R}_{R_1}(\alpha) < \mathcal{R}_{R_2}(\alpha)$, $\mathcal{R}_{R_1}(\alpha) > \mathcal{R}_{R_2}(\alpha)$).

3 PROBABILITIES BASED ON RL-NUMBERS

Let us note $[0,1]_{RL} \subseteq \mathbb{R}_{RL}$ the set of RL-real numbers verifying $0 \leq R \leq 1\ \forall R \in [0,1]_{RL}$. We introduce the notion of RL-probability space as follows:

Definition 3.1 *A RL-probability space is a triple* (X, Σ, P) *where*

1. X *is a crisp set.*
2. Σ *is a collection of imprecise events defined by RL-representations on* X*, closed under complement and countable unions, and verifying* $X \in \Sigma$.
3. $P : \Sigma \to [0,1]_{RL}$ *verifying Kolmogorov's Axioms, i.e.,* $P(X) = 1$, $P(E) \geq 0\ \forall E \in \Sigma$, *and for any finite collection of disjoint representations* $E_1, \ldots, E_n$, $P(E_1 \cup \cdots \cup E_n) = \sum_{i=1}^{n} P(E_i)$.

In the previous definition, P is a RL-probability measure. In particular, RL-probability measures can be obtained from ordinary probability measures in an easy way as follows: let (U, F, p) be a (crisp) probability space and let U_F^{RL} be the set of all the RL-representations on U such that for any $A \in U_F^{RL}$, $\Omega_A \subseteq F$. By the properties of F, U_F^{RL} is closed under complement and countable unions, and $U \in U_F^{RL}$. For any $A \in U_F^{RL}$, let $P(A) \in [0,1]_{RL}$ defined by $\rho_{P(A)}(\alpha) = p(\rho_A(\alpha))\ \forall \alpha \in \Lambda_A$. It is easy to show that (U, U_F^{RL}, P) is a RL-probability space and P is a RL-probability measure.

Also easy to show, RL-probabilities verify the addition law $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and the property $P(\neg A) = 1 - P(A)$. Finally, by definition 3.1, it is obvious that crisp probabilities are a particular case of RL-probabilities since crisp events are particular cases of imprecise events as described by RL-representations.

The following example illustrates some of the concepts introduced in this section: Let A and B be atomic imprecise events defined by the following fuzzy sets on $U = \{u_1, \ldots, u_5\}$:

$$\begin{aligned} A &= 1/u_1 + 0.8/u_2 + 0.5/u_3 + 0.4/u_5 \\ B &= 0.9/u_1 + 0.6/u_3 + 0.5/u_4 \end{aligned}$$

Then we have $\Lambda_A = \Lambda_{\neg A} = \{1, 0.8, 0.5, 0.4\}$ and $\Lambda_B = \Lambda_{\neg B} = \{1, 0.9, 0.6, 0.5\}$. The RL-set for any operation between A and B is $\Lambda_A \cup \Lambda_B$. Table 1 shows the RL-representation of A, $\neg A$, B, $A \wedge B$, $A \vee B$, and $A \wedge \neg B$. Assuming that every u_i is equally probable, table 2 shows the corresponding RL-probabilities of the imprecise events $p(A)$, $p(\neg A)$, $p(B)$, $p(A \wedge B)$, $p(A \vee B)$, and $p(A \wedge \neg B)$. It is easy to check that the usual arithmetic relations between probabilities are preserved in each restriction level and, hence, they are preserved by the arithmetic of RL-numbers. For example, $p(A) + p(\neg A) = 1$.

The concept of conditional probability is straightforwardly extended as follows:

Definition 3.2 *Let* A *and* B *be two imprecise events defined on* U *by RL-representations* $A = (\Lambda_A, \rho_A)$ *and* $B = (\Lambda_B, \rho_B)$ *with* $\rho_{p(B)} > 0\ \forall \alpha \in (0,1]$. *The RL-probability of* A *given* B *in* U *is*

$$p(A|B) = \frac{p(A \wedge B)}{p(B)} \tag{13}$$

Proposition 3.1 *Let* A *and* B *be two imprecise events defined on* U *by RL-representations* $A = (\Lambda_A, \rho_A)$ *and* $B = (\Lambda_B, \rho_B)$ *with* $\rho_{p(B)} > 0\ \forall \alpha \in (0,1]$. *Then*

$$\rho_{p(A|B)}(\alpha) = \frac{\rho_{p(A \wedge B)}(\alpha)}{\rho_{p(B)}(\alpha)} \tag{14}$$

Proof: Immediate by definition of quotient of RL-numbers. □

As an example, the RL-probability $p(B|A)$ is shown in the last column of table 2. The RL-probability $p(A|B)$ is undefined since $\rho_{p(B)}(1) = 0$.

Definition 3.3 *Let* A *and* B *be two imprecise events defined on* U *by RL-representations* $A = (\Lambda_A, \rho_A)$ *and* $B = (\Lambda_B, \rho_B)$. *We say* A *and* B *are* independent *in* U *iff* $p(A \wedge B) = p(A)p(B)$.

4 STATISTICAL TESTS BASED ON RL-PROBABILITIES

Statistical tests are performed on a data sample. The starting point of our proposal is to consider that the available data is imprecise in the sense that it is described by a RL-representation. This includes RL-numbers as values of numerical variables, and imprecise events defined by RL-representations (in particular, fuzzy sets), among others.

In order to extend a certain statistical test to a sample of imprecise data, we follow the same basic idea employed for extending crisp operations to RL-representations: working on each level separately, considering as the set of levels the union of all the sets of levels employed to describe the data. The main ideas of the approach are:

- The crisp test is performed in each restriction level. Since each level corresponds to a crisp ver-

α	$\rho_A(\alpha)$	$\rho_{\neg A}(\alpha)$	$\rho_B(\alpha)$	$\rho_{A\wedge B}(\alpha)$	$\rho_{A\vee B}(\alpha)$	$\rho_{A\wedge\neg B}(\alpha)$
1	$\{u_1\}$	$\{u_2,u_3,u_4,u_5\}$	$\emptyset$	$\emptyset$	$\{u_1\}$	$\{u_1\}$
0.9	$\{u_1\}$	$\{u_2,u_3,u_4,u_5\}$	$\{u_1\}$	$\{u_1\}$	$\{u_1\}$	$\emptyset$
0.8	$\{u_1,u_2\}$	$\{u_3,u_4,u_5\}$	$\{u_1\}$	$\{u_1\}$	$\{u_1,u_2\}$	$\{u_2\}$
0.6	$\{u_1,u_2,u_3\}$	$\{u_4,u_5\}$	$\{u_1,u_3\}$	$\{u_1,u_3\}$	$\{u_1,u_2,u_3\}$	$\{u_2\}$
0.5	$\{u_1,u_2,u_3\}$	$\{u_4,u_5\}$	$\{u_1,u_3,u_4\}$	$\{u_1,u_3\}$	$\{u_1,u_2,u_3,u_4\}$	$\{u_2\}$
0.4	$\{u_1,u_2,u_3,u_5\}$	$\{u_4\}$	$\{u_1,u_3,u_4\}$	$\{u_1,u_3\}$	U	$\{u_2,u_5\}$

Table 1: RL-representation of several imprecise events derived from two atomic properties A and B.

α	$\rho_{p(A)}(\alpha)$	$\rho_{p(\neg A)}(\alpha)$	$\rho_{p(B)}(\alpha)$	$\rho_{p(A\wedge B)}(\alpha)$	$\rho_{p(A\vee B)}(\alpha)$	$\rho_{p(A\wedge\neg B)}(\alpha)$	$\rho_{p(B\|A)}$
1	1/5	4/5	0	0	1/5	1/5	0
0.9	1/5	4/5	1/5	1/5	1/5	0	1
0.8	2/5	3/5	1/5	1/5	2/5	1/5	1/2
0.6	3/5	2/5	2/5	2/5	3/5	1/5	2/3
0.5	3/5	2/5	3/5	2/5	4/5	1/5	2/3
0.4	4/5	1/5	3/5	2/5	1	2/5	1/2

Table 2: RL-probabilities in $U = \{u_1, \ldots, u_5\}$ of the imprecise events in table 1 and the conditional RL-probability $p(B|A)$.

sion of the data, the test in each level is an ordinary test on crisp data.

- The usual notions of sample, probability distribution, statistic, etc. are extended to RL-sample, RL-probability distribution, RL- statistic, etc. by levels. As an example, a RL-sample is an assignment of samples to levels.

- The parameters can be the same or different in each level; in the latter case, the notion of RL-parameter value appears as an assignment of different crisp parameter values to each level.

- In each level we obtain a crisp result (reject or not reject). As a whole, the result is a RL- decision, i.e., an assignment of decisions to each restriction level. Notice that the RL-decision can be a crisp decision even for imprecise data (when the result is the same in all the levels).

- In order to provide a more informative result, we propose to summarize the RL- decision by calculating the evidence supporting each crisp decision. This is obtained as the basic probability assignment of the RL-decision. The RL- decision is totally imprecise if the evidence is of the form 0.5/reject + 0.5/not reject; otherwise, there is more evidence in favor of one of the results.

In order to illustrate our approach, let us consider that A and B are two imprecise events and $U = \{u_1, \ldots, u_5\}$ is a sample of data. The fulfilment of A and B in U is that of table 1. We want to perform a chi-square independence test between A and B. Of course, this sample is too small to perform a reliable test, so we have repeated the data five times in order to obtain a larger sample with the same frequencies in each restriction level. Table 3 shows the contingency table in each level (RL-contingency table). Notice that the addition of rows and columns yield the correct marginal probabilities since we are using RL-representation and negation instead of fuzzy negation.

The chi-square test rejected the null hypothesis of independence in levels 0.9, 0.8 and 0.4 only. In the rest of levels, the null hypothesis was not rejected. If this result is going to be employed for further calculations, we should proceed operating in each level. On the contrary, this result can be summarized and offered to an user in terms of evidence as follows: the basic probability assignment to restriction levels 0.9, 0.8 and 0.4 is 0.1, 0.2 and 0.4, respectively. Hence, the evidence supporting the reject decision is 0.7, whilst the evidence against rejecting is 0.3.

5 CONCLUSIONS

Probability is a very important concept that is on the basis of statistics. In order to extend statistics to the case of imprecise information, probabilities of imprecise events are crucial. RL-probabilities are able to represent consistently the (imprecise) probability of imprecise events. The basic idea, coming from the notion of RL-representation, is that the usual crisp representations of properties, operations and measures are employed in a set of restriction levels. When applied to atomic properties represented by fuzzy sets,

$\alpha = 1$	B	$\neg B$
A	0	1/5
$\neg A$	0	4/5

$\alpha = 0.9$	B	$\neg B$
A	1/5	0
$\neg A$	0	4/5

$\alpha = 0.8$	B	$\neg B$
A	1/5	1/5
$\neg A$	0	3/5

$\alpha = 0.6$	B	$\neg B$
A	2/5	1/5
$\neg A$	0	2/5

$\alpha = 0.5$	B	$\neg B$
A	2/5	1/5
$\neg A$	1/5	1/5

$\alpha = 0.4$	B	$\neg B$
A	2/5	2/5
$\neg A$	1/5	0

Table 3: RL-contingency table of events A and B on U as given in table 1.

the crisp representation in each level of a property is an α-cut of the fuzzy set. Crisp measures like cardinality are applied on every restriction level, providing a crisp value of the measure in each level. The set of measures for all the levels is a RL-value. RL-numbers and RL-probabilities are examples of this approach. Operations are also applied in each level. Let us remark that RL-representations correspond to fuzzy sets only for atomic properties in general.

A RL-number associates a single number to each restriction level. In [4] we discuss about the relationship between RL-numbers and fuzzy numbers. In our view, both are complementary. Fuzzy numbers correspond to RL-intervals and define concepts as restrictions, whilst RL-numbers correspond to measures on imprecise properties. This discussion has been presented in [2] as well. Fuzzy numbers are better in order to translate the result of an operation to an user, since they are easier to understand; on the contrary, RL-numbers are more difficult to understand, but are better representations and allow us to extend arithmetic operations in an easy way and, more important, preserving the ordinary properties of arithmetic. In addition, arithmetic operations on RL-numbers does not necessarily increases the imprecision of the results (the result can even be crisp!); the maximum imprecision is determined by the number of restriction levels considered and its associated degree. In [4] we propose to use RL-numbers and operate in each restriction level until a final result must be given to the user, and then to approximate the RL-number by a RL-interval corresponding to a fuzzy number.

RL-numbers and probabilities have been also introduced in [2] under the name *gradual numbers*. Both approaches start from a very similar point of view, with slight differences that affect mainly to the way negation is considered in [2] (negation is coherent with fuzzy negation, i.e., gradual representations are completely equivalent to fuzzy sets both for atomic and derived properties). However, for the representation of imprecise quantities, this makes no difference. On the other hand, the fuzzy set corresponding to a "gradual representation" is obtained by a max-sup operation on the set of representatives (called *gradual elements*), whilst our "fuzzy summary" is obtained on the basis of a basic probability assignment, yielding a different result in general. This in fact affects the way imprecise quantities are approximated by linguistic terms. We shall study this in detail in a future work.

An interesting feature of RL-probabilities for the representation of the probability of imprecise events is that imprecision and uncertainty are clearly separated. The imprecision is represented in the number of different crisp representatives of the property that appear in the representation. In particular, a property represented by the same crisp set in all the restriction levels is a property that can be represented by a single restriction level and, hence, is a crisp property. On its turn, probabilistic uncertainty can be represented in each restriction level.

Finally, following the same ideas underlying the definition of RL-objects, we propose an approach to statistical testing on imprecise data. The usual mathematical objects involved in a test are extended to their corresponding RL counterparts as an assignment of objects to restriction levels. This way, we consider the notions of RL-sample, RL-probability distribution, RL-statistic, RL-parameter, RL-decision, etc. The test is performed in each level, obtaining a RL-decision, that can be later summarized by the addition of the evidence supporting each crisp decision (reject or not reject) in the set of levels, via the basic probability assignment associated to the levels. Though we are losing information, this summary is easier to understand.

References

[1] James J. Buckley. *Fuzzy Probability and Statistics*, volume 196 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin Heidelberg, 2006.

[2] D. Dubois and H. Prade. Gradual elements in a fuzzy set. *Soft Computing*, 12:165–175, 2008.

[3] D. Sánchez, M. Delgado, and M.A. Vila. A restriction level approach to the representation of imprecise properties. In *IPMU 2008*, Accepted for publication.

[4] D. Sánchez, M. Delgado, and M.A. Vila. RL-numbers: An alternative to fuzzy numbers for the representation of imprecise quantities. In *WCCI 2008*, Accepted for publication.

UN MODELO PARA MEDIR LA N-CONTRADICCIÓN DE CONJUNTOS BORROSOS*

Elena E. Castiñeira Carmen Torres-Blanc Susana Cubillo

Dpto. Matemática Aplicada, Facultad de Informática, Universidad Politécnica de Madrid
28660, Boadilla del Monte, Madrid
{ecastineira, ctorres, scubillo}@fi.upm.es

Resumen

En este artículo, se revisan los conceptos de conjuntos borrosos autocontradictorios y contradictorios respecto de una negación, extendiéndolos al marco, más general, de las negaciones menos restrictivas, en lugar de considerar sólo negaciones fuertes. A continuación, se establecen los axiomas mínimos que debe de verificar una función para que pueda ser considerada una medida de N-autocontradicción, y análogamente para una medida de N-contradicción, estableciéndose algunas relaciones entre ambos tipos de medidas. Finalmente, se introducen los axiomas para modelar la continuidad de las medidas cuando los conjuntos borrosos "crecen", y análogamente, cuando "decrecen"; asimismo se proporcionan sendas familias de medidas de contradicción, de ambos tipos, que satisfacen una u otra propiedad.

Palabras Clave: Contradicción, medidas de contradicción, negaciones, medidas semicontinuas.

1 INTRODUCCIÓN

El estudio de la contradicción tiene interés no sólo desde el punto de vista teórico, sino también desde el práctico, dada la importancia que puede alcanzar la obtención de consecuencias contradictorias en los procesos de inferencia. Estas preocupaciones llevaron a los autores de [11] y [12] a introducir los conceptos de conjuntos borrosos autocontradictorios y contradictorios. Más tarde, en [2, 3, 8, 9] se propusieron formas de medir el grado en que esta contradicción se produce. En [5] se plantearon los conceptos de medida de autocontradicción y medida de contradicción de forma axiomática. Aunque en dicho artículo se comenzaron a desarrollar las ideas principales que creemos que involucra este concepto, la profundización de su estudio nos ha llevado a una revisión del modelo allí propuesto, separando claramente los axiomas que modelizan el concepto de medida de contradicción (incluso variando alguno de ellos) con los que modelizan la continuidad, añadiendo nuevos axiomas y, finalmente, ampliando el marco de estudio, de forma que se extienda el modelo a medidas, no sólo de autocontradicción y contradicción, sino también a medidas donde la contradicción depende de una negación N, esto es, medidas de N-autocontradicción y N-contradicción.

Los conceptos de medida de N-autocontradicción y medida de N-contradicción, aunque fueron introducidos en [4] allí apenas se estudiaron. En el presente trabajo se tratan en mayor profundidad y en un marco más general.

Finalizamos esta introducción comentando la estructura de este artículo. En la siguiente sección, sección 2, se revisan los conceptos de conjuntos borrosos autocontradictorios y contradictorios respecto de una negación, extendiéndolos al marco, más general, de las funciones de negación menos restrictivas, en lugar de considerar sólo negaciones fuertes como se venía haciendo hasta ahora. A continuación, se establecen los axiomas mínimos que debe de verificar una función para que pueda entenderse como medida de N-autocontradicción, después se hace lo propio para una medida de N-contradicción, y finalmente, se determina alguna relación entre ambos tipos de medidas. En la sección tercera, se introducen los axiomas que pueden modelar cierto tipo de continuidad, concretamente, la continuidad cuando los conjunto borrosos "crecen", semicontinuidad desde abajo, y cuando los conjuntos "decrecen", semicontinuidad desde arriba. Este tipo de continuidad se ha dado en llamar completa, para diferenciarla de otros tipos que se pueden definir imponiendo condiciones a los conjuntos borro-

* Este trabajo está parcialmente subvencionado por el Proyecto TIN 2005-08943-C02-01 (CICYT).

sos, pero aquí no se tratarán pues es el tópico que las autoras está investigando en la actualidad. En todos los casos se proporcionan familias de medidas verificando estas propiedades.

2 CONTRADICCIÓN Y MEDIDAS DE CONTRADICCIÓN RESPECTO A UNA NEGACIÓN

Puesto que el estudio que se va a realizar utiliza negaciones, comenzamos recordando este concepto, así como algunas propiedades que puede verificar.

Definición 2.1 [6, 10, 13] Sea la función $N : [0,1] \to [0,1]$, y consideremos las siguientes propiedades que puede verificar:

(n1) $N(0) = 1$ y $N(1) = 0$.

(n2) N es decreciente.

(n3) N es estrictamente decreciente y continua.

(n4) $a \leq N^2(a) = N(N(a))$ para todo $a \in [0,1]$.

(n5) $N^2(a) \leq a$ para todo $a \in [0,1]$.

Se dice que N es: una *función de negación* o simplemente una *negación* si verifica los axiomas (n1) y (n2); una *negación estricta* si verifica (n1) y (n3); una *negación ordinaria* si verifica (n1), (n2) y (n4); una *negación débil* si verifica (n1), (n2) y (n5); y es una *negación fuerte o involución* si verifica (n1), (n2), (n4) y (n5).

Dado que una involución es biyectiva, entonces también es una negación estricta. Por otra parte, existen negaciones que son estrictas y no son ordinarias ni débiles. Por ejemplo, la función $N : [0,1] \to [0,1]$ definida, para cada $a \in [0,1]$, por

$$N(a) = \begin{cases} 1 - \dfrac{3}{2}\,a, & \text{si } a \leq \dfrac{1}{2}; \\ \dfrac{1-a}{2}\,a, & \text{si } a > \dfrac{1}{2}; \end{cases}$$

es una negación estricta, no ordinaria ni débil.

Denotaremos por $N_\wedge, N_\vee : [0,1] \to [0,1]$ las negaciones que presentan los casos extremos, definidas por

$$N_\wedge(a) = \begin{cases} 1, & \text{si } a = 0; \\ 0, & \text{si } a \neq 0; \end{cases} \qquad N_\vee(a) = \begin{cases} 1, & \text{si } a \neq 1; \\ 0, & \text{si } a = 1. \end{cases}$$

Se verifica que $N_\wedge$ es una negación ordinaria y $N_\vee$ es una negación débil.

2.1 CONJUNTOS N-AUTOCONTRADICTORIOS Y CONJUNTOS N-CONTRADICTORIOS

Definición 2.2 Sea un universo $X \neq \emptyset$ y sea N una negación. Se dice que un conjunto borroso sobre X, o alternativamente su función de pertenencia $\mu \in [0,1]^X$, es *N-autocontradictorio* o *autocontradictorio respecto a N*, si para todo $x \in X$ se verifica que $\mu(x) \leq N(\mu(x))$.

Observación 2.3 Sean $X \neq \emptyset$ y $\mu \in [0,1]^X$. Se tiene:

1) Si N_1, N_2 son negaciones verificando que $N_1 \leq N_2$, entonces si μ es N_1-autocontradictorio también es N_2-autocontradictorio.

2) μ es $N_\wedge$-autocontradictorio si y sólo si $\mu = \mu_\emptyset$, donde $\mu_\emptyset$ representa la función de pertenencia del conjunto vacío ($\mu_\emptyset(x) = 0,\ \forall x \in X$).

3) μ es $N_\vee$-autocontradictorio si y sólo si $\mu(x) < 1$ $\forall x \in X$. Es decir, los únicos conjuntos borrosos que no son $N_\vee$-autocontradictorios son los *normalizados*.

4) Si N es una negación continua, μ es N-autocontradictorio si y sólo si $\operatorname{Sup}_{x \in X} \mu(x) \leq \alpha_N$, donde α_N es el punto fijo de la negación.

Esta observación sugiere la siguiente definición un poco más restrictiva, que permitirá recuperar las definiciones de [11, 12] cuando el marco de trabajo a considerar se restringe al conjunto de negaciones fuertes.

Definición 2.4 Dado $X \neq \emptyset$, se dice que un conjunto borroso sobre X, o alternativamente su función de pertenencia $\mu \in [0,1]^X$, es *autocontradictorio* si verifica que $\operatorname{Sup}_{x \in X} \mu(x) < 1$.

Si $\mu \in [0,1]^X$ es autocontradictorio, entonces existen infinitas negaciones respecto de las cuales es autocontradictorio. En efecto, basta considerar, para cualquier α tal que $\operatorname{Sup}_{x \in X} \mu(x) \leq \alpha < 1$, una negación continua N_α con punto fijo α.

Definición 2.5 Sean $X \neq \emptyset$ y N una negación. Dados dos conjuntos borrosos cuyas funciones de pertenencia son $\mu, \sigma \in [0,1]^X$, se dice que μ *es N-contradictorio con σ* si $\mu(x) \leq N(\sigma(x))$ para todo $x \in X$. Si μ es N-contradictorio con σ y σ es N-contradictorio con μ, entonces se dice que son *N-contradictorios*.

Observación 2.6 Sean $X \neq \emptyset$ y $\mu, \sigma \in [0,1]^X$:

1) Si N es una negación ordinaria y $\mu, \sigma \in [0,1]^X$, se verifica que μ es N-contradictorio con σ si y sólo si μ y σ son N-contradictorios.

2) Si N_1, N_2 son negaciones tales que $N_1 \leq N_2$, se verifica que si μ es N_1-contradictorio con σ también es N_2-contradictorio con σ.

3) μ es $N_\wedge$-contradictorio con σ si y sólo si para cada $x \in X$ se verifica que $\mu(x) = 0$ ó $\sigma(x) = 0$, es decir, $\text{Min}(\mu, \sigma) = \mu_\emptyset$.

4) μ es $N_\vee$-contradictorio con σ si y sólo si para cada $x \in X$ se verifica que $\mu(x) = 0$ ó $\sigma(x) < 1$.

5) Si N es una negación fuerte y φ es el automorfismo de orden que la genera (es decir, φ es una biyección creciente del intervalo $[0,1]$ tal que $N(a) = \varphi^{-1}(1 - \varphi(a))$, $\forall a \in [0,1]$, véase [10]), entonces μ y σ son N-contradictorios si y sólo si $\varphi(\mu(x)) + \varphi(\sigma(x)) \leq 1$ para todo $x \in X$.

Definición 2.7 Sea N una función de negación. Si $X \neq \emptyset$, se dice que un conjunto borroso sobre X, o alternativamente su función de pertenencia $\mu \in [0,1]^X$, es *N-normal* si verifica que $N\left(\sup_{x \in X} \mu(x)\right) \leq \sup_{x \in X} \mu(x)$.

Dado que es bastante usual denotar $I = [0,1]$, llamaremos $I_N^X = \{\mu \in [0,1]^X \mid \mu \text{ es } N\text{-normal}\}$

Observación 2.8 Sean $X \neq \emptyset$ y $\mu \in [0,1]^X$:

1) Si N_1, N_2 son negaciones verificando que $N_1 \leq N_2$, entonces $I_{N_2}^X \subset I_{N_1}^X$.

2) $\mu \in I_{N_\wedge}^X$ si y sólo si $\mu \neq \mu_\emptyset$, lo que es equivalente a que $\sup_{x \in X} \mu(x) > 0$.

3) $\mu \in I_{N_\vee}^X$ si y sólo si $\sup_{x \in X} \mu(x) = 1$, es decir, si μ es normal.

4) Si N es una negación continua, $\mu \in I_N^X$ si y sólo si $\sup_{x \in X} \mu(x) \geq \alpha_N$, donde α_N es el punto fijo de N.

2.2 MEDIDAS DE N-AUTOCONTRADICCIÓN: $\mathcal{A}_N(X)$

En el marco de la teoría de los conjuntos borrosos existen infinidad de conjuntos autocontradictorios, pero en la teoría de conjuntos clásica, el único conjunto autocontradictorio es el vacío; por lo que cualquier función susceptible de medir contradicción debería tomar el valor máximo sobre el conjunto vacío (axioma (ai) de la siguiente definición) y el mínimo sobre cualquier conjunto clásico no vacío. Este último requisito se cumple si además exigimos, como es razonable, que todo conjunto no N-autocontradictorio tenga esa medida mínima. Para formular esta exigencia, teniendo en cuenta que μ no es N-autocontradictorio si existe un elemento x tal que $\mu(x) > N(\mu(x))$, en [4] se consideró que si $\mu(x) \geq N(\mu(x))$ para algún x, la autocontradicción de μ debería ser 0, lo que se puede justificar de la siguiente manera: si existe algún x para el que $\mu(x) - N(\mu(x)) > 0$, por muy pequeño que sea este valor, la medida de μ debe ser 0, por lo que también es razonable considerar que sea 0 si $\mu(x) - N(\mu(x)) = 0$. Aun así, en la definición que aquí se presenta, por consideraciones similares, se modifica la exigencia de que exista un x tal que $\mu(x) \geq N(\mu(x))$ imponiendo que esta desigualdad la cumpla el $\sup_{x \in X} \mu(x)$ (axioma (aii)). Finalmente, también será necesario establecer, de algún modo, cómo debe variar una medida de N-autocontradicción en función de los conjuntos. Para ello, observemos que si μ, σ son N-autocontradictorios y $\mu \leq \sigma$, entonces $\mu \leq \sigma \leq N \circ \sigma \leq N \circ \mu$, por lo que resulta evidente que σ está más "cerca" de ser no N-autocontradictorio que μ, por tanto, su medida debería ser menor que la de μ. Así, se tiene:

Definición 2.9 Sean $X \neq \emptyset$ y N una negación. Se llama *medida de N-autocontradicción* sobre los conjuntos borrosos de X, $\mathcal{F}(X)$, a cualquier aplicación $A_N : [0,1]^X \to [0,1]$ que verifica las siguientes condiciones:

(ai) $A_N(\mu_\emptyset) = 1$.

(aii) Si $\mu \in I_N^X$, entonces $A_N(\mu) = 0$.

(aiii) Antimonotonía: si $\mu, \sigma \in [0,1]^X$ verifican que $\mu \leq \sigma$, entonces $A_N(\sigma) \leq A_N(\mu)$.

Denotaremos al conjunto de medidas de N-autocontradicción sobre los conjuntos borrosos de X por $\mathcal{A}_N(X)$.

Obsérvese que si cambiamos el axioma (aii) por otro (aii') donde se imponga la condición de que la medida sea 0 para los conjuntos normales, entonces se mediría cuán autocontradictorio es un conjunto borroso, que también coincide con cuán $N_\vee$-autocontradictorio es; por ello, a tales aplicaciones las denominaremos simplemente *medidas de autocontradicción*. En realidad, esta definición fue introducida en [5] como medida de autocontradicción débil, pero aquí no se utiliza tal denominación, pues se consideran aparte los axiomas que modelizan la continuidad, como ya ha sido comentado.

Observación 2.10 Sean $X \neq \emptyset$ y N_1, N_2 negaciones, si $N_1 \leq N_2$, entonces $\mathcal{A}_{N_1}(X) \subset \mathcal{A}_{N_2}(X)$.

Ejemplo 2.11 Sea $X \neq \emptyset$, entonces:

1) Para toda negación N se puede considerar la siguiente medida de N-autocontradicción trivial, definida para cada $\mu \in [0,1]^X$ por

$$A_0^N(\mu) = \begin{cases} 0, & \text{si } \mu \text{ es } N\text{-normal;} \\ 1, & \text{en otro caso.} \end{cases}$$

Obviamente, para cualquier $A_N \in \mathcal{A}_N(X)$ se verifica que $A_0^{N_\wedge} \leq A_N \leq A_0^{N_\vee}$. Por otra parte, $A_0^{N_\wedge}$ es medida de N-autocontradicción para toda negación N.

2) La aplicación definida para cada $\mu \in [0,1]^X$ por $A(\mu) = 1 - \underset{x\in X}{\text{Sup}}\,\mu(x)$ es una medida de $N_\vee$-autocontradicción o medida de autocontradicción.

3) Para cada negación continua N, las aplicaciones definidas para cada $\mu \in [0,1]^X$ por

$$A_1^N(\mu) = \frac{1}{\alpha_N}\,\text{Max}\left(0, \alpha_N - \underset{x\in X}{\text{Sup}}\,\mu(x)\right),$$

$$A_2^N(\mu) = N\left(\text{Min}\left(1, \frac{1}{\alpha_N}\,\underset{x\in X}{\text{Sup}}\,\mu(x)\right)\right),$$

donde α_N es el punto fijo de N, son medidas de N-autocontradicción.

2.3 MEDIDAS DE N-CONTRADICCIÓN: $\mathcal{C}_N(X)$

En este apartado establecemos los requisitos mínimos que una función debe satisfacer para que modele la contradicción, respecto de una negación dada N, entre dos conjuntos borrosos; es decir, para medir cuán N-contradictorios son. Como el conjunto vacío, "el menor" de todos los conjuntos, está contenido en su complementario, que es el "mayor" de todos, la N-contradicción del conjunto vacío consigo mismo debe ser la mayor de todas. Por otra parte, como queremos medir con qué grado se satisfacen a la vez las desigualdades $\mu \le N \circ \sigma$ (i.e. μ es N-contradictorio con σ) y $\sigma \le N \circ \mu$ (i.e. σ es N-contradictorio con μ), obviamente debemos de exigir simetría. Además, razonando de forma similar al caso del axioma (aii) para medidas de N-autocontradicción, se justifica que impongamos que, si $N(\underset{x\in X}{\text{Sup}}\,\mu(x)) \le \underset{x\in X}{\text{Sup}}\,\mu(x)$ ó $N(\underset{x\in X}{\text{Sup}}\,\sigma(x)) \le \underset{x\in X}{\text{Sup}}\,\sigma(x)$, la N-contradicción entre ambos sea 0. Finalmente, del mismo modo que para las medidas de N-autocontradicción se justificaría la antimonotonía. Así, se tiene:

Definición 2.12 Sean $X \neq \emptyset$ y N una negación. Se llama *medida de N-contradicción* sobre $\mathcal{F}(X)$, a cualquier aplicación $C_N : [0,1]^X \times [0,1]^X \to [0,1]$ que verifica las siguientes condiciones:

(ci) $C_N(\mu_\emptyset, \mu_\emptyset) = 1$.

(cii) Dados $\mu, \sigma \in [0,1]^X$, si $\mu \in I_N^X$ ó $\sigma \in I_N^X$, entonces $C_N(\mu, \sigma) = 0$.

(ciii) Simetría: $C_N(\mu, \sigma) = C_N(\sigma, \mu)$, $\forall \mu, \sigma \in [0,1]^X$.

(civ) Antimonotonía: si $\mu, \sigma \in [0,1]^X$ tales que $\mu \le \sigma$, entonces $C_N(\sigma, \nu) \le C_N(\mu, \nu)$, $\forall\, \nu \in [0,1]^X$.

Denotaremos al conjunto de medidas de N-contradicción sobre los conjuntos borrosos de X por $\mathcal{C}_N(X)$.

Al igual que sucedía con la autocontradicción, si C es una medida de $N_\vee$-contradicción, diremos que es una *medida de contradicción*.

Observación 2.13 Si N_1, N_2 son negaciones tales que $N_1 \le N_2$ entonces $\mathcal{C}_{N_1}(X) \subset \mathcal{C}_{N_2}(X)$.

Ejemplo 2.14 Sea $X \neq \emptyset$, se tiene:

1) Para toda negación N se puede considerar la siguiente medida de N-contradicción trivial, definida para cada $(\mu, \sigma) \in [0,1]^X \times [0,1]^X$ por

$$C_0^N(\mu, \sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ 1, & \text{en otro caso.} \end{cases}$$

Obviamente, para cualquier $C_N \in \mathcal{C}_N(X)$ se verifica que $C_0^{N_\wedge} \le C_N \le C_0^{N_\vee}$. Además $C_0^{N_\wedge}$ es medida de N-contradicción para toda negación N.

2) La aplicación definida para cada $(\mu, \sigma) \in [0,1]^X \times [0,1]^X$ por

$$C(\mu, \sigma) = \begin{cases} 0, & \text{si } \mu \text{ ó } \sigma \text{ normal;} \\ 1 - \frac{1}{2}\,\underset{x\in X}{\text{Sup}}(\mu(x) + \sigma(x)), & \text{en otro caso.} \end{cases}$$

es una medida de contradicción.

3) Para cada negación continua N, sea α_N su punto fijo, entonces las aplicaciones definidas para cada $(\mu, \sigma) \in [0,1]^X \times [0,1]^X$ por

$$C_1^N(\mu, \sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ \text{Max}\left(0, \frac{\alpha_N - \frac{1}{2}\,\underset{x\in X}{\text{Sup}}(\mu(x)+\sigma(x))}{\alpha_N}\right), & \text{e.o.c.} \end{cases}$$

$$C_2^N(\mu, \sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ N\left(\text{Min}\left(1, \frac{\underset{x\in X}{\text{Sup}}(\mu(x)+\sigma(x))}{2\alpha_N}\right)\right), & \text{e.o.c.} \end{cases}$$

son medidas de N-contradicción.

2.4 ALGUNAS RELACIONES ENTRE $\mathcal{A}_N(X)$ Y $\mathcal{C}_N(X)$

Finalizamos esta sección estableciendo dos resultados que relacionan las medidas de N-autocontradicción y de N-contradicción. Antes recordemos que una t-norma es una operación binaria en $[0,1]$ asociativa, conmutativa, monótona respecto al orden usual en los números reales y con 1 como elemento neutro (por ejemplo, véase [1] o [7]).

Proposición 2.15 Sean $X \neq \emptyset$ y N una negación. Si $C_N \in \mathcal{C}_N(X)$, entonces la aplicación $A_N : [0,1]^X \to [0,1]$, definida para cada $\mu \in [0,1]^X$ por $A_N(\mu) = C_N(\mu, \mu)$, es una medida de N-autocontradicción.

Proposición 2.16 Sean $X \neq \emptyset$ y N una negación. Si $A_1, A_2 \in \mathcal{A}_N(X)$, entonces para toda t-norma T , la aplicación $C_N : [0,1]^X \times [0,1]^X \to [0,1]$, definida para cada $(\mu, \sigma) \in [0,1]^X \times [0,1]^X$ por

$$C_N(\mu, \sigma) = T(A_1(\mu), A_2(\sigma)),$$

es una medida de N-contradicción.

3 MEDIDAS COMPLETAMENTE SEMICONTINUAS

Los axiomas de medidas de contradicción, respecto de una negación, no garantizan que las funciones que los verifiquen tomen sus valores de forma gradual, como se puede apreciar en cada medida A_0^N ó C_0^N. Por tanto, si se quiere exigir algún tipo de continuidad a las medidas de contradicción, habrá que introducir nuevos axiomas para modelizarla.

Para definir los conceptos de continuidad se considera el retículo acotado y completo $([0,1]^X, \vee, \wedge)$, donde las operaciones binarias $\vee$ y $\wedge$ se definen, para cada $\mu, \sigma \in [0,1]^X$, como los elementos $\mu \vee \sigma, \mu \wedge \sigma \in [0,1]^X$ dados para cada $x \in X$ por $(\mu \vee \sigma)(x) = \mathrm{Sup}(\mu(x), \sigma(x))$ y $(\mu \wedge \sigma)(x) = \mathrm{Inf}(\mu(x), \sigma(x))$, respectivamente.

3.1 MEDIDAS DE N-AUTOCONTRADICCIÓN COMPLETAMENTE SEMICONTINUAS

Definición 3.1 Sean $X \neq \emptyset$ y N una negación. Si $A \in \mathcal{A}_N(X)$, entonces:

1) A es *completamente $\vee$-semicontinua*, o *completamente semicontinua desde abajo*, si para todo $\{\mu_i\}_{i\in\mathcal{I}} \subset [0,1]^X$, donde $\mathcal{I}$ es un conjunto arbitrario de índices, se verifica que

$$\operatorname*{Inf}_{i\in\mathcal{I}} A(\mu_i) = A\left(\bigvee_{i\in\mathcal{I}} \mu_i\right), \qquad \text{(aiv)}$$

2) A es *completamente $\wedge$-semicontinua*, o *completamente semicontinua desde arriba*, si para todo $\{\mu_i\}_{i\in\mathcal{I}} \subset [0,1]^X \setminus I_N^X$

$$\operatorname*{Sup}_{i\in\mathcal{I}} A(\mu_i) = A\left(\bigwedge_{i\in\mathcal{I}} \mu_i\right), \qquad \text{(av)}$$

Denotaremos al conjunto de las medidas de N-autocontradicción completamente $\vee$-semicontinuas por $\mathcal{A}_N^{\vee csc}(X)$, y al de las completamente $\wedge$-semicontinuas por $\mathcal{A}_N^{\wedge csc}(X)$.

Observación 3.2 Nótese que la ecuación (aiv) deben verificarla todas las familias de conjuntos borrosos, mientras que (av) deben verificarlos los conjuntos borrosos que no son N-normales. La razón de esta asimetría está en que no puede existir ninguna medida de N-autocontradicción, con $N \neq N_\vee$, verificando (av) para cualquier familia de conjuntos siempre que el universo tenga más de un elemento. En efecto, sean $x_0 \in X$ y $A \in \mathcal{A}_N(X)$; consideremos

$$\mu(x) = \begin{cases} \alpha, & \text{si } x = x_0; \\ 0, & \text{si } x \neq x_0; \end{cases} \quad \sigma(x) = \begin{cases} 0, & \text{si } x = x_0; \\ \alpha, & \text{si } x \neq x_0, \end{cases}$$

siendo $\alpha \in (0,1)$ tal que $N(\alpha) < \alpha$. Entonces μ y σ son ambos N-normales, por lo que $A(\mu) = A(\sigma) = 0$. Sin embargo, $A(\mu \wedge \sigma) = 1$ dado que $\mu \wedge \sigma = \mu_\emptyset$, así $\mathrm{Sup}\{A(\mu), A(\sigma)\} < A(\mu \wedge \sigma)$.

Proposición 3.3 Sean $X \neq \emptyset$ y N una negación continua con punto fijo α_N. Consideremos para cada $p \in (0, \alpha_N]$ las siguientes funciones definidas, para cada $\mu \in [0,1]^X$, por:

$$A_1^p(\mu) = \frac{1}{p} \operatorname{Max}\left(0, p - \operatorname*{Sup}_{x\in X} \mu(x)\right),$$

$$A_2^p(\mu) = N\left(\operatorname{Min}\left(1, \frac{1}{p} \operatorname*{Sup}_{x\in X} \mu(x)\right)\right).$$

Se verifica que $\{A_1^p, A_2^p\}_{p\in(0,\alpha_N]} \subset \mathcal{A}_N^{\vee csc}(X)$ y $A_1^p, A_2^p \notin \mathcal{A}_N^{\wedge csc}(X),\ \forall p \in (0, \alpha_N]$.

Proposición 3.4 Sean $X \neq \emptyset$ y N una negación continua con punto fijo α_N. Consideremos para cada $p \in (0, \alpha_N]$ las siguientes funciones definidas, para cada $\mu \in [0,1]^X$, por:

$$\tilde{A}_1^p(\mu) = \begin{cases} 0, & \text{si } \mu \in I_N^X \\ \frac{1}{p} \operatorname{Max}\left(0, p - \operatorname*{Inf}_{x\in X} \mu(x)\right), & \text{e.o.c.} \end{cases}$$

$$\tilde{A}_2^p(\mu) = \begin{cases} 0, & \text{si } \mu \in I_N^X \\ \frac{1}{p} N\left(\operatorname{Min}\left(1, \frac{1}{p} \operatorname*{Inf}_{x\in X} \mu(x)\right)\right), & \text{e.o.c.} \end{cases}$$

Se verifica que $\{\tilde{A}_1^p, \tilde{A}_2^p\}_{p\in(0,\alpha_N]} \subset \mathcal{A}_N^{\wedge csc}(X)$ y $\tilde{A}_1^p, \tilde{A}_2^p \notin \mathcal{A}_N^{\vee csc}(X),\ \forall p \in (0, \alpha_N]$.

3.2 MEDIDAS DE N-CONTRADICCIÓN COMPLETAMENTE SEMICONTINUAS

Definición 3.5 Sean $X \neq \emptyset$ y N una negación. Si $C \in \mathcal{C}_N(X)$, entonces:

1) C es *completamente $\vee$-semicontinua*, o *completamente semicontinua desde abajo*, si para todo $\mu \in [0,1]^X$ y para toda familia $\{\mu_i\}_{i\in\mathcal{I}} \subset [0,1]^X$,

$$\operatorname*{Inf}_{i\in\mathcal{I}} C(\mu_i, \mu) = C\left(\bigvee_{i\in\mathcal{I}} \mu_i, \mu\right), \qquad \text{(civ)}$$

2) C es *completamente $\wedge$-semicontinua*, o *completamente semicontinua desde arriba*, si para todo $\mu \in [0,1]^X$ y para toda familia $\{\mu_i\}_{i\in\mathcal{I}} \subset [0,1]^X \setminus I_N^X$,

$$\operatorname*{Sup}_{i\in\mathcal{I}} C(\mu_i,\mu) = C\left(\bigwedge_{i\in\mathcal{I}}\mu_i,\mu\right), \qquad \text{(cv)}$$

Denotaremos al conjunto de las medidas de N-contradicción que son completamente $\vee$-semicontinuas por $\mathcal{C}_N^{\vee csc}(X)$, y al de las que son completamente $\wedge$-semicontinuas por $\mathcal{C}_N^{\wedge csc}(X)$.

La exclusión de los conjuntos N-normales en la parte 2) de esta definición se justifican de manera similar al caso de las medidas de N-autocontradicción.

Proposición 3.6 Sean $X \neq \emptyset$ y N una negación continua con punto fijo α_N. Consideremos para cada $p \in (0,\alpha_N]$ las siguientes funciones definidas, para cada $(\mu,\sigma) \in [0,1]^X \times [0,1]^X$, por:

$$C_1^p(\mu,\sigma) = \operatorname{Max}\left(0, \operatorname{Min}\left(\frac{p - \operatorname*{Sup}_{x\in X}\mu(x)}{p}, \frac{p - \operatorname*{Sup}_{x\in X}\sigma(x)}{p}\right)\right)$$

$$C_2^p(\mu,\sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ N\left(\operatorname{Min}\left(1, \frac{\operatorname*{Sup}_{x\in X}(\mu(x)+\sigma(x))}{2p}\right)\right), & \text{e.o.c.} \end{cases}$$

Se verifica que $\{C_1^p, C_2^p\}_{p\in(0,\alpha_N]} \subset \mathcal{C}_N^{\vee csc}(X)$ y $C_1^p, C_2^p \notin \mathcal{C}_N^{\wedge csc}(X)$, $\forall p \in (0,\alpha_N]$.

Proposición 3.7 Sean $X \neq \emptyset$ y N una negación continua con punto fijo α_N. Consideremos para cada $p \in (0,\alpha_N]$ las siguientes funciones definidas, para cada $(\mu,\sigma) \in [0,1]^X \times [0,1]^X$, por:

$$\tilde{C}_1^p(\mu,\sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ \operatorname{Max}\left(0, \operatorname{Max}\left(\frac{p - \operatorname*{Inf}_{x\in X}\mu(x)}{p}, \frac{p - \operatorname*{Inf}_{x\in X}\sigma(x)}{p}\right)\right) & \text{e.o.c.} \end{cases}$$

$$\tilde{C}_2^p(\mu,\sigma) = \begin{cases} 0, & \text{si } \mu \in I_N^X \text{ ó } \sigma \in I_N^X; \\ N\left(\operatorname{Min}\left(1, \frac{\operatorname*{Inf}_{x\in X}(\mu(x)+\sigma(x))}{2p}\right)\right), & \text{e.o.c.} \end{cases}$$

Se verifica que $\{\tilde{C}_1^p, \tilde{C}_2^p\}_{p\in(0,\alpha_N]} \subset \mathcal{C}_N^{\wedge csc}(X)$ y $\tilde{C}_1^p, \tilde{C}_2^p \notin \mathcal{C}_N^{\vee csc}(X)$, $\forall p \in (0,\alpha_N]$.

CONCLUSIONES

En este artículo se ha establecido y estudiado un modelo axiomático para medir la contradicción de conjuntos borrosos respecto a un negación dada. Aunque ya había sido hecho una propuesta en [5] para medir la contradicción, no se trataba ésta respecto a una negación fija. Por lo tanto, el modelo aquí propuesto es más amplio y además recupera al anterior, con la ventaja de que proporciona el marco adecuado para el estudio de la contradicción para cada teoría borrosa dada. Por otra parte, se ha mejorado sustancialmente la parte del modelo dedicada a la continuidad, ya que hasta ahora sólo se había considerado y estudiado la continuidad desde una sola vertiente, atendiendo al "crecimiento" de los conjuntos borrosos, mientras que aquí se ha introducido y tratado también el caso en que los conjuntos "decrecen".

Desde un punto de vista estrictamente teórico, entre otros, señalemos los siguientes temas de interés como futuras líneas de investigación: el establecimiento y estudio de diferentes tipos de continuidad; la exploración de métodos que permitan la agregación de las medidas de contradicción, así como las propiedades que son estables para dichos mecanismos; la relación con otras medidas borrosas, en particular, con las medidas de incompatibilidad; la búsqueda de métodos generales para la construcción de medidas, siendo relevante investigar el uso de uninormas en el caso de que los conjuntos borrosos estén definidos sobre universos finitos. Por otra parte, desde el punto de vista de la aplicación, estas medidas ofrecen un gran potencial como herramienta, siendo de interés investigar su uso para la evaluación de los resultados en los procesos de inferencia borrosos, o en algoritmos de clasificación adaptativa, a fin de optimizar los resultados.

Referencias

[1] C. Alsina, M. Frank, B. Schweizer. *Associative Functions: Triangular Norms and Copulas*, World Scientific (Singapore, Thailand), 2006.

[2] E. Castiñeira, S. Cubillo, S. Bellido. Degrees of Contradiction in Fuzzy Sets Theory, *Proceedings of IX Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Annecy (France), Pág. 171-176, 2002.

[3] E. Castiñeira, S. Cubillo, S. Bellido. Contradicción entre dos conjuntos, *Actas del XI Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)*, León, Pág. 379-383, 2002.

[4] E. Castiñeira, S. Cubillo, Pedro L. Fernández. Medidas de incompatibilidad entre conjuntos borrosos, *Actas del XIII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF)*, Ciudad Real, Pág. 65-70, 2006.

[5] S. Cubillo, E. Castiñeira. Measuring contradiction in fuzzy logic. *International Journal of General Systems*, **34** (1), Pág. 39-59, 2005.

[6] F. Esteva, X. Domingo. Sobre funciones de negación en [0,1]. *Stochastica*, **IV** (2), Pág. 141-165, 1980.

[7] E.P. Klement, R. Mesiar, E. Pap. *Triangular Norms*, Kluwer Academic Publishers (Dordrecht, Netherlands), 2000.

[8] C. Torres, E. Castiñeira, S. Cubillo, V. Zarzosa. A Geometrical Interpretation to define contradiction degrees between two fuzzy sets Ref. *Int. J. Information Theories and Applications*, **12**(2), Pág. 127-134, 2005.

[9] C. Torres, S. Cubillo, E. Castiñeira. Contradiction versus self-contradiction in Fuzzy Logic. *Int. J. Information Theories and Applications*, **14**(4), Pág. 331-338, 2007.

[10] E. Trillas. Sobre funciones de negación en la teoría de conjuntos difusos. *Stochastica*, **III** (1), Pág. 47-60, 1979. Reimpreso (version en inglés) en *Advances of Fuzzy Logic* (eds. S. Barro et altri-Universidad de Santiago de Compostela), Pág. 31-43, 1998.

[11] E. Trillas, C. Alsina, J. Jacas. On Contradiction in Fuzzy Logic. *Soft Computing*, **3** (4), Pág. 197-199, 1999.

[12] E. Trillas, S. Cubillo. On non-contradicictory Input/Ouput couples in Zadeh's CRI. *Proceedings of the 18th International Meeting of the North American Fuzzy Information Processing Society (NAFIPS)*, Pág. 28-32, New York, 1999.

[13] S. Weber. A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. *Fuzzy Sets and Systems*, **11**, Pág. 115-134, 1983

Reasoning about Fuzzy Functional Dependencies*

Pablo Cordero Angel Mora Manuel Enciso Gabriel Aguilera I.P. de Guzmán

E.T.S.I. Informática. Universidad de Málaga.
{pcordero,enciso}@uma.es,{amora,gabri,guzman}@ctima.uma.es,

Abstract

In literature, there exists several logics to specify and manipulate functional dependencies (FDs) and fuzzy FDs (FFDs). Nevertheless, their inference systems are suitable to illustrate FD and FFD semantics but they may not be used as a formal base to develop automated deduction methods. In this work, we use the correct and complete axiomatic system $\mathcal{S}^e$, with a new FFD inference rule, named *Substitution Rule*. This rule allows us to introduce the *Deduction Theorem for FFD* and is the key for solving the FFD implication problem.

Keywords: fuzzy functional dependencies, logic, implication problem.

1 Introduction

Constraints are often used to guide the design of noteworthy relational schema for the sake of the database consistence and therefore to avoid the problems of redundances, anomalies,etc. This statement is valid for any extension of the classical relational model. During three consecutive decades, Raju and Majumdar [19], Cubero and Vila [6], Tyagi et.al. [23],etc. have studied the fuzzy models and which constraints are more appropriated to extend the well studied relational database theory to their fuzzy database.

Several approaches to the definition of fuzzy functional dependency [6, 18, 19, 22, 23] (FFD) are proposed in the literature. In the same way as the concept of functional dependency (FD) correspond to the notion of partial function, it should be desired that the concept of FFD would correspond to the notion of fuzzy partial function. The definitions proposed in [19, 22, 23] fit in this idea.

There exists a wide range of dependencies. Some of them were investigated in the past (FDs [4, 9], Multivalued Dependencies [10], etc.) and others are still being studied today (XML FDs [13], FFDs [22], etc.). Moreover, each dependency definition is usually followed by its corresponding logic. In [19, 23] the authors propose Armstrong's axioms as a useful tool for reasoning with FFD, but these inferences rule has not been used successfully in automated deduction. The reason is that this inference system were created to explain dependency semantics more than to design an automated deduction system. In fact, in [23] the authors propose the classical closure algorithm to solve the implication problem and don't directly use Armstrong's axioms.

In [5] a novel logic (**SL**) equivalent to classical Armstrong's axioms was presented. The core of **SL** is the *Substitution Rule* [1]. The definition of **SL** introduces, for the first time, interesting solutions to database problems, which are solved using logic-based automated deduction methods [1, 15, 16].

In this work, we illustrate how the *Substitution Rule* can be considered to reasoning with *FFDs*. First, we extend the language of FFD Logic to allow empty left hand side formulae and we use this new well formed formulae as *goals* to be satisfied in a given FFD theory. The substitution rule is used to build a novel substitution algorithm directly based on the inference system. We emphasize that we solve the implication problem for FFDs reducing the set of FFDs with the substitution rule.

This work opens the door to the management of FFD constraints in an efficient and intelligent way and is organized as follows: Section 2 introduces some defi-

*Partially supported by Spanish DGI project TIN2007-65819 and Junta de Andalucia project TIC115.

[1] The primitive inference system doesn't contain the transitivity rule and is a rule derived in **SL** (see [5]).

nitions of FFDs. In section 3 we show how previous FD logics reason about FFDs and the limitations of these logics to manipulate automatically a set of FFDs. We introduce the problem that we solve in this paper, named "*the FFD implication problem*". In section 4 we propose a new automated deduction method to solve the FFD implication problem and finally we establish several conclusions and future works in section 5.

2 Fuzzy Functional Dependencies

Since the 80's, some authors have extended the Relational Model to include imprecise/fuzzy information [6, 18, 19, 23]. A fuzzy extension of the Codd Relational Model also require to consider an extension of the integrity constraints that may involve fuzzy concepts. In classical relational databases the study of the integrity constraints have been developed successfully in the past. The FD is considered the first step of this studies, since the definition of FD play an outstanding role in the Relational Model. The FD is very close to the notion of *key* and it is used in other areas like knowledge discovery, query optimization, database design, etc.

The classical FDs specify an association between the values of the attributes of a relation which allows to include, in a formal way, constraints like the following: *if two employees have the same qualification then they have the same salary.* As J. Paredaens et al. say in [17]: "they are constraint in the description of a database in order to ensure that the instances we might obtain are meaningful". Thus, if the Functional Dependency $X \mapsto Y$ holds in a relation R, where X, Y are subset of attributes, then for any instance of R if two tuples agree on X, then they also agree on Y.

When a fuzzy extension of the notion of FD are introduced, some authors consider that the data becomes fuzzy while others consider that the dependency may be imprecise. In [19] the author introduces a definition of FFD to model fuzzy information such as *employees with similar experience must have similar salaries*, which may be consider a precise dependency, while in [22] they propose a FFD which allows to model expressions such as *the intelligence level of a person more or less determines the degree of success.*

Each definition of FFD treats a particular aspect of the fuzziness. We emphasize the definition of [19] because the introduction of an operator (a fuzzy resemblance measure EQ for comparing domain values) made possible a generalization of the FDs and to the their definition of FFDs: A FFD $X \rightsquigarrow Y$ holds in a fuzzy relation r, if for all tuples t_1 and t_2 of r, we have that

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \mu_{EQ}(t_1[Y], t_2[Y])$$

where μ is the membership function of the fuzzy relation [2].

A great variety of FFD definitions can be found in the literature. In [23] the authors summarize some of them and study the fuzziness and imprecision degree, both in the data and in the dependency of the FFD definitions. This study shows a definition [22] which has a high degree of fuzziness and it is a natural extension of the classical FD concept:

$X \rightarrow_F^{\Theta} Y$, holds in a fuzzy relation if, for every pair of tuples t_1 and t_2 then $Conf((t_1, t_2)[X]) \geqslant min(\Theta, Conf((t1, t2)[Y]), \Theta \in [0, 1]$.

If $\Theta = 0$, there is no valid dependency and if $\Theta = 1$ then imprecise FFD become precise FFD; otherwise, imprecise FFDs exist. Here $Conf(t_1, t_2)[X]$ is the degree of closeness between two tuples t_1 and t_2 projected over attributes X in a fuzzy relation instance and is called the conformance.

When $\Theta = 1$ then, the definition of [22] generalizes the one presented in [19]. In this work, we deal with this second definition. An extension of our study to consider more general definitions will be made in a future work.

3 Reasoning about FFDs

Several axiomatic systems has been published [9, 17] that are equivalent to Amstrong's axiomatic system. These logics may be considered as formal tools to formally explain how to deduce a dependency from a given set of dependencies. However, in [11, 12], the authors state the limitations to manage dependencies using logic, and they claim for efficient automatic computation methods. On the other hand, in [12, 14, 20], indirect methods for the manipulation of dependencies are used as an alternative.

Specifically, we are interested in solving the implication problem by using directly the logic. Unfortunately, all the classical axiomatic systems are not suitable tools to develop automated deduction techniques, because all of them are based on the transitive rule.

Instead of that, in literature, there are several algorithms to solve the implication problem in polynomial time (see [3, 7] for further details). Nevertheless, these efficient methods have a very important disadvantage: they do not allow to give an explanation about the answer. When we use an *indirect* method we are not

[2] μ denotes the *more o less equal* relation.

able to translate the final solution into an inference chain to explain the answer in terms of the inference system. This limits the use of these indirect methods in artificial intelligence environments.

In [5] we introduce a new logic **SL** equivalent to the classical logics (its axiomatic system is equivalent to the Armstrong's system), but with an important difference: the transitive rule has been replaced by a novel rule. This is the unique axiomatic system in the literature that have not the transitive rule as primitive rule. The novel rule named substitution rule was introduced with the idea to remove redundancy in sets of dependencies. This characteristic make this logic suitable to be used for directly reasoning. Other interesting characteristic of this rule is that can be considered as a transformation of *equivalence* [15].

In [15] we proved that **SL** axiomatic system is equivalent to other well known axiomatic systems [3, 9, 12, 17] and thus, all Paredaens derived rules are derived rules in **SL**.

To solve the implication problem we present an extension of **SL** logic that we will name $\mathbf{SL}^e$ logic. This basically consists in to extend the language of FFD Logic to allow empty left hand side formulae:

Definition 3.1 *The* $\mathbf{SL}^e$ *logic is the pair* $(\mathcal{L}^e, \mathcal{S})$ *where* $\mathcal{L}^e = \{X \rightsquigarrow Y \mid X, Y \in 2^\Omega\}$ *and* Ω *is a set of attributes. The axiomatic system* $\mathcal{S}$ *consists in:*

Axiom scheme*: if* $Y \subseteq X \neq \varnothing$ *then*

$$\vdash_{\mathcal{S}} X \rightsquigarrow Y$$

Fragmentation rule*: if* $Y' \subseteq Y$ *then*

$$X \rightsquigarrow Y \vdash_{\mathcal{S}} X \rightsquigarrow Y'$$

Composition rule*:*

$$X \rightsquigarrow Y,\ U \rightsquigarrow V \vdash_{\mathcal{S}} XU \rightsquigarrow YV$$

Substitution rule*: if* $X \subseteq U$ *and* $X \cap Y = \varnothing$ *then*

$$X \rightsquigarrow Y,\ U \rightsquigarrow V \vdash_{\mathcal{S}} (U\text{-}Y) \rightsquigarrow (V\text{-}Y)$$

From now on, to use the standard notation used by database community, $\top$ will denote the empty set of attributes and XY the union of attributes.

The following theorem ensures the correctness and completeness of the method presented in the next section.

Theorem 3.2 (Deduction Theorem for FFD) [3] *Given* $\Gamma \subseteq \mathcal{L}$*, we have the following equivalence:*

1. *For all* $X, U, V \in 2^\Omega$*,*

$$\Gamma \cup \{\top \rightsquigarrow X\} \vdash_{\mathcal{S}^e} U \rightsquigarrow V \text{ iff } \Gamma \vdash_{\mathcal{S}} UX \rightsquigarrow V$$

And, in particular:

2. *For all* $X, Y \in 2^\Omega$*. the following equivalence is stated:*

$$\Gamma \cup \{\top \rightsquigarrow X\} \vdash_{\mathcal{S}^e} \top \rightsquigarrow Y \text{ iff } \Gamma \vdash_{\mathcal{S}} X \rightsquigarrow Y$$

4 A new automated deduction method

Now we are interested in the benefits of the extension that we have just consider.

In this section, a novel technique for applying in a systematic way the system $\mathbf{SL}^e$is introduced. With this aim, three rewriting rules of simplification are defined using the symbol $\rightarrowtail$ where $\Gamma \rightarrowtail \Gamma'$ means that all the elements in Γ must be replaced by all the elements in Γ'.

Definition 4.1 *Given* $X, U, V \in 2^\Omega$*.*

SC *Simplification: If* $U \subseteq X$ *then* $\{\top \rightsquigarrow X, U \rightsquigarrow V\} \rightarrowtail \{\top \rightsquigarrow XV\}$

SA *Simplification: If* $V \subseteq X$ *then* $\{\top \rightsquigarrow X, U \rightsquigarrow V\} \rightarrowtail \{\top \rightsquigarrow X\}$

S *Simplification:* $\{\top \rightsquigarrow X, U \mapsto V\} \rightarrowtail \{\top \rightsquigarrow X, U - X \rightsquigarrow V - X\}$

Lemma 4.2 *Let* Γ *and* Γ' *be two sets of FFDs. If* Γ' *is obtained from* Γ *applying the rewriting rules of simplification introduced in definition 4.1 then*

$$\Gamma \equiv_{\mathcal{S}^e} \Gamma'$$

Proof.

SC Simplification:

$$\{\top \rightsquigarrow X, U \rightsquigarrow V\} \overset{1}{\equiv}_{\mathcal{S}^e} \{\top \rightsquigarrow X, U - X \rightsquigarrow V - X\}$$

$$\overset{2}{\equiv}_{\mathcal{S}^e} \{\top \rightsquigarrow XV\}$$

Substitution rule is applied in 1 and, since $U \subseteq X$, *Composition* rule is applied in 2.

SA Simplification:

$$\{\top \rightsquigarrow X, U \rightsquigarrow V\} \overset{3}{\equiv}_{\mathcal{S}^e} \{\top \rightsquigarrow X, U - X \rightsquigarrow V - X\}$$

$$\overset{4}{\equiv}_{\mathcal{S}^e} \{\top \rightsquigarrow X\}$$

Substitution rule is applied in 3 and, since $V \subseteq X$, the *Axiom* is applied in 4.

[3] See the proof of this theorem in [8].

$\top\rightsquigarrow X$	*Simp.Rule*	$ad\rightsquigarrow c$	$b\rightsquigarrow e$	$be\rightsquigarrow cg$	$bc\rightsquigarrow g$	$c\rightsquigarrow a$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow bd}$	S	$\underline{ad\rightsquigarrow c}$	$b\rightsquigarrow e$	$be\rightsquigarrow cg$	$bc\rightsquigarrow g$	$c\rightsquigarrow a$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow bd}$	SC	$a\rightsquigarrow c$	$\underline{b\rightsquigarrow e}$	$be\rightsquigarrow cg$	$bc\rightsquigarrow g$	$c\rightsquigarrow a$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow bde}$	SC	$a\rightsquigarrow c$		$\underline{be\rightsquigarrow cg}$	$bc\rightsquigarrow g$	$c\rightsquigarrow a$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow bcdeg}$	SA	$a\rightsquigarrow c$			$\underline{bc\rightsquigarrow g}$	$c\rightsquigarrow a$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow bcdeg}$	SC	$a\rightsquigarrow c$				$\underline{c\rightsquigarrow a}$	$cd\rightsquigarrow b$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow abcdeg}$	SA	$a\rightsquigarrow c$					$\underline{cd\rightsquigarrow b}$	$cf\rightsquigarrow bh$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow abcdeg}$	S	$a\rightsquigarrow c$						$\underline{cf\rightsquigarrow bh}$	$cg\rightsquigarrow af$
$\underline{\top\rightsquigarrow abcdeg}$	SC	$a\rightsquigarrow c$						$f\rightsquigarrow h$	$\underline{cg\rightsquigarrow af}$
$\underline{\top\rightsquigarrow abcdefg}$	SA	$\underline{a\rightsquigarrow c}$						$f\rightsquigarrow h$	
$\underline{\top\rightsquigarrow abcdefg}$	SC							$\underline{f\rightsquigarrow h}$	
$\underline{\top\rightsquigarrow abcdefgh}$									

Figure 1: Table of the example 1

S Simplification:

$$\{\top\rightsquigarrow X, U\rightsquigarrow V\} \overset{5}{\equiv}_{\mathcal{S}^e} \{\top\rightsquigarrow X, U-X\rightsquigarrow V-X\}$$

Substitution rule is applied in 5.

□

Theorem 4.3 *Given $\Gamma \subseteq \mathcal{L}$ and $X\rightsquigarrow Y \in \mathcal{L}$. If Γ' is obtained from $\Gamma \cup \{\top\rightsquigarrow X\}$ applying the rewriting rules of simplification introduced in definition 4.1 while these rules can be applied then there exists a unique $\top\rightsquigarrow Z \in \Gamma'$ with $X \subseteq Z$ and*

$$\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y \quad \textit{if and only if} \quad Y \subseteq Z$$

Proof. Firstly, there exist $\top\rightsquigarrow Z \in \Gamma'$ with $X \subseteq Z$ because we apply the rewriting rules to $\Gamma \cup \{\top\rightsquigarrow X\}$ and, when these rules modify X, X increases.

The uniqueness of $\top\rightsquigarrow Z$ is ensured by **SC** rule.

$Y \subseteq Z$ implies $\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y$ is obtained using Theorem 3.2, Lemma 4.2 and the fragmentation rule.

Conversely, the following steps prove that $\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y$ implies that $Y \subseteq Z$. Let Γ'' be $\Gamma' - \{\top\rightsquigarrow Z\}$:

1. Since $\top\rightsquigarrow Z$ is unique, $\Gamma'' \subseteq \mathcal{L}$.
2. If $\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y$ then Theorem 3.2 ensures that $\Gamma' \vdash_{\mathcal{S}^e} \top\rightsquigarrow Y$ and, from $\Gamma' = \{\top\rightsquigarrow Z\} \cup \Gamma''$, (1) and Theorem 3.2, $\Gamma'' \vdash_{\mathcal{S}} Z\rightsquigarrow Y$ is obtained.
3. If $U\rightsquigarrow V \in \Gamma''$ then $U \cap Z = \varnothing$ and $V \cap Z = \varnothing$, since otherwise **S** rule of simplification could be applied.
4. If $\Gamma'' \vdash_{\mathcal{S}} Z\rightsquigarrow Y$ then $Y \subseteq Z$ because, due to (3), $Z\rightsquigarrow Y$ must be an axiom.

□

The above theorem states the method to determinate if $\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y$. The solution arose from adding the goal $\top\rightsquigarrow X$ to Γ, rendering an initial Γ'. Then, rewriting rules of simplification are applied to Γ' obtaining $\{\top\rightsquigarrow Z\} \cup \Gamma''$. Finally, $\Gamma \vdash_{\mathcal{S}} X\rightsquigarrow Y$ if and only if $Y \subseteq Z$.

Below, we solve the implication problem using our new methodology:

Example 4.1 *Let $\Gamma = \{ad\rightsquigarrow c, b\rightsquigarrow eh,\ be\rightsquigarrow c, bc\rightsquigarrow d,\ c\rightsquigarrow a,\ cd\rightsquigarrow b,\ ce\rightsquigarrow af,\ cf\rightsquigarrow bdh\}$ be this set of FFDs.*

In order to know whether $\Gamma \vdash bd\rightsquigarrow ah$, firstly we initialize $\Gamma' = \Gamma \cup \{\top\rightsquigarrow bd\}$ rendering: $\Gamma' = \{\top\rightsquigarrow bd, ad\rightsquigarrow c, b\rightsquigarrow eh, be\rightsquigarrow c, bc\rightsquigarrow d, c\rightsquigarrow a, cd\rightsquigarrow b, ce\rightsquigarrow af, cf\rightsquigarrow bdh\}$

The table in figure 1 shows step by step how the rewriting rules of simplification are applied. Note that the underscore points the FFD that is being reduced. The second column shows the applied rule:

Since $ah \subseteq abcdefgh$ and by Theorem 4.3, the following deduction is obtained:

$$\Gamma \models bd\rightsquigarrow ah$$

A novel algorithm for solving the implication problem using rules of simplification defined below is shown in figure *2*. The algorithm simply adds $\top\rightsquigarrow X$ and, in an exhaustive way applies the rules of simplification based on the theoretical study (Theorem 3.2).

$$\texttt{Implies?}(\Gamma, X \to Y) = \begin{cases} Yes, & \text{if } Y \subseteq \texttt{Closure}(X, \texttt{nil}, \Gamma, 1); \\ No, & \text{otherwise.} \end{cases}$$

$$\texttt{Closure}(X, \Gamma_1, \Gamma_2, b) = \begin{cases} X, & \text{if } \Gamma_2 = \texttt{nil} \text{ or } b = 0; \\ \texttt{Closure}(\texttt{Simplify}(X, \Gamma_1, \Gamma_2, 0)), & \text{otherwise.} \end{cases}$$

$$\texttt{Simplify}(X, \Gamma_1, \texttt{nil}, b) = (X, \texttt{nil}, \Gamma_1, b)$$

$$\texttt{Simplify}(X, \Gamma_1, U \to V :: \Gamma_2, b) = \\ = \begin{cases} \texttt{Simplify}(X, \Gamma_1, \Gamma_2, b), & \text{if } V \subseteq X; \\ \texttt{Simplify}(XV, \Gamma_1, \Gamma_2, 1), & \text{if } U \subseteq X \text{ and } V \not\subseteq X; \\ \texttt{Simplify}(X, U\text{-}X \to V\text{-}X :: \Gamma_1, \Gamma_2, b), & \text{otherwise.} \end{cases}$$

Figure 2: Algorithm to solve implication problem

Since every step adds at least one attribute, in the worst case, the "Closure" loop is repeated at most $|\mathcal{A}|$ times. The "Simplify" loop is repeated at most $|\Gamma|$ times. Consequently, the complexity of the algorithm is $O(|\mathcal{A}||\Gamma|)$.

We emphasize the following characteristics of the algorithm :

- The algorithm has the same complexity as the previous algorithms [7, 17] cited in literature, namely linear with regard to the input.

- Contrary to these previous algorithms, our algorithm has a solid base, since it uses the **SL** logic. Consequently proofs and explanations are given automatically by the algorithm applying directly the logic **SL**. Namely, the trace shown in column *Simp.Rule* reflects the rules of simplification based on **SL** logic that must be applied in order to prove the implication and the order in which the rules need to be applied.

5 Conclusions and future work

We have illustrated the difficulties of directly using other previous logics for dependencies to face up to the implication problem. We have shown how Substitution rules allow the development of automatic deduction methods.

None of the classical logics for dependencies can solve the implication problem efficiently without using indirect methods. The algorithm we have proposed in this paper has the same complexity as typical indirect methods but using directly FFD **SL**. Thus, we can reason and we may built explanations. So, this new algorithm is more appropriate to be used in an artificial intelligence environment.

The most efficient algorithms that solve the implication problem were based on complex indirect techniques instead of using theoretical advantages. The use of these indirect techniques obscures the nature of the method. On the contrary, the structure of our algorithm is notably simple.

In a future work, we will use **SL** to solve some interesting problems in the fuzzy databases, like removing redundancy, normalization, canonical closure, etc.

We will also study the extension of the technique we have shown in this work to more general definitions of FFDs, for example considering fuzziness not only in the data but in the dependency itself [22].

References

[1] Gabriel Aguilera, Pablo Cordero, Manuel Enciso, Angel Mora, and Inmaculada P. de Guzmán. A non-explosive treatment of functional dependencies using rewriting logic. *Lecture Notes in Artificial Intelligence*, 3171: 31–40, 2004.

[2] William W. Armstrong. Dependency structures of data base relationships. *Proc. IFIP Congress. North Holland, Amsterdam*, pages 580–583, 1974.

[3] Paolo Atzeni and Valeria De Antonellis. Relational Database Theory. *The Benjamin/Cummings Publishing Company Inc.*, 1993.

[4] Edgar F. Codd. Recent investigations into relational data base systems. *IFIP Congress, Estocolmo, Suecia*, 1974.

[5] Pablo Cordero, Manuel Enciso, Inmaculada P. de Guzmán, and Angel Mora. Slfd logic: Elimination of data redundancy in knowledge representation. *Lecture Notes in Artificial Intelligence 2527, Springer-Verlag*, pages 141–150, 2002.

[6] J.C. Cubero, M.A. Vila. A new definition of fuzzy functional dependency in fuzzy relational databases. *Internat. J. Intell. Systems*, 9 (5): 441–448, 1994.

[7] Jim Diederich and Jack Milton. New methods and fast algorithms for database normalization. *ACM Transactions on Database Systems*, 13 (3):339–365, 1988.

[8] Manuel Enciso, Gabriel Aguilera, Pablo Cordero, Angel Mora, and Inmaculada P. de Guzmán. Logic-based Functional Dependencies Programming. *Technical Reports. Departamento de Matemática Aplica. Universidad de Málaga. http://enciso.lcc.uma.es/TechnicalReports/ip.pdf* , 2008

[9] Ronald Fagin. Functional dependencies in a relational database and propositional logic. *IBM. Journal of research and development*, 21 (6):534–544, 1977.

[10] Ronald Fagin. Multivalued dependencies and a new normal form for relational databases. *ACM TODS 2*, 1977.

[11] J. W. Guan and D. A. Bell. Rough computational methods for information systems. *Artificial Intelligence*, 105 1 (2):77–103, 1998.

[12] Toshihide Ibaraki, Alexander Kogan, and Kazuhisa Makino. Functional dependencies in horn theories. *Artificial Intelligence*, 108 1-2:1–30, 1999.

[13] M. L. Lee, T. W. Ling, and W. L. Low. Designing functional dependencies for xml. *Lecture Notes in Computer Science, Springer-Verlag*, 2287:124–141, 2002.

[14] Heikki Mannila and Kari-Jouko Raiha. Algorithms for inferring functional dependencies from relations. *Data and Knowledge Engineering*, 12 (1):83–99, 1994.

[15] Angel Mora, Manuel Enciso, Pablo Cordero, and Inmaculada P. de Guzmán. The functional dependence implication problem: optimality and minimality. an efficient preprocessing transformation based on the substitution paradigm. *Lecture Notes in Artificial Intelligence, Springer-Verlag*, 3040, 2004.

[16] Angel Mora, Gabriel Aguilera, Manuel Enciso, Pablo Cordero, and Inmaculada P. de Guzmán. A new closure algorithm based in logic: SLFD-Closure versus classical closures. *Inteligencia Artificial, Revista Iberoamericana de IA*, 31 (10): 31–40, 2006

[17] Jan Paredaens, Paul De Bra, Marc Gyssens, and Dirk Van Van Gucht. The structure of the relational database model. *EATCS Monographs on Theoretical Computer Science*, 1989.

[18] H. Prade, C. Testemale. Generalizing database realtional algebra for the treatment of incomplete or uncertain information and vague queries. *Information Sciences*, 34: 115–143, 1984.

[19] H.V.S.V.N. Raju, A.K. Mazumdar. Fuzzy dependencies and losless join descomposition of fuzzy relational database systems. *ACM Transaction of Database Systems*, 13 (2): 129–166, 1988.

[20] Iztok Savnik and Peter A. Flach. Bottom-up induction of functional dependencies from relations. *Proc. of AAAI-93 Workshop: Knowledge Discovery in Databases*, pages 174–185, 1993.

[21] E. Sciore. A complete axiomatization of full join dependencies. *J. ACM*, 29 (2):373–393, 1982.

[22] Mustafa Ilker Sozat and Adnan Yazici. A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. *Fuzzy Sets and Systems*, 117 (2):161–181, 2001.

[23] B.K. Tyagi, A. Sharfuddin, R.N. Dutta, Devendra K. Tayal. A complete axiomatization of fuzzy functional dependencies using fuzzy function. *Fuzzy Set and Systems*, 151: 363–379, 2005.

CONGRUENCE RELATIONS ON HYPERGROUPOIDS AND ND-GROUPOIDS

P. Cordero, G. Gutiérrez, J. Martínez, M. Ojeda-Aciego, I. de las Peñas
Depto. Matemática Aplicada. Universidad de Málaga
pcordero@uma.es

Abstract

In this work we prove that the set of congruences on an nd-groupoid under suitable conditions is a complete lattice which is a sublattice of the lattice of equivalence relations on the nd-groupoid. The study of these conditions allowed to construct a counterexample to the statement that the set of (fuzzy) congruences on a hypergroupoid is a complete lattice.

Keywords: Hypergroupoid, nd-groupoid, congruence, L-fuzzy sets.

1 Introduction

This paper follows the current trend of providing suitable fuzzifications of crisp concepts, as a theoretical tool to the development of new method of reasoning under uncertainty, imprecision and lack of information. Moreover, although originally fuzzy sets were presented as mappings with codomain $[0, 1]$, the unit interval was soon replaced by more general structures, for instance a complete lattice, as in the L-fuzzy sets introduced by Goguen [7].

This paper follows one of our research lines which is aimed at investigating L-fuzzy sets where L has the structure of a multilattice [3]. Roughly speaking, a multilattice is an algebraic structure in which the restrictions imposed on a lattice, namely, the "existence of least upper bounds and greatest lower bounds" are relaxed to the "existence of minimal upper bounds and maximal lower bounds".

Recently, Cordero *et al.* [9, 10] proposed an alternative algebraic definition of multilattice which is more closely related to that of lattice, allowing for natural definitions of related structures such that multisemilattices and, in addition, is better suited for applications. For instance, Medina *et al.* [12] developed a general approach to fuzzy logic programming in which the authors use a multilattice as underlying set of truth-values for the logic.

Attending to the description given above, the main difference that one notices when working with multilattices is that the operators which compute suprema and infima are no longer single-valued, since there may be several multi-suprema or multi-infima, or may be none. This immediately leads to the theory of hyperstructures, that is, set-valued operations.

Hyperstructure theory was introduced in [11] when Marty defined hypergroups, began to analyse their properties and applied them to groups, rational fractions and algebraic functions. Nowadays, a number of different hyperstructures are widely studied from the theoretical point of view and for their applications to many subjects of pure and applied mathematics.

In this paper, we will focus on the most general hyperstructures, namely hypergroupoid and nd-groupoid. In a nutshell, let us recall that a hypergroupoid is simply a binary operator $H \times H \to \mathcal{P}(H) \setminus \{\varnothing\}$, whereas a non-deterministic groupoid (nd-groupoid, for short) is a binary operator $H \times H \to \mathcal{P}(H)$, that is, a hypergroupoid n which the restriction of the codomain being the non-empty subsets is dropped.

Our interest in nd-groupoids and hypergroupoids arises from the fact that, in a multilattice, the operators which compute the multi-suprema and multi-infima are precisely nd-groupoids or, if we have for granted that at least a multi-supremum always exists, a hypergroupoid. Most of the results will be stated mainly in terms of multisemilattices.

Several papers have been published on the lattice of fuzzy congruences on different algebraic structures [1, 5, 6, 14, 15], and in [4] we initiated research in this

direction. Specifically, we focused on the theory of congruences on a multilattice, as this is a necessary step prior to considering the multilattice-based generalization of the concept of L-fuzzy congruence. In this paper, we consider the particular case of the congruences on an nd-groupoid.

The fact that the structure of nd-groupoid is simpler than that of a multilattice does not necessarily mean that the theory is simpler as well. Specifically, the set of congruences on an nd-groupoid is not a lattice unless we assume some extra properties. This problem led us to review some related literature and, as a result, we found one counter-example in the context of congruences on a hypergroupoid.

2 Preliminaries

The concept of multilattice [3] was introduced by Benado in 1954, and is an extension of the concept of lattice by means the so-called multi-suprema and multi-infima, which are formally defined below:

Definition 2.1 *Given $(M, \leq)$ a partially ordered set (poset, hereafter) and $B \subseteq M$, a **multi-supremum of** B is a minimal element of the set of upper bounds of B and $\text{Multi-sup}(B)$ denote the set of all the multi-suprema of B. Dually, we define the **multi-infima**.*

Now, we introduce the definition of multilattice.

Definition 2.2 *A poset $(M, \leq)$ is said to be a **multi-semilattice** if it satisfies that, for all $a, b, x \in M$ with $a \leq x$ and $b \leq x$, there exists $z \in \text{Multi-sup}(\{a, b\})$ such that $z \leq x$.*

*Similarly to what happens in the theory of lattices, a poset $(M, \leq)$ is said to be a **multilattice** if it is a multisemilattice and also its dual $(M, \geq)$.*

Note that the definition is consistent with the existence of two incomparable elements *without* any multisupremum. In other words, $\text{Multi-sup}(\{a, b\})$ can be empty.

Let us move now to the context of congruences on different algebraic structures.

It is well-known that given a lattice L, the set of all congruence relations on L, partially ordered by set inclusion, is a complete lattice [8]. In the recent literature, several authors have presented different approaches to fuzzy congruence relations on some algebraic structures [5, 6, 13, 15]. The most general frame which almost includes multilattices seems to be that of a hypergroupoid. In [2], the notion of fuzzy congruence relation on hypergroupoid is introduced.

Note that we will assume the following notational convention: If $X, Y \subseteq H$ then $X \equiv Y$ denotes that, for all $x \in X$, there exists $y \in Y$ such that $x \equiv y$ and for all $y \in Y$ there exists $x \in X$ such that $x \equiv y$.

Definition 2.3 *Let $(H, \cdot)$ be a hypergroupoid, that is, a mapping $\cdot : H \times H \to \mathcal{P}(H) \setminus \{\varnothing\}$. A **congruence** relation on H is an equivalence relation $\equiv$ such that for all $a, b, c \in H$, if $a \equiv b$, then $ac \equiv bc$.*

Definition 2.4 *(Zadeh, [16]) Let H be a nonempty set. A **fuzzy relation** ρ on H is a fuzzy subset of $H \times H$ (i.e. ρ is a function from $H \times H$ to $[0, 1]$). Now, ρ is said to be a **fuzzy equivalence** relation if*

1. *$\rho(x, x) = \sup_{y,z \in H} \rho(y, z)$ (fuzzy reflexivity),*
2. *$\rho(y, x) = \rho(x, y)$ (fuzzy symmetry),*
3. *$\rho(x, y) \geq \sup_{z \in H} \min(\rho(x, z), \rho(z, y))$ (fuzzy transitivity).*

Definition 2.5 *(Bakhshi and Borzooei [2]) Let $(H, \cdot)$ be a hypergroupoid and ρ a fuzzy relation on H.*

1. *ρ is **fuzzy left** (right) **compatible** if for all $u \in ax$ ($u \in xa$) there exists $v \in ay$ ($v \in ya$) and for all $v \in ay$ ($v \in ya$) there exists $u \in ax$ ($u \in xa$) such that $\rho(u, v) \geq \rho(x, y)$, for all $x, y, a \in H$ and **fuzzy compatible** if it is both fuzzy left and right compatible.*
2. *ρ is a **fuzzy** (left, right) **congruence** relation on H if it is a fuzzy (left, right) compatible equivalence relation on H.*

Under the additional assumption of commutativity with respect to the usual composition of binary relations, Bakhshi and Borzooei [2], apparently proved that the set of all fuzzy congruence relations on an hypergroupoid $(H, \cdot)$ is a complete lattice where, in particular, the infimum of two congruences is its intersection. The following example proves that, without additional hypotheses, the statement is false even in a crisp setting.

Example 2.1 *Let H be the set $\{a, b, c, u_0, u_1, v_0, v_1\}$ provided with a commutative hyperoperation $*$ which is defined as follows:*

$$a * a = a * b = b * b = \{a, b\}; \quad a * c = \{u_0, u_1\};$$

$$b * c = \{v_0, v_1\} \quad \text{and } x * y = \{c\}, \quad \text{elsewhere}$$

Consider $R, S : H \times H \to \{0, 1\}$ two binary relations, where R is the least equivalence relation containing $\{(a, b), (u_0, v_0), (u_1, v_1)\}$ and S the least equivalence relation containing $\{(a, b), (u_0, v_1), (u_1, v_0)\}$. A simple (but tedious) check shows that $R \circ S = S \circ R$ and they are both compatible with the hyperoperation

$*$, *therefore, both* R *and* S *are congruence relations. However, one can check that the intersection* $R \cap S$ *is not a congruence relation.*

The previous example motivated the search for a sufficient condition which granted the structure of complete lattice for the set of congruences on a hypergroupoid and, by extension, on an nd-groupoid.

The idea underlying this missing property comes from a previous work by the authors [4], in which we studied congruence relations on a multilattice M as a natural inception for the fuzzy case. In that paper, we proved that the set of equivalence relations compatible with both operations of a multilattice forms a complete lattice, *under a suitable form of distributivity.*

In the following section we extend the concept of hypergroupoid by formally introducing and studying the properties of nd-groupoids, together with a suitable generalization of the property of distributivity, albeit in a framework with just one operation!!

3 Congruence relations in nd-groupoids

We are specially interested in a generalization of hypergroupoid that we will call non deterministic groupoid (nd-groupoid, for short).

We consider a set A endowed with an **nd-operation**, $\cdot : A \times A \to \mathcal{P}(A)$ and thus the pair $(A, \cdot)$ is called **nd-groupoid**. Notice that the definition allows the assignment of the empty set to a pair of elements, that is $a \cdot b = \varnothing$, this mere fact, albeit simple, represents an important difference with hypergroupoids.

Notation: As usual, we will use the notational conventions:

- If $a \in A$ and $X \subseteq A$ $aX = \{ax \mid x \in X\}$ and $Xa = \{xa \mid x \in X\}$. In particular, $a\varnothing = \varnothing a = \varnothing$
- We will use multiplicative notation and, thus, the dot is omitted.
- When the result of the nd-operation is a singleton, we will often omit the braces.

As stated in the introduction, our interest in extending the concept of hypergroupoid is justified by the algebraic characterization of multilattices and multisemilattices, since the operators for multi-suprema and multi-infima are both examples of nd-groupoids.

With this idea in mind, we introduce below the extension to the framework of nd-groupoids of some well-known properties. Assume that $(A, \cdot)$ is an nd-groupoid:

- **Idempotency**: $aa = a$ for all $a \in A$.
- **Commutativity**: $ab = ba$ for all $a, b \in A$.
- **Left m-associativity**: $(ab)c \subseteq a(bc)$ when $ab = b$, for all $a, b, c \in A$.
- **Right m-associativity**: $a(bc) \subseteq (ab)c$ when $bc = c$, for all $a, b, c \in A$.
- **m-associativity**: if it is left and right m-associative.

Note that the prefix 'm' has its origin in the concept of multilattice.

We will focus our interest in the binary relation usually named *natural ordering*, which is defined by

$$a \leq b \text{ if and only if } ab = b$$

In general, this relation is not an ordering, but we have sufficient conditions which guarantee the properties of an ordering. Specifically, it is reflexive if the nd-groupoid is idempotent, the relation is antisymmetric if the nd-groupoid is commutative and, finally, it is transitive if the nd-groupoid is m-associative.

The two following properties of nd-groupoids have an important role in multilattice theory:

- C_1: $c \in ab$ implies that $a \leq c$ and $b \leq c$.
- C_2: $c, d \in ab$ and $c \leq d$ imply that $c = d$.

These two properties are named **comparability**. Similarly to lattice theory, we can define algebraically the concept of multisemilattice as an nd-groupoid that satisfies idempotency, commutativity, m-associativity and comparability laws. Both definitions of multisemilattice can be proved to be equivalent considering $a \cdot b = \text{Multi-sup}\{a, b\}$ and $\leq$ the natural ordering (see [10, Theorem 2.11]).

In the following, we will consider congruence relations on an nd-groupoid. This concept is defined in the same manner as in Definition 2.3.

The following result is an immediate consequence from the definition.

Lemma 3.1 *Let* $(A, \cdot)$ *be an idempotent nd-groupoid and* $\equiv$ *be a congruence relation. If* $a \equiv b$ *then* $\varnothing \neq ab \equiv a$.

Theorem 3.2 *Let* $(A, \cdot)$ *be an nd-groupoid satisfying idempotency and property* C_1. *An equivalence relation* $\equiv$ *is a congruence if and only if the following holds:*

$$\forall a, b, c \in A, \text{ if } a \leq b \text{ and } a \equiv b, \text{ then } ac \equiv bc.$$

Proof The necessity is obvious, thus we will just prove the sufficiency.

If $a \equiv b$ then, by Lemma 3.1, there exists $z \in ab$ such that $a \equiv z \equiv b$. Property C_1 ensures that $a \leq z$ and $b \leq z$ and then, by the condition, $ac \equiv zc \equiv bc$. □

In the rest of the paper, we focus on the search of properties that ensure the condition of the previous theorem.

Proposition 3.3 *Let $(A, \cdot)$ be an m-associative nd-groupoid that satisfies C_1 and, for $a, b, c \in A$, consider $a \leq b$ and $z \in bc$:*

1. *There exists $w \in ac$ such that $w \leq z$.*
2. *Furthermore, if $(A, \cdot)$ is commutative and C_2 holds, then $a \equiv b$ implies that every element w as in the previous item satisfies that $w \equiv z$.*

Proof

1. By hypothesis $a \leq b$ and, by C_1, since $z \in bc$, we obtain $b \leq z$. Therefore $a \leq z$ because, by m-associativity of the nd-operation $\cdot$, the relation $\leq$ is transitive.

   Now, again by C_1, since $z \in bc$, $c \leq z$ and, by m-associativity, $z = az = a(cz) \subseteq (ac)z$. In particular, we have that $z \in (ac)z$, this implies the existence of $w \in ac$ such that $z = wz$, that is, $w \leq z$.

2. Consider $w \in ac$ such that $wz = z$ (i.e., $w \leq z$). By C_1 we have that $a \leq w$ and $w = aw \equiv bw$. Now, it is sufficient to prove that $z \in bw$.

   Since $b \leq z$ and $w \leq z$, by m-associativity, $z = bz = b(wz) \subseteq (bw)z$, that is, $z \in (bw)z$. Consider $z' \in bw$ such that $z' \leq z$. By C_1, $b \leq z'$ and $c \leq w \leq z'$. Once again, m-associativity ensures that there exists $z'' \in bc$ such that $z'' \leq z'$ and therefore $z'' \leq z$. Now, by C_2, $z'' = z$. From $z \leq z'$ and $z' \leq z$, by commutativity, the relation $\leq$ is antisymmetric and, hence,$z = z' \in bw$. □

In order to prove the converse result, we need to introduce the following definition.

Definition 3.4 *An nd-operation $\cdot$ in a set A is said to be **m-distributive** when, for all $a, b, c \in A$, if $a \leq b$ and $w \in ac$, then $bw \cap bc \neq \varnothing$.*

The justification of this name is that a multilattice $(A, \vee, \wedge)$ in which both operations are m-distributive satisfies the following property: for all $a, b \in A$ with $a \leq b$ and $c \in A$:

1. $(a \wedge b) \vee c \subseteq (a \vee c) \wedge (b \vee c)$
2. $(a \vee b) \wedge c \subseteq (a \wedge c) \vee (b \wedge c)$

In fact, the two latter conditions are equivalent to the m-distributivity of $\wedge$ and $\vee$, as the following result shows.

Proposition 3.5 *Let $(A, \vee, \wedge)$ be a multilattice and $a, b, c \in A$. The following conditions are equivalent:*

1. *If $a \leq b$ and $w \in a \vee c$, then $b \vee w \cap b \vee c \neq \varnothing$.*
2. *If $a \leq b$, then $(a \wedge b) \vee c \subseteq (a \vee c) \wedge (b \vee c)$.*

Proof

- $(1 \Rightarrow 2)$ If $w \in (a \wedge b) \vee c = a \vee c$ then, by hypothesis 1, there exists $u \in b \vee w \cap b \vee c$. Since $u \in b \vee w$, by property C_1, we have $w \leq u$ and, therefore, $w = w \wedge u \subseteq (a \vee c) \wedge (b \vee c)$.
- $(2 \Rightarrow 1)$ Conversely, assume $w \in a \vee c$, since $a \leq b$, we have $a = a \wedge b$ and, by hypothesis 2 we have $w \in (a \wedge b) \vee c \subseteq (a \vee c) \wedge (b \vee c)$. Thus, there exists an element $u \in a \vee c$ and $v \in b \vee c$ such that $w \in u \wedge v$. Moreover,[1] by C_1, $w \leq v$ and $b \leq v$. By m-associativity we have now that $v = b \vee v = b \vee (w \vee v) \subseteq (b \vee w) \vee v$. This means that there exists $v' \in b \vee w$ such that $v = v' \vee v$ (that is, $v' \leq v$). Finally, we will prove that $v \leq v'$ and thus $v = v' \in b \vee w \cap b \vee c$ as desired.

  By C_1 we have that $b \leq v'$ and $c \leq w \leq v'$ which, by transitivity, implies $c \leq v'$. This means that v' is an upper bound of b and c and, by definition of multilattice, there should exist $v'' \in b \vee c$ such that $v'' \leq v'$. But we already know that $v' \leq v$, hence $v'' \leq v' \leq v$, but as v and v'' are multi-suprema, the only consistent possibility is that $v = v' = v''$. □

Proposition 3.6 *Let $(M, \cdot)$ be an m-distributive nd-groupoid that satisfies C_1 and $a, b, c \in M$. If $a \leq b$ and $w \in ac$ then there exists $z \in bc$ such that $w \leq z$.*

Proof By m-distributivity, from $a \leq b$ and $w \in ac$, we obtain that there exists $z \in bw \cap bc$. Now, by C_1 $w \leq z$. □

Notice that the properties required as hypotheses of Proposition 3.6 and Proposition 3.3 are those of a multisemilattice without idempotency. The following result, stated in terms of a multisemilattice, is a straightforward consequence of these two propositions.

[1] Recall that, as stated in the definition of multilattice, the order used considered for $\wedge$ is $\geq$, this explains the following use of property C_1 .

Proposition 3.7 *Let $(M, \cdot)$ be an m-distributive multisemilattice, $\equiv$ be a congruence relation and $a, b, c \in M$. If $a \leq b$, $a \equiv b$, $w \in ac$ and $z \in bc$ with $w \leq z$ then $w \equiv z$.*

Now, we have all the required properties and lemmas needed in order to face the main goal of this paper, namely, to prove that the set of congruences of an nd-groupoid is a complete lattice.

It is well-known that, for every set A, the set of equivalence relations on A, $Eq(A)$, with the inclusion ordering (in the powerset of $A \times A$) is a complete lattice in which the infimum is the meet and the supremum is the transitive closure of the join. This generalization of distributivity will be proved to be a sufficient condition for the set of congruence relations being a complete lattice.

Theorem 3.8 *The set of the congruence relations in an m-distributive multisemilattice M, $Con(M)$, is a sublattice of $Eq(M)$ and, moreover is a complete lattice wrt the inclusion ordering.*

Proof Let $\{\equiv_i\}_{i\in\Lambda}$ be a set of congruence relations in M, consider $\equiv_\cap$ to be their intersection.

From Theorem 3.2 we have just to check that, for all $a, b, c \in M$, $a \leq b$ and $a \equiv_\cap b$ imply that $ac \equiv_\cap bc$.

From Proposition 3.3, if $z \in bc$ then there exists $w \in ac$ such that $w \leq z$ and, for all $w \in ac$ with $w \leq z$ and all $i \in \Lambda$, $w \equiv_i z$ (so $w \equiv_\cap z$).

Conversely, from Proposition 3.6 and Proposition 3.7, if $w \in ac$ then there exists $z \in bc$ such that $w \leq z$ and, for all $z \in bc$ with $w \leq z$ and all $i \in \Lambda$, $w \equiv_i z$ (so $w \equiv_\cap z$).

The proof for the transitive closure of union follows by a routine calculation. □

4 Conclusions and future work

We have proved that the set of congruence relations on an nd-groupoid is a complete lattice which is a sublattice of the lattice its equivalence relations. The main step in the proof of this statement is the introduction of the concept of m-distributivity, which allowed for translating the ideas used in the context of congruences on multilattices.

As future work on this research line, our plan is to keep investigating new or analogue results concerning congruences on generalized algebraic structures, specially in a non-deterministic sense. Then, continuing towards our main aim, studying the computational properties of multilattices as a suitable algebraic structure on which found an extended theory of fuzzy structures, we will try to lift our results from a crisp to a fuzzy setting.

Acknowledgements

This paper has been partially supported by projects TIN2006-15455-C03-01, TIN2007-65819 (Science Ministry of Spain) and P06-FQM-02049 (Junta de Andalucía).

References

[1] M. Bakhshi and R. Borzooei. Lattice structure on fuzzy congruence relations of a hypergroupoid. *Information Sciences*, 177:3305–3313, 2007.

[2] M. Bakhshi and R. A. Borzooei. Lattice structure on fuzzy congruence relations of a hypergroupoid. *Inform. Sci.*, 177(16):3305–3313, 2007.

[3] M. Benado. Les ensembles partiellement ordonnés et le théorème de raffinement de Schreier. I. *Čehoslovack. Mat. Ž.*, 4(79):105–129, 1954.

[4] P. Cordero, G. Gutiérrez, J. Martínez, M. Ojeda-Aciego, and I. de las Peñas. Congruence relations on multilattices. In *The 8th International FLINS Conference on Computational Intelligence in Decision and Control*, 2008.

[5] P. Das. Lattice of fuzzy congruences in inverse semigroups. *Fuzzy Sets and Systems*, 91(3):399–408, 1997.

[6] T. Dutta and B. Biswas. On fuzzy congruence of a near-ring module. *Fuzzy Sets and Systems*, 112(2):399–408, 2000.

[7] J. Goguen. L-fuzzy sets. *J. Math. Anal. Appl.*, 18:145–174, 1967.

[8] G. Grätzer. *General lattice theory*. Birkhäuser, second edition, 1998.

[9] J. Martínez, G. Gutiérrez, I. P. de Guzmán, and P. Cordero. Generalizations of lattices via non-deterministic operators. *Discrete Math.*, 295(1-3):107–141, 2005.

[10] J. Martínez, G. Gutiérrez, I. P. de Guzmán, and P. Cordero. Multilattices via multisemilattices. In *Topics in applied and theoretical mathematics and computer science*, pages 238–248. WSEAS, 2001.

[11] F. Marty. Sur une généralisation de la notion de groupe. In *8th Congress Math. Scandinaves*, pages 45–49, Stockholm, 1934.

[12] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Fuzzy logic programming via multilattices. *Fuzzy Sets and Systems*, 158(6):674–688, 2007.

[13] V. Murali. Fuzzy equivalence relations. *Fuzzy Sets and Systems*, 30(2):155–163, 1989.

[14] V. Murali. Fuzzy congruence relations. *Fuzzy Sets and Systems*, 41(3):359–369, 1991.

[15] Y. Tan. Fuzzy congruences on a regular semigroup. *Fuzzy Sets and Systems*, 117(3):399–408, 2001.

[16] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

RATIONAL COMPLETENESS RESULTS FOR PROMINENT PROPOSITIONAL FUZZY LOGICS WITH TRUTH-CONSTANTS

Francesc Esteva[1] **Lluís Godo**[1] **Carles Noguera**[2]

[1] IIIA - CSIC, 08193 Bellaterra, Spain
{*esteva,godo*}*@iiia.csic.es*

[2] Dept. of Math. and Computer Science, University of Siena, 53100 Siena, Italy
cnoguera@iiia.csic.es

Abstract

In this paper we consider expansions of Łukasiewiz, Product, Gödel and Nilpotent Minimum logics with truth-constants for an algebra of rational truth-values. We study the semantics for these logics given by chains defined over the rational unit interval and the completeness properties they provide, with a special attention to the completeness with respect to the canonical chain (i.e. the algebra where each truth-constant is interpreted in its corresponding truth-value).

Keywords: t-norm based fuzzy logics, truth-constants, rational semantics, completeness.

1 INTRODUCTION

Fuzzy logics are the logics corresponding to the notion of comparative truth, i.e. the paradigm where the classical truth-values are extended by adding intermediate ones which are comparable by using some total ordering. The expansions of fuzzy logics with truth-constants have been proposed as a means to deal explicitly in the language with the intermediate truth-values. They first appeared already in the 70s when Pavelka [13] introduced a propositional many-valued logical system which turned out to be equivalent to the expansion of Łukasiewicz logic by adding into the language a truth-constant $\overline{r}$ for each *real* $r \in [0,1]$, together with a number of additional axioms. Hájek in [11] simplified in significant form Pavelka's system by showing that the same results could be obtained by adding just a truth-constant for each *rational* $r \in [0,1]$. Expansions of other t-norm based fuzzy logics with countable sets of truth-values and their completeness properties with respect to the standard semantics over the real unit interval have been studied in [4, 6, 8, 9, 14]. Nevertheless, although the power of rational-chain semantics for fuzzy logics was been shown in [3, 7, 2], it had not been considered for logics with truth-constants yet. Thus, in this paper we will study rational completeness properties for fuzzy logics with truth-constants.

2 PRELIMINARIES

The basic logic in this framework is the (propositional) Monoidal t-norm based logic MTL [5], with primitive connectives & (multiplicative conjunction), $\rightarrow$ (implication), $\wedge$ (additive conjunction) and the truth-constant $\overline{0}$. MTL is in fact the logic of left-continuous t-norms and their residua [12], in the sense that the set of its theorems is exactly $\bigcap\{Taut(*) \mid * \text{ is a left-continuous t-norm}\}$, where $Taut(*)$ denotes the set of tautologies when interpreting respectively &, $\rightarrow$ and $\wedge$ by $*$, its residuum $\Rightarrow$ and the min operation. In this setting, we denote by $[0,1]_*$ the standard MTL-chain defined by the left-continuous t-norm $*$ and its residuum $\Rightarrow$, i. e. $[0,1]_* = \langle [0,1], *, \Rightarrow, \min, \max, 0, 1\rangle$, and by L_* the axiomatic extension of MTL whose equivalent algebraic semantics is the variety generated by $[0,1]_*$.

In this paper we will mainly focus on four prominent examples of these logics, namely Gödel (G),

Łukasiewicz (Ł), Product (Π), and Nilpotent Minimum (NM) logics, corresponding respectively to the cases when $*$ is the minimum t-norm, the Łukasiewicz t-norm, the product t-norm, or the nilpotent minimum t-norm (see [11, 5] for their axiomatics and further details).

Given a logic L_* and a class $\mathbb{K}$ of L_*-chains, one defines three completeness properties:

- L_* has the property of *strong* $\mathbb{K}$-*completeness*, S$\mathbb{K}$C for short, if for every set of formulae Γ and every formula φ, $\Gamma \vdash_{\mathrm{L}_*} \varphi$ iff $\Gamma \models_{\mathbb{K}} \varphi$.
- L_* has the property of *finite strong* $\mathbb{K}$-*completeness*, FS$\mathbb{K}$C for short, if for every *finite* set of formulae Γ and every formula φ, $\Gamma \vdash_{\mathrm{L}_*} \varphi$ iff $\Gamma \models_{\mathbb{K}} \varphi$.
- L_* has the property of $\mathbb{K}$-*completeness*, $\mathbb{K}$C for short, if for every formula φ, $\vdash_{\mathrm{L}_*} \varphi$ iff $\models_{\mathbb{K}} \varphi$.

If $\mathbb{K}$ is the class of all chains over the real unit interval $[0,1]$ we use the notation $\mathcal{R}$C and call the properties *standard completeness*, while if it is the class of all chains over the rational unit interval $[0,1]^{\mathbb{Q}} = [0,1] \cap \mathbb{Q}$ we use the notation $\mathcal{Q}$C. The standard and rational completeness properties for the above considered logics are in Table 1.

Theorem 2.1 ([2])**.** *Let $*$ be a left-continuous t-norm. If L_* has the* FS$\mathcal{R}$C*, then it has the* S$\mathcal{Q}$C*.*

	Ł	Π	G	NM
$\mathcal{R}$C	Yes	Yes	Yes	Yes
FS$\mathcal{R}$C	Yes	Yes	Yes	Yes
S$\mathcal{R}$C	No	No	Yes	Yes
$\mathcal{Q}$C	Yes	Yes	Yes	Yes
FS$\mathcal{Q}$C	Yes	Yes	Yes	Yes
S$\mathcal{Q}$C	Yes	Yes	Yes	Yes

Table 1: Standard and rational completeness properties for propositional fuzzy logics.

2.1 ADDING TRUTH-CONSTANTS

Given a left-continuous t-norm $*$ let $\mathcal{C}$ be a *countable* subalgebra of $[0,1]_*$. Then, $\mathrm{L}_*(\mathcal{C})$ is the propositional fuzzy logic defined as follows:

(i) the language of $\mathrm{L}_*(\mathcal{C})$ is the one of L_* expanded with a new propositional constant $\overline{r}$ for each $r \in C \setminus \{0,1\}$,

(ii) the axioms and rules of $\mathrm{L}_*(\mathcal{C})$ are those of L_* plus the book-keeping axioms:

$$\overline{r} \& \overline{s} \leftrightarrow \overline{r * s}$$
$$(\overline{r} \to \overline{s}) \leftrightarrow \overline{r \Rightarrow s}$$

for each $r, s \in C$.

Its algebraic counterpart, the $\mathrm{L}_*(\mathcal{C})$-algebras, are the expansions of L_*-algebras with nullary functions $\overline{r}^{\mathcal{A}}$ (one for each $r \in C$) satisfying the book-keeping axioms, i.e. for every $r, s \in C$ the following identities hold:

$$\overline{r}^{\mathcal{A}} \& \overline{s}^{\mathcal{A}} = \overline{r * s}^{\mathcal{A}}$$
$$\overline{r}^{\mathcal{A}} \to \overline{s}^{\mathcal{A}} = \overline{r \Rightarrow s}^{\mathcal{A}}.$$

$\mathrm{L}_*(\mathcal{C})$-chains defined over the real unit interval $[0,1]$ are called *standard*. Among them, there is one which reflects the intended standard semantics, the so-called *canonical standard* $\mathrm{L}_*(\mathcal{C})$-chain $[0,1]_{\mathrm{L}_*(\mathcal{C})}$ which is the standard chain where the truth-constants are interpreted by their defining values, i.e. one has $\overline{r}^{\mathcal{A}} = r$ for all $r \in C$ whenever $\mathcal{A} = [0,1]_{\mathrm{L}_*(\mathcal{C})}$. It is worth to point out that for a logic $\mathrm{L}_*(\mathcal{C})$ there may exist multiple standard chains as soon as there exist different ways of interpreting the truth-constants on $[0,1]$ respecting the book-keeping axioms. Indeed, in the standard chains of Ł, Π, G and NM logics, the only possible interpretations are of the following type: for each proper filter F of $\mathcal{C}$ (a non-empty upper subset closed under the t-norm and not containing 0), truth-constants admit the following interpretation:

$$\overline{r}^{\mathcal{A}} = \begin{cases} 1, & \text{if } r \in F \\ 0, & \text{if } \neg r \in F \\ r, & \text{otherwise} \end{cases}$$

The resulting standard chain is denoted $[0,1]^F_{\mathrm{L}_*(\mathcal{C})}$. In the case of Ł the only proper filter is the trivial one $\{1\}$. In the case of Π there are two: $(0,1]$ and $\{1\}$. For G (resp. NM) there are many: $[c,1]$ for

every $c > 0$ (resp. $c > \frac{1}{2}$) and $(c, 1]$ for every $c \geq 0$ (resp. $c \geq \frac{1}{2}$).

Completeness properties w.r.t. both the class of all standard chains and the canonical standard chain have been studied in the literature and solved for the logics we are considering (see Table 2). In the cases where the standard completeness fails one can try to improve the situation by restricting the logic to the so called *evaluated language*, i.e. formulae $\overline{r} \to \varphi$ where no additional truth-constant occurs in φ. In the case of NM one must require that the value r is *positive*, i.e. $r > \frac{1}{2}$. Table 3 summarizes these completeness results.

	Ł($\mathcal{C}$)	Π($\mathcal{C}$)	G($\mathcal{C}$)	NM($\mathcal{C}$)
$\mathcal{R}$C	Yes	Yes	Yes	Yes
FS$\mathcal{R}$C	Yes	Yes	Yes	Yes
S$\mathcal{R}$C	No	No	Yes	Yes
Can$\mathcal{R}$C	Yes	Yes	Yes	Yes
CanFS$\mathcal{R}$C	Yes	No	No	No
CanS$\mathcal{R}$C	No	No	No	No

Table 2: Standard completeness properties for propositional fuzzy logics with truth-constants.

	Ł($\mathcal{C}$)	Π($\mathcal{C}$)	G($\mathcal{C}$)	NM($\mathcal{C}$)
$\mathcal{R}$C	Yes	Yes	Yes	Yes
FS$\mathcal{R}$C	Yes	Yes	Yes	Yes
S$\mathcal{R}$C	No	No	Yes	Yes
Can$\mathcal{R}$C	Yes	Yes	Yes	Yes
CanFS$\mathcal{R}$C	Yes	Yes	Yes	Yes
CanS$\mathcal{R}$C	No	No	No	No

Table 3: Standard completeness properties for propositional fuzzy logics with truth-constants restricted to (positively) evaluated formulae.

3 NEW RESULTS: THE GENERAL CASE

Let $*$ be a left-continuous t-norm and $\Rightarrow$ its residuum such that the rational unit interval $[0,1]^{\mathbb{Q}}$ is closed under the operations $*$ and $\Rightarrow$. Let $[0,1]^{\mathbb{Q}}_*$ be the L_*-chain defined by the restriction of $*$ and $\Rightarrow$ to $[0,1]^{\mathbb{Q}}$. Let $\mathcal{C}$ be a countable subalgebra of $[0,1]^{\mathbb{Q}}_*$ and consider the logic $\mathrm{L}_*(\mathcal{C})$. Now $\mathrm{L}_*(\mathcal{C})$-chains defined over the rational unit interval are called *rational chains* and among them, the one which reflects the intended rational semantics is the so-called *canonical rational* $\mathrm{L}_*(\mathcal{C})$-chain

$$[0,1]^{\mathbb{Q}}_{\mathrm{L}_*(\mathcal{C})} = \langle [0,1], *, \Rightarrow, \min, \max, \langle r : r \in C \rangle \rangle,$$

i.e. the rational chain over $[0,1]^{\mathbb{Q}}_*$ where the truth-constants are interpreted by their defining values.

As already mentioned, in this paper we restrict ourselves to the logics $\mathrm{L}_*(\mathcal{C})$ defined by four prominent t-norms (Gödel, Łukasiewicz, Product and Nilpotent Minimum), and study for them completeness properties with respect to all rational $\mathrm{L}_*(\mathcal{C})$-chains ($\mathcal{Q}C, FS\mathcal{Q}$C and S$\mathcal{Q}$C) and with respect to the canonical rational $\mathrm{L}_*(\mathcal{C})$-chain (Can$\mathcal{Q}$C, CanFS$\mathcal{Q}$C and CanS$\mathcal{Q}$C). In fact, all the logics we are considering obviously enjoy the $\mathcal{Q}$C, FS$\mathcal{Q}$C and S$\mathcal{Q}$C by virtue of Theorem 2.1, since all of them enjoy the FS$\mathcal{R}$C (see Table 2). We start the study of canonical completeness properties for the Łukasiewicz-based logics Ł($\mathcal{C}$).

Theorem 3.1. *For every countable $\mathcal{C} \subseteq [0,1]^{\mathbb{Q}}_{Ł}$, the logic Ł($\mathcal{C}$) enjoys the* Can FS$\mathcal{Q}$C.

Proof. Suppose that for some arbitrary set of formulae we have $\varphi_1, \ldots, \varphi_n \vdash_{\mathrm{Ł}(\mathcal{C})} \psi$. We must prove that $\varphi_1, \ldots, \varphi_n \not\models_{[0,1]^{\mathbb{Q}}_{\mathrm{Ł}(\mathcal{C})}} \psi$. On one hand, by the FS$\mathcal{R}$C, there is an evaluation e over $[0,1]_{\mathrm{Ł}(\mathcal{C})}$ such that $e(\varphi_1) = \ldots = e(\varphi_n) = 1$ and $e(\psi) < 1$. On the other hand, as proved in [1][1], we know that $[0,1]_{\mathrm{Ł}(\mathcal{C})}$ is partially embeddable into $[0,1]^{\mathbb{Q}}_{\mathrm{Ł}(\mathcal{C})}$. Therefore, by embedding into $[0,1]^{\mathbb{Q}}_{\mathrm{Ł}(\mathcal{C})}$ the values of all the subformulae of the formulae involved, we obtain an evaluation over the rational canonical chain such that is a model of $\{\varphi_1, \ldots, \varphi_n\}$ while it is not a model of ψ. □

However, this property fails for the remaining logics we are considering. Indeed, suppose that $*$ is the product, Gödel or Nilpotent Minimum t-norm and $\mathcal{C} \subseteq [0,1]^{\mathbb{Q}}_*$, let F be a non-trivial proper filter of $\mathcal{C}$ and take $r \in F \setminus \{1\}$. Then:

- $(p \to q) \to \overline{r} \models_{[0,1]^{\mathbb{Q}}_{\mathrm{L}_*(\mathcal{C})}} q \to p$

[1] Actually the result we use from [1] comes from a translation of the paper in Russian [10].

- $(p \to q) \to \overline{r} \not\models_{([0,1]^{\mathbb{Q}})^F_{\mathrm{L}_*(\mathcal{C})}} q \to p$

Therefore, there is an entailment which holds for the canonical rational chain, but not for all the chains. Thus none of these logics enjoy the CanFS$\mathcal{Q}$C. We turn now to the Can$\mathcal{Q}$C.

Theorem 3.2. *For every countable* $\mathcal{C} \subseteq [0,1]^{\mathbb{Q}}_{\Pi}$, *the logic* $\Pi(\mathcal{C})$ *enjoys the* Can$\mathcal{Q}$C.

Proof. For this proof we need to introduce some notation. Given $x = \langle x_1, \ldots, x_n \rangle \in \mathbb{R}^n$ and $\delta = \langle \delta_1, \ldots, \delta_n \rangle \in (\mathbb{R}_+)^n$, we define the set $E_\delta(x) = \{\langle y_1, \ldots, y_n \rangle \in \mathbb{R}^n \mid x_i = y_i \text{ if } x_i \in \mathbb{Q}, \text{ and } y_i \in (x_i - \delta_i, x_i + \delta_i) \text{ if } x_i \notin \mathbb{Q}\}$.

Suppose that $\not\vdash_{\Pi(\mathcal{C})} \varphi$. Assume further that the variables of φ are among $\{p_1, \ldots, p_n\}$. By the Can$\mathcal{R}$C, there is an evaluation e on $[0,1]_{\Pi(\mathcal{C})}$ such that $e(\varphi) < 1$. We prove by induction that for every subformula ψ of φ:

1. If $e(\psi) = 0$, then there is $E_\delta(e(p_1), \ldots, e(p_n))$ such that for every evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$, if $\langle v(p_1), \ldots, v(p_n) \rangle \in E_\delta(e(p_1), \ldots, e(p_n))$ then $v(\psi) = 0$.
2. If $e(\psi) \neq 0$, then for every $\varepsilon > 0$ there is $E_\delta(e(p_1), \ldots, e(p_n))$ such that for every evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$, if $\langle v(p_1), \ldots, v(p_n) \rangle \in E_\delta(e(p_1), \ldots, e(p_n))$ then $\mid v(\psi) - e(\psi) \mid < \varepsilon$.

From this it easily follows that there is an evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$ such that $v(\varphi) < 1$, and thus the theorem is proved. □

Theorem 3.3. *For every* $\mathcal{C}$, *the logics* $\mathrm{G}(\mathcal{C})$ *and* $\mathrm{NM}(\mathcal{C})$ *enjoy the* Can$\mathcal{Q}$C.

Proof. It can be proved just in the same way as we proved in [8] that these logics enjoy the Can$\mathcal{R}$C. We rewrite the proof here for the reader's convenience. Assume that $*$ is the Gödel or Nilpotent Minimum t-norm. Suppose φ is a tautology with respect to $[0,1]^{\mathbb{Q}}_{\mathrm{L}_*(\mathcal{C})}$. We will prove that φ is also a tautology with respect to $([0,1]^{\mathbb{Q}})^F_{\mathrm{L}_*(\mathcal{C})}$ for each filter F of $\mathcal{C}$, which, due to the $\mathcal{Q}$C, implies that $\vdash_{\mathrm{L}_*(\mathcal{C})} \varphi$. Let e be an interpretation over the chain $([0,1]^{\mathbb{Q}})^F_{\mathrm{L}_*(\mathcal{C})}$. Suppose that $\mathcal{A}$ is the finite algebra generated by $\{e(\psi) \mid \psi$ subformula of $\varphi\}$ and $\alpha = \min\{r \in F \mid \overline{r}$ occurs in $\varphi\}$. Take $f : [0,1]^{\mathbb{Q}} \to [0,1]^{\mathbb{Q}}$ such that $f : (\neg\alpha, \alpha)^{\mathbb{Q}} \to (0,1)^{\mathbb{Q}}$ is a bijection, $f(r) = r$ for all $r \notin F$ such that $\neg r \notin F$ and $\overline{r}$ occurs in φ. So defined, f is a homomorphism from $\mathcal{A}$ into the canonical rational chain. Then define an evaluation e' over the canonical rational chain by $e'(p) = f^{-1}(e(p))$ if p is a propositional variable that appears in φ and $e'(p) = 1$ otherwise. Since φ is a tautology for the canonical rational chain, $e'(\varphi) = 1$. Take the algebra $[0,1]^{\mathbb{Q}}_*/F_\alpha$ where F_α is the filter $[\alpha, 1] \cap C$. This algebra is isomorphic to $[0,1]^{\mathbb{Q}}_*$. Define the evaluation e'' on the quotient algebra obtained from e' and it obviously satisfies $e''(\varphi) = [1]_{F_\alpha}$. Now a simple computation shows that the algebra $\mathcal{B}$ generated by $\{e''(\psi) \mid \psi$ subformula of $\varphi\}$ is isomorphic to $\mathcal{A}$ and $e''(\varphi)$ over the quotient algebra corresponds to $e(\varphi)$ over the chain $([0,1]^{\mathbb{Q}})^F_{\mathrm{L}_*(\mathcal{C})}$, and hence $e(\varphi) = 1$. □

The new results proved in this section are gathered in Table 4.

	Ł($\mathcal{C}$)	$\Pi(\mathcal{C})$	G($\mathcal{C}$)	NM($\mathcal{C}$)
$\mathcal{Q}$C	Yes	Yes	Yes	Yes
FS$\mathcal{Q}$C	Yes	Yes	Yes	Yes
S$\mathcal{Q}$C	Yes	Yes	Yes	Yes
Can$\mathcal{Q}$C	Yes	Yes	Yes	Yes
CanFS$\mathcal{Q}$C	Yes	No	No	No
CanS$\mathcal{Q}$C	No	No	No	No

Table 4: Rational completeness properties for propositional fuzzy logics with truth-constants.

4 NEW RESULTS: THE CASE OF EVALUATED FORMULAS

In this section we focus on completeness properties for the restriction of the logics to (positively) evaluated formulae. As regards to completeness properties with respect to the class of all rational chains there is nothing to say: all three properties hold for evaluated formulae because they hold in general for all formulae. Thus, we only need to examine the restricted canonical completeness properties.

All the logics under our scope fail to satisfy the

CanS$\mathcal{Q}$C restricted to (positively) evaluated formulae as it can be seen by the following counterexample (already used in [4] for standard semantics in the continuous t-norm case). Let $\Gamma = \{\overline{(\frac{n}{n+1})} \to \varphi \mid n \in \mathbb{N}\}$. For every logic $\mathrm{L}_*(\mathcal{C})$ we have $\Gamma \models_{[0,1]^{\mathbb{Q}}_{\mathrm{L}_*(\mathcal{C})}} \varphi$. But if $\Gamma \vdash_{\mathrm{L}_*(\mathcal{C})} \varphi$, then, by finitariness, there would exist $n_0 \in \mathbb{N}$ such that $\overline{(\frac{n_0}{n_0+1})} \to \varphi \vdash_{\mathrm{L}_*(\mathcal{C})} \varphi$, hence, we would have $\overline{(\frac{n_0}{n_0+1})} \to \varphi \models_{[0,1]^{\mathbb{Q}}_{\mathrm{L}_*(\mathcal{C})}} \varphi$; a contradiction.

In spite of this negative result which does not improve the situation with respect to the unrestricted CanS$\mathcal{Q}$C, we can still show several positive results for the CanFS$\mathcal{Q}$C.

Theorem 4.1. *For every countable $\mathcal{C} \subseteq [0,1]^{\mathbb{Q}}_{\Pi}$, the logic $\Pi(\mathcal{C})$ enjoys the* CanFS$\mathcal{Q}$C *restricted to evaluated formulae.*

Proof. Assume that $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \cup \{\overline{s} \to \psi\}$ is a finite set of evaluated formulae such that $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \nvdash_{\Pi(\mathcal{C})} \overline{s} \to \psi$. We must prove that $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \not\models_{[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}} \overline{s} \to \psi$. By the CanFS$\mathcal{R}$C restricted to evaluated formulae, there is an evaluation e on $[0,1]_{\Pi(\mathcal{C})}$ such that for every $i \in \{1, \ldots, n\}$, $e(\overline{r_i} \to \varphi_i) = 1$ and $e(\overline{s} \to \psi) = 1$, i.e. $s > e(\psi)$ and $r_i \leq \varphi_i$ for every i. Without loss of generality we can assume that $r_i < e(\varphi_i)$ for every i (if it is not the case, we choose any positive real number α such that for every i, $r_i \leq e(\varphi_i) < e(\varphi_i)^\alpha$ and $s > e(\psi)^\alpha > e(\psi)$ and take instead of e the evaluation $e'(x) = e(x)^\alpha$). Then we use the same trick as in the proof of Theorem 3.2 showing by induction that for every subformula ψ of φ (we assume the variables in φ are among $\{p_1, \ldots, p_n\}$):

1. If $e(\psi) = 0$ then there is $E_\delta(e(p_1), \ldots, e(p_n))$ such that for every evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$, if $\langle v(p_1), \ldots, v(p_n)\rangle \in E_\delta(e(p_1), \ldots, e(p_n))$ then $v(\psi) = 0$.

2. If $e(\psi) \neq 0$ then for every $\varepsilon > 0$ there is $E_\delta(e(p_1), \ldots, e(p_n))$ such that for every evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$, if $\langle v(p_1), \ldots, v(p_n)\rangle \in E_\delta(e(p_1), \ldots, e(p_n))$ then $\mid v(\psi) - e(\psi) \mid < \varepsilon$.

Therefore, there is an evaluation v on $[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}$ that maps to 1 the premises while maps the conclusion to some lower value and hence $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \not\models_{[0,1]^{\mathbb{Q}}_{\Pi(\mathcal{C})}} \overline{s} \to \psi$. □

For the two remaining cases the result is obtained following a reasoning analogous to that of [6] for the standard semantics.

Lemma 4.2. *Let $a \in (0,1]^{\mathbb{Q}}$ and $b \in (\frac{1}{2}, 1]^{\mathbb{Q}}$ and define a pair of mappings $f_a, f^b : [0,1]^{\mathbb{Q}} \to [0,1]^{\mathbb{Q}}$ as follows:*

$$f_a(x) = \begin{cases} 1, & \text{if } x \geq a \\ x, & \text{otherwise} \end{cases}$$

$$f^b(x) = \begin{cases} 1, & \text{if } x \geq b \\ 0, & \text{if } x \leq 1 - b \\ x, & \text{otherwise} \end{cases}$$

Then f_a is a homomorphism with respect to the Gödel truth functions, and f^b is a homomorphism with respect to the Nilpotent Minimum truth functions. Therefore, if e is a evaluation over $[0,1]^{\mathbb{Q}}_{\mathrm{G}(\mathcal{C})}$ (resp. over $[0,1]^{\mathbb{Q}}_{\mathrm{NM}(\mathcal{C})}$) then $e_a = f_a \circ e$ (resp. $e^b = f^b \circ e$) is also an evaluation over the same algebra.

Theorem 4.3. *For every countable $\mathcal{C} \subseteq [0,1]^{\mathbb{Q}}_{\Pi}$, the logics $\mathrm{G}(\mathcal{C})$ and $\mathrm{NM}(\mathcal{C})$ enjoy the* CanFS$\mathcal{Q}$C *restricted to evaluated formulae.*

Proof. Consider first the case of Gödel logic. Assume again that $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \cup \{\overline{s} \to \psi\}$ is a finite set of evaluated formulae such that $\{\overline{r_i} \to \varphi_i \mid i = 1, \ldots, n\} \nvdash_{\mathrm{G}(\mathcal{C})} \overline{s} \to \psi$. By the Deduction-detachment Theorem we have $\nvdash_{\mathrm{G}(\mathcal{C})} \bigwedge_{i=1}^n (\overline{r}_i \to \varphi) \to (\overline{s} \to \psi)$, and hence by the Can$\mathcal{Q}$C, there is an evaluation e on the rational canonical chain which is not a model of $(\bigwedge_{i=1}^n (\overline{r}_i \to \varphi) \to (\overline{s} \to \psi)$. We must find another evaluation e' which is model of $\{\overline{r}_i \to \varphi_i \mid i = 1, \ldots, n\}$ and not of $\overline{s} \to \psi$.

If e is a model of every $\overline{r}_i \to \varphi_i$, then we can take $e' = e$ and the problem is solved. Otherwise, there exists some $1 \leq j \leq n$ for which $r_j > e(\varphi_j)$ and thus $e(\overline{r}_j \to \varphi_j) = e(\varphi_j) < 1$. Let $J = \{j \mid r_j > e(\varphi_j)\}$ and let $a = e(\bigwedge_{i=1}^n \overline{r}_i \to \varphi_i) = \min\{e(\varphi_j) \mid j \in J\}$. Then the evaluation e' such that $e' = e_a$ over the propositional variables does the job. Namely, by the previous lemma, over Gödel formulae we have $e' = e_a \geq e$, so e' is still a model

of $\overline{r}_i \to \varphi_i$ for every $i \in \{1, \ldots, n\} \setminus J$. But now, $e'(\varphi_j) = 1$ for every $j \in J$, so e' is also a model of $\{\overline{r}_i \to \varphi_i \mid i = 1, \ldots, n\}$. On the other hand, since $e(\bigwedge_{i=1}^{n}(\overline{r}_i \to \varphi) \to (\overline{s} \to \psi)) < 1$, it must be $s > e(\psi)$ and $a = e(\bigwedge_{i=1}^{n}(\overline{r}_i \to \varphi_i)) > e(\psi)$. Now, by the previous lemma, $e'(\psi) = e_a(\psi) = e(\psi)$, hence $e'(\overline{s} \to \psi) = e(\overline{s} \to \psi) < 1$.

For the case of Nilpotent Minimum logic the proof runs analogously by using the evaluations e^b defined in the previous lemma. □

The obtained results are summarized in Table 5.

	Ł($\mathcal{C}$)	Π($\mathcal{C}$)	G($\mathcal{C}$)	NM($\mathcal{C}$)
$\mathcal{Q}$C	Yes	Yes	Yes	Yes
FS$\mathcal{Q}$C	Yes	Yes	Yes	Yes
S$\mathcal{Q}$C	Yes	Yes	Yes	Yes
Can$\mathcal{Q}$C	Yes	Yes	Yes	Yes
CanFS$\mathcal{Q}$C	Yes	Yes	Yes	Yes
CanS$\mathcal{Q}$C	No	No	No	No

Table 5: Rational completeness properties for propositional fuzzy logics with truth-constants restricted to (positively) evaluated formulae.

5 CONCLUSIONS

In this paper we have proposed the rational semantics for fuzzy logics expanded with truth-constants as a new topic for research. The collection of first results we have presented show the interest of the approach, as the rational semantics has demonstrated to provide better completeness properties for the main propositional fuzzy logics. We plan to extend these results in forthcoming papers by addressing the following problems:

(i) extend the investigation on rational completeness properties to wider classes of logics based on continuous and weak nilpotent t-norms;

(ii) study rational completeness properties for first-order predicate fuzzy logics, as it has been done for the standard semantics in [9]; and

(iii) investigate rational completeness results for expansions of logics with the projection connective Δ, both in the propositional and the first-order case.

Acknowledgements

The authors acknowledge partial support by the Spanish project MULOG2 (TIN2007-68005-C04), including FEDER funds of the European Union, and by the Generalitat de Catalunya grant 2005-SGR-00093. The third author also acknowledges partial support from the grant 2006-BP-A-10043 of the Departament d'Educació i Universitats of the Generalitat de Catalunya.

References

[1] R. Cignoli. Notes of the course "Álgebras de las lógicas borrosas", Univ. of Buenos Aires.

[2] P. Cintula, F. Esteva, J. Gispert, L. Godo, F. Montagna and C. Noguera. Distinguished algebraic semantics for t-norm based fuzzy logics: methods and algebraic equivalencies. Submitted, 2007.

[3] F. Esteva, J. Gispert, L. Godo and F. Montagna. On the standard and Rational Completeness of some axiomatic extensions of Monoidal t-norm Based Logic, *Studia Logica*, 71: 199–226, 2002.

[4] F. Esteva, J. Gispert, L. Godo and C. Noguera. Adding truth-constants to logics of a continuous t-norm: axiomatization and completeness results, *Fuzzy Sets and Systems* 158: 597–618, 2007.

[5] F. Esteva and L. Godo. Monoidal t-norm based logic: Towards a logic for left-continuous t-norms, *Fuzzy Sets and Systems*, 124: 271–288, 2001.

[6] F. Esteva, L. Godo and C. Noguera. On rational weak nilpotent minimum logics, *Journal of Multiple-Valued Logic and Soft Computing* 12, Number 1-2, pp. 9-32, 2006.

[7] F. Esteva, L. Godo and C. Noguera. Real, rational and finite chain semantics for fuzzy logics, *Actas del XII congreso español sobre tecnologías y lógica fuzzy ESTYLF 2006*, pp. 83-88, Ciudad Real, 2006.

[8] F. Esteva, L. Godo and C. Noguera. On expansions of t-norm based logics with truth-constants. To appear in the book *Fuzzy Logics and Related Structures* (S. Gottwald, P. Hájek, U. Höhle and E.P. Klement eds.), 2007.

[9] F. Esteva, L. Godo and C. Noguera. On completeness results for predicate Lukasiewicz, Product, Gödel, and Nilpotent Minimum logics expanded with truth-constants, *Mathware and Soft Computing* 14: 233-246, 2007.

[10] Y. S. Gurevich and A. I. Kokorin, Universal equivalence of ordered abelian groups (in Russian), *Algebra and Logic*, 2: 37-39, 1963.

[11] P. Hájek. *Metamathematics of Fuzzy Logic*, volume 4 of *Trends in Logic-Studia Logica Library*. Dordrecht/Boston/London, 1998.

[12] S. Jenei and F. Montagna. A proof of standard completeness for Esteva and Godo's logic MTL. *Studia Logica*, 70:183–192, 2002.

[13] J. Pavelka. On Fuzzy Logic I, II, III, *Zeitschrift fur Math. Logik und Grundlagen der Math*, 25:45–52, 119–134, 447–464, 1979.

[14] P. Savický, R. Cignoli, F. Esteva, L. Godo and C. Noguera. On product logic with truth-constants, *Journal of Logic and Computation* 16: 205-225, 2006.

Estudio de la ley de importación para QL y D-implicaciones

Margarita Mas[1] **Miquel Monserrat**[1] **Joan Torrens**[1]

[1] Departamento de Ciencias Matemáticas e Informática
Universidad de las Islas Baleares
Carretera de Valldemossa, Km. 7,5
07122 Palma de Mallorca
{dmimmg0, dmimma0, dmijts0}@uib.es

Resumen

En esta comunicación se estudia la ley de importación para las QL-implicaciones y las D-implicaciones, tanto en el caso de funciones de implicación definidas sobre el intervalo unidad como en el caso de implicaciones discretas (definidas sobre una cadena finita). Se caracterizan las implicaciones que verifican esta propiedad en el caso continuo (suave para las implicaciones discretas) y se estudian algunos casos no continuos.

Palabras Clave: QL-implicaciones, D-implicaciones, Ley de importación, Implicaciones discretas, Continuidad, Suavidad.

1 INTRODUCCIÓN

Las funciones de implicación desempeñan un papel esencial en la lógica borrosa y el razonamiento aproximado, así como en la mayoría de campos en los que esta teoría es aplicada (ver por ejemplo [13]). Muchas propiedades relativas a estas funciones de implicación o implicaciones (borrosas) han sido ampliamente estudiadas por diversos autores. Una de ellas es la llamada *ley de importación*,

$$(p \wedge q) \to r \equiv p \to (q \to r),$$

o equivalentemente en el contexto de la lógica borrosa,

$$I(T(x,y),z) = I(x,I(y,z)) \qquad (1)$$

para cualesquiera $x, y, z \in [0,1]$, donde T es una t-norma e I es una implicación.

Debido a la conmutatividad de T, la ley de importación asegura que la implicación correspondiente ha de verificar necesariamente el principio de intercambio que es una propiedad crucial para implicaciones borrosas. Además, algunas posibles aplicaciones de la ley de importación han sido puestas de manifiesto en el artículo [2].

A pesar de su interés, sólo unos pocos resultados sobre ella son conocidos en la literatura. Concretamente, un caso particular (tomando como T el producto) fue inicialmente estudiada en [18] donde se propuso como un axioma más en la definición de implicación. En [1], fue estudiada conjuntamente con la ley distributiva

$$I(x,T(y,z)) \equiv T(I(x,y),I(x,z))$$

caracterizando las implicaciones que satisfacen ambas propiedades al mismo tiempo. Recientemente en [2] la ley de importación fue estudiada con detalle para algunas clases de implicaciones, incluyendo las S, R y QL-implicaciones derivadas de t-normas y t-conormas. Sin embargo en este último caso sólo se consiguieron algunos resultados parciales.

En esta comunicación queremos completar estos resultados para QL-implicaciones y ampliar este estudio a las D-implicaciones, tanto en el contexto de implicaciones definidas en el intervalo $[0,1]$, como en el caso de implicaciones discretas, es decir, definidas sobre una cadena finita. En ambos casos, caracterizaremos las QL y D-implicaciones que satisfacen la ley de importación con algunas condiciones de continuidad (suavidad en el caso discreto) y discutiremos algunos casos particulares en ausencia de continuidad.

2 QL Y D-IMPLICACIONES

Daremos por conocida la teoría básica de t-normas y t-conormas (ver por ejemplo [8]). Recordaremos solamente algunos hechos y definiciones sobre los QL y D-operadores derivados a partir de t-normas y t-conormas.

Definición 1 *Una operación* $I : [0,1] \times [0,1] \to [0,1]$

se dice que es un operador de implicación o simplemente una implicación si verifica:

I1) I es decreciente en la primera variable y creciente en la segunda.

I2) $I(0,0) = I(1,1) = 1 \quad e \quad I(1,0) = 0.$

Notemos que de la definición se deriva inmediatamente que $I(0,x) = 1$ y $I(x,1) = 1$ para todo $x \in [0,1]$ mientras que los valores de $I(x,0)$ e $I(1,x)$ no se derivan de la definición.

Tres propiedades especialmente importantes en implicaciones son:

i) Contraposición respecto de una negación fuerte N:

$$I(x,y) = I(N(y), N(x)), \quad x, y \in [0,1].$$

ii) Principio de intercambio:

$$I(x, I(y,z)) = I(y, I(x,z)), \quad x, y, z \in [0,1].$$

iii) Principio de neutralidad (por la izquierda):

$$I(1,y) = y \qquad y \in [0,1].$$

Definición 2 *Sean T una t-norma, S una t-conorma y N una negación fuerte. Llamaremos*

i) QL-operador *a toda operación binaria sobre $[0,1]$ definida por*

$$I_{S,N,T}(x,y) = S(N(x), T(x,y)) \tag{2}$$

para cualesquiera $x, y \in [0,1]$.

ii) D-operador a toda operación binaria sobre $[0,1]$ definida por

$$I^{S,N,T}(x,y) = S(T(N(x), N(y)), y) \tag{3}$$

para cualesquiera $x, y \in [0,1]$.

Es conocido que ni los QL ni los D-operadores así definidos son siempre implicaciones y que está abierta todavía una caracterización de las S, N, T para las cuales el correspondiente QL o D-operador es una implicación según la definición 1. Sí son conocidas, en cambio, condiciones necesarias por un lado y suficientes por el otro (ver [11]). Recordaremos aquí algunas, pero antes establezcamos ciertas notaciones. Se sabe que cualquier negación fuerte es en realidad isomorfa a la negación standard dada por la asignación $x \to 1 - x$ (ver [16]) y consecuentemente viene dada por

$$N_\varphi(x) = \varphi^{-1}(1 - \varphi(x)) \text{ para todo } x \in [0,1]$$

donde $\varphi : [0,1] \to [0,1]$ es una biyección creciente.

Dada una operación binaria $S : [0,1]^2 \to [0,1]$, denotaremos por S_φ la φ-conjugada de S, es decir,

$$S_\varphi(x,y) = \varphi^{-1}(S(\varphi(x), \varphi(y)))$$

para todos $x, y \in [0,1]$. Con estas notaciones, una condición necesaria para que el QL-operador $I_{S,N,T}$ (igualmente el D-operador $I^{S,N,T}$) sea una implicación (y en tal caso la llamaremos QL y D-implicación respectivamente) es que se verifique

$$S(x, N(x)) = 1 \quad \text{para todo} \quad x \in [0,1]. \tag{4}$$

Además, si S es una t-conorma continua, dicha condición implica la existencia de una biyección creciente φ de forma que $S = (S_{\mathbf{L}})_\varphi$, donde $S_{\mathbf{L}}$ denota la t-conorma de Łukasiewicz y $N \geq N_\varphi$ (ver [11]).

Si bien en [2] el autor pudo caracterizar todas las R y S-implicaciones que verifican la ley de importación con una t-norma T_1, en el caso de QL-implicaciones sólo unos resultados parciales son conocidos (ver el siguiente teorema) y en el caso de las D-implicaciones no existen estudios previos.

Teorema 1 *Sea $\varphi : [0,1] \to [0,1]$ una biyección creciente, $S = (T_{\mathbf{L}})_\varphi$, $N = N_\varphi$ y T una t-norma. Si T es el mínimo o la φ-conjugada de la t-norma producto o de la t-norma de Łukasiewicz, el correspondiente I_{QL} es siempre una QL-implicación y,*

i) Si T es la φ-conjugada de la t-norma de Łukasiewicz, entonces I_{QL} verifica la ley de importación con una t-norma T' si y sólo si T' es el mínimo.

ii) Si T es la φ-conjugada de la t-norma producto, entonces I_{QL} verifica la ley de importación con una t-norma T' si y sólo si $T' = T$.

iii) Si T es el mínimo, entonces I_{QL} verifica la ley de importación con una t-norma T' si y sólo si T' es la φ-conjugada de la t-norma de Łukasiewicz.

Recordemos también las implicaciones fuertes o S-implicaciones, que están muy relacionadas tanto con las QL como con las D-implicaciones, así como su comportamiento con la ley de importación. Dada una t-conorma S y una negación fuerte N, la S-implicación derivada viene dada por

$$I_{S,N}(x,y) = S(N(x), y)) \quad \text{para todos} \quad x, y \in [0,1].$$

La siguiente caracterización de las S-implicaciones será utilizada más adelante. Fue dada inicialmente en [17], ver también [4] y [13] para la versión actual.

Teorema 2 *Sea $I : [0,1]^2 \to [0,1]$ una función. I es una S-implicación si y sólo si verifica $I1$), $I2$), los principios de intercambio y neutralidad y contraposición respecto de una negación fuerte N.*

Finalmente, la ley de importación para las S-implicaciones ha sido estudiada en [2] obteniendo el siguiente resultado.

Teorema 3 *Una S-implicación $I_{S,N}$ verifica la ley de importación con una t-norma T si y sólo si T es la N-dual de S.*

3 LEY DE IMPORTACIÓN EN EL CASO $[0,1]$

Empezaremos con algunos resultados generales para QL y D-operadores. Notemos primero que ambos tipos de operadores están muy relacionados siendo uno la contraposición del otro respecto de la negación N. Esto es:

$$I_{S,N,T}(x,y) = I^{S,N,T}(N(y),N(x))$$

para todos $x, y \in [0,1]$. Consecuentemente, tenemos que los QL y D-operadores $I_{S,N,T}$ y $I^{S,N,T}$ verifican el principio de contraposición respecto de N si y sólo si coinciden, esto es, si $I_{S,N,T} = I^{S,N,T}$. Esta propiedad está muy relacionada también con la ley de importación como se ve en el siguiente resultado.

Proposición 1 *Sean S una t-conorma, T una t-norma y N una negación fuerte. Si el operador $I_{S,N,T}$ (o $I^{S,N,T}$) verifica la ley de importación con una t-norma T_1 entonces verifica contraposición respecto de N.*

Demostración: Notemos que

$$I_{S,N,T}(x,0) = S(N(x),T(x,0)) = S(N(x),0) = N(x)$$

y análogamente

$$I^{S,N,T}(x,0) = S(T(N(x),1)0) = N(x).$$

Si denotamos por I cualquiera de los dos operadores tenemos que, si I verifica la ley de importación con una t-norma T_1, verifica también el principio de intercambio y entonces,

$$I(N(y),N(x)) = I(N(y),I(x,0)) =$$

$$I(x,I(N(y),0)) = I(x,y).$$

■

Corolario 1 *Sean S una t-conorma, T una t-norma y N una negación fuerte. Entonces, $I_{S,N,T}$ verifica la ley de importación con una t-norma T_1 si y sólo si la verifica $I^{S,N,T}$, en cuyo caso ambos operadores coinciden. Si además $I_{S,N,T} = I^{S,N,T}$ es una implicación, necesariamente coincide con una S-implicación.*

Demostración: Por el teorema anterior, si $I_{S,N,T}$ verifica la ley de importación con una t-norma T_1 entonces verifica contraposición respecto de N y por tanto coincide con $I^{S,N,T}$. El mismo razonamiento es válido si $I^{S,N,T}$ verifica la ley de importación con una t-norma T_1.

Si además $I_{S,N,T} = I^{S,N,T}$ es una implicación verifica $I1$) e $I2$), por lo que utilizando el Teorema 2 sólo hace falta ver que verifican el principio de neutralidad. Efectivamente,

$$I_{S,N,T}(1,y) = S(0,T(1,y)) = y$$

y análogamente

$$I^{S,N,T}(1,y) = S(T(0,N(y)),y) = S(0,y) = y.$$

■

A partir de este corolario es suficiente estudiar la ley de importación para QL-operadores y automáticamente se tiene solucionado también el caso de D-operadores.

3.1 Caso en que $S = (S_{\mathbf{L}})_\varphi$ y $N = N_\varphi$.

Como se ha comentado anteriormente para que un QL o un D-operador sea una implicación, la t-conorma S ha de verificar la condición (4). En el caso en que S es una t-conorma continua, dicha condición implica la existencia de una biyección creciente φ de forma que $S = (S_{\mathbf{L}})_\varphi$ y $N \geq N_\varphi$.

En esta sección vamos a estudiar el caso en que $N = N_\varphi$. Para esta negación unos cálculos sencillos demuestran que los QL y D-operadores se escriben respectivamente como,

$$I_{S,N,T}(x,y) = \varphi^{-1}(1 - \varphi(x) + \varphi(T(x,y)))$$

y

$$I^{S,N,T}(x,y) = \varphi^{-1}(\varphi(T(N_\varphi(x),N_\varphi(y))) + \varphi(y))$$

para todos $x, y \in [0,1]$, respectivamente. Dado que ambos operadores quedan totalmente determinados por la función φ y la t-norma T, en todo este apartado denotaremos los QL y D-operadores correspondientes por $I_{\varphi,T}$ y $I^{\varphi,T}$ respectivamente.

En este caso podemos completar el Teorema 1 con respecto a una t-norma cualquiera. Para ello recordemos cómo son las t-normas de Frank.

Definición 3 *(Ver por ejemplo [8]) Una t-norma T es una t-norma de Frank si pertenece a la familia $(T_\lambda)_{\lambda\in[0,+\infty]}$, dada por*

$$T_\lambda(x,y)=\begin{cases}\min(x,y) & si\ \lambda=0\\ T_{\mathbf{P}}(x,y) & si\ \lambda=1\\ T_{\mathbf{L}}(x,y) & si\ \lambda=\infty\\ log_\lambda\big(1+\frac{(\lambda^x-1)(\lambda^y-1)}{\lambda-1}\big) & en\ otro\ caso\end{cases}$$

donde $T_{\mathbf{P}}$ denota la t-norma producto.

Teorema 4 *Sea $\varphi:[0,1]\to[0,1]$ una biyección creciente y T una t-norma continua. Entonces $I_{\varphi,T}$ (o $I^{\varphi,T}$) verifica la ley de importación con una t-norma T_1 si y sólo si T es la φ-conjugada de una t-norma de Frank T_λ. En este caso, $I_{\varphi,T}$ e $I^{\varphi,T}$ son siempre implicaciones, coinciden $I_{\varphi,T}=I^{\varphi,T}$ y verifican la ley de importación con una t-norma T_1 si y sólo si T_1 es la φ-conjugada de la t-norma de Frank con subíndice $1/\lambda$, con el convenio*

$$1/0=+\infty \quad y \quad 1/+\infty=0.$$

Demostración: Si $I_{\varphi,T}$ verifica la ley de importación con la t-norma T_1, verifica el principio de intercambio y entonces T es la φ-conjugada de una t-norma de Frank T_λ y $I_{\varphi,T}$ es una implicación (ver [11], Corolario 1 y Proposición 9). Lo mismo ocurre con el operador $I^{\varphi,T}$.

Recíprocamente, si T es la φ-conjugada de una t-norma de Frank T_λ entonces $I_{\varphi,T}$ coincide con la S-implicación derivada de la t-conorma φ-conjugada de $S_{1/\lambda}$ y la negación clásica $1-j$ (dada por $x\to 1-x$), donde $S_{1/\lambda}$ es la t-conorma de Frank $(1-j)$-dual de $T_{1/\lambda}$ (ver [3], Teorema 4). Aplicando ahora el Teorema 3, $I_{\varphi,T}$ verifica la ley de importación con la t-norma T_1 si y sólo si T_1 es la N-dual de la t-conorma φ-conjugada de $S_{1/\lambda}$. Esto es:

$$T_1(x,y)=N(\varphi^{-1}(S_{1/\lambda}(N(\varphi(x)),N(\varphi(y))))) =$$

$$\varphi^{-1}(1-S_{1/\lambda}(1-\varphi(x),1-\varphi(y)))=\varphi^{-1}(T_{1/\lambda}(\varphi(x),\varphi(y)))$$

y por tanto T_1 es la φ-conjugada de $T_{1/\lambda}$. ∎

Podemos hacer los siguientes comentarios referentes al teorema anterior:

1. Notemos que para los valores $\lambda=0,1,+\infty$ obtenemos los tres casos particulares estudiados en el Teorema 1 para QL-implicaciones.
2. Según resultados obtenidos en [11], se deduce que, para los QL y D-operadores considerados en esta sección, la ley de importación y el principio de intercambio son equivalentes.
3. La equivalencia anterior es también cierta para S-implicaciones y para R-implicaciones derivadas de t-normas continuas por la izquierda (ver [2]).
4. Más aún, en todos los casos conocidos en los que la ley de importación se verifica, existe una única t-norma T para la cual se satisface.

A partir de los comentarios anteriores surgen de forma inmediata las siguientes cuestiones:

- ¿Es cierta siempre la equivalencia entre la ley de importación y el principio de intercambio? Es decir, si una implicación I verifica el principio de intercambio, ¿existe una t-norma T_1 para la cual I verifica la ley de importación?

- ¿Es cierta la unicidad de la t-norma T_1 para la ley de importación?

Ambas cuestiones fueron ya planteadas en [7] como problemas abiertos. Se ha visto ya que la primera tiene respuesta negativa (ver [12]) puesto que existen implicaciones derivadas de uninormas que verifican intercambio y para las cuales no existe ninguna t-norma con la que se verifique la ley de importación. Sin embargo, sí que verifica dicha ley con una uninorma conjuntiva, lo que traslada el problema a la siguiente cuestión:

- Si una implicación I verifica el principio de intercambio, ¿existe una uninorma U para la cual I verifica la ley de importación?

3.2 Caso en que S no es continua.

En el caso en que S no es continua, existen muchas t-conormas que verifican la condición (4), por ejemplo la t-conorma máximo nilpotente. Dada una negación fuerte N, dicha t-conorma viene dada por

$$\max_N(x,y)=\begin{cases}\max(x,y) & \text{si } y<N(x)\\ 1 & \text{si } y\geq N(x).\end{cases}$$

En este caso tenemos el siguiente resultado referente a la ley de importación.

Teorema 5 *Sea N una negación fuerte, $\max_N$ el máximo nilpotente correspondiente, T una t-norma y $I_{N,T}$ ($I^{N,T}$) el correspondiente QL-operador (D-operador). Entonces, $I_{N,T}$ ($I^{N,T}$) verifica la ley de importación con una t-norma T_1 si y sólo si T es el mínimo. En este caso, $I_{N,T}=I^{N,T}=I$ donde I es la S-implicación dada por*

$$I(x,y)=\begin{cases}1 & si\ x\leq y\\ \max(N(x),y) & si\ x>y\end{cases} \qquad (5)$$

y verifica la ley de importación con una t-norma T_1 si y sólo si T_1 es el mínimo nilpotente $\min_N$.

Demostración: Por la Proposición 1 sabemos que si $I_{N,T}$ (o $I^{N,T}$) verifica la ley de importación con una t-norma T_1, también verifica contraposición respecto

de N. Pero entonces necesariamente T ha de ser el mínimo (ver [3], Teorema 5).

Recíprocamente, si T es el mínimo tenemos,

$$I_{N,T}(x,y) = \max_N(N(x), \min(x,y)) = \max_N(N(x), y))$$

mientras que

$$I^{N,T}(x,y) = \max_N(\min(N(x), N(y)), y) = \max_N(N(x), y)).$$

Es decir, ambas $I_{N,T}$ y $I^{N,T}$ coinciden con la S-implicación derivada de $\max_N$ y N (o sea, la implicación dada por la ecuación (5)). Así, el resultado se sigue del Teorema 3 y del hecho que la N-dual de $\max_N$ es el mínimo nilpotente $\min_N$. ∎

Existe una familia parametrizada de t-conormas que verifican la condición (4) y que generalizan el máximo nilpotente. Dada una negación fuerte N y $a \in [0,1]$ tal que $a \leq N(a)$, denotamos por $S_{N,a}$ la t-conorma dada por $S_{N,a}(x,y) =$

$$\begin{cases} 1 & \text{si } \ y \geq N(x) \\ a + (N(a)-a)S_0\left(\frac{x-a}{N(a)-a}, \frac{y-a}{N(a)-a}\right) & \text{si } x,y \in [a, N(a)] \\ & \text{y } y < N(x) \\ \max(x,y) & \text{otro caso} \end{cases} \tag{6}$$

donde S_0 es una t-conorma nilpotente con negación asociada N_0 dada por

$$N_0(x) = \frac{N((N(a)-a)x + a) - a}{N(a) - a}$$

para todo $x \in [0,1]$. Esta familia está formada por las t-conormas N-duales de la familia de t-normas caracterizada en [6]. Notemos que esta familia incluye tanto la propia t-conorma S_1 (caso $a = 0$) como el máximo nilpotente (caso $a = s$ siendo s el punto fijo de N).

Veamos ahora que los QL y D-operadores derivados de una t-conorma $S_{N,a}$, la negación N y la t-norma mínimo verifican la ley de importación.

Proposición 2 *Sea N una negación fuerte, $a \in [0,1]$ tal que $a \leq N(a)$ e I el QL-operador (D-operador) derivado de $S_{N,a}, N$ y la t-norma mínimo. Entonces, I verifica la ley de importación con una t-norma T_1 si y sólo si T_1 es la N-dual de $S_{N,a}$.*

Demostración: Es fácil ver que tanto el QL-operador como el D-operador vienen dados, en este caso, por

$$I(x,y) = \begin{cases} 1 & \text{si } x \leq y \\ S_{N,a}(N(x), y) & \text{si } y < x. \end{cases}$$

Notemos que, cuando $y \geq x$, también se tiene $S_{N,a}(N(x), y) = 1$, por lo que

$$I(x,y) = S_{N,a}(N(x), y) \quad \text{para todos } x, y \in [0,1].$$

Es decir, tanto el QL-operador como el D-operador coinciden con la S-implicación derivada de $S_{N,a}$ y N y el resultado es consecuencia otra vez del Teorema 3. ∎

Queda pendiente estudiar si existen otras t-normas T, diferentes del mínimo, para las cuales el QL o D-operador derivado de $S_{N,a}, N$ y T verifican la ley de importación.

4 LEY DE IMPORTACIÓN EN EL CASO DISCRETO

Este mismo estudio puede realizarse en el caso discreto, esto es, para operadores lógicos definidos en una cadena finita, campo cada día más en alza debido a sus aplicaciones en la computación con palabras. Para el estudio de operadores lógicos todas las cadenas finitas con el mismo número de elementos son equivalentes (ver [15]) por lo que trabajaremos con la cadena más simple de $n+1$ elementos:

$$L = \{0, 1, 2 \ldots, n-1, n\}$$

donde $n \geq 1$.

Recordemos que sobre L existe una y sólo una negación fuerte N, dada por

$$N(x) = n - x \quad \text{para todo } x \in L. \tag{7}$$

Para el estudio de t-normas y t-conormas sobre L se introduce la definición de suavidad. Dicho concepto es importante porque viene a ser el equivalente a la continuidad en el caso de $[0,1]$ y es también equivalente a la condición de divisibilidad ($x \leq y$ si y sólo si existe $z \in L$ tal que $T(y,z) = x$), ver por ejemplo [15]. La siguiente definición está adaptada de [5] (ver también [15]).

Definición 4 *Una función $f : L \to L$ se dice que es suave si verifica*

$$|f(x) - f(x-1)| \leq 1 \quad \textit{para todo} \quad x \in L \textit{ con } x \geq 1.$$

Un operador binario F sobre L se dice que es suave si lo es en cada variable.

Las t-normas suaves han sido caracterizadas como sumas ordinales de t-normas arquimedianas (tales que $T(x,x) < x$ para todo $x \in L \setminus \{0, n\}$) como sigue.

Proposición 3 (Ver [14] o [15]). *Existe una y sólo una t-norma suave arquimediana sobre L dada por*

$$T(x,y) = \max\{0, x+y-n\} \quad \textit{para todos } x, y \in L \tag{8}$$

conocida como la t-norma de Łukasiewicz. Además, dado un subconjunto J de L conteniendo $0, n$, existe

una y sólo una t-norma suave T_J sobre L cuyo conjunto de elementos idempotentes es J. De hecho, si

$$J = \{0 = i_0 < i_1 < \ldots < i_{m-1} < i_m = n\}$$

entonces T_J viene dada por $T_J(x,y) =$

$$\begin{cases} \max\{i_k, x+y-i_{k+1}\} & \text{si existe } i_k \in J \text{ tal que} \\ & x,y \in [i_k, i_{k+1}] \\ \min\{x,y\} & \text{en otro caso} \end{cases} \tag{9}$$

Las t-conormas suaves admiten un teorema de caracterización idéntico al de t-normas que puede ser fácilmente deducido por N-dualidad. Recordemos que la expresión de la única t-conorma arquimediana suave es la siguiente:

$$S(x,y) = \min\{n, x+y\} \quad \text{para todos } x,y \in L \tag{10}$$

y que también es conocida como la t-conorma de Łukasiewicz.

A partir de t-normas y t-conormas se pueden definir sobre L los QL y D-operadores de forma totalmente análoga al caso $[0,1]$. Dado que disponemos de una única negación fuerte sobre N, las expresiones en este caso serán:

$$I_{S,T}(x,y) = S(n-x, T(x,y)) \quad \text{para todos} \quad x,y \in L$$

para los QL-operadores y

$$I^{S,T}(x,y) = S(T(n-x, n-y), y) \quad \text{para todos} \quad x,y \in L$$

para los D-operadores. De forma similar al caso $[0,1]$ tenemos el siguiente resultado para QL y D-operadores en general.

Teorema 6 *Sean S una t-conorma y T una t-norma sobre L. Entonces, $I_{S,T}$ verifica la ley de importación con una t-norma T_1 si y sólo si la verifica $I^{S,T}$, en cuyo caso ambos operadores coinciden. Si además $I_{S,T} = I^{S,T}$ es una implicación, necesariamente coincide con una S-implicación.*

El estudio de cuándo los QL y D-operadores son implicaciones, según la definición 1 adaptada al caso L, ha sido realizado en [10]. Allí se demuestra que una condición necesaria para que dichos operadores sean implicaciones es que

$$S(x, n-x) = n \quad \text{para todo} \quad x \in L \tag{11}$$

y esta condición es también suficiente en el caso en que T y S sean suaves. Es más, en este caso dicha condición es equivalente a que S sea la t-conorma arquimediana dada por (10). De esta forma las QL y D-implicaciones correspondientes pueden escribirse sencillamente como

$$I_T(x,y) = n - x + T(x,y), \quad I^T(x,y) = y + T(n-x, n-y)$$

para todos $x,y \in L$ respectivamente. Además se demuestra en [10] que, si $T = T_J$ con

$$J = \{0 = i_0 < i_1 < \ldots < i_{m-1} < i_m = n\},$$

entonces se tiene

$$I_{T_J} = I^{T_{N(J)}}$$

donde

$$N(J) = \{0 < n - i_{m-1} < \ldots < n - i_1 < n\}.$$

En la Figura 1 puede verse la estructura de las QL-implicaciones y de las D-implicaciones para cada una de las t-normas suaves T_J.

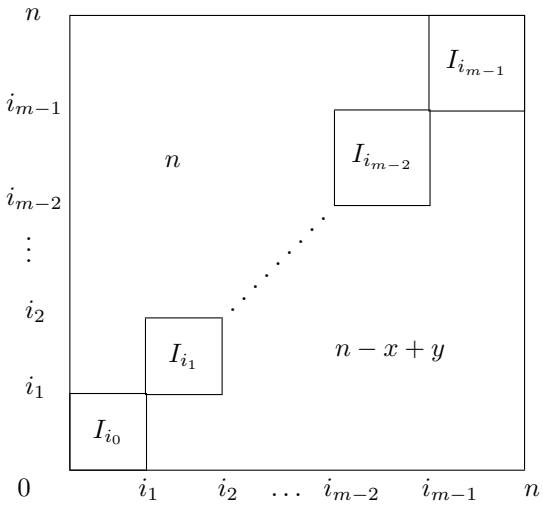


Figura 1. Estructura de la QL-implicación I^{T_J} donde $J = \{0 = i_0 < i_1 < \ldots < i_{m-1} < i_m = n\}$. Para cada $j = 0, ..., m-1$, I_{i_j} viene dada por $I_{i_j}(x,y) = \max\{n-x+i_j, n+y-i_{j+1}\}$ para todos $x,y \in [i_j, i_{j+1}]$. Notemos que ésta es también la estructura de la D-implicación correspondiente a la t-norma suave $T_{N(J)}$.

Así tenemos el siguiente resultado en el caso suave[1].

Teorema 7 *Sean T una t-norma y S una t-conorma suaves sobre L. Entonces, la QL-implicación I_T (D-implicación I^T) asociada verifica la ley de importación con una t-norma T_1 si y sólo si T es el mínimo o la t-norma arquimediana dada por (8). Además, en estos casos I_T e I^T coinciden y,*

[1]Para el caso discreto no incluimos las demostraciones por falta de espacio. Sin embargo, el lector podrá deducirlas fácilmente, puesto que son muy similares a las del caso $[0,1]$, utilizando en esta ocasión resultados de las QL y D-implicaciones discretas obtenidos en [10]

1. *Si* $T = \min$, $I_T = I^T$ *es la implicación de Łukasiewicz* $I(x,y) = \min(n, n-x+y)$ *y verifica la ley de importación con* T_1 *si y sólo si* T_1 *es la t-norma dada por (8).*

2. *Si* T *viene dada por (8),* $I_T = I^T$ *es la implicación de Kleene-Dienes* $I(x,y) = \max(n-x, y)$ *y verifica la ley de importación con* T_1 *si y sólo si* $T_1 = \min$.

4.1 Caso no suave

En el caso no suave, la condición (11) sigue siendo necesaria para la ley de importación, pero en este caso existen muchas otras t-conormas que la verifican. Una familia de t-conormas que la verifica es la dual (respecto de la única negación fuerte sobre L) de la familia de t-normas caracterizada en [9]. Concretamente, para cada $a \in L$ tal que $n - a \leq a$ tenemos la t-conorma dada por

$$S_a(x,y) = \begin{cases} n & \text{si } \ x+y \geq n \\ x+y-n+a & \text{si } \ x+y < n \\ & \text{y } \ n-a \leq x,y \leq a \\ \max\{x,y\} & \text{otro caso.} \end{cases} \tag{12}$$

Notemos que S_a no es suave excepto en el caso $a = n$ en el cual S_n coincide con la t-conorma de Łukasiewicz. Por otra parte, el máximo nilpotente se obtiene en el otro caso extremo en el cual $n-a = \lfloor n/2 \rfloor$ donde $\lfloor n/2 \rfloor$ representa la parte entera de $n/2$.

Ejemplo 1 *Sean* $I_{\min,S_a}$ *y* $I^{\min,S_a}$ *el QL y el D-operador derivados de la t-conorma* S_a *y el mínimo. Es fácil ver que ambos operadores son en realidad implicaciones, coinciden y vienen dados por:* $I_{\min,S_a}(x,y) = I^{\min,S_a}(x,y) =$

$$\begin{cases} n & si \ \ x \leq y \\ a+y-x & si \ \ n-a \leq y < x \leq a \\ \max\{n-x, y\} & otro \ caso \end{cases} \tag{13}$$

Las estructuras de la t-conorma S_a y la QL y D-implicación asociada a S_a y el mínimo pueden verse en la Figura 2.

Para esta familia tenemos el siguiente resultado análogo al caso $[0,1]$.

Proposición 4 *Sea* $a \in L$ *tal que* $n - a = \lfloor n/2 \rfloor$. *Entonces* $I_{\min,S_a} = I^{\min,S_a}$ *verifica la ley de importación con una t-norma* T_1 *si y sólo si* T_1 *es la dual de* S_a *respecto de la única negación fuerte sobre* L, $N(x) = n - x$.

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto nacional del Ministerio de Educación y Ciencia MTM2006-05540 y el proyecto del Gobierno de las Islas Baleares PCTIB-2005GC1-07.

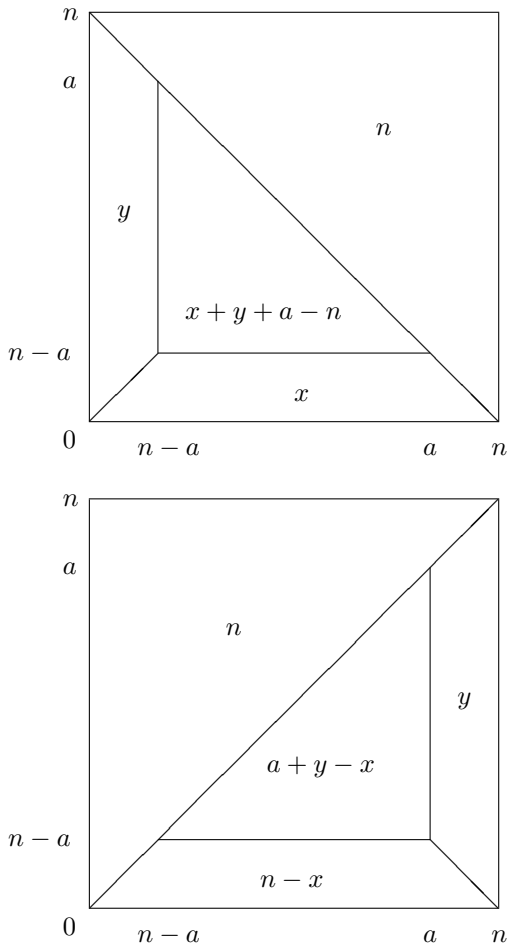


Figura 2. La t-conorma S_a (arriba) y la QL y D-implicación asociada $I_{\min,S_a} = I^{\min,S_a}$ (abajo).

Referencias

[1] M. Baczyński. On a class of distributive fuzzy implications, *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, Vol. 9, Pág. 229-238, 2001.

[2] J. Balasubramaniam (2008), On the law of importation $(x \wedge y) \to z \equiv (x \to (y \to z))$ in fuzzy logic, *IEEE Transactions on Fuzzy Systems*, vol. 16, 130-144.

[3] J.C. Fodor (1995), Contrapositive symmetry on fuzzy implications, *Fuzzy Sets and Systems*, Vol. 69, 141–156.

[4] J.C. Fodor and M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support, *Theory and Decision Library, serie D: System*

theory, Knowledge Engineering and Problem Solving, Kluwer Academic Publishers, Dordrecht, 1994.

[5] Ll. Godo and C. Sierra, A new approach to connective generation in the framework of expert systems using fuzzy logic, Proceedings of the XVIIIth ISMVL, Palma de Mallorca, 1988, pp. 157-162.

[6] S. Jenei (2000), New family of triangular norms via contrapositive symmetrization of residuated implications, *Fuzzy Sets and Systems*, Vol. 110, 157-174.

[7] E.P. Klement and R. Mesiar (2006), Open problems posed at the eighth international conference on fuzzy set theory and aplications (FSTA 2006, Liptovský Ján, Slovakia), *Kybernetika*, **42**, pp. 225-235.

[8] E.P. Klement, R. Mesiar and E. Pap, Triangular norms, Kluwer Academic Publishers, Dordrecht (2000).

[9] M. Mas, M. Monserrat and J. Torrens (2004), S-implications and R-implications on a finite chain, *Kybernetika*, **40**, pp. 3-20.

[10] M. Mas, M. Monserrat and J. Torrens (2005), On two types of discrete implications, *International Journal of Approximate Reasoning*, vol. 40, 262-279.

[11] M. Mas, M. Monserrat and J. Torrens (2006), QL-implications versus D-implications, *Kybernetika*, vol. 42 (3), 351-366.

[12] M. Mas, M. Monserrat and J. Torrens (2008), On the law of importation for some kinds of fuzzy implications derived from uninorms, in *Proceedings of IPMU-08*, to appear.

[13] M. Mas, M. Monserrat, J. Torrens and E. Trillas (2007), A survey on fuzzy implication functions, *IEEE Transactions on Fuzzy Systems*, vol. 15(6), 1107-1121.

[14] G. Mayor and J. Torrens (1993), On a class of operators for expert systems, *Int. J. of Intelligent Systems*, **8** N. 7, pp. 771-778.

[15] G. Mayor and J. Torrens (2005), Triangular norms in discrete settings, in E.P. Klement and R. Mesiar, Editors, *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier, pp. 193-236.

[16] E. Trillas (1979), Sobre funciones de negación en la teoría de conjuntos difusos, *Stochastica*, vol 3, 47-60.

[17] E. Trillas and L. Valverde (1981), On some functionally expressable implications for fuzzy set theory, in *Proceedings of 3rd International Seminar on Fuzzy Set Theory*, E.P. Klement, Ed., Linz, Austria, 1981, pp. 173-190.

[18] I. Türksen, V. Kreinovich and R. Yager (1998), A new class of fuzzy implications. Axioms of fuzzy implications revisited, *Fuzzy Sets and Systems*, Vol. 100, 267-272.

On the fuzzy law $(\mu \to \mu') \to \mu' = \mu_1$

Enric Trillas Itziar García-Honrado Eloy Renedo

European Centre for Soft-Computing. Mieres, Asturias (España)
{enric.trillas, itziar.garcia, eloy.renedo}@softcomputing.es

Abstract

This paper is devoted to study the law in its title when the arrows are represented by different T-conditional functions.

1 Introduction

The law with fuzzy sets $(\mu \to \mu') \to \mu' = \mu_1$, comes from the tautology $\vdash (A \to \sim A) \to \sim A$, called *reductio ad absurdum* principle, and proved by Jan Łukasiewicz in his axiomatic formal system for propositional logic ([3])[1]. Before Łukasiewicz, Whitehead and Russell ([6]) also presented the equivalent principle $\vdash (\sim A \to A) \to A$.

The law $(A \to A') \to A' = 1$ is verified in all boolean algebras with the classical material conditional, and in the orthomodular lattices with the Sasaki's and the Dishkant's arrows, but not in any lattice with the intersection operator. Indeed,

- With $A \to B = A' + B$, it is $(A' + A')' + A' = A + A' = 1$
- With $A \to B = A' + A{\cdot}B$, it is $(A \to A') \to A' = (A' + A{\cdot}A') \to A' = A' \to A' = A + A'{\cdot}A' = A + A' = 1$. Similar result follows with $A \to B = B + A'{\cdot}B'$
- Nevertheless, with $A \to B = A \cdot B$, it is $(A \cdot A') \cdot A' = 0$.

[1]Let us remark that it does not be confused with the expression 'reductio to absurd', the argumentative method widely used in mathematics, whose logic schema can be stated as $B, \sim A \to \sim B \vdash A$ (see [4] for a fuzzy-like interpretation of this schema) that, in the light of deduction theorem, can be expressed as $\sim A \to \sim B \vdash B \to A$ or $\vdash (\sim A \to \sim B) \to (B \to A)$, formula proposed as the Axiom of Contraposition in the mentioned Łukasiewicz's system.

The generalization of that arrow functions is given, in fuzzy logic, by operators $\to$, functionally expressible by numerical functions $J : [0,1] \times [0,1] \to [0,1]$, that is, such that $(\mu \to \sigma)(x,y) = J(\mu(x), \sigma(y))$.

Among the most widely used models to represent functions J, there are the following,

1. *Residuated or R-implications*:
   $J_T(a,b) = sup\{z \in [0,1] : T(a,z) \leq b\}$. T is a continuous t-norm.
2. *Strong or S-implications*:
   $J(a,b) = S(N(a),b)$. S is a continuous t-conorm, and N is a strong negation.
3. *Quantum Logic or QM-operators*:
   $J(a,b) = S(N(a), T(a,b))$, where S and T are a continuous t-conorm and t-norm, respectively, and N is a strong negation.
4. *Dishkant or D-operators*:
   $J(a,b) = S(b, T(N(a), N(b)))$, where S and T are a continuous t-conorm and t-norm, respectively, and N is a strong negation.
5. *Mamdani-Larsen or ML-operators*:
   $J(a,b) = T(\varphi_1(a), \varphi_2(b))$, where φ_1 is an order automorphism on $[0,1]$ and $\varphi_2 : [0,1] \to [0,1]$ is a non-null contractive mapping (i.e., $\varphi_2(a) \leq a$ for all $a \in [0,1]$), and T is a continuous t-norm.

Any negation function N can be written in the form $N_\varphi(a) = \varphi^{-1}(1 - \varphi(a))$, for some order automorphism φ of the unit interval $[0,1]$.

Functions 1 and 2 come from the classical conditional, and 3, 4 and 5 from the Sasaki, Dishkant and intersection arrow, respectively.

Remark 1 *Functions 1 and 2 verify the characteristic implication's properties (see, for example, [1]), and so they can be named 'implications' properly.*

With fuzzy sets μ in $[0,1]^X$, we can deal with the equation $(\mu \to \mu') \to \mu' = \mu_1$ ($\mu_1 = 1$, for all $x \in X$), once translated into the functional equation $J(J(a, N(a)), N(a)) = 1$, for all a in $[0,1]$, and a T-conditional J.

The equation does not hold with J an R-implication, or an S-implication ([5]). The equation does not hold also with $J(a,b) = T(a,b)$, a Mamdani-Larsen min-conditional ([5]). In general, formula $(A \to A') \to A' = A_1$ also does not hold with QM-operators, for example, with the so-called Early-Zadeh one, $J(a,b) = \max(1-x, \min(a,b))$, it results $J(J(0.3, 0.7), 0.7) = 0.7 \neq 1$.

This paper just poses the question if the formula $(\mu \to \mu') \to \mu' = \mu_1$ can hold with two different T-conditionals, that is, it tries to search for solutions of the equation $J_2(J_1(a, N(a)), N(a)) = 1$, for two T_i-conditionals J_i $(i = 1, 2)$.

That is, the aim of this paper is to find solutions for the Pexider's equation

$$J_2(J_1(a, N(a)), N(a)) = 1, \tag{1}$$

combining different T-conditionals for J_2 and J_1, such as R-implication, S-implication, Mamdani operators, QM-operators and D-operators. To completely solve this equation is obviously difficult if not imposible.

2 Refusal of some J_1 and J_2

Let us begin by finding which of the five above listed operators cannot ever be used either for J_1 or J_2.

2.1 J_2 neither can be a Mamdani operator nor a force operator

This is the first possibility to refuse, since if $J_2(a,b) = T(a,b)$, for any t-norm T, we have $J_2(J_1(a, N(a)), N(a)) = T(J_1(a, N(a)), N(a)) \leq N(a)$, and $N(a) < 1$ for all $x \neq 0$. So, (1) is not verified.

It is impossible to get $T(a,b) = 1$, when $a, b < 1$, so it can be also refused the case that J_2 is a force implication ([2]), $J_2(a,b) = T(a, E(a,b))$ with E a T-indistinguishability, or $J_2(a,b) = T(a, 1-d(a,b))$ with d a distance.

2.2 J_2 cannot be a D-operator with the t-conorms Max or Prod*

Using as J_2 a D-operator, the formula (1) becomes in

$$S(N(a), T(N_1(J_1(a, N(a))), N_1(N(a)))) = 1.$$

For $x \in (0,1)$ it is , $N_1(N(a)) < 1$, so,

$$T(N_1(J_1(a, N(a))), N_1(N(a))) \leq$$

$$\min(N_1(J_1(a, N(a))), N_1(N(a))) \leq N_1(N(a))) < 1.$$

And since $N(a) < 1$ and S is an strict t-conorm (Max or Prod*), it is $S(N(a), T(N_1(J_1(a, N(a))), N_1(N(a)))) < 1$.

Further on it would be proven that using a D-operator with the t-conorm of Łukasiewicz we get some solutions for (1).

2.3 J_1 cannot be an R-implication

Let x such that $x < N(a)$ and $x \neq 0$, then $J_1(a, N(a)) = 1$, since J_1 is a R-implication. So, $J_2(J_1(a, N(a)), N(a)) = J_2(1, N(a)) = N(a)$ for all T-conditional J_2, that is 1 only for $x = 0$.

2.4 J_1 cannot be always an S-implication with $S = W^*_\varphi$ or $S = Prod^*_\varphi$

In this case (1) becomes

$$J_2(S(N_1(a), N(a)), N(a)) = 1.$$

- Taking $S = W^*_\varphi$, $N_1 = N = N_\varphi$ and $x = \varphi^{-1}(0.5)$, we have $J_1(a, N(a)) = \varphi^{-1}(\min(1, 0.5 + 0.5)) = 1$. So, $J_2(J_1(a, N(a)), N(a)) = N(a) = \varphi^{-1}(0.5) \neq 1$.
- Taking $S = Prod^*_\varphi$ and $N = N_1 = 1 - id$, as $J_1(a, N(a)) = 1 - (\varphi(a))^2$ it is easy to find examples for each type of J_2 that do not satisfy (1).

Below we will see that with S=Max some solutions are found.

2.5 J_1 cannot be always a QM-operator with the t-conorm $Prod^*$ or W^* and t-norm Min or $Prod$

Taking J_1 a QM-operator, we have the equation

$$J_2(S(N_1(a), T(a, N(a))), N(a)) = 1.$$

- Taking $S = Prod^*$, $T = \min$ and $N = N_1 = 1 - id$, (1) is not true for any J_2, since it is easy to find negative examples for each type of operator. For example with J_2 a R-implication$J_2(J_1(a, N(a)), N(a)) = J_2(N(a) + x - N(a) \cdot x, N(a)) < 1$, since $N(a) + x - N(a) \cdot x > N(a)$.
- Taking $S = Prod^*$, $T = Prod$ and $N = N_1 = 1 - id$, it is also easy to find many cases in which (1) does not hold.

- Taking $S = W^*_\varphi$, $T = \min_\varphi$, $N = N_1 = 1 - id$ and $x = \varphi^{-1}(0.5)$, we get $J_1(a, N(a)) = \varphi^{-1}\min(1, \varphi(N(a)) + \varphi(\min_\varphi(a, N(a)))) = \min(1, 0.5 + 0.5) = 1$. Therefore $J_2(J_1(a, N(a)), N(a)) = N(a) = 0.5 \neq 1$.

- Taking $S = W^*$, $T = \min$ and $N = N_1 = 1 - id$, there are also very simple examples in which (1) does not hold.

Then, only in some cases (1) is verified for $S = W^*$ and $T = W$.

3 Some solutions

From the cases above we get an indication of which operators we should not use in order that equation (1) does hold. But, combining J_1 (that can be a Mamdani-operator, some S-implication, D-operator or QM-operator) with J_2 (that can be an R-implication, an S-implication some QM-operator or D-operator) we get the following solutions.

3.1 If J_1 is a Mamdani-Larsen operator

In this case equation (1) turns into

$$J_2(T(a, N(a)), N(a)) = 1, \tag{2}$$

and it holds in the following cases:

- J_2 is any R-implication.

  $J^R(T(a, N(a)), N(a)) = 1$, since $T(a, N(a)) \leq N(a)$ for any T and N.

- J_2 is the W^*_φ-S-implication, $J_2(a, b) = W^*_\varphi(N_1(a), b)$, and $N_1 \geq N_\varphi$.

  Indeed, $\varphi^{-1}(\min(1, \varphi(N_1(T(a, N(a)))) + \varphi(N(a)))) = 1 \Rightarrow N_1(T(a, N(a))) \geq \varphi^{-1}(1 - \varphi(N(a)) = N_\varphi(N(a))$ that holds for all x in $[0, 1]$ if $N_\varphi \leq N_1$.

- J_2 is the QM-operator, $Q(x, y) = W^*_\varphi(N_1(a), \min(a, b))$, or the D-operator, $D(a, b) = W^*_\varphi(b, \min(N_1(a), N_1(b)))$, and $N_\varphi \leq N_1$.

  From equality (2) it follows, $W^*_\varphi(N_1(T(a, N(a))), \min(T(a, N(a)), N(a))) = W^*_\varphi(N_1(T(a, N(a))), T(a, N(a))) = 1$, if $N_1 \geq N_\varphi$.

  Analogously,

  $W^*_\varphi(N(a), \min(N_1(T(a, N(a))), N_1(N(a)))) = W^*_\varphi(N(a), N_1(N(a))) = 1$, that holds if $N_1 \geq N_\varphi$

In the particular case in which J_1 is of the Mamdani-Larsen type with $T = W$, there are the following additional solutions. The equation (2) particularizes to

$$J_2(W_\varphi(a, N(a)), N(a)) = 1 \tag{3}$$

- J_2 is a S-implication, and $N \leq N_\varphi$.

  Taking $J_2(a, b) = S(N_1(a), b)$, the equation (3) gives $S(N_1(W_\varphi(a, N(a))), N(a)) = 1$.

  1) If S is either a strict t-conorm or Max, since in $(0, 1]$ is $N(a) \neq 1$, then $N_1(W_\varphi(a, N(a))) = 1$, that is, $W_\varphi(a, N(a)) = 0$ and, therefore, $N \leq N_\varphi$.

  2) If S is a non-strict archimedean t-conorm, from the above conditions follows that $W^*_\psi(N_1(W_\varphi(a, N(a))), N(a)) = W^*_\psi(1, N(a)) = 1$.

- J_2 is a QM-operator and $N \leq N_\varphi$.

  Now, with $J_2(a, b) = S_\psi(N_1(a), T(a, b))$ the equation (3) is

  $S_\psi(N_1(W_\varphi(a, N(a)), T(W_\varphi(a, N(a)), N(a)))) = 1$.

  If S is either a strict t-conorm or Max, since in $(0, 1]$ it is $N(a) < 1$, then $T(W_\varphi(a, N(a)), N(a))) < 1$, and, therefore, $W_\varphi(a, N(a)) = 0$ that is, $N \leq N_\varphi$.

  From the above conditions, $S(N(W_\varphi(a, N(a)), T(W_\varphi(a, N(a)), N(a)))) = 1$, for any continuous t-conorm S.

3.2 If J_1 is a Kleene-Dienes implication

Taking $J_1(a, b) = \max(N_1(a), b)$, the Equation (1) is

$$J_2(\max(N_1(a), N(a)), N(a)) = 1, \tag{4}$$

and it will hold in following cases,

- J_2 is an R-implication and $N_1 \leq N$.

  Provided that $N_1 \leq N$, it is $\max(N_1(a), N(a)) = N(a)$, therefore $J_R(N(a), N(a)) = 1$.

- J_2 is the QM-operator, $J_2(x, y) = W^*_\varphi(N_2(x), T(x, y))$, and $N_2 \geq N_\varphi$.

  Provided that $J_2(a, b) = S(N_2(a), T(a, b))$, the equation (4) gives $S(N_2(\max(N_1(a), N(a))), T(\max(N_1(a), N(a)), N(a))) = 1$.

  Notice that $N_2(\max(N_1(a), N(a))) \leq N_2(N(a))$ and $T(\max(N_1(a), N(a)), N(a))) \leq N(a)$, therefore $1 \leq S(N_2(N(a)), N(a)$, which forces $S = W^*_\varphi$ with .

3.3 If J_1 is a QM-operator

The substitution of J_1 by a QM operator in the equation (1) gives

$$J_2(S(N_1(a), T(a, N(a))), N(a)) = 1 \qquad (5)$$

and two groups of solutions can be found

3.3.1 $S = \mathbf{Max}$

the equation (5) is therefore

$$J_2(\max(N_1(a), T(a, N(a))), N(a)) = 1 \qquad (6)$$

- J_2 is an R-implication and $N_1 \leq N$.

  $J_R(\max(N_1(a), T(a, N(a)), N(a)) = 1$, since $T(a, N(a))$ and $N_1(a)$ are both less or equal than $N(a)$ and so $\max(N_1(a), T(a, N(a)) \leq N(a)$.

- J_2 is a QM-operator such a way $J_2(x, y) = S_2(N_2(x), T_2(x, y))$. The equation (6) turns into $S_2(N_2(\max(N_1(a), T(a, N(a)))), T_2(\max(N_1(a), T(a, N(a))), N(a))) = 1$.

  Both with $N_1(a) \leq T(a, N(a))$ and vice versa, follows that the necessary condition is $S_2 = W^*_\varphi$ and $N_2 \geq N_\varphi$

  A particular solution is obtained when $\varphi = id$, and $N_2 = N_1 = N$, with such values $T_2 = \min$

- J_2 is a D-operator such a way $J_2(x, y) = S_2(y, T_2(N_2(x), N_2(y)))$.

  An analogous set of solutions as in the above case of QM is obtained.

3.3.2 $T = W$

If J_1 is a QM-operator with $T = W$, J_2 could be a R-implication. Let us suppose that $J_R(S_\psi(N(a), W_\varphi(a, N(a)), N(a)) = 1$, in the case of $S_\psi = Prod^*_\psi$, from $J_R(Prod^*_\psi(N(a), W_\varphi(a, N(a)), N(a)) = 1$ follows $Prod^*_\psi(N(a), W_\varphi(a, N(a)) \leq N(a)$ that is $Prod^*_\psi(N(a), W_\varphi(a, N(a)) \leq N(a)$, which its only solution for all x is $W_\varphi(a, N(a)) = 0$, that is, $N \leq N_\varphi$. In this way, $J_R(S_\psi(N(a), 0), N(a)) = J_R(N(a), N(a)) = 1$.

From now on, in this subsection we suppose N_1(negation of J_1)= N_2(negation of J_2)= $N \leq 1 - id$. So, $J_1(a, N(a)) = S(N_1(a), \max(0, x + N(a) - 1)) = S(N(a), 0) = N(a)$.

- Under the condition $N_1 = N_2 = N \leq 1 - id$ with J_2 an S-implication becomes in to $S(N(N(a)), N(a)) = 1$, and implies that $S = W^*_\varphi$ and $N \geq N_\varphi$.

- Under these condition equation (1) could be verified, using J_2 a Q operator. We have $J_2(J_1(a, N(a)), N(a)) = S(N(a), T(a, x))$, and for $S(N(a), T(a, x)) = 1$, it is necessary $S = W^*_\varphi$, $N \geq N_\varphi$ and $T = Min$.

- It is analogous to prove that equation (1) could be also true for J_2 a D operator with $S = W^*_\varphi$, $N \geq N_\varphi$ and $T = Min$.

3.4 IfJ_1 is a D-operator

In this case the equation becomes

$$J_1(a, N(a)) = S(N(a), T(N_1(a), N_1(N(a)))).$$

3.4.1 S=Max

If J_1 is a D-operator with the t-conorm Max, we get $J_1(a, N(a)) = \max(N(a), T(N_1(a), N_1(N(a))))$ and if $N_1 \leq N$, $J_1(a, N(a)) = N(a)$, since, $J_1(a, N(a)) = \max(N(a), T(N_1(N(a)), N_1(a))) \leq \max(N(a), N_1(a)) = N(a)$, and $J_1(a, N(a)) = \max(N(a), T(N_1(N(a)), N_1(a))) \geq N(a)$.

- With the hypothesis $N_1 \leq N$, it can be repeated the proof in which the case of J_1 is a Kleene-Dienes operator, to show that the law is verified if J_2 is any R-implication.

- With the above hypothesis and with J_2 an S-implication, the equation (1) is $S(N_2(N(a)), N(a)) = 1$, that implies $S = W_\varphi$ and $N_2 \geq N_\varphi$.

- The law can be also verified if J_2 is a D-operator, with the t-conorm of Łukasiewicz and the t-norm minimum, it is a necessary condition.
  With J_2 a D-operator, the equation becomes $S(N(a), T(N_2(\max(N(a), T(N_1(N(a)), N_1(a)))), N_2(N(a)))) = 1$ and if $N_1 \leq N$

  $$S(N(a), T(N_2(N(a), N_2(N(a)))) = 1,$$

  so $T(N_2(N(a)), N_2(N(a))) \geq N_2(N(a))$, therefore T must be the t-norm minimum, and it is $S = W^*$ and $N_2 \geq N$.

- There is an analogous proof if J_2 is a QM-operator with the t-conorm of Łukasiewicz and t-norm minimum. Since if it is supposed $N_1 \leq N$, we get the equation

  $$S(N_2(N(a)), T(N(a), N(a))) = 1,$$

  and then $T = \min$, $S = W^*$ and $N_2 \geq N$.

3.4.2 J_1 is a D-operator, with T=W

We suppose N_1(negation of J_1)= N_2(negation of J_2)= $N \leq 1-id$. So, $J_1(a, N(a)) = S(N(a), \max(0, N_1(a)+ N_1(N(a)) - 1)) = S(N(a), 0) = N(a)$.

- Under these conditions $J_2(J_1(a, N(a)), N(a)) = J_2(N(a), N(a))$ and obviously, if J_2 is a R-implication, we get $J_2(N(a), N(a)) = 1$.
- As it was shown above, if $N_1 = N_2 = N \leq 1 - id$, it is $J_1(a, N(a)) = N(a)$. So, equation (1) with J_2 an S-implication, becomes $S(N_2(N(a)), N(a)) = 1$, and that implies $S = W^*_\varphi$ and $N_2 \geq N_\varphi$.
- Equation (1) could be verified, using J_2 a Q operator. We have $J_2(J_1(a, N(a)), N(a)) = S(N(a), T(a, x))$, and for $S(N(a), T(a, x)) = 1$, it is necessary $S = W^*_\varphi$, $N \geq N_\varphi$ and $T = Min$.
- It is analogous to prove that equation (1) could be also true for J_2 a D operator with $S = W^*_\varphi$, $N \geq N_\varphi$ and $T = Min$.

4 Conclusion

From long ago, functional equations are a tool of a great interest in the study of standard theories of fuzzy sets on a universe of discourse X, $([0,1]^X, T, S, N)$, where the t-norm T stands for 'and', the t-conorm S for 'or', and the strong negation N stands for 'not'. This tool allows to obtain fuzzy solutions for logical formulas coming from algebraic logic, namely boolean and orthomodular algebras. It gives a good enough mathematical way to deal with the fuzzy modeling of linguistic sentences, in order to capture and manage its meaning.

Extending this method to conditional functions J, once they are built whether by means of connectives *and, or* and *not*, that is, t-norms, t-conorms and negation functions or not, as in the case of R-implications with $T \neq W_\varphi$, the meaning of conditional 'if ...then...' sentences can be modeled.

As in the case of standard theories, when more than one function J is possible in the concrete formula in consideration, this tool allows to obtain Pexider-like solutions. This is the case of the problem initiated in this paper of which a lot of solutions are here presented. Table 1 collects the solutions founded in the particular case $\varphi = id$ and all strong negations equal to $N_{id} = 1 - id$.

Acknowledgement

This paper is partially supported by CICYT (Spain) under project TIN2005-08943-C02-01.

Table 1:

J1 \ J2	S(Max) ó S(Prod)	S(W*)	D(W*,W) ó D(W*,Pr)	D (W*,Min)	Q(W*,W) ó Q(W*,Pr)	Q (W*,Min)	R
Mamdani Min ó Prod		×		×		×	×
Mamdani W	×	×	×	×	×	×	×
S(Max)		×		×		×	×
D(Max,T)		×		×		×	×
D(S,W)		×		×		×	×
Q(Max,T)		×		×		×	×
Q(S,W)		×		×		×	×

Bibliography

[1] C. Alsina, M.J. Frank, and B. Scheiwzer. *Associative Functions: Triangular Norms And Copulas.* World Scientific Publishing, Singapour, 2006.

[2] C. del Campo. *Contribución al estudio y clasificación de las funciones de implicación borrosas.* PhD thesis, Univ. Politécnica de Madrid, Madrid, 2002. (In Spanish).

[3] J. Lukasiewicz. *Elements of mathematical logic.* Pergamon Press, 1963.

[4] E. Trillas and C. Alsina. On some schemes of reasoning in fuzzy logic. Technical report, European Centre for Soft Computing, 2007.

[5] E. Trillas, M. Mas, M. Monserrat, and J. Torrens. On the representation of fuzzy if then rules. Forthcoming in *Int. Jour. of Approx. Reasoning.*

[6] A.N. Whitehead and B. Russell. *Principia Mathematica.* Cambridge University Press, 1912.

Una mejora del modelo genético iterativo de SLAVE

Antonio González Raúl Pérez

Dpto. Ciencias de la Computación e Inteligencia Artificial
E.T.S. Ingenierías de Informática y Telecomunicación
Universidad de Granada
18071-Granada (Spain)
{A.Gonzalez, Raul_Perez}@decsai.ugr.es

Resumen

El modelo genético iterativo ha demostrado su utilidad en el desarrollo de algoritmos de aprendizaje de reglas difusas. En este trabajo[1], extendemos el esquema iterativo del algoritmo SLAVE para aprender una regla completa en cada iteración con la intención de reducir el tiempo necesario para el aprendizaje. El trabajo analiza esta extensión y presenta una amplio estudio experimental que muestra la mejora de la extensión propuesta.

Palabras Clave: Aprendizaje, Clasificación, Modelo genético iterativo, Algoritmos Genéticos

1 El modelo genético iterativo

La descomposición de un problema en subproblemas más simples es una técnica de resolución que permite reducir la complejidad de un gran número de problemas. Esta técnica ha dado lugar a la definición de multitud de modelos de resolución de problemas, tanto dentro como fuera del área de la inteligencia artificial. El algoritmo de recubrimiento secuencial o *sequential covering algorithm* [10] es un claro exponente de esta técnica dentro del campo del aprendizaje automático y es la base de algoritmos de aprendizaje muy conocidos.

El modelo genético iterativo [5, 6] puede verse como la adaptación de esa técnica, incorporando un algoritmo genético como mecanismo de búsqueda. Frente a los modelos de aprendizaje más conocidos que usan algoritmos genéticos, este modelo permite una importante reducción en el espacio de búsqueda que implica una reducción a su vez del tiempo requerido en el aprendizaje. Algunos algoritmos conocidos que siguen este modelo son SIA [14], MOGUL [2] o SLAVE [7].

En este trabajo presentamos NSLV, un algoritmo de aprendizaje de reglas difusas que extiende el modelo iterativo de SLAVE para aprender en cada iteración una regla completa. SLAVE utiliza un esquema iterativo donde se fija previamente la clase a aprender, y el algoritmo genético busca el mejor antecedente para esta clase. Este proceso requiere establecer el orden en el que las clases se van a aprender, lo que provoca un sesgo en el aprendizaje. Además, SLAVE implementa un mecanismo inmerso de selección de características. Este mecanismo requiere estimar la relevancia a priori de cada variable predictiva antes de cada iteración. Dicha estimación es costosa cuando el número de variables o ejemplos es elevado. En la extensión que se propone en NSLV, el esquema iterativo sólo realiza una vez dicha estimación. El nuevo esquema iterativo propuesto junto con la modificación del algoritmo genético nos permitirá reducir el tiempo requerido para el aprendizaje, manteniendo la capacidad de predicción y el nivel de comprensibilidad del conjunto de reglas.

En la sección 2 presentamos una breve descripción del algoritmo SLAVE. A continuación describimos NSLV, una evolución de SLAVE, que propone un nuevo esquema genético iterativo y un algoritmo genético diferente al de SLAVE, orientado a aprender reglas completas. En la sección 4 se realiza un estudio experimental exhaustivo con un gran número de bases de datos y un amplio número de algoritmos de aprendizaje para comparar el comportamiento de NSLV. Por último se presentan las conclusiones de este trabajo.

[1]Este trabajo ha sido financiado parcialmente por el Proyecto de Excelencia de la Junta de Andalucía P06-TIC1670 y por el Proyecto del Plan Nacional de I+D+i TIN2007-66367.

2 Una breve descripción de SLAVE

SLAVE [5, 7, 8] es un algoritmo de aprendizaje inductivo de reglas difusas que sigue el esquema básico del algoritmo de recubrimiento secuencial. La tarea principal de este algoritmo, como ilustra la figura 1, consiste en dada una clase, encontrar el mejor antecedente que describe las relaciones existentes entre las variables y la clase fijada. Esta tarea se formaliza como un problema de búsqueda, y de entre los distintos algoritmos que son aplicables, hemos utilizado un algoritmo genético.

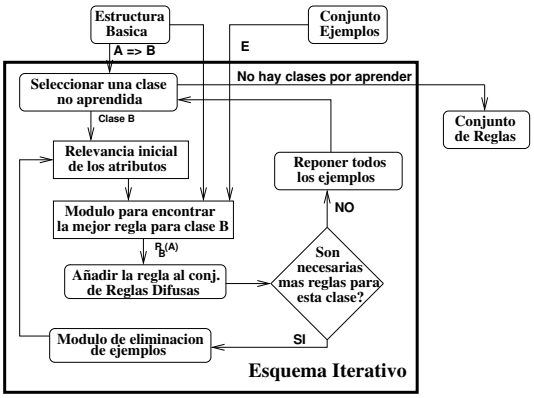


Figura 1: Esquema Iterativo de SLAVE.

Para la búsqueda del mejor antecedente, SLAVE utiliza un algoritmo genético que trabaja a dos niveles [8]. El primero de estos niveles, busca las variables relevantes para la descripción de la clase, mientras el segundo nivel busca la mejor asignación de valores a estas variables.

El primero de los niveles, denominado **nivel de variable** codifica la relevancia o irrelevancia de una variable para una regla particular. En este nivel se usa una codificación real con n+1 componentes (siendo n el número de variables implicadas en el problema). El elemento i-ésimo contiene un valor entre 0 y 1 que representa el grado de relevancia de la variable i-ésima con respecto a la clase. El valor n+1 es un valor entre 0 y 1 que representa un umbral de activación, de manera que la variable i-ésima se considerará relevante para el antecedente de la regla si su valoración es mayor o igual que este umbral de activación. En otro caso, la variable se considerará como irrelevante y será eliminada de la descripción de la regla. En la generación de la población inicial, se calcula una estimación de la relevancia a priori de cada variable con respecto a la clase usando una medida de información [8]. El proceso evolutivo cambiará estos valores iniciales para obtener una mejor estimación de la relevancia de cada una de las variables. Los operadores genéticos definidos para evolucionar estos valores son el operador de cruce uniforme y el operador de mutación no uniforme.

El segundo de los niveles denominado **nivel de valor** representa las asignaciones de valores a las variables. En este nivel se ha utilizado codificación binaria que expresa la presencia o ausencia de la asignación a una variable de cada uno de los valores de su dominio. SLAVE usa el modelo de regla difusa DNF que permite asignar un subconjunto de valores de su dominio a cada una de las variables. Los operadores genéticos definidos sobre este nivel son el cruce sobre dos puntos y la mutación uniforme.

El objetivo del algoritmo genético es, dada una clase, seleccionar del mejor antecedente. La medida fundamental utilizada para valorar una regla se basa en las condiciones de consistencia y completitud. En [7] se proponen definiciones que suavizan y adaptan las definiciones clásicas de estos conceptos para trabajar con reglas difusas. Además, en la evaluación de las reglas, también se tienen en cuenta dos funciones que miden la simplicidad y la comprensibilidad de las mismas [8]. La función de evaluación del algoritmo genético combina las medidas anteriores utilizando una evaluación funcional lexicográfica, es decir, hay establecido un orden de prioridad en la evaluación de los criterios.

El algoritmo genético sigue un modelo generacional con elitismo y con operador de selección basado en la ruleta ponderada donde la probabilidad de los individuos se calcula siguiendo un *ranking* lineal. El genético termina cuando la mejor solución encontrada no se mejoró en n generaciones. Este individuo se devuelve como el mejor antecedente.

3 NSLV: Una extensión del modelo genético iterativo de SLAVE

En esta sección describimos el algoritmo de aprendizaje NSLV. Este algoritmo hereda la mayor parte de la funcionalidad del algoritmo SLAVE aunque extendiendo su modelo genético iterativo para aprender una regla completa en cada iteración.

Como muestra la figura 1, SLAVE fija la clase a aprender y utiliza el algoritmo genético para encontrar el mejor antecedente para dicha clase. Este hecho crea una dependencia entre el conocimiento aprendido y el orden en el que se aprenden las clases. La extensión que aquí se propone da lugar a un nuevo algoritmo que denominaremos NSLV, que reformula tanto el esquema iterativo como el algoritmo genético para eliminar la dependencia y extraer la mejor regla posible en cada iteración sin tener que fijar la clase. El nuevo esquema de aprendizaje se muestra en la figura 2.

Así, la idea clave del nuevo esquema es aprender una

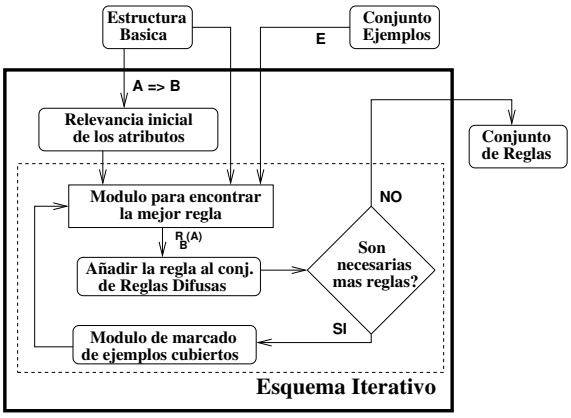


Figura 2: Esquema Iterativo de NSLV.

regla completa en cada iteración. Con este fin, el esquema redefine tanto el módulo de penalización de reglas como la condición de terminación del proceso. El módulo de penalización no puede eliminar los ejemplos cubiertos positivamente por la regla seleccionada como hace SLAVE. En NSLV, este módulo marca dichos ejemplos para que no afecten positivamente a ninguna de las reglas que se aprendan con posterioridad; únicamente afectarán negativamente sobre las reglas que difieren de las clases de los ejemplos. La condición de finalización del proceso iterativo se define en base al grado de completitud. Si la inclusión de la nueva regla seleccionada no mejora el grado de completitud, el proceso termina devolviendo como solución el conjunto formado por la agregación de las reglas encontradas excepto la última.

En NSLV, el módulo para encontrar la mejor regla, busca reglas completas y no sólo el antecedente de la regla. Esto implica dos cambios significativos de su algoritmo genético en relación con el que se propone en SLAVE: **(a)** la representación de los individuos de la población para incluir la codificación de la clase, **(b)** la inclusión de mecanismos genéticos para mantener alta la diversidad de la población y evitar que la evolución haga desaparecer todos los individuos de alguna de las clases a aprender.

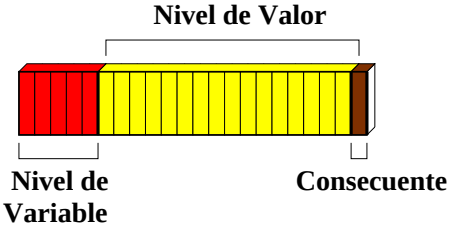


Figura 3: Representación genética de un individuo.

NSLV usa la codificación que muestra la figura 3 para representar cada individuo de la población. Como se puede observar, se mantienen los niveles de variable y de valor de la codificación utilizada en SLAVE. A diferencia de este, se añade un nuevo nivel con un único valor que identifica el consecuente de la regla que usa codificación entera y sobre él hay definido únicamente un operador de mutación uniforme.

Ya que los algoritmos genéticos convergen habitualmente hacia una única solución, es necesario asegurar que durante el proceso de búsqueda individuos de todas las clases permanezcan en la población. Con este fin, replanteamos el objetivo del algoritmo genético como un mecanismo que busca la mejor regla posible para cada una de las clases. Así, usando una estrategia elitista clásica aplicada a cada una de las clases, conseguimos que al menos los m mejores individuos de cada una de las clases se mantengan en la población. El parámetro m se fija durante el proceso de experimentación. Este proceso permite mantener alta diversidad en el genotipo de los individuos, por lo que los cruces entre individuos dan un buen grado de exploración de soluciones. Adicionalmente, con la intención de controlar el elitismo en las clases, se cambia el algoritmo genético de generacional a estacionario. En nuestro caso particular, hace que el proceso de selección de individuos consista en cruzar y mutar dos individuos para generar cuatro descendientes y sustituir los dos mejores de ellos por los dos peores de la población, de tal manera que se preserve que todas las clases tienen al menos m individuos. En la experimentación que se ha realizado en este trabajo hemos tomado m=*2*.

Mantener los mejores representantes de cada clase en la población, conlleva un beneficio adicional. Mientras en SLAVE, tras cada iteración, era necesario recalcular la estimación de la relevancia inicial de cada atributo para definir el nivel de variable en la población inicial, siendo este un proceso costoso en tiempo cuando el número de ejemplos y variables es elevado, en este nuevo enfoque, la población inicial de la siguiente iteración es la población final de la iteración anterior. Ya que esta población tiene los mejores individuos encontrados hasta el momento, es un buen punto de partida para la nueva búsqueda. Esto, por si mismo, implica una reducción en tiempo, pero dicha reducción aumenta al no ser necesario recalcular la relevancia inicial.

NSLV al igual que SLAVE devuelve, en cada iteración, una única solución, el mejor individuo global de toda la población.

Además de las modificaciones anteriores, NSLV usa el modelo de regla difusa DNF incluyendo un peso en

cada regla. Ese peso, un valor entre 0 y 1, es el ratio del número de ejemplos cubiertos de la clase sobre el número total de ejemplos que cubre la regla. Este factor entra en juego en el proceso de inferencia multiplicando el grado de adaptación del ejemplo sobre el antecedente de la regla. La regla seleccionada en la inferencia es aquella que hace máximo ese producto entre adaptación y peso. Si se produce un conflicto, se dispara la regla que tenga mayor peso. Si el conflicto se mantiene, se disparará la regla que se aprendió primero.

4 Resultados Experimentales

En esta sección comparamos el comportamiento de la nueva propuesta con la versión del algoritmo de SLAVE que se propone en [8]. También comparamos los resultados obtenidos con los de algunos de los algoritmos de aprendizaje más conocidos. Para la realización del estudio experimental hemos utilizado 20 base de datos obtenidas del UCI Repository of Machine Learning Databases and Domain Theories[11] que representan una amplia variedad de problemas que deben afrontar los algoritmos de aprendizaje.

Tabla 1: Las bases de datos usadas en la experimentación. En cada fila se muestra, para cada una de las bases de datos, el número de ejemplos (#E), atributos (#A), atributos continuos (#C), atributos nominales (#N), clases (#Cl) y si hay valores perdidos (P).

base de datos	#E	#A	#C	#N	#Cl	P
annealing	898	38	6	32	5	No
automobile	205	25	15	10	6	Si
balance	625	4	4	0	3	No
bupa	345	6	6	0	2	No
contracep.	1473	9	2	7	3	No
horse colic	368	22	7	15	2	Si
glass	214	9	9	0	6	No
heart-c	303	13	6	7	2	Si
heart-s	270	13	13	0	2	No
iris	150	4	4	0	3	No
pima	768	8	7	1	2	No
sonar	208	60	60	0	2	No
tao	888	2	2	0	2	No
thyroid	215	5	5	0	3	No
vehicle	846	18	18	0	4	No
w.breast-c	699	9	0	9	2	Si
w.diag.breast-c	569	30	30	0	2	No
wine	178	13	13	0	3	No
w.prog.breast-c	198	33	33	0	2	Si
zoo	101	17	0	17	7	No

Tanto NSLV como SLAVE discretizan los dominios de las variables continuas usando conjuntos difusos, de manera que cada uno de estos dominios se define usando cinco etiquetas difusas uniformemente distribuidas sobre su rango de definición. En ambos caso se utiliza el tratamiento de valores perdidos que se describe en [7]. Los resultados que se presentan en este trabajo se han obtenido aplicando validación cruzada sobre diez particiones en cada base de datos manteniendo la misma semilla, es decir, todos los algoritmos estudiados han trabajado sobre las mismas particiones. En los siguientes apartados comparamos los resultados obtenidos con el nuevo algoritmo de aprendizaje.

4.1 Comparando NSLV con SLAVE

En esta experimentación hemos centrado nuestro estudio en cuatro parámetros: acierto sobre los ejemplos vistos (train), la capacidad de predicción (test), el número total de condiciones que aparecen en el conjunto final de reglas (conds) y el tiempo necesitado para el aprendizaje (time). En todos los parámetros se muestra la media obtenida tras la validación cruzada.

Los cuatro parámetros estudiados nos muestran tres aspectos relevantes para evaluar la calidad del conocimiento obtenido por el algoritmo de aprendizaje. Los dos primeros nos indican la capacidad que tiene el conjunto de reglas aprendido para representar y generalizar las relaciones existentes en los ejemplos. El tercer parámetro nos indica la complejidad del conocimiento extraído. En este sentido, valores bajos del número total de condiciones indican descripciones simples y cercanas a lo que los seres humanos somos capaces de entender. El último parámetro nos da una estimación de la eficiencia del algoritmo.

Los resultados obtenidos sobre estos cuatro parámetros en los dos algoritmos se muestran en la tabla 2. Para analizar los resultados, aplicamos el test de Bonferroni-Dunn [3]. Este test define que un clasificador tiene un comportamiento significativamente diferente en comparación con otro, si sus correspondientes puestos medios (última fila, etiquetada como *Rnk*, de la tabla 2) difieren al menos en una distancia crítica (DC). Esta distancia crítica se define como:

$$DC = q_\alpha \sqrt{\frac{n_c(n_c+1)}{6n_{db}}}$$

donde n_c es el número de clasificadores, n_{db} es el número de bases de datos usadas en la experimentación y q_α es el valor crítico basado en el criterio de aceptación de hipótesis de Student a nivel α. Para $\alpha = 0.05$ el valor crítico $q_\alpha = 1.960$ y teniendo en cuenta el resto de datos de la experimentación, el valor de $DC = 0.438$.

A partir de este valor de DC, detectamos que existen diferencias significativas en los parámetros *train*

Tabla 2: Los resultados obtenidos con SLAVE y NSLV.

	train		test		conds		time	
	SLAVE	NSLV	SLAVE	NSLV	SLAVE	NSLV	SLAVE	NSLV
annealing	97.9(2)	98.9(1)	96.7(2)	98.5(1)	20.6(1)	23.2(2)	529.7(2)	157.9(1)
automobile	91.3(2)	95.3(1)	65.3(2)	75.7(1)	77.5(2)	75.7(1)	322.5(2)	93.1(1)
balance	86.4(1)	84.7(2)	73.4(1)	73.1(2)	60.6(2)	31.3(1)	173.7(2)	30.4(1)
bupa	65.6(2)	65.9(1)	59.9(1)	57.9(2)	23.3(2)	16.1(1)	52.1(2)	15.3(1)
contracep.	56.4(2)	58.0(1)	46.0(1)	44.3(2)	195.5(1)	230.1(2)	1428.7(2)	840.6(1)
horse colic	89.7(2)	89.9(1)	82.8(1)	81.8(2)	19.3(2)	19.1(1)	132.8(2)	36.5(1)
glass	70.5(2)	80.2(1)	56.2(2)	65.2(1)	49.5(2)	47.7(1)	165.7(2)	49.1(1)
heart-c	87.1(2)	88.0(1)	77.8(2)	78.1(1)	22.3(1)	23.3(2)	68.3(2)	27.6(1)
heart-s	88.5(1)	87.9(2)	76.2(1)	75.5(2)	26.1(2)	23.5(1)	63.9(2)	21.3(1)
iris	94.6(2)	96.6(1)	93.9(2)	96.0(1)	4.5(2)	3.7(1)	7.5(2)	2.1(1)
pima	78.6(2)	81.6(1)	72.2(2)	74.8(1)	50.7(2)	44.3(1)	277.3(2)	126.5(1)
sonar	86.8(2)	93.9(1)	72.5(1)	72.1(2)	40.6(1)	49.7(2)	166.3(2)	64.7(1)
tao	81.3(2)	83.1(1)	82.1(2)	82.8(1)	4.0(1)	7.0(2)	39(2)	19(1)
thyroid	93.6(2)	94.0(1)	91.2(1)	90.2(2)	13.5(2)	11.5(1)	25.2(2)	7.5(1)
vehicle	74.7(2)	83.5(1)	66.4(2)	66.7(1)	164.9(1)	185.3(2)	1537.3(2)	810.6(1)
w.breast-c	97.7(1)	97.4(2)	95.8(1)	95.4(2)	18.1(2)	13.0(1)	117.3(2)	27.3(1)
w.diag.breast-c	93.7(2)	95.9(1)	92.6(2)	93.8(1)	14.3(1)	18.6(2)	238.3(2)	84.5(1)
wine	97.1(2)	97.2(1)	92.6(2)	93.8(1)	10.9(1)	14.5(2)	33.5(2)	13(1)
w.prog.breast-c	84.9(1)	81.2(2)	73.3(2)	74.2(1)	70.7(2)	11.3(1)	160.2(2)	13.7(1)
zoo	99.7(2)	100.0(1)	95.8(1)	94.9(2)	10.0(1)	15.0(2)	23.7(2)	4.1(1)
Rnk	1.8	1.2	1.55	1.45	1.55	1.45	2	1

y *time*. Este hecho refleja que NSLV tiene mayor capacidad para representar los ejemplos de entrenamiento, pero necesitando menos tiempo de aprendizaje. En comparación con SLAVE, NSLV produce una reducción media en tiempo del 300%. El hecho de que no se detecten diferencias en los otros dos parámetros nos demuestra que la mejora en *train* no es debida a un sobreaprendizaje del conjunto de entrenamiento, ni ocasiona un aumento de la complejidad del conocimiento obtenido, ya que los resultados en *test* y *conds* se mantienen.

A la luz de este estudio, podemos concluir que NSLV mantiene el comportamiento de SLAVE en cuanto a capacidad de predicción y complejidad del conocimiento y mejora significativa el tiempo de respuesta del algoritmo.

4.2 Comparando NSLV con otros algoritmos

En esta sección comparamos experimentalmente el comportamiento de NSLV con otros algoritmos que siguen otros paradigmas de aprendizaje. Los algoritmos son los siguientes:

C4.5 [13] basado en árboles de decisión, extiende el algoritmo ID3 para trabajar con variables continuas y con valores perdidos.

IBk [1] basado en el algoritmo del vecino más cercano. Clasifica un ejemplo con la clase mayoritaria de sus k vecinos más cercanos. En nuestra experimentación el valor de k=5.

N.Bayes [9] es un clasificador probabilístico que estima los parámetros de un modelo Bayesiano.

Part [4] es una arquitectura de aprendizaje que aplica técnicas de aprendizaje de reglas *divide y vencerás* para crear un clasificador.

SMO [12] es una implementación de *support vector machines*. En nuestra experimentación hemos usados SMO con núcleo polinomial de orden 3 (SMOpol) y SMO con núcleo gausiano (SMOgau).

Todos los algoritmos se han ejecutado usando WEKA [15] y en todos los casos, se ha usado la configuración recomendada en dicha plataforma. Los resultados obtenidos sobre capacidad de predicción (*test*) se muestran en la tabla 3. La última fila de esta tabla, *Rnk*, muestra la posición media de cada algoritmo.

Aplicando el mismo test estadístico que se realizó en el apartado anterior, para un valor de $\alpha = 0.05$, el valor de DC es 1.80.

Con este valor de DC no se encuentran diferencias significativas entre los resultados obtenidos por NSLV y los obtenidos por el resto de los algoritmos estudiados. Este hecho pone de manifiesto que el algoritmo aquí propuesto presenta una capacidad de predicción comparable a la que ofrecen otros algoritmos conocidos que representan otros paradigmas de aprendizaje.

Tabla 3: Comparación de NSLV con algoritmos de aprendizaje.

	C4.5	IB5	N.Bayes	Part	SMOpol	SMOgau	NSLV
annealing	98.9(2)	97.3(5)	86.3(7)	98.6(3)	99.3(1)	91.9(6)	98.5(4)
automobile	80.9(1)	64.0(5)	58.8(6)	74.4(4)	78.1(2)	45.6(7)	75.7(3)
balance	77.4(6)	88.2(4)	90.6(2)	82.9(5)	91.2(1)	88.3(3)	73.1(7)
bupa	62.3(2)	58.9(4)	56.0(7)	67.6(1)	60.0(3)	58.0(5)	57.9(6)
contracep.	52.6(1)	46.5(5)	50.6(2)	50.0(3)	48.7(4)	42.7(7)	44.3(6)
horse colic	85.3(1)	81.5(5)	78.2(6)	84.5(2)	75.6(7)	82.4(3)	81.8(4)
glass	66.1(2.5)	64.7(5)	48.9(6)	66.6(1)	66.1(2.5)	35.6(7)	65.2(4)
heart-c	78.4(5)	83.2(1)	82.8(2)	74.2(7)	78.6(4)	82.5(3)	78.1(6)
heart-s	79.3(5)	80.7(3)	83.3(1)	80.0(4)	78.9(6)	82.6(2)	75.5(7)
iris	94.0(4.5)	96.0(2)	96.0(2)	94.0(4.5)	92.7(7)	93.3(6)	96.0(2)
pima	74.2(5)	73.3(6)	75.8(3)	78.9(1)	76.7(2)	65.1(7)	74.8(4)
sonar	71.0(5)	84.0(2)	69.7(6)	74.4(3)	85.5(1)	69.3(7)	72.1(4)
tao	95.9(2)	97.1(1)	81.0(7)	94.3(3)	84.2(4)	83.6(5)	82.8(6)
thyroid	94.9(2)	94.8(3)	97.2(1)	94.3(4)	88.9(6)	69.8(7)	90.2(5)
vehicle	71.1(3)	68.9(4)	46.3(6)	73.4(2)	83.3(1)	41.7(7)	66.7(5)
w.breast-c	95.1(7)	97.1(1)	96.1(3.5)	95.7(5)	96.4(2)	96.1(3.5)	95.4(6)
w.diag.breast-c	94.4(4)	96.8(2)	93.1(6)	94.5(3)	97.6(1)	92.9(7)	93.8(5)
wine	93.9(4)	96.7(3)	97.2(2)	93.3(6)	97.7(1)	39.9(7)	93.8(5)
w.prog.breast-c	71.6(5)	78.8(2)	69.4(7)	70.0(6)	81.2(1)	73.0(4)	74.2(3)
zoo	92.8(5)	90.5(6)	94.5(3)	93.8(4)	97.8(1)	76.0(7)	94.9(2)
Rnk	3.60	3.45	4.28	3.58	2.88	5.55	4.6

5 Conclusiones

En este trabajo hemos descrito el algoritmo NSLV, una evolución del algoritmo SLAVE, que mejora el tiempo de aprendizaje manteniendo la capacidad de predicción y la complejidad del conocimiento obtenido. Tomando como base SLAVE, NSLV propone un nuevo esquema iterativo y un nuevo algoritmo genético con el fin de aprender una regla completa en cada iteración y eliminar la dependencia creada entre el conocimiento obtenido y el orden en el que se aprenden las clases. Los resultados experimentales muestran que NSLV proporciona una mejora significativa del 300% de reducción en tiempo de aprendizaje en comparación con SLAVE y que sus resultados en capacidad de predicción son comparables a los que ofrecen otros algoritmos conocidos de aprendizaje.

Referencias

[1] D.W. Aha, D.F Kibler, M.K. Albert, Instance-based learning algorithms. Marchine Learning, 6(1) pp. 37-66, 1991.

[2] O. Cordón, M.J. del Jesús, F. Herrera, Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods, Int. Journal of Intelligent System, vol. 13, pp. 1025-1053, 1998.

[3] O.J. Dunn, Multiple comparisons among means. Journal of the American Statistical Association, 56 pp. 52-64, 1961.

[4] E. Frank, I.H. Witten, Generating accurate rule sets without global optimization. In Proceedings of the 15th International Conference on Machine Learning, pp. 144-151. Morgan Kaufmann, 1998.

[5] A. González, R. Pérez, J.L. Verdegay, Learning the structure of a fuzzy rule: a genetic approach, Proc. EUFIT'93, vol. 2, pp. 814-819, 1993. Además en Fuzzy System and Artificial Intelligence, vol. 3, no. 1, pp. 57-70, 1994.

[6] A. González, F. Herrera, Multi-stage genetic fuzzy systems based on the iterative rule learning approach, Mathware and Soft Computing vol. 4, no. 3 pp. 233-249, 1997.

[7] A. González, R. Pérez, SLAVE: a genetic learning system based on an iterative approach, IEEE Transaction on Fuzzy System, vol. 27, no. 2, pp. 176-191, 1999.

[8] A. González, R. Pérez, Selection of relevant features in fuzzy genetic learning algorithm, IEEE Transaction on Systems, Man and Cybernetics, Vol. 31(3), 417-425, 2001.

[9] G.H. John, P. Langley, Estimating continuous distributions in bayesian classifiers. In 11th Conference

on Uncertainty in Artificial Intelligence, pp. 338-345. Morgan Kaufmann, 1995.

[10] T. Mitchell, Machine Learning, Ed. McGraw-Hill, 1997.

[11] P.M. Murphy, D.W. Aha, UCI Repository of machine learning databases, Irvine, CA, University of California, Departament of Information and Computer Science, (1998) [http://www.ics.uci.edu/ mlearn/MLRepository.html]

[12] J. Platt, Fast training of support vector machines using sequential minimal optimization. In Advances in Kernel Methods - Support Vector Machine. MIT Press, 1998.

[13] J.R. Quinlan, *C4.5 Program for Machine Learning*, Morgan Kaufmann, 1993.

[14] G. Venturini, SIA: a supervised inductive algorithm with genetic search for learning attributes based concepts, Machine Learning: ECML-93, pp. 280-296, 1993.

[15] I.H. Witten, E. Frank, Data mining: practical machine learning tools and techniques. Morgan Kaufmann, San Francisco, 2nd edition, 2005.

APRENDIZAJE COOPERATIVO DE LA BASE DE REGLAS Y EL SISTEMA DE INFERENCIA MEDIANTE ALGORITMOS GENÉTICOS MULTIOBJETIVO

A.A. Márquez **F.A. Márquez** **A. Peregrín**

Dept. de Tecnologías de la Información. Universidad de Huelva 21071 – Huelva,
{amarquez, alfredo.marquez, peregrin}@dti.uhu.es

Resumen

En este trabajo se propone utilizar Algoritmos Genéticos Multiobjetivo para el aprendizaje de Sistemas Basados en Reglas Difusas Lingüísticos con un buen equilibrio entre precisión e Interpretabilidad basándose en el aprendizaje cooperativo entre los parámetros de los operadores difusos adaptativos en el Sistema de Inferencia y las reglas de la Base de Reglas a través de un mecanismo de selección. Para ello, se utilizan dos algoritmos bien conocidos como son SPEA2 y NSGA-II, de modo que se obtendrán soluciones en la zona del pareto a fin de seleccionar aquellas que mejor equilibrio tengan entre el número de reglas y la precisión en dichos sistemas.

Palabras Clave: Modelado Difuso Lingüístico, Interpretabilidad, Algoritmos Genéticos Multiobjetivo, Sistema de Inferencia Adaptativo, Selección de Reglas.

1 INTRODUCCION

La Interpretabilidad y la Precisión son dos requerimientos generalmente contradictorios en el diseño de Modelos Difusos [3]. La interpretabilidad es la capacidad de expresar el comportamiento del sistema real de una forma entendible mientras que la precisión es la capacidad de representar fielmente el comportamiento del sistema real.

Las tareas fundamentales en el diseño de Sistemas Difusos Lingüísticos para una aplicación concreta son dos: la obtención de la Base de Conocimiento (BC), y la selección de los operadores del Sistema de Inferencia y la Interfaz de Defuzzificación. Si bien la BC es el elemento fundamental que define el comportamiento del sistema, recientemente, cuando se trata de conseguir un buen equilibrio entre precisión e interpretabilidad, la selección de los operadores ha captado una relevancia creciente. El uso de operadores adaptativos tanto en el Sistema de Inferencia [1] como en la Interfaz de Defuzzificación [10], o en ambos de modo simultáneo [16], ha demostrado ofrecer mejoras importantes en la precisión para cada aplicación. En [19] se propuso un modelo evolutivo en el que la Base de Reglas (BR) y los conectivos del Sistema de Inferencia son aprendidos de forma simultánea, buscándose la mejor precisión posible en base a un buen nivel de cooperación entre ambos elementos. Los Algoritmos Genéticos empleados en dicho trabajo se muestran como una herramienta muy adecuada para este tipo de problemas de optimización.

En este trabajo se propone en primer lugar un modelo evolutivo en el que se aprenden los parámetros correspondientes al Sistema de Inferencia Adaptativo y se le añade un mecanismo de selección de reglas [4] con objeto de obtener una base de reglas con mayor compacidad, y por tanto favoreciendo la interpretabilidad del sistema al tiempo que se optimiza la precisión. En este sentido se han utilizado dos métodos de generación inicial de la Base de Reglas como son WM[21] y el método COR[2].

En segundo lugar, se propone un modelo evolutivo basado en Algoritmos Genéticos Multiobjetivo (AGMO) [5] con objeto de optimizar al mismo tiempo dos objetivos: la precisión del sistema y el número de reglas. Se obtiene así un conjunto de soluciones con distinto equilibrio en el citado marco del compromiso entre precisión e interpretabilidad, expresada ésta última en función del número de reglas, considerándose mejor cuanto menor es el número de las mismas.

El trabajo se estructura de la siguiente forma: en la Sección 2 se describe el Sistema de Inferencia Adaptativo, sus componentes y efectos. A continuación en la Sección 3, se muestra el uso de AGMO en el problema del equilibrio entre precisión e interpretabilidad, se describen dos de los modelos más extendidos que son empleados en este trabajo SPEA2 y NSGA-II y se muestran los operadores genéticos considerados. La

Sección 4 se dedica a un estudio experimental en el que se emplean estas técnicas anteriormente descritas en dos problemas reales, y por último el trabajo termina mostrando las conclusiones al respecto del estudio desarrollado.

2 SISTEMA DE INFERENCIA ADAPTATIVO

En esta sección se describe la notación empleada para los Sistemas Basados en Reglas Difusas (SBRDs), así como la construcción del Sistema de Inferencia adaptativo utilizado.

Los SBRDs lingüísticos utilizan reglas SI - ENTONCES de la siguiente forma:

$$R_i : Si\ X_{i1}\ \ es\ A_{i1}\ \ y \ldots y\ X_{im}\ es\ A_{im}\ entonces\ Y\ es\ B_i$$

con i = 1 hasta N reglas, siendo X_{i1} hasta X_{im} las entradas e Y la salida, y con A_{i1} hasta A_{im} y B_i los antecedentes y consecuentes respectivamente.

La expresión de la Regla Composicional de Inferencia en el modelado difuso con fuzzificación puntual es mostrada en la expresión (1), donde $\mu_{B'}$ es la función de pertenencia del consecuente inferido, I(·) es el operador de implicación, C(·) el operador de conjunción, μ_{Ai} al tratarse de fuzzificación puntual son los puntos de corte de las entradas discretas ($x_{1,\ldots,}x_n$) del sistema con las funciones de pertenencia de los antecedentes de las reglas, y μ_B el consecuente de la regla.

$$\mu_{B'}(y) = I\,(C\,(\mu_{A1}(x_1), \ldots, \mu_{An}(x_n)), \mu_B(y)), \qquad (1)$$

Por tanto, el sistema de inferencia emplea dos operadores: el de conjunción, C(·) para calcular el grado de emparejamiento, y el de implicación, I(·) para calcular el consecuente inferido.

Estos dos operadores son pues susceptibles de utilizar a priori de forma parametrizada para diseñar el Sistema de Inferencia Adaptativo, como se estudió en [1 y 15]. Sin embargo, en los citados trabajos se concluye que la influencia del operador de implicación parametrizado es menor por lo que en este trabajo se empleará el Sistema de Inferencia Adaptativo basado sólo en el operador de conjunción parametrizado.

Las t-normas son una familia de operadores generalmente utilizados en el diseño de modelos difusos, tanto para el operador de conjunción como para de implicación [14] ofreciendo buen comportamiento [9]. En [20] se encuentran varias t-normas parametrizadas de las cuales ha sido seleccionada para este trabajo la de Dubois, cuya expresión se muestra en (2), como operador de conjunción adaptativo por ser la más eficiente computacionalmente, y ofrecer un comportamiento comprendido entre la del mínimo y la del producto algebraico en función del valor de su parámetro (con α = 0 y α = 1 respectivamente) que son las que mejores resultados ofrecen [8].

$$\underset{(0\le\alpha\le1)}{\text{Dubois}} \qquad T_{\text{Dubois}}(x,y,\alpha) = \frac{x\cdot y}{\text{Max}(x,y,\alpha)} \qquad (2)$$

Se podrían considerar dos modelos de operador de conjunción adaptativo dependiendo de si se utiliza un único parámetro α, o distintos parámetros, α_i, uno para cada regla de la BC. Un sólo parámetro permite ajustar el comportamiento del conectivo globalmente mientras que los parámetros individuales lo hacen particularmente para cada regla de la BC. El modelo con un sólo parámetro, como es de esperar, ofrece peores resultados [1], por lo que en este trabajo se utiliza el modelo con mayor flexibilidad que dota de un parámetro para cada regla. La t-norma de Dubois actúa como la del mínimo cuando α = 0, mientras que cuando α = 1 lo hace como la del producto. Cuando los valores de la t-norma de Dubois están comprendidos entre 0 y 1, ésta continua comportándose como el mínimo salvo cuando todos los *cruces* con los antecedentes se encuentran por debajo de α, tomando valores entre el mínimo y el producto, y provocando un efecto similar al conocido mecanismo de concentración [18], que consiste en alterar el significado lingüístico de la estructura de la regla modificando el significado de la etiqueta. Por tanto, la t-norma parametrizada de Dubois conecta con el mínimo en aquellos casos en los que los cruces con los antecedentes son más significativos, mientras que los demás los lleva a cabo con un valor comprendido entre el mínimo y el producto algebraico.

3 USO DE ALGORITMOS GENÉTICOS MULTIOBJETIVO EN EL PROBLEMA DE LA INTERPRETABILIDAD-PRECISION. SPEA2 Y NSGA-II

Los AGMO generan una familia de soluciones válidas donde cada solución tiende a satisfacer varios objetivos de tal forma que unas soluciones potencian un objetivo más que a otros. Por esta razón, los AGMO han sido también aplicados para encontrar un buen balance en el problema de interpretabilidad-precisión en los SBRDs, donde cada solución en la frontera del pareto representa una opción en este equilibrio (Figura 1).

En las próximas subsecciones, en primer lugar presentaremos las bases de SPEA2 [22] y NSGA-II [11], dos de los AGMO más representativos de los pertenecientes a la segunda generación, caracterizados por utilizar el elitismo como mecanismo estándar. Posteriormente describimos el esquema de codificación y

los operadores utilizados para la implementación de los algoritmos.

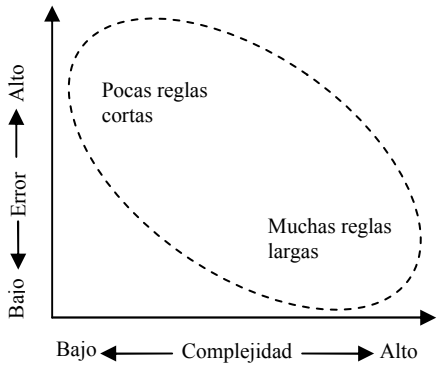


Figura 1: Balance entre precisión e interpretabilidad

3.1 SPEA2

El algoritmo SPEA2 [22] (*Strength Pareto Evolutionary Algorithm for multiobjective optimization*) es una de las técnicas mas utilizadas en la resolución de problemas multiobjetivo. Éste se diferencia de otros AGMO en varios aspectos, entre los cuales hay dos de gran importancia:

- Incorpora una estrategia fina de asignación del *fitness* que considera, para cada individuo, el número de los individuos que domina y el número de los individuos por los cuales es dominado.
- Utiliza la técnica del vecino más cercano para la valoración de la densidad, dirigiendo la búsqueda de forma más eficiente.

Según la descripción de los autores en [22] el bucle principal de SPEA2 consta de los siguientes pasos:

Entrada: N (tamaño de la población),
$\overline{N}$ *(tamaño de la población externa),*
T (máximo número de generaciones).
Salida: A (conjunto de no dominados).

1. Crear la población inicial P_0 y una población externa vacía $\overline{P_0} = \varnothing$;.

2. Calcular el *fitness* de los individuos de P_t y $\overline{P_t}$.

3. Copiar los no dominados de $P_t \cup \overline{P_t}$ en $\overline{P_{t+1}}$. Si $|\overline{P_{t+1}}| > \overline{N}$ se aplica el operador de truncamiento. Si $|\overline{P_{t+1}}| < \overline{N}$ se rellena con dominados de $P_t \cup \overline{P_t}$.

4. Si t ≥ T, se devuelve A y acaba la ejecución.

5. Aplicar torneo binario con reemplazamiento en $\overline{P_{t+1}}$ hasta completar la población de padres.

6. Aplicar cruce y mutación para construir P_{t+1}. Volver al Paso 2 con t = t + 1.

3.2 NSGA-II

El algoritmo NSGA-II [11] es otro de los AGMO más conocidos y frecuentemente utilizados en la literatura para la resolución de problemas multiobjetivos. Como en otros algoritmos evolutivos, primeramente el NSGA-II genera una población inicial. La población descendiente se genera desde la población actual a través de la selección, el cruce y la mutación. La siguiente generación se construye desde la población actual y la descendiente hasta que se cumpla la condición de parada.

El NSGA-II posee dos características que lo hacen uno de los principales y más importantes AGMO. Una es la asignación del *fitness* basada en el *ranking* del pareto y en el operador de *crowding*, y la otra es el procedimiento de actualización de cada generación a través del elitismo.

Considerando estos dos componentes previamente definidos y las descripciones de sus autores en [11], NSGA-II consiste en los siguientes pasos:

1. Una población combinada R_t se forma con la población inicial de la Población padre P_t y la población descendiente Q_t (inicialmente vacía).

2. Generar todas las fronteras de soluciones no dominadas $F = (F_1, F_2, ...)$ de R_t.

3. Inicializar $P_{t+1} = 0$ e $i = 1$.

4. Repetir hasta que la solución padre es completada.

5. Calcular la distancia de *crowding* en F_i.

6. Incluir *i-th* individuos no-dominados desde la población padre.

7. Comprobar la siguiente frontera para su inclusión.

8. Ordenar en orden descendiente usando el operador de *crowding* para comparar.

9. Elegir el primer (N − | P_{t+1} |) elementos de F_i.

10. Aplicar cruce y mutación para crear la nueva población Q_{t+1}.

11. Incrementar el contador de generación.

3.3 ESQUEMA DE CODIFICACIÓN

En este trabajo se utiliza un esquema de codificación doble:

$$C_C + C_S$$

donde C_C codifica los parámetros del operador de conjunción y C_S a la selección de reglas.

- C_C utiliza un esquema de codificación real, siendo m parámetros α_i, uno para cada R_i de la BR. Cada uno toma valores en el intervalo [0,1], es decir, entre la t-norma del mínimo y la del producto algebraico

- La parte C_S está formada por cadenas binarias de longitud m (numero inicial de reglas). Cada gen podrá tomar los valores '1' o '0' indicando si la correspondiente regla es o no seleccionada:

$$C_S = (c_{S1}, \ldots, c_{Sm}) \mid c_{Si} \in \{0, 1\}$$

La población inicial se obtiene de la siguiente forma:

- En la parte C_C se les asigna a todos el valor inicial 0 para que la t-norma adaptativa de Dubois equivalga inicialmente a la t-norma del mínimo.
- Para la parte C_S, a todos los genes se les asigna el valor 1 inicialmente en todos los individuos de la población inicial.

3.4 OPERADORES DE CRUCE Y MUTACIÓN

El operador de cruce utilizado depende de la parte del cromosoma donde sea aplicado:

- En la del operador de conjunción se utiliza BLX-0.5 [13].
- En la de selección de reglas se utiliza el cruce HUX [12].

Finalmente, se generan cuatro descendientes combinando dos hijos de la parte C_S con dos hijos de la parte C_C (los dos mejores reemplazan a los padres). El operador de mutación cambia aleatoriamente el valor de un gen tanto en la parte C_S como en la parte C_C con probabilidad P_m.

4 ESTUDIO EXPERIMENTAL

Para analizar el funcionamiento del método propuesto, se ha escogido dos problemas reales correspondientes a un problema de distribución del sistema eléctrico español [6], consistente en el cálculo de los costes óptimos teóricos de la línea de media tensión en núcleos urbano, en primer lugar. En segundo lugar, se aborda un problema característico descrito en la literatura especializada, el problema de degustación del arroz [17]. Se trata de dar un valor a determinados platos de arroz según sus características (sabrosidad, apariencia, sabor, viscosidad y dureza).

Los métodos estudiados se muestran en la Tabla 1. Dos métodos de generación de la BR inicial han sido utilizados: WM [21] es utilizado como método simple para la generación de la BC inicial; y el método COR (reglas que cooperan) [2], que es un método *ad hoc* guiado por ejemplos, que proporciona conjuntos de reglas particularmente precisas dado que no selecciona la mejor regla individual en cada caso sino que prefiere aquellas que en conjunto, ofrezcan una BC con un mejor comportamiento global. Este método maneja un conjunto de etiquetas de consecuentes, uno para cada regla, obteniendo la BR como un problema de optimización combinatoria.

Tabla 1: Métodos analizados.

Ref.	Método	Descripción
[21]	**WM**	Wang & Mendel
[17]	**COR**	Método COR
[4]	**WM/COR+S**	Selección de Reglas
[16]	**WM/COR+C**	Ajuste del Sistema de Inferencia
-	**WM/COR+S-C**	Selección de Reglas y Ajuste
[22]	**SPEA2**	Algoritmo SPEA2
[11]	**NSGA-II**	Algoritmo NSGA-II

Los modelos llamados C y S realizan el ajuste de parámetros en la inferencia y en selección de reglas respectivamente. El modelo S-C es el que selecciona reglas y aprende el operador de conjunción simultáneamente. Todos ellos a excepción de los dos últimos de la tabla, consideran como objetivo único la precisión del modelo. Los dos restantes situados en la parte inferior de la tabla son los AGMO, SPEA2 y NSGA-II (ambos realizan selección de reglas y ajuste de parámetros con dos objetivos, precisión medido por el tradicional Error Cuadrático Medio (ECM), e interpretabilidad en base al número de reglas).

Las particiones lingüísticas consideradas tendrán términos lingüísticos con forma triangular y con el número con el que los métodos presentan el mejor comportamiento). Se utiliza el centro de gravedad ponderado por el grado de emparejamiento como operador de defuzzificación y la t-norma del mínimo como operador de implicación. Los valores de los parámetros utilizados en todos los experimentos presentados son: 200 de longitud de la población, 61 de longitud de la población externa (para el caso del algoritmo SPEA2) y 200 de longitud de la población, 200 de longitud de la descendiente (para el caso del NSGA-II), 1000 generaciones y 0.2 como probabilidad de mutación por cromosoma.

4.1 DESCRIPCIÓN DE LOS PROBLEMAS

Estimar los costos de mantenimiento de la red eléctrica de media tensión de una ciudad [6] es un problema complejo

pero interesante, dado que es difícil obtener una medición real y por ello la consideración de modelos resulta muy útil, ya que estas estimaciones permiten a las compañías eléctricas justificar sus gastos. Además, el modelo obtenido debe poder explicar cómo se calcula un valor específico para una determinada ciudad. El objetivo pues será el de relacionar el costo de mantenimiento con cuatro características: suma de las longitudes de todas las calles de la ciudad, área total de la ciudad, área ocupada por edificios, y energía suministrada a la ciudad. Para ello se dispone de estimaciones de costos de mantenimiento basadas en un modelo de una red eléctrica óptima para cada ciudad, en una muestra de 1.059 ciudades. El dominio de las variables de entrada y salida fue dividido en cinco etiquetas {MP, P, M, G, MG}.

Utilizamos una BR con 65 reglas lingüísticas obtenidas con el método de WM [21] y 65 por el método COR[2] desde un conjunto de datos de entrenamiento con el 80% de los datos disponibles, es decir, 847 de las 1059 ciudades, tomados de modo aleatorio. Se ha utilizado un modelo de validación cruzada con 5 particiones de datos, esto es, 5 particiones aleatorias al 20% (4 de ellas con 212 ejemplos y una con 211), y la combinación de cuatro particiones (80%) como entrenamiento y la partición restante (20%) como test. Así se tienen 5 particiones al 80% y 20% en entrenamiento y test respectivamente.

En cuanto al problema de degustación del arroz, se realiza una calificación del sabor del arroz llevando a cabo una evaluación subjetiva llamada test sensorial. En este test, un grupo de expertos, generalmente compuesto por 24 personas, evalúan el arroz acorde aun conjunto de características asociadas a él.

Debido a la cantidad de variables relevantes, el problema del análisis del sabor en el arroz se vuelve bastante complejo, requiriendo así el diseño de un modelo que represente las relaciones existentes no lineales. En el presente trabajo se ha utilizado el conjunto de datos representado en [17]. Dicho conjunto está compuesto por 105 tablas de datos conteniendo la evaluación subjetiva de seis variables (las cinco mencionadas y la estimación global de la clase de arroz), realizada por expertos sobre los tipos distintos de arroz cultivados en Japón (por ejemplo, Sasanishiki, Akita-Komachi, etc.).

Las seis variables se han normalizado, tomando así valores en un intervalo real. Debido al pequeño número de ejemplos que se han utilizado, existe un alto riesgo de sesgo en el proceso de aprendizaje. Por ello, se ha obtenido aleatoriamente varias particiones del conjunto de datos mencionados (71% para entrenamiento y 29% para test). En este sentido se han utilizado 10 particiones para el conjunto de entrenamiento y de test con 75 y 30 ejemplos respectivamente.

Por último mencionar que se han utilizado dos etiquetas para todo el dominio de las variables lingüísticas.

4.2 RESULTADOS Y ANÁLISIS

Para cada una de las cinco particiones de los datos del problema eléctrico realizadas se ejecutan los modelos evolutivos seis veces, mientras que para el arroz se ejecutan 3 veces de modo que se llevan a cabo un total de 30 ejecuciones cuyos resultados medios se muestran en la Tabla 2 y 3, respectivamente. En el caso de los métodos con enfoque multiobjetivo (los cuatro últimos), las medias han sido obtenidas considerando la solución más precisa de cada pareto obtenido. Los algoritmos propuestos se han comparado con varios métodos guiados por un único objetivo.

Los resultados obtenidos se muestran en las Tablas 2 y 3, donde #R representa el número de reglas, ECM_{ent} y ECM_{prueba} el error medio en entrenamiento y prueba, σ la desviación típica y t-test los resultados de aplicar un test t-student (con 95% de confianza) al mejor resultado medio de la correspondiente columna comparado 1 a 1 con el resto de resultados medios. La interpretación de esta última columna es:

* Indica el resultado con la mejor media.
\+ Indica un comportamiento significativamente peor que el mejor.

Observando los resultados presentados en la Tabla 2 y 3 podemos destacar los siguientes hechos:

- Ambos modelos multiobjetivo, SPEA2 y NSGA-II, obtienen soluciones con mejor precisión tanto en entrenamiento como en prueba, además de una reducción importante en el número de reglas respecto a los modelos que consideran como único objetivo la precisión.
- Los modelos WM+S-C y COR+S-C, de un único objetivo basado en la precisión, tienen mayor número de reglas que los modelos multiobjetivo, pero su nivel de precisión alcanza también buen nivel.

Los resultados muestran que los AGMOs permiten seleccionar un subconjunto de reglas que cooperan mejorando a la vez en precisión. Esta idea ya se apunta en los resultados del modelo WM+S y COR+S, pero combinándolo con el aprendizaje del operador de conjunción en un modelo Multiobjetivo, se obtiene un sistema en el que existe cooperación entre la BR y el Sistema de Inferencia, con mejor equilibrio de precisión e interpretabilidad.

En la Figura 2 se muestra la formación del pareto durante la ejecución de cada algoritmo SPEA2 y NSGA-II, en el que puede apreciarse la frontera del pareto en cada generación.

Tabla 2: Resultados obtenidos por los métodos estudiados para el problema eléctrico

Modelo	#R	ECM_{ent}	σ_{ent}	t-test	ECM_{prueba}	σ_{prueba}	t-test
WM	65	56135,75	1321,42	+	56359,42	2375,47	+
COR	65	50710,80	1120,12	+	54584,76	2156,89	+
WM+S	40,9	41517,01	1012,66	+	44064,67	4174,48	+
WM+C	65	34362,32	512,78	+	37077,15	2164,00	+
WM+S-C	52,5	31196,26	954,09	+	54747,97	6386,45	+
COR+S	44,7	40763,48	1152,30	+	43228,38	3982,34	+
COR+C	65	31236,59	1581,95	+	34762,19	4449,19	+
COR+S-C	47,0	28530,80	1529,25	+	48836,66	5905,10	+
WM+SPEA2	**39,7**	30045,83	896,77	+	34021,08	4055,76	+
WM+NSGA-II	40,3	29588,02	983,88	+	33271,00	3502,89	+
COR+SPEA2	43,5	**27582,24**	1128,20	*	31700,19	2933,96	=
COR+NSGA-II	42,7	28200,27	1414,49	=	**31242,74**	2793,76	*

Tabla 3: Resultados obtenidos por los métodos estudiados para el problema arroz

Modelo	#R	ECM_{ent}	σ_{ent}	t-test	ECM_{prueba}	σ_{prueba}	t-test
WM	15	0,013284	0,003161	+	0,013119	0,004468	+
COR	15	0,007979	0,000984	+	0,008244	0,001103	+
WM+S	**4,4**	0,003672	0,000323	+	0,004630	0,000999	+
WM+C	15	0,002320	0,000321	+	0,001989	0,000352	=
WM+S-C	8,8	0,001752	0,000414	+	0,004268	0,003020	+
COR+S	5,4	0,003400	0,000231	+	0,004005	0,000733	+
COR+C	15	0,001578	0,000534	+	**0,001957**	0,000418	*
COR+S-C	11,0	0,001386	0,000212	=	0,003620	0,002109	+
WM+SPEA2	6,7	0,001578	0,000282	+	0,002408	0,001214	+
WM+NSGA-II	6,5	0,001608	0,000305	+	0,002381	0,001182	+
COR+SPEA2	9,5	**0,001312**	0,000184	*	0,002092	0,000464	=
COR+NSGA-II	9,2	0,001348	0,000175	=	0,002118	0,000354	=

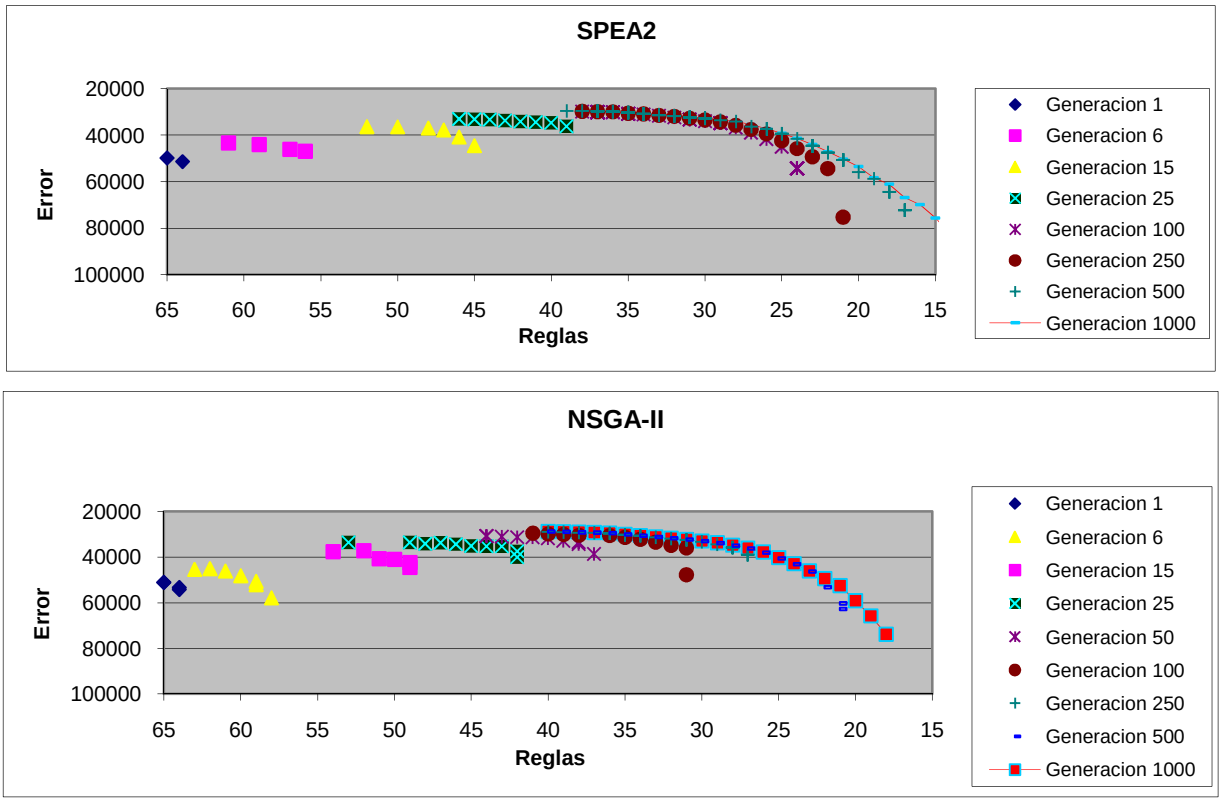


Figura 2: Frentes del pareto del SPEA2 y del NSGA-II para el problema eléctrico.

5 CONCLUSIONES

Los modelos difusos lingüísticos en los que la cooperación entre la BR y el Sistema de Inferencia son optimizados mediante algoritmos evolutivos con el doble objetivo de la precisión y la interpretabilidad en base al número de reglas constituyen una política de diseño útil ya que los resultados del estudio experimental han mostrado que además de ofrecer un conjunto de soluciones con diferente equilibrio entre las citadas características, ofrecen modelos más compactos y precisos que aquellos que utilizan como único objetivo simplemente la precisión.

No obstante, observando las Figuras 1 y 2, consideramos ampliar el estudio tratando de centrar el algoritmo evolutivo multiobjetivo en la zona central de la frontera del pareto, ya que los casos con gran cantidad de reglas y gran precisión, o pequeño número de reglas y escasa precisión son ambos menos interesantes. Centrar y reducir el espacio de búsqueda consideramos que mejoraría los resultados ya que concentraría el esfuerzo de la búsqueda en la zona de soluciones de más interés.

Agradecimientos

Este trabajo está soportado por el proyecto TIN2005-08386-C05-01 del Ministerio de Educación y Ciencia y los proyectos P05-TIC-00531 y P07-TIC-03179 de la Consejería de Innovación Ciencia y Empresa de la Junta de Andalucía.

Referencias

[1] J. Alcala-Fdez, F. Herrera, F. Márquez, A. Peregrín. Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *International Journal of Intelligent Systems* 22:9, Pág. 1035-1064, 2007

[2] J. Casillas, O. Cordón, F. Herrera. COR: A methodology to improve ad hoc data-driven linguistic rule learning methods by inducing cooperation among rules. *IEEE Transactions on Systems, Man and Cybernetics —Part B: Cybernetics*, 32:4, Pág. 526–537, 2002

[3] J. Casillas, O. Cordón, F. Herrera, L. Magdalena(Eds). Accuracy improvements in linguistic fuzzy modeling. *Springer-Verlag*, 2003.

[4] J. Casillas, O. Cordón, M.J. del Jesus, F. Herrera. Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. *IEEE Trans. on Fuzzy Systems* 13:1, Pág.13-29, 2005.

[5] C.A. Coello Coello, D.A. Van Veldhuizen, G.B. Lamont. Evolutionary algorithms for solving multi-objective problems. *Kluwer Academic Publishers*, 2002.

[6] O. Cordón, F. Herrera, L. Sánchez. Solving electrical distribution problems using hybrid evolutionary data analysis techniques. *Applied Intelligence* 10, Pág.5-24, 1999.

[7] J.S. Cho, D.J. Park. Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network. *Journal of Intelligent and Fuzzy Systems,* 8, Pág. 99-100, 2000.

[8] O. Cordón, F. Herrera, A. Peregrín. Applicability of the Fuzzy Operators in the Design of Fuzzy Logic Controllers. *Fuzzy Sets and Systems.* 86, Pág. 15-41, 1997.

[9] O. Cordón, F. Herrera. A. Peregrín. Searching for Basic Properties Obtaining Robust Implication Operators in Fuzzy Control. *Fuzzy Sets and* Systems**,** 111, Pág. 237-251, 2000.

[10] O. Cordón, F. Herrera, F. Márquez, A. Peregrín. A Study on the Evolutionary Adaptive Defuzzification Methods in Fuzzy Modelling. *Internacional Journal of Hybrid Intelligent Systems*, 1:1, Pág. 36-48, 2004.

[11] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolutionary Computation* 6:2, Pág.182-197, 2002.

[12] L.J. Eshelman. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *Foundations of genetic Algorithms* 1, Pág. 265-283, 1991.

[13] L.J. Eshelman, J.D. Schaffer. Real-coded genetic algorithms and interval-schemata. *Foundations of Genetic Algorithms* 2, Pág.187-202, 1993.

[14] M. Gupta, J. Qi. Theory of T-norms and Fuzzy Inference Methods. *Fuzzy Sets and Systems*, 40, Pág. 431-450, 1991.

[15] F. Herrera, F. Márquez, A. Peregrín. Genetic Adaptation of Rule Connectives and Conjunction Operators in Fuzzy Rule Based System: An Experimental Comparative. *Third International Conference of the European Society for Fuzzy Logic and Technology Study*. Pág. 100-104, 2003

[16] F. Herrera, F. Márquez, A. Peregrín. Cooperación de Reglas Difusas basada en la Adaptación Evolutiva de los Operadores de Conjunción y los Defuzzificadores. *Coms. del XII Congr. Español sobre Tecnologías y Lógica Fuzzy.* Pág. 231-236, 2004.

[17] H. Ishibuchi, K. Nozaki, H. Tanaka, Y. Hosaka, M. Matsuda. Empirical study on learning in fuzzy systems by rice taste analysis. *Fuzzy Sets Syst*., 64, Pág. 129–144, 1994.

[18] B.D. Liu, C.Y. Chen C-Y., J.Y. Tsao. Design of adaptive fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms. *IEEE*

Transactions on Systems, Man, and Cybernetics, 31, nº 1, Pág. 32-53, 2001.

[19] F.A. Márquez, A. Peregrín y F. Herrera Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Linguistic Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*. 15:6 Pág. 1162-1178, 2007.

[20] M. Mizumoto. Pictorial Representations of Fuzzy Connectives, Part I: Cases of T-norms, T-conorms and Averaging Operators. *Fuzzy Sets and Systems,* 31, Pág. 217-242, 1989.

[21] L.X. Wang, J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man, and Cybernetics* 22:6, Pág.1414-1427, 1992.

[22] E. Zitzler, M. Laumanns, L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. *Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems* (EUROGEN'01), Pág.95-100, 2001.

APRENDIZAJE EVOLUTIVO DE SISTEMAS DIFUSOS LINGÜÍSTICOS BASADO EN LA COOPERACIÓN ENTRE LOS OPERADORES Y LAS FUNCIONES DE PERTENENCIA

A.A. Márquez **F.A. Márquez** **A. Peregrín**

Dept. de Tecnologías de la Información. Universidad de Huelva 21071 – Huelva,
{amarquez, alfredo.marquez, peregrin}@dti.uhu.es

Resumen

Este trabajo presenta un modelo evolutivo que aprende de forma simultánea los parámetros de los operadores difusos adaptativos del Sistema de Inferencia y la Defuzzificación y el ajuste de las funciones de pertenencia con el propósito de obtener un mayor nivel de cooperación entre ambos elementos: Base de Conocimiento y operadores difusos, que se traduzca en una mejor precisión del modelo manteniendo un buen nivel de interpretabilidad. Por tanto, este trabajo propone un mecanismo para que se pueda dar una sinergia positiva entre dos elementos de diseño diferentes que siendo aprendidos de forma simultánea ofrecen mejores resultados en las aplicaciones.

Palabras Clave: Modelado Difuso Lingüístico, Interpretabilidad, Defuzzificación Adaptativa, Sistema de Inferencia Adaptativo, Ajuste de las Particiones Difusas.

1 INTRODUCCION

El Modelado Difuso Lingüístico (MDL) [3] permite la obtención de modelos más fácilmente interpretables por el ser humano. Sin embargo, la precisión es una propiedad contradictoria con la citada interpretabilidad, por lo que en sistemas complejos, siendo ambas propiedades deseables, suele ser necesario llegar a un compromiso entre ellas. Por este motivo diferentes autores han realizado diferentes propuestas en las que han tratado de dotar al MDL de una mayor precisión preservando el mayor grado posible de interpretabilidad.

Generalmente, de las dos tareas básicas de diseño que son la obtención de la Base de Conocimiento (BC) y la elección de los operadores del Sistema de Inferencia y la Interfaz de Defuzzificación, la primera es la más ampliamente estudiada.

En relación con el aprendizaje de la BC, una de las técnicas más conocidas para mejorar la precisión de los Sistemas Basados en Reglas Difusas (SBRDs) es el llamado ajuste de las funciones de pertenencia contenidas en la Base de Datos (BD) que consiste en el refinamiento de la misma, manteniendo la Base de Reglas (BR) fija. Concretamente, el proceso de ajuste consiste en una variación en la forma y en la ubicación de su universo de discurso de las funciones de pertenencia predefinidas.

En cuanto a la elección de los operadores difusos, como se indicaba anteriormente existen menos aportaciones, pero recientemente su relevancia es creciente en tanto se ha revelado como un recurso útil. El uso de operadores adaptativos tanto en el Sistema de Inferencia [1] como en la Interfaz de Defuzzificación [8], o en ambos de modo simultáneo [12] ha demostrado ofrecer mejoras significativas en la precisión de diversas aplicaciones.

En [14] se muestra un modelo que aprende tanto la BR, partiendo de la BD fija, como los operadores adaptativos en el Sistema de Inferencia e Interfaz de Defuzzificación, obteniéndose sistemas con un alto nivel de cooperación entre las reglas y los operadores, que se traduce en una mejora significativa de la precisión.

En este trabajo se propone un modelo que al igual que el referido en el párrafo anterior, aprende elementos tanto de la BC como del Sistema de Inferencia y Defuzzificación, pero en este caso eligiendo el ajuste de las funciones de pertenencia que se almacenan en la BD, en un mismo modelo evolutivo, esto es, favoreciendo la cooperación entre los operadores del Sistema de Inferencia y las funciones de pertenencia al ser aprendidos de forma simultánea.

El trabajo se organiza de la siguiente forma: En la Sección 2 se introducen los operadores difusos adaptativos. En la Sección 3 se describe en primer lugar la metodología de aprendizaje del ajuste de la las

funciones de pertenencia mediante algoritmos evolutivos, y posteriormente la propuesta del modelo evolutivo cooperativo que se plantea en este trabajo. En la Sección 4 se desarrolla un breve estudio experimental, y finalmente se incluye una Sección dedicada a las conclusiones.

2 OPERADORES DIFUSOS ADAPTATIVOS

En esta sección se describen los fundamentos del Sistema de Inferencia y de la Interfaz de Defuzzificación adaptativos que serán utilizados posteriormente en la siguiente sección para el modelo propuesto.

2.1. OPERADOR DE CONJUNCIÓN ADAPTATIVO

Los SBRDs para modelado utilizan reglas de tipo SI - ENTONCES de la forma siguiente:

$$R_i : Si\ X_{i1}\ es\ A_{i1}\ y \ldots y\ X_{im}\ es\ A_{im}\ entonces\ Y\ es\ B_i$$

con i = 1 hasta N reglas, siendo X_{i1} hasta X_{im} las entradas e Y la salida, y con A_{i1} hasta A_{im} y B_i los antecedentes y consecuentes respectivamente.

La expresión de la Regla Composicional de Inferencia en el modelado difuso con fuzzificación puntual es la que se muestra en la expresión (1), donde $\mu_{B'}$ es la función de pertenencia del consecuente inferido, I(·) es el operador de implicación, C(·) el operador de conjunción, μ_{Ai} al tratarse de fuzzificación puntual son los puntos de corte de las entradas discretas ($x_{1,\ldots,}x_n$) del sistema con las funciones de pertenencia de los antecedentes de las reglas, y μ_B el consecuente de la regla.

$$\mu_{B'}(y) = I\,(C\,(\mu_{A1}(x_1), \ldots, \mu_{An}(x_n)), \mu_B(y)), \qquad (1)$$

Por tanto, el sistema de inferencia emplea dos operadores: el de conjunción, C(·) para calcular el grado de emparejamiento, y el de implicación, I(·) para calcular a continuación el consecuente inferido.

En [1 y 11] se desarrollaron estudios sobre los diferentes métodos de introducir parámetros en el Sistema de Inferencia, observándose menor influencia en el uso de operadores de implicación adaptativos frente al de los de conjunción. Por este motivo, en este trabajo se utiliza la parametrización del Sistema de Inferencia exclusivamente mediante el operador de conjunción.

Las T-normas son una familia de operadores frecuentemente utilizados en el diseño de sistemas difusos, modelando tanto el operador de conjunción como el de implicación [10] y ofreciendo buen comportamiento [7]. En este trabajo, el Sistema de Inferencia Adaptativo se basa en utilizar las t-normas parametrizadas de Dubois, Dombi y Frank recogidas en [15], cuyas expresiones se muestran en (2), (3) y (4).

Dubois ($0 \le \alpha \le 1$)
$$T_{Dubois}(x, y, \alpha) = \frac{x \cdot y}{Max(x, y, \alpha)} \qquad (2)$$

Dombi ($\alpha > 0$)
$$T_{Dombi}(x, y, \alpha) = \frac{1}{1 + \sqrt[\alpha]{\left(\frac{1-x}{x}\right)^{\alpha} + \left(\frac{1-y}{y}\right)^{\alpha}}} \qquad (3)$$

Frank ($\alpha > 0), (\alpha \neq 1$)
$$T_{Frank}(x, y, \alpha) = \log_{\alpha}\left[1 + \frac{(\alpha^x - 1)(\alpha^y - 1)}{\alpha - 1}\right] \qquad (4)$$

La Tabla 1 muestra la relación entre las cinco t-normas clásicas y las t-normas parametrizadas anteriormente citadas dependiendo del valor del parámetro α.

Tabla 1: Relación entre las t-normas clásicas y las t-normas parametrizadas dependiendo del valor del parámetro.

	$T_{Mín}$	T_{Ham}	T_{Alg}	T_{Einst}	T_{Acot}	$T_{Drást}$
T_{Frank}	$\to 0$		$\to 1$		∞	
T_{Dombi}	∞	1				$\to 0$
T_{Dubois}	0		1			

Se podrían considerar dos modelos distintos dependiendo de si se opta por utilizar un sólo parámetro α para toda la expresión del operador de conjunción, o distintos parámetros α_i, uno para cada regla de la BR. Un sólo parámetro permite ajustar globalmente el comportamiento del conectivo mientras que en el caso de los parámetros individuales, éstos ajustan dicho comportamiento localmente para cada regla. El modelo de un sólo parámetro ofrece peores resultados prácticos en cuanto a precisión [1], por lo que en este trabajo se utiliza el modelo de un parámetro para cada regla.

La utilización de operadores de conjunción adaptativos modula la influencia del grado de emparejamiento atenuándolo o acentuándolo. Este cambio no es lineal debido a que las expresiones de las t-normas adaptativas no son lineales. El efecto que producen es pues semejante a un mecanismo para modificar el significado lingüístico de la estructura de las reglas [13], que consiste en relajar la estructura de la regla cambiando el significado de las etiquetas involucradas.

En este caso, el parámetro en el operador de conjunción juega el mismo papel que cambiar las funciones de pertenencia asociadas con los antecedentes de las reglas. Estos antecedentes cambian de la misma forma que lo hace el operador o función aplicada. Las t-normas parametrizadas como operadores de conjunción sólo producen efectos de concentración.

2.2. MÉTODO DE DEFUZZIFICACIÓN ADAPTATIVO

La expresión formal (5) se emplea generalmente para generar métodos de defuzzificación adaptativos,

$$y_0 = \frac{\sum_{i}^{N} f(h_i) \cdot V_i}{\sum_{i}^{N} f(h_i)}, \qquad (5)$$

donde h_i es el grado de emparejamiento, $f(h_i)$ es por tanto un término funcional del grado de emparejamiento y V_i es el máximo valor (MV_i) o el centro de gravedad (CG_i). Se trata de la expresión de un método de defuzzificación que actúa en Modo B, es decir, primero defuzzifica la aportación individual de la inferencia obtenida con cada regla y posteriormente se computa una suma ponderada.

El término funcional puede utilizar un solo parámetro β, o bien un parámetro por cada regla de la BC, β_i, que es la opción utilizada en este trabajo por obtener mejores resultados prácticos [8]. Por otro lado, el término funcional puede ser de tipo producto o potencia, como se muestra en las expresiones (6) y (7).

$$f(h_i) = h_i \cdot \beta_i, \quad f(h_i) = h_i^{\beta_i}. \qquad (6), (7)$$

La combinación de estos dos funcionales con el MV o el CG genera cuatro métodos de defuzzificación. El empleo del funcional de tipo potencia con un parámetro para cada regla es equivalente a una de las técnicas empleadas para realizar la modificación lingüística de la estructura de la regla [13]. Por otro lado, el funcional de tipo producto tiene un efecto equivalente al uso de reglas con pesos [5]. A continuación se muestra un conjunto de reglas con pesos, donde éstos son denominados por w_i:

Si X_{11} es A_{11} y... y X_{1n} es A_{1n} entonces Y es B_1 con w_1
Si X_{21} es A_{21} y... y X_{2n} es A_{2n} entonces Y es B_2 con w_2
...
Si X_{n1} es A_{n1} y... y X_{nn} es A_{nn} entonces Y es B_n con w_n

En este trabajo se ha optado por el funcional de tipo producto con un parámetro para cada regla por sus buenos resultados empíricos mostrados en [8], y utilizando CG, cuya expresión es (8).

$$y_0 = \frac{\sum_{i}^{N} h_i \cdot \beta_i \cdot CG_i}{\sum_{i}^{N} h_i \cdot \beta_i}, \qquad (8)$$

Este método de defuzzificación es equivalente al conocido WCOA cuando los parámetros se igualan a la unidad.

3. MODELO EVOLUTIVO COOPERATIVO DE APRENDIZAJE

Como se ha indicado con anterioridad, el aprendizaje de la BC de un SBRD se basa generalmente en la existencia de una BD previamente definida [16]. Generalmente, la BD es obtenida fijando tanto un número de etiquetas para cada variable como los parámetros de dichas etiquetas y distribuyéndolas de modo uniforme en el universo de discurso de la variable. La BD tiene una gran influencia en el comportamiento del SBRD, y prefijarlas para obtener la BR condiciona en gran medida la bondad final de la BC resultante.

Otros enfoques realizan un ajuste a posteriori de la BD [6]. Para ello, aprenden los parámetros o la forma de las funciones de pertenencia, mejorando la precisión del SBRD. No obstante, el proceso de aprendizaje está sesgado por la BD y BR iniciales, quedando fijos aspectos tales como la granularidad de las particiones, o el número y la estructura de las reglas.

En el presente trabajo, en una primera aproximación, hemos tomado este segundo enfoque. Además, con el fin de minimizar la pérdida de interpretabilidad que se produce al realizar el aprendizaje de las etiquetas, los márgenes de variación de éstas se han sesgado para que sólo puedan variar en cierta medida, tal como se comentará posteriormente en la descripción de las asignaciones de valores iniciales de esta parte del cromosoma.

El modelo cooperativo presentado en este trabajo utiliza un modelo evolutivo basado en un Algoritmo Genético (AG) de tipo CHC [9] cuyo cromosoma tiene un esquema de codificación triple (CS_T+CS_C+CS_D), donde CS_T codifica las funciones de pertenencia, CS_C los parámetros α_i del operador de conjunción, y CS_D los parámetros β_i del método de defuzzificación.

La parte del cromosoma CS_T utiliza un esquema con codificación real, siendo m_i el número de etiquetas de cada una de las n variables que comprenden la BD,

$$C_i = (a_i^1, b_i^1, c_i^1, \ldots, a_i^m, b_i^m, c_i^m),\ i = 1, \ldots, n$$

$$CS_T = C_1 C_2 \ldots C_n$$

Cada gen toma valores en un cierto intervalo el cual es obtenido a partir de la BD inicial, de tal forma que, por cada etiqueta lingüística $C_i^j = a^j, b^j, c^j$, el intervalo de variación se calcula de la siguiente forma:

$$\left[I^l_{a_j}, I^r_{a_j}\right] = \left[a_j - \left(b_j - a_j\right)/2,\ a_j + \left(b_j - a_j\right)/2\right]$$
$$\left[I^l_{b_j}, I^r_{b_j}\right] = \left[b_j - \left(b_j - a_j\right)/2,\ b_j + \left(c_j - b_j\right)/2\right]$$
$$\left[I^l_{c_j}, I^r_{c_j}\right] = \left[c_j - \left(c_j - b_j\right)/2,\ c_j + \left(c_j - b_j\right)/2\right]$$

La zona del cromosoma dedicada a los parámetros de los operadores, tiene dos partes:

- La del operador de conjunción, CS_C, que utiliza N parámetros α_i (genes), uno para cada R_i de la BR. Cada gen toma valores en el intervalo [0,1] para la t-norma de Dubois, (0,10] para Dombi y (0,100] para Frank.

- La del método de defuzzficación, CS_D, que utiliza N parámetros β_i, uno para cada R_i de la BR. Cada gen en este caso toma valores en el intervalo [0,10], suficiente para actuar como peso de esa regla.

El AG elegido es un CHC [9] considerado como un modelo evolutivo con un buen balance entre diversidad y convergencia en problemas con grandes espacios de búsqueda.

En cada generación, el algoritmo CHC utiliza una población de padres de tamaño M para generar una población intermedia de M individuos, los cuales son aleatoriamente emparejados y utilizados para generar M' descendientes. Se mantiene entonces una competición por la supervivencia, donde los mejores M cromosomas de las poblaciones de padres y descendientes son seleccionados para constituir la nueva generación.

En la fase de recombinación no se aplica ningún proceso de mutación. En su lugar, cuando la población converge o el proceso de búsqueda deja de progresar adecuadamente (el umbral de cruce llega a 0 y no se generan nuevos descendientes), la población se reinicia. El cromosoma que representa la mejor solución hasta ese momento se utiliza como patrón para generar la nueva población en un cromosoma, mientras que el resto es aleatorio.

Aunque el algoritmo CHC fue concebido para cromosomas con codificación binaria, existen versiones para su uso con cromosomas con codificación real como la empleada en la parte correspondiente a los parámetros de este trabajo, la cual utiliza el BLX-α como operador de cruce (α=0.5). El cálculo de la distancia de Hamming se realiza convirtiendo los genes reales en una representación en cadenas y hallando la distancia de Hamming teniendo en cuenta si difiere o no cada uno de los caracteres. Sólo aquellas cadenas con una distancia mayor del umbral son combinados. El umbral se inicializa a L/4 siendo L la longitud de la cadena o cromosoma. Cuando ningún descendiente es insertado en la nueva población, el umbral se reduce en una unidad.

El tamaño de la población es de 50 individuos aleatoriamente inicializados, salvo uno, con la siguiente configuración:

- La parte de la BD (funciones de pertenencia), CS_T, con las funciones de pertenencia inicialmente almacenadas en la BD original como el primer individuo, y el resto como se ha indicado se generan aleatoriamente en el correspondiente intervalo de variación.

- La parte de parámetros de los operadores difusos:

  - Operador de conjunción, CS_C , con los N genes inicializados a 0 para la t-norma de Dubois (equivalente a la t-norma del mínimo), a 1 en caso de la de Dombi y en 0.5 en caso de la de Frank.

  - Método de defuzzificación, CS_D, con los N genes que representan los pesos de las reglas inicializados a 1, con objeto de partir con el WCOA.

El modelo evolutivo utiliza un solo umbral, tanto para la parte del cromosoma del ajuste de la BD como para la parte de los operadores difusos.

La función objetivo utilizada es el clásico Error Cuadrático Medio (ECM), mostrado en (9).

$$ECM\ (S[i])_B = \frac{\frac{1}{2}\sum_{k=1}^{r}(y_k - S[i](x_k))^2}{P} \qquad (9)$$

donde S[i] es el modelo difuso cuyo Sistema de Inferencia utiliza la t-norma de Dubois, Dombi o Frank (i=1,…,3) como operador de conjunción y la t-norma del mínimo como operador de implicación, y el método de defuzzificación WCOA adaptativo con pesos, utilizando una BD concreta decodificada. Esta medida utiliza un conjunto de datos para evaluación constituido por P pares de datos numéricos $Z_k=(x_k, y_k)$, k=1,...,P, siendo x_k los valores de las variables de entrada y y_k los correspondientes a las variables de salida.

4. ESTUDIO EXPERIMENTAL

En esta Sección se propone un estudio experimental en el que se utiliza una aplicación real, y se compara la precisión obtenida con distintos modelos difusos cuyos resultados se someten a un estudio estadístico con objeto de mostrar lo significativo de las mejoras de precisión obtenidas con el modelo cooperativo propuesto.

4.1. DESCRIPCIÓN DE LA APLICACIÓN

La aplicación considerada es el problema de distribución del sistema eléctrico español descrito en [4], que consiste en el cálculo de los costes óptimos

teóricos de la línea de media tensión en núcleos urbanos.

Para ello se dispuso de un conjunto de datos de 1059 ciudades que constan de cuatro variables de entrada y una de salida. El dominio de las variables de entrada y salida fue dividido en cinco etiquetas {MP, P, M, G, MG}.

Utilizamos una BR con 65 reglas lingüísticas obtenidas con el método de WM [16] a partir de un conjunto de datos de entrenamiento con el 80% de los datos disponibles, es decir, 847 de las 1059 ciudades, tomados de modo aleatorio.

Se ha utilizado un modelo de validación cruzada con 5 particiones de datos, esto es, 5 particiones aleatorias al 20% (4 de ellas con 212 ejemplos y una con 211), y la combinación de cuatro particiones (80%) como entrenamiento y la partición restante (20%) como test. Así se tienen 5 particiones al 80% y 20% en entrenamiento y test respectivamente.

4.2 METODOLOGÍA DE COMPARACIÓN

La Tabla 2 resume los modelos que se han construido con objeto de comparar su precisión. La primera línea contiene el de referencia, WM. A continuación los tres modelos adaptativos evolutivos, esto es, el modelo con aprendizaje de los operadores difusos (WM + C-D), el modelo con aprendizaje de la BD, es decir, las funciones de pertenencia (WM + T), y el modelo anterior sobre el que se aprenden a posteriori los parámetros de los operadores difusos (WM + T + C-D). Por último aparece el modelo cooperativo presentado en este trabajo (WM + T-C-D), en el que el aprendizaje de las funciones de pertenencia y de los operadores difusos se realiza al mismo tiempo.

Tabla 2. Modelos Difusos del Estudio Experimental

Abreviatura	Aprendizaje Funciones Pertenencia	Conjun. Adapt.	Defuzz. Adapt.
Modelo de Referencia			
WM			
Modelos Adaptativos Evolutivos Op. Difusos			
WM + C_{DUBOIS}-D		●	●
WM + C_{DOMBI}-D		●	●
WM + C_{FRANK}-D		●	●
Modelo Adaptativo Evolutivo BD			
WM + T	●		
Modelos Adaptativos Evolutivos BD + Op. Difusos			
WM + T + C_{DUBOIS} -D	●	●	●
WM + T + C_{DOMBI} -D	●	●	●
WM + T + C_{FRANK} -D	●	●	●
Modelo Evolutivo Cooperativo propuesto			
WM + T- C_{DUBOIS} -D	●	●	●
WM + T- C_{DOMBI} -D	●	●	●
WM + T- C_{FRANK} -D	●	●	●

En los modelos donde existe aprendizaje de los operadores difusos se presenta uno para cada t-norma adaptativa utilizada como conjunción (Dubois, Dombi y Frank).

Para cada una de las 5 particiones generadas, cada modelo evolutivo se ha ejecutado 6 veces, dando lugar a 30 ejecuciones de cada modelo, resultado de la combinación de 6 semillas para el generador de números aleatorios sobre las 5 particiones diferentes de datos obtenidas.

Los distintos modelos se han comparado utilizando el la medida del ECM (9) que se introdujo en la Sección 3, si bien se computa la media aritmética de los 30 resultados. La Tabla 3 muestra las evaluaciones realizadas por el modelo evolutivo CHC para cada modelo difuso.

Tabla 3. Evaluaciones realizadas por el algoritmo evolutivo CHC utilizado

Modelo	Evaluaciones
WM + T	30.000
WM + C-D	30.000
WM + T + C-D	60.000
WM + T-C-D	60.000

4.3 RESULTADOS Y ANÁLISIS

En esta sección se muestran los resultados obtenidos y el análisis de los mismos. En primer lugar los correspondientes al ECM de los modelos difusos utilizados, y en segundo lugar los resultados resumidos del análisis estadístico.

4.3.1 RESULTADOS GENERALES

La Tabla 4 muestra los ECMs de los distintos modelos difusos construidos.

Tabla 4. ECM obtenido con los diferentes modelos para el problema de distribución del sistema eléctrico.

Modelo	$ECM_{entrenamiento}$	ECM_{prueba}
WM	56135,75	56359,42
WM + C_{DUBOIS}-D	22561,77	25492,77
WM + C_{DOMBI}-D	28627,74	31558,00
WM + C_{FRANK}-D	25691,23	28895,61
WM + T	14962,14	18122,23
WM + T + C_{DUBOIS} –D	8435,83	12203,10
WM + T + C_{DOMBI} –D	9193,79	13020,94
WM + T + C_{FRANK} –D	14759,64	18922,81
WM + T- C_{DUBOIS} –D	**5990,18**	**9101,50**
WM + T- C_{DOMBI} –D	7193,25	10913,21
WM + T- C_{FRANK} –D	7124,48	9990,57

En ella podemos observar:

A) En cuanto a los modelos adaptativos evolutivos (WM + T, WM + C-D y WM + T-C-D):

- Como es bien conocido, el ajuste de las funciones de pertenencia es una herramienta muy útil para mejorar de forma importante la precisión del modelo.
- La mejora que presenta aplicar, al resultado del ajuste de las funciones de pertenencia, el aprendizaje de los operadores difusos es también muy relevante, por lo que el ajuste de las funciones de pertenencia y el aprendizaje de los operadores son dos métodos distintos pero compatibles para mejorar la precisión.
- Entre las t-normas parametrizadas, la de Dubois, además de ser computacionalmente más eficiente, es la que mejores resultados obtiene, tal como se observó en [7].

B) En cuanto al modelo de aprendizaje cooperativo evolutivo, cuyos resultados se encuentran en las tres últimas filas (WM + T-C-D):

- La cooperación entre ambos elementos, operadores y funciones de pertenencia, se manifiesta presentando la mejor precisión del estudio experimental.

4.3.2 ESTUDIO ESTADÍSTICO

Para comparar los resultados obtenidos con los diferentes modelos y confirmar que las mejoras encontradas son significativas, se ha realizado el estudio estadístico que se describe a continuación. En primer lugar se realizó el análisis ANOVA de un factor [2] para cada modelo, siendo el factor los modelos utilizados en los test de los conjuntos de datos. Por extensión, no se han incluido las tablas en este trabajo. Dado que se encontraron diferencias significativas para todos los modelos con respecto a los valores medios asociados, se realizó el test de Tamhane [2] con un coeficiente de confianza del 95%, debido a que la hipótesis de igualdad de varianzas ha sido rechazada en los análisis de todos los modelos.

La Tabla 5 resume los resultados obtenidos en el test de Tamhane para cada aplicación. El signo (+) significa que el modelo de la fila seleccionada mejora al modelo de la columna seleccionada, mientras que el signo (-) significa lo contrario. El signo (≈) significa que son similares. Las tablas deben leerse comenzando por la fila, por ejemplo: *en la Tabla 6, el modelo* WM + T-C_{DUBOIS}-D *(fila 9) mejora al modelo WM + T (columna 3).*

Tabla 5. Resumen del test estadístico de Tamhane para el problema de distribución del sistema eléctrico

	WM	WM +C-D_{DUBOIS}	WM +C-D_{DOMBI}	WM +C-D_{FRANK}	WM + T	WM + T +C-D_{DUBOIS}	WM + T +C-D_{DOMBI}	WM + T +C-D_{FRANK}	WM + T - C-D_{DUBOIS}	WM + T - C-D_{DOMBI}	WM + T - C-D_{FRANK}
WM		-	-	-	-	-	-	-	-	-	-
WM + C_{DUBOIS}-D	+		≈	≈	-	-	-	-	-	-	-
WM + C_{DOMBI}-D	+	≈		≈	-	-	-	-	-	-	-
WM + C_{FRANK}-D	+	≈	≈		-	-	-	-	-	-	-
WM + T	+	+	+	+		-	-	-	-	-	-
WM + T + C_{DUBOIS} -D	+	+	+	+	+		≈	≈	-	-	-
WM + T + C_{DOMBI} -D	+	+	+	+	+	≈		≈	-	-	-
WM + T + C_{FRANK} -D	+	+	+	+	+	≈	≈		-	-	-
WM + T- C_{DUBOIS} -D	+	+	+	+	+	+	+	+		≈	≈
WM + T- C_{DOMBI} -D	+	+	+	+	+	+	+	+	≈		≈
WM + T- C_{FRANK} -D	+	+	+	+	+	+	+	+	≈	≈	

5. CONCLUSIONES

En este trabajo se ha propuesto un modelo que consigue la cooperación entre las funciones de pertenencia descritas en la BD y los operadores difusos basada en el aprendizaje evolutivo de ambos elementos a la vez. Las descripciones de las etiquetas lingüísticas y la forma en la que deben trabajar los operadores difusos, conjunción y método de defuzzificación para cada regla, son elementos distintos, y el hecho de encontrarse en un mismo cromosoma del algoritmo evolutivo sólo busca la cooperación en el aprendizaje de ambos elementos.

Este trabajo pues pone de manifiesto de nuevo que en el marco del compromiso entre precisión e interpretabilidad, el ajuste de los operadores difusos es también un elemento valioso, y que combinado con el aprendizaje de los elementos de la BC, en este caso con las funciones de pertenencia, se obtienen mejoras significativas en la precisión gracias a la sinergia positiva que adquieren.

El estudio se ha centrado en un método para mejorar la precisión sin introducir a nuestro juicio pérdidas importantes en la interpretabilidad, si bien cuantificarla es un problema abierto y complejo. Igualmente, se tratará de ampliar el estudio con otras aplicaciones que permitan observar los efectos del modelo propuesto en otros casos.

Agradecimientos

Este trabajo está soportado por el proyecto TIN2005-08386-C05-01 del Ministerio de Educación y Ciencia y los proyectos P05-TIC-00531 y P07-TIC-03179 de la Consejería de Innovación Ciencia y Empresa de la Junta de Andalucía.

Referencias

[1] J. Alcalá-Fdez, F. Herrera, F. Márquez, A. Peregrín. Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *International Journal of Intelligent Systems* 22:9, Pág. 1035-1064, 2007

[2] T. W. Anderson. An introduction to multivariate statistical analysis. *John Wiley and Sons*, 1984.

[3] J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds). Accuracy improvements in linguistic fuzzy modeling. *Springer-Verlag*, 2003.

[4] O. Cordón, F. Herrera, L. Sánchez. Solving electrical distribution problems using hybrid evolutionary data analysis techniques. *Applied Intelligence* 10, Pág.5-24, 1999.

[5] J.S. Cho, D.J. Park. Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network. *Journal of Intelligent and Fuzzy Systems,* 8, Pág. 99-100, 2000.

[6] O. Cordón, F. Herrera. A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples. *Int. J. of Approximate Reasoning* 17:4, Pág.369-407, 1997.

[7] O. Cordón, F. Herrera, A. Peregrín. Applicability of the Fuzzy Operators in the Design of Fuzzy Logic Controllers. *Fuzzy Sets and Systems.* 86, Pág. 15-41, 1997.

[8] O. Cordón, F. Herrera, F. Márquez, A. Peregrín. A Study on the Evolutionary Adaptive Defuzzification Methods in Fuzzy Modelling. *Internacional Journal of Hybrid Intelligent Systems*, 1:1, Pág. 36-48, 2004.

[9] L.J. Eshelman. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *Foundations of genetic Algorithms* 1, Pág.265-283, 1991.

[10] M. Gupta, J. Qi. Theory of T-norms and Fuzzy Inference Methods. *Fuzzy Sets and Systems*, 40, Pág. 431-450, 1991.

[11] F. Herrera, F. Márquez, A. Peregrín. Genetic Adaptation of Rule Connectives and Conjunction Operators in Fuzzy Rule Based System: An Experimental Comparative. *Third International Conference of the European Society for Fuzzy Logic and Technology Study*. Pág. 100-104, 2003

[12] F. Herrera, F. Márquez, A. Peregrín. Cooperación de Reglas Difusas basada en la Adapación Evolutiva de los Operadores de Conjunción y los Defuzzificadores. *Coms. del XII Congr. Español sobre Tecnologías y Lógica Fuzzy.* Pág. 231-236, 2004.

[13] B.D. Liu, C.Y. Chen C-Y., J.Y. Tsao. Design of adaptative fuzzy logic controller based on linguistic-hedge concepts and genetic algorithtms. *IEEE Transactions on Systems, Man, and Cybernetics*, 31:1, Pág. 32-53, 2001.

[14] F.A. Márquez, A. Peregrín y F. Herrera Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Lingüistic Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*. 15:6 Pág. 1162-1178, 2007.

[15] M. Mizumoto. Pictorial Representations of Fuzzy Connectives, Part I: Cases of T-norms, T-conorms and Averaging Operators. *Fuzzy Sets and Systems,* 31, Pág. 217-242, 1989.

[16] L.X. Wang, J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man, and Cybernetics* 22:6, Pág.1414-1427, 1992.

Influencia de la granularidad y las medidas de calidad en el modelo evolutivo SDIGA de Descubrimiento de Subgrupos

Pedro González [1] **Cristóbal J. Carmona** [1] **María José del Jesus** [1] **Francisco Herrera** [2]

[1] Depto. Informática, Universidad de Jaén, Jaén, 23071, España, {pglez,ccarmona,mjjesus}@ujaen.es
[2] Depto. CCIA, ETSIIT, Universidad de Granada, Granada, 18071, España, herrera@decsai.ugr.es

Resumen

En este trabajo presentamos un estudio preliminar sobre la influencia de distintos aspectos en el comportamiento del modelo evolutivo SDIGA de extracción de reglas difusas de descripción de subgrupos. El estudio se centra en la influencia de las medidas de calidad utilizadas por el algoritmo, el tipo de regla con el que se representa el conocimiento extraído y la granularidad seleccionada para las variables continuas. Se determina que los mejores resultados se obtienen utilizando como medidas de calidad para SDIGA la confianza difusa, el soporte nítido sobre los ejemplos de la clase y el interés, que el tipo de regla a utilizar se puede dejar a elección del experto, y que la granularidad más adecuada para el conjunto de bases de ejemplos estudiado es 3 ó 5 etiquetas.

Palabras Clave: Descubrimiento de Subgrupos, Minería de Datos, Reglas difusas, Algoritmos Genéticos.

1 INTRODUCCIÓN

La extracción de conocimiento se puede abordar desde dos perspectivas distintas: mediante un proceso de *inducción predictiva*, en el que se intenta obtener conocimiento para clasificación o predicción, o mediante un proceso *inducción descriptiva* cuyo objetivo fundamental es descubrir conocimiento de interés dentro de los datos. El descubrimiento de subgrupos [7][12] es un tipo de inducción descriptiva en el que, dado un conjunto de datos y una propiedad de los mismos que tenga interés para el usuario, encontrar (y describir) subgrupos relevantes respecto a esa propiedad. Un algoritmo de descubrimiento de subgrupos debe extraer reglas que representen de forma simbólica el conocimiento y que sean lo suficientemente sencillas y descriptivas como para ser utilizadas por el usuario final.

Así, la interpretabilidad del conocimiento extraído es uno de los aspectos fundamentales para cualquier algoritmo de descubrimiento de subgrupos. La lógica difusa tiene una gran afinidad con la forma en que representamos el conocimiento humano; esto hace que su uso en reglas con variables cuantitativas facilite la interpretabilidad de las mismas, su diseño, o la incorporación de conocimiento cualitativo sobre el problema [13] [4]. En la literatura especializada existen distintos modelos para la extracción de reglas de descripción de subgrupos [10] [6] [3], pero hasta donde conocemos solo los trabajos de Del Jesus et al. [5] utilizan reglas difusas como forma de representación del conocimiento.

En este estudio preliminar, nuestro objetivo es analizar la influencia de distintos aspectos sobre los resultados de los algoritmos de extracción de reglas difusas de descripción de subgrupos, como el tipo de reglas utilizadas, la granularidad, o las medidas de calidad utilizadas para la evaluación de las reglas obtenidas. Para ello se ha realizado una experimentación sobre distintos conjuntos de datos disponibles en el repositorio UCI[1] [2].

El trabajo se organiza de la siguiente forma: En la Sección 2 se describe el descubrimiento de subgrupos, y en la Sección 3 el algoritmo evolutivo SDIGA de inducción descriptiva de reglas difusas de descripción de subgrupos. El estudio experimental realizado se describe en la sección 4, y en la Sección 5 se muestran las conclusiones obtenidas.

2 DESCUBRIMIENTO DE SUBGRUPOS

Las técnicas de descubrimiento de subgrupos generan modelos basados en reglas cuya finalidad es descriptiva, empleando una perspectiva predictiva para obtenerlos [10]. Las reglas utilizadas en la tarea de descubrimiento

[1] www.ics.uci.edu/~mlearn/MLRepository.html

de subgrupos tienen la forma *Cond → Clase*, donde la propiedad de interés para el descubrimiento de subgrupos es el valor de la *Clase* que aparece en el consecuente de la regla [8], y el antecedente es una conjunción de variables (parejas atributo-valor) seleccionadas entre las variables del conjunto de datos.

El concepto de descubrimiento de subgrupos fue formulado inicialmente por Klösgen [7] y Wrobel [12], como "dado un conjunto de datos y una propiedad de esos datos que sea de interés para el usuario, buscar subgrupos que sean de mayor interés para el usuario". En este sentido, se dice que un subgrupo es interesante cuando tiene una distribución estadística inusual respecto a la propiedad en la que estamos interesados. El objetivo es descubrir propiedades características de subgrupos construyendo reglas individuales sencillas, altamente significativas y con soporte alto.

Uno de los aspectos más importantes a tener en cuenta en cualquier enfoque de inducción de reglas que describan subgrupos es la elección de las medidas de calidad a utilizar, tanto para seleccionar las reglas, como para valorar los resultados obtenidos en el proceso. A continuación se describen distintas medidas de calidad utilizadas en la tarea de descubrimiento de subgrupos.

- *Cobertura* [10]: mide el porcentaje de ejemplos cubiertos por una regla, y se define como:

$$Cob(R_i) = Cob(Cond_i \rightarrow Clase) = p(Cond_i) = \frac{n(Cond_i)}{N} \quad (1)$$

  donde $n(Cond_i)$ es el número de ejemplos que verifican la condición $Cond_i$, N es el número total de ejemplos y R_i denota la i-ésima regla. La cobertura media para un conjunto de reglas (*COB*) se calcula como la media de la cobertura de las reglas individuales.

- *Relevancia* [7]: se calcula en términos de su razón de verosimilitud (que mide la diferencia entre la distribución de probabilidad de la clase en el conjunto de ejemplos de entrenamiento cubiertos por la regla y la distribución de probabilidad de la clase en el conjunto de todos los ejemplos de entrenamiento), a través de la siguiente expresión:

$$Rel(R_i) = 2 \cdot \sum_j n(Clase_j.Cond_i) \cdot \log \frac{n(Clase_j.Cond_i)}{n(Clase_j) \cdot p(Cond_i)} \quad (2)$$

  donde para cada clase *j*, $n(Clase_j.Cond_i)$ es el número de instancias de la clase en el conjunto que cumplen el antecedente de la regla, $n(Clase_j)$ es el número de instancias de la clase, y $p(Cond_i)$ se utiliza como factor normalizador. La relevancia media del conjunto de reglas (*REL*) se calcula como la media de la relevancia de las reglas individuales.

- *Atipicidad* [9]: se mide como la precisión relativa ponderada de una regla (*WRAcc*, *weighted relative accuracy*), con la siguiente expresión:

$$Ati(R_i) = \frac{n(Cond_i)}{N} \cdot \left(\frac{n(Clase.Cond_i)}{n(Cond_i)} - \frac{n(Clase)}{N} \right) \quad (3)$$

  La atipicidad del conjunto de reglas (*ATI*) se mide como la media de la atipicidad de las reglas individuales.

- *Precisión predictiva*: mide la calidad del conjunto de reglas y se define como el porcentaje de instancias correctamente clasificadas. Para un problema binario de clasificación, la precisión se calcula como:

$$PREC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

  PREC mide la precisión del conjunto de reglas tanto sobre los ejemplos positivos como los negativos, mientras que la precisión de una regla (definida como $Prec(Cond \rightarrow Clase) = p(Clase|Cond)$) mide la precisión de una única regla sólo sobre los ejemplos positivos.

- *Soporte (o completitud)* [10]: se define como la frecuencia de ejemplos positivos cubiertos, de la forma:

$$Sop-N(R_i) = \frac{n(Clase.Cond_i)}{N} \quad (5)$$

  donde $n(Clase.Cond_i)$ es el número de ejemplos cubiertos por $Cond_i$ pertenecientes a *Clase*.
  También se suele calcular el soporte como el número de ejemplos de la clase cubiertos por la regla entre el número de ejemplos de la clase:

$$Sop_c\text{–}N(R_i) = \frac{n(Clase.Cond_i)}{n(Clase)} \quad (6)$$

  El soporte global del conjunto de reglas se calcula considerando la tasa de aciertos positivos para la unión de subgrupos:

$$SOP-N = \frac{1}{N} \sum_{Clase_j} n(Clase_j \cdot \bigvee_{Cond_i \rightarrow Clase_j} Cond_i) \quad (7)$$

  La expresión de la sumatoria corresponde al número de ejemplos distintos que verifican al menos una de las reglas del conjunto, por lo que los ejemplos que verifican varias reglas sólo se cuentan una vez.

- *Confianza*: definida como el porcentaje de ejemplos positivos cubiertos; es decir, el número de ejemplos que cumplen el antecedente y el consecuente de la

regla dividido entre el número de ejemplos que cumplen el antecedente [1]:

$$Conf-N(R_i) = \frac{n(Clase.Cond_i)}{n(Cond_i)} \quad (8)$$

La confianza del conjunto de reglas (*CONF-N*) se calcula como la media de la confianza de las reglas individuales.

3 ALGORITMO EVOLUTIVO SDIGA DE INDUCCIÓN DE REGLAS DIFUSAS DE DESCUBRIMIENTO DE SUBGRUPOS

El algoritmo SDIGA (Subgroup Discovery Iterative Genetic Algorithm) es un modelo evolutivo para la extracción de reglas difusas para la tarea de descubrimiento de subgrupos. Este algoritmo está descrito en detalle en [5], presentándose aquí de forma breve sus principales características.

En la tarea de descubrimiento de subgrupos hay un conjunto de variables descriptivas y una sola variable objetivo que describe los subgrupos. Como el objetivo es obtener un conjunto de reglas que describan subgrupos para todos los valores de la variable objetivo, el algoritmo genético (AG) de esta propuesta descubre reglas difusas donde el consecuente está prefijado a uno de los posibles valores de esta variable objetivo.

El modelo puede utilizar dos tipos de reglas: reglas *canónicas* en las que en el antecedente de las reglas sólo hay conjunciones de condiciones formadas por pares variable/valor, y reglas en *forma normal disyuntiva* (*DNF, Disjunctive Normal Form*) que admiten disyunciones en los valores de cada variable. Este tipo de reglas se pueden expresar como:

$$R_1 : \text{SI } (X_1 = LL_1^1 \text{ O } LL_1^3) \text{ Y } (X_7 = LL_7^1) \text{ ENTONCES } \quad Clase_j$$

donde LL_n^l es la etiqueta lingüística *l* de la variable *n.*

Cada solución candidata se codifica de acuerdo con el enfoque *"Cromosoma = Regla"*, representando sólo el antecedente de la regla (ya que todos los individuos de la población están asociados con el mismo valor de la variable objetivo). La información relativa a cada regla está almacenada en un cromosoma de longitud fija para el que se utiliza un modelo de representación entera (la i-ésima posición indica el valor adoptado por la i-ésima variable).

El núcleo de SDIGA es un AG que utiliza una etapa de post-procesamiento basada en una búsqueda local (un procedimiento de ascensión de colinas). El AG híbrido extrae una regla difusa sencilla e interpretable optimizando el soporte y la confianza. La etapa de post-procesamiento se aplica para incrementar la generalidad de las reglas extraídas.

Este AG híbrido se incluye en un proceso iterativo para la extracción de un conjunto de reglas que describen diferentes zonas (no necesariamente disjuntas) del espacio de búsqueda. Esto se consigue marcando las instancias cubiertas por la regla obtenida, de forma que se evita que una regla obtenida después cubra exactamente los mismos ejemplos que otra previamente extraída. De esta forma se obtienen reglas difusas diferentes, aunque no se excluye la posibilidad de que estén solapadas.

La función de evaluación del AG combina, según la siguiente expresión, tres factores: la confianza, el soporte y el grado de interés de la regla:

$$fitness(c) = \frac{\omega_1 \cdot Confianza(c) + \omega_2 \cdot Soporte(c) + \omega_3 \cdot Interés(c)}{\omega_1 + \omega_2 + \omega_3} \quad (9)$$

Estas medidas se definen de la siguiente forma:

- *Confianza*: se puede calcular mediante una expresión nítida, como en (8), o de forma difusa, como la suma del grado de pertenencia de los ejemplos de la clase a la zona determinada por el antecedente dividido entre la suma del grado de pertenencia de todos los ejemplos (independientemente de la clase a la que pertenezcan) a la misma zona:

$$Conf-D(R_i) = \frac{\sum_{E^k \in E \,/\, E^k \in Class_j} APC(E^k, R_i)}{\sum_{E^k \in E} APC(E^k, R_i)} \quad (10)$$

  donde *APC (Antecedent Part Compatibility)* es el grado de compatibilidad entre un ejemplo y el antecedente de una regla difusa, calculado según la expresión:

$$APC(E^k, R_i) = T(TC(\mu_{LL_1^1}(e_1^k),...,\mu_{LL_1^{l_1}}(e_1^k)),..., TC(\mu_{LL_{n_v}^1}(e_{n_v}^k),...,\mu_{LL_{n_v}^{l_{n_v}}}(e_{n_v}^k))) \quad (11)$$

  donde $LL_{n_v}^{l_{n_v}}$ es la etiqueta lingüística número l_{n_v} de la variable n_v, $\mu_{LL_{n_v}^{l_{n_v}}}(e_{n_v}^k)$ es el grado de pertenencia del valor de la variable n_v para el ejemplo E^s al conjunto difuso correspondiente a la etiqueta lingüística l_{n_v} para esta variable (n_v,), y *T* y *TC* son la t-norma y la t-conorma utilizadas para representar la unión e intersección difusa, respectivamente.
  La confianza difusa del conjunto de reglas (*CONF-D*) se calcula como la media de la confianza difusa de las reglas individuales.
- *Soporte*: se pude utilizar una expresión nítida, como en (5,6), o difusa, teniendo en cuenta el grado de pertenencia de los ejemplos de la clase a la zona determinada por el antecedente:

$$Sop\text{–}D(R_i) = \frac{\sum_{E^k \in E / E^k \in Class_j} APC(E^k, R_i)}{N} \quad (12)$$

donde *N* es el número total de ejemplos.
El soporte difuso del conjunto de reglas (*SOP-D*) se calcula como la media de las reglas individuales.

- *Interés*: el grado de interés se determina en esta propuesta objetivamente mediante el criterio de interés aportado por Noda y otros [11] en un proceso de modelado de dependencias. En la propuesta se utiliza sólo la parte referente al antecedente para el cálculo del interés, puesto que el consecuente está prefijado, según la expresión:

$$Int(R_i) = 1 - \left(\frac{\sum_{i=1}^{n} Ganancia(A_i)}{n \cdot \log_2 \left(|dom(G_k)| \right)} \right) \quad (13)$$

donde *Ganancia* es la ganancia de información, n es el número de variables que aparecen en el antecedente de la regla y |*dom*(*Gk*)| es la cardinalidad de la variable objetivo (el número de valores posibles para la variable considerada como clase).

El objetivo global de la función de evaluación es orientar la búsqueda hacia reglas que maximicen la precisión y la medida de interés, minimizando el número de ejemplos negativos y no cubiertos.

El AG utiliza un modelo de reproducción de estado estacionario modificado que intenta obtener un equilibrio entre convergencia y diversidad. La recombinación se lleva a cabo mediante un operador de cruce en dos puntos y un operador de mutación aleatorio sesgado que potencia la diversidad de la población.

4 ESTUDIO EXPERIMENTAL

En esta sección se describe la metodología utilizada para el desarrollo de la experimentación de este estudio. Se muestran y analizan los resultados obtenidos para cada uno de los aspectos estudiados.

4.1. Características de la experimentación

Para analizar el comportamiento de SDIGA, se va a realizar una experimentación con distintas bases de ejemplos sintéticas disponibles en el repositorio UCI [2]. Las características principales de estos conjuntos de datos se muestran en la Tabla 1, en la que para cada conjunto de datos, N_{Var} representa el número de variables, $N_{Var\text{-}D}$ el número de variables discretas, $N_{Var\text{-}C}$ el número de variables continuas, N_{Cl} el número de clases, y N_{Ej} el número de ejemplos del conjunto.

Respecto a los aspectos a analizar, en primer lugar trataremos las medidas de calidad utilizadas en la función de evaluación, un aspecto muy importante en los algoritmos de descubrimiento de subgrupos. Distintos autores han utilizado distintas medidas objetivas de calidad, pero no hay un consenso sobre cuál es la mejor medida a utilizar.
Un segundo aspecto a estudiar es la influencia del tipo de reglas difusas a utilizar para representar el conocimiento.

Por último, la calidad de los resultados obtenidos por el algoritmo depende de la idoneidad de la partición difusa utilizada, tanto en tipo de conjuntos difusos, como en número, como en definición de los mismos, y en el caso de una partición uniforme (por ausencia de conocimiento experto), de la granularidad considerada.

Tabla 1: Propiedades de los conjuntos de datos del repositorio UCI utilizados

NOMBRE	N_{Var}	$N_{Var\text{-}D}$	$N_{Var\text{-}C}$	N_{Cl}	N_{Ej}
Australian	14	8	6	2	690
Breast-w	9	9	0	2	699
Bridges	7	4	3	2	102
Diabetes	8	0	8	2	768
Echo	6	1	5	2	131
German	20	13	7	2	1000
Heart	13	6	7	2	270
Hepatitis	19	13	6	2	155
Hypothyroid	25	18	7	2	3163
Ionosphere	34	0	34	2	351
Iris	4	0	4	3	150
Tic-tac-toe	9	9	0	2	958
Vote	16	16	0	2	435
Balance	4	0	4	3	625
Car	6	6	0	4	1728
Glass	9	0	9	6	214
Wine	13	0	13	3	178

Por lo tanto, se va a analizar el comportamiento del algoritmo utilizando distintos tipos de reglas, distintas medidas de calidad y distintas granularidades en las variables continuas. De esta forma, se ha realizado la experimentación para:

- 3 versiones de SDIGA: un esquema nítido (SDIGA-N), que utiliza las medidas de confianza nítida (*Conf-N*, como se define en (8)), soporte nítido (*Sop-N*, como se define en (5)) e interés (*Int*, como se define en (13)); un esquema difuso (SDIGA-D), que utiliza las medidas de confianza difusa (*Conf-D* (10)), soporte difuso (*Sop-D* (12)) e interés (*Int* (13)); y un esquema híbrido (SDIGA-H), que utiliza las medidas de confianza difusa (*Conf-D* (10)), soporte nítido sobre los ejemplos de la clase (Sop_c-*N* (6) e interés (*Int* (13)).
- 2 tipos de modelos de representación del conocimiento; es decir, 2 tipos de reglas: DNF y no DNF (o canónicas).
- Distinto número de etiquetas para las variables continuas: 3, 5 y 7 etiquetas.

Los experimentos se han llevado a cabo mediante validación cruzada con 10 particiones. Para cada base de ejemplos, y para cada algoritmo, por ser no determinista, se han realizado 5 ejecuciones y se muestran las medias de los resultados de estas 5 ejecuciones. Para estas experimentaciones, se ha utilizado un tamaño de la población de 100, 10.000 evaluaciones en cada ejecución del AG: 10.000, una probabilidad de mutación de 0,01, y como pesos para la función de adaptación: 0,4 para el soporte, 0,3 para la confianza y 0,1 para el interés.

Para analizar los resultados obtenidos, la Tabla 2 detalla los resultados promedios de los 17 conjuntos de datos estudiados. Esta tabla refleja las medias para las 5 ejecuciones de cada base de ejemplos, y muestra los resultados de las distintas medidas de calidad incluidas en la función de adaptación y del resto de medidas consideradas en la bibliografía especializada, cuyo contenido es el siguiente:

- El tipo de regla extraída (canónica o DNF).
- La versión del algoritmo SDIGA utilizado para la generación de reglas.
- El número de etiquetas para las variables continuas (N_{Et}), el número medio de reglas obtenidas (N_{Reg}) y el número medio de variables por regla (N_{Var}).
- Los valores medios de las medidas de calidad para los conjuntos de reglas: cobertura (*COB*), relevancia (*REL*), atipicidad (*ATIP*), precisión (*PREC*), soporte nítido sobre ejemplos de la clase (SOP_c-*N*), soporte nítido (*SOP-N*), soporte difuso (*SOP-D*), confianza difusa (*CNF-D*) y confianza nítida (*CNF-N*).

4.2. Análisis de resultados

La Tabla 2 muestra que para el conjunto de bases de ejemplos estudiado, los mejores resultados se obtienen en general (tanto para reglas canónicas como reglas DNF) con las medidas de calidad utilizadas en el esquema híbrido (SDIGA-H): confianza difusa (Conf-D), soporte nítido sobre los ejemplos de la clase (Sop_c-*N*) e interés (*Int*). De hecho, con el esquema evolutivo que utiliza esas tres medidas de calidad, no sólo se obtienen conjuntos de reglas con mejores valores en ellas, sino también para el resto de medidas.

Se observa que la versión SDIGA-N (con medidas nítidas) obtiene peores resultados y además genera siempre un mayor número de reglas. La utilización de las medidas de calidad de soporte y confianza difusas (SDIGA-D) aporta mejores resultados sólo en algunas bases de ejemplos con todas sus variables continuas y con variables objetivo con más de dos clases. No obstante, nuestra propuesta tiene como objetivo obtener reglas difusas, nítidas o mixtas en función del tipo de variables de cada problema, por lo que la combinación híbrida (soporte nítido sobre ejemplos de la clase y confianza difusa) es la más adecuada, obteniendo buenos resultados.

El análisis realizado sobre distintas medidas de calidad determina, bajo nuestro punto de vista, que en un proceso de descubrimiento de subgrupos es más adecuado el uso de medidas de:

- Exactitud en la descripción, como la confianza (en sus distintas definiciones) y la relevancia.

Tabla 2: Resultados por medidas de calidad

Regla	Versión	N_{Et}	N_{Reg}	N_{Var}	COB	REL	ATIP	PREC	SOP_c-N	SOP-N	SOP-D	CNF-D	CNF-N
Canónica	SDIGA-N	3	9,25	3,35	0,122	3,082	0,028	0,532	0,292	0,623	0,092	0,573	0,525
		5	10,48	3,28	0,120	2,614	0,022	0,516	0,241	0,610	0,089	0,529	0,473
		7	8,69	3,03	0,115	2,542	0,021	0,505	0,223	0,613	0,084	0,518	0,470
	SDIGA-D	3	5,66	3,12	**0,278**	4,581	**0,050**	0,550	0,455	0,720	**0,199**	0,527	0,489
		5	6,79	3,01	0,240	4,138	0,045	0,490	0,372	0,760	0,177	0,451	0,424
		7	6,99	**2,95**	0,243	4,050	0,041	0,477	0,371	**0,762**	0,172	0,436	0,413
	SDIGA-H	3	**5,29**	3,54	0,231	**4,958**	0,049	**0,601**	**0,530**	0,580	0,175	**0,627**	**0,583**
		5	6,57	3,23	0,237	4,803	0,045	0,563	0,438	0,647	0,182	0,547	0,512
		7	6,71	3,13	0,245	4,303	0,041	0,537	0,414	0,680	0,187	0,519	0,484
DNF	SDIGA-N	3	13,94	3,51	0,191	3,337	0,027	0,488	0,362	0,779	0,116	0,527	0,508
		5	14,36	4,14	0,191	3,610	0,029	0,549	0,367	0,735	0,121	0,594	0,558
		7	17,38	4,31	0,191	3,633	0,028	0,557	0,365	0,751	0,113	0,605	0,554
	SDIGA-D	3	5,83	**3,08**	0,231	5,993	0,046	0,416	0,407	0,837	0,138	0,399	0,380
		5	6,05	3,34	0,228	6,219	0,051	0,445	0,410	**0,841**	0,140	0,431	0,408
		7	6,04	3,51	0,243	6,377	0,048	0,461	0,437	0,840	0,144	0,448	0,421
	SDIGA-H	3	5,67	4,52	**0,253**	**7,275**	**0,055**	**0,611**	**0,667**	0,584	**0,188**	**0,622**	**0,601**
		5	5,71	5,09	0,236	7,242	0,053	0,601	0,616	0,560	0,170	0,613	0,591
		7	**5,62**	5,19	0,211	6,802	0,042	0,572	0,573	0,557	0,150	0,600	0,559

Tabla 3: Resultados por tipo de regla y granularidad para SDIGA-H

Tipo regla	N_{Et}	N_{Reg}	N_{Var}	COB	REL	ATIP	PREC	SOP_c-N	SOP-N	SOP-D	CNF-D	CNF-N
Canónica	3 etiq	**5,29**	3,54	0,231	4,958	0,049	0,601	0,530	0,580	0,175	**0,627**	0,583
	5 etiq	6,57	3,23	0,237	4,803	0,045	0,563	0,438	0,647	0,182	0,547	0,512
	7 etiq	6,71	**3,13**	0,245	4,303	0,041	0,537	0,414	**0,680**	0,187	0,519	0,484
DNF	3 etiq	5,67	4,52	**0,253**	**7,275**	**0,055**	**0,611**	**0,667**	0,584	**0,188**	0,622	**0,601**
	5 etiq	5,71	5,09	0,236	7,242	0,053	0,601	0,616	0,560	0,170	0,613	0,591
	7 etiq	5,62	5,19	0,211	6,802	0,042	0,572	0,573	0,557	0,150	0,600	0,559

- Generalidad de la descripción, pero considerando ejemplos bien descritos (ejemplos positivos), como el soporte.
- Novedad en la descripción, como el interés o la atipicidad.

Una vez determinada la mejor combinación de medidas, para el modelo SDIGA, la Tabla 3 muestra los resultados obtenidos solamente por las experimentaciones realizadas sobre el esquema híbrido SDIGA-H, para analizar el comportamiento del algoritmo en función del tipo de regla utilizada para expresar el conocimiento extraído. En esta tabla se puede observar que los resultados medios muestran una ligera superioridad de las reglas DNF respecto a las canónicas.

Por último, respecto al análisis de la granularidad de las variables continuas, se puede observar que se obtienen mejores resultados con 3 y 5 etiquetas. Cuando se consideran 7 etiquetas, este parece un nivel elevado de granularidad. La elección entre 3 y 5 puede depender de las características concretas del problema a resolver.

5 CONCLUSIONES

En este trabajo, se presenta un estudio sobre la influencia de distintos aspectos en los resultados obtenidos por el algoritmo SDIGA de inducción reglas difusas de descripción de subgrupos. Se trata de un estudio preliminar en el que se han utilizado distintas bases de ejemplos sintéticas habitualmente utilizadas en el análisis de los algoritmos de descubrimiento de subgrupos.

Este estudio ha determinado que la versión de SDIGA que utiliza como medidas de calidad el soporte nítido sobre ejemplos de la clase, la confianza difusa y el interés es la más adecuada. Por otro lado, las reglas DNF obtienen resultados ligeramente mejores que las reglas canónicas y ofrecen una estructura más flexible que permite expresar el conocimiento extraído de forma más descriptiva.

Como trabajos futuros, pretendemos ampliar este estudio con la inclusión en la función de evaluación de SDIGA de otras medidas de calidad como la relevancia o la atipicidad. Por otro lado, se pretende desarrollar una versión multiobjetivo del algoritmo y ampliar este estudio con sus resultados.

Agradecimientos

Este trabajo ha sido financiado por el MCYT a través de los proyectos TIN2005-08386-C05-01 y TIN2005-08386-C05-03.

Referencias

[1] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H. y Verkamo, I. Fast discovery of association rules. *Advances in Knowledge Discovery and Data Mining*, Fayyad, U., et al., Editors. AAAI Press: Menlo Park. Pág. 307–328, 1996.

[2] Asuncion, A. y Newman, D.J. UCI machine learning repository (http://www.ics.uci.edu/~mlearn/MLRepository.html). Univ. California, Irvine, School of Information and Computer Sciences, 2007.

[3] Atzmueller, M. y Puppe, F. SD-Map - A fast algorithm for exhaustive subgroup discovery. *10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2006)*. Springer, 2006.

[4] Au, W.H. y Chan, K.C.C. An effective algorithm for discovering fuzzy rules in relational databases. *IEEE International Conference on Fuzzy Systems (Fuzz IEEE'98)*. Pág. 1314-1319, 1998.

[5] del Jesus, M.J., González, P., Herrera, F. y Mesonero, M. Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A Case Study in Marketing. *IEEE Transactions on Fuzzy Systems*, 15(4). Pág. 578-592, 2007.

[6] Kavsek, B. y Lavrac, N., APRIORI-SD: Adapting association rule learning to subgroup discovery. *Applied Artificial Intelligence*, 20. Pág. 543–583, 2006.

[7] Klösgen, W. Explora: A multipattern and multistrategy discovery assistant. *Advances in Knowledge Discovery and Data Mining*. Pág. 249-271, 1996.

[8] Lavrac, N., Cestnik, B., Gamberger, D. y Flach, P. Decision support through subgroup discovery: Three case studies and the lessons learned. *Machine Learning*, 57(1-2). Pág. 115-143, 2004.

[9] Lavrac, N., Flach, P. y Zupan, B. Rule evaluation measures: A unifying view. *Inductive Logic Programming*. Pág. 174-185, 1999.

[10] Lavrac, N., Kavsek, B., Flach, P. y Todorovski, L. Subgroup discovery with CN2-SD. *Journal of Machine Learning Research*, 5. Pág 153-188, 2004.

[11] Noda, E., Freitas, A.A. y, Lopes, H.S., Discovering Interesting Prediction Rules with a Genetic Algorithm. *Congress on Evolutionary Computation*. Pág. 1322-1329, 1999.

[12] Wrobel, S. An algorithm for multi-relational discovery of subgroups. In Proc. of the First European Conference on Principles of Data Mining and Knowledge Discovery, pp. 78-87, 1997.

[13] Zadeh, L.A. The concept of a linguistic variable and its applications to approximate reasoning, Parts I, II, III. *Information Sciences*, 8-9. Pág. 199-249, 301-357, 43-80, 1975.

Algoritmo de hormigas mejorado para la descripción de modelos difusos interpretables mediante reglas con excepciones

P. Carmona[1] **J.L. Castro**[1]

[1] Depto. de Ingeniería de Sistemas Informáticos y Telemáticos, UEx, pablo@unex.es
[2] Depto. de Ciencias de la Computación e Inteligencia Artificial, UGr, castro@decsai.ugr.es

Resumen

En un trabajo previo, los autores proponen la ampliación de la sintaxis de las reglas difusas con nuevos predicados y con reglas excepcionales, así como el uso de un algoritmo de hormigas para la descripción optimal de un modelo difuso inicial. Se presentan aquí diversas extensiones a dicho algoritmo para mejorar la interpretabilidad de los modelos difusos obtenidos y se analiza experimentalmente la ganancia obtenida.

Palabras Clave: Modelado difuso, algoritmos de hormigas, interpretabilidad.

1 Introducción

Una de las formas de mejorar la interpretabilidad de un modelo difuso consiste en tratar de identificar reglas tan generales como sea posible [4], de forma que cada regla cubra el mayor número de ejemplos de entrenamiento y, de este modo, el tamaño de la base de reglas disminuya. Además, la extensión de la sintaxis de las reglas con nuevos predicados diferentes del habitual *es-igual-a* permite generar reglas más compactas. Adicionalmente, los autores han propuesto el uso de reglas excepcionales para aumentar aún más la interpretabilidad de modelos descritos mediante reglas difusas [2, 3].

En un trabajo previo [1], se propone un método para la búsqueda de un conjunto de reglas generales y reglas excepcionales que describan un modelo difuso inicial. Para ello, se empleó un algoritmo de hormigas [7], método de optimización global que se basa en el comportamiento emergente que surge de la búsqueda cooperativa realizada por un conjunto de agentes denominados hormigas artificiales.

Se proponen aquí diversas extensiones a dicho trabajo para mejorar la interpretabilidad del modelo obtenido. La primera sustituye el modelo de hormigas AS [5] empleado en la propuesta original por el modelo ACS [6], que mejora al anterior. La segunda extensión se basa en relajar las restricciones de un paso válido, aumentando el espacio de soluciones válidas y, por tanto, la flexibilidad de las soluciones obtenidas. La tercera, añade una búsqueda local para refinar cada solución alcanzada. Finalmente, una cuarta extensión elimina reglas que subsumen, no ya en otras reglas, sino en la base de reglas en su conjunto.

2 Reglas simples, compuestas y excepcionales

Se consideran sistemas con mltiples entradas y una salida (Multiple-Input–Single-Output, MISO) con n variables de entrada $\mathbf{X} = \{X_1, \ldots, X_n\}$ definidas sobre el universo del discurso $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$, y una variable de salida Y definida sobre el universo del discurso $\mathcal{Y}$. El dominio difuso de X_i se denota $\widetilde{\mathcal{X}}_i = \{LX_{i,1}, \ldots, LX_{i,p_i}\}$, donde $LX_{i,j}$ representa la etiqueta lingística del j-ésimo valor difuso. Análogamente, $\widetilde{\mathcal{Y}} = \{LY_1, \ldots, LY_q\}$ es el dominio difuso de salida.

Habitualmente, las reglas de un modelo difuso MISO contienen en el antecedente un predicado para cada variable de entrada que la asocia a una etiqueta de su dominio difuso. A estas reglas, que denominaremos *reglas simples*, las denotamos como:

$$R^{i_1 \ldots i_n}_{LY^i} : LX_{1,i_1}, \ldots, LX_{n,i_n} \rightarrow LY^i. \qquad (1)$$

Para mejorar la compactación de las reglas, se puede asociar más de una etiqueta a cada variable de entrada en el antecedente, así como utilizar otros predicados además del habitual *es-igual-a*. A estas reglas, que denominaremos *reglas compuestas*, las denotamos como:

$$R^i \ : \ [\neg]X_1 \ op_i \ SX^i_1, \ \ldots, \ [\neg]X_n \ op_n \ SX^i_n \rightarrow LY^i$$

donde $op_i \in \{=, \leq, \geq, \div, \in\}$ ($\div$ significa *entre*), $\neg$ representa al operador de negación y donde cada SX_j^i es una etiqueta si $op_i \in \{=, \leq, \geq\}$, dos si op_i es $\div$, y dos o más etiquetas si op_i es $\in$.

Es más, la compactación del modelo difuso puede aún mejorarse mediante el concepto de *reglas excepcionales*, introducido por los autores en [2, 3]. Una regla excepcional está asociada a otra regla — la *regla exceptuada*— y excluye el disparo de ésta de una porción del espacio de entrada que cubre, siendo la regla excepcional la que se aplica en su lugar. Además, para expresar de forma aún más compacta la regla excepcional, ésta aprovecha el contexto introducido por el antecedente de la regla exceptuada.

Así, dado el dominio difuso $\{NL, NS, Z, PS, PL\}$ para todas las variables, las reglas exceptuada y excepcional

$$R^1 : X_1 \in \{NL, Z, PS\}, X_2 \div [NS, PS] \rightarrow NL$$
$$\text{exc: } X_1 \geq Z, X_3 = Z \rightarrow NS$$

agrupan a $3 \times 3 \times 5 = 45$ reglas simples y equivalen a las reglas compuestas

$$R^1 : X_1 = NL, X_2 \div [NS, PS] \rightarrow NL,$$
$$R^2 : X_1 \in \{Z, PS\}, X_2 \div [NS, PS], \neg X_3 = Z \rightarrow NL,$$
$$R^3 : X_1 \in \{Z, PS\}, X_2 \div [NS, PS], X_3 = Z \rightarrow NS.$$

3 El algoritmo de hormigas original

A continuación se presenta el algoritmo propuesto en [1], que busca la mejor transformación de un conjunto de reglas iniciales (CRI) en un conjunto de reglas compuestas (CRC) con posibles reglas excepcionales.[1]

Una hormiga puede realizar pasos de *inclusión*, que añaden reglas del CRI al CRC, o de *amplificación*, que extienden la cobertura de una regla del CRC añadiendo una etiqueta a una de sus premisas en el antecedente. La amplificación que la etiqueta $LX_{j,k}$ a R^i se denota $\langle i, j, k\rangle$, mientras que la inclusión de la i-ésima regla inicial (RI) al CRC se denota $\langle i, 0, 0\rangle$ y da lugar a la regla compuesta R^i.

3.1 Construcción de soluciones

Para construir una solución, la hormiga se sitúa en una RI aleatoria, incluyéndola así en su CRC, y sucesivamente selecciona sus pasos entre los *válidos* para cada estado hasta cubrir positivamente todas las RIs. Un paso es válido:

1. Si es de inclusión, cuando la RI aún no esté cubierta positivamente por el CRC.
2. Si es de amplificación de una regla compuesta R^i:
   (a) si R^i no es excepcional, cuando:
      i. la zona de amplificación (ZdA) —la nueva zona cubierta por $R^{i'}$ (R^i amplificada)— no cubre solo RIs negativamente, y
      ii. cada R^j que solapa con $R^{i'}$ no es excepcional de otra regla, está completamente cubierta por $R^{i'}$ y, o bien los consecuentes de sendas reglas coinciden, o difieren y R^j no cubre RIs negativamente.
   (b) Si R^i es excepcional de R^j, cuando la ZdA está contenida en R^j, no cubre RIs negativamente ni solapa con otra regla excepcional.

Tras cada amplificación, las reglas que subsumen en R^i se eliminan y la conflictivas que solapan con ella se convierten en excepcionales.

3.2 Información heurística

Para guiar la búsqueda hacia caminos con pasos prometedores se emplea la función heurística

$$\eta = \gamma \frac{Reg_{cob}}{Reg_{max}} + (1-\gamma)\frac{Reg_+/(Reg_- + 1)}{Reg_{max}}, \qquad (2)$$

donde Reg_{cob} es el número de regiones de entrada cubiertas por la ZdA, Reg_+ (Reg_-) es el número de RIs cubiertas positivamente (negativamente) por la ZdA y Reg_{max} es el número máximo de regiones de entrada cubiertas por la ZdA cuando se amplifica la variable considerada en el paso (para $s : \langle i, j, k\rangle$, $Reg_{max} = \prod_{l \neq j} p_l$)[2].

La ecuación (2) se evalúa sobre el estado en el que se encuentra la hormiga justo antes de realizar el paso.

3.3 Actualización de feromonas

Esta sección se refiere a las feromonas depositadas por una hormiga tras finalizar su recorrido. Dada una solución (CRC final), se evalúa cada regla R^i incluida en ella y se deposita en cada paso $\langle i, \cdot, \cdot\rangle$ del recorrido de la hormiga una cantidad de feromonas proporcional a dicho valor (en este sentido, las reglas exceptuadas y las excepcionales se tratan de igual modo).

La valoración de una regla evalúa su simplicidad sintáctica. Para ello, se define la complejidad de cada

[1] Para poder comparar los resultados experimentales, se han incluido algunos cambios menores como parte del algoritmo original (concretamente, el uso de la negación, la expresión de las reglas excepcionales aprovechando el contexto introducido por la regla exceptuada y la inicialización de feromonas utilizada en la Sección 3.3).

[2] Para $s : \langle i, 0, 0\rangle$, $Reg_{max} = \prod_l p_l / \min_l p_l$ y (2) se divide por 2 para favorecer las amplificaciones.

premisa como

$$c(P_j) = \begin{cases} 0 & \text{si es la premisa } \textit{vacía} \\ 3 & \text{si } X_j(=,\leq,\geq)LX_{j,k} \\ 4 & \text{si } X_j \div [LX_{j,k}, LX_{j,k+l}] \\ l+2 & \text{si } X_j \in \{LX_{j,k_1}, \ldots, LX_{j,k_l}\}, \end{cases} \quad (3)$$

se evalua cada premisa de la regla P_j y su negación $\neg P_j$ (donde $c(\neg P_j) = c(P_j) + 1$) con objeto de considerar la negación y se utiliza en el resto de ecuaciones que se describen a continuación la que resulte más simple (es decir, la de menor valor). A continuación, se define la complejidad de una regla como

$$C(R^i) = \min\left(1, \frac{E + \sum_j c([\neg]P_j)}{\sum_j p_j + 1}\right), \quad (4)$$

donde E es el número de reglas excepcionales de R^i, y su interpretabilidad se define como

$$I(R^i) = \delta(1 - C(R^i)) + (1-\delta)\frac{Reg_{cob}}{Reg_{max}}, \quad (5)$$

donde $\delta \in [0,1]$ es un parámetro, Reg_{cob} es el número de regiones de entrada cubiertas por R^i excepto las cubiertas por sus reglas excepcionales y $Reg_{max} = \prod_l p_l$.

Además, dado que lo adecuado de un paso también depende de los pasos previos, la cantidad de feromonas asociadas a un paso $s : \langle i, \cdot, \cdot \rangle$ dado por la k-ésima hormiga se calcula como

$$\Delta\tau_s^k = I(R^i) \times \frac{L - l + 1}{L}, \quad (6)$$

donde L es la longitud del recorrido y l la posición de $\langle i, \cdot, \cdot \rangle$ en el mismo.

Finalmente, la medida utilizada para seleccionar la mejor solución se define como

$$I(BR) = \frac{NRI - \sum_i C(R^i)}{NRI}, \quad (7)$$

donde NRI es el número de RIs y representa la máxima complejidad del modelo (número máximo de reglas, cada una con complejidad máxima).

La cantidad de feromonas inicial es igual a

$$\tau_0 = IPR_{voraz}/NR_{voraz} \quad (8)$$

donde IPR_{voraz} es la interpretabilidad media de las reglas difusas obtenidas con un algoritmo voraz igual al algoritmo de hormigas (sin las extensiones de 4.3 y 4.4) eligiendo siempre el paso de mayor valor heurístico, y NR_{voraz} es el número de reglas de dicho modelo difuso.

4 Extensiones al algoritmo original

4.1 Modelo ACS

La primera extensión consiste en sustituir el modelo AS [5] utilizado en el algoritmo original, por el modelo mejorado ACS propuesto por Dorigo y Gambardella en [6]. Las novedades de este modelo son:

- Se sustituye la regla proporcional aleatoria de selección de un paso por una regla pseudoaleatoria donde un nuevo parámetro q_0 establece la probabilidad de seleccionar el mejor paso en lugar de utilizar la regla proporcional aleatoria original.
- Tras cada ciclo, se actualizan sólo las feromonas de los pasos de la hormiga m que obtuvo la mejor solución hasta el momento, $\tau_s^m = (1-\alpha)\cdot\tau_s^m + \alpha\cdot\Delta\tau_s^m$, en lugar de los pasos de todas las hormigas del ciclo.
- Durante el recorrido de cada hormiga k, disminuyen las feromonas de cada paso dado s de la forma $\tau_s^k = (1-\rho)\cdot\tau_s^k + \rho\cdot\tau_0$, reduciendo la probabilidad de que otras hormigas elijan ese paso.

4.2 Solapamiento entre reglas

En la Sección 3.1, sólo se permite que una regla en amplificación solape con otra si, tras la amplificación, la última queda cubierta completamente por la primera (condición 2.(a)ii). Un modo de incrementar la capacidad de encontrar soluciones más interpretables consiste en permitir que reglas con el mismo consecuente puedan solapar parcialmente. Aunque esto puede provocar coberturas redundantes del espacio de entrada, el resultado será positivo si estas redundancias suponen una simplificación de las reglas implicadas. Dado que el algoritmo busca esa simplicidad, es de esperar que la metaheurística tenderá a evitar por sí misma las variantes indeseadas.

Para implementar esta extensión, la definición de paso válido se redefine como sigue (las modificaciones aparecen *resaltadas*):

1. ...
2. Si es de amplificación de una regla compuesta R^i:
   (a) si R^i no es excepcional, cuando:
      i. ...
      ii. cada R^j que solapa con $R^{i'}$:
         - es consistente con $R^{i'}$ y, o bien está completamente cubierta por $R^{i'}$ *o bien la zona de solapamiento no cubre RIs negativamente*, o

- es conflictiva con $R^{i'}$, está completamente cubierta por ella, no es excepcional de otra regla y no cubre RIs negativamente.

(b) Si R^i es excepcional de R^j, cuando la ZdA está contenida en R^j, no cubre RIs negativamente ni *solapa con ninguna regla en conflicto con* R^i *excepto* R^j.

Aunque esta extensión flexibiliza la definición original, se impide el solapamiento entre reglas consistentes en regiones donde cubran RIs conflictivas, con objeto de evitar reglas excepcionales de más de una regla.

4.3 Subsunción respecto a la base de reglas

Si se incluye la extensión de la Sección 4.2, pueden surgir reglas que, aunque no subsuman en una regla individual, sí lo hagan en un conjunto de dos o más reglas del modelo final. Por ello, otra extensión consiste en lanzar un proceso para eliminar tales reglas.

Durante este proceso, las reglas exceptuadas se analizan antes que sus reglas excepcionales, de modo que si se elimina alguna regla exceptuada, sus reglas excepcionales se conviertan en no excepcionales. Una regla exceptuada subsumirá en la base de reglas si las regiones cubiertas por su antecedente, exceptuando las cubiertas por sus reglas excepcionales, están también cubiertas por otras reglas del CRC.

4.4 Búsqueda local

En (2), se favorece la amplitud de la cobertura de una regla incluso aunque la amplificación no proporcione un beneficio inmediato, con la esperanza de que futuras amplificaciones sobre la misma premisa sí lo consigan. Esto puede dar lugar a reglas finales con etiquetas innecesarias que, no sólo no proporcionen ningún beneficio, sino que incluso incrementen la complejidad injustificadamente. Por ello, se introduce aquí una búsqueda local para explorar cada regla final y analizar la idoneidad de cada etiqueta. Concretamente, la búsqueda local consiste en:

Para cada regla R^i del modelo final:
  Para cada paso $s : \langle i, j, k\rangle$ del recorrido:
    Sea $R^{i'}$ igual a R^i sin $LX_{j,k}$
    Si $\langle conds\rangle$ entonces
      Reemplazar R^i con $R^{i'}$
      Eliminar el paso s del recorrido

donde $\langle conds\rangle$ son las condiciones para considerar $LX_{j,k}$ innecesario y comprenden:

1. $I(R^{i'}) \geq I(R^i)$
2. el CRI sigue cubierto tras sustituir R^i por $R^{i'}$
3. si R^i es una regla exceptuada, ninguna de sus excepciones cubre la ZdA delimitada por s
4. si se permite el solapamiento (Sección 4.2) y R^i es una regla excepcional, la ZdA sólo está cubierta por reglas excepcionales de R^i.

Esta búsqueda local se lanza después de construir cada solución y antes actualizar las feromonas. Además, si se incluye la extensión de la Sección 4.3, se eliminará la regla resultante si subsume en la base de reglas.

5 Resultados experimentales

Se realizaron experimentos para determinar la efectividad de las mejoras presentadas en la Sección 4. Para ello, se obtuvieron los resultados con el método original, con la extensión presentada en la Sección 4.1 (para determinar el beneficio derivado del uso del modelo ACS) y con todas las extensiones presentadas en la Sección 4 aplicadas conjuntamente. Cada configuración del algoritmo se ejecutó 5 veces y se analizaron los valores medios.

Los parámetros empleados fueron $\gamma = 0.25$, $\delta = 0.5$ y un número de hormigas $M = 5$. Para la regla proporcional aleatoria se empleó el parámetro $\beta = 1$. En el modelo AS, se empleó el parámetro de evaporación $\rho = 0.1$ y para el modelo ACS los parámetros $\alpha = \rho = 0.1$ y $q_0 = 0.9$. El número máximo de ciclos se estableció en 80 y se consideró como condición de convergencia para finalizar la ejecución del algoritmo que todas las hormigas en un ciclo construyeran el mismo CRC.

Se diseñó un generador automático para proporcionar bases de reglas compuestas con excepciones con diferentes dimensionalidades (número de variables de entrada, n) y cardinalidades para los dominios difusos (la misma cardinalidad p para todos los dominios). Este generador selecciona aleatoriamente un conjunto de etiquetas para cada premisa en el antecedente de las reglas con la restricción de que el modelo final no contenga reglas conflictivas ni subsumidas, tras lo cual se exceptúan algunas de las reglas generando aleatoriamente reglas excepcionales compuestas.

Se emplearon las configuraciones $[2, 9]$ y $[3, 7]$ para $[n, p]$ y para cada una se generaron tres bases de reglas (BR^1, BR^2 y BR^3).

La Tabla 1 muestra resumidamente los resultados obtenidos. Para cada configuración $[n, p]$ se muestra, en la primera fila, los datos relativos a cada modelo difuso preliminar sobre el que se aplica cada método

Tabla 1: Resultados obtenidos con las distintas versiones del método propuesto

	BR_1	BR_2	BR_3
$[2,9]$	(4.40)⟨10.0⟩{49}	(3.75) ⟨9.0⟩ {47}	(3.85) ⟨9.0⟩ {40}
AS	(4.39)⟨10.8⟩	(3.61) ⟨9.4⟩	(3.08) ⟨8.2⟩
ACS	(3.92)⟨10.6⟩	(3.38) ⟨9.0⟩	(2.85) ⟨8.2⟩
$ACS + E_*$	(3.38)⟨11.0⟩	(3.29) ⟨9.0⟩	(2.70) ⟨8.2⟩
$[3,7]$	(4.50)⟨10.0⟩{88}	(5.88)⟨13.0⟩{145}	(5.54)⟨12.0⟩{98}
AS	(5.25)⟨12.8⟩	(6.32)⟨17.0⟩	(4.47)⟨11.6⟩
ACS	(4.88)⟨12.0⟩	(6.18)⟨13.6⟩	(4.07)⟨11.0⟩
$ACS + E_*$	(4.27)⟨12.0⟩	(5.76)⟨14.0⟩	(3.75)⟨11.0⟩

y, en el resto de filas, los resultados promediados relativos a las distintas configuraciones de dicho método: AS, el algoritmo de hormigas original presentado en la Sección 3 y que utiliza el modelo AS; ACS, que sustituye el modelo AS por el modelo ACS introducido en la Sección 4.1; y $ACS + E_*$ que añade al método anterior el resto de extensiones propuestas en la Sección 4, permitiendo el solapamiento entre reglas compuestas, aplicando la búsqueda local y considerando la subsunción de reglas respecto a la base de reglas junto con la nueva definición de paso válido.

Esta tabla muestra, para cada modelo difuso al que se aplica el algoritmo de hormigas, la siguiente información:

- entre paréntesis, la complejidad acumulada tanto del modelo difuso descrito mediante reglas compuestas. Esta medida equivale a la suma de la complejidad de sus reglas y, por tanto, es inversa a la medida de la interpretabilidad de una base de reglas definida en (7),
- entre paréntesis angulares, el número de reglas compuestas que describen al modelo difuso, y
- entre llaves, el número de RIs simples del modelo difuso proporcionado como entrada del algoritmo de hormigas.

Cabe resaltar que el algoritmo de hormigas no parte del conjunto de reglas compuestas generadas automáticamente, sino del conjunto de reglas simples que contiene. A este respecto, aunque no puede asegurarse que el generador de bases de reglas compuestas proporcione el conjunto óptimo de reglas que cubre a las RIs subyacentes, al menos genera una descripción razonablemente buena si atendemos a la relación número de reglas simples vs. número de reglas compuestas (incluyendo las reglas excepcionales). Por ejemplo, para $[n,p]$ igual a $[3,7]$ las relaciones son 88/10, 145/13 y 98/12 (veáse la tabla). Por ello, se marcó como objetivo conseguir con el algoritmo de hormigas soluciones tan interpretables como las generadas automáticamente.

En la Tabla 1 pueden observarse las mejoras en interpretabilidad obtenidas con las diversas extensiones al algoritmo original, consiguiendo reducir sucesivamente la complejidad de los modelos difusos resultantes, aunque a veces esto signifique un aumento en el número de reglas compuestas que los describen (más reglas, pero más simples). La versión $ACS + E_*$ consigue, no sólo igualar el nivel de interpretabilidad de los modelos difusos preliminares utilizados como entrada al algoritmo, sino reducirlo en todos los casos. Como se apuntó anteriormente, este resultado es más significativo si se tiene en cuenta que el algoritmo parte, no del modelo difuso descrito mediante reglas compuestas, sino del conjunto de reglas simples que lo constituyen.

6 Conclusiones

En este trabajo se presentan diversas extensiones sobre el algoritmo de hormigas presentado por los autores en [1], centradas en incrementar la interpretabilidad de los resultados obtenidos.

Con el objetivo de evaluar la idoneidad de las extensiones propuestas, se desarrolló un generador de modelos difusos aleatorios descritos mediante reglas compuestas con posibles reglas excepcionales. Dado que dichos modelos, si bien no son óptimos en general, proporcionan una buena descripción en términos de número de reglas simples frente a número de reglas compuestas, consideramos como objetivo conseguir descripciones tan interpretables como las de dichos modelos difusos generados automáticamente.

Como conclusión, tanto el uso del modelo ACS como el resto de extensiones proporcionan beneficios respecto al algoritmo original, siendo éstos máximos cuando se aplican todas las extensiones conjuntamente. En este último caso, no sólo se consiguen soluciones tan interpretables como las de los modelos difusos generados automáticamente, sino que en todos casos se mejora

su interpretabilidad. Este resultado es tanto más significativo si se observa que la entrada al algoritmo de hormigas propuesto no es la base de reglas descrita mediante reglas compuestas sino las reglas simples que lo constituyen.

Referencias

[1] P. Carmona and J.L. Castro. *Analysis and Design of Intelligent Systems Using Soft Computing Techniques*, volume 41 of *Advances in Soft Computing*, chapter An Ant Colony Optimization plug-in to enhance the interpretability of fuzzy rule bases with exceptions, pages 436–444. Springer-Verlag, Heidelberg, Germany, 2007.

[2] P. Carmona, J.L. Castro, and J.M. Zurita. FRIwE: Fuzzy rule identification with exceptions. *IEEE Trans. Fuzzy Syst.*, 12(1):140–151, 2004.

[3] P. Carmona, J.L. Castro, and J.M. Zurita. Learning maximal structure fuzzy rules with exceptions. *Fuzzy Sets Syst.*, 146(1):63–77, 2004.

[4] J.L. Castro, J.J. Castro-Schez, and J.M. Zurita. Learning maximal structure rules in fuzzy logic for knowledge acquisition in expert systems. *Fuzzy Sets Syst.*, 101:331–342, 1999.

[5] M. Dorigo, A. Colorni, and V. Maniezzo. The ant system: Optimization by a colony of cooperating agents. *IEEE Trans. Syst. Man Cybern. B*, 26(1):29–41, 1996.

[6] M. Dorigo and L.M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evol. Comput.*, 1(1):53–66, 1997.

[7] M. Dorigo and T. Stützle. *Ant Colony Optimization*. MIT Press, 2004.

MODELO DE DECISIÓN DE COLAS BASADO EN LAS CADENAS DE MARKOV CON ESTADOS FUZZY

David de la Fuente García [1] **María José Pardo Santiago** [2]

[1] Universidad de Oviedo, david@uniovi.es
[2] Universidad del País Vasco, mjose.pardo@ehu.es

Resumen

En este trabajo diseñamos un sistema de colas determinando la mejor política a seguir en las decisiones sobre propaganda, tomando como base las cadenas de Markov con estados fuzzy. Para ello, en primer lugar definimos y calculamos las cadenas de Markov con estado inicial y estado final fuzzy y, en segundo lugar, ilustraremos con un ejemplo, y a través de la solución por programación lineal del proceso de decisión markoviano, los resultados teóricos previamente obtenidos.

Palabras Clave: Modelos de Decisión de Colas, Cadenas de Markov con Estados Fuzzy.

1 INTRODUCCIÓN

Un proceso de decisión Markoviano puede ser usado para describir un sistema dinámico el cual evoluciona en el tiempo conforme a los efectos simultáneos de las leyes probabilísticas de movimiento y la secuencia de decisiones tomadas. El sistema se observa en los instantes de tiempo $t = 0,1,2,...$, y es clasificado en uno de los M estados del sistema. Después de cada observación se toma una de las K posibles decisiones. Ello implica que cuando el sistema se encuentra en un estado particular y se toma una decisión, el sistema se mueve a un nuevo estado con una probabilidad de transición conocida y se incurre en un coste, y esto para toda combinación de estados y decisiones. Una política es un conjunto de decisiones tomadas para cada uno de los estados del sistema. En general hay K^M políticas. El problema es encontrar la mejor política de manera que el coste promedio por unidad de tiempo sea mínimo.

Los procesos de decisión Markovianos representan una útil herramienta para modelar y encontrar la política óptima en una amplia clase de sistemas. Son aplicados en áreas tales como teoría de colas, inventarios, mantenimiento, programación dinámica probabilística,...

En este trabajo diseñamos un sistema de colas con capacidad limitada $M/M/1/N$ (Winston [5]) a través de la determinación de la mejor política a seguir en las decisiones sobre propaganda, las cuales son tomadas teniendo en cuenta al número de clientes en cola en el instante de la toma de decisiones. Así, si por ejemplo la longitud de cola es larga la decisión podrá ser efectuar o no propaganda para aumentar la tasa de llegadas, tal decisión dependerá de los costes que conlleve la realización de publicidad.

Para llevar a cabo el proceso de optimización y encontrar la mejor política a seguir en cada estado de la Cadena de Markov que determina el sistema de colas vamos a utilizar la solución por programación lineal del proceso de decisión markoviano (Taha [4]). Esta solución conlleva la formulación de un programa lineal. Para un sistema con M estados y K alternativas de decisión, la formulación del programa lineal asociado debe incluir $M+1$ restricciones y MK variables, el cual tenderá a ser muy extenso para valores de M y K grandes. Por ejemplo, si el sistema cuenta con una capacidad limitada a 15 unidades, el número de estados es $M = N+1 = 16$, y suponiendo que existan dos alternativas de decisión, efectuar propaganda o no, entonces $K = 2$, luego el sistema de programación lineal que permite resolver para cada estado de la Cadena de Markov la decisión más adecuada consta de 32 variables y 17 restricciones. Para reducir este nivel de cálculo hemos desarrollado las Cadenas de Markov con estados fuzzy, de manera que en el problema en estudio se pueden considerar, por ejemplo, los estados fuzzy "longitud corta de cola", "longitud media de cola" y "longitud larga de cola" y para resolver la optimización del sistema de colas a cada estado fuzzy se le asocia una alternativa de decisión (efectuar publicidad o no) de forma que el sistema de programación lineal, dado que el modelo cuenta ahora con 3 estados ($M = 3$) y las alternativas de decisión son también 2 ($K = 2$) consta de 6 variables y 4 restricciones, una cantidad de variables y restricciones sensiblemente inferior a las obtenidas sin la aplicación de estados fuzzy sobre la Cadena de Markov.

Para utilizar las Cadenas de Markov con estados fuzzy en primer lugar en este trabajo hemos desarrollado la teoría que relaciona los estados no fuzzy con los estados fuzzy, y que permite calcular las probabilidades de transición de la Cadena de Markov con estados fuzzy conocidas las probabilidades de transición de la Cadena de Markov con estados no-fuzzy, así como calcular las probabilidades de estado estable de la Cadena de Markov con estados fuzzy mediante las probabilidades de estado estable con estados no-fuzzy.

A continuación, en el trabajo describimos brevemente la solución del problema de decisión markoviano por programación lineal. Y, por último, ilustramos con un ejemplo la aplicación práctica de los resultados teóricos obtenidos.

Para mantener la longitud del presente artículo dentro de los límites todas las demostraciones y pruebas han sido omitidas, sólo se establecen los resultados.

2 CADENAS DE MARKOV CON ESTADOS FUZZY

2.1. PRELIMINARES

2.1.1. Probabilidad de eventos fuzzy

Sea (Ω, Λ, P) el espacio probabilístico, donde Ω denota el espacio muestral, Λ una σ-álgebra sobre Ω y P una medida de probabilidad. Un conjunto fuzzy A en Ω define un *evento fuzzy*. Sea $\mu_A(\cdot)$ la función de pertenencia del evento fuzzy A. Entonces la probabilidad del evento fuzzy A y la probabilidad condicional del evento fuzzy A, dado un evento fuzzy B son definidas por Zadeh [7] como:

$$P(A) = \int_\Omega \mu_A(\omega) dP \qquad \mu_A(\omega) : \Omega \to [0,1] \quad (1)$$

$$P(A/B) = \frac{P(A \cdot B)}{P(B)} \qquad P(B) > 0 \quad (2)$$

El producto de dos eventos fuzzy A y B se define como (Zimmermann [8]):

$$A \cdot B \leftrightarrow \mu_{A \cdot B} = \mu_A \cdot \mu_B \quad (3)$$

2.1.2. Probabilidades de las Cadenas de Markov

Sea X_t el estado del sistema en el instante t. Se considera una cadena de Markov con estado finito en la que la matriz de probabilidades de transición es:

$$\boldsymbol{P} = \left[p_{ij} \right] \qquad \forall i, j \in \{ 0,1,...,N \} \quad (4)$$

donde p_{ij} representa las probabilidades de transición desde el estado i al estado j de un paso, $p_{ij} \geq 0, \ \forall i, j$:

$$p_{ij} = P\{X_{t+1} = j / X_t = i\} = P\{X_1 = j / X_0 = i\} \quad (5)$$

con $\sum_{j=0}^{N} p_{ij} = 1$. Y la matriz de transición de probabilidad de r pasos es, donde p_{ij}^r es la probabilidad de transición desde el estado i al estado j, en r pasos:

$$\boldsymbol{P}^{\boldsymbol{r}} = \left[p_{ij}^r \right] \qquad p_{ij}^r \geq 0, \forall i, j \in \{ 0,1,...,N \} \quad (6)$$

$$p_{ij}^r = P\{X_{t+r} = j / X_t = i\} = P\{X_r = j / X_0 = i\} \quad (7)$$

Los elementos de la matriz $\boldsymbol{P}^{\boldsymbol{r}}$ se obtienen de forma directa con:

$$\boldsymbol{P}^{\boldsymbol{r}} = (\boldsymbol{P})^r \quad (8)$$

Por último, sea $p_i = P\{X_0 = i\}$, $\forall i \in \{ 0,1,...,N \}$, es decir. p_i son las probabilidades de estado inicial de la cadena de Markov, con $\sum_{i=0}^{N} p_i = 1$.

Si la cadena de Markov es irreducible ergódica, entonces se verifica:

$$\lim_{r \to \infty} p_{ij}^r = \pi_j \qquad \forall j \in \{ 0,1,...,N \} \quad (9)$$

en donde π_j son las probabilidades de estado estable.

2.2. PROBABILIDADES DE LAS CADENAS DE MARKOV CON ESTADOS FUZZY

Sobre los estados del sistema definimos una partición fuzzy[1] (Dubois y Prade [1]), es decir, un conjunto de estados fuzzy $\{ \tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_n \}$, de manera que cada subconjunto fuzzy $\tilde{A}_i$, $i \in \{ 1,2,...,n \}$ representa un estado o evento fuzzy de la cadena de Markov inicial.

Definición 1: Se define la *probabilidad de estado inicial fuzzy*, $P(\tilde{A}_i) = P(X_0 = \tilde{A}_i)$ a través la probabilidad crisp de un evento fuzzy calculada mediante la integral de *Lebesgue-Stieltjes* (Zadeh [7]):

[1] Sea $X = \{x_1, x_2, ..., x_n\}$ un conjunto dado. Una partición fuzzy de X es una familia de subconjuntos de X, denotados por $A = \{ \tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_N \}$ $(\forall i = \{1,...,N\}, \tilde{A}_i \neq \varnothing, \tilde{A}_i \neq X)$ con las correspondientes funciones de pertenencia $\mu_{\tilde{A}_1}, ..., \mu_{\tilde{A}_N}$ las cuales satisfacen:

$$\sum_{i=1}^{N} \mu_{\tilde{A}_i}(x_r) = 1 \qquad \forall x_r \in X \qquad r = \{1,...,n\}$$

$$P\left(\tilde{A}_i\right) = P\left\{\tilde{X}_0 = \tilde{A}_i\right\} = \sum_{s=0}^{N} P\left\{X_0 = s\right\} \mu_{\tilde{A}_i}(s) = \sum_{s=0}^{N} p_s \mu_{\tilde{A}_i}(s)$$

Proposición 1: La probabilidad condicional del estado fuzzy $\tilde{A}_j$ dado el estado inicial *m,* con $j \in \{1,...,n\}$ y $m \in \{0,...,N\}$, es:

$$P\left(\tilde{A}_j / m\right) = P\left\{\tilde{X}_1 = \tilde{A}_j / X_0 = m\right\} = \sum_{s=0}^{N} p_{ms} \mu_{\tilde{A}_j}(s) \quad (10)$$

y representa la *probabilidad de transición a estado fuzzy* (de un paso). La demostración es directa desde (1) y (2).

Proposición 2: La *probabilidad de transición a estado fuzzy de r pasos* es:

$$P^r\left(\tilde{A}_j / m\right) = P\left\{\tilde{X}_r = \tilde{A}_j / X_0 = m\right\} = \sum_{s=0}^{N} p_{ms}^r \mu_{\tilde{A}_j}(s) \quad (11)$$

De igual forma, la demostración es directa desde (1) y (2).

Definición 2: Se define la *cadena de Markov de estado final fuzzy*, de *r* pasos, mediante la matriz:

$$\bar{\boldsymbol{P}}^{\boldsymbol{r}} = \left[P^r\left(\tilde{A}_j / m\right)\right] \quad (12)$$

Esta matriz proporciona la probabilidad de transición del estado inicial *m*, ($m \in \{0,...,N\}$) a un estado final fuzzy $\tilde{A}_j$, ($j \in \{1,...,n\}$), en *r* pasos.

Proposición 3: La *probabilidad condicional del evento fuzzy* $\tilde{A}_j$ *dado el evento fuzzy* $\tilde{A}_i$, $i, j \in \{1,...,n\}$ es función de una combinación lineal de probabilidades $P\left(\tilde{A}_j / m\right)$, de la forma:

$$P\left(\tilde{A}_j / \tilde{A}_i\right) = P\left\{\tilde{X}_1 = \tilde{A}_j / \tilde{X}_0 = \tilde{A}_i\right\} = \sum_{m=0}^{N} P\left(\tilde{A}_j / m\right) \frac{p_m \mu_{\tilde{A}_i}(m)}{P\left(\tilde{A}_i\right)}$$

y representa la *probabilidad de transición de estado inicial fuzzy a estado final fuzzy* (de un paso). La demostración es directa desde (1), (2) y (3).

Proposición 4: La *probabilidad de transición entre estados fuzzy de r pasos* es:

$$P^r\left(\tilde{A}_j / \tilde{A}_i\right) = P\left\{X_r = \tilde{A}_j / X_0 = \tilde{A}_i\right\} = \sum_{m=0}^{N} P^r\left(\tilde{A}_j / m\right) \frac{p_m \mu_{\tilde{A}_i}(m)}{P\left(\tilde{A}_i\right)}$$

La demostración es directa desde (7) y (11).

Definición 3 Se define la *cadena de Markov de estado inicial fuzzy y estado final fuzzy*, de *r* pasos, mediante la matriz:

$$\tilde{\boldsymbol{P}}^{\boldsymbol{r}} = \left[P^r\left(\tilde{A}_j / \tilde{A}_i\right)\right] \quad (13)$$

Esta matriz proporciona la probabilidad de transición del estado inicial fuzzy $\tilde{A}_i$ al estado final fuzzy $\tilde{A}_j$, en *r* pasos, $i, j \in \{1,...,n\}$.

Las matrices $\bar{\boldsymbol{P}}^{\boldsymbol{r}}$ y $\tilde{\boldsymbol{P}}^{\boldsymbol{r}}$ son estocásticas dado que la suma de cada una de sus filas es 1.

En base a estas definiciones y proposiciones proponemos el siguiente teorema:

Teorema: Las probabilidades de estado estable de la cadena de Markov con estado final fuzzy $\pi_{\tilde{A}_j}$, $j \in \{1,...,n\}$, son:

$$\pi_{\tilde{A}_j} = \sum_{s=0}^{N} \pi_s \mu_{\tilde{A}_j}(s) \quad (14)$$

es decir, $\pi_{\tilde{A}_j}$ es la probabilidad del evento fuzzy $\tilde{A}_j$ y está calculada en función de la distribución de probabilidad de estado estable π_s. El valor de $\pi_{\tilde{A}_j}$ no depende del estado inicial de la cadena de Markov, y es el mismo para los *m* estados iniciales crisp, $m \in \{0,...,N\}$ y para los $\tilde{A}_i$ estados iniciales fuzzy, $i \in \{1,2,...,n\}$. Además, el resultado (14) verifica la propiedad $\sum_{j=1}^{n} \pi_{\tilde{A}_j} = 1$.

La demostración es inmediata desde las definiciones y proposiciones establecidas previamente.

Para terminar queremos hacer notar que el cálculo de las matrices $\bar{\boldsymbol{P}}^{\boldsymbol{r}}$ y $\tilde{\boldsymbol{P}}^{\boldsymbol{r}}$ se simplifica considerablemente utilizando el cálculo matricial. Para ello definimos las matrices $\boldsymbol{Q}$ y $\boldsymbol{S}$ de la siguiente forma:

$$\boldsymbol{Q} = \begin{bmatrix} \mu_{\tilde{A}_1}(0) & \mu_{\tilde{A}_2}(0) & \cdots & \mu_{\tilde{A}_n}(0) \\ \mu_{\tilde{A}_1}(1) & \mu_{\tilde{A}_2}(1) & \cdots & \mu_{\tilde{A}_n}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{\tilde{A}_1}(N) & \mu_{\tilde{A}_2}(N) & \cdots & \mu_{\tilde{A}_n}(N) \end{bmatrix} \quad (15)$$

Esta matriz contiene los valores de las funciones de pertenencia de la partición fuzzy que determina los estados fuzzy del sistema $\left\{\tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_n\right\}$.

$$\boldsymbol{S} = \begin{bmatrix} \frac{p_0 \mu_{\tilde{A}_1}(0)}{P\left(\tilde{A}_1\right)} & \frac{p_1 \mu_{\tilde{A}_1}(1)}{P\left(\tilde{A}_1\right)} & \cdots & \frac{p_N \mu_{\tilde{A}_1}(N)}{P\left(\tilde{A}_1\right)} \\ \frac{p_0 \mu_{\tilde{A}_2}(0)}{P\left(\tilde{A}_2\right)} & \frac{p_1 \mu_{\tilde{A}_2}(1)}{P\left(\tilde{A}_2\right)} & \cdots & \frac{p_N \mu_{\tilde{A}_2}(N)}{P\left(\tilde{A}_2\right)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_0 \mu_{\tilde{A}_n}(0)}{P\left(\tilde{A}_n\right)} & \frac{p_1 \mu_{\tilde{A}_n}(1)}{P\left(\tilde{A}_n\right)} & \cdots & \frac{p_N \mu_{\tilde{A}_n}(N)}{P\left(\tilde{A}_n\right)} \end{bmatrix} \quad (16)$$

Con las matrices $\boldsymbol{Q}$ y $\boldsymbol{S}$ se llega a:

$$\bar{\boldsymbol{P}}^{\boldsymbol{r}} = \boldsymbol{P}^{\boldsymbol{r}} \boldsymbol{Q} \quad (17)$$

$$\tilde{\boldsymbol{P}}^{\boldsymbol{r}} = \boldsymbol{S}\bar{\boldsymbol{P}}^{\boldsymbol{r}} = \boldsymbol{S}\boldsymbol{P}^{\boldsymbol{r}}\boldsymbol{Q} \tag{18}$$

Y si definimos las matrices $\boldsymbol{\Pi}$ y $\tilde{\boldsymbol{\Pi}}$:

$$\boldsymbol{\Pi} = \begin{bmatrix} \pi_0 & \pi_1 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \cdots & \pi_N \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \cdots & \pi_N \end{bmatrix} \in M_{(n,N+1)} \tag{19}$$

$$\tilde{\boldsymbol{\Pi}} = \begin{bmatrix} \pi_{\tilde{A}_1} & \pi_{\tilde{A}_2} & \cdots & \pi_{\tilde{A}_n} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{\tilde{A}_1} & \pi_{\tilde{A}_2} & \cdots & \pi_{\tilde{A}_n} \end{bmatrix} \in M_{(n,n)} \tag{20}$$

entonces $\tilde{\boldsymbol{\Pi}} = \boldsymbol{\Pi}\boldsymbol{Q}$.

3 SOLUCIÓN DEL PROBLEMA DE DECISIÓN MARKOVIANO

En esta sección describimos brevemente la solución por programación lineal del proceso de decisión markoviano (Taha [4]). Para ello definimos las siguientes notaciones:

$p_{ij}(k)$: probabilidad de transición del sistema del estado *i* al estado *j* cuando se toma la decisión *k*.

C_{ik}: coste esperado en el que se incurre durante la próxima transición si el sistema está en el estado *i* y se toma la decisión *k*.

y_{ik}: probabilidad condicional de elegir la alternativa *k* dado que el sistema está en el estado *i*.

C: coste esperado por transición siguiendo una política *R*.

Una política *R* es un vector de decisiones $d_i(R)$ cuando el sistema está en el estado *i*. Así, *R* queda caracterizado por el vector $d(R) = (d_1(R), ..., d_M(R))$. Alternativamente, *R* también puede ser caracterizado asignando valores $D_{ik} = 0$ ó 1 en la matriz, en la cual cada columna debe contener un único 1 y todos los demás elementos iguales a 0:

$$\begin{array}{c} \\ \text{Estados} \end{array} \begin{array}{c} \text{Decisiones} \\ \begin{array}{cccc} 1 & 2 & \cdots & K \end{array} \\ \begin{array}{c} 1 \\ 2 \\ \cdots \\ M \end{array} \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1K} \\ D_{21} & D_{22} & \cdots & D_{2K} \\ \cdots & \cdots & \cdots & \cdots \\ D_{M1} & D_{M2} & \cdots & D_{MK} \end{bmatrix} \end{array}$$

El conjunto de decisiones D_{ik} y las variables y_{ik} están relacionados por la ecuación:

$$D_{ik} = \frac{y_{ik}}{\sum_{k=1}^{K} y_{ik}} \tag{21}$$

Con todo, el problema de decisión markoviano queda formulado mediante el siguiente programa lineal:

$$\begin{aligned} \text{Minimizar} \quad & C = \sum_{i=1}^{M}\sum_{k=1}^{K} C_{ik}\, y_{ik} \\ \text{sujeto a} \quad & \sum_{i=1}^{M}\sum_{k=1}^{K} y_{ik} = 1 \\ & \sum_{k=1}^{K} y_{jk} - \sum_{i=1}^{M}\sum_{k=1}^{K} y_{ik}\, p_{ij}(k) = 0 \\ & y_{ik} \geq 0,\ j = 1,...,M;\ i = 1,...,M;\ k = 1,...,K \end{aligned}$$

Con la solución de este programa lineal se obtienen los valores de y_{ik}, y con (21) son obtenidos los resultados de las decisiones a tomar dependiendo el estado *i*, D_{ik}.

Está demostrado que y_{ik} es estrictamente positivo para un único *k* para cada $i = 1,...,M$. Esto significa que la solución óptima automáticamente garantiza que $D_{ik} = 1$ para un único *k* para cada *i*.

Hay que tener en cuenta que para un problema con *M* estados y *K* alternativas de decisión, la formulación del programa lineal asociado debe incluir $M+1$ restricciones y MK variables, el cual tenderá a ser muy extenso para valores de *M* y *K* grandes. Y es en esta situación en donde las Cadenas de Markov con estados fuzzy pueden reducir considerablemente la complejidad computacional, proporcionando resultados más acordes con las circunstancias reales.

4 EJEMPLO NUMÉRICO

El problema está relacionado con las decisiones de publicidad a tomar en un sistema de colas, y consiste en decidir si efectuar o no publicidad teniendo en cuenta la longitud de cola en el sistema. El sistema de colas en estudio es del modelo $M/M/1/N$ y cuenta con un proceso de entrada de Poisson y tiempos de servicio acordes a una distribución exponencial, la capacidad del sistema es de 10 unidades. El estudio del comportamiento del sistema evidencia que la tasa de llegadas es $\lambda = 0.4$ cuando no se toma la decisión de efectuar publicidad y $\lambda = 0.6$ cuando hay publicidad. La tasa de servicios depende de la longitud de la cola de manera que si la longitud de unidades en cola es inferior a 5 unidades entonces la tasa de servicio es de $\mu = 0.3$, y si la longitud de unidades en cola es 5 ó más, en ese momento el servicio se refuerza aumentando la tasa de servicio a

$\mu = 0.5$. Dada la capacidad del sistema, los estados de la cadena de Markov asociada son $X = \{0,1,...,10\}$.

Hay dos alternativas de decisión, (1) no publicidad ($k = 1$) y (2) publicidad ($k = 2$).Si denotamos por P la matriz de probabilidad de transición asociada con *X*, entonces dados los valores de λ y μ, es:

$P =$

State	0	1	2	⋯	8	9	10
0	$\bar{\lambda}$	λ	0	⋮	0	0	0
1	$\bar{\lambda}\cdot\mu$	$\lambda\cdot\mu+\bar{\lambda}\cdot\bar{\mu}$	$\lambda\cdot\bar{\mu}$	⋮	0	0	0
2	0	$\bar{\lambda}\cdot\mu$	$\lambda\cdot\mu+\bar{\lambda}\cdot\bar{\mu}$	⋮	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
9	0	0	0	⋮	$\bar{\lambda}\cdot\mu$	$\lambda\cdot\mu+\bar{\lambda}\cdot\bar{\mu}$	$\lambda\cdot\bar{\mu}$
10	0	0	0	⋮	0	μ	$\bar{\mu}$

Y si denotamos por P_k la matriz de probabilidad de transición asociada con *X* para la decisión *k*, entonces las matrices P_k para $k = 1, 2$ son:

$$P_1 = \begin{bmatrix} 0.6 & 0.4 & 0 & \vdots & 0 & 0 & 0 \\ 0.18 & 0.54 & 0.28 & \vdots & 0 & 0 & 0 \\ 0 & 0.18 & 0.54 & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & 0.5 & 0.2 & 0 \\ 0 & 0 & 0 & \vdots & 0.3 & 0.5 & 0.2 \\ 0 & 0 & 0 & \vdots & 0 & 0.5 & 0.5 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0.4 & 0.6 & 0 & \vdots & 0 & 0 & 0 \\ 0.12 & 0.46 & 0.42 & \vdots & 0 & 0 & 0 \\ 0 & 0.12 & 0.46 & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & 0.5 & 0.3 & 0 \\ 0 & 0 & 0 & \vdots & 0.2 & 0.5 & 0.3 \\ 0 & 0 & 0 & \vdots & 0 & 0.5 & 0.5 \end{bmatrix}$$

Para encontrar la solución por programación lineal del proceso de decisión planteado el programa lineal debe contener 12 restricciones y 22 variables. Este método de cálculo puede ser sensiblemente reducido si los estados del sistema son fuzzy. Así, en el sistema las decisiones serán asociadas con tres estados fuzzy denotados $\tilde{A}_1, \tilde{A}_2$ y $\tilde{A}_3$ que corresponden a "longitud corta de cola", "longitud media de cola" y "longitud larga de cola", respectivamente. Y sobre los 11 estados del sistema se establece la siguiente partición fuzzy que relaciona los estados iniciales del sistema con los estados fuzzy, definidos mediante la matriz $\boldsymbol{Q}$:

$$\boldsymbol{Q}^T = \begin{bmatrix} 1 & 1 & 0.8 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.6 & 1 & 1 & 1 & 0.6 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 & 0.8 & 1 & 1 \end{bmatrix}$$

Hay que reseñar que los estados del sistema, *X* son conocidos en los problemas específicos de modelos de decisión de colas. Sin embargo hay que decidir los estados fuzzy correspondientes y sus funciones de pertenencia, para poder desarrollar la metodología propuesta. Esto puede hacerse a través del método formulado por Yager [6] o el método de Dubois et al. [2] para la construcción de funciones de pertenencia a partir de datos obtenidos de diversas fuentes o de expertos.

Se han calculado los costes esperados para las dos decisiones y los tres estados fuzzy y se muestran en la tabla 1.

Tabla 1: Costes esperados C_{ik}

Estados fuzzy	$k = 1$	$k = 2$
$\tilde{A}_1$	10	3
$\tilde{A}_2$	7	5
$\tilde{A}_3$	4	7

Supondremos que las probabilidades de estado inicial p_i son todas iguales, luego $p_i = 1/11$. Con este dato se obtiene que las probabilidades de estado inicial fuzzy son:

$$P(\tilde{A}_1) = 3.2/11 \qquad P(\tilde{A}_2) = 4.6/11 \qquad P(\tilde{A}_3) = 3.2/11$$

Y la matriz $\boldsymbol{S}$:

$$\boldsymbol{S} = \begin{bmatrix} \frac{1}{3.2} & \frac{1}{3.2} & \frac{0.8}{3.2} & \frac{0.4}{3.2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{0.2}{4.6} & \frac{0.6}{4.6} & \frac{1}{4.6} & \frac{1}{4.6} & \frac{1}{4.6} & \frac{0.6}{4.6} & \frac{0.2}{4.6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{0.4}{3.2} & \frac{0.8}{3.2} & \frac{1}{3.2} & \frac{1}{3.2} \end{bmatrix}$$

Las matrices de probabilidad transición correspondientes a los estados fuzzy $\tilde{A}_1, \tilde{A}_2$ y $\tilde{A}_3$ para la decisión *k*, $\tilde{\boldsymbol{P}}_{\boldsymbol{k}} = \left[P_k\left(\tilde{A}_j / \tilde{A}_i\right) \right]$ se obtienen usando (18): $\tilde{\boldsymbol{P}}_{\boldsymbol{k}} = \boldsymbol{S}\boldsymbol{P}_{\boldsymbol{k}}\boldsymbol{Q}$:

$$\tilde{P}_1 = \begin{bmatrix} 0.8335 & 0.1665 & 0 \\ 0.0941 & 0.8103 & 0.0956 \\ 0 & 0.1687 & 0.8313 \end{bmatrix}$$

$$\tilde{P}_2 = \begin{bmatrix} 0.7978 & 0.2022 & 0 \\ 0.0755 & 0.8071 & 0.1174 \\ 0 & 0.1375 & 0.8625 \end{bmatrix}$$

Con todo, el programa lineal a plantear para encontrar la solución al problema de si efectuar o no publicidad

teniendo en cuenta la longitud de cola en el sistema, consta de 4 restricciones y 6 variables, y es:

$$\text{Minimizar} \quad C = 10y_{11} + 3y_{12} + 7y_{21} + 5y_{22} + 4y_{31} + 7y_{32}$$

sujeto a

$$y_{11} + y_{12} + y_{21} + y_{22} + y_{31} + y_{32} = 1$$

$$y_{11} + y_{12} - \\ -\left(0.8335y_{11} + .07978y_{12} + 0.0941y_{21} + 0.0755y_{22}\right) = 0$$

$$y_{21} + y_{22} - \left(0.1665y_{11} + 0.2022y_{12} + 0.8103y_{21} + \right. \\ \left. +0.8071y_{22} + 0.1687y_{31} + 0.1375y_{32}\right) = 0$$

$$y_{31} + y_{32} - \\ -\left(0.0956y_{21} + 0.1174y_{22} + 0.8313y_{31} + 0.8625y_{32}\right) = 0$$

$$y_{ik} \geq 0, \quad i = 1,2,3; \quad k = 1,2$$

La solución óptima es $y_{11} = y_{21} = y_{32} = 0$ y $y_{12} = 0.1804$, $y_{22} = 0.4834$ y $y_{31} = 0.3362$. Luego utilizando (21) es:

$$\begin{array}{cc} & \text{Decisiones} \\ & 1 \quad 2 \\ \text{Estados} & \begin{array}{c} \tilde{A}_1 \\ \tilde{A}_2 \\ \tilde{A}_3 \end{array} \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

Así, la política óptima selecciona la alternativa 2 ($k = 2$), es decir efectuar propaganda cuando el sistema se encuentra en los estados fuzzy $\tilde{A}_1$ y $\tilde{A}_2$ (longitud corta y media de cola) y la alternativa 1 ($k = 1$), es decir no efectuar propaganda cuando el sistema se encuentra en el estado fuzzy $\tilde{A}_3$ (longitud larga de cola), siendo el valor óptimo de *C* igual a 3.697.

Conclusiones

A través del concepto de probabilidad de eventos fuzzy hemos adaptado los procesos de decisión Markovianos a las situaciones en las cuales los estados pueden ser considerados fuzzy, bien porque los estados del sistema no puedan ser medidos con precisión, bien porque los estados son perfectamente medibles y observables pero el número de estados es tan amplio que las decisiones no pueden ser asociadas con los estados exactos el sistema. Así, tomando como punto de partida la probabilidad de eventos fuzzy en este trabajo hemos desarrollado la teoría de las Cadenas de Markov con estados fuzzy calculando, por un lado la matriz de probabilidad de transición entre estados fuzzy y, por otro lado, las probabilidades de estado estable de la Cadena de Markov con estados fuzzy.

Si el sistema en estudio cuenta con un amplio número de estados la solución del problema de decisión markoviano por programación lineal proporciona un sistema con un gran número de variables y restricciones. Mediante el uso de las Cadenas de Markov con estados fuzzy, tal como hemos mostrado en el ejemplo, el sistema a resolver es reducido considerablemente lo que conlleva una disminución de la complejidad computacional así como de la toma de decisiones ya que ésta queda limitada a sólo tres estados (longitud corta, media o larga de cola).

Si bien en este trabajo nos hemos limitado a la aplicación de las Cadenas de Markov con estados fuzzy en la Teoría de Colas, éstas pueden ser utilizadas tanto en problemas de inventarios como en mantenimiento de sistemas. Además, las Cadenas de Markov con estados fuzzy pueden ser utilizadas para la optimización y diseño de un sistema de líneas de flujo con capacidad limitada y datos fuzzy tal como se muestra en de la Fuente y Pardo [3].

Referencias

[1] D. Dubois, H. Prade. Fuzzy Sets and Systems: Theory and Applications. Academic Press, Boston, 1980.

[2] D. Dubois, H. Prade, R.R. Yager. Merging fuzzy information. En: J. Bezdek, D. Dubois, H. Prade (Eds.), Fuzzy Sets in Approximate Reasoning and Information Systems. Kluwer, Boston 1999, pág. 335-401.

[3] D. de la Fuente, M.J. Pardo. Calculation of the degree of customer satisfaction for the optimization and design of a flow line system with finite capacity and fuzzy data. XIV Congress of International Association for Fuzzy Set Management and Economy, SIGEF07. Poiana Braşov, Romania, pág. 540-558, 2007.

[4] H.A. Taha. Operations Research: An Introduction (seventh ed.). Prentice-Hall, New Jersey, 2003.

[5] W.L. Winston. Operations Research: Applications and Algorithms (fourth ed.). Duxbury Press, Boston, 2003.

[6] R.R. Yager. Level sets for membership evaluation of fuzzy subsets. En R.R. Yager (Ed.), Fuzzy Sets and Possibility Theory: Recent Developments. Pergamon Press, Oxford, 1982, pág. 90-97.

[7] L.A. Zadeh. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications* 23, pág. 421-427, 1968.

[8] H.J. Zimmermann. Fuzzy Set Theory and its applications (fourth ed.). Kluwer, Boston, 2001.

MODELADO DIFUSO DE UN SISTEMA FÍSICO

Matilde Santos[1] **Victoria López**[1] **Luis Garmendia**[2]

[1]Dpto. Arquitectura de Computadores y Automática
[2]Dpto. Ingeniería del Software e I. A.
Facultad de Informática, Universidad Complutense de Madrid, Spain,
msantos@dacya.ucm.es
vlopez@fdi.ucm.es

Resumen

En este trabajo se presenta una aproximación al modelado de un sistema físico de segundo orden mediante reglas difusas. La respuesta del sistema se ha dividido en dos tramos de forma explícita y para cada uno de ellos se ha definido y simulado un sistema difuso que reproduce el comportamiento del proceso en esa fase. El modelo se basa en el error.

Palabras Clave: Modelo difuso, sistema físico, control basado en modelo.

1 INTRODUCCIÓN

Los controladores difusos se han aplicado ampliamente al control de sistemas complejos ya que no requieren un modelo matemático preciso de los mismos [9]. Sin embargo, estos sistemas difusos seleccionan una acción de control basándose en las predicciones de un modelo implícito [6], [8]. Es decir, el conocimiento del comportamiento dinámico del proceso es necesario para formular las reglas heurísticas que especifican la señal de control.

En este trabajo se muestra el diseño cualitativo de un modelo difuso de un sistema físico, obteniéndolo de forma explícita mediante reglas. La utilidad de este tipo de modelos estriba en que proporcionan una representación del comportamiento del mismo sin necesidad de la información y el esfuerzo que requieren las técnicas de modelado clásicas y permiten entender mejor el funcionamiento del sistema [7]. Además puede ser la base de un controlador difuso basado en modelo (Figura 1) ya que el modelo difuso de una planta consiste en definitiva en un conjunto de reglas que representan la dinámica inversa del controlador [1].

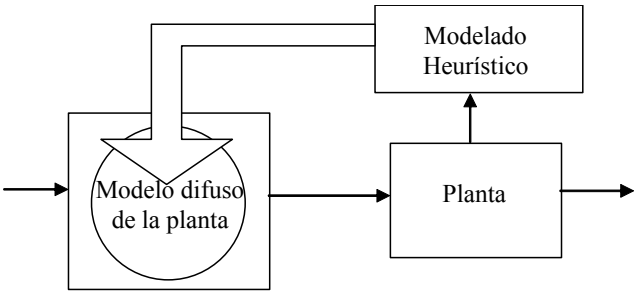


Figura 1: Esquema de control basado en modelo difuso

Un sistema adecuado sobre el que desarrollar este tipo de modelos puede ser cualquier proceso físico (mecánico, electrónico, etc.). La representación de un sistema físico mediante técnicas de modelado da lugar a un sistema de ecuaciones diferenciales. La mayoría de estos modelos pueden simplificarse y reducirse a modelos de primer o segundo orden, o a una combinación de éstos. Por ese motivo en este trabajo se trata con un sistema cuya respuesta es la correspondiente a un segundo orden, lo que permite generalizar el estudio a un amplio rango de sistemas reales.

El artículo se estructura de la siguiente manera. Tras la introducción, la sección 2 describe el desarrollo del modelo de un sistema físico que servirá de base para definir el modelo difuso. En la sección 3, tras comentar brevemente algunos conceptos del control con lógica difusa, se presenta el modelo del sistema basado en reglas difusas así como resultados de simulación. La última sección recoge las conclusiones y trabajos futuros.

2. RELACIONES FUNCIONALES DEL SISTEMA

Hay muchos procesos físicos que tienen una respuesta que puede ser descrita como un sistema de segundo orden. Algunos ejemplos bien conocidos son el de un

amortiguador, dos depósitos acoplados, un circuito RLC, etc. Se podrían introducir retardos o no linealidades al sistema siempre que el comportamiento global del mismo no varíe de forma drástica (por ejemplo, que no resulte inestable o de fase mínima), sin perder generalidad.

Existen diversas formas de representar un sistema para su simulación y el análisis de su respuesta. Las dos grandes vías de obtener esta representación son la identificación y el modelado. El modelado de un sistema se basa en el comportamiento físico del mismo, y utiliza junto con ciertas hipótesis, las leyes de la naturaleza que describen su funcionamiento. Por eso esta representación resulta especialmente apta para obtener el conocimiento necesario para describir en forma de reglas su comportamiento.

Para ejemplificar la obtención de un modelo difuso de un sistema de segundo orden y observar la dependencia de las variables que definen su comportamiento, partimos de un sistema real compuesto por un muelle al que está ligada una masa que se desplaza en un entorno con rozamiento. Este sistema, que representa un amortiguador, se puede expresar como una ecuación diferencial en función de las variables físicas que intervienen en el proceso, que viene descrito por las leyes de la mecánica. Considerando para el mismo una masa *M*, un coeficiente de fricción *f* y una constante del muelle *K*, para una entrada de fuerza *u(t)* la ecuación diferencial que describe su comportamiento es,

$$M\frac{d^2y(t)}{dt^2} + f\frac{dy(t)}{dt} + Ky(t) = u(t) \qquad (1)$$

siendo *y(t)* la señal de salida que corresponde al desplazamiento de la masa en el eje de movimiento.

Aplicamos la trasformada de Laplace a esta ecuación (1) para obtener la función de transferencia *G(s)*, que relaciona la entrada y la salida del mismo en ese dominio de Laplace, puesto que esta representación va a permitir generalizar el modelo:

$$\frac{Y(s)}{U(s)} = \frac{1}{Ms^2 + fs + K} = G(s) \qquad (2)$$

donde hemos supuesto condiciones iniciales nulas, es decir, sistema en reposo.

Es bien conocido que la formulación general para cualquier sistema de segundo orden en el dominio de Laplace viene dada por (3),

$$\frac{Y(s)}{U(s)} = \frac{K.e^{-s.t_d}}{(\tau_1 s+1)(\tau_2 s+1)} \qquad (3)$$

donde Y (s) es la señal de salida o respuesta.
U(s) es la señal de entrada
K es la ganancia del sistema
t_d es el tiempo de retardo
τ_1 y τ_2 son las constantes del sistema (del tiempo de proceso).

Suponiendo que el tiempo de retardo es muy pequeño, $t_d \approx 0$, y que $\tau_1 >> \tau_2$, la ecuación de la planta de un sistema de segundo orden se reduce a:

$$G(s) = \frac{K}{(\tau_1 s+1)(\tau_2 s+1)} \qquad (4)$$

Lo que permite establecer las siguientes relaciones funcionales entre los parámetros físicos del sistema real que sirve como ejemplo y la función de transferencia de un sistema de segundo orden:

$$\tau_1\tau_2 = \frac{M}{K}$$
$$\tau_1 + \tau_2 = \frac{f}{K}$$
$$Kp = \frac{1}{K}$$

o lo que es equivalente, la ecuación del sistema se puede reformular

$$\tau_1\tau_2\ddot{y}(t) + (\tau_1 + \tau_2)\dot{y}(t) + y(t) = Kp.u(t)$$

De esta forma podemos relacionar las dos constantes de tiempo que determinan la respuesta con parámetros físicos del sistema y esto nos permite entender mejor cómo varía la respuesta al modificar estos valores.

Para la aplicación computacional de un controlador es necesario tener un modelo discreto en el tiempo. Esta representación discreta se puede expresar en función de aproximaciones (backward o forward) de la derivada. El tiempo de muestreo debe ser seleccionado atendiendo a las dos constantes de tiempo que definen la respuesta del sistema (si es de segundo orden como en este caso). Se han utilizado las siguientes aproximaciones para discretizar las derivadas:

$$\left.\frac{dy}{dt}\right|_t = \frac{y(t+1) - y(t)}{T} \qquad (5)$$

$$\left.\frac{d^2y}{d^2t}\right|_t = \frac{y(t+2) - 2y(t+1) + y(t)}{T^2} \qquad (6)$$

donde T el periodo de muestreo.

En el caso del sistema físico que nos ocupa, la aplicación de esas aproximaciones da lugar a la expresión:

$$y(t+2)=\left(2-\frac{Tf}{M}\right)y(t+1)+\left(\frac{T(f-TK)}{M}-1\right)y(t)+\frac{T^2}{M}u(t) \quad (7)$$

Suponiendo que la referencia que se pretende alcanzar es el valor estacionario del proceso y_{ss}, la salida y(t) puede reemplazarse por el error e(t). De esta forma se expresa el comportamiento del sistema en función de la diferencia entre la respuesta actual y el valor en el estado estacionario o valor deseado (8).

$$e(t)=y(t)-y(t_{ss}) \quad (8)$$

donde $y(t_{ss}) = y_{ss}$ = constante.

Despejando y(t) en (8) y sustituyéndolo en la ecuación (7) se obtiene:

$$e(t+2)=\left(2-\frac{Tf}{M}\right)e(t+1)+\left(\frac{T(f-TK)}{M}-1\right)(e(t)-y_{ss})+\frac{T^2}{M}u(t) \quad (9)$$

o lo que es equivalente, en general:

$$e(t+2)=f\left(e(t+1),e(t),u(t)\right) \quad (10)$$

En el tratamiento de este sistema se va a trabajar con la primera derivada, es decir, con el valor de la señal en el siguiente instante de tiempo o próximo estado. Para ello integramos (10) de forma que trabajamos con la formulación incremental, dando lugar a:

$$e(t+1)=f\left(e(t),e(t-1),u(t-1)\right) \quad (11)$$

por lo que el próximo estado dependerá del valor actual y de los valores anteriores de esa variable de estado, y del valor anterior de la entrada.

Para el ejemplo del sistema que se está desarrollando, se reformula en función de estas variables como:

$$e(t+1)=\left(2-\frac{Tf}{M}\right)e(t)+\left(\frac{T(f-TK)}{M}-1\right)e(t-1)+\frac{T^2}{M}u(t-1) \quad (12)$$

El sistema que se va a enunciar a continuación mediante reglas difusas representa en general a cualquier sistema de segundo orden, y por lo tanto a éste sistema físico real así como a un amplio rango de procesos. Estas ecuaciones nos han permitido delimitar las relaciones entre las variables del sistema y expresarlo en función del error.

3. MODELO BASADO EN LÓGICA DIFUSA

En 1965 Lofti A. Zadeh presentó su artículo "Fuzzy sets" [10] donde introduce la teoría sobre conjuntos difusos. La primera aplicación de la lógica borrosa al campo del control data de 1974 [4], y desde entonces han sido muy numerosas sus aplicaciones en la automática en muy diversos ámbitos.

En control basado en lógica difusa se trabaja con variables lingüísticas que describen de forma cualitativa las variables del sistema. En el caso de un sistema como el que se propone, éstas serían la respuesta del sistema *y(t)* que se puede enunciar, como hemos visto, en función del error *e(t)*, de sus valores anteriores y futuros. Por lo tanto denominaremos a las variables lingüísticas: valor_anterior, valor_actual, y valor_próximo.

Estas variables tienen un valor analógico concreto que hay que convertir a difuso. Para ello se definen para cada una de ella unos conjuntos difusos representados por funciones de pertenencia distribuidas a lo largo de sus correspondientes universos de discurso.

Posteriormente, una vez definidas las variables difusas, la relación funcional que modela el sistema puede ser descrita mediante reglas de producción del tipo "*If … then …*", donde los antecedentes y consecuentes de esas reglas corresponden a valores de pertenencia de las variables de entrada o salida, respectivamente, a los conjuntos difusos [5]. Para obtener la salida del sistema y representarla se aplica la operación inversa a la difusificación.

En este trabajo se ha utilizado el Centro del Área como método de desborrosificación; las operaciones de máx-prod para la inferencia porque da buenos resultado en el campo del control [2], y Mamdani para convertir de valore nítidos a difusos las entradas.

Se ha simulado el proceso con el paquete computacional Matlab© que tiene una herramienta específica para lógica borrosa, implementando sistemas difusos tipo Mandani (Figura 2).

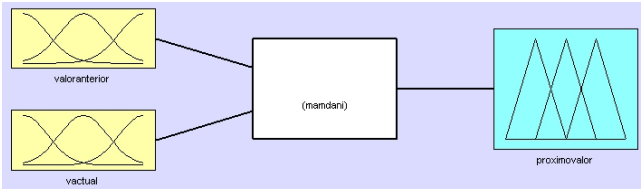


Figura 2: Sistema Difuso empleado para el modelo

Para explicitar el modelo, la respuesta del sistema de segundo orden se ha dividido en dos secciones (Figura 3). El primer tramo corresponde al tiempo de subida, es decir, a la respuesta desde el instante inicial en el que se excita el sistema hasta que ésta alcanza el primer pico en amplitud. Esta partición es arbitraria puesto que se podría considerar también como tiempo de subida el tramo de la respuesta hasta que ésta alcanza por primera vez el valor que tomará en el estacionario, es decir, cuando cruza la referencia. Hay diversas convenciones para definir el tiempo de subida, y aquí hemos adoptado una de ellas que resultara fácil de plasma en reglas.

La segunda sección representa las oscilaciones de la respuesta del sistema desde el punto anterior hasta que alcanza el estado estacionario, e incluye a éste; por lo tanto desde el pico de la máxima oscilación hasta que se estabiliza a un valor constante.

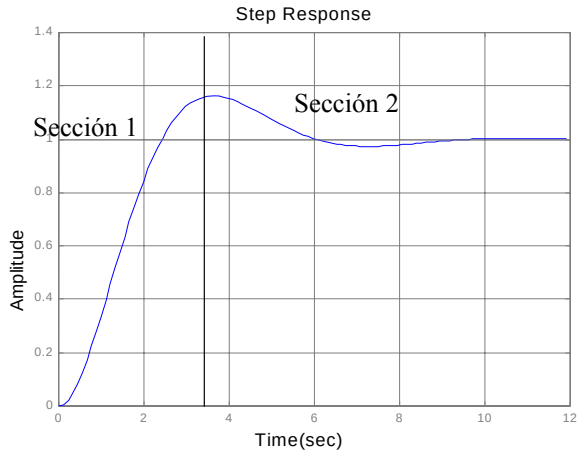


Figura 3: Secciones de la respuesta del sistema de segundo orden

Podrían modelarse otras características de la respuesta como el retardo, si se hubiera considerado ese efecto, es decir, el tiempo durante el cual el sistema ha recibido la entrada y aún no ha respondido variando el valor de su respuesta.

3.1. MODELO DE LA SECCIÓN 1 DE LA RESPUESTA

Para modelar el primer tramo de la respuesta, cuando ésta es monótona y creciente, se ha definido un sistema difuso con una variable de entrada, "valor_actual", y una variable de salida, "valor_próximo". Para la entrada se han definido tres conjuntos difusos gaussianos que corresponden a las etiquetas MuyBajo, Medio, Alto (Figura 4). Para la salida se trabaja con cinco conjuntos también gaussianos que corresponden a las etiquetas MuyBajo, Bajo, Medio, Unitario, MuyAlto (Figura 5). Están uniformemente distribuidos en el universo de discurso de las variables.

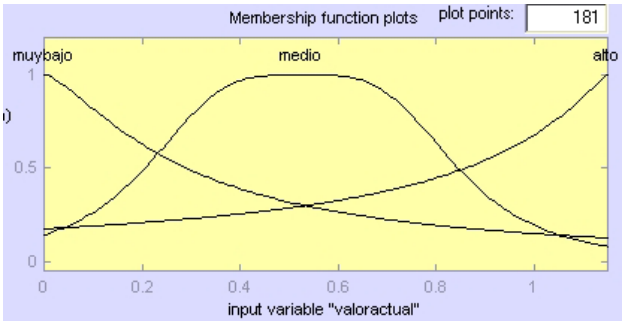


Figura 4: Conjuntos difusos de la variable de entrada.

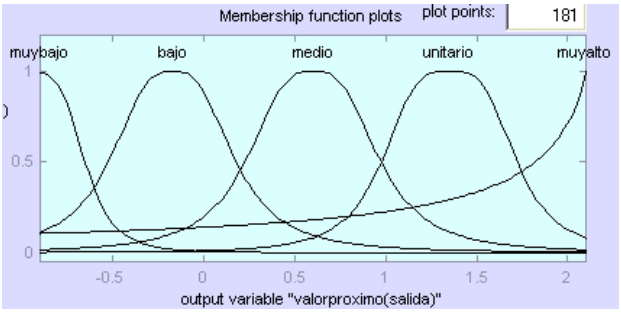


Figura 5: Conjuntos difusos de la variable de salida.

El comportamiento cualitativo del tramo primero respecto al error tendrá dos fases, según haya superado o no el valor de la referencia, que harán que cambie el signo del error. Es decir, habrá una primera fase en la que la salida aún no ha alcanzado al valor del estacionario y la respuesta está por debajo de la referencia; y una segunda fase en la cual se empieza a producir una sobre-elongación cuando la salida supera la referencia.

Este comportamiento debe reflejarse en la reglas. Por lo tanto, si la respuesta del sistema es más pequeña que la señal de referencia entonces el error actual es positivo (medio, unitario o alto), y lo que queremos conseguir es que el próximo valor sea también positivo. Si la respuesta supera la señal de referencia deseada, el error es negativo (muy bajo o bajo) por lo que se intenta que la velocidad del cambio disminuya (Figuras 6 y 7). El resultado de la simulación de este tramo de la respuesta con este sistema difuso se muestra en Figura 8.

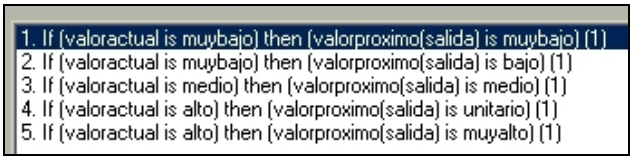


Figura 6: Tabla de reglas difusas para la sección 1

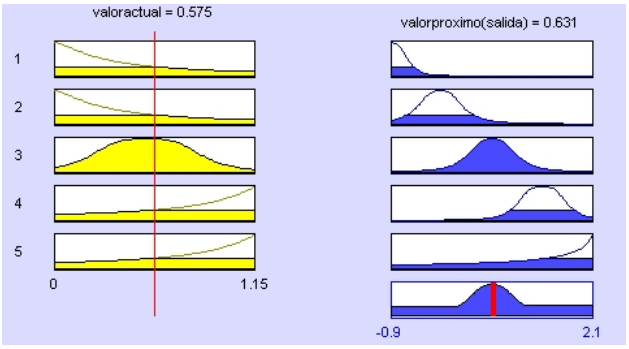


Figura 7: Actuación de las reglas de la sección 1

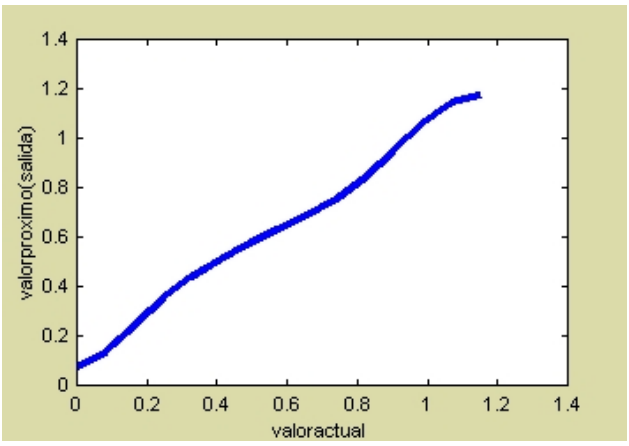


Figura 8: Modelo difuso del primer tramo del sistema de segundo orden.

Se puede observar que este modelo difuso sigue bien la trayectoria creciente propia de un sistema de segundo orden durante su transitorio.

3.2 MODELO DE LA SECCIÓN 2 DE LA RESPUESTA

Para modelar la sección 2 de la respuesta (Figura 3) se ha trabajado con las variables "valor_anterior" y "valor_actual" como entradas al sistema difuso, que infieren el "próximo valor" como salida. Para cada una de las entradas se emplean tres conjuntos difusos definidos por funciones triangulares que corresponden a las etiquetas: P: positivo, Z: cero, y N: negativo (Figura 9).

Para la salida se han definido cinco conjuntos triangulares asignados a: GN: negativo grande, MN: negativo medio, Z: cero, MP: positivo medio, GP: positivo grande (Figura 10). Con esto conseguimos mayor granularidad y por tanto mayor precisión a la salida, como es habitual. Los conjuntos difusos están uniformemente distribuidos en los correspondientes universos de discurso.

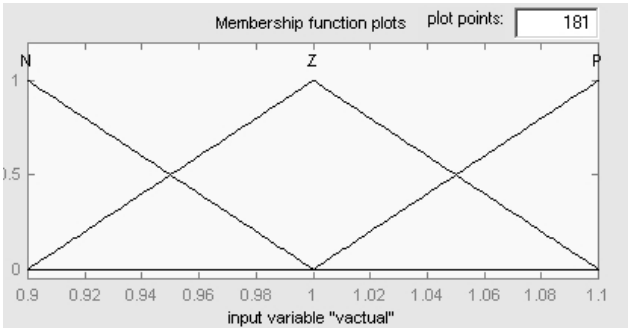


Figura 9: Conjuntos difusos de las variables de entrada.

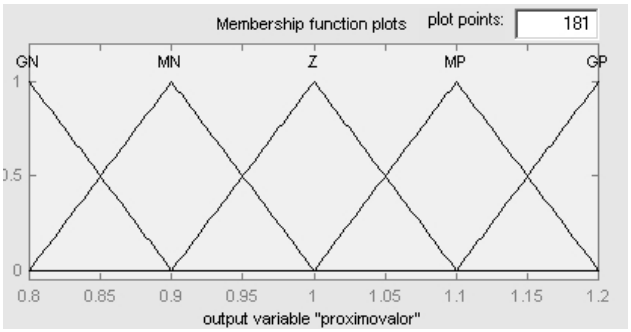


Figura 10: Conjuntos difusos de la variable de salida.

El comportamiento cualitativo de la respuesta en este tramo es más complejo ya que hay fases en las que la respuesta está por encima de la referencia (por ejemplo, la parte inicial), y a partir de ese momento el error va cambiando de signo según se producen las oscilaciones (Figura 3).

Las reglas que lo definen recogen el conocimiento del tipo: si el error anterior es negativo y el error actual es cercano al cero o ligeramente positivo, el factor de amortiguamiento se debe incrementar por lo que el próximo valor de la salida será positivo. En la figura 11 se muestra la tabla de reglas que se ha implementado, en la figura 12 se representa cómo actúan y en la figura 13 la superficie de control generada por este sistema difuso.

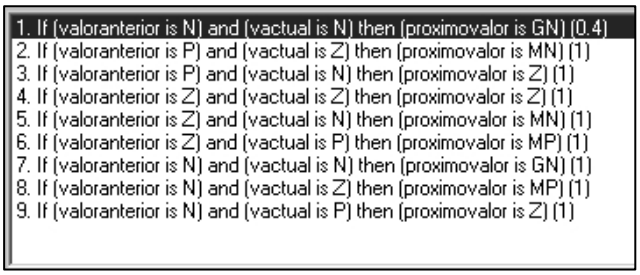


Figura 11: Tabla de reglas difusas para la sección 2

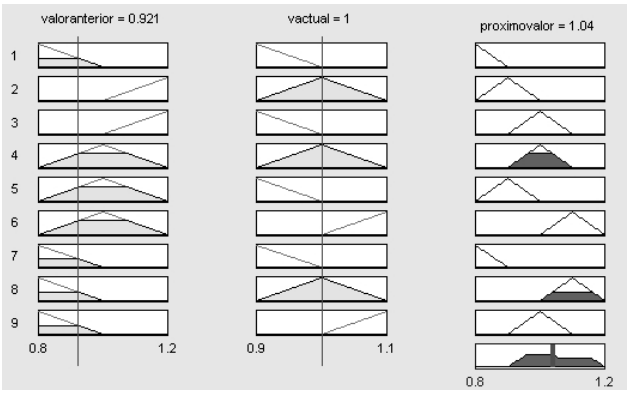


Figura 12: Actuación de las reglas de la sección 2

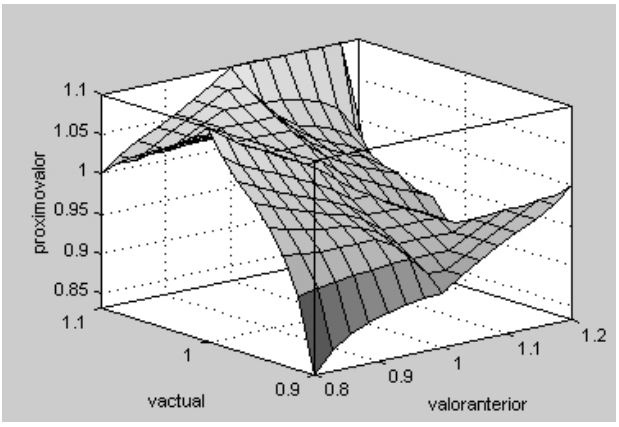


Figura 13: Superficie generada por el sistema difuso para la sección 2 del modelo

Este tipo de modelos basados en reglas difusas son aptos para usarlos en un esquema de controlador fuzzy basado en modelo [1], [3].

5. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se ha desarrollado una primera aproximación al modelo difuso de un sistema físico de segundo orden en base al conocimiento cualitativo de su respuesta. Se basa en el error y en función de estas variables permite predecir el valor de la respuesta del sistema en el tiempo.

El modelo se explicita en forma de reglas con antecedentes y consecuentes difusos, aplicando consideraciones lógicas para generarlas.

Como trabajo futuro en esta línea del modelado cualitativo se pretende implementar modelos que consideren más variables como antecedente de las reglas, para definir de forma más precisa el comportamiento de la respuesta.

Otra posible ampliación del mismo sería desarrollar un control fuzzy adaptativo de sistemas de orden elevado [2].

Agradecimientos

Los autores quieren agradecer la ayuda del Proyecto de Innovación Educativa y Mejora de la Calidad Docente de la UCM nº 56/2007.

Referencias

[1] Czogala, E., Pedrycz, W.: On Identification in fuzzy systems and its applications in control problems, *Fuzzy Sets and Systems*, 6, 73-83, 1981.

[2] Graham, B.P., Newell, R.B.: Fuzzy adaptive control of a first-order process. *Fuzzy Sets and Systems*, 31, 47-65, 1989.

[3] Graham, B.P., Newell, R.B.: An adaptive fuzzy model-based controller. IJCAI, *Fuzzy Logic and Fuzzy Control*, 56-66, Sydney, Australia, 1991.

[4] Mandani, E.H.: *Application of fuzzy algorithms for control of simple dynamic plant*. Proceedings IEE, Vol. 121, No. 12, December, pp.1585-1588, 1974.

[5] Oh, S., Pedrycz, W., Park, B.: Hybrid identification of fuzzy rule-based models. *Int. Journal of Intelligent Systems*, vol. 17 (1), 77-103, 2002.

[6] Pedrycz, W.: Fuzzy Dynamic Systems. *IJCAI, Fuzzy Logic and Fuzzy Control*, 37-44, Sydney, Australia 1991.

[7] Pedrycz, W.: Modelling with fuzzy sets in fuzzy control. *Fuzzy Days*, 3-34, 1992.

[8] Santos, M., López, R. de la Cruz, J.M.: A Neuro-Fuzzy Approach to Fast Ferry Vertical Motion Modelling, *Engineering Application of Artificial Intelligence,* vol. 19, pág. 313-321, 2006.

[9] Sugeno, M.: An introductory survey of fuzzy control. *Information Sciences*, n. 36, 59-83, 1985.

[10] Zadeh, L.A.: Fuzzy sets. *Proceedings of the IEEE (Information and Control)*, Vol:8: 338-353, 1965.

APLICACIONES DE OPERADORES MORFOLÓGICOS BASADOS EN UNINORMAS

Manuel González Hidalgo[1] **Arnau Mir Torres**[2] **Daniel Ruiz Aguilera**[2] **Joan Torrens Sastre**[2]

[1] Grupo de Gráficos, Visión e Inteligencia Artificial, Departamento de Matemáticas y Informática, Universidad de las Islas Baleares, {manuel.gonzalez}@uib.es

[2] Grupo de Lógica Borrosa y Fusión de la Información, Departamento de Matemáticas y Informática, Universidad de las Islas Baleares, {arnau.mir,dmijts0,daniel.ruiz}@uib.es

Resumen

En este trabajo se presenta una continuación del estudio realizado en [11] sobre una morfología matemática basada en uninormas conjuntivas continuas por la izquierda. Se presentan resultados experimentales sobre las transformaciones Top-Hat en dicha morfología, usadas para destacar determinadas componentes de la imagen, y sobre el filtrado alterno, usado en la eliminación de ruido.

Palabras Clave: Morfología matemática, Top-Hat, filtros alternos, uninormas, uninormas representables e idempotentes.

1. INTRODUCCIÓN

La morfología matemática borrosa es una generalización de la morfología binaria [19] usando técnicas de conjuntos borrosos [7], [2], [3], [5]. Las herramientas básicas de la morfología matemática son las llamadas operaciones morfológicas que se aplican a una imagen A, modificándola a través de un "elemento estructurante" B, cuyo tamaño y forma se eligen para estudiar la estructura de la imagen A. Las operaciones morfológicas básicas son la erosión, la dilatación, la apertura y el cierre. Otras operaciones morfológicas borrosas y extensiones han sido introducidas por otros autores, [16] y [5].

Los operadores borrosos usados para construir una morfología matemática son los conjuntores y los implicadores. Entre ellos, los más usados son las t-normas continuas y sus implicaciones residuales. Recientemente las uninormas conjuntivas, como caso particular de conjuntores, han sido usadas también en este ámbito [4], [10] y [11]. En estos dos últimos trabajos se usan dos tipos de uninormas conjuntivas para obtener una morfología matemática con "buenas" propiedades, incluyendo dualidad entre los operadores morfológicos.

Este trabajo se puede considerar una continuación del trabajo [11] donde se comparaban los resultados de la morfología basada en uninormas con la basada en t-normas y el "umbra approach". En dicho trabajo se comprobaba cómo con las uninormas se detectan mejor los contornos de las imágenes que con las otras aproximaciones. Se demuestran también en [11] propiedades de idempotencia e idempotencia generalizada para los operadores de apertura y cierre basados en uninormas. Así, en este trabajo pretendemos extender este estudio a los operadores "Top-Hat" (residuos entre aperturas y cierres) basados en uninormas, así como a los llamados filtros alternos (composiciones alternas de aperturas y cierres). El Top-Hat se utiliza para destacar determinadas componentes de la imagen, mientras que los filtros alternos son usados para la eliminación de ruido.

La importancia de los gradientes morfológicos radica en el hecho de que siguen siendo objeto de estudio y de aplicaciones. Así, en [12] se propone un algoritmo direccional de detección de transiciones basado en residuos morfológicos, considerando elementos estructurantes lineales. En el año 2002, Q.H. Wu y otros en [6] diseñan, sobre la base de la morfología matemática, un nuevo detector de bordes suaves en regiones oscuras llamado "Pseudo Top-Hat" el cual proporciona, en el caso de la clásica imagen del cámara de video (ver figura 1), un desarrollo mucho mejor que el obtenido por otros detectores de bordes. En el año 2005, en [22] se usan gradientes morfológicos basados en aperturas y cierres para detectar bordes en imágenes médicas de tipo CT alteradas por ruido. Los resultados obtenidos son comparados con otros detectores de contornos, mejorando los resultados. Más recientemente, en [13] y en [17] se aplican operadores gradiente de nuevo para la detección de bordes y la segmentación de imágenes, respectivamente. Así, en [13] se desarrolla un detector de bordes para la obtención de bordes "finos" en regiones o imágenes con bajo contraste. Se reali-

zan pruebas con varias imágenes, incluso algunas de ellas alteradas por ruido tipo "sal y pimienta" y ruido gaussiano de media cero y varianza 0.02, obteniendo buenos resultados en comparación con los otros cuatro métodos con los que trabajan. En [17] se diseñan operadores morfológicos multiescala basados en gradientes, enfocándolos a la segmentación de imágenes de Tomografía Axial Computerizada (TAC). Todo esto pone de manifiesto que los gradientes morfológicos siguen siendo de interés y útiles en el análisis y procesamiento de imágenes.

En cuanto a la eliminación y suavizado del ruido podemos decir que es una de las tareas fundamentales del análisis y procesamiento de imágenes como una fase previa de la visión artificial, pues muchas técnicas de interpretación, medida, segmentación, detección de estructuras, etc., fallarán en presencia de ruido. En [22], [13] los operadores gradientes diseñados son robustos en presencia de ruido. Mucho esfuerzo se ha dedicado, y se sigue dedicando, en la literatura al tema del suavizado y eliminación del ruido, tanto desde el punto de vista "crisp" como "fuzzy". Algunos ejemplos se pueden ver en [20], [14] y [8]. Más concretamente podemos encontrar un buen trabajo sobre filtros borrosos en [15] y más recientemente en [21] se presenta un trabajo sobre reducción de ruido. El diseño de filtros básicos y rápidos sigue siendo objeto de interés. En este trabajo estudiaremos la viabilidad del uso de filtros alternos basados en la apertura y cierre de la morfología matemática borrosa basada en uninormas para la eliminación y suavizado del ruido.

Este trabajo está organizado de la siguiente manera: en la sección siguiente se introducen los operadores morfológicos borrosos definidos a partir de uninormas conjuntivas continuas por la izquierda, junto con sus propiedades básicas. En la sección 3 se analizan el "Top hat" y el filtrado basados en esta morfología definida a partir de uninormas y se presentan algunos resultados experimentales junto con las interpretaciones de los mismos. Terminamos con la sección 4 dedicada a las conclusiones y trabajo futuro.

2. OPERADORES MORFOLÓGICOS BORROSOS

Supondremos conocidas las nociones básicas sobre uninormas usadas en este trabajo que, en todo caso, pueden hallarse en [9], [1] y [18]. Usaremos la notación siguiente: $\mathcal{I}$ será una implicación, $\mathcal{C}$ una conjunción, $\mathcal{N}$ una negación fuerte, U una uninorma conjuntiva con elemento neutro e, $\mathcal{I}_U$, su implicación residual, A una imagen con niveles de gris y B un elemento estructurante con niveles de gris.

Vamos a definir operadores morfológicos borrosos siguiendo las ideas de [7]. El método consiste en generalizar las operaciones lógicas (conjunción e implicaciones booleanas) a operadores borrosos. Consideramos que una imagen n-dimensional és una función $R^n \longrightarrow [0,1]$. Los valores de la imagen deben estar en $[0,1]$ para que ésta sea considerada un objeto borroso. A continuación pasamos a explicar dicho método a partir de las definiciones y proposiciones siguientes:

Definición. La dilatación borrosa $D_{\mathcal{C}}(A,B)$ y la erosión borrosa $E_{\mathcal{I}}(A,B)$ de la imagen A usando el elemento estructurante B son las imágenes siguientes:

$$D_{\mathcal{C}}(A,B)(y) = \sup_x \mathcal{C}(B(x-y), A(x)), \quad (1)$$

$$E_{\mathcal{I}}(A,B)(y) = \inf_x \mathcal{I}(B(x-y), A(x)). \quad (2)$$

Dada una imagen borrosa B, definimos la reflexión $-B$ como $-B(y) = B(-y)$, para toda $y \in R^n$. Así mismo, dadas dos imágenes B_1, B_2, diremos que $B_1 \subseteq B_2$ cuando $B_1(y) \leq B_2(y)$ para toda $y \in R^n$.

Definición. El cierre borroso $C_{\mathcal{C},\mathcal{I}}(A,B)$ y la apertura borrosa $O_{\mathcal{C},\mathcal{I}}(A,B)$ de la imagen A usando el elemento estructurante B son las imágenes siguientes:

$$C_{\mathcal{C},\mathcal{I}}(A,B)(y) = E_{\mathcal{I}}(D_{\mathcal{C}}(A,B), -B)(y), \quad (3)$$

$$O_{\mathcal{C},\mathcal{I}}(A,B)(y) = D_{\mathcal{C}}(E_{\mathcal{I}}(A,B), -B)(y). \quad (4)$$

En este trabajo utilizaremos como conjuntores dos tipos de uninormas conjuntivas, continuas por la izquierda y como implicadores, sus implicaciones residuales. Concretamente estos dos tipos de uninormas son:

- Las representables. Sea $e \in]0,1[$ y $h : [0,1] :\longrightarrow [-\infty, \infty]$ una función estrictamente creciente y continua, con $h(0) = -\infty$, $h(e) = 1$ y $h(1) = \infty$. Entonces,

  $U_h(x,y) =$

  $$\begin{cases} h^{-1}(h(x)+h(y)), & \text{si } (x,y) \notin \{(1,0),(0,1)\}, \\ 0, & \text{en caso contrario,} \end{cases}$$

  es una uninorma representable, conjuntiva con elemento neutro e (ver [9]) y su implicador residual $\mathcal{I}_{U_h}$ viene dado por

  $\mathcal{I}_{U_h}(x,y) =$

  $$\begin{cases} h^{-1}(h(x)-h(y)), & \text{si } (x,y) \notin \{(0,0),(1,1)\}, \\ 1, & \text{en caso contrario.} \end{cases}$$

  Además, U_h verifica autodualidad (excepto en los puntos (0,1) y (1,0)) respecto de la negación fuerte $\mathcal{N}(x) = h^{-1}(-h(x))$, (ver [1]).

- Un determinado tipo de idempotentes. Sea $\mathcal{N}$ una negación fuerte cualquiera. La función dada por:

$$U^{\mathcal{N}}(x,y) = \begin{cases} \text{mín}(x,y), & \text{si } y \leq \mathcal{N}(x), \\ \text{máx}(x,y), & \text{en caso contrario,} \end{cases}$$

es una uninorma idempotente, conjuntiva y continua por la izquierda. Su implicación residual viene dada por (ver [18])

$$\mathcal{I}_{U^{\mathcal{N}}}(x,y) = \begin{cases} \text{mín}(\mathcal{N}(x),y), & \text{si } y < (x), \\ \text{máx}(\mathcal{N}(x),y), & \text{si } y \geq (x). \end{cases}$$

Estos dos tipos de uninormas conjuntivas garantizan la mayoría de las "buenas" propiedades algebraicas y morfológicas de los operadores morfológicos derivados ([10], [1], [18]). Entre las cuales destacamos las siguientes (en todas las propiedades, U es una uninorma conjuntiva, continua por la izquierda, con elemento neutro $e \in]0,1[$, $\mathcal{I}_U$ su implicador residual, A una imagen de niveles de gris y B un elemento estructurante tipo imagen con niveles de gris)

- La dilatación borrosa D_U es creciente en ambos argumentos, la erosión borrosa E_U es creciente en el primer argumento y decreciente en el segundo, el cierre borroso $C_{U,\mathcal{I}_U}$ y la apertura borrosa $O_{U,\mathcal{I}_U}$ son ambos crecientes en el primer argumento.

- Si $B(0) = e$, se cumple que la dilatación es extensiva y la erosión antiextensiva. Es decir:

$$E_{\mathcal{I}_U}(A,B) \subseteq A \subseteq D_U(A,B).$$

- Así mismo, el cierre borroso es extensivo mientras que la apertura borrosa es anti-extensiva:

$$O_{U,\mathcal{I}_U} \subseteq A \subseteq C_{U,\mathcal{I}_U}(A,B).$$

- El cierre y la apertura borrosas son idempotentes; o sea,

$$\begin{aligned} C_{U,\mathcal{I}_U}(C_{U,\mathcal{I}_U}(A,B),B) &= C_{U,\mathcal{I}_U}(A,B), \\ O_{U,\mathcal{I}_U}(O_{U,\mathcal{I}_U}(A,B),B) &= O_{U,\mathcal{I}_U}(A,B). \end{aligned}$$

- Si $B(0) = e$, entonces

$$\begin{aligned} E_{\mathcal{I}_U}(A,B) &\subseteq O_{U,\mathcal{I}_U}(A,B) \subseteq A \\ &\subseteq C_{U,\mathcal{I}_U}(A,B) \subseteq D_U(A,B). \end{aligned}$$

- Si U es una uninorma de alguno de los dos tipos aquí considerados, se verifican además las propiedades de dualidad entre los operadores morfológicos: erosión y dilatación, apertura y cierre.

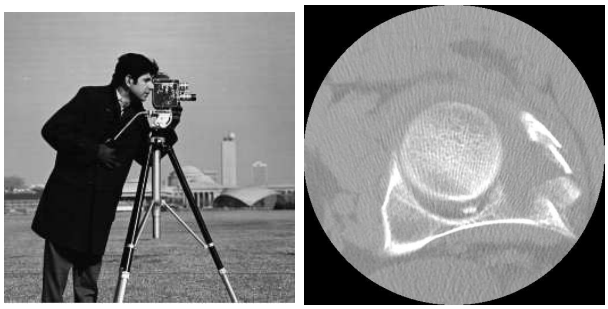


Figura 1: Imágenes originales usadas en experimentos.

3. RESIDUOS Y FILTROS BÁSICOS

En los experimentos que mostraremos, salvo que se indique explícitamente otra cosa, se utiliza un elemento estructurante de tamaño 3×3 (ver [11]) sobre la imagen considerada. La uninorma idempotente $U^{\mathcal{N}}$ con $\mathcal{N}(x) = 1 - x$ y la uninorma representable U_h con $h(x) = \ln\left(\frac{x}{1-x}\right)$. Se comparan los resultados con los obtenidos usando la t-norma de Łuckasiewicz y el desarrollo clásico basado en el *umbra approach* (ver [16]). En particular, el elemento estructurante usado por los operadores morfológicos borrosos vendrá representado por la matriz

$$B = e \cdot \begin{pmatrix} 0{,}86 & 0{,}86 & 0{,}86 \\ 0{,}86 & 1{,}00 & 0{,}86 \\ 0{,}86 & 0{,}86 & 0{,}86 \end{pmatrix}$$

donde e es el elemento neutro de la uninorma considerada. La figura 1 y la figura 6 muestran algunas de las imágenes originales (representadas por A en los operadores morfológicos borrosos) usadas en los experimentos. En la figura 1 se muestran dos imágenes en niveles de grises, de tamaño 256×256 la clásica imagen del cámara de cine y, de tamaño 512×512 la imagen medica de la cadera.

El residuo de dos operaciones o transformaciones morfológicas es su diferencia. El primer residuo que se puede definir en morfología matemática es el residuo gradiente morfológico, utilizado como un detector de contornos y como una primera aproximación a la segmentación morfológica. El residuo gradiente morfológico, conocido en morfología como gradiente de Beucher, es la diferencia entre una dilatación y una erosión, una dilatación y la imagen original o una diferencia entre la imagen original y su erosión. Cuando realizamos la dilatación y la erosión de un objeto borroso A por el elemento estructurante B considerando como conjuntor la uninorma U y como implicador su implicador

Figura 2: Izquierda: cierre. Derecha: apertura. Arriba, usando una uninorma idempotente. Abajo, usando una uninorma representable.

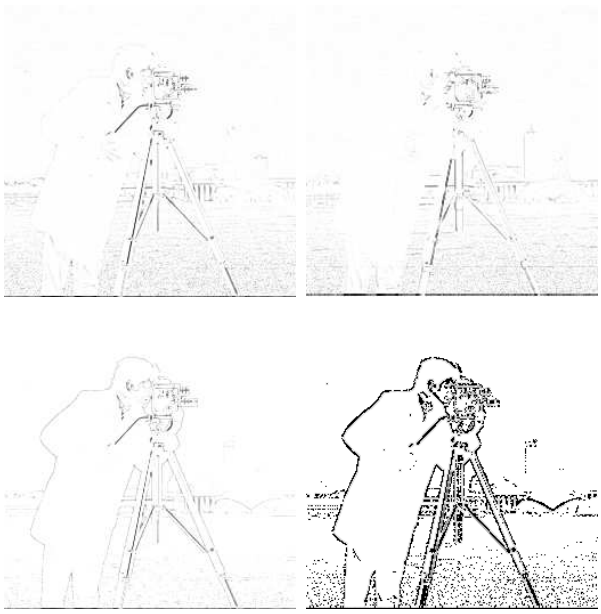


Figura 3: Izquierda: Top-Hat. Derecha: Top-Hat dual. Arriba, usando el *umbra approach*. Abajo, t-norma de Łuckasiewicz.

residual $\mathcal{I}_{\mathcal{U}}$, los gradientes vienen dados por:

$$\nabla^{-}_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = A \setminus E_{\mathcal{I}_{\mathcal{U}}}(A,B), \quad (5)$$

$$\nabla^{+}_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = D_U(A,B) \setminus A, \quad (6)$$

$$\nabla_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = D_U(A,B) \setminus E_{\mathcal{I}_{\mathcal{U}}}(A,B), \quad (7)$$

donde la expresión (5) representa el gradiente por erosión, (6) el gradiente por dilatación y (7) el gradiente simétrico.

El objetivo de los gradientes es la detección de bordes (contornos) y/o perímetro de los objetos presentes en la imagen. La elección correcta del elemento estructurante y de la variante del gradiente dependerá de la geometría de los objetos. En [11] se puede ver la aplicación del gradiente simétrico $\nabla_{U,\mathcal{I}_{\mathcal{U}}}$ a la detección de bordes, observando que intensifica la detección de los mismos, si comparamos con el gradiente derivado a partir de t-normas o el clásico. Desde el punto de vista de la inclusión borrosa, se satisface que $\nabla^{-}_{U,\mathcal{I}_{\mathcal{U}}}(A,B) \subseteq \nabla_{U,\mathcal{I}_{\mathcal{U}}}(A,B)$ y $\nabla_{U,\mathcal{I}_{\mathcal{U}}}(A,B) \supseteq \nabla^{+}_{U,\mathcal{I}_{\mathcal{U}}}(A,B)$.

La apertura y el cierre, tal cual los hemos definido en la sección anterior, son los filtros morfológicos más elementales, a los que llamaremos *filtros básicos*. Concretamente, como ya hemos indicado en la sección anterior, la apertura es un filtro morfológico antiextensivo y el cierre es un filtro morfológico extensivo. Así, en un primer paso se pueden usar para eliminar objetos no deseados presentes en una imagen. Se puede observar que la apertura de una imagen en niveles de gris por un elemento estructurante, elimina las formas claras de tamaño menor que el elemento, oscureciendo los objetos claros. El cierre morfológico ayuda a eliminar estructuras oscuras menores en tamaño al elemento estructurante, atenuando los objetos oscuros. El tamaño y la forma de los elementos estructurantes empleados en la apertura deben ser acordes con las estructuras de la imagen que se desean eliminar. En ocasiones, tamaños elevados de elementos estructurantes eliminarán formas indeseables en una imagen, pero también pueden eliminar otras formas y, desde luego, afectarán al resto de estructuras. Tamaños reducidos serán óptimos cuando las imágenes contengan pequeños detalles. En la figura 2 se puede observar el efecto del cierre y de la apertura sobre la imagen clásica del cámara de cine, usando la uninorma idempotente (en la parte superior) y la uninorma representable (en la parte inferior).

La apertura y el cierre son transformaciones morfológicas y son susceptibles de calcular los residuos asociados, que reciben el nombre de transformaciones *top-hat*. La transformación de top-hat, inicialmente presentada por F. Meyer en 1977 (ver [19]), consiste en descubrir aquellas estructuras de la imagen que han sido eliminadas en el filtrado por apertura o cierre. Con la elección de un elemento estructurante de forma, tamaño y orientación adecuado, es posible filtrar la imagen y eliminar determinados elementos en la imagen original. Una operación de diferencia entre el mapa original y el filtrado aumenta considerablemente el contraste de las zonas eliminadas.

Las transformaciones top-hat se definen como un re-

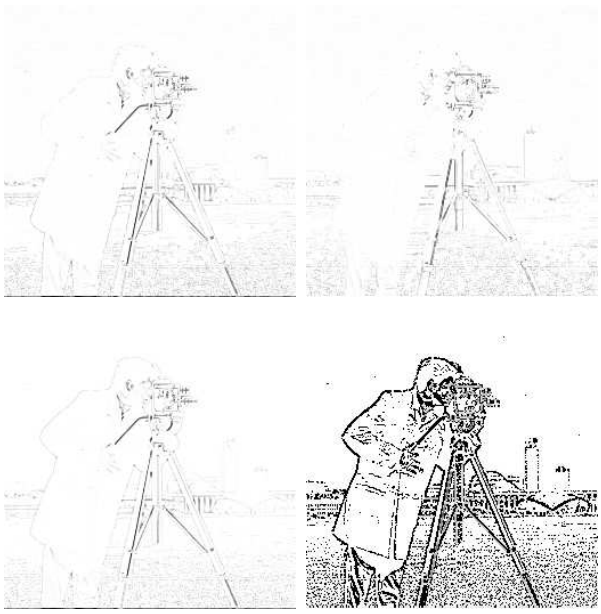


Figura 4: Izquierda: Top-Hat. Derecha: Top-Hat dual. Arriba, usando uninorma idempotente. Abajo, uninorma representable.

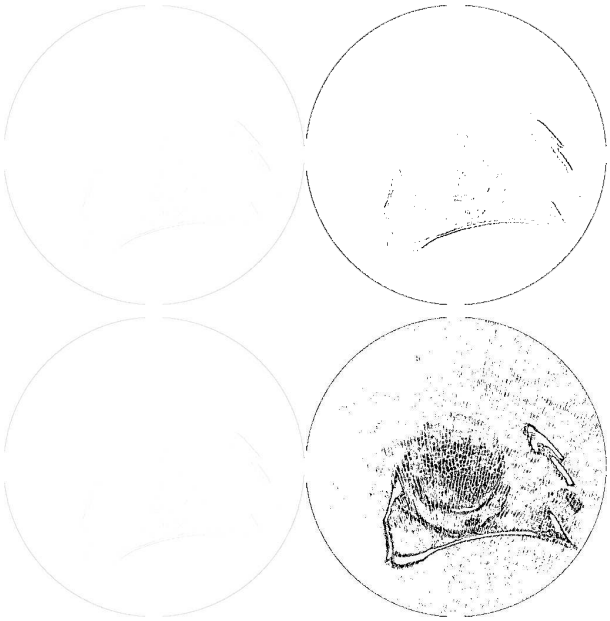


Figura 5: Izquierda: Top-Hat. Derecha: Top-Hat dual. Arriba, usando t-norma de Łuckasiewicz. Abajo, uninorma representable.

siduo entre la identidad y la apertura, recibiendo el nombre de top-hat (top-hat por apertura o top-hat blanco), o bien entre el cierre y la identidad, recibiendo el nombre de top-hat dual (top-hat por cierre o top-hat negro). Vienen determinadas por las expresiones:

$$\rho_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = A \setminus O_{U,\mathcal{I}_{\mathcal{U}}}(A,B) \text{ (top-hat)}$$
$$\rho^{d}_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = C_{U,\mathcal{I}_{\mathcal{U}}}(A,B) \setminus A \text{ (top-hat dual)}$$

El top-hat destaca los objetos claros que han sido eliminados por la apertura, se usa para extraer componentes contrastadas respecto del fondo, mientras que el top-hat dual extrae las componentes oscuras que han sido eliminadas por el cierre. Típicamente el top-hat elimina las tendencias suaves, realizando de este modo un realce de contraste. La detección de objetos brillantes en una imagen se logra con un top-hat por apertura, realizando lo mismo con un top-hat dual en el caso de los objetos oscuros. El tamaño del elemento estructurante ha de ser un factor determinante a la hora de realizar las transformaciones top-hat.

El top-hat es idempotente (pero no creciente) y antiextensiva. El top-hat dual no es creciente ni idempotente.

En las figuras 3 y 4 se pueden ver los contornos obtenidos por el top-hat y el top-hat dual sobre la imagen del cámara de cine. La información recogida en el caso de la uninorma idempotente es similar a la del caso clásico, y en el caso de la uninorma representable se obtiene (con el dual) más información que en el resto de casos. En la figura 5 se muestran ambos top-hats

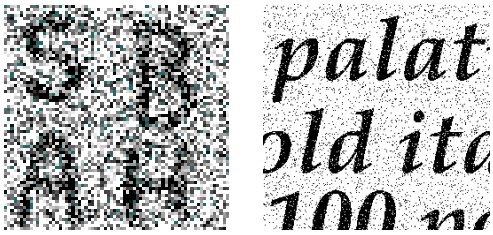


Figura 6: Imágenes originales con ruido.

para la imagen de la cadera, observando que la uninorma representable proporciona más información.

Hemos dicho un poco más arriba que la apertura y el cierre son los filtros básicos de la mofología matemática borrosa. Para la construcción de nuevos filtros morfológicos ([19]) se parte de los básicos y mediante composición o combinación con otras operaciones se implementan los filtros alternos, los filtros alternos secuenciales, el operador de centro, de contraste, etc. Las aplicaciones básicas de los filtros morfológicos son la atenuación del ruido y la extracción selectiva de objetos en la imagen. Ambas aplicaciones son importantes en visión artificial, pues técnicas de interpretación o medida fallarán en presencia de ruido, mientras que la selección de estructuras y objetos es fundamental en los procesos de segmentación.

Los filtros más inmediatos de construir a partir de la apertura y el cierre son los llamados *filtros alternos*. Supongamos que $\xi(A,B)$ y $\psi(A,B)$ representan la

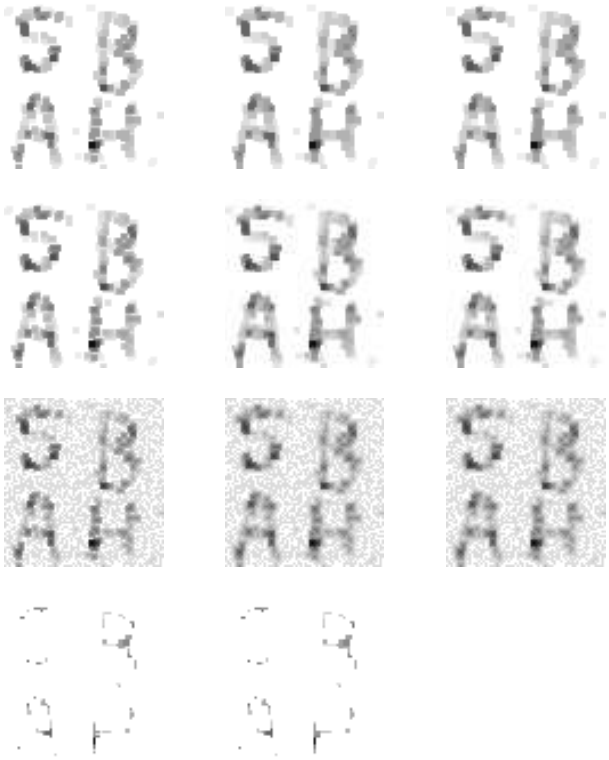


Figura 7: De izquierda a derecha: filtrado de la imagen por ξ, $\xi\psi$ y $\psi\xi\psi$, respectivamente. De arriba hacía abajo: uninorma idempotente, representable, t-norma de Łuckasiewicz y el *umbra approach*.

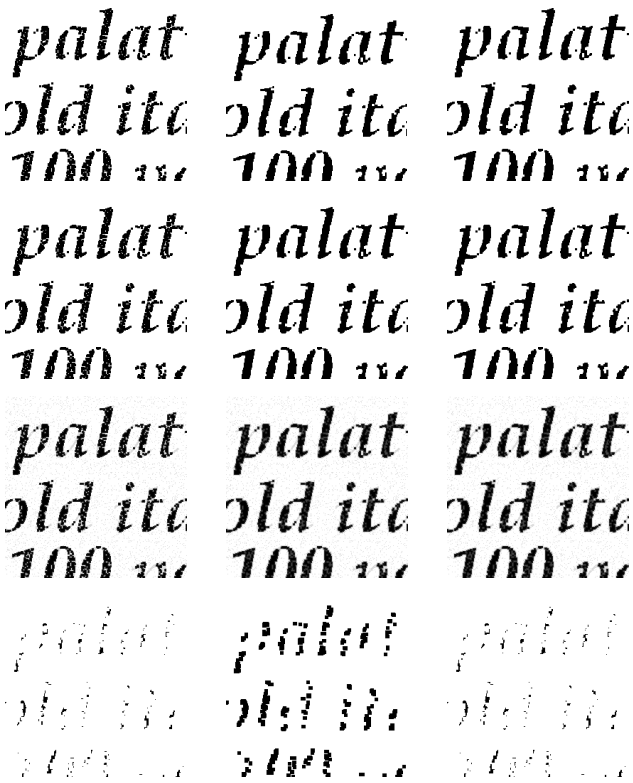


Figura 8: De izquierda a derecha: filtrado de la imagen por ξ, $\xi\psi$ y $\psi\xi\psi$, respectivamente. De arriba hacía abajo: uninorma idempotente, representable, t-norma de Łuckasiewicz y el *umbra approach*.

apertura y el cierre, respectivamente, de un objeto borroso A por el elemento estructurante B considerando como conjuntor la uninorma U y como implicador su implicador residual $\mathcal{I}_{\mathcal{U}}$. Por combinación de estos dos filtros se pueden generar cuatro filtros idempotentes y crecientes: $\xi\psi$, $\psi\xi$, $\psi\xi\psi$ y $\xi\psi\xi$, teniendo propiedades similares a los filtros morfológicos clásicos. Debido a la propiedad de idempotencia, la composición de más de tres operadores no va a proporcionar un nuevo filtro, salvo que se cambie el elemento estructurante.

En la figura 6 podemos ver dos imágenes binarias a las que se les ha añadido ruido. En la imagen de la izquierda, de tamaño 70×74, el 60 % de los puntos ha sido sustituido por ruido aleatorio en niveles de gris[1]. En la imagen de la derecha, de tamaño 256×256, se ha añadido ruido gaussiano aditivo. La figura 7 muestra los resultados obtenidos al aplicar, de izquierda a derecha los filtros alternos ψ, $\xi\psi$ y $\psi\xi\psi$, respectivamente. De arriba hacía abajo se ven los resultados obtenidos usando los diferentes enfoques. Podemos observar cómo en el caso de las uninormas se elimina el ruido de la imagen de tal forma que la imagen *original* es discernible. La misma estructura, se puede observar en la figura 8.

Figura 9: De izquierda a derecha filtrado de la imagen por $\psi\xi$, una uninorma idempotente, la t-norma de Łuckasiewicz y el *umbra approach*.

En la figura 9 se puede observar el filtrado de la imagen usando el filtro $\psi\xi$ y considerando la uninorma idempotente (muy similar al obtenido con la uninorma representable), la t-norma de Łuckasiewicz y el *umbra approach*. El filtrado obtenido con la estructura $\psi\xi\psi$, es muy similar, salvo en el caso clásico en que la imagen queda más emborronada y en el de la t-norma un poco más suavizado el ruido.

[1]Esta imagen ha sido cogida de la página web de J. A. Sethian, http://math.berkeley.edu/~sethian/, de la sección de eliminación de ruido de la parte de aplicaciones.

Figura 10: Izquierda: Imagen original. Derecha: imagen con ruido añadido tipo "sal y pimienta" .

En la figura 10 podemos ver una imagen de tamaño 308×242 a la que se le ha añadido ruido "sal y pimienta" de parámetro 0.02. En la figura 11 podemos ver los resultados obtenidos usando un filtro tipo $\psi\,\xi\,\psi$. En la parte superior izquierda se ha usado la t-norma de Łuckasiewicz y se puede ver que permanece gran parte del ruido. A su derecha se presenta el filtrado con el *umbra approach*, se puede ver que permanece algo del ruido tipo *sal*, sobre todo en el interior de la monedas y sobre sus contornos. En la parte inferior se observan los resultados usando uninormas. A la izquierda con la uninorma idempotente y, a la derecha con una uninorma representable de generador $h(x) = \frac{x-0,5}{x(1-x)}$. En ambos casos se mejoran los resultados y desaparece la práctica totalidad del ruido.

Es posible realizar una extensión de la transformación básica de top-hat construyendo un residuo entre la imagen original y un filtro alterno de apertura y cierre. Por ejemplo, para la eliminación de ruido denso y simétrico, una simple apertura no servirá para reducir los pixels oscuros, lo mismo ocurrirá con un cierre y los pixels de ruido claros. La composición del filtro $\xi\,\psi$ podrá eliminar el ruido de cualquier tonalidad, con lo que el top-hat resultante será el residuo:

$$\rho^{e}_{U,\mathcal{I}_{\mathcal{U}}}(A,B) = A \setminus \xi\left(\psi(A,B),B\right).$$

Los filtros creados por diferencia se han utilizado para el control industrial de piezas con oclusiones parciales e incluso para la eliminación de ruido.

4. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo hemos visto cómo los filtros alternos, basados en los operadores morfológicos de apertura y cierre usando uninormas, pueden eliminar el ruido de una imagen conservando las estructuras. Mejorando los resultados enfrente del uso de los operadores basados en t-normas o el *umbra approach*, y obteniendo imágenes muy similares a las imágenes originales sin ruido. Podemos decir, que los filtros alternos basados en uninormas pueden ser adecuados para la eliminación de ruido. La construcción de nuevos filtros será objeto de un trabajo futuro. También hemos mostrado el uso de la apertura y el cierre en la detección de bordes. Falta estudiar el efecto de la elección del elemento estructurante en las aplicaciones mostradas. En los experimentos mostrados hemos usado elementos estructurante isotrópicos; eso es, tienen el mismo efecto en todas las direcciones dónde se aplican. Creemos que en determinadas imágenes, el uso de elementos estructurantes anisotrópicos, mejoraría bastante los resultados obtenidos.

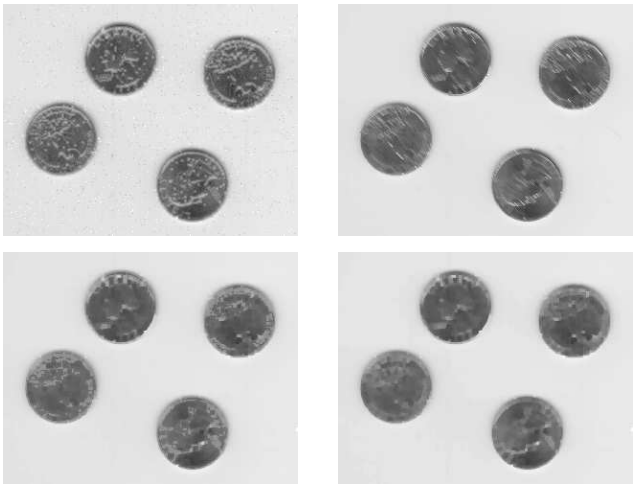


Figura 11: De izquierda a derecha y de arriba hacía abajo filtrado de la imagen por $\psi\,\xi\,\psi$, usando la t-norma de Łuckasiewicz, el *umbra approach*, una uninorma idempotente y una uninorma representable.

Otro estudio a realizar de mayor envergadura es aplicar las técnicas mostradas a un conjunto más ámplio de imágenes con niveles de gris e, incluso, imágenes en color.

Agradecimientos

Este trabajo ha sido subvencionado por el proyecto del Gobierno Español MTM2006-05540.

Referencias

[1] B. De Baets and J.C. Fodor. Residual operators of uninorms. *Soft Computing*, 3:89–100, 1999.

[2] B. De Baets, E.E. Kerre, and M. Gupta. The fundamentals of fuzzy mathematical morfologies part i: basics concepts. *International Journal of General Systems*, 23:155–171, 1995.

[3] B. De Baets, E.E. Kerre, and M. Gupta. The fundamentals of fuzzy mathematical morfologies part ii: idempotence, convexity and decomposi-

tion. *International Journal of General Systems*, 23:307–322, 1995.

[4] B. De Baets, N. Kwasnikowska, and E. Kerre. Fuzzy morphology based on uninorms. In *Proceedings of the seventh IFSA World Congress*, pages 215–220, Prague, 1997.

[5] I. Bloch and H. Maître. Fuzzy mathematical morphologies: a comparative study. *Pattern Recognition*, 28:1341–1387, 1995.

[6] T. Chen, Q. H. Wu, R. Rahmani-Torkaman, and J. Hughes. A pseudo top-hat mathematical morphological approach to edge detection in dark regions. *Pattern Recognition*, 35(1):199–210, 2002.

[7] B. De Baets. Fuzzy morphology: A logical approach. In B. M. Ayyub and M. M. Gupta, editors, *Uncertainty Analysis in Engineering and Science: Fuzzy Logic, Statistics, and Neural Network Approach*, pages 53–68. Kluwer Academic Publishers, Norwell, 1997.

[8] F. Farbiz and M. B. Menhaj. A fuzzy logic control based approach for image filtering. In Etienne E. Kerre and Mike Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 7, pages 194–221. Physica-Verlag, New York, 2000.

[9] J.C. Fodor, R.R. Yager, and A. Rybalov. Structure of uninorms. *Int. J. Uncertainty, Fuzziness, Knowledge-Based Systems*, 5:411–427, 1997.

[10] M. González, D. Ruiz-Aguilera, and J. Torrens. Algebraic properties of fuzzy morphological operators based on uninorms. In *Artificial Intelligence Research and Development*, volume 100 of *Frontiers in Artificial Intelligence and Applications*, pages 27–38, Amsterdam, 2003. IOS Press.

[11] M. González-Hidalgo, A. Mir-Torres, D. Ruiz-Aguilera, and J. Torrens. Fuzzy morphology based on uninorms: Image edge-detection. opening and closing. In João Tavares and Natal Jorge, editors, *Computational Vision and Medical Image Processing*, pages 127–133, London, 2008. Taylor & Francis Group.

[12] S. Jamil Ferzoli Guimarães, N. Jerônimo Leite, M. Couprie, and A. de Albuquerque Araújo. A directional and parametrized transition detection algorithm based on morphological residues. In *Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI'02*, pages 261–268, Washington, DC, USA, 2002. IEEE Computer Society.

[13] J.-A. Jiang, C.-L. Chuang, Y.-L. Lu, and C.-S-Fahn. Mathematical-morphology-based edge detectors for detection of thin edges in low-contrast regions. *IET Image Processing*, 1(3):269–277, 2007.

[14] Chang-Shing Lee and Yau-Hwang Kuo. Adaptative fuzzy filter and its application to image enhancement. In Etienne E. Kerre and Mike Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 6, pages 172–193. Physica-Verlag, New York, 2000.

[15] M.Ñachtegael, D. Van der Weken, D. Van De Ville, and E.E. Kerre, editors. *Fuzzy filters for image processing*, volume 122 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag, 2003.

[16] M.Ñachtegael and E.E. Kerre. Classical and fuzzy approaches towards mathematical morphology. In Etienne E. Kerre and Mike Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 1, pages 3–57. Physica-Verlag, New York, 2000.

[17] J. I. Pastore, E. Moler, and V. Ballarin. Multiscale morphological operators and geodesic distance applied to computer axial tomography segmentation. *IEEE Latin America Transactions*, 5(1):28–31, 2007.

[18] D. Ruiz-Aguilera and J. Torrens. Residual implications and co-implications from idempotent uninorms. *Kybernetika*, 40:21–38, 2004.

[19] J. Serra. *Image analysis and mathematical morphology, vols. 1, 2.* Academic Press, London, 1982,1988.

[20] Hamid R. Tizhoosh. Fuzzy image enhancement: An overview. In Etienne E. Kerre and Mike Nachtegael, editors, *Fuzzy techniques in image processing*, number 52 in Studies in Fuzziness and Soft Computing, chapter 5, pages 137–171. Physica-Verlag, New York, 2000.

[21] D. Van De Ville, M.Ñachtegael, D. Van der Weken, E.E. Kerre, and W. Philips. Noise reduction by fuzzy images filtering. *IEEE Transactions on Fuzzy Systems*, 11(4):429–436, 2003.

[22] Zhao Yu-qian, Gui Wei-hua, Chen Zhen-cheng, Tang Jing-tian, and Li Ling-yun. Medical images edge detection based on mathematical morphology. In *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 6492–6495, Shanghai, China, 2005.

ESTRUCTURA DE LAS SIMILARIDADES

Ramón Gonzalez del Campo [1] **Luis Garmendia** [2] **Jordi Recasens** [3]

[1] SIC. Facultad de Informática, rgonzale@estad.ucm.es
[2] DISIA. Facultad de Informática. UCM, lgarmend@fdi.ucm.es
[3] Universitat Politècnica de Catalunya. j.recasens@upc.edu

Resumen

En este artículo se define formalmente el concepto de estructura de similaridad, se cuenta el número de estructuras de similaridades hasta dimensión 5, se propone una nomenclatura y un algoritmo que asigna una estructura a cada similaridad.

Palabras Clave: Relación borrosa, Min-transitividad, Similaridad.

1 INTRODUCCIÓN

Las relaciones borrosas de similaridad permiten dar grados de parecido entre los elementos de un universo y encuentran numerosas aplicaciones en métodos de clustering y de inferencia borrosa. Sus alfa-cortes son relaciones de equivalencia clásicas, por lo que ofrecen numerosas "rejillas" o particiones de un universo. Por otro lado son Min-preórdenes [1], por lo que proporcionan conclusiones según Tarski cuando se utilizan para hacer inferencia mediante la regla composicional de inferencia.
En el estudio de estas importantes relaciones, nos damos cuanta de que muchas de ellas se parecen mucho y otras no. Nos parece muy definir lo que vamos a llamar "estructura" de similaridad, que nos ayude a entenderlas y a clasificarlas según a qué estructura pertenezcan.
Este artículo comienza definiendo el concepto de estructura como clases de equivalencia de una relación de equivalencia RE, que relaciona dos similaridades si tienen la misma estructura. Para definir RE definimos otras dos relaciones de equivalencia, Rπ y RT, que relaciona dos similaridades si se pueden transformar una en otra mediante una permutación y mediante una relación monótona estricta.
Hemos estudiado y contado las estructuras de las similaridades hasta la dimensión 5

2 PRELIMINARES

Sea $E=\{e_1,...,e_n\}$ un conjunto finito. Una relación difusa es una transformación $R:E \times E \rightarrow [0,1]$. El grado de relación para los elementos e_i y e_j, $R(e_i,e_j)$, se denota como r_{ij}.

Definición 2.1: Un operador *binario* $T:[0,1] \times [0,1] \rightarrow [0,1]$ es una t-norma o norma triangular [4,2] si satisface:

- $T(1,x)=x$
- $T(y,x)=T(x,y)$
- $T(x,T(y,z))=T(T(x,y),z)$
- Si $x \leq x'$ $y \leq y'$ entonces $T(x,y) \leq T(x',y')$

Definición 2.2: Dada una t-norma T, una relación difusa $R:E \times E \rightarrow [0,1]$ es T-transitiva si $T(R(a,b),R(a,c)) \leq R(a,c)$ $\forall\ a,b,c \in E$. Esto es, si $T(r_{ik},r_{kj}) \leq r_{ij}$ $\forall$ $i,j,k:1..n$.

Definición 2.3: Dada una relación difusa, R, es una Proximidad si safisface:

- Reflexiva: $r_{ii}=1$
- Simétrica: $r_{ij}=r_{ji}$ $\forall\ 1 \leq i, j \leq n$

Si además R es una relación T-transitiva, se denomina T-indistinguibilidad [6,7]. Si la T-norma es el mínimo, a esta T-indistinguibilidad se le llama Similaridad [8].

3 RELACIONES ENTRE RELACIONES BORROSAS

Definición 3.1: Sean A y B dos matrices de dimensión n, están relacionadas por R_T si y solo si $\exists$ T tal que $T: A \rightarrow B$, $T(a_{ij})=b_{ij}$ tal que es monótona estricta.

Lema 3.1: R_T es una relación de equivalencia.
Demostración:

- Es reflexiva con T igual a la función identidad.
- Es simétrica: si $A\ R_T\ B$ entonces $B\ R_T^{-1} A$.
- Es transitiva: si $A\ R_T B$ y $B\ R_{T'}\ C$ entonces: $A\ R_{T' \cdot T}\ C$.

El número de clases de equivalencia, S^n/R_T están contados en [5].

Lema 3.2: Dada una matriz de una similaridad A, si permutamos las filas y las correspondientes columnas, denotado por $B=P_\pi(A)$, obtenemos una matriz que representa a una similaridad.

Demostración:

Una permutación, $\pi=\{\pi_i\}\ \forall\ i:1..n$, consiste en reetiquetar los índices de la siguiente forma: $i \rightarrow \pi_i$.

- Es reflexiva: $b_{ii}=a_{\pi i\ \pi i}=1$
- Es simétrica: $b_{ij}=a_{\pi i\ \pi j}=a_{\pi j\ \pi i}=b_{ji}$
- Es min-transitiva: $min(b_{ik},b_{kj}\leq b_{ij})$ ya que $min(a_{\pi i\pi k},a_{\pi k\ \pi j})\leq a_{\pi i\ \pi j}$

Definición 3.2: Sean A y B dos matrices de dimensión n, están relacionadas por R_π si y solo si existe una permutación π tal que $B=P_\pi(A)$.

Lema 3.3: R_π es una relación de equivalencia.

Demostración:

- Es reflexiva: A R_π A pues $A=P_\pi(A)$ con $\pi=(1..n)$.
- Es simétrica: A R_πB $\leftrightarrow$ B=P_π(A) $\leftrightarrow$ P_π (B)=P_π (P_π(A)) puesto que: A=P_π(P_π(A)).
- Es transitiva: A Rπ B, B Rπ C $\rightarrow$ A=P_π(B),B=P_π(C) $\rightarrow$ A=P_π ($P_{\pi'}$(C)) $\rightarrow$ A=$P_{\pi\cdot\pi'}$(C) .

La siguiente relación de equivalencia indica si dos similaridades tienen la misma "estructura".

Lema 3.4: R_E es una relación de equivalencia.

Demostración:

- Es reflexiva: AR_EA tomando $\pi=(1..n)$ ya que $AR_\pi A$ y $A\ R_{Id}\ A$.
- Es simétrica por ser R_T y R_π simétricas.
- Es transitiva por ser R_T y R_π transitivas.

Definición 3.4: Sea S^n el conjunto de similaridades de dimensión n. Cada elemento de partes de $R_E(S^n/R_E)$ es una estructura de similaridad de dimensión n.

4 CONTEO DE ESTRUCTURAS DE SIMILARIDAD

Para investigar el número de estructuras de similaridad nos basamos en el siguiente lema:

Lema 4.1(Teorema de Lee [3]): Sean C y D dos relaciones borrosas t y:

$$E(t;C,D)=\begin{pmatrix} C & F^T \\ F & D \end{pmatrix}$$

Si C y D son similaridades entonces $E(t;C,D)=E$ también es una similaridad $\forall t \in [0,min(min(C),min(D))]$.

Para cada matriz, S^n, que representa una similaridad de dimensión n existe una descomposición:

$$S^n=P_\pi(E(t;C_{n1xn1},D_{n2xn2}))$$

Para todas las similaridades de dimensión n investigamos el número de estructuras diferentes a partir de las descomposiciones posibles: $E(t;C,D)$.

Casos:

n=2:

Solamente existe una estructura:

$$\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$$

n=3:

Las descomposiciones con ($n_1=1,n_2=2$) y ($n_1=2,n_2=1$) representan a la misma estructura bajo la permutación $\pi=(3,2,1)$.

$$E(t;C,D)=\begin{pmatrix} \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} & t \\ t & 1 \end{pmatrix}$$

donde $a \geq t$. Por lo tanto, existen 2 estructuras:

$$\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix} \text{ y } \begin{pmatrix} 1 & a & t \\ a & 1 & t \\ t & t & 1 \end{pmatrix}$$

donde $a > t$.

n=4:

- Las descomposiciones con $(n_1=1,n_2=3)$ y $(n_1=3,n_2=1)$ representan la misma estructura bajo la permutación $\pi=(4,2,3,1)$. Las similaridades de dimensión 4 son de dos posibles formas:

- Descomposición con $(n_1=3,n_2=1)$:

$$\begin{pmatrix} \begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix} & t \\ t\ \ t\ \ t & 1 \end{pmatrix}$$

con $a \geq b \geq t$. Sustituyendo cada desigualdad por una desigualdad estricta o una igualdad obtenemos 4 posibles estructuras.

- Descomposición con $(n_1=2,n_2=2)$:

$$\begin{pmatrix} \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} & \begin{matrix} t & t \\ t & t \end{matrix} \\ \begin{matrix} t & t \\ t & t \end{matrix} & \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix} \end{pmatrix}$$

donde $a \geq t$ y $b \geq t$ ($a \in C$ y $b \in D$).

Los casos con $b=t$ ya se han tenido en cuenta en la descomposición anterior. Tenemos 2 estructuras adicionales.

- En total, para $n=4$ existen 6 estructuras.

n=5:
Las descomposiciones con $(n_1=1,n_2=4)$ y $(n_1=4,n_2=1)$ representan la misma estructura bajo la permutación $\pi=(5,2,3,4,1)$.
Las descomposiciones con $(n_1=2,n_2=3)$: sean $a \in C$ y $b, c \in D$. Entonces se verifica:
Combinaciones posibles de valores de a,b,c:

- $a \geq b \geq c \geq t \rightarrow 2^3$ casos distintos
- $b \geq a \geq c \geq t \rightarrow 2^2$ casos distintos (porque $a=b$ ya se ha tratado)
- $b \geq c \geq a \geq t \rightarrow 2$ casos distintos adicionales (porque $a=c$ ya se ha tratado y si $a=t$ obtenemos una estructura ya tratada en la descomposición anterior).

- En total, tenemos 22 estructuras distintas para $n=5$.

4.1 HACIA UNA NOMENCLATURA DE LAS ESTRUCTURAS

Sea S cualquier similaridad de dimensión n.
Consideremos la descomposición:

$$E(t;C,D)= \begin{pmatrix} C & F^T \\ F & D \end{pmatrix}$$

Sea z el número de elementos de F y F^T. Se verifica:

$$n^2=n_1^2+ n_2^2+2n_1xn_2$$
$$z=2n_1xn_2$$

Sea $I=\{1..n_1+n_2-1\}$.
Denotamos cada estructura mediante la expresión: $S^{(n,z)}_{\{qi\}\{pi\}}$ donde:
- n es la dimensión, z es el número de elementos de las submatrices F y F^T.
- q_i es el número de veces que aparece el valor i. Para los elementos de C se cumple: $\sum_{i=1} q_i \leq n_1^2-n_1$. El valor del puente, q_t, es el primer valor para el que se verifica $\sum_{i=1} q_i > n_1^2-n_1$. Para los elementos de D se cumple:

$$\sum_{i=qt+1} q_i \leq n_2^2-n_2$$

Sea $\{p_i\}$ la lista de los valores distintos de S, además: $p_j \neq 1$, tal que si $k<j$ la magnitud del valor k es mayor que la magnitud del valor j. Si $\{p_i\}=\{1..k\}$ con $k \leq n_1+n_2-1$ la omitimos.

Ejemplos:

- $\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$ se denota por $S^{(2,2)}$.
- $\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$ y $\begin{pmatrix} 1 & a & t \\ a & 1 & t \\ t & t & 1 \end{pmatrix}$ se denotan por $S^{(3,4)}_{\{6\}}$ y $S^{(3,4)}_{\{2,4\}}$ respectivamente.
- $\begin{pmatrix} 1 & a & b & t \\ a & 1 & b & t \\ b & b & 1 & t \\ t & t & t & 1 \end{pmatrix}$ con $a>b>t$ se denota $S^{(4,8)}_{\{2,4,8\}}$.

Si $a>b=t$ entonces: $S^{(4,8)}_{\{2,12\}}$

- $\begin{pmatrix} \begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix} & t \\ t & \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} \end{pmatrix}$ con $a>c>b=t$ se denota $S^{(5,12)}_{\{2,16,2\}\{1,3,2\}}$.

Este es el ejemplo más pequeño en el que podemos apreciar que la relación entre los valores de las similaridades C y D es relevante. En la matriz anterior si $c>a$, por ejemplo, no existe una función T monótona estricta que con la misma matriz en la que se dé $c<a$. Son estructuras diferentes: $S^{(5,12)}_{\{2,16,2\}\{3,1,2\}}$ y $S^{(5,12)}_{\{2,16,2\}\{1,3,2\}}$.

4.2 ALGORITMO DE CLASIFICACIÓN DE SIMILARIDADES

Podemos clasificar fácilmente cualquier similaridad de dimensión n, S, en alguna estructura de dimensión n con la que esté R_E relacionada, mediante el siguiente algoritmo.

Pasos:

- Sea $_{bi0\,j0}=t$ el mínimo valor.
- Sean $I = \{i \text{ tal que } i \in N \text{ y } b_{ij0}=t\}$, y $I' = N \setminus I$.
- Sea el vector π tal que: $\pi=I \cup I'$.
- Para cada $p,q \in \{1..n\}$ hacer: $a_{pq}=b_{\pi[p]\,\pi[q]}$.
- $z=2card(I)x\,card(I')$
- Sea $\{c_i\}$ la lista de las magnitudes de los valores de C y $\{d_i\}$ la lista de valores de D, ambos ordenados decrecientemente.
- Sea $\{a_j\}$ la lista de valores de S ordenados decrecientemente por su magnitud.
- Sea la lista $\{nc_i\}$ tal que nc_i es el número de veces que aparece el valor i en C y la lista $\{nd_i\}$ tal que nd_i es el número de veces que aparece el valor j en D.
- Para cada $p,q \in I$ hacer: si $a_{pq}=c_i$ incrementa nc_i.
- Para cada $p,q \in I'$ hacer: si $a_{pq}=d_i$ incrementa nd_i.
- Pertenece a $S^{n,z}_{\{\{nci\},z,\{ndi\}\}\,\{ai\}}$

Ejemplo:
Sea la similaridad, S, con $n=4$:

$$\begin{pmatrix} 1 & 0,4 & 0,7 & 0,4 \\ 0,4 & 1 & 0,4 & 0,6 \\ 0,7 & 0,4 & 1 & 0,4 \\ 0,4 & 0,6 & 0,4 & 1 \end{pmatrix}$$

Pasos:
- El mínimo valor es: $b_{12}=0,4$.
- $I= \{1,3\}$, $I'= \{2,4\}$.
- $\pi= \{1,3,2,4\}$.
- Obtenemos:

$$\begin{pmatrix} 1 & 0,7 & 0,4 & 0,4 \\ 0,7 & 1 & 0,4 & 0,4 \\ 0,4 & 0,4 & 1 & 0,6 \\ 0,4 & 0,4 & 0,6 & 1 \end{pmatrix}$$

Esta similaridad se clasifica en $S^{4,8}_{\{2,8,2\}\{1,2\}}$.

5 ÁRBOLES JERÁRQUICOS Y ESTRUCTURAS

Una de las razones del interés en el estudio de las similaridades es que generan árboles de particiones indexados. En efecto, para cada entrada de a de la similaridad S definida en el universo X, la relación $\approx$ crisp: $x_i \approx x_j$ si y sólo si $S(x_i,x_j) \geq a$ es una partición P_a de X y si $a \geq b$, entonces P_a es un refinamiento de P_b.

Definición 5.1:
Un árbol es un grafo sin ciclos.
Los nodos con una sola arista se llaman hojas.
La distancia entre dos nodos es el número de aristas que los separan.

Definición 5.2:
Un árbol jerárquico es un árbol con un nodo privilegiado A tal que:
- La distancia d entre A y las hojas del árbol es constante.
- El cardinal de los nodos a distancia k de A es menor que el cardinal de los nodos a distancia $k+1$ de A para todo $k=1,...,d$.

Es inmediato comprobar los siguientes resultados.

Proposición 5.3:
Dos relaciones de similaridad generan el mismo árbol (no indexado) si y sólo si son equivalentes por R_T.

Proposición 5.4:
Dos relaciones de similaridad generan el mismo árbol indexado salvo permutaciones de sus ramas si y sólo si son equivalentes por R_π.

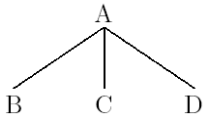
Proposición 5.5:
Dos relaciones de similaridad generan el mismo árbol (no indexado) salvo permutaciones de sus ramas si y sólo si son equivalentes por R_E.

La búsqueda de estructuras de similaridades es pues equivalente a hallar todos los árboles jerárquicos salvo permutaciones de sus ramas.

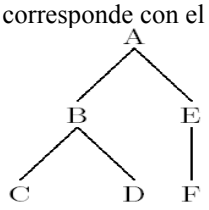
En la representación de las estructuras mediante árboles jerárquicos:
- cada nodo representa una similaridad,
- cada nodo hoja representa un objeto del universo X.
- las aristas que unen nodos de dos niveles consecutivos representan el valor de la relación de los objetos y/o similaridades.

Ejemplos:

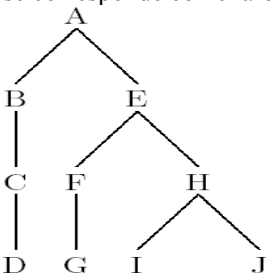
- La similaridad $S^{(3,4)}_{\{6\}}$: $\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$ se corresponde con el árbol:



- La similaridad $S^{(3,4)}_{\{2,4\}}$: $\begin{pmatrix} 1 & a & b \\ a & 1 & b \\ b & b & 1 \end{pmatrix}$ se corresponde con el árbol:



- La similaridad $S^{(4,8)}_{\{2,4,8\}}$: $\begin{pmatrix} 1 & a & b & t \\ a & 1 & b & t \\ b & b & 1 & t \\ t & t & t & 1 \end{pmatrix}$ se corresponde con el árbol:



6 CONCLUSIONES Y TRABAJO FUTURO

Una vez hemos definido formalmente el concepto de estructura, hemos encontrado una notación adecuada para representar cada estructura y hemos encontrado una algoritmo que decide a que

estructura pertenece una similaridad, permitiendo una clasificación de las similaridades según su estructura y un mejor entendimiento de las mismas que podría ser considerado para estudiar el comportamiento de las similaridades en distintas aplicaciones según su forma.

Seguimos buscando una fórmula, seguramente recursiva, que nos cuente el número de estructuras de similaridades de dimensión mayor a 5, y un algoritmo que genere un representante de dichas estructuras para cualquier dimensión, que podría acercarnos una manera de representación, a una mejor comprensión y como generador universal de todas las similaridades finitas.

Referencias

[1] J. Elorza y P. Burillo. On the relation between fuzzy preorders and fuzzy consequence operators. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 7(3):219{234, 1999}.

[2] J. Jacas y J. Recasens. Fuzzy t-transitive relations: eigenvectors and generators. Fuzzy Sets Syst., 72(2):147{154, 1995}.

[3] Hsuan-Shih Lee. An optimal algorithm for computing the max-min transitive closure of a fuzzy similarity matrix. Fuzzy Sets and Systems, 123(1):129{136, 2001}.

[4] Sklar A. Schweizer, B. Probabilistic Metric Spaces. Dover Publications, 1984.

[5] Riera T. How similarity matrices are? Stochastica, 2(4), 1978.

[6] Valverde L. Trillas, E. An inquiry into indistinguishability operators, in Aspects of Vagueness. Reidel Pubs, 1984.

[7] L. Valverde. On the structure of f-indistinguishability operators. Fuzzy Sets and Systems 17, pages 313{328, 1985}.

[8] L. A. Zadeh. Similarity relations and fuzzy orderings. Inform. Sci. 3, pages 177{200, 1971}.

Aproximación de Proximidades por Similitudes

L. Garmendia
Ingeniería del Software e Inteligencia Artificial
Facultad de Informática
Universidad Complutense de Madrid
C/ Profesor José García Santesmases, s/n
28040 Madrid
lgarmend@fdi.ucm.es

J. Recasens
Secció Matemàtiques i Informàtica
ETS Arquitectura del Vallès
Universitat Politècnica de Catalunya
Pere Serra 1-15
08190 Sant Cugat del Vallès
j.recasens@upc.edu

Resumen

En este trabajo se presenta un algoritmo para hallar una relación de similitud próxima a una relación de proximidad (i.e.: una relación de tolerancia) R dada. La relación de similitud obtenida es más próxima a R que su clausura transitiva o cualquier opening transitivo de R.

Palabras Clave: Relación de Proximidad, Relación de Tolerancia, Relación de Similitud, Clausura Transitiva, Opening Transitivo.

1 Introducción

Las relaciones de similitud son operadores de T-indistinguibilidad respecto a la t-norma Mínimo [9]. Son relaciones borrosas interesantes con diversas aplicaciones en Técnicas de Clasificación y Cluster analysis porque sus α-cortes son particiones del universo de discurso y al aumentar α las correspondientes particiones refinan a las anteriores. Dicho de otro modo, las relaciones de similitud generan árboles jerárquicos indexados. Recíprocamente, a partir de un árbol jerárquico indexado se puede generar una relación de similitud.

En muchas situaciones, los datos de un problema se dan en forma de una relación de proximidad en el universo de discurso. Se parte entonces de una relación reflexiva y simétrica. Si se necesita clasificar los objetos del universo mediante un árbol jerárquico, esta relación debe ser reemplazada por una relación de similitud lo más cercana a la original como sea posible.

El método más común de hallar una relación de similitud cercana a una relación de proximidad R dada es calcular su clausura transitiva, que es una relación mayor o igual que R, o buscar un opening transitivo de R, es decir, una relación de similitud maximal entre las que son menores o iguales que R.

Este trabajo presenta un algoritmo de simple ejecución para hallar una relación de similitud a partir de una relación de proximidad R dada que es más cercana a R que su clausura transitiva o cualquiera de sus openings transitivos. En contraste con la clausura transitiva y los openings transitivos, en general algunas entradas de la similitud obtenida son mayores y otras menores que las correspondientes entradas de la relación de proximidad.

En [3] se presentaron un par de métodos parra hallar buenas aproximaciones de una relación de proximidad por operadores de T-indistinguibilidad cuando T es una t-norma arquimediana continua. Estos métodos no funcionan para la t-norma Mínimo (i.e.: relaciones de similitud). Con los resultados de este trabajo podemos obtener buenas aproximaciones para las t-normas más usuales: Łukasiewicz, Producto y Mínimo.

2 Preliminares

Esta sección presenta algunos resultados sobre operadores de T-indistinguibilidad que serán necesarios a lo largo del trabajo.

Definición 2.1 *La residuación* $\overrightarrow{T}$ *de una t-norma* T *se define para todo* $x, y \in [0,1]$ *mediante*

$$\overrightarrow{T}(x|y) = sup\{\alpha \in [0,1] \mid T(x,\alpha) \leq y).$$

Definición 2.2 *El operador de* T*-indistinguibilidad natural* E_T *asociado a una t-norma* T *dada es la relación borrosa en* [0,1] *definida para todo* $x, y \in [0,1]$ *mediante*

$$E_T(x,y) = T(\overrightarrow{T}(x|y), \overrightarrow{T}(y|x)).$$

E_T es de hecho un tipo especial de operador de T-indistinguibilidad (Definición 2.3) [1] y en el contexto de

la Lógica Borrosa, donde T representa la conjunción, E_T se interpreta como la biimplicación asociada a T [4].

Definición 2.3 *Dada una t-norma T, un operador de T-indistinguibilidad E en un conjunto X es una relación borrosa $E : X \times X \to [0,1]$ que para todo $x, y, z \in X$ satisface*

1. $E(x,x) = 1$ *(Reflexividad)*
2. $E(x,y) = E(y,x)$ *(Simetría)*
3. $T(E(x,y), E(y,z)) \leq E(x,z)$ *(T-transitividad).*

Ejemplo 2.4

1. *Si T es la t-norma de Łukasiewicz, entonces $E_T(x,y) = 1 - |x-y|$ para todo $x, y \in [0,1]$.*
2. *Si T es la t-norma Producto, entonces $E_T(x,y) = Min(\frac{x}{y}, \frac{y}{x})$ para todo $x, y \in [0,1]$ donde $\frac{z}{0} = 1$.*
3. *Si T es la t-norma Mínimo, entonces* $E_T(x,y) = \begin{cases} Min(x,y) & \text{si } x \neq y \\ 1 & \text{en otro caso.} \end{cases}$

Teorema 2.5 *Teorema de Representación* [8]. *Sea R una relación borrosa en un conjunto X y T una t-norma continua. R es una relación de T-indistinguibilidad si, y sólo si, existe una familia $(h_i)_{i \in I}$ de subconjuntos borrosos de X tal que para todo $x, y \in X$*

$$R(x,y) = inf_{i \in I} E_T(h_i(x), h_i(y)).$$

$(h_i)_{i \in I}$ se llama una familia de generadores de R.

En particular, dada una relación de proximidad R en X (i.e. una relación borrosa reflexiva y simétrica), podemos construir el operador de T-indistinguibilidad $\underline{R}$ generado por el conjunto de las columnas de R (i.e. los subconjuntos borrosos $R(x,\cdot)$, $x \in X$).

Proposición 2.6 $\underline{R} \leq R$.

Definición 2.7 *Sea R una relación o matriz de proximidad en X y T una t-norma continua. La clausura T-transitiva $\overline{R}$ de R es el menor operador de T-indistinguibilidad en X que cumple $R \leq \overline{R}$.*

Definición 2.8 *Sean R y S dos relaciones borrosas de X y T una t-norma continua. El producto Sup-T de R y S es la relación borrosa $R \circ S$ on X definida para todo $x, y \in X$ mediante*

$$(R \circ S)(x,y) = sup_{z \in X} T(R(x,z), S(z,y)).$$

Gracias a que el producto Sup-T es asociativo para t-normas continuas, podemos definir la potencia n-ésima R_T^n de una relación borrosa R para cualquier $n \in N$:

$$R_T^n = \overbrace{R \circ ... \circ R}^{n \ veces}.$$

Definición 2.9 *Sea R una relación borrosa en un conjunto X y T una t-norma continua. La clausura T-transitiva de R es la relación borrosa*

$$R_T = sup_{n \in N} R_T^n.$$

Proposición 2.10 *Sea R una relación de proximidad en un conjunto finito X de cardinal n. Entonces*

$$R_T = sup_{s \in \{1,...,n-1\}} R_T^s.$$

Definición 2.11 *Sea R una relación de proximidad en un universo X y T una t-norma. Un operador de T-indistinguibilidad E de X es un opening T-transitivo de R si, y sólo si, $E \leq R$ y E es maximal en el conjunto de operadores de T-indistinguibilidad menores o iguales que R.*

3 El algoritmo

En esta sección se dará un algoritmo que genera una relación de similitud E cercana a una relación de proximidad R dada. La idea es partir de su clausura transitiva o de una relación de similitud menor o igual que R y modificar sus valores para hallar una mejor aproximación de R.

Proposición 3.1 [5] *Sea E una relación de similitud en un universo finito X de cardinal n. El número de entradas diferentes de la matriz que representa a E es menor o igual que n.*

El algoritmo para aproximar relaciones de proximidad mediante relaciones de similitud se basa en el siguiente resultado.

Proposición 3.2 *Sea E una relación de similitud en un universo finito X de cardinal n y $a_1 < a_2 < ... < a_k = 1$ ($k \leq n$) las entradas de E. Si reemplazamos las entradas por $a'_1 \leq a'_2 \leq ... \leq a'_k = 1$ respectivamente, se obtiene una nueva relación de similitud en X.*

Demostración

La reflexividad y la simetría de la nueva matriz son triviales.

Transitividad: Si $Min(a_i, a_j) \leq a_k$, entonces $Min(a'_i, a'_j) \leq a'_k$.

Ejemplo 3.3

$$E = \begin{pmatrix} 1 & 0.2 & 0.2 & 0.2 \\ 0.2 & 1 & 0.3 & 0.3 \\ 0.2 & 0.3 & 1 & 0.4 \\ 0.2 & 0.3 & 0.4 & 1 \end{pmatrix}$$

es una relación de similitud. Si reemplazamos 0.2, 0.3, 0.4, 1 por 0.1, 0.5, 0.8, 1 respectivamente, obtenemos la nueva relación de similitud

$$E' = \begin{pmatrix} 1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 1 & 0.5 & 0.5 \\ 0.1 & 0.5 & 1 & 0.8 \\ 0.1 & 0.5 & 0.8 & 1 \end{pmatrix}$$

La idea que se sigue para hallar una relación de similitud cercana a una relación de proximidad R dada es sencilla. Podemos calcular la clausura *Min*-transitiva $\overline{R}$ de R y luego modificar las entradas de $\overline{R}$ para minimizar alguna distancia dada a R o para maximizar alguna medida de similitud a R. Sin duda, también podemos calcular primero un opening transitivo o la relación de similitud obtenida de R mediante el Teorema de Representación en vez de la clausura *Min*-transitiva.

Sin embargo, el método no es tan directo tal y como muestran los siguientes dos ejemplos.

Ejemplo 3.4 *Consideremos la relación de proximidad con matriz*

$$R = \begin{pmatrix} 1 & 0.7 & 0.3 & 1 \\ 0.7 & 1 & 0.4 & 0.7 \\ 0.3 & 0.4 & 1 & 0.8 \\ 1 & 0.7 & 0.8 & 1 \end{pmatrix}.$$

Su clausura Min-transitiva es

$$\overline{R} = \begin{pmatrix} 1 & 0.7 & 0.8 & 1 \\ 0.7 & 1 & 0.7 & 0.7 \\ 0.8 & 0.7 & 1 & 0.8 \\ 1 & 0.7 & 0.8 & 1 \end{pmatrix}.$$

Si reemplazamos las entradas con valores 0.7 y 0.8 de $\overline{R}$ por a y b respectivamente, para minimizar la distancia euclídea d entre R y la nueva matriz debemos minimizar

$$\begin{aligned} f(a,b) &= (a-0.7)^2 + (a-0.4)^2 + (a-0.7)^2 \\ &\quad +(b-0.8)^2 + (b-0.3)^2. \end{aligned}$$

$$\frac{\partial f}{\partial a} = 2(a-0.7) + 2(a-0.4) + 2(a-0.7) = 0$$

$$\frac{\partial f}{\partial b} = 2(b-0.8) + 2(b-0.3) = 0$$

y

$$a = \frac{0.7+0.7+0.4}{3} = 0.6$$

$$b = \frac{0.8+0.3}{2} = 0.55$$

obteniendo

$$\begin{pmatrix} 1 & 0.6 & 0.55 & 1 \\ 0.6 & 1 & 0.6 & 0.6 \\ 0.55 & 0.6 & 1 & 0.55 \\ 1 & 0.6 & 0.55 & 1 \end{pmatrix}.$$

Esta matriz no es Min-transitiva y por tanto no es una relación de similitud.

El mismo problema puede presentarse con aproximaciones por debajo tal y como muestra el siguiente ejemplo.

Ejemplo 3.5 *De la matriz de proximidad*

$$R = \begin{pmatrix} 1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 1 & 0.7 & 0.8 \\ 0.3 & 0.7 & 1 & 0.5 \\ 0.4 & 0.8 & 0.5 & 1 \end{pmatrix}$$

se obtiene la matriz $\underline{R}$ mediante el Teorema de Representación.

$$\underline{R} = \begin{pmatrix} 1 & 0.2 & 0.2 & 0.2 \\ 0.2 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.3 \\ 0.2 & 0.2 & 0.3 & 1 \end{pmatrix}$$

Si reemplazamos las entradas con valores 0.2 y 0.3 de $\underline{R}$ por a y b respectivamente, para minimizar la distancia euclídea d entre R y la nueva matriz debemos minimizar

$$\begin{aligned} f(a,b) &= (a-0.2)^2 + (a-0.3)^2 + (a-0.4)^2 \\ &\quad + (a-0.7)^2 + (a-0.8)^2 + (b-0.3)^2. \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial a} &= 2(a-0.2) + 2(a-0.3) + 2(a-0.4) \\ &\quad +2(a-0.7) + 2(a-0.8) = 0 \\ \frac{\partial f}{\partial b} &= 2(b-0.3) = 0 \end{aligned}$$

y

$$a = \frac{0.2+0.3+0.4+0.7+0.8}{5} = 0.48$$

$$b = 0.3.$$

La matriz

$$\begin{pmatrix} 1 & 0.48 & 0.48 & 0.48 \\ 0.48 & 1 & 0.48 & 0.48 \\ 0.48 & 0.48 & 1 & 0.3 \\ 1 & 0.48 & 0.3 & 1 \end{pmatrix}$$

no es Min-transitiva y por tanto no es una relación de similitud.

Una posible solución en estos casos es reemplazar las entradas "'mal ordenadas"' por un único valor. Por ejemplo, en el Ejemplo 3.4 las entradas 0.6 y 0.55 se pueden reemplazar por $\frac{3\cdot 0.6+2\cdot 0.55}{5} = 0.58$ y en el Ejemplo 3.5 0.48 y 0.3 por $\frac{5\cdot 0.48+1\cdot 0.3}{6} = 0.45$. Nótese que gracias al siguiente lema estos nuevos valores son los que minimizan la distancia entre la matriz obtenida y R.

Lema 3.6 *Sea $P = (x_1, x_2, ..., x_n) \in R^n$. El punto Q más próximo a P respecto a la distancia euclídea de la forma $(a, a, ..., a)$ cumple $a = \frac{x_1+x_2+...+x_n}{n}$.*

El algoritmo que proponemos para hallar una relación de similitud cercana a una relación de proximidad R dada en un universo finito de cardinal n es el siguiente.

Algoritmo 3.7

1. *Calcular la clausura Min-transitiva $\overline{R}$ o una aproximación inferior de R.*
2. *Ordenar las entradas $a_1 < a_2 < ... < a_k = 1$ $(k \leq n)$ de $\overline{R}$.*
3. *Reemplazar cada a_i por la media a'_i de las entradas de R que ocupan el mismo lugar que los a_i.*
4. IF *$a'_1 \leq a'_2 \leq ... \leq a'_k = 1$, entonces la relación de similitud E se obtiene reemplazando las entradas $a_1, a_2, ..., a_k$ de $\overline{R}$ por $a'_1 \leq a'_2 \leq ... \leq a'_k$ respectivamente.*
5. ELSE, *para cada cadena maximal $C = \{a'_i, a'_{i+1}, ..., a'_{i+j}\}$ con $a'_i > a'_{i+j}$, reemplazar todos los elementos de C por su media aritmética ponderada a_C, pesando cada a'_l de C por el número de entradas de R que corresponden a a'_l. Reemplazando los elementos de C por a_C en E se obtiene la relación de similitud deseada.*

4 Conclusiones

En este trabajo hemos presentado un algoritmo para hallar aproximaciones de relaciones de proximidad por relaciones de similitud que en general son mejores que su clausura transitiva o sus openings transitivos. Por ejemplo, la distancia euclídea entre la proximidad del Ejemplo 3.4 y su clausura transitiva es 0.825, mientras que su distancia a la matriz obtenida usando el algoritmo propuesto es 0.795. La distancia euclídea entre la proximidad del Ejemplo 3.5 y la similitud obtenida mediante el Teorema de Representación es 1.183, y su distancia a la matriz obtenida usando el algoritmo propuesto es 0.768.

La sencillez del algoritmo permite su aplicación a problemas reales sin necesidad de aumentar la complejidad o tiempo de ejecución respecto a los métodos clásicos.

Los resultados de este trabajo junto con los de [3] permiten hallar buenas aproximaciones de una relación de proximidad mediante un operador de T-indistinguibilidad para las t-normas más usuales: *Mínimo* y t-normas arquimedianas continuas (incluidas la t-norma de *Łukasiewicz* y la t-norma *Producto*).

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos números TIN2006-14311 y TIN2006-06190.

Referencias

[1] D. Boixader, J. Jacas, J. Recasens (2000). Fuzzy Equivalence Relations: Advanced Material. En Dubois, Prade Eds. *Fundamentals of Fuzzy Sets*, Kluwer, 261-290.

[2] J. Fodor, M. Roubens (1995). Structure of transitive valued binary relations. *Math. Social Sci.* 30, 71–94.

[3] L. Garmendia, J. Recasens (2007). Approximation of Proximities by Aggregating T-indistinguishability Operators. Mathware & Soft Computing 14, 171-181.

[4] P. Hájek (1998) *Metamathematics of Fuzzy Logic*. Kluwer. Dordrecht.

[5] J. Jacas (1990). Similarity relations - the calculation of minimal generating families. *Fuzzy Sets and Systems* 35, 151–162.

[6] E. P. Klement, R. Mesiar, E. Pap (2000). *Triangular norms*. Kluwer. Dordrecht.

[7] B. Schweizer, A. Sklar (1983) *Probabilistic Metric Spaces*. North-Holland. Amsterdam.

[8] L.Valverde (1985). On the structure of F-indistinguishability operators, *Fuzzy Sets and Systems* 17, 313-328.

[9] L.A.Zadeh (1971). Similarity relations and fuzzy orderings, *Information Science* 3, 177-200.

FUZZY CLOSING OPERATORS AND THEIR COHERENCE AS FUZZY CONSEQUENCE OPERATORS

J. Elorza[1] **R. Fuentes-González**[2] **J. Bragard**[1] **P. Burillo**[2]

[1] Dept. Physics and Applied Mathematics. University of Navarra, {jelorza,jbragard}@unav.es
[2] Dept. Automation and Computer Science. Public University of Navarra, {rfuentes,pburillo}@unavarra.es

Abstract

In a previous paper [7] we explored the notion of coherent fuzzy consequence operator. It is well-known that the operator induced by a fuzzy preorder through Zadeh's compositional rule is always a coherent fuzzy consequence operator. It is also known that the relation induced by a fuzzy consequence operator is a fuzzy preorder if such operator is coherent [5]. Moreover fuzzy closing operators of mathematical morphology can be considered as fuzzy consequence operators. The aim of this paper is to analyze the coherence of such operators and also to consider their induced relations.

Keywords: Fuzzy Preorder, Fuzzy Consequence Operator, Fuzzy Closing Operator, Coherent Operators, Approximate Reasoning, Mathematical Morphology, Image Processing.

1 INTRODUCTION

Consequence Operators in classical logic were introduced by A. Tarski in 1930 [17]. According to Tarski, a logic is just a set of propositions with a consequence operator. The relationship between consequence operators and preorders is well known [3].

Concepts of Fuzzy Preorder and Fuzzy Consequence Operator (FCO for short) are essential on fuzzy logic. These notions have been defined as a natural generalization of the classical ones:

Given a non-empty universal set X which will represent a set of propositions and a t-norm $*$, a fuzzy (binary) relation R on X (fuzzy subset of $X \times X$) is called a fuzzy $*$-preorder if it satisfies:

(R1) $R(x,x) = 1 \; \forall \; x \in X$ (reflexivity)

(R2) $R(x,z) \geq R(x,y) * R(y,z) \; \forall \; x,y,z \in X$ ($*$-transitivity).

If R is also symmetric, this is, $R(x,y) = R(y,x)$ for all $x, y \in X$ then it is called a fuzzy $*$-similarity, fuzzy $*$-indistinguishability or, equivalently, fuzzy $*$-equivalence. If R is only reflexive and symmetric, we will say that it is a fuzzy tolerance.

Fix a complete lattice L which will be the range of the memberships of the fuzzy subsets of X, J. Pavelka introduced in 1979 the concept of FCO on X in fuzzy logic [14] extending the concept of consequence operator in Tarski's sense in a natural way.

A function $C : L^X \longrightarrow L^X$ is a FCO on X if it satisfies:

(C1) $\mu \subset C(\mu)$ for all fuzzy subset $\mu \in L^X$ (inclusion)

(C2) $\mu_1 \subset \mu_2 \Longrightarrow C(\mu_1) \subset C(\mu_2)$ for all $\mu_1, \mu_2 \in L^X$ (monotony)

(C3) $C(C(\mu)) \subset C(\mu)$ for all $\mu \in L^X$ (idempotence)

Notice that, under the inclusion axiom, (C3) may be written equivalently as (C3') $C(C(\mu)) = C(\mu) \; \forall \; \mu \in L^X$.

These operators, called closure operators in the algebraic, logical and topological contexts, have been studied extensively [1],[2],[9],[19]. During the last and current decades, these operators have been also studied in the context of the fuzzy logic taking the chain $L = [0,1]$ as a special case [4],[5],[7],[8],[11].

In 1991, J.L. Castro and E. Trillas proved [5] the following result: If R is a fuzzy $*$-preorder then the operator C_R^* between fuzzy subsets of X given by the

max-$*$ Zadeh's compositional rule

$$C_R^*(\mu) = \mu \circ^* R \tag{1}$$

is a fuzzy consequence operator (induced by R), where

$$(\mu \circ^* R)(x) = \sup_{w \in X} \{\mu(w) * R(w,x)\} \tag{2}$$

We proved [7] that the converse of the previous result is also true: R is a $*$-preorder if and only if C_R^* is a FCO. Thus the only possibility to obtain FCOs from relations using Zadeh's max-$*$ compositional rule is to work with $*$-preorders.

For the inverse process (inducing fuzzy preorders from FCOs), J.L. Castro and E. Trillas added the coherence axiom to the FCO concept [5]. A FCO C is called $*$-coherent if $\mu(a) * C(\varphi_a)(x) \leq C(\mu)(x)$ for all $\mu \in [0,1]^X$ and for all $(a,x) \in X \times X$, where $\varphi_a(t) = \varphi_{\{a\}}(t)$ is the crisp membership of the singleton $\{a\}$.

It is proven [5] that if C is a $*$-coherent FCO then the relation defined by

$$R_C(x,y) = C(\varphi_x)(y) \tag{3}$$

is a fuzzy $*$-preorder (induced by C). It is also proven that the operator C_R^* is $*$-coherent.

It is easy to extend the notion of coherence to operators $C : L^X \longrightarrow L^X$ between fuzzy subsets in L^X. This is: a FCO C is called $*$-coherent if $\mu(a) * C(\varphi_a)(x) \leq C(\mu)(x)$ for all $\mu \in L^X$ and for all $(a,x) \in X \times X$, where $\varphi_a(t) = \begin{cases} 1 & \text{if } t = a \\ 0 & \text{if } t \neq a \end{cases}$, being in this case 1 and 0 the greatest and least element of the lattice L, respectively. Then, all the previous results hold.

On the other hand, the classical morphological operators of image processing have been introduced by G. Matheron an J. Serra in [12],[13],[17]. These operators have been studied recently, see for instance [10],[18]. In [10] they are studied in a new context: obtaining relevant information in fuzzy relational systems.

In this context, the most used morphological operators are erosion, dilation, opening and closing that are given respectively by $\varepsilon_R(\mu) = R^{op} \lhd \mu$, $\delta_R(\mu) = R \circ \mu$, $\alpha_R(\mu) = R \circ (R^{op} \lhd \mu)$ and $\beta_R(\mu) = R \lhd (R^{op} \circ \mu)$, where R is the relation acting as structurant element and μ is in the class of all fuzzy subsets of the universal set X or $X \times X$. R^{op} denotes the inverse relation of R, $R^{op}(x,y) = R(y,x)$ and the operators $\circ$, $\lhd$ are defined as usual:

$(R \circ \mu)(x) = \sup_{w \in X} \{R(x,w) * \mu(w)\}$, $(R \lhd \mu)(x) = \inf_{w \in X} \{R(x,w) \rightsquigarrow \mu(w)\}$ for some t-norm $*$ and for some implication operator $\rightsquigarrow$.

Notice that the operator $\circ$ and possibly the operators $\lhd$ and $\rightsquigarrow$ depend on the t-norm $*$. However, when there is no confusion, in order to alleviate the notation we omit the symbol $*$.

From now on, $*$ will be a left-continuous t-norm, that is, $\alpha * (\sup M) = \sup(\alpha * M)$ for all $\alpha \in L$, $M \subset L$ and $\rightsquigarrow_*$ will denote the residuated implication induced by the t-norm $*$: $\alpha \rightsquigarrow_* \beta = \sup\{w \in L / \alpha * w \leq \beta\}$. It is also called the pseudoinverse application of $*$.

Both in the crisp and in the fuzzy cases, opening and closing operators with an isotropic structuring element are used in image processing to eliminate specific image details smaller than the structuring element. The global shape of the objects is not distorted.

In particular, a closing operator connects objects that are close to each other, fills up small holes and smooths the object outline by filling up narrow gulfs. Meanings of *near*, *small* and *narrow* are related to the size and the shape of the structuring element.

In this way, the concept of fuzzy closing operator φ_R^* induced by a fuzzy relation is relevant in fuzzy mathematical morphology, defined in [10] as follows. For any fuzzy relation R on X,

$$\varphi_R^*(\mu) = R^{op} \lhd_{\rightsquigarrow_*} (R \circ^* \mu) \tag{4}$$

where $R \circ^* \mu$ is given by

$(R \circ^* \mu)(x) = \sup_{w \in X} \{R(x,w) * \mu(w)\}$ and

$(R \lhd_{\rightsquigarrow_*} \mu)(x) = \inf_{w \in X} \{R(x,w) \rightsquigarrow_* \mu(w)\}$.

It is well-know that φ_R^* is a closure operator, that is, a fuzzy consequence operator.

In Section 2 we analyze the coherence of fuzzy closing operators as fuzzy consequence operators.

In Section 3 we analyze some properties of the preorders induced from fuzzy closing operators.

Recall three properties of t-norms, infimum and residuated implications which will be used in Section 2:

$$x * \inf_{i \in I} \{y_i\} \leq \inf_{i \in I} \{x * y_i\} \tag{5}$$

$$x * (y \rightsquigarrow_* z) \leq y \rightsquigarrow_* (x * z) \tag{6}$$

$$y_1 \leq y_2 \ \ implies \ \ x \rightsquigarrow_* y_1 \leq x \rightsquigarrow_* y_2 \tag{7}$$

The following property will be used in Section 3:

$$x_1 \leq x_2 \ \ implies \ \ x_2 \rightsquigarrow_* y \leq x_1 \rightsquigarrow_* y \tag{8}$$

2 COHERENCE OF FUZZY CLOSING MORPHOLOGICAL OPERATORS

In this section, we will prove that φ_R^* is always an operator $*$-coherent as fuzzy consequence operator.

Notice that if for any $x \in X$, $\tilde{x}_\alpha$ denotes the fuzzy subset $\tilde{x}_\alpha(t) = \begin{cases} \alpha & \text{if } t = x \\ 0 & \text{if } t \neq x \end{cases}$ then $\varphi_a = \tilde{a}_1$. In the following, we make use of this notation.

Lemma 1.-*Let R be any fuzzy relation on X, $\mu \in L^X$ and $x \in X$, then:*

$$\varphi_R^*(\mu)(x) = \inf_{w \in X} \{R^{op}(x,w) \rightsquigarrow_* C^*_{R^{op}}(\mu)(w)\}$$

Proof.- It is straightforward from (1) and commutativity of the t-norm $*$. □

Lemma 2.-*Let R be any fuzzy relation on X, $\mu \in L^X$ and $x \in X$, then:*

$$C_R^*(\tilde{a}_1)(x) = R(a,x)$$

Proof.- Given $\mu \in L^X$ and $a, x \in X$: $C_R^*(\tilde{a}_1)(x) = (\mu \circ^* R)(x) = \sup_{w \in X} \{\tilde{a}_1(w) * R(w,x)\}$. Observe that this supremum is taken in $w = a$. Thus: $C_R^*(\tilde{a}_1)(x) = \tilde{a}_1(a) * R(a,x) = 1 * R(a,x) = R(a,x)$. □

Theorem 1.-*Let R be any fuzzy relation on X. Then the fuzzy closing operator $\varphi_R^* : L^X \longrightarrow L^X$ induced by R by means of (4) is a $*$-coherent fuzzy consequence operator .*

Proof.- It is enough to prove the coherence condition. Given $\mu \in L^X$ and $a, x \in X$. From Lemma 1:

$$\mu(a) * \varphi_R^*(\tilde{a}_1)(x) =$$

$$\mu(a) * \inf_{w \in X} \{R^{op}(x,w) \rightsquigarrow_* C^*_{R^{op}}(\tilde{a}_1)(w)\}$$

From Lemma 2:

$$\mu(a) * \varphi_R^*(\tilde{a}_1)(x) =$$

$$\mu(a) * \inf_{w \in X} \{R^{op}(x,w) \rightsquigarrow_* R(w,a)\} =$$

$$\mu(a) * \inf_{w \in X} \{R(w,x) \rightsquigarrow_* R(w,a)\}$$

By properties (5) and (6):

$$\mu(a) * \varphi_R^*(\tilde{a}_1)(x) \leq$$

$$\inf_{w \in X} \big\{\mu(a) * \big(R(w,x) \rightsquigarrow_* R(w,a)\big)\big\} \leq$$

$$\inf_{w \in X} \big\{R(w,x) \rightsquigarrow_* \big(R(w,a) * \mu(a)\big)\big\}$$

Notice that

$$C^*_{R^{op}}(\mu)(w) = \sup_{u \in X} \{R(w,u) * \mu(u)\} \geq$$

$$R(w,a) * \mu(a)$$

Then by property (7):

$$\mu(a) * \varphi_R^*(\tilde{a}_1)(x) \leq$$

$$\inf_{w \in X} \{R(w,x) \rightsquigarrow_* C^*_{R^{op}}(\mu)(w)\} = \varphi_R^*(\mu)(x)$$

Therefore $\mu(a) * \varphi_R^*(\tilde{a}_1)(x) \leq \varphi_R^*(\mu)(x)$ and φ_R^* is $*$-coherent. □

3 RELATIONS INDUCED BY FUZZY CLOSING MORPHOLOGICAL OPERATORS

We have proven in the previous section that the fuzzy closing operator φ_R^* induced by a relation R by means of (4) is $*$-coherent.

As a consequence the relation induced by the fuzzy consequence operator operator φ_R^* by means of (3) is a $*$-preorder.

We now introduce a characterization of such induced relation in order to analyze some properties.

Recall that the $\lhd_{\rightsquigarrow_*}$ composition for two relations R and S on X is defined by $(R \lhd_{\rightsquigarrow_*} S)(x,y) = \inf_{w \in X} \{R(x,w) \rightsquigarrow_* S(w,y)\}$.

Lemma 3.-*Let R be any fuzzy relation on X and put $C = \varphi_R^*$. Then the $*$-preorder induced by C by means of (3), R_C is such that*

$$R_C = (R^{op} \lhd_{\rightsquigarrow_*} R)^{op} \tag{9}$$

Proof.- $(R^{op} \lhd_{\rightsquigarrow_*} R)^{op}(x,y) = (R^{op} \lhd_{\rightsquigarrow_*} R)(y,x) = \inf_{w \in X} \{R(w,y) \rightsquigarrow_* R(w,x)\}$. It is the same that $R_C(x,y)$. In fact, $R_C(x,y) = \varphi_R^*(\tilde{x}_1)(y)$.

By Lemmas 1,2:

$R_C(x,y) = \left[R^{op} \lhd_{\rightsquigarrow_*} (R \circ^* \tilde{x}_1)\right](y) = \inf_{w\in X}\{R(w,y) \rightsquigarrow_* R(w,x)\}$ □

Remark 1.- Notice that given any fuzzy relation R, $R^{op} \lhd_{\rightsquigarrow_*} R$ is always $*$-transitive even though it does not involve the $\circ$ operator.

In the following, we make use of the identity relation I on X, that is, $I(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$

Theorem 2.-*Let R be any fuzzy relation on X. Then R is reflexive if and only if $R^{op} \lhd_{\rightsquigarrow_*} R \leq R$.*

Proof.- If R is reflexive $I \leq R$ and $I = I^{op} \leq R^{op}$. Then by (8): $R = I \lhd_{\rightsquigarrow_*} R \geq R^{op} \lhd_{\rightsquigarrow_*} R$.

Conversely, if $R^{op} \lhd_{\rightsquigarrow_*} R \leq R$, for each $x \in X$:

$$R(x,x) \geq (R^{op} \lhd_{\rightsquigarrow_*} R)(x,x) =$$

$$\inf_{w\in X}\{R^{op}(x,w) \rightsquigarrow_* R(w,x)\} =$$

$$\inf_{w\in X}\{R(w,x) \rightsquigarrow_* R(w,x)\} = \inf_{w\in X}\{1\} = 1.$$

Thus $I \leq R$ and R is a reflexive relation. □

Remark 2.- In the case of a universe X with two elements, put $X = \{x,y\}$, then R is reflexive if and only if $R^{op} \lhd_{\rightsquigarrow_*} R = R$. In fact, if R reflexive:

$$(R^{op} \lhd_{\rightsquigarrow_*} R)(x,y) = \inf_{w\in X}\{R^{op}(x,w) \rightsquigarrow_* R(w,y)\} =$$

$$\inf_{w\in X}\{R(w,x) \rightsquigarrow_* R(w,y)\} =$$

$$\inf\{R(x,x) \rightsquigarrow_* R(x,y)\,,\, R(y,x) \rightsquigarrow_* R(y,y)\} =$$

$$\inf\{1 \rightsquigarrow_* R(x,y)\,,\, R(y,x) \rightsquigarrow_* 1\} =$$

$\inf\{R(x,y)\,,\,1\} = R(x,y)$ and $R^{op} \lhd_{\rightsquigarrow_*} R = R$. The converse is straightforward from Theorem 2.

Notice that it is easy to prove as in the proof of previous theorem that $R^{op} \lhd_{\rightsquigarrow_*} R \leq R^{op} \Longrightarrow I \leq R^{op} \Longrightarrow I \leq R$. Then:

Remark 3.- The condition $R^{op} \lhd_{\rightsquigarrow_*} R \leq R^{op}$ also implies the reflexivity of relation R. However, in this case, the converse does not hold as show the following easy example.

Example.- Consider the universe $X = \{x,y\}$ and the fuzzy reflexive relation in $[0,1]^{X\times X}$ given by $R = \begin{pmatrix} 1 & 0.3 \\ 0.1 & 1 \end{pmatrix}$. Then $R^{op} = \begin{pmatrix} 1 & 0.1 \\ 0.3 & 1 \end{pmatrix}$ and from Remark 2, $R^{op} \lhd_{\rightsquigarrow_*} R = R = \begin{pmatrix} 1 & 0.3 \\ 0.1 & 1 \end{pmatrix} \not\leq \begin{pmatrix} 1 & 0.1 \\ 0.3 & 1 \end{pmatrix} = R^{op}$.

For the following results, recall that R_C denotes the $*$-preorder induced by C by means of (3), this is, $R_C(x,y) = C(\varphi_x)(y)$, where $C = \varphi^*_R$.

Corollary 1.- *Let R be any fuzzy relation on X. Then R is reflexive if and only if $R_C \leq R^{op}$.*

Proof.- It is straightforward from Theorem 2 and Lemma 3. □

Corollary 2.-*Let R be any fuzzy reflexive and symmetric relation on X, this is, a tolerance. Then $R_C \leq R$.*

Proof.- If R is symmetric, $R^{op} = R$. Then it is straightforward from Theorem 2. □

Recall that a fuzzy relation R is called irreflexive if $R(x,x) = 0$ for all $x \in X$.

For the following result we will consider a negation such that $\alpha' = \alpha \rightsquigarrow_* 0$.

Notice that in this case, since $(R \rightsquigarrow_* Q)^{op} = R^{op} \rightsquigarrow_* Q^{op}$, we get: $(R^{op})' = (R')^{op}$.

Corollary 3.-*If R is an irreflexive relation on X then $R_C \leq R'$.*

Proof.- If R is irreflexive, $R \leq I'$. Then by (7): $R^{op} \lhd_{\rightsquigarrow_*} R \leq R^{op} \lhd_{\rightsquigarrow_*} I'$. Moreover $(R \lhd_{\rightsquigarrow_*} I')(x,y) = \inf_{w\in X}\{R(x,w) \rightsquigarrow_* I'(w,y)\}$. This infimum is taken in $w = y$. Thus $(R \lhd_{\rightsquigarrow_*} I')(x,y) = R(x,y) \rightsquigarrow_* 0 = R'(x,y)$.
Hence $R^{op} \lhd_{\rightsquigarrow_*} R \leq (R^{op})' = (R')^{op}$. Therefore by Lemma 3: $R_C = (R^{op} \lhd_{\rightsquigarrow_*} R)^{op} \leq ((R')^{op})^{op} = R'$ □

Observe that the previous results in this paragraph about reflexivity and symmetry do not make use of the t-norm $*$ since these properties do not depend on $*$.

Theorem 3.-*Let R be any fuzzy relation on X. Then R is $*$-transitive if and only if $R \leq R^{op} \lhd_{\rightsquigarrow_*} R$.*

Proof.- Prove that for all Q, R, S fuzzy relations on X:

$$Q \circ^* R \leq S \Longleftrightarrow R \leq Q^{op} \lhd_{\rightsquigarrow_*} S \qquad (10)$$

In fact, $Q \circ^* R \leq S \iff \sup_{y \in X} \{Q(x,y) * R(y,z)\} \leq S(x,z)$ for all x, $z \in X \iff Q(x,y) * R(y,z) \leq S(x,z)$ for all x, y, $z \in X \iff R(y,z) \leq Q(x,y) \rightsquigarrow_* S(x,z)$ for all x, y, $z \in X$ due to the adjointness property of residuated lattices. Hence $Q \circ^* R \leq S \iff R(y,z) \leq Q^{op}(y,x) \rightsquigarrow_* S(x,z)$ for all x, y, $z \in X$. Taking infimum in x, the previous statement is equivalent to $R(y,z) \leq \inf_{x \in X} \{Q^{op}(y,x) \rightsquigarrow_* S(x,z)\}$ for all y, $z \in X \iff R(y,z) \leq (Q^{op} \lhd_{\rightsquigarrow_*} S)(y,z)$ for all y, $z \in X \iff R \leq Q^{op} \lhd_{\rightsquigarrow_*} S$. In particular for $R = S = Q$:

$$R \circ^* R \leq R \iff R \leq R^{op} \lhd_{\rightsquigarrow_*} R$$

Therefore: R is $*$-transitive $\iff R \leq R^{op} \lhd_{\rightsquigarrow_*} R$ □

Remark 4.- Again, we have shown a result about the $*$-transitivity which does not involve the $\circ$ operator.

Corollary 4.-*Let R be any fuzzy relation on X. Then R is an $*$-preorder if and only if $R = R^{op} \lhd_{\rightsquigarrow_*} R$.*

Proof.- It is enough to consider Theorems 2 and 3 □

Remark 5.- Using Lemma 3, the previous Corollary can be written: R is an $*$-preorder if and only if $R_C = R^{op}$.

Corollary 5.-*Let R be any fuzzy reflexive, symmetric and $*$-transitive relation on X, this is, a fuzzy $*$-equivalence. Then $R_C = R$.*

Proof.- If R is symmetric, $R^{op} = R$. Then it is straightforward from Remark 5. □

Remark 6.- In this case (R fuzzy $*$-equivalence), it holds $R_C = R \lhd_{\rightsquigarrow_*} R = R$.

Now, we will show that the converse of Corollary 5 is also held.

Theorem 4.-*Let R be any fuzzy relation, then R is a fuzzy $*$-equivalence if and only if $R_C = R$.*

Proof.- We will prove only that if $R_C = R$ then R is a fuzzy $*$-equivalence.

(i) Reflexivity: Recall that R_C is always a fuzzy $*$-preorder. In particular R_C is reflexive. Hence, if $R_C = R$, R is also reflexive.

(ii) Symmetry: $R_C = R \iff (R^{op} \lhd_{\rightsquigarrow_*} R)^{op} = R \iff R^{op} = R^{op} \lhd_{\rightsquigarrow_*} R \implies R^{op} \leq R^{op} \lhd_{\rightsquigarrow_*} R$. Taking $Q = R$, $R = R^{op}$ and $S = R$ in (10) of Theorem 3: $R^{op} \leq R^{op} \lhd_{\rightsquigarrow_*} R$, this is,

$$R \circ^* R^{op} \leq R$$

On the other hand, taking the inverse of the previous inequality $(R \circ^* R^{op})^{op} \leq R^{op}$. As $(R \circ^* S)^{op} = S^{op} \circ^* R^{op}$, we get: $(R^{op})^{op} \circ^* R^{op} \leq R^{op}$, this is,

$$R \circ^* R^{op} \leq R^{op}$$

Therefore $R \circ^* R^{op} \leq \inf\{R, R^{op}\}$, this is, $(R \circ^* R^{op})(x,y) \leq \inf\{R(x,y), R(y,x)\}$ for all x, $y \in X$, this is, $\sup_{w \in X} \{R(x,w) * R(y,w)\}$, this is, $R(x,w) * R(y,w) \leq \inf\{R(x,y), R(y,x)\}$ for all x, y, $w \in X$.

In particular, for $w = x$: $R(y,x) = 1 * R(y,x) = R(x,x) * R(y,x) \leq \inf\{R(x,y), R(y,x)\}$ due to the reflexivity of R.

and analogously for $w = y$: $R(x,y) = R(x,y) * 1 = R(x,y) * R(y,y) \leq \inf\{R(x,y), R(y,x)\}$ for all x, $y \in X$.

In consequence, $R(x,y) \leq \inf\{R(x,y), R(y,x)\}$ and $R(x,y) \leq \inf\{R(x,y), R(y,x)\}$. Hence $R(y,x) \leq R(x,y)$ and $R(x,y) \leq R(y,x)$, this is, $R(x,y) = R(y,x)$ for all x, $y \in X$. Therefore $R = R^{op}$ and R is symmetric.

(iii) $*$-transitivity: In the previous paragraph, we have proven that if $R_C = R$ then R is symmetric and $R^{op} \lhd_{\rightsquigarrow_*} R = R^{op}$. Therefore $R^{op} \lhd_{\rightsquigarrow_*} R = R$. From Theorem 3, R is $*$-transitive. □

We now show an application to fuzzy relational equations.

It is known that the fuzzy relation equation $R \circ^* X = S$ has a solution if and only if $X_0 = R^{op} \lhd_{\rightsquigarrow_*} S$ is a solution and in this case it is the greatest solution, see [6],[15].

Corollary 6.-*The fuzzy relation equation $R \circ^* X = R$ has a solution if and only if $X_0 = R^{op} \lhd_{\rightsquigarrow_*} R = (R_C)^{op}$ is a solution. In this case it is the greatest solution.*

Proof.- It is enough to consider the case $R = S$, the previous paragraph and Lemma 3. □

Notice that X_0 can be seen on one hand as the inverse of the fuzzy relation induced by the fuzzy closing operator that comes from R and on the other hand as the erosion of R with itself.

4 CONCLUSIONS AND FUTURE WORKS

We have connected the class of fuzzy consequence operators in the context of approximate reasoning with the class of fuzzy closing operators used in image processing acting as filters. We do not know of any previ-

ous relationship between both classes in the literature apart that both are closure operators.

More precisely, we have shown that given any fuzzy relation R used as the structurant element, the fuzzy closing operator $C = \varphi_R^*$ induced by such relation is always a coherent fuzzy consequence operator. Therefore such operator C induces always a preorder R_C. This result is the key for obtaining the other properties.

We have obtained the following representation of the induced relation: $R_C = (R^{op} \lhd_{\rightsquigarrow_*} R)^{op}$ from which other properties can be easily deduced.

Given a fuzzy relation R on an universe X, we have shown the following characterizations: R is reflexive if and only if $R \geq R^{op} \lhd_{\rightsquigarrow_*} R$ and R is $*$-transitive if and only if $R \leq R^{op} \lhd_{\rightsquigarrow_*} R$. In consequence R is an $*$-preorder if and only if $R = R^{op} \lhd_{\rightsquigarrow_*} R$. It should be notice that the characterization about the transitivity does not involve the $\circ$ operator.

Moreover, we have shown the following results about the preorder R_C induced by the fuzzy closing operator φ_R^*: If R is reflexive then $R_C \leq R^{op}$. If it is also symmetric then $R_C \leq R$ and it is also $*$-transitive, this is, a fuzzy $*$-equivalence relation then $R_C = R$. We have proven the converse of this result that involves reflexivity, symmetry and $*$-transitivity, this is, R is a fuzzy $*$-equivalence if and only if $R_C = R$. We have proven also that if R is irreflexive then $R_C \leq R'$ where $\alpha' = \alpha \rightsquigarrow 0$.

Finally, as an application to fuzzy relational equations, we have shown that the fuzzy relational equation $R \circ^* X = R$ has a solution if and only if $X_0 = R^{op} \lhd_{\rightsquigarrow_*} R = (R_C)^{op}$ is a solution and in this case it is the greatest solution.

In future works we will use the connection established in this paper between both classes of operators to translate well known properties from the field of approximate reasoning to the one of fuzzy relational systems. In particular we will explore its applications to image processing.

Ackonowledgements

We thank both anonymous referees for their valuable suggestions that have improved the final release of this work.

References

[1] R. Bělohlávek. Fuzzy closure operators. *Journal of Mathematical Analysis and Appl.* Vol 262, pp. 473-489, 2001.

[2] R. Bělohlávek. Fuzzy closure operators II. *Soft Computing* Vol 7. No 1, pp, 53-64, 2002.

[3] J.L. Castro. Fuzzy logics as families of bivaluated logics. *Fuzzy sets and systems* 64, pp. 321-332, 1994.

[4] J.L. Castro, M. Delgado, E. Trillas. Inducing Implication Relations. *International Journal of Approximate Reasoning.* Vol.10, pp. pp. 235-250, 1994.

[5] J.L. Castro, E. Trillas. Tarski's fuzzy consequences. *Proc. of the Internat. Fuzzy Engineering Symp'91.* Vol.1, pp. 70-81, 1991.

[6] A. Di Nola, A. Lettieri. Relation equations in residuated lattices. *Rendiconti del Circolo matematico di Palermo*, s.II, XXXVIII, pp. 246-256, 1989.

[7] J. Elorza, P. Burillo. On the relation between fuzzy preorders and fuzzy consequence operators. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems.* Vol.7. No 3, pp. 219-234, 1999.

[8] J. Elorza, P. Burillo. Connecting fuzzy preorders, fuzzy consequence operators and fuzzy closure and co-closure systems. *Fuzzy sets and systems.* Vol.139. No 3, pp. 601-613, 2003.

[9] F. Esteva, P. García, L. Godo, R.O Rodríguez. Fuzzy Approximation Relations, Modal Structures and Possibilistic Logic. *Mathware & Soft Computing*, Vol. V, n 2-3, pp. 151-166, 1998.

[10] R. Fuentes-González, P. Burillo, N. Frago. Técnicas de Morfología Matemática en el tratamiento de Sistemas Relacionales Difusos *Actas del congreso Estylf 2002 sobre tecnologías y lógica fuzzy*, pp. 321-326, 2002.

[11] R. Jacas, J. Recasens. Preórdenes Duales, Operadores de Clausura y Contextos Borrosos *Actas del congreso Estylf 2004 sobre tecnologías y lógica fuzzy*, pp. 125-128, 2004.

[12] G. Matheron. Eléments pour une théorie des milieux poreux. *Masson.* Paris. 1967

[13] G. Matheron. Random Sets and Integral Geometry. *Wiley.* NY. 1975

[14] J. Pavelka. On Fuzzy Logic I. *Zeitschr. f. Math. Logik und Grundlagen d. Math.* Bd.25, pp. 45-52, 1979.

[15] K. Peeva, Y. Kyosev. Algorithm for solving max-product fuzzy relational equations. *Soft Computing* Vol 11. No 7, pp, 593-605, 2007.

[16] A. Tarski. Logic, semantics, metamathematics. Clarendon Press, Oxford, 1956.

[17] J. Serra. Image Analysis and Mathematical Morphology. *Academic Press*. Vol I (fourth printing, 1990) and Vol II (second printing, 1992).

[18] M.E. Valle, P. Sussner. A general framework for fuzzy morphological associative memories. *Fuzzy sets and systems*. Vol.159. No 7, pp. 747-768, 2008.

[19] M. Ward. The closure operators of a lattice. *Annals of Mathematics*. Vol.43. N.2, pp. 191-196, 1940.

Modelling conjunctions by ordinal sums*

Itziar García-Honrado* **Enric Trillas*** **Eloy Renedo***

* European Center for Soft-Computing. Mieres, Asturias (España), {itziar.garcia, enric.trillas, eloy.renedo}@softcomputing.es

Abstract

This paper shows two examples which require the use of ordinal sums in order to model a conjunction by a continuous t-norm.

keywords: Ordinal sum of t-norms, idempotency.

1 Introduction

It is a fact that fuzzy logic contributed so much to wide the knowledge of t-norms and t-conorms. Although practically only the t-norms *min* and *prod* are used in the applications, in a lot of theoretical papers both *min* and the archimedean t-norms are taken into account to obtain mathematical results. But for what concerns the family of ordinal sums of archimedean t-norms, neither in theoretical nor in applied communications are they considered, except in some few cases. Hence, it is currently avoided the case in which only some statements, but not all, are to be considered conjunctively idempotent.
Would it be the case that 'x is P' is idempotent with $\mu_P(x) \neq 0, 1$, and 'x is Q' is not idempotent, to keep $\mu_{P\&P}(x) = T(\mu_P(x), \mu_P(x)) = \mu_P(x)$, and $T(\mu_Q(x), \mu_Q(x)) < \mu_Q(x)$, there is the possibility of taking T as a ordinal sum of archimedean t-norms. In the way from fuzzy logic towards computation with words, where more complex linguistic expressions will be considered, ordinal sums will surely grow in importance.
This short paper only tries to show two simple examples where the use of ordinal sums seems to be reasonable.

*This paper is partially supported by CICYT (Spain) under project TIN2005-08943-C02-01

2 Basic concepts

This paper deals with continuous t-norms, T ([AFS06]), that is with mappings $T : [0,1] \times [0,1] \to [0,1]$ verifying the following properties,

- Continuity in both variables
- Associativity
  $T(T(x,y),z) = T(x,T(y,z)),\ \forall x,y,z \in [0,1]$
- Commutativity $T(x,y) = T(y,x)\ \forall x \in [0,1]$
- Monotonicity $T(x,y) \leq T(z,t)$ if $x \leq z, y \leq t$
- $T(x,1) = x,\ \forall x \in [0,1]$
- $T(x,0) = 0,\ \forall x \in [0,1]$

A t-norm T verifying $T(a,a) < a,\ \forall a \in (0,1)$ is an archimedean t-norm, for example, both *prod* and W are archimedean, but *min* is not.

An *order-automorphism* of the unit interval is a continuous and strictly increasing function $\varphi : [0,1] \to [0,1]$ verifying the boundary conditions $\varphi(0) = 0$ and $\varphi(1) = 1$.

Given a t-norm and an order-automorphism φ, the operator T_φ, defined by $T_\varphi = \varphi^{-1} \circ T \circ (\varphi \times \varphi)$, is also a t-norm that is continuous if and only if T is continuous. Given a t-norm T, the set of all the t-norms T_φ, where φ is any order-automorphism, is denoted $F(T)$ and is called *the family of* T, $F(T) = \{T_\varphi : \varphi \text{ is an order-automorphism}\}$.

Continuous t-norms are characterized as follows ([PTGR07]):

A function $T : [0,1] \times [0,1] \to [0,1]$ is a continuous t-norm if and only if it belongs to one of the following families,

- The min family, $F(min)$, which contains only the minimum t-norm min.
- The Product family, $F(Prod)$, composed of all the t-norms of the form $Prod_\varphi = \varphi^{-1} \circ Prod \circ (\varphi \times \varphi)$, where $Prod(x, y) = x \cdot y$ is the product t-norm and φ is any order-automorphism.
- The Łukasiewicz family, $F(W)$, made of all the t-norms of the form $W_\varphi = \varphi^{-1} \circ W \circ (\varphi \times \varphi)$, where $W(x, y) = max(0, x + y - 1)$ is the Łukasiewicz t-norm and φ is any order-automorphism.
- An ordinal sum of archimedean t-norms.
  Let I be a finite or infinite numerable set, $\{(a_i, b_i)/i \in I\}$ be a collection of open subintervals in $[0, 1]$ such that $(a_i, b_i) \cap (a_j, b_j) = \emptyset$, $\forall i \neq j$, and T_i for $i = 1, 2, ..., n$ an archimedean t-norm. So, the ordinal sum of these t-norms is

$$T(x,y) = \begin{cases} a_i + (b_i - a_i) \cdot T_i(\frac{x-a_i}{b_i-a_i}, \frac{y-a_i}{b_i-a_i}), \\ \qquad\qquad \text{if } (x,y) \in [a_i, b_i]^2 \\ \min(x,y), \quad \text{otherwise} \end{cases}$$

A t-norm, T, is said to be strictly monotonic, whenever, if $x > 0$ and $y < z$ then $T(y, x) < T(z, x)$ ([KMP00]). Any T in $F(Prod)$ is an strictly monotonic t-norm, but min and all the t-norms in $F(W)$ are not. An ordinal sum of archimedean t-norms T is not strictly monotonic if $T \notin F(Prod)$. So, there are the following posible situations,

- $T \in F(W)$, so T is not an strictly monotonic t-norm.
- The collection of open intervals, to define the ordinal sum of archimedean t-norms, contains at least two intervals, let them be $(a_1, a_2), (b_1, b_2)$ with $a_2 < b_1$. Taking $x \in (a_1, a_2)$ and $y, z \in (b_1, b_2)$ with $y < z$, it is $T(y, x) = \min(y, x) = x = \min(z, x) = T(z, x)$. So, T is not an strictly monotonic t-norm.

3 Modelling with ordinal sums

3.1 Introduction

Let us suppose we are working with statements coming from a family of predicates in the set X, and that for one of them it is known that,

- 'x_1 is P' does coincide with 'x_1 is $P\&P$'
- 'x_2 is P' does coincide with 'x_2 is $P\&P$'
- 'x_3 is P' does not coincide with 'x_3 is $P\&P$'

Provided we work in an standard algebra (T, S, N), it is $\mu_{P\&P} = T \circ (\mu_P \times \mu_P)$, and hence

- $\mu_{P\&P}(x_1) = T(\mu_P(x_1), \mu_P(x_1)) = \mu_P(x_1)$
- $\mu_{P\&P}(x_2) = T(\mu_P(x_2), \mu_P(x_2)) = \mu_P(x_2)$
- $\mu_{P\&P}(x_3) = T(\mu_P(x_3), \mu_P(x_3)) < \mu_P(x_3)$

Calling $\mu_P(x_1) = a$, $\mu_P(x_2) = b$, $\mu_P(x_3) = c$, and provided $a < b < c$, the t-norm T should be an ordinal sum. To define it, the collection of open intervals must contain intervals of which a and b are extremes. In this case if a and b are the only idempotent points, the collection of open intervals is {(0,a),(a,b), (b,1)}, and

$$T(x,y) = \begin{cases} a \cdot T_1(\frac{x}{a}, \frac{y}{a}), & \text{if } (x,y) \in [0,a]^2 \\ a + (b-a) \cdot T_2(\frac{x-a}{b-a}, \frac{y-a}{b-a}), & \text{if } (x,y) \in [a,b]^2 \\ b + (1-b) \cdot T_3(\frac{x-b}{1-b}, \frac{y-b}{1-b}), & \text{if } (x,y) \in [b,1]^2 \\ \min(x,y), & \text{otherwise,} \end{cases}$$

where $T_1, T_2, T_3 \in F(W) \cup F(Prod)$. The choice of these t-norms depend on each situation. For example if $\mu_{P\&P}(x) > 0$ when $\mu_P(x) > 0$, it is $T_1, T_2, T_3 \notin F(W)$.

Taking $T_1, T_2, T_3 = Prod$, $a = 0.3$, $b = 0.7$ and $c = 0.8$, it is gotten the surface in figure(1).

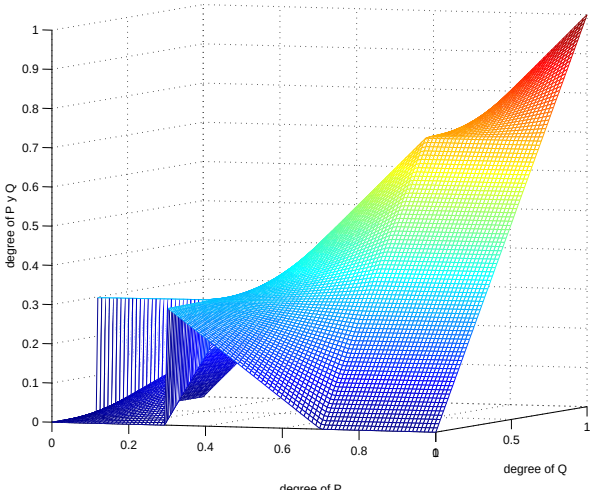


Figure 1:

To collect the idempotency of a predicate P, we pay attention only to the restriction of T to the main diagonal $\{(x, x) : x \in [0, 1]\}$. So, the t-norm can be shortened by an strictly monotonic curve described by the pairs $(\mu_P(x), T(\mu_P(x), \mu_P(x)))$ (figure 2).

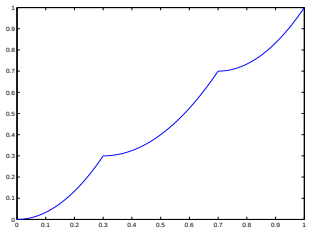


Figure 2:

3.2 A case with just one idempotent value

Let us consider the following semantical situation in which 'x is P' with numerical degrees of truth $\mu_P(x) \in [0,1]$, and the interpretation,

- 'x is P' is false, if $\mu_P(x) = 0$
- 'x is P' is true, if $\mu_P(x) = 1$
- 'x is P' is more false than true, if $\mu_P(x) \in (0, 0.5)$
- 'x is P' is more true than false, if $\mu_P(x) \in (0.5, 1)$
- 'x is P' is semantically undefined, if $\mu_P(x) = 0.5$

In this sense, it seems reasonable that when two statements 'x is P'and 'x is Q' do verify $\mu_P(x) = 0.5$ and $\mu_Q(x) = 0.5$, the decision $\mu_{P\&Q}(x) = 0.5$ will be taken. In this case, and in a standard algebra of fuzzy sets given by a triplet (T, S, N), it is reasonable to consider,

1. $\mu_{P\&Q}(x) = T(\mu_P(x), \mu_Q(x)) = T(0.5, 0.5) = 0.5$, that is, that the conjunctive statement 'x is $P\&Q$' is also semantically undefined.
2. If $\mu_P(x)$ and $\mu_Q(x)$ are a little greater than 0.5, for example, if $\mu_P(x) = \mu_Q(x) = 0.5001$, there is not enough evidence to consider $\mu_{P\&Q}(x) \neq 0.5$.
3. If both statements have a degree near 0, then $\mu_{P\&Q}(x) = 0$, for example if $\mu_P(x) = \mu_Q(x) = 0.001$, then $\mu_{P\&Q}(x) = 0$.

These last items warrant the existence of some non-idempotent points, as well as that 0.5 is idempotent. Then, in a standard algebra, the only possibility for T is to be an ordinal sum of archimedean t-norms, for which 0.5 is the only idempotente value. Hence, the family of open non-pairwise disjoint intervals should be $\{(0, 0.5), (0.5, 1)\}$. Then,

$$T(x,y) = \begin{cases} 0.5 \cdot T_1(\frac{x}{0.5}, \frac{y}{0.5}), & \text{if } (x,y) \in [0,0.5]^2 \\ 0.5 + (0.5) \cdot T_2(\frac{x-0.5}{0.5}, \frac{y-0.5}{0.5}), & \text{if } (x,y) \in [0.5,1]^2 \\ \min(x,y), & \text{otherwise} \end{cases}$$

To get $T(x,x) = 0.5$ with x a little great that 0.5, an to get $T(x,x) = 0$ for x closer to 0, it cannot be $T_1, T_2 \in F(prod)$. So, T_1 and T_2 should be in $F(W)$, and a solution is reached with $\varphi = id_{[0,1]}$,

$$T(x,y) = \begin{cases} 0.5 \cdot \max(\frac{x}{0.5} + \frac{y}{0.5} - 1, 0), & \text{if } (x,y) \in [0,0.5]^2 \\ 0.5 + (0.5) \cdot \max(\frac{x-0.5}{0.5} + \frac{y-0.5}{0.5} - 1, 0), & \text{if } (x,y) \in [0.5,1]^2 \\ \min(x,y), & \text{otherwise} \end{cases}$$

$$T(x,y) = \begin{cases} 0.5 \cdot \max(\frac{x}{0.5} + \frac{y}{0.5} - 1, 0), & \text{if } (x,y) \in [0,0.5]^2 \\ 0.5 + (0.5) \cdot \max(\frac{x-0.5}{0.5} + \frac{y-0.5}{0.5} - 1, 0), & \text{if } (x,y) \in [0.5,1]^2 \\ \min(x,y), & \text{otherwise} \end{cases}$$

Figure(3) is the graphic representation of this t-norm T, and it shows that $T(x,y) = 0$ for $0 \leq x + y \leq 0.5$ and $T(x,y) = 0.5$ for $1 \leq x + y \leq 1.5$.

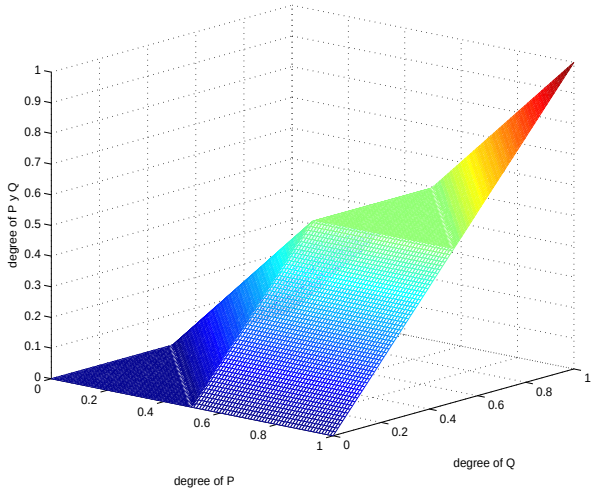


Figure 3:

4 Conclusion

This paper broaches some cases in which ordinal sums of archimedean t-norms are useful.
These t-norms allow to have a set of finite or infinite numerable points that are idempotent, something that does not happen with the t-norm *min*, for which all points in $[0,1]$ are idempotent, nor with t-norms in $F(W)$ and $F(Prod)$ that have not any idempotent point different from 0 or 1.

Acknowledgement

Authors thanks Dr. Sergio Guadarrama (ECSC) for his help in the preparation of this paper.

Bibliography

[AFS06] C. Alsina, M. J. Frank, and B. Schweizer. *Associative Functions. Triangular Norms and Copulas*. World Scientific, Singapore, 2006.

[KMP00] E.P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Trends in Logic. Kluwer Academic Publishers, Dordrecht, 2000.

[PTGR07] A. Pradera, E. Trillas, S. Guadarrama, and E. Renedo. *On fuzzy set theories*, chapter 2, pages 15–47. Springer, 2007.

Uso de Ontologías para facilitar las tareas de Extracción y Representación de Conocimiento en el diseño de Sistemas Basados en Reglas Borrosas

José M. Alonso[1], Andrés Muñoz[2], Juan Botía[2], Luis Magdalena[1], Antonio F. Gómez-Skarmeta[2]

[1] European Centre for Soft Computing, {jose.alonso,luis.magdalena}@softcomputing.es
[2] Universidad de Murcia, {amunoz,juanbot,skarmeta}@um.es

Resumen

El modelado borroso, en su etapa de elicitación de conocimiento, necesita forzosamente de un experto en lógica borrosa. Esto se debe a que es ineludible, desde el principio, manejar correctamente las etiquetas lingüísticas con las que se definen las magnitudes a considerar en las reglas que pasarán a ser los correspondientes modelos borrosos del sistema. Con este trabajo pretendemos relegar la actividad de dicho experto a etapas posteriores, donde es realmente necesario. Para ello presentamos una herramienta que permite, mediante la representación visual de los conceptos a usar en las reglas, un diseño intuitivo de bases de reglas borrosas expertas, haciendo transparente para el diseñador el hecho de que están basadas en el formalismo de la lógica borrosa.

Palabras Clave: Modelado borroso, ontologías, herramientas software de libre distribución.

1 INTRODUCCIÓN

En los últimos años se ha extendido el uso del término Web Semántica [5] en contraposición a Web Sintáctica, que hace referencia a la web actual. Esta última está pensada y diseñada para que la información sea entendida fácilmente por personas, pero la forma en que se estructura dicha información impide que sea procesable automáticamente por máquinas. Esto dificulta la gestión de la ingente cantidad de datos disponible. En la Web Semántica, en cambio, se etiquetan las páginas web de acuerdo con convenios estandarizados para eliminar ambigüedades y hacer posible la aplicación de técnicas de inteligencia artificial para efectuar razonamientos lógicos sobre la información.

Los sistemas de razonamiento clásicos sólo manejan dos valores de verdad (1/0, todo/nada, verdadero/falso, etc.). Desgraciadamente, la mayoría de conceptos representados en la Web Semántica, como la mayoría de situaciones de la vida real, involucran imprecisión e incertidumbre, y en consecuencia no son manejables por sistemas deductivos clásicos. Por ejemplo, supóngase un escenario representado por técnicas de la Web Semántica en una aplicación domótica donde ciertos dispositivos de aire acondicionado en un edificio son controlados por un sistema inteligente [14]. El comportamiento de este sistema establece que si la temperatura es mayor de 28°C, se deben encender los dispositivos; en cambio, cuando la temperatura es menor de 27°C, el aire acondicionado se apaga. Si la temperatura en el edificio oscila entre 26°C y 29°C debido a este comportamiento, los dispositivos estarán incesantemente enciendiéndose y apagándose. Con la introducción de imprecisión se podría evitar este problema, como se verá a continuación.

En 1965, Zadeh presentó las bases de la Teoría de Conjuntos Borrosos [19], que incluye a la teoría de conjuntos clásica como un caso particular, y estableció una relación directa entre los conceptos de función de pertenencia y grado de verdad característicos de los conjuntos borrosos, y una correspondencia entre las operaciones de conjuntos y los operadores lógicos. La sintaxis de esta lógica amplía la clásica con la introducción de modificadores lingüísticos (*muy*, *algo*, etc.) y la posibilidad de definir hechos y funciones borrosas, a partir de los cuales el sistema es capaz de realizar inferencias. Uno de los objetivos de esta lógica es proporcionar las bases del razonamiento aproximado que hace posible manejar conceptos con imprecisión e incertidumbre. Siguiendo con el ejemplo anterior, los comportamientos se definirían ahora como *Si la temperatura es* alta*, encender aire acondicionado* y *Si la temperatura es* normal*, apagar aire acondicionado.*

Las etiquetas lingüísticas *alta* y *normal* definen un rango de temperaturas impreciso que evita el constante encendido/apagado de los dispositivos. El concepto de conjunto borroso parte de la idea de representar conceptos imprecisos, muy comunes en el lenguaje natural y que se pueden manejar con la lógica borrosa, pero que son inmanejables con la lógica clásica. A la hora de resolver problemas, las personas razonamos utilizando prototipos y etiquetas lingüísticas con las que intentamos modelar y describir la realidad.

En el modelado de sistemas se manejan dos tipos de conocimiento: Experto e inducido. El conocimiento experto sobre un problema específico es difícil de formalizar, ya que parte del mismo puede residir en el subconsciente de la persona. Como resultado, la adquisición de conocimiento llega a ser un proceso tedioso, difícil de estructurar y automatizar, se convierte en un cuello de botella [10]. A su vez, el conocimiento inducido consiste en conocimiento extraído de forma automática a partir de datos (*Machine learning*). Obviamente, el experto sigue siendo necesario, especialmente en la generación de los datos, y también en la validación del sistema, pero su participación se reduce significativamente. En algunas aplicaciones en las que los datos se generan de forma automática sólo se acude al experto para evaluar el resultado final. No obstante, la adquisición de datos experimentales es costosa, en tiempo y dinero, ya que es necesario repetir gran cantidad de experimentos bajo las mismas condiciones y esto no siempre es factible.

La expresividad semántica de la lógica borrosa, mediante el uso de reglas [13] y variables lingüísticas [20], es muy próxima al lenguaje natural empleado por los humanos. Esto ayuda en la tarea de adquisición de conocimiento a realizar como parte del proceso de modelado borroso, es decir, el modelado de Sistemas Basados en Reglas Borrosas. Pero además, tanto el conocimiento procedente de expertos, como aquel que se puede extraer de datos, se pueden formalizar de acuerdo con un mismo tipo de representación. Esto hace posible que los procesos a realizar sobre la base de conocimiento (análisis, simplificación, optimización, etc.) sean independientes del mecanismo utilizado para extraer dicho conocimiento.

En las etapas iniciales del proceso de modelado el conocimiento experto debería permanecer en un alto nivel de abstracción, quedando ocultos los detalles relativos a la lógica borrosa. Téngase en cuenta que el experto es la persona que más sabe sobre el problema a modelar, es un experto de dominio, pero no tiene por qué saber nada de lógica borrosa. Para ello, el uso de una representación visual de los conceptos a usar en las reglas permitirá un diseño intuitivo y trasparente.

El objetivo de este trabajo pretende mostrar cómo se pueden combinar dos herramientas que han sido diseñadas para dar respuestas a los desafíos del modelado de conocimiento, tanto en el campo de la Web Semántica, como en el de modelado borroso. Por un lado, se muestra un sistema de autoría de reglas sobre ontologías de Web Semántica para modelar comportamientos como los del escenario del aire acondicionado. Por otro, se describe una metodología para construir sistemas borrosos robustos de forma intuitiva. Combinando ambos elementos, se intenta obtener un modelo borroso basado en Web Semántica.

El resto del artículo se ha estructurado de la siguiente forma. La sección 2 presenta una herramienta para editar reglas (no borrosas) a partir de conceptos definidos en ontologías. La sección 3 describe una herramienta para modelado borroso. A continuación, la sección 4 explica, a través de un ejemplo, cómo la combinación de ambas herramientas permite hacer más sencilla la tarea de extracción y representación de conocimiento en el proceso de modelado borroso. En la sección 5 se hace una revisión de herramientas que trabajan con ontologías, modelado borroso, y ontologías borrosas. Por último, la sección 6 cierra el artículo con unas breves conclusiones, apuntando trabajos futuros.

2 MODELADO VISUAL BASADO EN WEB SEMÁNTICA

El ciclo de vida para generar reglas que modelan conocimiento [4] está compuesto por las etapas de edición, testeo, depuración y validación de las reglas. El experto sobre el dominio debe estar presente en todas ellas, con la consiguiente carga de tiempo que implica todo el proceso de generación. Esta carga se ve altamente agravada cuando la base de conocimiento, las reglas definidas o ambas aumentan. Además, programar las reglas a mano no es una tarea sencilla cuando el usuario del sistema no está familiarizado con este tipo de programación.

Teniendo en cuenta esta dificultad, se puede justificar la necesidad de disponer de herramientas que permitan controlar el ciclo de vida de las reglas de conocimiento, implementando el desarrollo de cada etapa de una forma sencilla e intuitiva. Las herramientas capaces de abstraer al usuario de la lógica subyacente y del modelo interno de programación de las reglas son especialmente interesantes. Es decir, la única habilidad que el usuario necesita poseer es un aceptable nivel de conocimiento del dominio y de cómo está representado, en este caso por medio de RDF Triples [12]. RDF es un lenguaje de modelado basado en XML, y utiliza el formato de triplas (s, p, o) para representar información sobre un dominio, donde s es el

sujeto que representa cualquier recurso del dominio, *p* es un *predicado* que expresa una propiedad sobre el sujeto, y *o* es el *objeto* que contiene el valor de la propiedad. Estos objetos pueden ser a su vez otros recursos o valores de datos básicos. Todos los elementos de la tripla están precedidos por una URI que los identifica unívocamente. El ejemplo de la figura 1 utiliza el formato de grafos RDF para describir una habitación de un edificio que contiene un dispositivo de aire acondicionado y una temperatura de 28°C. Como puede observarse, las elipses representan recursos, bien como sujetos (*#habitación3*) o bien como objetos (*#AireAcondicionadoLG*), mientras que las flechas dirigidas denotan predicados, relacionando recursos entre sí. Por otra parte, un cuadrado que encierra un valor significa que es un tipo de dato básico. La URI *"http://www.um.es/FI"* desgina el dominio de una Facultad de Informática, y *"http://www.w3.org/...rdf-syntax-ns"* es el esquema de la ontología RDF, que en este caso se utiliza para especificar el tipo del sujeto de la tripla.

Figura 1: El recurso *habitación 3* es el sujeto de las triplas RDF, y los predicados *tipo, tiene* y *temperatura* toman valores en el concepto *Habitación*, el recurso *AireAcondicionadoLG* y el valor de datos 28°C, respectivamente.

El resultado de estas consideraciones es ORE [14], una herramienta diseñada para cubrir el ciclo de generación de reglas ontológicas, es decir, reglas configuradas por elementos que forman parte de una ontología de la Web Semántica. Sus características más resaltables son la *sencillez* a la hora de manejar la herramienta, mediante el uso de interfaces gráficas intuitivas para definir reglas con la técnica *drag and drop*, y su *versatilidad* para trabajar con distintos tipos de razonadores genéricos sobre los que ejecutar las reglas. ORE[1] está basado en componentes de la Web Semántica, tales como OWL [9] y SWRL [11], los lenguajes que representan el conocimiento sobre el dominio y las reglas ontológicas, respectivamente. Por otro lado, ORE puede utilizar diversos motores de inferencia, que son los encargados de ejecutar las reglas definidas sobre la base de conocimiento. En este trabajo hemos seleccionado Jena [8] y Pellet [17] como razonadores para llevar cabo los procesos de inferencia sobre ontologías y reglas.

[1]http://sourceforge.net/projects/ore/

OWL es un lenguaje basado en RDF para modelar ontologías, más potente que éste pero siguiendo el esquema de triplas anterior. El objetivo principal de OWL está orientado a ofrecer una representación de cualquier dominio de forma estandarizada y formal para que pueda ser automáticamente interpretada por computadoras e integrada en la Web, permitiendo que este conocimiento sea compartido fácilmente. Además de describir los conceptos de un dominio y las relaciones entre éstos, es posible explicitar ciertos detalles que refinan ambos elementos. Por ejemplo, se pueden establecer jerarquías de especialización/generalización entre conceptos, expresar conceptos disjuntos, o bien establecer restricciones de cardinalidad entre las relaciones.

El uso de un lenguaje formal en las ontologías OWL tiene un valor añadido al permitir procesos de inferencia sobre el dominio modelado. En esta propuesta se ha adoptado la versión OWL-DL, basada en Lógica de Descripción (DL) [3]. Básicamente existen dos tipos de razonamiento aplicables: el razonamiento *axiomático* y el razonamiento *basado en reglas*. El primero extrae información oculta (o que no ha sido explicitada) del esquema de la ontología formado por los conceptos, relaciones y restricciones. Por ejemplo, si el concepto B es subclase de A, y el concepto C es subclase de B, pero no existe ninguna relación entre A y C, este tipo de razonamiento es capaz de concluir que C también es subclase de A. El segundo tipo de razonamiento deriva nuevos hechos a partir de un conjunto de reglas formadas por elementos ontológicos, como veremos a continuación, y un proceso de inferencia deductivo. Tomando el ejemplo del aire acondicionado de la introducción, las siguientes expresiones de Lógica de Descripción definen los conceptos *Habitación* y *DispAC* (dispositivo de aire acondicionado) como subclases de *Espacio_Habitable* y *Dispostivo*, respectivamente, junto con sus propiedades:

$$Habitación \sqsubseteq Espacio_Habitable \qquad \exists tiene.Dispositivo \quad \exists temperatura.Entero$$

$$DispAC \sqsubseteq Dispositivo \qquad \exists estado.String$$

Por otra parte, SWRL (Semantic Web Rule Language) es el lenguaje utilizado en ORE para expresar reglas ontológicas. Este lenguaje abstracto es una combinación de OWL-DL y RuleML, confiriendo de esta manera una mayor facilidad para construir reglas a partir de elementos OWL. El formato de las reglas SWRL consiste en una implicación entre una con-

junción de antecedentes y una conjunción de consecuentes. Ambas conjunciones están formadas por los siguientes *átomos* básicos: Predicados unarios ($C(x)$) o binarios ($P(x, y)$) de la ontología y *Built-ins*, donde C es un concepto, P representa relaciones entre conceptos, y x, y son variables, instancias de conceptos o valores de datos. Las *Built-ins* son funciones que ofrecen diferentes operaciones de tipo matemático, de comparación, etc. Cuando todos los átomos del antecedente se cumplen, entonces se deviran los hechos del consecuente, que pueden ser añadidos finalmente al dominio. A continuación se muestra un ejemplo de regla expresada en SWRL a partir del escenario del aire acondicionado y los elementos OWL definidos anteriormente. Esta regla enciende el aire acondicionado si la temperatura de la habitación es mayor de 28°C. Observar que la función comparativa *mayorQue* es una *Built-in* utilizada en el antecedente de la regla.

$$Habitacion(?h) \wedge DispAC(?x) \wedge tiene(?h, ?x) \wedge temperatura(?h, ?t) \wedge mayorQue(?t, 28) \Rightarrow estado(?x, \text{“}encendido\text{”})$$

Esta regla se puede interpretar en forma de regla lingüística: **Si** una *Habitación* tiene un *DispAC* y su *temperatura* ambiente es *mayorQue* 28°C, entonces el *DispAC* debe estar *“encendido”*.

La utilidad de ORE en este proceso radica en mostrar visualmente la ontología que modela el dominio de forma que puedan editarse reglas de forma intuitiva, arrastrando los elementos OWL que la componen como átomos básicos de los antecedentes o consecuentes. Una vez editadas, las reglas se pueden testear, depurar y validar en ORE de forma visual siguiendo una serie de pasos guiados.

3 MODELADO BORROSO

El modelado borroso ha sido una línea de investigación muy activa durante más de 40 años. Entre 1965 y 1990, la atención se centró en construir modelos formados por pequeños grupos de reglas lingüísticas sencillas a partir de conocimiento experto. Sin embargo, para abordar problemas complejos el conocimiento experto no es suficiente. Por ello, la incorporación al modelado borroso de técnicas de aprendizaje automático a partir de datos se produjo de forma natural entre 1990 y 2000. Con ello cambió el tipo de sistemas construidos, se pasó a generar sistemas formados por muchas reglas complejas capaces de alcanzar una gran precisión. Pero a partir del año 2000, los ingenieros del conocimiento se dieron cuenta de que era necesario echar la vista atrás y recuperar la interpretabilidad de los sistemas borrosos. Como resultado, actualmente se plantea un nuevo enfoque, consistente en construir sistemas precisos e interpretables que combinen conocimiento experto y conocimiento extraído de forma automática a partir de datos experimentales.

Siguiendo este razonamiento, HILK (*Highly Interpretable Linguistic Knowledge*) es una nueva metodología concebida para construir sistemas borrosos compactos y robustos [2]. Su objetivo es conseguir que el resultado final sea lo más fácil de interpretar posible, sin descuidar la precisión. Para ello se combinan dos formas de conocimiento, experto e inducido, claramente diferenciadas. Se pretende que el conocimiento inducido sirva para completar el conocimiento proporcionado por un experto. Al combinar ambas fuentes de conocimiento nos podemos encontrar con dos problemas principales. En primer lugar pueden aparecer problemas de coherencia (redundancia, contradicción, etc.) que es necesario analizar y resolver. En segundo lugar, el tamaño de la base de conocimiento crece y esto empeora la interpretabilidad. Para resolver este problema se propone un algoritmo de simplificación, cuyo objetivo es convertir la base de conocimiento resultante de la integración en un sistema más interpretable. Esto se hace manteniendo, y si es posible aumentando, la precisión. A continuación, se analizan varias estrategias de optimización que puedan incrementar la precisión, sin alterar la interpretabilidad. Por último, se estudia la completitud de la base de reglas obtenida. En resumen, los sistemas construidos se caracterizan por:

1. Semántica global. El primer paso consiste en caracterizar las variables de entrada y salida del sistema, mediante el uso de términos lingüísticos y particiones borrosas fuertes.
2. Definición del comportamiento del sistema a través de reglas lingüísticas de la forma **Si** *condición* **Entonces** *conclusión*. Formalmente estas reglas son equivalentes a las manejadas por ORE. Sin embargo, en este caso tanto las condiciones como las conclusiones consisten en proposiciones lingüísticas de la forma X *es* A. Como todas las reglas usan los mismos términos lingüísticos definidos por los mismos conjuntos borrosos, la comparación de reglas se puede hacer a nivel lingüístico para detectar y corregir redundancias e incoherencias.

El nuevo proceso de modelado propuesto en [2] está implementado en KBCT[2] (*Knowledge Base Configuration Tool*) [1], una herramienta software distribuida como código abierto bajo licencia GPL. Es una herramienta multiplataforma, con una interfaz gráfica

[2] http://www.mat.upm.es/projects/advocate/kbct.htm

amigable, para facilitar el proceso de extracción y representación de conocimiento en sistemas basados en lógica borrosa. El conocimiento se formaliza en varios niveles de abstracción permitiendo un uso intuitivo tanto por usuarios noveles, con conocimientos reducidos de lógica borrosa como por usuarios expertos. La herramienta es capaz de construir particiones y reglas de forma automática a partir de datos, y su posterior integración con el conocimiento extraído de expertos de dominio. El usuario es quien decide el nivel de detalle al que quiere interactuar.

4 COMBINANDO EL MODELADO BASADO EN WEB SEMÁNTICA Y EL MODELADO BORROSO

Los dos enfoques presentados en las secciones previas se pueden combinar con dos objetivos:

1. **Introducir el uso de ontologías en las fases iniciales del modelado borroso**. Para ello se puede usar ORE para definir una base de conocimiento borrosa experta (sencilla) que se pueda usar con KBCT para añadir conocimiento inducido. Serviría para definir las variables de entrada/salida y algunas reglas básicas.

2. **Introducir la lógica borrosa en la Web Semántica para manejar incertidumbre y conceptos imprecisos, pudiendo realizar razonamientos aproximados**. Esto se puede conseguir incorporando a ORE un motor de inferencias borroso, que pueda usar bases de conocimiento construidas con KBCT, de forma que ORE pueda hacer razonamientos de tipo aproximado.

En este enfoque van a convivir tanto reglas que contienen conocimiento borroso como reglas que sólo contienen conocimiento del dominio. A las reglas que no contienen conocimiento borroso se les conoce como *reglas crisp*. Igualmente, al razonamiento que únicamente utiliza este tipo de reglas se le llama *razonamiento crisp*.

Ilustraremos la combinación de ORE y KBCT con un ejemplo sencillo: IRIS. Se trata de un problema de clasificación de plantas muy conocido (los datos proceden de la UCI[3], Universidad de California, Irvine) y disponible online. Está perfectamente caracterizado, a partir de cuatro variables de entrada (anchura y longitud de sépalo y pétalo) se pretende clasificar tres tipos de planta iris (setosa, versicolor, virginica).

[3]El conjunto de datos consta de 150 casos extraídos de http://www.ics.uci.edu/-mlearn/MLSummary.html

Supongamos que nos proponen construir un sistema para clasificar los tipos de planta IRIS, y que es la primera vez que se aborda este problema de forma que la información del UCI relativa al mismo no existiese. Los pasos a seguir serían los siguientes (ver figura 2):

1. Buscamos una ontología existente en el dominio de interés de nuestra aplicación. En este caso, clasificación de plantas.

2. La ontología seleccionada se borrosifica, es decir, se combina con una ontología borrosa que permite caracterizar cada concepto con el tipo de variable (categórica, lógica, numérica), el rango (min, max), el número de etiquetas (términos lingüísticos), y el nombre de las etiquetas. También es posible dar valores prototipo, por ejmplo un voltaje habitual son 5V, la temperatura humana habitual es 36°, 39° es una temperatura alta, etc.

3. Acudimos a hablar con un experto en clasificación de plantas IRIS para que nos ayude a formalizar el problema. Lo habitual es concertar una entrevista para que el experto nos describa el problema y nos diga las variables que considera importantes, y si es posible alguna regla básica. Para ello utilizamos ORE que muestra al experto de forma gráfica los conceptos incluidos en la ontología del dominio. Una interfaz visual permite al experto seleccionar los conceptos (y atributos borrosos) sin necesidad de que tenga conocimientos relativos a lenguajes de ontologías (ni lógica borrosa). También permite definir reglas básicas que combinen los conceptos seleccionados previamente para describir el comportamiento global del sistema. En nuestro ejemplo, la información proporcionada por el experto sería la siguiente: Existen 3 tipos de planta IRIS (setosa, versicolor y virginica) y para distinguirlas hay que tener en cuenta cuatro elementos (anchura y longitud de sépalo y pétalo), que a su vez están caracterizadas en la tabla 1.

Tabla 1: Información del experto sobre IRIS.

Variable	**Rango de interés**	**Términos lingüísticos**
Longitud de Sépalo	[4.3, 7.9]	pequeña, media, grande
Anchura de Sépalo	[2, 4.4]	pequeña, media, grande
Longitud de Pétalo	[1, 6.9]	muy pequeña, pequeña, media, grande, muy grande
Anchura de Pétalo	[0.1, 2.5]	muy pequeña, pequeña, media, grande, muy grande
Clase	-	setosa, versicolor, virginica

Adicionalmente el experto nos da 3 reglas muy sencillas:

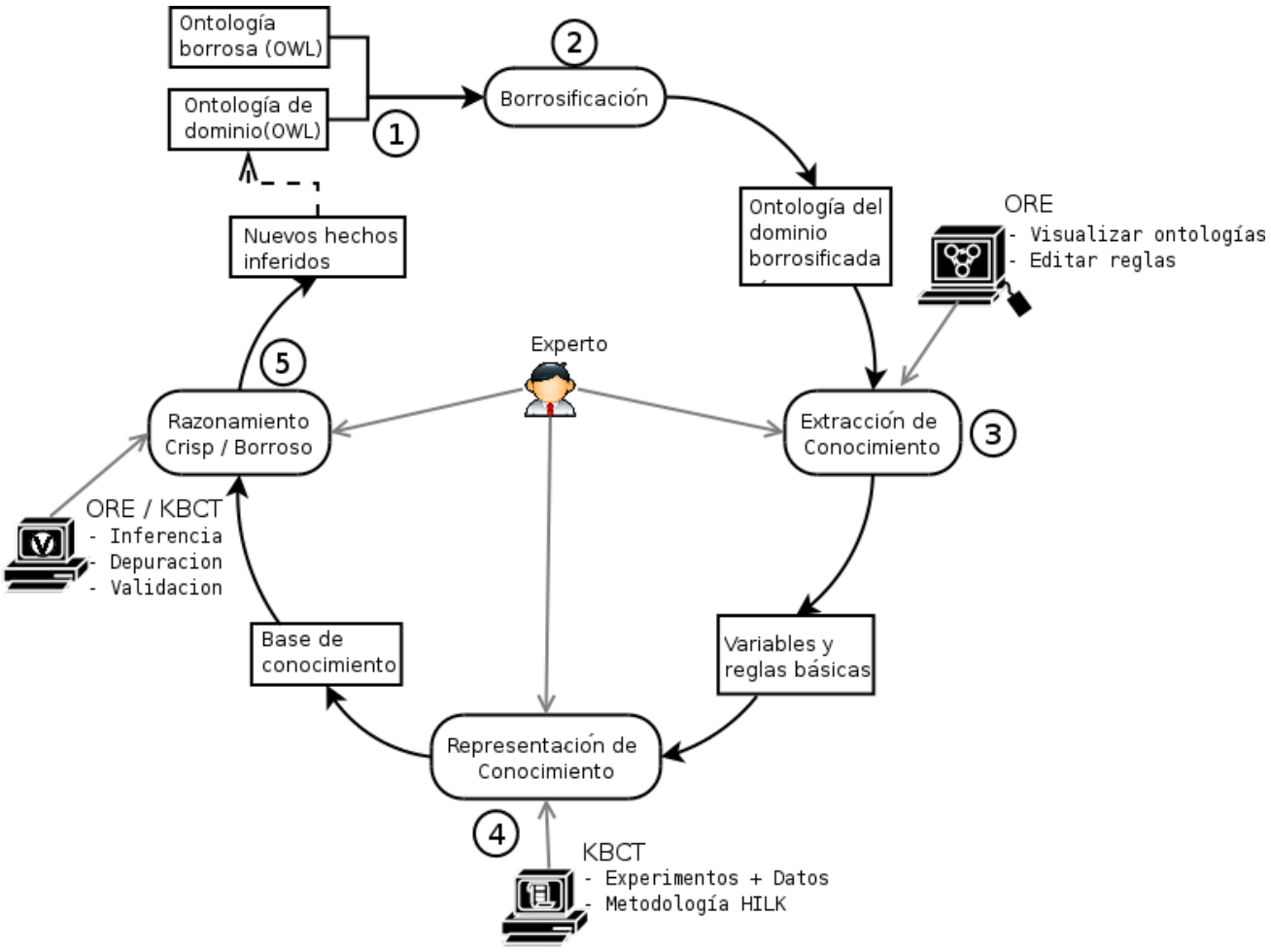


Figura 2: Metodología para combinar ontologías de la Web Semántica con modelado borroso. El último paso realiza razonamiento aproximado con ORE y KBCT.

R1: **Si** Anchura de Pétalo es muy pequeña
Entonces IRIS es setosa

R2: **Si** Anchura de Pétalo es media
Entonces IRIS es versicolor

R3: **Si** Anchura de Pétalo es grande o muy grande **y** Longitud de Pétalo es grande o muy grande
Entonces IRIS es virginica

Aunque el experto considera que las cuatro variables son necesarias, sólo usa en las reglas las 2 que considera más importantes, quizás porque no sabe cómo formalizar la interacción de todas ellas.

4. Con la información capturada usando ORE podemos inicializar una base de conocimiento experta que será refinada y completada con KBCT. El primer paso consiste en utilizar las indicaciones del experto para realizar experimentos que nos proporcionen datos significativos. En nuestro ejemplo habría que tomar varias muestras de cada tipo de planta y clasificarlas. Supongamos que como resultado obtenemos los 150 casos clasificados que hemos obtenido de la UCI. A partir de estos datos los algoritmos implementados en KBCT permiten construir reglas de forma automática que se pueden combinar con las definidas por el experto.

5. La base de conocimiento final construida con KBCT puede ser utilizada por ORE o KBCT para realizar razonamientos de tipo aproximado. En este tipo de razonamiento es posible utilizar ya tanto reglas *crisp* como reglas borrosas. Por último, los nuevos hechos generados (en este caso, las etiquetas de clasificación de las plantas) pueden ser añadidos a la ontología del dominio para completar el ciclo (representado en la figura 2 mediante la flecha discontinua).

5 TRABAJOS RELACIONADOS

Con el desarrollo de la Web Semántica han ido apareciendo diversas herramientas para trabajar sobre ontologías. Protégé [15] es una herramienta bastante conocida que permite al usuario definir ontologías a través de un plug-in OWL. Estas ontologías pueden ser exportadas a varios formatos, incluyendo RDF(S), OWL, y XML Schema. La edición y ejecución de reglas SWRL también puede llevarse a cabo con esta herramienta, sin embargo el ciclo de vida de generación de reglas no está suficientemente guiado y es extremadamente difícil de usar sin tener un conocimiento lógico y de modelos de ontología profundo. Protégé puede ser considerado como una herramienta útil y efectiva para la etapa de modelado del conocimiento adquirido en el dominio. Las ontologías resultantes de este proceso son una posible entrada a ORE para editar, depurar y validar reglas. De esta forma, ambas herramientas pueden ser usadas conjuntamente para gestionar el ciclo completo de extracción y generación de conocimiento.

Por otra parte, SweetRules [16] es un conjunto integrado de herramientas pensado para procesar ontologías y reglas ontológicas. Los servicios ofrecidos incluyen análisis de ontologías, inferencia y traducción de reglas a diversos formatos. A pesar de que el proceso de generación de reglas está contemplado en SweetRules, sólo se centra en la edición de reglas, sin un proceso guiado para los procesos de testeo o depuración como en ORE.

Por último hay que reseñar que la introducción de la lógica borrosa en la Web Semántica ha despertado gran expectación por el potencial que encierra. Se ha intentado flexibilizar la representación de conceptos mediante la definición de ontologías borrosas [6]. Sin embargo, una vez estudiados los aspectos teóricos para generar ontologías borrosas de forma automática [18], falta extender los editores y razonadores. Como parte del proyecto KAON [7] se está desarrollando un editor para ontologías borrosas. No obstante, más allá de la introducción de la lógica borrosa en la Web Semántica la principal novedad de nuestro trabajo es proponer el enfoque contrario: *Aprovechar la representación de conceptos mediante ontologías para hacer más sencilla la fase de adquisición de conocimiento de los modelos borrosos tradicionales.*

6 CONCLUSIONES Y TRABAJO FUTURO

El uso de lógica borrosa favorece la interpretabilidad del modelo construido y reduce el esfuerzo de extracción y representación de conocimiento, que normalmente es una tarea tediosa y constituye un cuello de botella en el proceso de modelado del sistema. En las primeras fases del proceso es recomendable ocultar al experto de dominio la complejidad de los detalles asociados a la lógica borrosa. Para ello la representación visual de los conceptos a usar en las reglas permite un diseño intuitivo y trasparente. En este artículo hemos introducido la combinación de dos herramientas, ORE y KBCT, orientadas a facilitar el modelado borroso en este sentido. La primera permite el modelado temprano de las reglas borrosas, visualmente, y la segunda consigue unas reglas más aproximadas al problema, sin abandonar la interpretabilidad inicial. Ambas herramientas se distribuyen en forma de software libre.

En la actualidad, estamos avanzando en la integración de ambas herramientas mediante la investigación de las implicaciones derivadas de plantear una inferencia borrosa visualmente, de tal forma que se enlace el resultado de este proceso con los conceptos disponibles en la ontología de dominio. Esto ayudará a interpretar adecuadamente el proceso de razonamiento borroso. Lo que, a su vez, contribuye a la construcción de conjuntos de reglas más interpretables.

Agradecimientos

Trabajos realizados en el contexto del Programa de Ayuda a los grupos de excelencia de la Fundación Séneca 04552/GERM/06 y gracias al Ministerio de Ciencia e Innovación mediante la beca AP2006-4154 del programa FPU.

Referencias

[1] José M. Alonso, L. Magdalena, and S. Guillaume. KBCT: Una herramienta para adquisición y representación de conocimiento usando lógica fuzzy. In *ESTYLF04*, pages 563–568, 2004.

[2] José M. Alonso, L. Magdalena, and S. Guillaume. HILK: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *International Journal of Intelligent Systems*, In Press 2008.

[3] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic handbook: theory, implementation, and applications.* Cambridge University Press, New York, NY, USA, 2003.

[4] Chitta Baral and Michael Gelfond. Logic programming and knowledge representation. *Journal of Logic Programming*, 19:73–148, 1994.

[5] Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. *Scientific American*, 2001.

[6] Ignacio J. Blanco, N. Marín, C. Martínez-Cruz, and M. Amparo Vila. About the use of ontologies for fuzzy knowledge representation. In *Joint EUSFLAT-LFA05*, pages 100–105, 2005.

[7] S. Calegari and D. Ciucci. Fuzzy ontology and Fuzzy-OWL in the KAON project. In *FUZZ-IEEE07*, pages 1420–1425, 2007.

[8] Jeremy J. Carroll, Ian Dickinson, Chris Dollin, Dave Reynolds, Andy Seaborne, and Kevin Wilkinson. Jena: implementing the Semantic Web recommendations. In *Proceedings of the 13th international World Wide Web conference*, pages 74–83. ACM Press, 2004.

[9] Mike Dean, Dan Connoll, Frank van Harmelen, James Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. Web ontology language (OWL). Technical report, W3C, 2004.

[10] F. Hayes-Roth, D. A. Waterman, and D. B. Lenat. *Building expert systems*. Addison-Wesley, 1983.

[11] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. SWRL: A Semantic Web rule language combining OWL and RuleML. Technical report, W3C, 2004.

[12] Graham Klyne and Jeremy J. Carroll. Resource Description Framework (RDF): Concepts and Abstract Syntax. *http://www.w3.org/TR/rdf-concepts/*, 2004.

[13] E. H. Mamdani. Application of fuzzy logic to approximate reasoning using linguistic systems. *IEEE Trans. on Computers*, 26(12):1182–1191, 1977.

[14] Andrés Muñoz, Antonia Vera, Juan A. Botía, and Antonio F. Gómez Skarmeta. Defining basic behaviours in ambient intelligence environments by means of rule-based programming with visual tools. In J. C. Augusto, editor, *1st Workshop of Artificial Intelligence Techniques for Ambient Intelligence. ECAI*, 2006.

[15] N. F. Noy, M. Sintek, S. Decker, M. Crubezy, R. W. Fergerson, and M. A. Musen. Creating Semantic Web contents with protege-2000. *IEEE Intelligent Systems*, 16(2):60–71, 2001.

[16] Martin O'Connor, Holger Knublauch, Samson Tu, Benjamin N. Grosof, Mike Dean, William Grosso, and Mark Musen. Supporting rule system interoperability on the Semantic Web with SWRL. In *Proceeding of 4th International Semantic Web Conference ISWC*, Ireland, Nov 2005.

[17] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics*, 5(2), 2007.

[18] Quan Thanh Tho, Siu Cheung Hui, A.C.M. Fong, and Tru Hoang Cao. Automatic fuzzy ontology generation for semantic web. *IEEE Trans. on Knowledge and Data Engineering*, 18(6):842–856, 2006.

[19] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[20] Lotfi A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Parts I, II, and III. Information Sciences*, 8, 8, 9:199–249, 301–357, 43–80, 1975.

A Decision Support System Based on Mobile Internet

I.J. Pérez[1] **S. Alonso**[2] **F.J. Cabrerizo**[1] **E. Herrera-Viedma**[1]

[1] Dept. of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain
{ijperez,cabrerizo,viedma }@decsai.ugr.es
[2] Dept. of Software Engineering, University of Granada, 18071 Granada, Spain
zerjioi@ugr.es

Abstract

The aim of this paper is to present a model of Mobile Decision Support System where the preferences about alternatives can be provided, by the experts using Mobile Internet devices such as mobile phones or PDAs. Experts' preferences can be represented using different preference elements: preference orderings, fuzzy preference relations, multiplicative preference relations and utility functions. A prototype of mobile decision support system has been developed to automatically support the decision process. As the decision support system is developed to make use of mobile technologies, the decision process is improved and can be carried out faster as it can be used anytime and anywhere.

Keywords: Group decision making, mobile internet, decision support system.

1 Introduction

A decision making process, consisting in deriving the best option from a feasible set, is present in just about every conceivable human task. As a result, the study of decision making is necessary and very important not only in Decision Theory but also in areas such as Management Science, Operations Research, Politics, Social Psychology, Artificial Intelligence, Soft Computing and so on.

It is obvious that the comparison of different actions according to their desirability in decision problems, in many cases cannot be done, by using a single criterion or an unique person. Thus, we interpret the decision process in the framework of group decision making (GDM) [11]. This has led to numerous evaluation schemes, and has become a major concern of research in decision making. Several authors have provided interesting results on GDM with the help of fuzzy theory, for example [6, 9, 11].

The central goal of decision support systems (DSS) is to process and provide suitable information in order to support individuals or organizations in decision making [7]. Therefore, appropriate mobile DSS could bridge the gap and provide additional value to users as timely information supply and supported decision making should provide a basis for more successful transactions [12].

The last decade has seen significant advances in the way human interact with technology. Users of computers are no longer constrained to the office desktop, but can have much more flexible access to technology almost anywhere and anytime. The spread of e-services and wireless or mobile devices has facilitated a new way of using computers, increased accessibility to data and, in turn, influenced the way in which users make decisions while on the move, the opportunities for decision support have been extended [5]. Users can make real-time decisions based on the most up-to-date data accessed via wireless devices, such as portable computers, mobile phones, and personal digital assistants (PDAs), which are usually carried all the time, and are not shared with others [2].

The aim of this paper is to present a prototype of mobile DSS to deal automatically with GDM problems based on Mobile Internet technologies. In order to build a flexible framework and to give a larger freedom degree to represent the preferences, we will use a GDM model in which the preferences can be provided in any of these four ways [3, 8]: *preference orderings, fuzzy preference relations, multiplicative preference relations or utility functions.* To solve GDM problems in this framework, firstly, we make the information uniform, using fuzzy preference relations as the main element on the uniform representation of the preferences, and

secondly, we apply the selection process and the consensus process to obtain a final solution with a high consensus degree. This mobile DSS implements the GDM model presented in [10], including some new improvements. Some of these improvements are directly derived from the advantages that mobile technologies provide,(anywhere and anytime use), and will allow the development of more dynamic GDM processes. In fact, at every stage of the decision process, users will be informed with updated data about current stage of the decision process, they will receive recommendations to help them to change their preferences, they will be able to send their updated preferences at any moment, thus improving the user participation in the GDM process.

In order to do this, the paper is set out as follows. The Mobile Internet technologies are presented in Section 2. Section 3 presents the prototype of mobile DSS. And, finally, Section 4 shows our conclusions.

2 Mobile Internet: Advantages and Limitations

Mobile communication systems are characterized by a variety of features. They differ from each other in the degree of their complexity, the level of the offered services and operation costs. The attributes of all mobile communication systems is the mobility of at least one of the connection users and the lack of wired connection of this user's terminal with the remaining part of system.

The Mobile Web refers to the World Wide Web as accessed from mobile devices such as cell phones, PDAs, and other portable gadgets connected to a public network. So, access to web services no longer requires a desktop computer.

With the fast increase of the Internet usage, the growing penetration of wireless devices, and the fast technological innovation, wireless technology shifts the world of wired internet to the wireless Mobile Internet, which is often referred to as M-Internet. The following list shows the different advantages that this technologies can provide.

1. Internet has provided an easy and effective way of delivering information and services to millions of users who are connected to wired network. Evidently, this wired network addresses two major constraints: time and place. These limitations have raised the issue of the mobile internet, which enables users to access information from any place at any moment using a mobile wireless device. The possibility to access to this kind services in wireless environments provides a great mobility to the users. This mobility can increase the productivity increasing agility of some tasks, can allow to save displacements and infrastructure's costs, can improve the business processes, can ease the decision making processes obtaining more dynamic and precise solutions, and even it can improve the offered services.

2. The mobile computing paradigm has several interesting and important applications for business, telecommunications, real-time control systems, remote operations and in accessing the Internet [5, 12, 13].

3. Recently, the fast technological innovation made it possible to provide secure, fast and quality communication through the wireless network. Mobile devices were a few years ago considered a luxury, but have at nowadays become a conventional communication tool. Moreover, the devices that used to deliver limited information are now able to provide a wide range of information and services such as e-mail, banking, entertainment and even games.

However, Mobile Web access today still suffers from some interoperability and usability problems. This is partly due to the small physical size of the screens of mobile devices and partly due to the incompatibility of many mobile devices with both computer operating systems and the format of much of the information available on the Internet.

To make use of M-Internet in the best way, several conditions need to be fulfilled. The first condition is the widespread use of mobile devices that connect individuals to the mobile network and the contents that provide useful information and services to users. In addition, the technological support in terms of speed, communication quality and security are also important in the development of the M-Internet [2].

Some of the limitations that current mobile services have to face are:

1. *Small screen size:* It is difficult or impossible to properly adapt text and graphics prepared for the standard size of a desktop computer screen with current information standards.

2. *Lack of windows:* On mobile web only one page can be displayed at a time, and pages usually can only be viewed in the sequence they were originally accessed.

3. *Navigation:*Usual mobile devices do not use a mouse like pointer, but rather simply an up and

down function for scrolling, thereby limiting the flexibility in navigation.

4. *Format of accessible pages:* Many sites that can be accessed on a desktop cannot on a mobile device. Many devices cannot show pages with a secured connection, Flash or other similar elements, PDFs, or video sites.

5. *Speed:* On most mobile devices, the speed of service is very slow, often slower than dial-up internet access.

6. *Size of messages:* Many devices have limits on the number of characters that can be sent in a single message.

The Mobile Web mainly uses lightweight pages written in Extensible Hypertext or Wireless Markup Language (XHTML) or (WML), to deliver content to mobile devices. However, new tools such as Macromedia's Flash Lite or Sun's J2ME enable the production of richer user interfaces customized for mobile devices.

3 A Prototype of Mobile DSS

While decision support systems have typically been associated with desktop systems and involve considerable processing, the development of new compact and mobile technologies provides new opportunities to develop this kind of DSS over M-Internet [1]. This section briefly describes the prototype of Mobile DSS which implements the GDM model with multiple elements of preference representation defined in [10] and uses the advantages of M-Internet technologies.

3.1 A GDM Problem with Different Elements of Preference Representation

In a GDM problem we have a finite set of alternatives. $X = \{x_1, x_2, \ldots, x_n\}$, $(n \geq 2)$ and these alternatives have to be classified from best to worst, using the information given by a set of experts, $E = \{e_1, e_2, \ldots, e_m\}$, $(m \geq 2)$. As each expert, $e_k \in E$, has his own ideas, attitudes, motivations and personality, it is quite natural to think that different experts will give their preferences in a different way. This leads us to assume that the experts' preferences over the set of alternatives, X, may be represented in different ways. In [3, 8] a GDM model assuming that the experts can present their preferences using any of the following elements of preference representation was presented:

- *A preference ordering of the alternatives.* In this case, an expert, e_k, gives his preferences on X as an individual preference ordering, $O^k = \{o^k(1), ..., o^k(n)\}$, where $o^k(\cdot)$ is a permutation function over the index set, $\{1, ..., n\}$, for the expert, e_k. Therefore, according to this point of view, an ordered vector of alternatives, from best to worst, is given.

- *A fuzzy preference relation.* With this representation, the expert's preferences on X is described by a fuzzy preference relation, $P^k \subset X\text{x}X$, with membership function, $\mu_{P^k} : X\text{x}X \rightarrow [0, 1]$, where $\mu_{P^k}(x_i, x_j) = p_{ij}^k$ denotes the preference degree of the alternative x_i over x_j.

- *A multiplicative preference relation.* In this case, the expert's preferences on X are described by a preference relation, $A^k \subset X\text{x}X$. The intensity of preference , a_{ij}^k, is measured using a ratio scale, particularly the 1 to 9 scale; $a_{ij}^k = 1$ indicates indifference between x_i and x_j, $a_{ij}^k = 9$ indicates that x_i is absolutely preferred to x_j, and $a_{ij}^k \in 2, 3, \ldots, 8$ indicates intermediate evaluations.

- *An utility function.* In this case, an expert, e_k, gives his preferences on X as a set of n utility values, $U^k = \{u_i^k,\ i = 1, ..., n\}$, $u_i^k \in [0, 1]$, where u_i^k represents the utility evaluation given by the expert e_k to the alternative x_i .

Usual resolution methods for a GDM problems are composed by two different processes [9] (see Figure 1):

1. *Consensus process:* Clearly, in any decision process, it is preferable that the experts reach a high degree of consensus on the solution set of alternatives. Thus, this process refers to how to obtain the maximum degree of consensus or agreement between the set of experts on the solution alternatives.

2. *Selection process:* This process consists in how to obtain the solution set of alternatives from the opinions on the alternatives given by the experts.

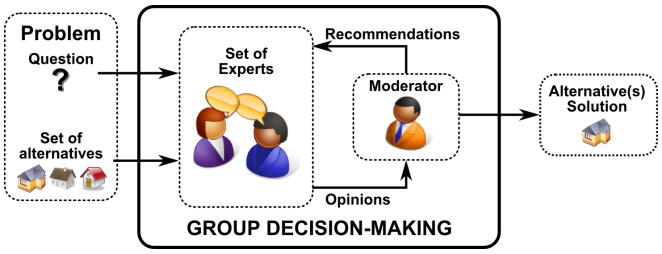


Figure 1: *Resolution process of a GDM*

As it was shown in [3, 8], when the information provided by a group of experts is assumed to be of a diverse nature, we need to make the information uniform before applying the consensus and selection processes in order to obtain the solution alternatives. Assuming the above fact, in [10] a GDM model was proposed, which incorporated i) a process to make the preferences uniform that used some transformation functions between the different preference elements and fuzzy preference relations as the base element of the uniform representation, ii) a selection process based on quantifier guided aggregation operators to obtain the ranking of solution alternatives, and iii) a consensus process with a feedback mechanism to help experts in the consensus reaching process. Its operation scheme is shown in Figure 2.

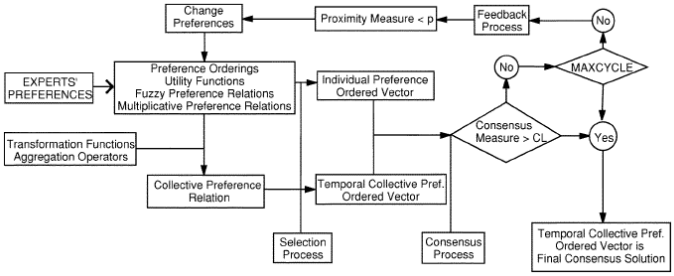


Figure 2: *Operation of GDM model with multiple elements of preference representation*

3.2 Prototype of the Mobile DSS

According to GDM model proposed in [10] the prototype of the Mobile DSS should to present an operation structure based three main modules, which receive/send information from/to the experts through M-Internet technologies (see Figure 3).

1. Uniform Information Module, that makes the information uniform using fuzzy preference relations as the base element of the uniform representation .
2. Selection Module, that returns the solution set of alternatives in each stage of decision process.
3. Consensus Module, that determines if the adequate agreement degree has been reached and the decision process must finish by applying the selection process, or, in other case, if we must do a new round of consensus using the feedback mechanism that generates recommendations to change preferences.

The user sends his/her preferences to Mobile DSS by means of a mobile device, and the system returns expert the final solution or recommendations or preferences of other experts or consensus levels, depending of stage of decision process. An important aspect is that the user-system interaction is done in anytime and anywhere and this facilitates expert's participation and the resolution of the decision process.

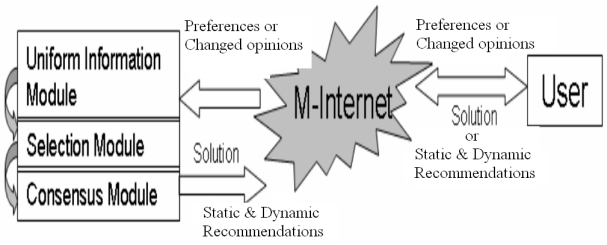


Figure 3: *Operation structure of prototype of Mobile DSS*

A DSS can be built of several ways, and the technology used determines how a DSS is developed [7]. There exist several infrastructures and platforms which can form the technological basis for a DSS: "Mainframe computers", "Client/server", "Internet and Web based systems", or "mobile devices and services" are some kinds of architectures [12].

A good architecture for our prototype of Mobile DSS, is a "Client/Server" architecture, where the client is a mobile device. The client/server paradigm is founded on the concept that clients (such a personal computers, or a mobile device) and servers (powerful personal computer) are both connected by a network enabling servers to provide different services for the clients. When a client sends a request to a server, this server processes the request and sends a response back to client. Furthermore, the technologies used to implement the prototype of Mobile DSS comprise java and java midlets for the client software, Php for the server functions and MySql for the database management.

In what follows, we describe in detail the client and server of prototype of Mobile DSS.

1. **Client**

   We suggest a thin client model, this model is a client device in the client-server architecture networks which depends primarily on the central server for processing activities, and mainly on converging input and output between the user and the remote server. Although this prototype is designed to operate on several mobile devices, such as GSM or UMTS mobile phones or PDA devices that uses mobile network infrastructures and mobile messaging services that fix the communication protocols to use, initially, we have implemented it for mobile phones with Internet con-

nection. The software installed in the client has to meet the next five requirements.

- Connection: The device can be connected to the network to send/receive information to the server.
- Authentication: The device will ask for an user and password to access the system (see Figure 4).

Figure 4: *Authentication and M-Internet connection*

- Interface: The device will have four different interfaces, one for each different element of preference representation (see Figures 5 and 6).

Figure 5: *Insertion of preferences*

- Feedback: When opinions must be modified, the device shows experts the recommendations and lets experts to send the new preferences (see Figure 7).

Figure 6: *Insertion of preferences*

Figure 7: *Recommendations and Final Solution*

- Output: At the end of the decision process, the device will show the set of solution alternatives (see Figure 7).

Some examples of use for this software can be decision making processes where the experts have not the possibility of gathering together, for example doctors of different countries that have to obtain a prediction for a disease or managers of different cities that have to choose the best applicant for a working place.

The client application complies with the MIDP 2.0 specifications. The J2ME wireless toolkit 2.2 provided by SUN was used in the development

phase. The wireless toolkit is a set of tools that provide J2ME developers with the emulation environments, documentation, and examples to develop MIDP-compliant applications. The application was later tested using the JAVA-enabled mobile phone on a GSM network using a GPRS-enabled SIM card. The MIDP application is packaged inside a JAVA archive (JAR) file, which contains the applications class and resource files. It is this JAR file that is actually downloaded to the physical device (mobile phone) along with the JAVA application descriptor file.

2. **Server**

The server is the most important part of the DSS. It implements the uniform information module, the consensus module, the selection module, the database that stores all data of problem as experts' data, alternative data, preferences, consensus measures, recommendations, consensus parameters, selection parameters, ect,. Therefore, the server has to carry out the following functions:

- **Verify the user messages and store the main information:** When a expert wants access to the system, he has to send a message through M-Internet using his mobile device. This message is composed by information of authentication (login and password) and his preferences about the problem, using any of these four ways: *preference orderings, fuzzy preference relations, multiplicative preference relations or utility functions.*
  The message is verified by the server, that check the login and password in the database using MD5 to encrypt the passwords. If the authentication process is correct, the rest of the information of the message is stored in the database, and the server check if all the experts have sent their preferences and the consensus stage can start.
- **Make uniform the experts' preferences:** Using transformation functions to convert all the kind of preferences into fuzzy preference relation, the server is able to make the information uniform, and save this information in the database, that will be used later by the selection module.
- **Calculate the set of solution alternatives:** Once the information is uniform, the server can apply the selection process to obtain a temporary solution of the problem. This process has two phases: i) Aggregation and ii) Exploitation. In the aggregation phase, the collective fuzzy preference relation is obtained. In the next phase, we apply the exploitation process to obtain the quantifier guided dominance degree of alternatives acting over the collective fuzzy preference relation, that allows to establish an order in the alternatives to obtain the ranking of temporary solutions, from best to worse. All this information about the temporary solution is saved in the database.
- **Calculate the consensus measures:** In this step, the consensus and proximity measures are calculated by the server and saved in the database.
- **Control the consensus stage:** If the consensus measure has reached the minimum consensus level defined as a parameter of the problem, the consensus process would be stopped and this temporary collective solution would be the final consensual solution. In other case, the consensus process should continue. In order to avoid that the collective solution does not converge after several discussion rounds the prototype incorporate a maximum number of rounds to develop, MAXCYCLE.
- **Generate the recommendations:** The ranking of experts according to the proximity of their individual solutions to the temporary collective solutions and the recommendations for the experts that must change their preferences are calculated in this step. This is a feedback mechanism and these results are saved in the database and sent to the experts' mobile devices for help them to change their preferences and in this way, the consensus process will be convergent and the solution will have a high consensus degree quickly.

4 Conclusions

We have presented a prototype of Mobile DSS for GDM problems based on multiple elements of preference representation, which uses the advantages of M-Internet technologies to improve the user satisfaction with the decision process and to develop decision processes in anytime and anywhere. We have used phones as device to send the experts' preferences but the structure of the prototype is designed to use any mobile device as PDAs.

In the future, we shall incorporate in the prototype tools to manage incomplete and inconsistent preferences.

Acknowledgements

This work has been supported

by the Research Projects SAINFOWEB-PAI00602 and TIN2007-61079.

References

[1] Aronson J, Liang T, Turban E (2005) Decision support systems and intelligent systems. Pearson, Upper Saddle River

[2] F. Burstein, J. Cowie, A. Zaslavsky and J. San Pedro, "Support for real-time decision making in mobile financial applications", *Handbook on Decision Support Systems* edited by Frada Burstein and Clyde W. Holsapple (2008) Springer.

[3] F. Chiclana, F. Herrera, and E. Herrera-Viedma, "Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations", *Fuzzy Sets and Syst.*, vol. 97, pp. 33-48, 1998.

[4] L. Cohen, L. Manion, and K. Morrison, *A Guide to Teaching Practice.* London: Routledge, 1996.

[5] J. Cowie and F. Burstein, "Quality of data model for supporting mobile decision making", *Decision Support Systems*, vol. 43, pp. 1675-1683, 2007.

[6] J. Fodor and M. Roubens, *Fuzzy Preference Modelling and Multicriteria Decision Support* Kluwer Academic Publishers, 1994.

[7] S. French and M. Turoff, "Decision Support Systems", *Communications of the ACM*, vol. 50, num. 3, pp. 39-40, 2007.

[8] F. Herrera, E. Herrera-Viedma, F. Chiclana, "Multiperson Decision Making Based on Multiplicative Preference Relations", *European Journal of Operational Research*, vol. 129, pp. 372-385, 2001.

[9] F. Herrera, E. Herrera-Viedma , and J.L. Verdegay, "A model of consensus in group decision making under linguistic assessments", *Fuzzy Sets and Syst.*, vol. 78, pp. 73-87, 1996.

[10] E. Herrera-Viedma, F. Herrera and F. Chiclana. A Consensus Model for Multiperson Decision Making with Different Preference Structures. IEEE Transactions on Systems, Man and Cybernetics. Part A: Systems and Man 32:3 (2002) 394-402

[11] L. Kitainick, *Fuzzy Decision Procedures with Binary Relations, Towards an Unified Theory.* Dordrecht: Kluwer Academic Publishers, 1993.

[12] J. Muntermann, "Mobile notification and Decision Support for private investors", is a chapter of the book *Event-Driven mobile Finantial Information Services*, pp. 51-92, Ed. DUV, 2008.

[13] U. Varsney and R. Vetter,"Emerging Mobile and Wireless Networks", *Communications of the ACM*, vol. 43, num. 6, pp. 73-81, june 2000.

Modelo de Evaluación Sensorial con Información Lingüística Multigranular para el Aceite de Oliva

M. Espinilla, L. Martínez, L. G. Pérez
Dpto. de Lenguajes y Sistemas Informáticos
Universidad de Jaén, 23071 - Jaén, España
mestevez, martin, lgonzaga@ujaen.es

J. Liu
School of Computing and Mathmatics
University of Ulster
j.liu@ulster.ac.uk

Resumen

Los procesos de evaluación se utilizan para la inspección de la calidad, la comercialización y otros campos en compañías industriales. En los procesos de evaluación sensorial la información proporcionada por los expertos implica siempre incertidumbre e imprecisión. Por lo que, el uso de la aproximación lingüística difusa para modelar esta información puede ayudar a los expertos a la hora de expresar sus percepciones. En los procesos de evaluación puede ocurrir que diferentes expertos tengan diferente grado de conocimiento sobre los elementos a evaluar; por lo que, el uso de escalas con distinta granularidad para modelar sus opiniones puede ser adecuada. En esta contribución se propone un modelo de evaluación sensorial basado en un modelo de análisis de decisión para problemas definidos en marcos lingüísticos multigranulares donde cada experto puede expresar sus valoraciones en conjuntos de etiquetas distintos según su grado de conocimiento para aplicarlo a la evaluación sensorial del aceite de oliva.

Palabras Clave: Evaluación Sensorial, Análisis de Decisión, Aceite de Oliva, Jerarquías Lingüísticas.

1. Introducción

La evaluación es un proceso cognitivo complejo que implica diversos mecanismos en los cuales es necesario identificar los elementos que van a ser evaluados, fijar el marco en el que se va a realizar la evaluación, recopilar la información y finalmente obtener una valoración de los elementos evaluados. El objetivo de cualquier proceso de evaluación es obtener una valoración de un artículo (producto, servicio, material, etc.) para mejorarlo o compararlo con otros artículos y así saber cuál es mejor. La información recopilada en los procesos de evaluación es proporcionada por un grupo de individuos previamente seleccionados y entrenados, llamado *panel de expertos*, donde cada uno expresa sus opiniones sobre los elementos a evaluar.

La evaluación sensorial [5, 9, 10, 11] es una disciplina de la evaluación donde la información proporcionada por el panel de expertos, es percibida por los órganos sensoriales de *la vista, el olfato, el oído, el gusto y el tacto*. Se considera que el aparato sensorial humano muestra grados de variación de sensibilidad de persona a persona y que la sensibilidad puede ser influenciada fácilmente por circunstancias externas o del medio [8]. Es por esto que uno de los grandes problemas asociados al análisis sensorial es conseguir que la respuesta humana sea precisa y se pueda determinar de modo cuantitativo. Este problema puede mejorarse modelando la información sensorial de forma cualitativa mediante etiquetas lingüísticas [9]. El uso de la aproximación lingüística difusa [12] ha producido buenos resultados a la hora de modelar la información cualitativa en diversos campos de aplicación [1, 3].

Como hemos visto los sentidos sensoriales muestran grados de variación en los humanos y, puede suceder que diferentes indicadores sensoriales de los objetos a evaluar se puedan analizar con diferente precisión. Por tanto, puede ocurrir que los expertos necesiten distintos grados de precisión o que unos indicadores se modelen mejor en un conjunto con pocos términos lingüísticos mientras que otros lo sean en conjuntos con más etiquetas. En tales situaciones, ofrecer un marco de evaluación flexible donde a los expertos se les permita expresarse en el conjunto de términos que se sientan más cómodos puede mejorar los resultados de la evaluación. Esta contribución propone un modelo de evaluación sensorial, basado en un esquema de análisis de decisión [4], con un marco de evaluación flexible que define un contexto lingüístico multigranular modelado mediante jerarquías lingüísticas [7] para mejorar la precisión de los procesos de computación con pala-

bras, que aparecen en el proceso de evaluación. Dicho modelo será aplicado a la evaluación sensorial del aceite de oliva.

Esta contribución se estructura como sigue, en la sección 2 revisamos brevemente varios conceptos necesarios para desarrollar el modelo. En la sección 3 presentamos el modelo de evaluación sensorial con escalas lingüísticas multigranulares. En la sección 4 mostramos su aplicación a la evaluación sensorial del aceite de oliva. En la sección 5 señalamos algunas conclusiones.

2. Información Lingüística: *Antecedentes*

En esta sección revisamos brevemente conceptos relacionados con la información lingüística y que serán utilizados en el desarrollo de nuestra propuesta, tales como, el enfoque lingüístico difuso, el modelado lingüístico basado en 2-tuplas y las jerarquías lingüísticas.

2.1. Enfoque Lingüístico Difuso

El enfoque lingüístico difuso representa información lingüística utilizando valores lingüísticos mediante variables lingüísticas [12].

Cuando trabajamos con información lingüística hay que seleccionar los descriptores lingüísticos adecuados y su semántica. Para llevar a cabo esta tarea, un aspecto importante a considerar es la granularidad de la incertidumbre, es decir, la capacidad de discriminación entre distintos valores de información. Valores típicos de cardinalidad en los modelos lingüísticos son valores impares, tales como 7 ó 9, donde el término medio representa una valoración de aproximadamente 0.5 y el resto de términos están distribuidos simétricamente a su alrededor [2]. Una posibilidad es generar directamente el conjunto de términos considerando todos los términos distribuidos en una escala sobre la cuál se define un orden total. Un ejemplo de un conjunto de 5 términos es: $S = \{s_0 : N, s_1 : L, s_2 : M, s_3 : H, s_4 : P\}$.

La semántica de los términos lingüísticos viene dada por números difusos definidos en el intervalo $[0, 1]$. Estas son valoraciones aproximadas, por lo que diferentes autores consideran que el uso de funciones de pertenencia paramétricas (triangulares o trapezoidales) es suficiente para capturar la vaguedad de los términos lingüísticos [2].

2.2. Modelado Lingüístico Basado en 2-tuplas

El modelo de representación lingüístico basado en 2-tuplas fue presentado en [6] para mejorar la precisión en los procesos de computación con palabras.

Este modelo se basa en el concepto de traslación simbólica. Sea $S = \{s_0, ..., s_g\}$ un conjunto de términos lingüísticos, y $\beta \in [0, g]$ un valor en el intervalo de granularidad de S.

Definición 1. *La Traslación Simbólica de un término lingüístico s_i es un número valorado en el intervalo [-.5,.5) que expresa la "diferencia de información" entre una cantidad de información expresada por el valor $\beta \in [0, g]$ obtenido en una operación simbólica y el valor entero más próximo, $i \in \{0, ..., g\}$, que indica el índice de la etiqueta lingüística, s_i, más cercana en S.*

A partir de este concepto la información lingüística se representa con un par de valores o 2-tupla, (s_i, α_i), donde $s_i \in S$ y $\alpha_i \in [-.5, .5)$. Este modelo de representación define un conjunto de funciones que facilitan las operaciones sobre 2-tuplas.

Definición 2. *Sea $S = \{s_0, ..., s_g\}$ un conjunto de términos lingüísticos y $\beta \in [0, g]$ un valor que representa el resultado de una operación simbólica, entonces la 2-tupla lingüística que expresa la información equivalente a β se obtiene usando la siguiente función:*

$$\Delta : [0, g] \longrightarrow S \times [-.5, .5)$$

$$\Delta(\beta) = (s_i, \alpha), \quad \begin{cases} s_i, & i = round(\beta) \\ \alpha = \beta - i, & \alpha \in [-.5, .5) \end{cases}$$

donde round es el operador usual de redondeo, s_i es la etiqueta con índice más cercano a β y α es el valor de la traslación simbólica.

Conviene señalar que Δ es bijectiva y $\Delta^{-1} : S \times [-.5, .5) \longrightarrow [0, g]$ es definida como $\Delta^{-1}(s_i, \alpha) = i + \alpha$. En este sentido, la 2-tupla $S \times [-.5, .5)$ queda identificada con el valor en el intervalo $[0, g]$.

Comentario 1: *A partir de las definiciones 1 y 2, la conversión de un término lingüístico en una 2-tupla consiste en añadir el valor cero como traslación simbólica: $s_i \in S \longrightarrow (s_i, 0)$.*

2.3. Jerarquías Lingüísticas

Las Jerarquías Lingüísticas se presentaron en [7] como una estructura capaz de mejorar la precisión en los procesos computacionales de información lingüística multigranular.

Una Jerarquía Lingüística está compuesta por un conjunto de niveles, donde cada nivel es un conjunto de

términos lingüísticos con distinta granularidad al resto de niveles de su jerarquía. Cada nivel de la jerarquía es notado como, $l(t, n(t))$, siendo:

1. t es un número que indica el nivel de la jerarquía.
2. $n(t)$ es la granularidad del conjunto lingüístico del nivel t.

Los niveles dentro de una jerarquía están ordenados de acuerdo a su granularidad, es decir, para dos niveles sucesivos t y $t+1$, $n(t+1) > n(t)$. Esto proporciona un refinamiento lingüístico del nivel anterior. Podemos definir una jerarquía lingüística, LH, como la unión de todos los niveles t: $LH = \bigcup_t l(t, n(t))$ donde el conjunto de etiquetas del nivel t es representado como $S^{n(t)}$. La construcción de una jerarquía lingüística sigue unas reglas definidas en [7]. De acuerdo a estas reglas la granularidad de un conjunto de términos del nivel $t+1$ se obtiene de su predecesor t como:

$$l(t, n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$$

Un ejemplo gráfico de una jerarquía lingüística puede verse en la Figura 1:

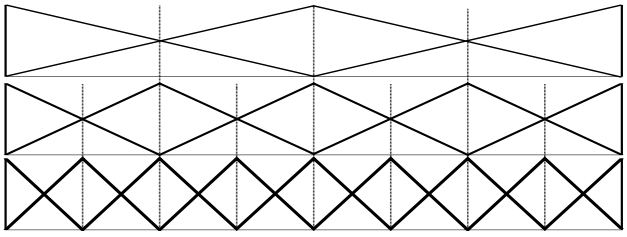


Figura 1: Jerarquía de 3, 5 y 9 etiquetas.

A continuación, podemos ver el conjunto de términos lingüísticos correspondiente a cada nivel

$$l(1,3) = S^3 = \{s_0^3, s_1^3, s_2^3\}$$
$$l(2,5) = S^5 = \{s_0^5, s_1^5, s_2^5, s_3^5, s_4^5\}$$
$$l(3,9) = S^9 = \{s_0^9, s_1^9, s_2^9, s_3^9, s_4^9, s_5^9, s_6^9, s_7^9, s_8^9\}$$

Las jerarquías lingüísticas permiten operar con etiquetas de distintos niveles sin pérdida de información mediante la transformación de etiquetas entre niveles de la jerarquía, mediante el uso del modelo computacional de la representación de 2-tuplas y la siguiente función de transformación.

Definición 3. *Sea $LH = \bigcup_t l(t, n(t))$ una jerarquía lingüística cuyos conjuntos de términos notaremos como $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$. La función de transformación de un término del nivel t a uno del nivel t' se define como:*

$$TF_{t'}^t : S^{n(t)} \times [-.5, .5) \longrightarrow S^{n(t')} \times [-.5, .5)$$

$$TF_{t'}^t(s_i^{n(t)}, \alpha^{n(t)}) = \Delta(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t') - 1)}{n(t) - 1})$$

Proposición 1[7]. *La función de transformación de términos entre distintos niveles de una jerarquía es biyectiva:*

$$TF_t^{t'}(TF_{t'}^t(s_i^{n(t)}, \alpha^{n(t)})) = (s_i^{n(t)}, \alpha^{n(t)})$$

Por lo tanto, nos garantiza la transformación sin pérdida de información.

3. Modelo de Evaluación Sensorial

En esta sección presentamos un modelo de evaluación sensorial donde cada experto puede expresar sus valoraciones en un conjunto de etiquetas distinto según su grado de conocimiento sobre el objeto evaluado. El modelo de evaluación propuesto está basado en el análisis de decisión [4]. Este análisis permite a los decisores tomar decisiones de una forma consistente ya que, hace un estudio detallado de las alternativas del problema, que es el objetivo de los procesos de *Evaluación*.

En [4] se muestran todas las fases de un proceso de análisis de decisión, en la Figura 2 mostramos las fases relacionadas con un modelo de evaluación:

- *Identificar los Objetos a Evaluar*. En cada problema se identifican sus objetos de interés.
- *Marco de Evaluación*. Aquí se definen la estructura del problema y el modelado de preferencias que usarán los expertos.
- *Recopilación de la información*. Los expertos expresan su preferencias mediante etiquetas lingüísticas en LH.
- *Valoración de los objetos*. Se realizan las operaciones necesarias para obtener una evaluación de los objetos evaluados.
- *Resultados de la evaluación*. Se estudian las valoraciones obtenidas.

A continuación, se presenta con mayor detalle las fases principales de nuestro modelo de evaluación sensorial.

3.1. Marco de Evaluación

En esta fase se define el esquema del problema de evaluación sensorial. En nuestro caso será un problema de Toma de Decisión Multi-Experto (TDME) donde

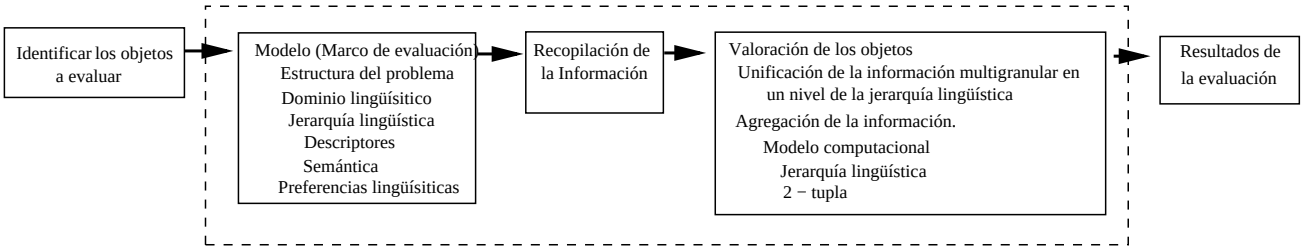


Figura 2: Fases de un análisis de decisión para un proceso de evaluación.

los expertos darán sus opiniones en vectores de utilidad expresadas en un contexto lingüístico multigranular basado en jerarquías lingüísticas.

El objetivo de la evaluación sensorial es evaluar los objetos según el conocimiento adquirido por el sistema sensorial de los expertos. La definición y terminología usada en nuestra propuesta es:

- Sea $E = \{e_1, \ldots, e_n\}$ el panel de expertos.
- Sea $X = \{x_1, ..., x_m\}$ objetos a evaluar.
- Sea $F = \{f_1, ..., f_h\}$ el conjunto de características sensoriales de los objetos.
- Cada experto, e_i, expresa sus preferencias en un conjunto de términos $S_i^{n(t)} \in LH$.

Este marco de evaluación permite que cada característica esté valorada en distintos conjuntos de etiquetas. Los expertos pueden elegir distinto número de etiquetas (nivel de la jerarquía lingüística) para evaluar la misma característica sensorial.

3.2. Recopilación de la Información

En esta fase se obtiene la intensidad de las percepciones de los expertos sobre cada una de las características de los objetos valorados. El experto e_i, proporciona las intensidades percibidas a través de los sentidos sensoriales en términos lingüísticos contenidos en LH con un vector de utilidad:

$$U_i = \{u_{11}^i, ..., u_{1h}^i, u_{21}^i, ..., u_{2h}^i, ..., u_{m1}^i, ..., u_{mh}^i\}$$

donde, $u_{jk}^i \in S_i^{n(t)}$, es la intensidad percibida por el experto, e_i, del atributo sensorial, f_k, del objeto, x_j.

La información recopilada es expresada en 2-tuplas, añadiendo el valor cero como traslación simbólica, tal y como se indica en el *Comentario 1* (sección 2.2).

3.3. Valoración de los Objetos

En esta fase se busca obtener un valor global de los objetos evaluados, atendiendo a las valoraciones proporcionadas por el panel de expertos. Dado que el contexto en que se definen estos problemas es multigranular esta fase se dividirá en los siguientes procesos:

3.3.1. Unificación de la Información Multigranular

En primer lugar, los vectores de utilidad, U_i, proporcionados por los expertos están valorados en distintos conjuntos de etiquetas. Para poder operar sobre ellos hay que unificar dicha información en un único dominio de expresión que denominaremos Conjunto Básico de Términos Lingüísticos (CBTL) notado como, t' y que puede ser cualquiera que pertenezca a la LH, ya que no se produce pérdida de información. Transformaremos las valoraciones de cada experto, e_i, a dicho nivel mediante $TF_{t'}^t$. Una vez unificada la información de entrada ésta estará expresada mediante 2-tuplas en el conjunto $S^{n(t')}$.

3.3.2. Agregación de la Información

El segundo proceso para obtener valoraciones de los objetos evaluados consiste en agregar la información unificada. Para ello utilizaremos operadores de agregación sobre 2-tuplas [6]. El proceso de agregación del modelo general consiste en dos etapas, aunque puede variar según cada problema.

1. *Evaluación colectiva para cada característica*: En primer lugar se calcula un valor colectivo, u_{jk}, para cada característica, f_k, usando un operador de agregación, OA_1, sobre las evaluaciones proporcionadas por los expertos, u_{jk}^i:

$$u_{jk} = OA_1(u_{jk}^1, ..., u_{jk}^n)$$

2. *Evaluación colectiva para cada objeto*: dado que el objetivo principal de la fase de valoración de los objetos es obtener una evaluación global, u_j, para cada objeto evaluado de acuerdo a las preferencias de todos los expertos. Aquí, se agregan los valores colectivos de cada característica para cada objeto, usando un operador de agregación, OA_2:

$$u_j = OA_2(u_{j1}, ..., u_{jh})$$

Dependiendo del objetivo del proceso de evaluación sensorial los operadores de agregación pueden variar, tanto en número como en los operadores a utilizar. En [6] se pueden encontrar operadores de agregación 2–tuplas: Media Aritmética para 2-tuplas, Media Ponderada...

4. Modelo de Evaluación Sensorial Aplicado al Aceite de Oliva

La calidad del aceite de oliva viene determinada, entre otras, por dos de sus características organolépticas (olor, sabor). De esta manera el aceite de oliva es uno de los pocos alimentos en los que la evaluación de sus características sensoriales es un requisito para evaluar su calidad (extra, virgen, corriente y lampante) según normativas de la Unión Europea y del Consejo Oleico Internacional (COI).

A continuación presentamos la aplicación del modelo de evaluación sensorial anterior para la evaluación de muestras de aceite de oliva.

4.1. Marco de Evaluación

El objetivo de la evaluación sensorial de una muestra de aceite de oliva es evaluar su calidad.

Supongamos un panel de expertos de 8 individuos, $E = \{e_1, \ldots, e_8\}$, que someterán a evaluación una muestra de aceite de oliva, $X = \{x_1\}$ de la que se evaluarán 7 atributos sensoriales $F = \{f_1, ..., f_7\}$ (6 atributos negativos: atrojado, moho, avinado, borras, metálico y rancio, y notaremos como $f_1, f_2, f_3, f_4, f_5, f_6$ respectivamente y un atributo positivo denominado frutado f_7) que determinan las características de cada variedad.

Los expertos suministrarán sus preferencias sobre los atributos de las muestras de aceite de oliva en diversos términos de la jerarquía lingüística mostrada en la Fig. 1. Específicamente, las preferencias de los atributos f_1, f_2 y f_3 en S^5 y el resto de atributos, $\{f_4, ..., f_7\}$, serán valorados en S^9, la sintaxis de los conjuntos lingüísticos de etiquetas es:

s_0^9=Excesivamente fuerte.
s_1^9=Muy fuerte.
s_2^9=Moderadamente fuerte.
s_3^9=Ligeramente fuerte.
s_4^9=Regular.
s_5^9=Ligeramente intenso.
s_6^9=Moderadamente intenso.
s_7^9=Poco intenso.
s_8^9=Inapreciable.

s_0^5=Excesivamente fuerte.
s_1^5=Moderadamente fuerte.
s_2^5=Regular.
s_3^5=Moderadamente intenso.
s_4^5=Inapreciable.

4.2. Recopilación de la Información

Los vectores de utilidad proporcionados por los expertos los podemos observar en la Tabla 1:

Tabla 1: Valoraciones del panel de expertos.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
e_1	s_0^5	s_0^5	s_0^5	s_0^9	s_0^9	s_0^9	s_1^9
e_2	s_0^5	s_0^5	s_0^5	s_0^9	s_0^9	s_1^9	s_1^9
e_3	s_0^5	s_1^5	s_0^5	s_0^9	s_1^9	s_2^9	s_2^9
e_4	s_0^5	s_1^5	s_0^5	s_0^9	s_0^9	s_1^9	s_2^9
e_5	s_0^5	s_1^5	s_0^5	s_0^9	s_0^9	s_0^9	s_2^9
e_6	s_0^5	s_1^5	s_0^5	s_1^9	s_0^9	s_1^9	s_2^9
e_7	s_0^5	s_0^5	s_0^5	s_0^9	s_0^9	s_0^9	s_1^9
e_8	s_0^5	s_0^5	s_0^5	s_0^9	s_0^9	s_0^9	s_1^9

4.3. Valoración de los Objetos

En esta fase se obtendrá una valoración global de la muestra de aceite de oliva evaluada.

4.3.1. Unificación de la Información Multigranular

Se selecciona un conjunto de términos lingüísticos para unificar la información lingüística multigranular. En este caso, elegiremos $l(3, 9)$ puesto que la mayoría de los expertos han expresado sus preferencias en S^9. Por lo tanto, unificamos los valores de preferencia en el CBTL, utilizando la función de transformación TF_3^2.

4.3.2. Agregación de la Información

Una vez expresada la información en un único dominio de expresión, se aplica un proceso de agregación que consta de dos etapas para obtener un valor global de la muestra evaluada:

1. *Evaluación colectiva para cada característica*: En esta etapa se calcula un valor colectivo,u_{jk}, para cada característica, f_k, usando el operador de agregación *mediana 2-tupla*, Me. ya que la normativa de Unión Europea y del COI proponen la mediana para agregar los valores de todos los expertos.

   Definición 4. *Sea* $X = \{(s_j, \alpha)_1, \ldots, (s_j, \alpha)_n\}$, $s_j \in S = \{s_0, ..., s_g\}$ *un conjunto ordenado de 2-tuplas, siendo* n *el número de 2-tuplas y* $(s_j, \alpha)_k$ *el k-ésimo elemento mayor del conjunto* X*, la mediana 2-tupla,* $Me(X)$*, se define como:*

$$Me(X) = \begin{cases} Me(X) = (s_j, \alpha)_{\frac{n+1}{2}} \; si \; n \; es \; impar \\ Me(X) = (s_j, \alpha)_{\frac{n}{2}} \; si \; n \; es \; par \end{cases}$$

Tabla 2: Valoraciones del panel de expertos en S^9.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
e_1	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$
e_2	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$	$(s_1^9,0)$
e_3	$(s_0^9,0)$	$(s_2^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$	$(s_2^9,0)$	$(s_2^9,0)$
e_4	$(s_0^9,0)$	$(s_2^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$	$(s_2^9,0)$
e_9	$(s_0^9,0)$	$(s_2^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_2^9,0)$
e_6	$(s_0^9,0)$	$(s_2^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$	$(s_2^9,0)$
e_7	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$
e_8	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,0)$

Tabla 3: Medianas 2-tupla en S^9.

f_1	f_2	f_3	f_4	f_5	f_6	f_7
$(s_0^9,0)$	$(s_1^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_0^9,0)$	$(s_1^9,-.5)$	$(s_2^9,-.5)$

Tabla 4: Evaluación del aceite de oliva.

Categoría	Mediana (Me) del defecto de mayor intensidad	Mediana del atributo frutado
Virgen Extra	$Me = (s_0^9, 0)$	$Me > (s_0^9, 0)$
Virgen	$(s_0^9, 0) < Me \leq (s_2^9, 0)$	$Me > (s_0^9, 0)$
Corriente	$(s_2^9, 0) < Me \leq (s_3^9, 0.1)$ o $Me \leq (s_2^9, 0)$	$Me > (s_0^9, 0)$
Lampante	$Me > (s_3^9, 0.1)$	

Cuando n es par, el valor de la Mediana no es único, $Me(X) \in [(s_j,\alpha)_{\frac{n}{2}}, (s_j,\alpha)_{\frac{n+1}{2}}]$. *Generalizando:*

$$Me(X) = \Delta(\frac{\Delta^{-1}(s_j,\alpha)_{\frac{n}{2}} + \Delta^{-1}(s_j,\alpha)_{\frac{n+1}{2}}}{2})$$

Y se obtienen los valores mostrados en la Tabla 3.

2. *Evaluación colectiva de la muestra*: Esta evaluación se obtiene a partir de las valoraciones colectivas del defecto percibido con mayor intensidad, en nuestro caso f_2, y del atributo frutado, tal y como se muestra en la Tabla 4. Ésta es la equivalente a la utilizada en el modelo cuantitativo actual propuesto según la normativa de la Unión Europea y del COI.

En nuestro ejemplo, el defecto con mayor valor colectivo es f_2 con una intensidad percibida de $(s_1^9, 0)$ y el atributo frutado posee una intensidad de $(s_2^9, -.5)$. Entonces, la valoración global de la muestra de aceite de oliva atendiendo a las valoraciones del panel de expertos y a la Tabla 4 es *Virgen*.

5. Conclusiones

La evaluación sensorial es un proceso de evaluación, en el cual, la información proporcionada por los expertos ha sido adquirida por medio de los órganos sensoriales, esta información es generalmente vaga e imprecisa y dependiendo del sentido utilizado el grado de conocimiento sobre las características evaluadas puede ser distinto. En esta contribución hemos propuesto un modelo de evaluación sensorial donde los expertos pueden expresar sus valoraciones de forma cualitativa mediante etiquetas lingüísticas. Además, el modelo ofrece un marco de modelado de preferencias flexible con escalas de distinta granularidad. Dicho modelo ha sido aplicado a la evaluación sensorial de una muestra de aceite de oliva.

Agradecimientos

Este artículo ha sido realizado con la ayuda de la "Secretaría de Estado de Universidades e Investigación del MEC PR2007-0250 del proyecto TIN2006-02121" y Fondos Feder.

Referencias

[1] B. Arfi. Fuzzy decision making in politics: A linguistic fuzzy-set approach (ILSA). *Political*

Analysis, 13(1):23–56, 2005.

[2] P.P. Bonissone and K.S. Decker. *Selecting Uncertainty Calculi and Granularity: An Experiment in Trading-Off Precision and Complexity.* In L.H. Kanal and J.F. Lemmer, Ed., Uncertainty in A.I. North-Holland, 1986. pp. 217-247.

[3] C.H. Cheng and Y. Lin. Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. *European Journal of Operational Research*, 142:174–186, 2002.

[4] R.T. Clemen. *Making Hard Decisions. An Introduction to Decision Analisys.* Duxbury Press, 1995.

[5] G.B. Dijksterhuis. *Multivariate Data Analysis in Sensory and Consumer Science, Food and Nutrition.* Press Inc. Trumbull, Connecticut, USA, 1997.

[6] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[7] F. Herrera and L. Martínez. A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in Multi-Expert decision-making. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 31(2):227–234, 2001.

[8] G Jellinek. Sensory evaluation of food. Theory and practice. 1985.

[9] L. Martínez. Sensory evaluation based on linguistic decision analysis. *International Journal of Aproximated Reasoning*, 44(2) 148–164, 2007.

[10] D. Ruan and X. Zeng (Eds.). *Sensory Evaluation: Methodologies and Applications.* Springer, 2004.

[11] H. Stone and J.L. Sidel. *Sensory Evaluation Practice.* Academic Press Inc., San Diego, CA, 1993.

[12] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences, Part I, II, III*, 8,8,9:199–249,301–357,43–80, 1975.

Estudio de la influencia de las medidas de complejidad de los datos en los Sistemas de Clasificación Basados en Reglas Difusas: Análisis de la Razón Discriminante de Fisher

Julián Luengo[1] **Salvador García**[1] **José Ramón Cano**[2] **Francisco Herrera**[1]

[1] Dept. CCIA, ETSIIT, Univ. Granada, Granada, 18071, España, {julianlm,salvagl,herrera}@decsai.ugr.es
[2] Dept. Ciencias de la Computación, CPS, Univ. Jaén, Linares, 23700, España, jrcano@ujaen.es

Resumen

El uso de Sistemas de Clasificación Basados en Reglas Difusas es bien conocido y tiene multitud de aplicaciones actualmente. No obstante, los problemas de clasificación pueden ser complejos, y esta complejidad ha sido tipificada y medida por diversos autores. La capacidad de predicción de un clasificador, y por tanto de un Sistema de Clasificación Basado en Reglas Difusas, posee una fuerte dependencia con la complejidad inherente de los datos usados.

En este trabajo presentamos un estudio preliminar sobre el comportamiento de los sistemas de clasificación basados en reglas difusas para la medida de complejidad denominada "razón discriminante de Fisher". Queremos estudiar en que medida las medidas de complejidad nos pueden aportar conocimiento a priori sobre el comportamiento de los sistemas basados en reglas difusas para estos problemas. Para ello comenzamos con esta primera medida y con un conjunto de bases de datos artificiales que se han generado para disponer de valores de la medida de Fisher en todo su rango, que nos permiten extraer unas primeras conclusiones y plantear diferentes trabajos a realizar de cara a tener un completo estudio para esta medida.

Palabras Clave: Clasificadores, Medidas de Complejidad, Medidas de Solapamiento, Sistemas de Clasificación Basados en Reglas Difusas.

1 INTRODUCCIÓN

Los Sistemas de Clasificación Basados en Reglas Difusas (SCBRD) [4] son métodos de clasificación que emplean reglas difusas para representar el conocimiento. Los SCBRDs se encuentran muy extendidos en la actualidad, con numerosas aplicaciones y estudios de su comportamiento y efectividad [9].

En la literatura especializada existen precedentes recientes que intentan caracterizar los datos sobre los que trabaja un método de clasificación utilizando las llamadas medidas de complejidad [1, 8]. Las medidas de complejidad se han aplicado en análisis similares para algunos Sistemas de Clasificación que toman como base las medidas de Ho y Basu[1]. Entre ellos podemos encontrar selección de características [6] y algoritmos de inducción de reglas[11, 3]. Sánchez et al. [16] analizan el efecto de la complejidad de datos en un clasificador K-Vecinos más Cercanos. Sin embargo, no se ha aplicado ningún análisis de este tipo a los SCBRDs.

Para efectuar este estudio preliminar utilizaremos una medida de complejidad para el *solapamiento* de las clases [8]: La razón discriminante de Fisher. Queremos analizar la relación existente entre los resultados obtenidos por los SCBRD y los valores de esta medida de solapamiento. Para ello, hemos analizado el comportamiento de modelos SCBRDs sencillos considerando conjuntos de datos generados de acuerdo a la medida establecida en [8]. A su vez, lo hemos comparado con un algoritmo clásico de clasificación como es C4.5 [15], de forma que sea posible ver las diferencias entre modelos de clasificación de diferentes familias.

Esta contribución se organiza de la siguiente manera. En la Sección 2 se describen los SCBRDs, y en particular el modelo usado en nuestro estudio. En la Sección 3 se define y describe la medida de solapamiento que emplearemos. Asimismo, en la Sección 4 se establece el marco experimental y los resultados obtenidos. Fi-

nalmente, en la Sección 5 expresamos las conclusiones de nuestro trabajo.

2 SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS DIFUSAS

Un SCBRDs está compuesto por una Base de Conocimiento (BC) y un Motor de Razonamiento Difuso (MRD) que, utilizando la información de la BC, determina una clase para cualquier patrón de datos admisible que llegue al sistema. La potencia del razonamiento aproximado reside en la posibilidad de obtención de un resultado (una clasificación) incluso cuando no tengamos compatibilidad exacta (con grado 1) entre el ejemplo y el antecedente de las reglas.

2.1 Base de Conocimiento

La BC está formada por dos componentes:

- La *Base de Datos* (BD), que contiene la definición de los conjuntos difusos asociados a los términos lingüísticos utilizados en la Base de Reglas.
- La *Base de Reglas* (BR), formada por un conjunto de reglas de clasificación

$$R = \{R_1, ..., R_L\} \tag{1}$$

En nuestro caso vamos a utilizar reglas difusas con una clase y un grado de certeza asociado a la clasificación para esa clase en el consecuente.

$$R_k: \quad \text{Si } X_1 \text{ es } A_1^k \text{ y } \ldots \text{ y } X_N \text{ es } A_N^k \text{ entonces } Y \text{ es } C_j \text{ con grado } r_k \tag{2}$$

Donde $X_1, \ldots, X_N$ son las variables asociadas a los diferentes atributos del sistema de clasificación, $A_1^k, \ldots, A_N^k$ son las etiquetas lingüísticas utilizadas para discretizar los dominios continuos de las variables, e Y es la variable que indica la clase C_j a la cual pertenece el patrón. r_k es el grado de certeza asociado a la clasificación de la clase C_j para los ejemplos pertenecientes al subespacio difuso delimitado por el antecedente de la regla.

2.2 Motor de Razonamiento Difuso

El MRD es un procedimiento de inferencia que utiliza la información de la BC para predecir una clase ante un ejemplo no clasificado. Tradicionalmente en la literatura especializada [5] se ha utilizado el MRD del máximo, también denominado MRD clásico o de la regla ganadora, que considera la clase indicada por una sola regla teniendo en cuenta el grado de asociación del consecuente de la regla sobre el ejemplo. Otros MRD que combinan la información aportada por todas las reglas que representan el conocimiento de la zona a la que pertenece el ejemplo se estudian en [5]. En este trabajo utilizaremos,el MRD clásico.

A continuación presentamos el modelo general de razonamiento difuso que combina la información proporcionada por las reglas difusas compatibles con el ejemplo.

En el proceso de clasificación del ejemplo $e = (e_1, \ldots, e_N)$, los pasos del modelo general de un MRD son los siguientes:

1. **Calcular el grado de emparejamiento del ejemplo con el antecedente de las reglas**.
2. **Calcular el grado de asociación del ejemplo a la clase consecuente de cada regla mediante una función de ponderación entre el grado de emparejamiento y el grado de certeza de la regla con la clase asociada**.
3. **Determinar el grado de asociación del ejemplo con las distintas clases**.
4. **Clasificación**. Para ello aplicaremos una función de decisión F sobre el grado de asociación del ejemplo con las clases que determinará, en base al criterio del máximo, la etiqueta de clase v a la que corresponda el mayor valor.

En el punto (3) es donde se distinguen los dos métodos mencionados.

2.3 Algoritmo de Wang y Mendel aplicado a problemas de clasificación

Para nuestra experimentación hemos utilizado la extensión del Algoritmo de Wang y Mendel [19] a problemas de clasificación realizada por Chi et al.[4]. Este método de diseño de SCBRDs determina las relaciones entre las variables del problema y establece una correspondencia entre el espacio de características y el de clases en un proceso que sigue esta serie de pasos:

1. *Establecimiento de las particiones lingüísticas.* Una vez determinado el dominio de variación de cada característica X_i, se determinan las particiones difusas.

2. *Generación de una regla difusa para cada ejemplo* $e^h = (e_1^h, \ldots, e_N^h, C_h)$. Para ello es necesario:

2.1 Calcular los grados de pertenencia del ejemplo e^h a las distintas regiones difusas.

2.2 Asignar el ejemplo e^h a la región difusa con mayor grado de pertenencia.

2.3 Generar una regla para el ejemplo, cuyo antecedente está determinado por la región difusa seleccionada y con la etiqueta de clase del ejemplo en el consecuente.

2.4 Calcular el grado de certeza. Para ello se determinará el cociente S_j/S, siendo S_j la suma del grado de pertenencia de los ejemplos de entrenamiento de la clase C_j a la región difusa determinada por el antecedente de la regla, y S la suma del grado de pertenencia a la misma región de todos los ejemplos independientemente de la clase a la que pertenezcan.

3 MEDIDAS DE COMPLEJIDAD

Los problemas de clasificación pueden ser difíciles por tres razones diferentes:

- Ciertos problemas son conocidos por tener un error de Bayes no nulo [7]. Algunas clases pueden ser ambiguas intrínsecamente o debido a medidas incorrectas de los atributos.
- Algunos problemas pueden presentar límites de decisión complejos, por lo que no se puede ofrecer una descripción compacta de los mismos [16].
- Muestras de tamaño reducido y la dispersión inducida por la alta dimensionalidad afecta a las reglas [13, 14].

El uso de medidas es reciente. En [8] Ho y Basu definen medidas de complejidad para conjuntos de datos de dos clases. Singh en [17] ofrece una revisión de medidas de complejidad y propone dos más. Sotoca et al. en [18] extienden algunas medidas de Ho y Basu para problemas multiclase. Para ello, analizan las medidas generalizadas en dos problemas clásicos de Selección de Prototipos.

En nuestro estudio, a partir de las medidas de Ho y Basu [8] hemos empleado la *razón discriminante de Fisher*. Se trata de una medida geométrica de solapamiento, cuya versión más sencilla indicada por Ho y Basu calcula cómo de separadas están dos clases de acuerdo a una característica en concreto. Compara la diferencia entre la media de las clases con la suma de la varianza de las clases. La razón discriminante de Fisher para un atributo se define como:

$$f = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \qquad (3)$$

donde μ_1, μ_2, σ_1 y σ_2 son las medias y varianzas de las dos clases respectivamente. Calculamos f para cada uno de los atributos y se toma el máximo como la medida F1. Valores pequeños indican que las clases tienen un alto grado de solapamiento. Las Figuras 1 a 4 muestran un ejemplo ilustrativo generado artificialmente con 2 variables en el rango $[0.0, 1.0]$ y dos clases como ejemplo.

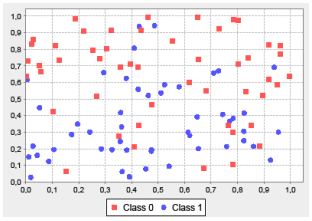


Figura 1: $F1 = 0.6994$

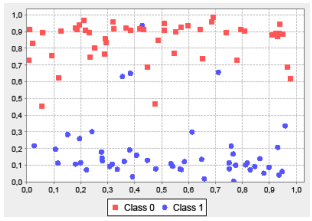


Figura 2: $F1 = 9.69$

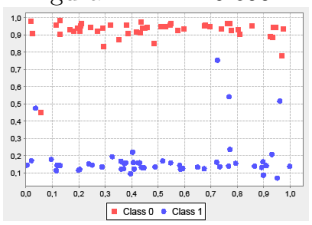


Figura 3: $F1 = 26.16$

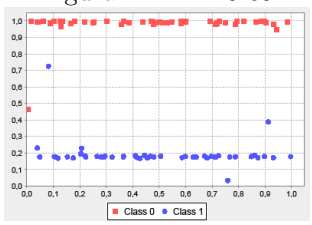


Figura 4: $F1 = 48.65$

4 ESTUDIO EXPERIMENTAL

En esta sección describiremos los datos y metodología usadas en la experimentación de nuestro estudio. Se mostrarán los resultados obtenidos y se analizarán obteniendo conclusiones en base a los mismos para cada una de las medidas de complejidad consideradas.

4.1 Marco Experimental

Para comprobar el comportamiento del método de Chi con diferentes valores de la medida F1, hemos empleado una serie de conjuntos de datos. Hemos considerado tres tipos de conjuntos de datos: con 2, 4 y 8 atributos de entrada. Cada tipo consta de cien conjuntos de datos distintos, en los cuales se incrementa progresivamente el valor de la medida F1, para un total de 300 conjuntos de datos. Estos conjuntos de datos disponen de atributos reales de entrada en el rango $[0.0, 1.0]$, y un atributo nominal de salida con dos posibles etiquetas diferentes (clases), existiendo un total de 100 ejemplos por conjunto de datos. De estos 100 ejemplos, hay 50 ejemplos para cada clase. En nuestros conjuntos, la medida F1 está en el intervalo $[0.699, 74.1]$. Partimos de una distribucin de ejemplos no totalmente ruidosa, cuyo valor F1 es 0.699. Conforme aumenta el valor de la medida de solapamiento,

las clases estarán identificadas por atributos menos solapados.

Para realizar un estudio comparativo vamos a utilizar un modelo de validación cruzada de orden 5 (5-fcv). Los métodos de clasificación se han ejecutado en cada uno de los 100 conjuntos de datos, de 2, 4 y 8 atributos, con este método de validación. La media de las tasas de acierto de las 5 particiones se toma como medida significativa del rendimiento del modelo.

Los modelos SCBRD empleados para medir y comparar su rendimiento son dos: un SCBRD Chi con 3 etiquetas por atributo y un SCBRD Chi con 5 etiquetas por atributo. Para los modelos de Chi emplearemos como *forma de combinación del grado de emparejamiento* y del *grado de certeza* la T-norma producto. El *tipo de inferencia* será el método clásico de la regla ganadora.

Asimismo, y con el fin de comparar con otro algoritmo clásico en el ámbito de clasificación de diferente naturaleza, hemos realizado el mismo estudio con los mismos datos y particiones para el árbol de decisión C4.5 [15]. Esto permitirá comprobar si los SCBRDs tienen comportamientos diferenciados con otras familias de clasificadores.

4.2 Resultados para F1: Razón discriminante de Fisher

El porcentaje de acierto en clasificación tanto en entrenamiento como en test para la medida F1 y cada uno de los tipos de conjuntos (con 2,4 y 8 atributos) se encuentra reflejado en las Figuras 5 a 7.

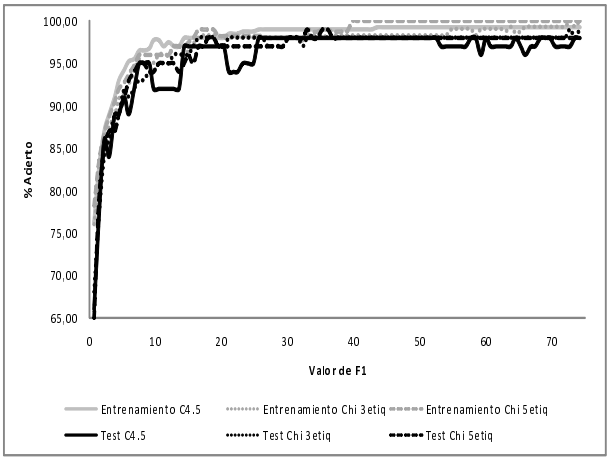


Figura 5: Porcentaje de acierto para F1 con 2 atr.

Para valores bajos de la medida F1 (alto solapamiento), tanto C4.5 como Chi con 3 etiquetas tienen un porcentaje de acierto bajo en entrenamiento, mientras que Chi con 5 etiquetas obtiene mejores resultados, con motivo del sobre-aprendizaje que se produce, por el alto número de reglas. Conforme el valor de F1 aumenta los tres métodos son capaces de mejorar su porcentaje de acierto en entrenamiento. En test para F1 queda patente que el valor en test mejora cuando aumenta la separabilidad. Para valores bajos de la medida el porcentaje de acierto en test es bajo.

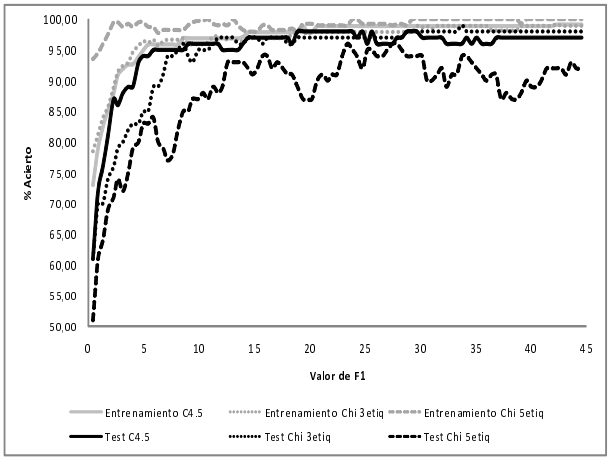


Figura 6: Porcentaje de acierto para F1 con 4 atr.

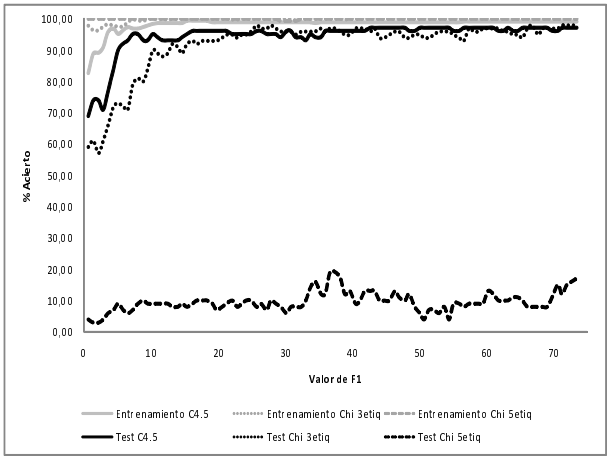


Figura 7: Porcentaje de acierto para F1 con 8 atr.

En las Figuras 8 a 10 se ha dispuesto la cantidad de nodos de C4.5 y de reglas de los métodos de Chi para cada conjunto de datos (ordenados según F1 de forma ascendente) según el número de atributos. Mientras que C4.5 disminuye el número de reglas al aumentar el valor de F1, los métodos de Chi mantienen un número alto de reglas para cubrir los ejemplos separados. Este alto número de reglas para Chi con 5 etiquetas es otro indicativo de su sobre-aprendizaje, sobre todo en el caso de 8 atributos donde siempre tiene 80 reglas (una

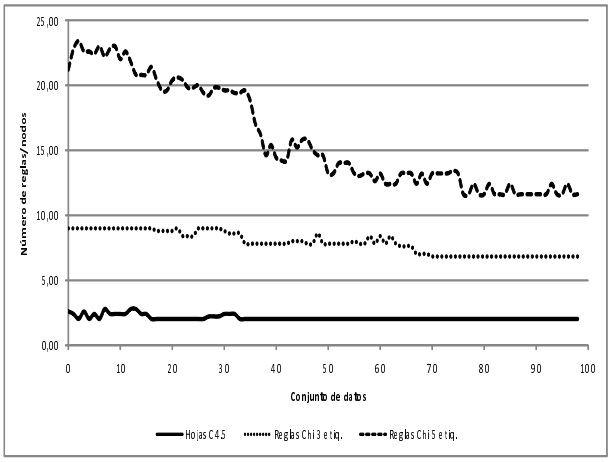


Figura 8: Evolución de reglas y nodos para 2 atributos

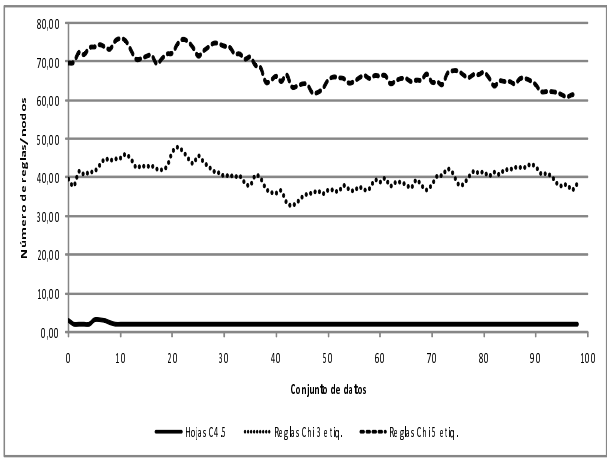


Figura 9: Evolución de reglas y nodos para 4 atributos

regla por ejemplo de entrenamiento). Chi con 3 etiquetas aumenta el número de reglas conforme aumenta también el número de atributos. C4.5 no obstante, obtiene siempre un número bajo de reglas sin importar el número de atributos considerado.

El uso de métodos de selección de características podría evitar el aumento del número de reglas (sobre todo en el caso de un alto número de atributos), puesto que C4.5 es capaz de seleccionar los atributos interesantes en su entrenamiento y es capaz de simplificar el número de reglas que produce.

4.3 Análisis de selección de características para F1: Razón discriminante de Fisher

Hemos analizado todas las combinaciones posibles de atributos a la hora de entrenar y evaluar los modelos Chi de 3 y 5 etiquetas para la medida F1 con la

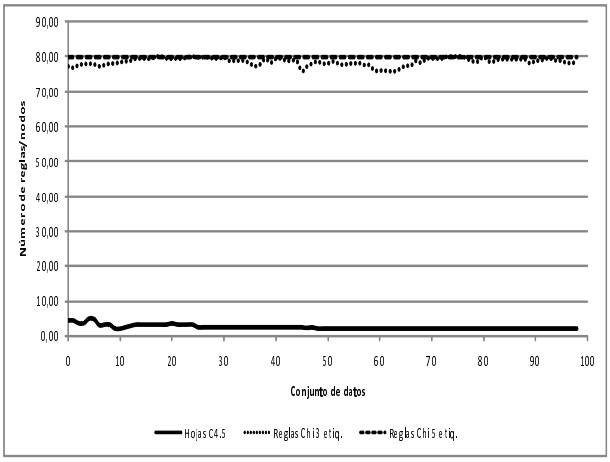


Figura 10: Evolución de reglas y nodos para 8 atributos

variante de 8 atributos. La mejor combinación se establece en base a dos criterios posibles: la que mejor precisión en entrenamiento tenga o la que mejor acierto en test ofrezca. De esta manera comprobamos el criterio del SCBRD de Chi para elegir la combinación de atributos que considera óptima con el mejor valor de entrenamiento (notada como [Mejor Entren.]), frente a la que realmente sería óptima (notada como [Mejor Test]) que obtiene mejor valor en test. Los resultados se encuentran en la Figura 11 para Chi con 3 etiquetas y en la Figura 12 para Chi con 5 etiquetas.

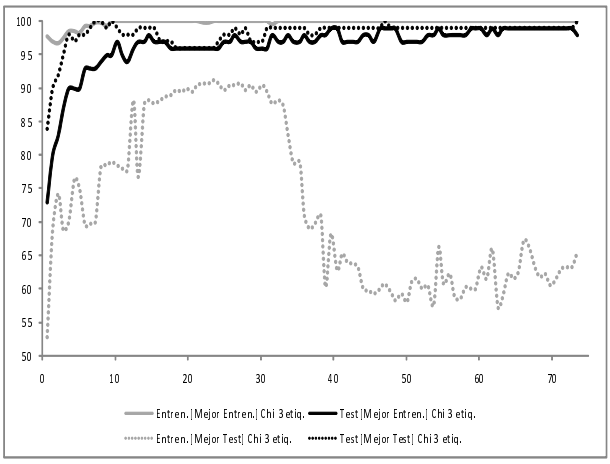


Figura 11: Porcentaje de acierto de Chi 3 etiq. para F1 con la mejor combinación de atributos

La combinación obtenida por el mejor entrenamiento ofrece sobre-aprendizaje para Chi con 5 etiquetas para este número de atributos, pero no en Chi con 3 etiquetas. El sobre-aprendizaje de Chi con 5 etiquetas se ve paliado con la mejor combinación de atributos para

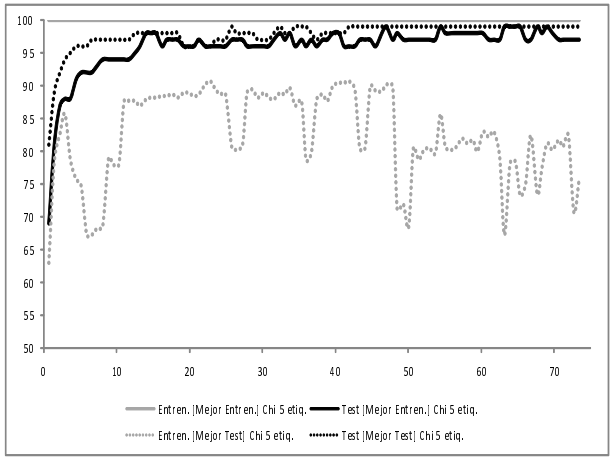


Figura 12: Porcentaje de acierto de Chi 5 etiq. para F1 con la mejor combinación de atributos

test, donde el valor de entrenamiento es similar al de Chi con 3 etiquetas para esta mejor combinación de atributos, y siempre por debajo del propio valor de test.

La combinación para el mejor entrenamiento produce valores de acierto en entrenamiento más altos, y siempre superiores a la combinación óptima para test. El porcentaje de acierto en test para ambos criterios es similar para ambos métodos, aunque el de la combinación óptima para test siempre es ligeramente mejor. Las razones de esta diferencia en entrenamiento podemos encontrarlas en la Figura 13, donde se muestra el número de medio de atributos para el cual cada estrategia (mejor entrenamiento o mejor test) ha sido óptima para cada uno de los 100 conjuntos de datos.

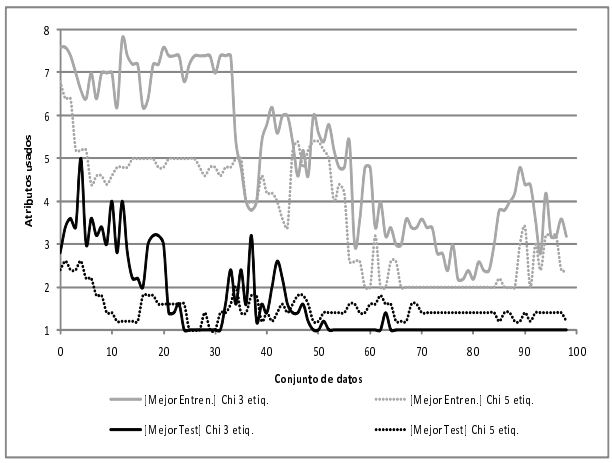


Figura 13: Media de atributos usados para F1 con 8 atr. con cada estrategia

Chi con 3 etiquetas emplea un mayor número de atributos que Chi con 5 etiquetas. Sin embargo, con la estrategia que obtendría el mejor valor de test se aprecia el hecho de que Chi con 3 etiquetas necesita menos atributos que Chi con 5 etiquetas (llegando incluso a usar sólo un atributo). Para obtener el mejor valor de test siempre se emplea un menor número de etiquetas que para obtener el mejor porcentaje de acierto en entrenamiento. Además, observamos que se disminuye la necesidad de utilizar todas las variables cuando nos alejamos de los valores bajos de la función F1 para ambas estrategias, esto permite evitar el sobre-aprendizaje y reafirma la necesidad de utilizar algoritmos de selección de características dentro del proceso de aprendizaje. Puede ser interesante el uso de métodos de selección de características basados en técnicas de filtro. No obstante, emplear la precisión del SCBRD Chi en entrenamiento para obtener la combinación óptima no es adecuado, y se plantea la necesidad de usar otros criterios.

5 CONCLUSIONES

En este trabajo se realiza un primer estudio sobre la influencia de la medida de complejidad llamada razón discriminante de Fisher para analizar el comportamiento de los SCBRDs. Observamos que el solapamiento de las clases (valor de F1 bajo) produce sobre-aprendizaje si utilizamos muchas etiquetas por variables (en este estudio 5 frente a 3) y que se mejora el comportamiento cuando aumenta el valor de F1 y por tanto la separabilidad de las clases. Esta primera experimentación también muestra que cuando aumenta la separabilidad puede ser conveniente el uso de técnicas de selección de características, y que para evitar el sobre-aprendizaje puede ser conveniente que estas técnicas sean de filtro.

Hemos de destacar que éste es un estudio muy preliminar donde hemos trabado con bases de datos artificiales con un numero de variables fijo en 5, un número fijo de instancias y cuatro clases. Este estudio es necesario extenderlo a bases de datos con diferente número de variables, instancias y clases, para poder obtener mas información sobre la utilidad de la medida F1.

Como trabajo a realizar para mejorar este estudio, estamos abordando los siguientes estudios:

a) Analizar el uso de técnicas de filtro, como los métodos LVF [12] y MIFS [2].

b) Analizar el comportamiento d e varios métodos de aprendizaje de reglas difusas, en particular aquellos que incluyan las selección de características dentro de las propias reglas, como el método propuesto en [10].

c) Estudiar posibles extensiones de la medida F1 atendiendo a otros parámetros como son el número de instancias, variables y clases que intervienen en el problema.

Finalmente hemos de indicar que extenderemos el presente estudio a otras medidas de complejidad presentes en la literatura especializada.

Agradecimientos

Este trabajo de investigación ha sido posible gracias a la subvención del proyecto TIN2005- 08386-C05-01.

Referencias

[1] M. Basu and T.K. Ho. Data complexity in pattern recognition,. *Springer*, 2006.

[2] R. Battiti. Using Mutual Information For Selection Features In Supervised Neural Net Learning. *IEEE Transactions on Neural Networks*, 5:4, Pág.537-550, 1994.

[3] E. Bernardó-Mansilla and T.K. Ho. Domain of competence of XCS classifier system in complexity measurement space. *IEEE Transactions on Evolutionary Computation*, 9:1, Pág.82-104, 2005.

[4] Z.Chi, H.Yan, and T.Pham. Fuzzy algorithms with applications to image processing and pattern recognition. *World Scientific*, 1996.

[5] O.Cordón, M.J. del Jesus, and F.Herrera. A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, 20:1, Pág.21-45, 1999.

[6] M. Dong and R. Kothari. Feature subset selection using a new definition of classifiability. *Pattern Recognition Letters*, 24:9-10, Pág.1215-1225, 2003.

[7] T.K. Ho and H.S. Baird. Large-scale simulation studies in image pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:10, Pág.1067-1079, 1997.

[8] T.K. Ho and M. Basu. Complexity measures of supervised classification problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:3, Pág.289-300, 2002.

[9] H. Ishibuchi, T. Nakashima and M. Nii. Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining. *Springer-Verlag*, 2004.

[10] H. Ishibuchi, T. Yamamoto and T. Nakashima. Hybridization of Fuzzy GBML Approaches for Pattern Classification Problems. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, 35:2, Pág.359-365, 2005.

[11] Y.-H. Li, M. Dong and R. Kothari. Classifiability-based omnivariate decision trees. *IEEE Transactions on Neural Networks*, 16:6, Pág.1547-1560, 2005.

[12] H. Liu, R. Setiono. Incremental Feature Selection. *Applied Intelligence*, 9:3, Pág.217-230, 1998.

[13] M. Liwicki and H. Bunke. Handwriting recognition of whiteboard notes - studying the influence of training set size and type. *International Journal of Pattern Recognition and Artificial Intelligence*, 21:1, Pág.83-98, 2007.

[14] X. Qiu and L. Wu. Nearest neighbour discriminant analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, 20:8, Pág.1245-1259, 2006.

[15] J.R. Quinlan. C4.5: Programs for Machine Learning. *Morgan Kaufmann Publishers*, 1993.

[16] J.S. Sánchez, R.A. Mollineda and J.M. Sotoca. An analysis of how training data complexity affects the nearest neighbours classifiers. *Pattern Analysis & Applications*, 10, Pág.189-201, 2007.

[17] S. Singh. Multiresolution estimates of classification complexity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:12, Pág.1534-1539, 2003.

[18] J.M. Sotoca, R.A. Mollineda and J.S. Sánchez. A meta-learning framework for pattern classification. *Revista Iberoamericana de Inteligencia Artificial*, 29, Pág.31-38, 2006.

[19] L.X. Wang and J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics*, 25:2, Pág.353-361, 1992.

A Study on the Evolution of Fuzzy Sets Based Research in Spain Based on Bibliometric Maps[1]

A.G. López-Herrera
Departamento de Informática,
Universidad de Jaén
23071 Jaén
aglopez@ujaen.es

M.J. Cobo, E. Herrera-Viedma, F. Herrera
Departamento de Ciencias de la
Computación e I.A.
Universidad de Granada
18071 Granada
{mjcobo, herrera, viedma}@decsai.ugr.es

Abstract

In this paper, a bibliometric study of the fuzzy sets theory field in Spain is presented. Bibliometric maps showing the associations between the main concepts in the field are provided for the periods 1978-1992, 1993-1997, 1998-2002 and 2003-2007. The maps provide insight into the structure of the fuzzy sets theory field in fuzzy Spanish community. They visualize the division of the field into several subfields, and they indicate the relations between these subfields. The maps are created by the co-word analysis.

Keywords: Bibliometric Studies, Fuzzy Sets Theory.

1 Introduction

Fuzzy sets theory, which was founded by Zadeh [10], has emerged as a powerful way of representing quantitatively and manipulating the imprecision in problems. Fuzzy sets or fuzzy numbers can appropriately represent imprecise parameters, and can be manipulated through different operations on fuzzy sets or fuzzy numbers. Since imprecise parameters are treated as imprecise values instead of precise ones, the process will be more powerful and its results more credible. Fuzzy sets theory has been studied extensively over the past 40 years (for more interest see [4, 6, 11]) and satisfactory applied to problems.

There is not any bibliometric study on fuzzy sets theory. As it is known the fuzzy Spanish community is a very important group in the EUSFLAT association.

[1] This work has been supported by the projects: FUZZY-LING, Cod. TIN2007-61079 and Proyecto de Excelencia de la Junta de Andalucia SAINFOWEB, Cod. 00602.

In this paper, a bibliometric study of the fuzzy sets theory in Spain is presented. It is done by means of bibliometric maps that show the associations between the main concepts studied in Spanish fuzzy community. The maps provide insight into the structure of the fuzzy sets theory field, visualize the division of the field into several subfields and indicate the relations between these subfields. More concretely, we will study the Spanish fuzzy sets theory field using the concept of longitudinal mapping, introduced by Garfield [5]. In longitudinal mapping, a series of chronologically sequential maps can be used to detect the advances of scientific knowledge and the evolution of the field over the last years. While maps of current data alone cannot predict where research will go, longitudinal maps can be useful indicators in the hands of informed analysts and domain experts. Analysts and domain experts can use longitudinal maps to forecast emerging trends for a subject domain. Since domain visualizations typically reference key works in a field, they are a good tool to enable the novice to become instantly familiar with a field through easy identification of key topics and theirs relationships. The bibliometric maps are created by the co-word analysis [2, 9].

This paper is organized as follows. Section 2 introduces the co-word analysis methodology. Section 3 describes the achieved results. Finally, in Section 4 some conclusions are drawn.

2 The Co-Words Analysis: Methodology

Co-word analysis is a content analysis technique that is effective in mapping the strength of association between information items in textual data. Co-word analysis reduces a space of descriptors (or keywords) to a set of network graphs that effectively illustrate the strongest associations between descriptors [2, 9].

This technique illustrates associations between keywords by constructing multiple networks that highlight

associations between keywords, and where associations between networks are possible.

The basic assumption in bibliometric mapping [1] is that each research field can be characterized by a list of the important keywords. Each publication in the field can in turn be characterized by a sub-list of these global keywords. Such sub-lists are like DNA fingerprints of these published articles. By matching keyword-based fingerprints, one can measure the similarity between a pair of publications. The more keywords two documents have in common, the more similar the two publications are, and the more likely they come from the same research area or research specialty at a higher level. Following the DNA metaphor, if two publications fingerprints are similar enough, they are bound to come from the same species.

According to Börner et al. [1], the process of constructing a bibliometric map can be divided into the following six steps: (1) collection of raw data, (2) selection of the type of item to analyze, (3) extraction of relevant information from the raw data, (4) calculation of similarities between items based on the extracted information, (5) positioning of items in a low-dimensional space based on the similarities, and (6) visualization of the low-dimensional space. We now discuss the way in which we implement each of these steps in this paper.

The first step in the process of bibliometric mapping is the collection of raw data. In this paper, the raw data consist of a corpus containing 1.524 papers, about fuzzy sets theory in Spain, are extracted from ISI Web of Science (ISIWoS) [8] with the query #1 on 4th April 2008:

```
query #1: (TI=fuzzy OR TS=(("fuzzy logic") OR ("fuzzy set*") OR
("fuzzy rule*") OR ("fuzzy syst*") OR ("fuzzy relat*") OR ("fuzzy
logic cont*") OR ("fuzzy set*") OR ("fuzzy prog*") OR ("fuzzy
model*") OR ("fuzzy data*") OR ("fuzzy preference*") OR ("fuzzy
linguistic") OR ("linguistic preferenc*") OR (fuzziness) OR ("fuzzy
neural*") OR ("neural fuzzy") OR ("fuzzy math") OR ("fuzzy goal") OR
("fuzzy inform*") OR ("fuzzy number*") OR ("fuzzy differen*") OR
("fuzzy languag*")) OR SO=(("FUZZY SETS AND SYSTEMS") OR ("IEEE
TRANSACTIONS ON FUZZY SYSTEMS"))) AND CU=SPAIN.
```

TI field is a search based on the "Title", TS field is a search based on the "Topic", and CU field is a search based on the "Country", Spain in this case. In Figure 1 the number of Spanish papers in ISIWoS from 1978 to 2008 are shown.

Four sets of data are collected, one for each studied period: 1978-1992, 1993-1997, 1998-2002 and 2003-2007. In this way, separate bibliometric maps can be constructed for each of the four periods. The number of papers per period are shown in Figure 2.

The second step in the process of bibliometric mapping is the selection of the type of item to analyze. According to Börner et al. [1], journals, papers, authors,

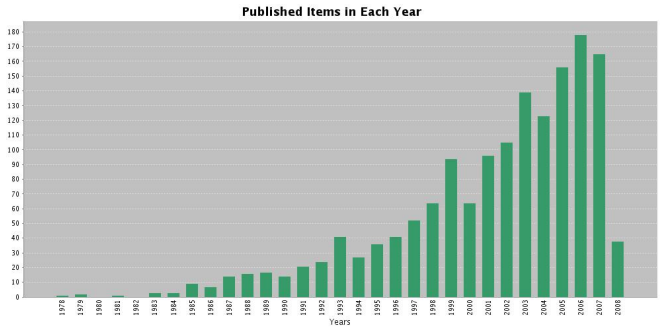


Figure 1: Spanish papers in ISIWoS from 1978 to 2008.

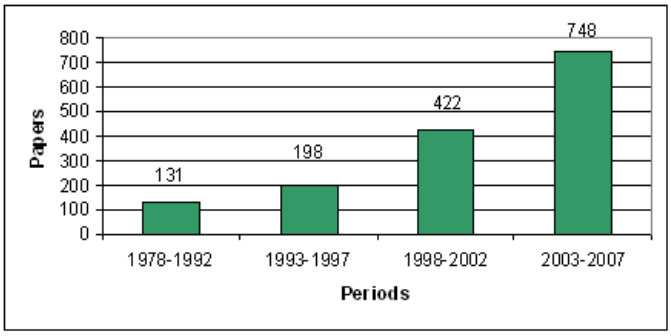


Figure 2: Spanish papers per period.

and descriptive terms or words are most commonly selected as the type of item to analyze. In this paper, we choose to analyze descriptive words, and more concretely, keywords. A bibliometric map showing the associations between keywords in a scientific field is referred to a keywords based map in this paper.

The third step in the process of bibliometric mapping is the extraction of relevant information from the raw data collected in the first step. In this paper, the relevant information consists of the co-occurrence frequencies of keywords. The co-occurrence frequency of two keywords is extracted by counting the number of papers in which the two keywords both occur in the keywords section[2] (see ISIWoS description sections [8]).

The fourth step in the process of bibliometric mapping is the calculation of similarities between items based on the information extracted in the third step. In this paper, similarities between items are calculated based on frequencies of keywords co-occurrences. When two keywords frequently occur together, they are said to be linked, and the intensity of the link is indicated by the equivalency index [7], $e_{ij} := \frac{c_{ij}^2}{c_i \cdot c_j}$, where

[2] If the keywords section is empty, the paper is not considered in the mapping process.

c_{ij} is the number of documents in which two keywords i and j co-occur and c_i and c_j represent the number of documents in which each one appears. When the keywords appear together, the equivalency index equals 1; when they are never associated, it equals 0. Once the links are quantified, by an algorithm called *simple centers*, groupings or themes, consisting of more strongly linked networks that represent the centers of interest of the researchers, are produced.

The simple centers algorithm uses two passes through the data to produce the desired networks. The first pass constructs the networks depicting the strongest associations, and links added in this pass are called internal links. The second pass adds to these networks links of weaker strengths that form associations between networks. The links added during the second pass are called external links.

Note that two keywords that appear infrequently in the corpus but always appear together will have larger strength values than keywords that appear many times in the corpus almost always together. Hence, possibly irrelevant or weak associations may dominate the network. A solution to this problem incorporated into the algorithm described in this section is to require that only the keyword pairs that exceed a minimum co-occurrence[3] are considered potential links while building networks during the first pass of the algorithm.

During the first pass, the link that has the largest strength is selected first, its nodes becoming the starting nodes of the first pass-1 network. Other links and their corresponding nodes are added to the graph using a breadth-first search on the strength of the links (i.e. the strongest link connecting a node that is not in any graph to the graph being constructed is added first), until there are no more links that exceed the co-occurrence threshold, or a maximum pass-1 link limit is exceeded. The next network is generated in a similar manner starting with the link with the largest strength that is not in any existing graph.

Networks are interconnected by pass-2 links. The *centrality* of a network measures the degree of interaction to other networks and is defined as: $c = 10 * \sum e_{kh}$, with k a keyword belonging to the theme and h a keyword belonging to other themes.

The *density* of a network measures the internal strength of the network and is defined as: $d = 100\frac{\sum e_{ij}}{w}$, with i and j keywords belonging to the theme and w el number of keywords in the theme.

Isolated Networks are those that have low centrality values. *Principal Networks* are those that have high centrality and high density values.

As described in [2], the algorithm for the generation of the networks is as follows:

1. Select a minimum for the number of co-occurrences, c_{ij} , for keywords i and j, select maxima for the number of pass-1 links, and select maxima for the total (pass-1 and pass-2) links.
2. Start pass-1;
3. Generate the highest e_{ij} value from all possible keywords to begin a pass-1 network;
4. From that link, form other links in a breadth-first manner until no more links are possible due to the co-occurrence minima or to pass-1 link or node maxima. Remove all incorporated keywords from the list of subsequent available pass-1 keywords;
5. Repeat steps 3 and 4 until all pass-1 networks are formed; i.e., until no two remaining keyword pairs co-occur frequently enough to begin a network;
6. Start pass-2;
7. Restore all pass-1 keywords to the list of available keywords;
8. Start with the first pass-1 network.
9. Generate all links to pass-1 nodes in the current network to any pass-1 nodes having at least the minimal co-occurrences in descending order of e_ij value; stop when no remaining keyword pairs meet the co-occurrence minima, or when the total link maxima is met. Do not remove any keyword from the available list;
10. Select the next succeeding pass-1 network, and repeat step 9.

The fifth step in the process of bibliometric mapping is the positioning of items in a low-dimensional space based on the similarities calculated in the fourth step. In this paper, the low-dimensional space is referred to a keywords based map and only two-dimensional keywords based maps are considered. The two dimensions are centrality rank (cr) and density rank (dr), calculated as: $cr = \frac{rank_i^c}{N}$; $dr = \frac{rank_i^d}{N}$, where $rank_i^c$ is the position of the theme i in the theme list in ascending sort of centrality, and $rank_i^d$ is the position of the theme i in the theme list in ascending sort of density. N is the number of themes in the whole network. N is introduced to normalize in $[0, 1]$ the cr and dr values.

The sixth step in the process of bibliometric mapping is the visualization of the low-dimensional space that results from the fifth step. In our study, we use CoPalRed [3] computer program, which visualizes the networks in a strategic diagram. A strategic diagram of the network is presented in Figure 3. The abscissa axis is centrality, or the external cohesion index. It represents the most or least central position within the overall network. The ordinate axis is density, or the index of internal cohesion. It represents the conceptual development of the theme. Papers in the upper-right quadrant (quadrant 1) are considered motor themes of the speciality, given that they present strong centrality and high density. Those of the lower-right quadrant (quadrant 2) are the most general basic themes, although with internal development not as high as those

[3] As each period has different number of papers, different minimum co-occurrence values are used: 1, 2, 5 and 5 per periods 1978-1992, 1993-1997, 1998-2002 and 2003-2007 respectively.

of quadrant 1. The themes of quadrant 3 (upper-left) are very specialized, but peripheral in character, while those of the lower-left (quadrant 4), with low density and little centrality, mainly represent either emerging or disappearing themes.

CoPalRed visualizes a keywords based map by displaying for each keyword a label that indicates the location of the theme in the strategic diagram.

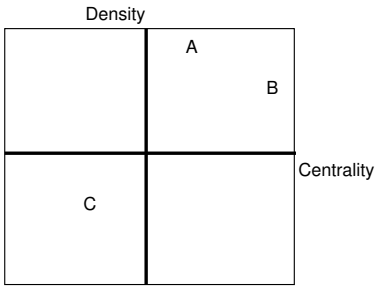


Figure 3: A example of strategic diagram.

As example, in the strategic diagram in Figure 3, the themes A, B and C are presented. The placement of theme B implies that this theme is quite related externally to concepts applicable to other themes that are conceptually closely related. In addition, the internal cohesion is also rather strong, and therefore we can consider B to be a *motor-theme* of the studied scientific field.

3 Visual Evolution of the Fuzzy Sets Theory Field in Spain

To analyze the evolution of the Fuzzy Sets Theory field over the last thirty years in Spain, strategic diagrams for the four studied subperiods (1978-1992, 1993-1997, 1998-2002 and 2003-2007) are shown in Figure 4. The volume of the spheres is proportional to the number of documents corresponding to each theme in each period. In the following the four periods are described.

In the first period (1978-1992), the longer one, in which only 131 papers were published, the most important themes in Spain, in relative weight in terms of number of documents, were: *fuzzy-integral*, *medicine*, *decision-making* and *t-norm* (see Figure 4).

Because of their strategic situation (upper-right quadrant), the themes *fuzzy-integral*, *medicine* and *t-norm* are supposed very conceptually developed, whereas, *decision-making* (lower-left quadrant) is considered as an incipient discipline, in development process.

In the second period (1993-1997), with 198 papers, *group-decision-making* was the theme more developed (upper-right quadrant). The theme was supported by

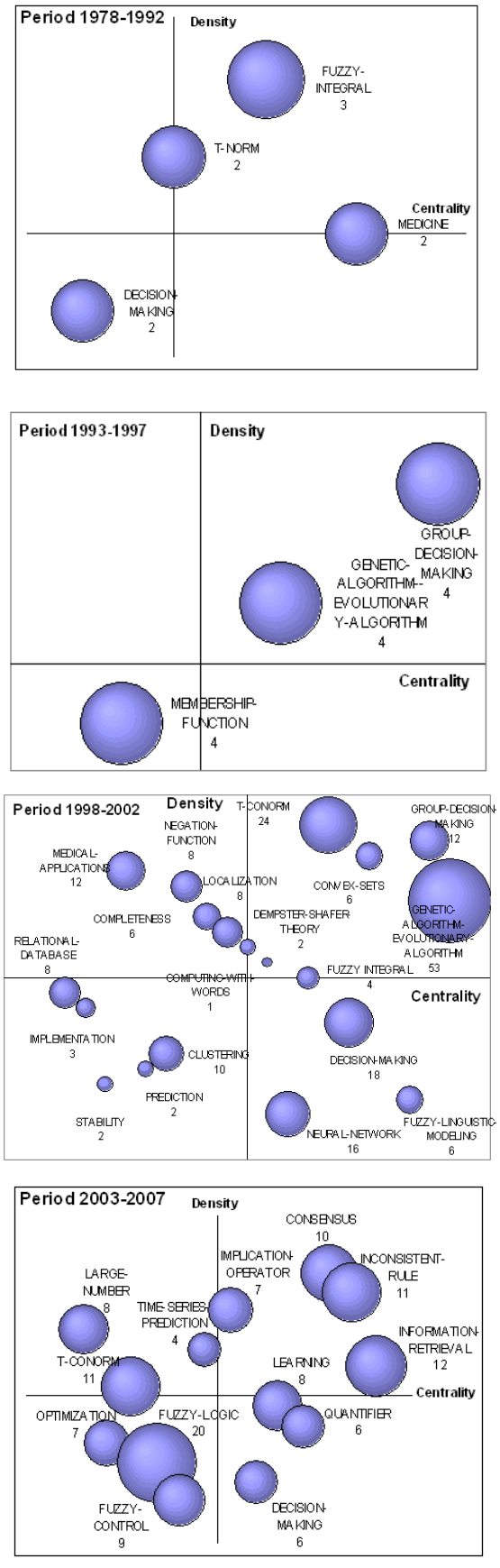


Figure 4: Strategic diagrams for the periods: 1978-1992, 1993-1997, 1998-2002 and 2003-2007.

12 papers[4].

In the last decade, from 1998 to 2007, a significative increment in published papers is observed (see Figures 1 and 2). In these years (periods 1998-2002 and 2003-2007) the number of interest topic of the Spanish fuzzy community has increased, and a set of different themes has been observed by CoPalRed (see Figure 4).

From 1998-2002, the most studied themes were: *genetic-algorithm-evolutionary-algorithm* (53 papers), *t-conorm* (24 papers), *group-decision-making* (12 papers), *decision-making* (18 papers) and *neural-network* (16 papers). All of them were strategically localized in the right quadrants (with high centrality indexes) of the strategic diagram for this period, i.e, they were quite related externally to concepts applicable to other themes that were conceptually closely related. In numeric terms, the macro-theme composed by the union of these five themes were supported by 168 different papers, almost 40% of the 422 published papers in this period.

From 2003 to 2007 the studied themes in Spain were those shown in the strategic diagram of Figure 4. In this last period, the principal themes, in number of papers, were *fuzzy-logic* (20 papers), *information-retrieval* (12 papers), *inconsistent-rule* (11 papers) and *consensus* (10 papers). In this period, we can appreciate did not exist so big themes than period 1998-2002. In this period the number of published papers (748) was more distributed around the themes.

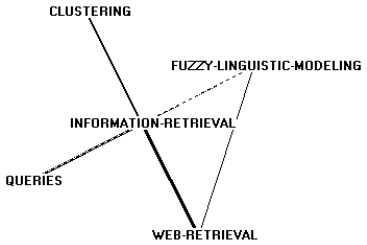


Figure 5: information-retrieval thematic network (Period 2003-2007).

From Figure 4 we can appreciate several conclusions in the evolution of the most long-lasting themes in the fuzzy sets based research in Spain:

- Theme *decision-making* and *t-norm* are the most long-lasting themes.
- *decision-making* has augmented its centrality ratio, moving from the lower-left quadrant to lower-right one. This indicates *decision-making* has consolidated as a general basic theme in the last ten years (see Figure 4).
- In the last period (2003-2007), *information-retrieval* has emerged as import theme in Spanish fuzzy community. It has became in one of the most studied themes with 4 papers (see thematic network in Figure 5).
- *t-conorm* has been re-localized to quadrant upper-left in the strategic diagram in the last five years. This movement indicates us that the theme *t-conorm* has lost "strength" or interest in the last five years in the Spanish Fuzzy Set research community, from a very dense and central theme in period 1998-2002 to peripheral theme in period 2003-2007. Its thematic network has stayed almost stable in the last ten years, having a strong union among the sub-themes *t-norm* and *uninorm*. See Figure 6.

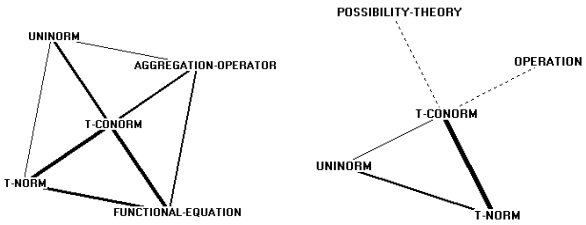


Figure 6: Theme *t-conorm*: Period 1998-2002 left, Period 2003-2007 right.

- Theme *group-decision-making* has been muting from the period 1993-1997. In the last period, 2003-2007, the thematic network has been rotated and the sub-theme *consensus* has emerged as main theme (see Figure 7). The number of papers has also varied, being the most fertile period 1998-2002 with 12 papers.
- Theme *genetic-algorithm–evolutionary-algorithm* has been rotated and split in two thematic network in the last five years. In this partition process, sub-themes *inconsistent-rule* and *learning* have flowered as main theme and a new thematic network has been created (see Figure 8).

Finally, other themes sited initially in the lower-left quadrant, have disappeared in these last years (period 2003-2007). This is the case of theme *relational-database* (see thematic network in Figure 9), which seems have vanished from period 1998-2002 (with 8 papers) to period 2003-2007, in which the keyword relational-database has been only used 3 times.

[4] CoPalRed assumes a paper belong to a theme when it presents at least 2 keywords of the theme.

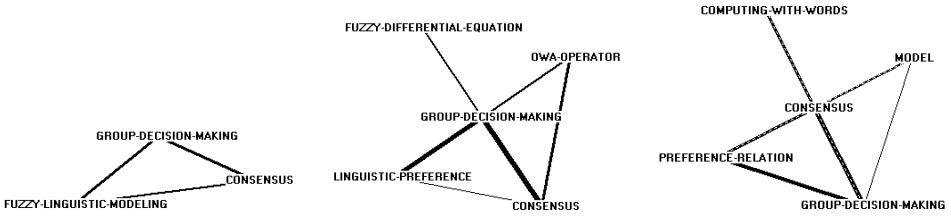


Figure 7: Evolution of the theme *GROUP-DECISION-MAKING*: Left) Period 1993-1997, Center) Period 1998-2002 and Right) Period 2003-2007.

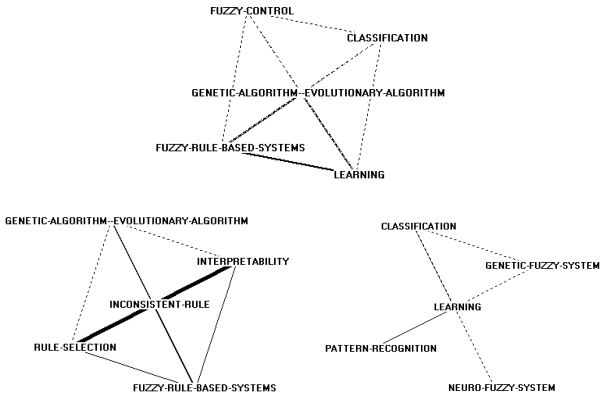


Figure 8: Themes *GENETIC-ALGORITHM-EVOLUTIONARY-ALGORITHM* (period 1998-2002), *INCONSISTENT-RULE* (period 2003-2007) and *LEARNING* (period 2003-2007).

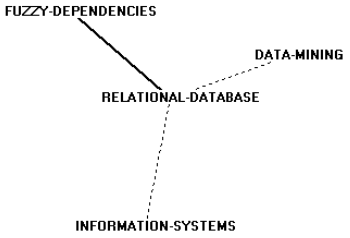


Figure 9: RELATIONAL-DATABASE thematic network (Period 1998-2002).

4 Conclusions

In this paper, we have presented a bibliometric study of the Fuzzy Sets Theory Field in Spain. More than 1.500 Spanish research papers have been processed. Based on our analysis, we have drawn the visual structure of the Spanish Fuzzy Sets Theory field.

We have been able to observe how several themes have emerged, disappeared, muted, re-localized or divided into other themes.

References

[1] K. Börner, C. Chen, and K.W. Boyack, *Visualizing knowledge domains*, Annual Review of Information Science and Technology **37** (2003), 179–255.

[2] N. Coulter, I. Monarch, and S. Konda, *Software engineering as seen through its research literature: A study in co-word analysis*, Journal of the American Society for Information Science **49** (1998), no. 13, 1206–1223.

[3] Grupo Evaluacin de la Ciencia y la Comunicacin Cientfica. Universidad de Granada, *http://ec3.ugr.es/*.

[4] D. Dubois and H. Prade, *Possibility theory: An approach to computerized processing of uncertainty*, Plenum Press: New York, 1985.

[5] E. Garfield, *Scientography: Mapping the tracks of science.*, Current Contents: Social & Behavioural Sciences **7** (1994), no. 45, 5–10.

[6] G.J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: Theory and applications*, Prentice Hall, 1995.

[7] B. Michelet, *L'analyse des associations*, Ph.D. thesis, University of Paris VII, 1988.

[8] ISI Web of Science, *http://scientific.thomson.com/products/wos/*.

[9] J. Whittaker, *Creativity and conformity in science: Titles, keywords, and co-word analysis*, Social Science in Science **19** (1989), 473–496.

[10] L.A. Zadeh, *Fuzzy sets*, Information and Control **8** (1965), no. 3, 338–353.

[11] H.J. Zimmermann, *Fuzzy sets: Theory and its applications*, Kluwer Academic, 1996.

IMPLEMENTACIÓN HARDWARE DE UNA TÉCNICA DE SEGMENTACIÓN DE IMÁGENES BASADA EN UMBRAL APLICANDO LÓGICA DIFUSA

Nashaat M. Hussein **Angel Barriga**

Instituto de Microelectrónica de Sevilla (CNM-CSIC)/Universidad de Sevilla
nashaat@imse.cnm.es
barriga@imse.cnm.es

Resumen

Esta comunicación describe una técnica para realizar la segmentación de imágenes basada en la umbralización binaria así como su implementación hardware. La técnica aplica lógica difusa en el cálculo del umbral. La metodología de diseño hardware empleada se basa en una descripción de alto nivel del comportamiento del sistema difuso. La base de conocimientos se describe mediante el lenguaje estándar de descripción de hardware VHDL. El criterio de diseño hardware ha sido el bajo coste del circuito y, de manera especial, la alta velocidad de procesado.

Palabras Clave: Segmentación de imágenes, umbralización de imágenes, descripción VHDL de sistemas difusos.

1 INTRODUCCIÓN

El primer paso en cualquier proceso de análisis de imagen es la segmentación. Mediante la segmentación se divide la imagen en las partes u objetos que la constituyen. Para ello la imagen se divide en regiones de las que se extraen una serie de parámetros que permiten clasificar las distintas regiones de la imagen en los distintos objetos que la constituyen. En el caso de considerar una única región la imagen se divide en objeto y fondo. El nivel al que se realiza esta subdivisión depende de la aplicación en particular, es decir, la segmentación terminará cuando se hayan detectado todos los objetos de interés para la aplicación. La segmentación es un paso imprescindible en diversos procesos de tratamiento de imagen. Entre otros, es necesaria para tomar medidas sobre una región, para realizar reconstrucciones tridimensionales de una zona de la imagen, para la clasificación o diagnóstico automático o para reducir la información de las imágenes.

Los algoritmos de segmentación de imagen generalmente se basan en dos propiedades básicas de los niveles de gris de la imagen: discontinuidad y similitud. Dentro de la primera categoría se intenta dividir la imagen basándonos en los cambios bruscos en el nivel de gris. Un ejemplo dentro de esta categoría es la detección de bordes en la imagen. Dentro de la segunda categoría se aplican técnicas de umbrales, crecimiento de regiones, y técnicas de división y fusión.

El problema más sencillo de segmentación se presenta cuando la imagen está formada por un sólo objeto que tiene intensidad luminosa homogénea sobre un fondo con un nivel de intensidad diferente. En este caso la imagen se puede segmentar en dos regiones utilizando una técnica basada en un parámetro umbral. Esta técnica de aplicar un valor de umbral es un método sencillo pero eficiente de separar los objetos del fondo. Ejemplos de aplicaciones en los que se emplea la umbralización son el análisis de documentos, procesado de mapas, imágenes de endoscopia, extracción de bordes, detección de objetos, etc. La elección del valor del umbral se puede hacer a partir del histograma. Si el fondo tiene también intensidad luminosa homogénea, entonces el histograma es bimodal y el umbral que se debe tomar es el que corresponde al mínimo local que está entre los dos máximos del histograma. En ocasiones no se distinguen bien las dos zonas modales (los máximos) y se debe aplicar una técnica de variación espacial del umbral.

Muchos de los algoritmos de umbralización se basan en una umbralización binaria. Los procedimientos de umbralización binario pueden extenderse a multi-nivel disponiendo de múltiples umbrales T_1, T_2,…,T_n para segmentar la imagen en $n+1$ regiones [1-3]. La umbralización multi-nivel basada en un histograma multidimensional aplica técnicas de crear agrupamientos (*pattern clustering*) para realizar la segmentación de imágenes.

Entre las técnicas de umbralización las que se basan en aplicar lógica difusa permiten resolver el problema de la información imprecisa de la imagen [5-7]. En esta

comunicación nos vamos a centrar en la umbralización binaria. La técnica que aplicamos se basa en aplicar lógica difusa. Nuestro interés se centra en la realización hardware del algoritmo. El artículo se organiza en cuatro secciones. En la sección 2 se discuten las estrategias de umbralización de imágenes. En la sección 3 se discute el método que se presenta en esta comunicación. En la sección 4 se realiza el diseño del circuito que realiza el cálculo del umbral.

2 TÉCNICAS DE UMBRALIZACIÓN DE IMÁGENES

Las técnicas basadas en umbralizar una imagen permiten clasificar los pixels en dos categorías (imagen en blanco y negro). Esta transformación se realiza para establecer una distinción entre los objetos de la imagen y el fondo. El modo de generar esta imagen binaria es realizando la comparación de los valores de los pixels con un umbral *T*, es decir cualquier valor inferior a este valor del umbral se considera un objeto mientras que los valores superiores al umbral corresponde al fondo. Así sea *G* el conjunto de valores de los pixels de la imagen:

$$G = \{0,1,...,L-1\} \begin{Bmatrix} 0 & negro \\ L-1 & blanco \end{Bmatrix}$$

Se define un umbral $T \in G$ de manera que se realiza la transformación que viene dada de la siguiente manera:

$$y_{i,j} = \begin{cases} 0 & si \quad x_{i,j} < T \\ L-1 & si \quad x_{i,j} > T \end{cases} \qquad (1)$$

donde $x_{i,j}$ es un pixel de la imagen original e $y_{i,j}$ es el pixel correspondiente a la imagen binaria. En el caso de una imagen monocolor en la que los pixels se codifican con 8 bits el rango de valores que toman los pixels corresponde al rango entre 0 y 255 (*L=256*). Es usual expresar dicho rango en valores normalizados entre 0 y 1. La figura 1 muestra un ejemplo para el caso de la imagen de Lena.

a) b)

Figura 1: a) Imagen de Lena, b) imagen binaria con *T=0.5*

Existen muchos métodos para seleccionar el valor del umbral *T*. En [8] se analizan 40 métodos de selección del umbral. Estos métodos se clasifican en seis categorías de acuerdo con la información que es empleada: métodos basados en la forma del histograma, agrupamientos, entropía, métodos basados en atributos de los objetos, métodos espaciales y locales.

Los métodos de cálculo del umbral más extendidos son los basados en el análisis del histograma. El histograma es una relación entre los niveles de grises de una imagen con la frecuencia de los valores de grises de la imagen. La figura 2 muestra el histograma de la imagen de Lena (fig 1a).

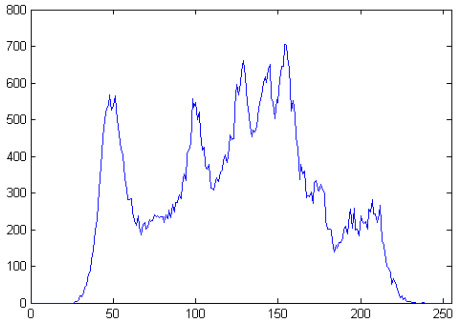


Figura 2: Histograma de la imagen de Lena.

Una técnica básica de cálculo del umbral se basa en la frecuencia de nivel de gris. En este caso el umbral *T* se calcula mediante la siguiente expresión:

$$T = \sum_{i=1}^{L} p_i i \qquad (2)$$

donde *i* es el nivel de gris que toma valores entre 1 y 256 para el caso de una codificación con 8 bits, p_i representa la frecuencia de nivel de gris (también denominada probabilidad de nivel de gris). Para una imagen con *n* pixels y n_i pixels con el nivel de gris *i*:

$$p_i = n_i / n \quad y \quad \sum_{i=1}^{L} p_i = 1 \qquad (3)$$

Las técnicas que aplican lógica difusa para el cálculo del umbral se basan principalmente en tres tipos de medidas de vaguedad [4]: entropía, medida de Kaufmann y medida de Yager.

La técnica basada en la entropía consiste en minimizar la dispersión del sistema. Así los pixels de la imagen se agrupan en dos clases correspondientes a los objetos y al fondo. Huang y Wang [5] consideran la media de los datos correspondientes a cada clase es μ_0 y μ_1. La función de pertenencia de cada clase viene definida por:

$$u_x(x)=\begin{cases}\dfrac{1}{1+\dfrac{|x-\mu_0|}{x_{\max}-x_{\min}}} & si \quad x<T\\ \dfrac{1}{1+\dfrac{|x-\mu_1|}{x_{\max}-x_{\min}}} & si \quad x>T\end{cases}$$

El cálculo del umbral *T* se basa en la entropía de un conjunto difuso, que se calcula usando la función de Shannon:

$$H_f(x)=-x\log x-(1-x)\log(1-x)$$

El umbral será aquél que minimice la entropía de los datos:

$$E(T)=\frac{1}{M}\sum_i H_f(\mu_x(i))h(i)$$

La medida de Kaufmann se define como [6]:

$$D(A)=\left[\sum_{x\in X}\left|\mu_A(x)-\mu_C(x)\right|^w\right]^{\frac{1}{w}}$$

Este método se basa en usar la distancia al conjunto *A* como métrica. Cuando *w=1* se utiliza la distancia de Hamming mientras que si *w=2* se trata de la distancia Euclidea.

El método de Yager [7] se basa en la distancia entre un conjunto difuso y su complementario. Así se basa en minimizar la siguiente función:

$$D_2(T)=\sqrt{\sum_i\left|\mu_x(i)-\mu_{\bar{x}}(i)\right|^2}$$

donde $\mu_{\bar{x}}(i)=1-\mu_x(i)$.

3 UMBRALIZACIÓN BASADA EN LÓGICA DIFUSA

La técnica que se propone en esta presentación consiste en aplicar lógica difusa en el cálculo del umbral *T*. Básicamente, desde un punto de vista formal, esta técnica se basa en el cálculo de la media aplicada al histograma de la imagen. Un aspecto ventajoso de esta técnica es que el mecanismo de cálculo mejora los tiempos de cómputo ya que tan sólo requiere recorrer una vez la imagen y permite calcular de manera directa el valor del umbral. Desde el punto de vista de la realización hardware se dispone de una arquitectura de elemento de procesado difuso de bajo coste que permite realizar la inferencia tal y como se verá en una próxima sección.

El sistema difuso tiene una entrada que corresponde al pixel que se va a evaluar y una salida que corresponde al resultado de la inferencia difusa. Una vez leída la imagen completa la salida muestra el valor del umbral *T*. Básicamente la operación que realiza el sistema difuso corresponde al cálculo del centro de gravedad del histograma de la imagen de acuerdo con la siguiente expresión:

$$T=\frac{\sum_{i=1}^{M}\sum_{j=1}^{R}\alpha_{ij}c_{ij}}{\sum_{i=1}^{M}\sum_{j=1}^{R}\alpha_{ij}} \qquad (4)$$

donde *T* es el umbral, *M* es el número de pixels de la imagen, *R* es el número de reglas del sistema difuso, *c* es el consecuente de cada regla y α es el grado de activación de la regla.

Para realizar la inferencia difusa se realiza un particionado del universo de discurso del histograma en un conjunto de *N* funciones de pertenencia equidistribuidas. La figura 3 muestra un ejemplo de particionado para *N=9*. Se han empleado funciones de pertenencia triangulares ya que son más fácilmente implementables en hardware. Dichas funciones tienen un solapamiento de 2 lo que permite limitar el número de reglas activas.

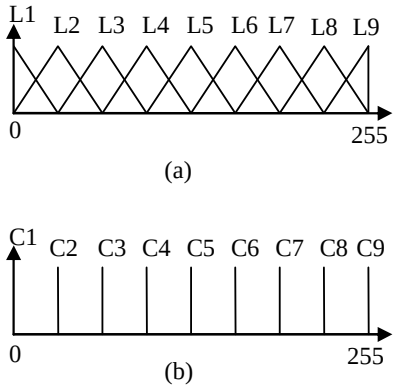


Figura 3: Funciones de pertenencia para *N=9*, a) antecedente, b) consecuente.

Las funciones de pertenencia del consecuente son funciones de tipo *singleton* equidistribuidas en el universo de discurso del histograma. El uso de funciones de

pertenencia de tipo *singleton* para el consecuente permite aplicar métodos de defuzzificación simplicados como es el caso de la media difusa. Este método de defuzzificación puede interpretarse como aquél en el que cada regla propone una conclusión con una determinado "peso" definido por su grado de activación. La acción global de las reglas se obtiene calculando el promedio de las diferentes conclusiones ponderadas por sus grados de activación. Estas características de procesado basado en reglas activas y método de defuzzificación simplificados permiten obtener realizaciones hardware de bajo coste y alta velocidad de operación.

La base de reglas del sistema de la figura 4 emplea las funciones de pertenencia mostradas en la figura 3. Observamos que el número de reglas viene dado por el número de funciones de pertenencia. La base de conocimiento (funciones de pertenencia y base de reglas) es común para todas las imágenes por lo que los valores pueden almacenarse en una memoria tipo ROM.

if x is L1 then c is C1;
if x is L2 then c is C2;
if x is L3 then c is C3;
if x is L4 then c is C4;
if x is L5 then c is C5;
if x is L6 then c is C6;
if x is L7 then c is C7;
if x is L8 then c is C8;
if x is L9 then c is C9;

Figura 4: Base de reglas para *N=9*.

Es posible optimizar la expresión de la ecuación (4) si el sistema está normalizado. En este caso la suma extendida a la base de reglas de los grados de activación toma el valor 1 ($\sum_{j=1}^{R} \alpha_{ij} = 1$). Por lo tanto la ecuación (4) se transforma en:

$$T = \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{R} \alpha_{ij} c_{ij} \qquad (5)$$

El sistema evalúa en cada instante de tiempo un pixel y realiza la inferencia de acuerdo con la base de reglas de la figura 4. La salida del sistema acumula el resultado correspondiente al numerador de la ecuación (5). La salida final se genera con el último pixel de la imagen que permite realizar la división.

4 IMPLEMENTACIÓN HARDWARE

A la hora de implementar el sistema difuso que genera el umbral se ha empleado la metodología de diseño descrita en [9]. Dicha metodología se basa en el lenguaje de descripción de hardware VHDL como la manera de describir y modelar el sistema en alto nivel. Con objeto de conseguir un modelado del comportamiento del módulo de inferencia difuso se ha optado por un estilo de descripción VHDL en el que se describen de manera independiente la estructura del sistema (conjuntos difusos, base de regla) y los operadores del sistema (conectivas, operaciones difusas). Esto permite independizar la descripción de la estructura del sistema de los mecanismos de procesado. La descripción del sistema difuso es sintetizable de manera que es posible generar la realización hardware haciendo uso de herramientas convencionales de síntesis de circuitos.

La figura 5 muestra la arquitectura VHDL del sistema difuso. La base de reglas se describe en el cuerpo de la arquitectura. Corresponde a la base de reglas de 9 reglas de la figura 4. Cada regla está constituida por dos componentes: antecedente de la regla y consecuente. El antecedente es una expresión de la variable de entrada relacionada con la etiqueta lingüística. El consecuente establece el valor lingüístico de la salida de la regla.

```
architecture knowledge_base of threshold is
 signal R: consec;
 -- MF for x
 constant L1: triangle:=((0,   0,  31),(0, 32));
 constant L2: triangle:=((0,  31,  63),(32,32));
 constant L3: triangle:=((31, 63,  95),(32,32));
 constant L4: triangle:=((63, 95, 127),(32,32));
 constant L5: triangle:=((95,127, 159),(32,32));
 constant L6: triangle:=((127,159,191),(32,32));
 constant L7: triangle:=((159,191,223),(32,32));
 constant L8: triangle:=((191,223,255),(32,32));
 constant L9: triangle:=((223,255,255),(32,0));

 --MF for z
 constant C1: integer := 0;
 constant C2: integer := 31;
 constant C3: integer := 63;
 constant C4: integer := 95;
 constant C5: integer := 127;
 constant C6: integer := 159;
 constant C7: integer := 191;
 constant C8: integer := 223;
 constant C9: integer := 255;
begin
   R(1) <= rule( (x = L1), C1 );
   R(2) <= rule( (x = L2), C2 );
   R(3) <= rule( (x = L3), C3 );
   R(4) <= rule( (x = L4), C4 );
   R(5) <= rule( (x = L5), C5 );
   R(6) <= rule( (x = L6), C6 );
   R(7) <= rule( (x = L7), C7 );
   R(8) <= rule( (x = L8), C8 );
   R(9) <= rule( (x = L9), C9 );

   Zout<=defuzz(R);
end knowledge_base;
```

Figura 5: Arquitectura VHDL de la base de conocimientos.

El mecanismo de procesado de la operación difusa '*is*' (=) y la inferencia '*then*' (*rule*(,)) no están definidas en la descripción VHDL. Solamente se ha descrito la estructura de la base de conocimiento. De esta manera la descripción es de alto nivel ya que no asume ningún criterio específico de implementación. Tan solo describe la base de conocimiento en términos de comportamiento de la base de reglas.

Las etiquetas lingüísticas representan el rango de valores de los conjuntos difusos dentro del universo de discurso para las variables entradas y salidas. Estas etiquetas se describen mediante funciones con objeto de procesar el grado de pertenencia de un cierto valor de la variable. Las funciones de pertenencia asociadas a las etiquetas lingüísticas pueden ser triangulares o trapezoidales. La zona declarativa de la arquitectura VHDL de la figura 5 muestra las definiciones de dichas funciones de pertenencia.

El tipo de dato *triangle* se encuentra definido en un paquete VHDL denominado "xfvhdlfunc". Dicho tipo contiene las definiciones de los puntos que definen un triangulo así como las pendientes. Por su parte la base de reglas expresa el conocimiento de la figura 4. La función *rule* también está definida en el paquete VHDL. Dicha función realiza la inferencia de la regla. El conjunto de reglas se evalúa de manera concurrente ya que las instrucciones de asignación de señal en el cuerpo de la arquitectura VHDL son concurrentes. El operador "=" también ha sido redefinido aprovechando las propiedades de sobrecarga de funciones en VHDL. La figura 6 muestra el paquete VHDL con las definiciones de los tipos de datos y las nuevas funciones.

```
PACKAGE xfvhdfunc IS
 type points is array (1 to 3) of integer;
 type slopes is array (1 to 2) of integer;

 type triangle is record
    point: points;
    slope: slopes;
 end record;
. . .
 -- Consequents
 type two_int is array (0 to 1) of integer;
 type consec is array (1 to RULES) of two_int;

 -- FUNCTIONS
 function "=" (x: integer; y: triangle) return integer;
 . . .
 function rule (x: integer; y: integer) return two_int;
 function defuzz(x: consec) return integer;
END xfvhdfunc;
```

Figura 6: Paquete VHDL con las definiciones de los tipos de datos y de las funciones.

Las funciones empleadas en la descripción del sistema difuso han sido descritas de acuerdo con las restricciones de VHDL para síntesis. Esto ha permitido generar el circuito que implementa el sistema difuso haciendo uso de herramientas convencionales de síntesis de circuitos. El resultado corresponde a un circuito combinacional que realiza la inferencia difusa. Con objeto de reducir el tiempo de procesado se ha replicado el sistema difuso para que pueda operar sobre 4 pixels en cada ciclo de reloj. De esta manera el circuito recibe un dato de 32 bits que procede de una memoria que almacena la imagen. La figura 7 muestra el símbolo del circuito del módulo de inferencia difuso. Dicho módulo está constituido por los bloques correspondientes a las 9 reglas. La salida de sistema difuso corresponde al numerador de la etapa de defuzzificación de acuerdo con la ecuación (5). La figura 7b muestra el esquema del circuito FIM y la etapa acumuladora que realiza la suma extendida a todos los pixels de la imagen.

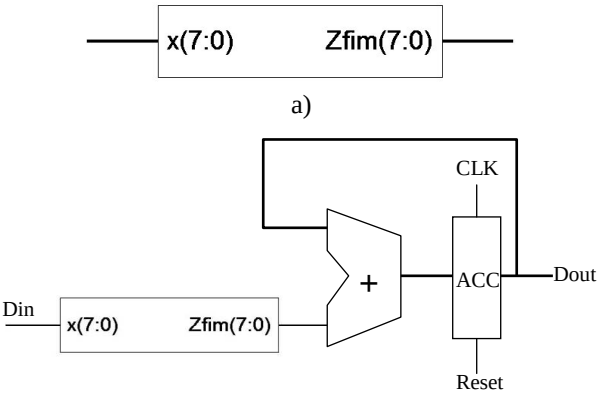


Figura 7: a) Símbolo del modulo de inferencia difuso (FIM), b) FIM y circuito acumulador.

El modulo de inferencia difuso junto con la etapa divisora generan el valor del umbral de la imagen. El circuito ha sido implementado sobre un dispositivo FPGA. El área ocupada sobre un FPGA de la familia Spartan3 de Xilinx ha sido de 1.180 *slices* que equivalen a un total de 56.845 puertas equivalentes. El circuito puede operar a una frecuencia de 137 MHz lo que permite procesar una imagen CIF (352x288 pixels) en 0,7 ms. En el caso de una imagen VGA (1024x768) el tiempo requerido para calcular el umbral les de 5,7 ms.

La figure 8 muestra algunos resultados para imágenes en formato CIF de 352x288 pixels. Los valores de los umbrales obtenidos se indican en cada caso.

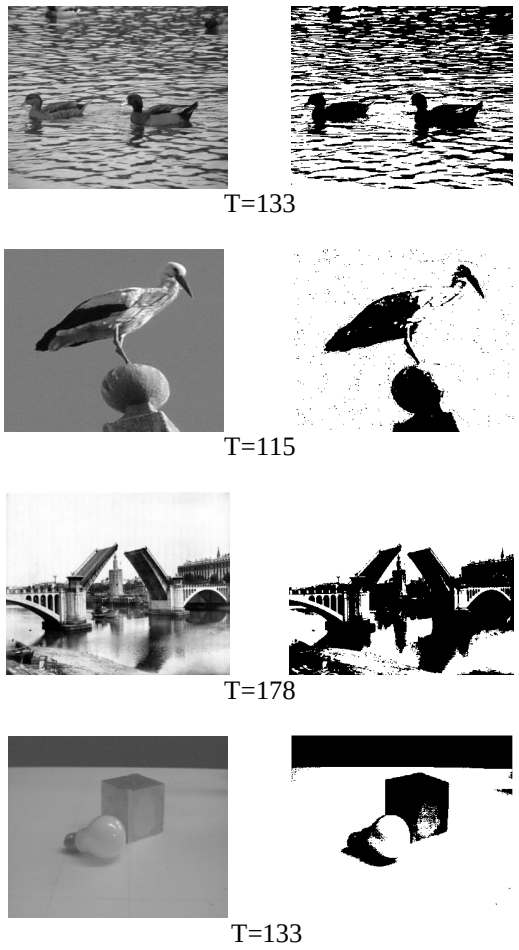


Figura 8: Imágenes de test y los resultados aplicando el método de calculo basado en lógica difusa.

5 CONCLUSIONES

Se ha propuesto una técnica de cálculo del umbral en imágenes basada en lógica difusa. Esta técnica permite realizar la segmentación binaria de imágenes. El principal objetivo del método de umbralización propuesto ha sido que sea eficiente desde el punto de vista de la implementación hardware en términos de coste y velocidad de procesado. Ello facilita el desarrollo de sistemas de análisis y procesado de imágenes en tiempo real que es el principal campo de aplicación de dicha técnica.

Agradecimientos

Este trabajo ha sido soportado en parte por el proyecto TEC2005-04359/MIC del Ministerio de Educación y Ciencia y por el proyecto TIC2006-635 de la Junta de Andalucía.

Referencias

1. P.-S., Liao, T.-S. Chen, P.C. Chung: A Fast Algorithm for Multilevel Thresholding. Journal of Information Science and Engineering, vol. 17, pp. 713--727 (2001)
2. L. Cao, Z.K. Shi and E.K.W. Chenp, Fast automatic multilevel thresholding method, Electronics Letters, vol. 38, no. 16, pp.868--870 (2002)
3. J.-T. Oh and W.-H. Kim, EWFCM Algorithm and Region-Based Multi-level Thresholding, LNAI 4223, pp. 864 -- 873 (2006)
4. Forero-Vargas, M.G., Rojas-Camacho, O., New formulation in image thresholding using fuzzy logic, 11th Portuguese Conf. on Pattern Recognition, pp. 117--124 (2000)
5. Huang, L.K., Wang, M.J., Image thresholding by minimizing the measure of fuzziness, Pattern Recognition, vol. 28, pp. 41--51, (1995)
6. Kaufmann, A., Introduction to the theory of fuzzy subsets, Academic Press (1975)
7. Yager, R.R., On the measure of fuzziness and negation. Part 1: membership in the unit interval. Int. Journal of Genet. Syst., vol. 5, pp. 221--229 (1979)
8. Sezgin, M., Sankur, B., Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging, 13, pp. 146--165 (2004)
9. A. Barriga, S. Sánchez Solano, Brox, P., A. Cabrera, I. Baturone, Modelling and Implementation of Fuzzy Systems based on VHDL, International Journal of Approximate Reasoning, vol. 41, issue 2, pp 164--278 (2006)

NUEVOS ALGORITMOS DE CLASIFICACIÓN INTEGRADOS EN XFUZZY 3

Francisco José Moreno Velo [1] **Iluminada Baturone** [2] **Santiago Sánchez Solano** [2] **Ángel Barriga** [2]

[1] Dpto Tecnologías de la Información (Universidad de Huelva), francisco.moreno@dti.uhu.es
[2] Instituto de Microelectrónica de Sevilla (CNM-CSIC), {lumi,santiago,barriga}@imse.cnm.es

Resumen

El entorno Xfuzzy 3 está formado por un amplio conjunto de herramientas dedicadas a dar soporte a las diferentes etapas del desarrollo de sistemas difusos. Entre estas herramientas se encuentra *Xfdm*, dedicada a la extracción de conocimiento difuso a partir de conjuntos de datos. Este trabajo presenta las últimas modificaciones realizadas en esta herramienta, que consisten en la integración de los algoritmos de clasificación difusos FuzzyID3 y FuzzConRI.

Palabras Clave: clasificadores difusos, algoritmos, herramientas CAD.

1 INTRODUCCIÓN

El entorno Xfuzzy [2] es una plataforma de desarrollo de sistemas difusos complejos. El entorno está formado por un conjunto de herramientas que dan soporte a las diferentes etapas de diseño de los sistemas difusos: edición, visualización, adquisición de conocimiento, optimización, simplificación, simulación, monitorización, síntesis software, síntesis hardware. Todas estas herramientas comparten una representación común del sistema en diseño basada en un lenguaje de especificación formal llamado XFL3. Xfuzzy soporta sistemas difusos complejos, con posibilidad de definir sistemas jerárquicos que pueden incluir módulos no difusos y expresiones lógicas complejas con modificadores lingüísticos y funciones difusas (operadores, funciones de pertenencia, métodos de *defuzzificación*) definidas por el usuario.

La herramienta Xfdm [1] está dedicada a la extracción automática de reglas difusas a partir de conjuntos de datos. Esta herramienta desarrolla los dos enfoques más utilizados en adquisición de reglas difusas: los basados en *clustering* y los basados en *grid*. El enfoque basado en *clustering* tiene como objetivo buscar grupos de datos entre los ejemplos, generando reglas difusas a partir de estos *clusters*. Las funciones de pertenencia de estas reglas se obtienen por proyección de los *clusters* sobre los universos de discurso de cada variable del sistema. Xfdm incluye algoritmos para generar *clusters* de forma esférica [3] y de forma elíptica con ejes libres [9] o paralelos a los ejes de coordenadas [8]. También incluye un algoritmo para generar los *clusters* de forma incremental [4]. Este enfoque tiene como ventaja que el número de reglas generadas es pequeño mientras que su desventaja principal es que las funciones de pertenencia generadas por proyección son difícilmente interpretables.

El enfoque basado en *grid* consiste en definir previamente las funciones de pertenencia de cada variable y encontrar las reglas utilizando estas funciones. El más conocido de los algoritmos que siguen este enfoque es el propuesto por Wang y Mendel [18] que recorre el conjunto de datos incluyendo entre las reglas a aquella que mejor cubre cada ejemplo. Los algoritmos basados en *grid* favorecen la interpretabilidad al utilizar funciones de pertenencia con un significado fijado previamente. Su principal desventaja es el gran número de reglas que generan. Para reducir el número de reglas algunos autores proponen escoger tan sólo las más significativas de las generadas por el algoritmo de Wang y Mendel [11][14], disminuyendo el tamaño del sistema generado al precio de aumentar el número de errores de clasificación del resultado.

La obtención de reglas lógicas para clasificar un conjunto de datos ha sido ampliamente tratada en el campo no difuso. Una de los enfoques en este campo es la inducción de árboles de decisión. Estos árboles permiten representar de forma compacta los procesos de decisión involucrados en las tareas de clasificación. En estos árboles las hojas representan la decisión final del proceso de clasificación (la clase seleccionada), los nodos internos describen la consulta acerca de un cierto atributo y las ramas de ese nodo se refieren a los diferentes valores que puede tomar dicho atributo. Los algoritmos de inducción de árboles de decisión se basan en seleccionar el atributo más adecuado para generar un nodo del árbol, en base al conjunto de ejemplos de entrada. El algoritmo CART utiliza el índice

Gini como medida para seleccionar el atributo. Por su parte, el algoritmo CHAID utiliza como medida el test de Pearson. El más conocido de estos algoritmos es ID3 [12], que utiliza como medida la entropía de información de cada atributo. El éxito de ID3 ha llevado a diversos autores a proponer versiones difusas de este algoritmo [5][15][16][19].

Entre los algoritmos clásicos dedicados a generar reglas de clasificación destaca la familia de algoritmos AQ [10]. La base de estos algoritmos consiste en seleccionar un ejemplo de clasificación positivo y utilizar sus valores para construir una regla consistente (es decir, que no cubra ningún ejemplo negativo) lo más general posible, repitiendo este procedimiento hasta cubrir a todos los ejemplos positivos. La consistencia de las reglas generadas provoca, sin embargo, que estos algoritmos sean muy sensibles al ruido. Para evitar estos problemas se propuso el algoritmo CN2 [6][7] en el que, en lugar de exigir consistencia a las reglas, estas se eligen utilizando criterios de representatividad estadística y de calidad. En el ámbito de los sistemas difusos, un enfoque similar se utiliza en el algoritmo FuzzConRI [17].

El algoritmo de Wang y Mendel y sus derivados generan reglas conjuntivas cuyos antecedentes incluyen valores de todas las variables de entrada del sistema. Esto provoca que el número de reglas crezca exponencialmente con el número de entradas. Por el contrario, los algoritmos FuzzyID3 y FuzzConRI potencian la generación de reglas en el que no todas las variables estén incluidas. Estas reglas cubren un espacio mucho mayor por lo que el resultado final es un sistema formado por un número de reglas más reducido, lo que mejora su interpretabilidad.

La estructura de este trabajo es la siguiente. El apartado 2 describe el algoritmo FuzzyID3, que genera árboles de decisión difusos. El apartado 3 describe el algoritmo FuzzConRI, que desarrolla una versión difusa del algoritmo clásico CN2. El apartado 4 presenta la implementación de los algoritmos anteriores en Xfuzzy. El apartado 5 muestra algunos resultados obtenidos por estos algoritmos sobre conjuntos de datos estándar. Las conclusiones de este trabajo se presentan en el apartado 6.

2 EL ALGORITMO FUZZYID3

El algoritmo ID3 forma parte de un conjunto de algoritmos de inducción de árboles de decisión conocidos como TDIDT (*Top Down Induction of Decision Trees*). El esquema general de estos algoritmos se presenta en la Figura 1.

Los algoritmos TDIDT construyen el árbol de decisión de forma recursiva. El caso base se produce cuando todos los ejemplos son de la misma clase (en cuyo caso se genera una hoja con dicha clase), cuando no quedan atributos (se genera una hoja con la clase mayoritaria) o cuando no existen ejemplos para clasificar (se genera una hoja con un valor por defecto). El caso general consiste en escoger el mejor atributo, atendiendo a una cierta función de evaluación, y generar un nodo interno con dicho atributo. Para generar las ramas de este nodo se aplica de nuevo el algoritmo al subconjunto de ejemplos de cada valor del atributo seleccionado. El algoritmo ID3 utiliza la ganancia de información para seleccionar el atributo:

```
function TDIDT(ejemplos, atributos, defecto)
returns Tree
 if ( atributos=∅ )
     return Hoja(clase-mayoritaria(ejemplos))
 if ( ejemplos=∅ )
     return Hoja(defecto)
 if (∀e ∈ ejemplos, e.clase = C)
     return Hoja(C)

 mejor ← escoger-atributo(ejemplos,atributos)
 nodo ← crear-nodo(mejor)
 resto-atr ← eliminar(atributos, mejor)
 foreach v ∈ valores(mejor)
   ejemplos-v ← extraer-ejemplos(ejemplos,mejor,v)
   rama ← TDIDT(ejemplos-v, resto-atr, defecto)
   añadir-rama(nodo,rama)
 return nodo
```

Figura 1. Esquema general de los algoritmos TDIDT

$$Ganancia(A,S) = Entropía(S) - Entropía(A,S) \quad (1)$$

$$Entropía(S) = -\sum_{c \in Clase} p_c \cdot \log_2 p_c \quad (2)$$

$$Entropía(A,S) = \sum_{a \in A} p_a \cdot Entropía(a,A,S_a) \quad (3)$$

$$Entropía(a,A,S_a) = -\sum_{c \in Clase} p_{ac} \cdot \log_2 p_{ac} \quad (4)$$

$$p_c = |S_c| / |S| \quad (5)$$

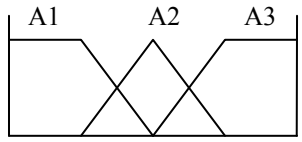
$$p_a = |S_a| / |S| \quad (6)$$

$$p_{ac} = |S_{ac}| / |S_a| \quad (7)$$

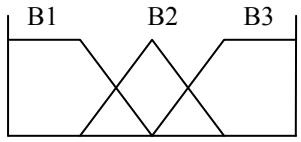
Donde S es el conjunto de ejemplos a estudiar, A es el atributo cuya ganancia de información se va a calcular, a se refiere a cada uno de los valores del atributo A, *Clase* es el atributo de clasificación, c se refiere a cada uno de los valores del atributo *Clase*, S_c es el subconjunto de S formado por los ejemplos de la clase c, S_a es el subconjunto de S formado por los ejemplos en los que el atributo A toma el valor a, S_{ac} es el subconjunto de S formado por los ejemplos en los que el atributo A toma el

Atributo A	Atributo B	Clase
35	29	c1
55	41	c2
69	60	c3

(a)



(b)



(c)

Activación	Atributo A			Atributo B			Clase		
μ	μA1	μA2	μA3	μB1	μB2	μB3	μc1	μc2	μc3
1.0	0.60	0.40	0	0.84	0.16	0	1.0	0	0
1.0	0	0.80	0.20	0.36	0.64	0	0	1.0	0
1.0	0	0.24	0.76	0	0.60	0.40	0	0	1.0

(d)

Activación	Atributo B			Clase		
μ	μB1	μB2	μB3	μc1	μc2	μc3
0.6	0.84	0.16	0	1.0	0	0
0.0	0.36	0.64	0	0	1.0	0
0.0	0	0.60	0.40	0	0	1.0

(e)

Activación	Atributo B			Clase		
μ	μB1	μB2	μB3	μc1	μc2	μc3
0.40	0.84	0.16	0	1.0	0	0
0.80	0.36	0.64	0	0	1.0	0
0.24	0	0.60	0.40	0	0	1.0

(f)

Activación	Atributo B			Clase		
μ	μB1	μB2	μB3	μc1	μc2	μc3
0.0	0.84	0.16	0	1.0	0	0
0.20	0.36	0.64	0	0	1.0	0
0.76	0	0.60	0.40	0	0	1.0

(g)

Figura 2: (a) Conjunto de datos original; (b,c) Funciones de pertenencia de los atributos A y B; (d) Conjunto de datos difusos; (e,f,g) Subconjuntos de datos difusos para los valores A1, A2 y A3 del atributo A.

valor *a* y el atributo *Clase* toma el valor *c*, y $|\cdot|$ indica el numero de elementos de un conjunto.

Una de las limitaciones más serias del algoritmo ID3 es que sólo está definido para atributos nominales. Cuando los atributos están definidos sobre dominios continuos es necesario discretizarlos definiendo intervalos en dichos dominios. La evolución del algoritmo ID3, denominada C4.5, discretiza cada atributo continuo A en los intervalos $(A \leq t)$ y $(A > t)$ siendo t el umbral que conduce al mejor valor de cierto criterio de división [13].

El algoritmo FuzzyID3 es una extensión del algoritmo ID3 orientado a problemas de clasificación que utilicen atributos continuos. El algoritmo se basa en describir los valores nominales de los atributos continuos por medio de conjuntos difusos en lugar de utilizar intervalos. Los conjuntos de datos a clasificar se convierten entonces en conjuntos de datos difusos (Figura 2) que contienen para cada instancia el grado de pertenencia a cada etiqueta de cada parámetro.

El resultado de FuzzyID3 es un árbol de decisión difuso. El funcionamiento de un árbol de decisión difuso es el siguiente. Cada nodo interno representa una consulta respecto al valor de un atributo. Las ramas se refieren a las distintas etiquetas lingüísticas del atributo, de manera que cada rama tiene un cierto grado de activación. Ante unos determinados valores de entrada, los grados de activación se van propagando por las ramas del árbol hasta determinar el grado de activación de cada hoja. Mientras que en un árbol de decisión no difuso sólo existe una hoja activa en cada caso, en un árbol de decisión difuso aparecen muchas hojas activas. Para obtener una salida concreta del clasificador difuso se toma la clase de la hoja más activa o se calcula el grado de activación agregado de cada clase y se toma la clase más activa.

Para transformar el algoritmo ID3 en un algoritmo difuso es necesario definir la ganancia de información difusa y la forma en la que se divide el conjunto de datos entre los valores de un atributo. Para definir la ganancia de información sobre un conjunto de datos difuso es necesario rescribir la forma en la que se calculan las diferentes proporciones con datos difusos:

$$p_c = \sum_e \mu \otimes \mu_c \Big/ \sum_{\theta \in Clase} \sum_e \mu \otimes \mu_\theta \quad (8)$$

$$p_a = \sum_e \mu \otimes \mu_a \Big/ \sum_{\alpha \in A} \sum_e \mu \otimes \mu_\alpha \quad (9)$$

$$p_{ac} = \sum_e \mu \otimes \mu_a \otimes \mu_c \Big/ \sum_{\theta \in Clase} \sum_e \mu \otimes \mu_a \otimes \mu_\theta \quad (10)$$

Donde e se refiere a cada uno de los ejemplos del conjunto de datos, μ es el grado de activación de cada ejemplo, μ_c es el grado de pertenencia del ejemplo a la case c, μ_θ es el grado de pertenencia del ejemplo a la case θ, μ_a es el grado de pertenencia del ejemplo a la etiqueta a del atributo A, μ_α es el grado de pertenencia del ejemplo a la etiqueta α del atributo A, y el operador $\otimes$ se refiere a la T-norma escogida para la conjunción.

Al dividir un conjunto de datos entre los valores de un atributo, cada elemento e del conjunto de datos es incluido en el subconjunto del valor a con el siguiente grado de activación:

$$\mu_{ea} = \mu \otimes \mu_a \quad (11)$$

Esta forma de dividir los conjuntos de datos tiende a generar subconjuntos con numerosas instancias con grado de activación pequeño. Para evitar esto se introduce un cierto umbral de activación por debajo del cual las instancias son eliminadas de los conjuntos de datos.

3 EL ALGORITMO FUZZCONRI

El algoritmo CN2 fue propuesto como forma de mejorar los problemas de los algoritmos AQ frente al ruido. El algoritmo realiza una búsqueda en estrella en el espacio de reglas conjuntivas considerando como potenciales selectores a todos los posibles valores de cada atributo. CN2 utiliza dos heurísticas para seleccionar las reglas. La primera es un test de significancia estadística que elimina aquellas reglas que no cubren un número suficiente de ejemplos. La segunda heurística es un test de bondad que selecciona a las reglas con mayor proporción de acierto. Inicialmente el algoritmo generaba listas de reglas ordenadas y utilizaba como heurísticas el factor de probabilidad (*likelihood ratio*) y la entropía de información. Posteriormente se propuso modificaciones al algoritmo para generar listas de reglas desordenadas y la entropía de información se sustituyó por el error laplaciano, con una considerable mejora del comportamiento del algoritmo. La Figura 3 muestra el esquema general de la versión mejorada de CN2.

El algoritmo tiene dos parámetros de control: el umbral de significancia y el tamaño máximo de la estrella, es decir, el número máximo de candidatos que se almacenan en la búsqueda.

```
function CN2(ejemplos,clases)
returns lista_de_reglas
lista_de_reglas ← ∅
foreach c ∈ clases
   P ← Positivos(ejemplos,c)
   N ← Negativos(ejemplos,c)
   reglas ← CN2ParaUnaClase(P,N,c)
   lista_de_reglas ← lista_de_reglas ∪ reglas
return lista_de_reglas

function CN2ParaUnaClase(P,N,c)
returns lista_de_reglas
lista_de_reglas ← ∅
regla ← EncontrarMejorRegla(P,N,c)
while regla ≠ null
   lista_de_reglas ← lista_de_reglas ∪ regla
   P ← EliminarCubiertos(P,regla)
   regla ← EncontrarMejorRegla(P,N,c)
return lista_de_reglas

function EncontrarMejorRegla(P,N,c)
returns regla
mgr ← "if(true) then c"
star ← { mgr }
mejor ← null
while star ≠ ∅
   newstar ← ∅
   foreach regla ∈ star, sel ∈ selectores
      newr ← AñadirSelector(regla,sel)
      if( Significancia(P,N,newr) > umbral
      & Bondad(P,N,newr) > Bondad(P,N,mejor))
      then
         mejor ← newr
         newstar ← newstar ∪ newr
         if(Tamaño(newstar) > maximo)
         then newstar ← EliminarPeor(newstar)
   star ← newstar
return mejor
```

Figura 3. Esquema general del algoritmo CN2.

El algoritmo FuzzConRI es una adaptación de CN2 para tratar conjuntos de datos difusos. El algoritmo introduce un corte-α para discriminar entre los ejemplos positivos y negativos de cada clase, de manera que para cada clase c todos los ejemplos cuyo μ_c sea menor que el corte-α se consideran negativos. Este corte-α se utiliza también para eliminar los elementos cubiertos por una regla, que serían aquellos para los que dicha regla alcanza un grado de activación superior a dicho corte-α. FuzzConRI requiere también adaptar las heurísticas de significancia y bondad de las reglas generadas. Los autores consideran la significancia estadística como una heurística opcional y en sus trabajos utilizan para esta función el porcentaje de

ejemplos positivos cubiertos por la regla. Por su parte, para evaluar la bondad de las reglas proponen utilizar la función de precisión difusa (*fuzzy accuracy*) que viene dada por la siguiente expresión.

$$A(r) = \sum_{p \in P} \mu_r(p) - \sum_{n \in N} \mu_r(n) \qquad (12)$$

Donde $\mu_r(e)$ representa el grado de activación de la regla *r* para la instancia *e*.

4 LA IMPLEMENTACIÓN EN XFUZZY

Los algoritmos FuzzyID3 y FuzzConRI se han añadido al entorno Xfuzzy como parte de los algoritmos integrados en la herramienta Xfdm. La Figura 4 muestra la ventana principal de esta herramienta. Desde esta ventana es posible seleccionar el algoritmo a utilizar en el proceso de inducción de reglas, el conjunto de datos a utilizar, la forma de las variables de entrada de la base de reglas a extraer y el estilo del sistema a generar.

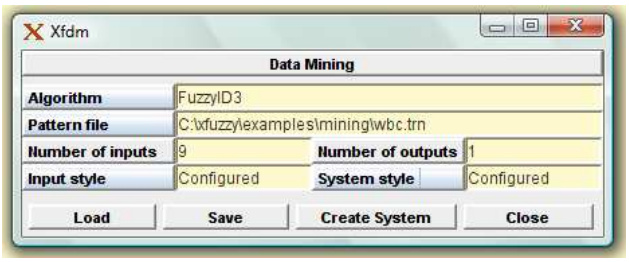


Figura 4: Ventana principal de Xfdm.

Cada algoritmo de extracción de reglas difusas dispone de una ventana de configuración en la que se introducen los valores de los parámetros de control del algoritmo. La Figura 5 muestra la ventana de selección del algoritmo FuzzyID3 que contiene como único parámetro de configuración el umbral de activación mínimo de las instancias a considerar por el algoritmo.

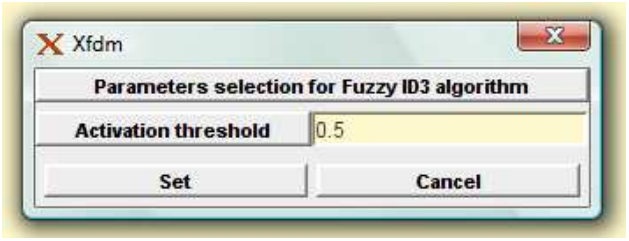


Figura 5: Configuración del algoritmo FuzzyID3.

Por su parte, el algoritmo FuzzConRI tiene como parámetros de control el tamaño máximo de la estrella (conjunto de reglas que se mantienen en memoria), el corte-α (que se utiliza para construir N y P y para eliminar las instancias cubiertas por una regla) y el umbral de significancia. La Figura 6 muestra la ventana de configuración de este algoritmo.

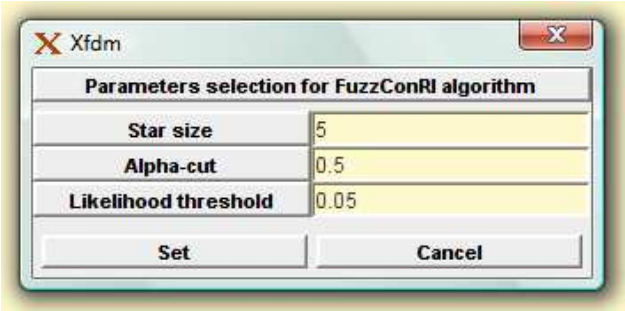


Figura 6: Configuración del algoritmo FuzzConRI.

La aplicación de estos algoritmos requiere definir las funciones de pertenencia de las variables de entrada de las reglas. La Figura 7 muestra la ventana de configuración de las variables de entrada. En esta ventana se puede seleccionar el número de funciones de pertenencia de las entradas, la forma de estas funciones o el universo de discurso de las variables.

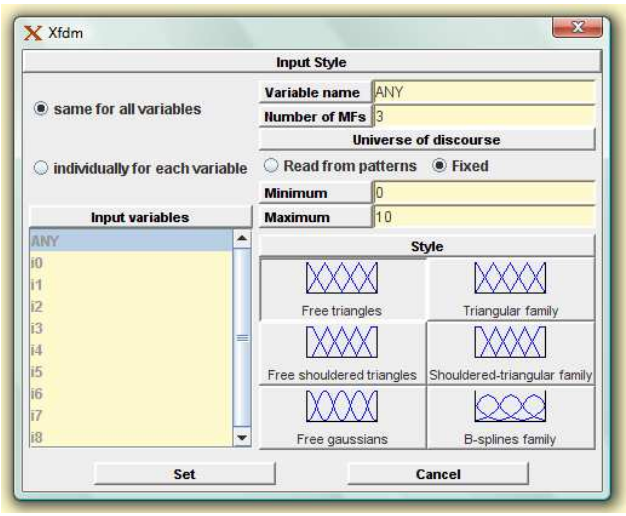


Figura 7: Configuración de las variables de entrada.

Para finalizar el proceso de configuración de los algoritmos es necesario definir el tipo de base de reglas que se va a generar (de clasificación en nuestro caso) y la T-norma a utilizar como conjunción. La Figura 8 muestra la ventana de Xfdm dedicada a seleccionar estos aspectos.

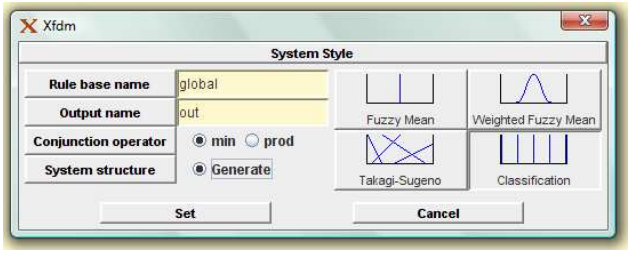


Figura 8: Configuración del estilo del sistema.

5 ALGUNOS EJEMPLOS

A continuación se presentan los resultados de un estudio preliminar sobre el comportamiento de estos dos algoritmos. Para desarrollar este estudio se han utilizado varias bases de datos de clasificación con atributos de entrada numéricos. Todas estas bases de datos han sido extraídas del repositorio UCI. En el estudio se han considerado los resultados de los algoritmos sobre las bases de datos completas. Para obtener conclusiones más firmes sería necesario realizar un estudio más profundo basado en evaluación por validación cruzada.

En primer lugar se ha realizado un estudio sobre el efecto de los diferentes parámetros de control de los algoritmos FuzzyID3 y FuzzConRI. Para ello se han utilizado dos bases de datos conocidas: Pima, que contiene 768 patrones de clasificación entre pacientes de diabetes de esta tribu indígena en base a 8 atributos numéricos, y Winsconsin Breast Cancer, que contiene 683 patrones de clasificación entre tumores malignos y benignos en base a 9 atributos numéricos. En este estudio se ha representado cada atributo de entrada por medio de 3 etiquetas como las mostradas en Fig. 2.b y se ha escogido el mínimo como T-norma.

La Tabla 1 muestra los resultados obtenidos por el algoritmo FuzzyID3 sobre estas bases de datos para distintos valores del umbral de activación. Este parámetro se utiliza para eliminar de los subconjuntos de datos aquellos ejemplos cuyo grado de activación no supere este umbral. Un valor pequeño del umbral provoca que la mayoría de los ejemplos no sean eliminados y que un mismo ejemplo se utilice en los subconjuntos de datos de diferentes ramas del árbol de decisión. Por el contrario, un valor alto del umbral provoca que muchos ejemplos sean eliminados de los subconjuntos y que algunos de ellos sean eliminados de todos los subconjuntos de datos de las diferentes ramas del árbol. En la tabla se observa que los mejores resultados se obtienen en el entorno de un umbral 0.5. Con respecto al número de reglas se observa que al aumentar el umbral de activación el número de reglas generadas disminuye. Esto se debe a que un umbral mayor conduce a subconjuntos de datos más pequeños que generan árboles de decisión menos profundos.

Tabla 1. Resultados de FuzzyID3

	Pima		WBC	
Umbral	# Reglas	%Error	# Reglas	%Error
0.0	1649	25.78	961	2.20
0.2	901	24.09	239	1.02
0.4	423	21.61	105	0.34
0.5	299	19.27	105	0.34
0.6	193	26.95	41	2.63
0.8	55	36.20	17	5.42

La Tabla 2 muestra los resultados que se obtienen con el algoritmo FuzzConRI para distintos valores de sus parámetros de configuración. Con respecto al tamaño de la estrella la tabla muestra que su influencia sobre el resultado del algoritmo es pequeña. Un tamaño de la estrella inferior a 5 puede provocar un aumento de la tasa de error, mientras que valores más altos en el tamaño de la estrella no consiguen mejorar los resultados. Por su parte, los valores superiores a 0.5 en el corte-α producen un aumento considerable en el número de reglas generadas y en la tasa de error. Al considerar valores altos de este parámetro la mayoría de los ejemplos se consideran no cubiertos por las reglas generadas lo que obliga a seguir buscando reglas. Un valor en el corte-α inferior a 0.5 provoca que se eliminen muchos ejemplos del conjunto P y que el número de reglas generadas sea menor, aunque también produce una tasa de error más alta al perder ejemplos que aportan información al conjunto de datos. Por último, los valores altos del umbral de significancia (10%) provocan una disminución del número de reglas a costa de un leve incremento en la tasa de error. La razón es que a las reglas generadas por el algoritmo se les exige un grado de significancia superior al umbral de manera que si este umbral es alto son pocas las reglas que consiguen superarlo.

Tabla 2. Resultados de FuzzConRI

			Pima		WBC	
Star	Corte-α	Sign.	# R.	%Err	# R.	%Err
5	0.5	0.10	34	27.99	28	7.47
5	0.5	0.05	42	27.86	34	7.47
5	0.5	0.01	59	27.86	34	7.47
5	0.3	0.05	46	28.51	24	7.47
5	0.5	0.05	42	27.86	34	7.47
5	0.7	0.05	72	31.38	40	50.95
3	0.5	0.05	42	28.25	34	7.47
5	0.5	0.05	42	27.86	34	7.47
10	0.5	0.05	43	27.86	33	7.47

La Tabla 3 contiene los resultados obtenidos por los algoritmos FuzzyID3, FuzzConRI y Wang&Mendel sobre 7 bases de datos de atributos numéricos extraídas del repositorio UCI. En todos los casos se ha considerado que cada atributo se describe por medio de 3 etiquetas como las mostradas en la Figura 2.b y se ha escogido el mínimo como T-norma. Para la aplicación del algoritmo FuzzyID3 se ha utilizado un umbral de activación de 0.5. Las pruebas realizadas con el algoritmo FuzzConRI utilizan una estrella de tamaño 5, un valor de 0.5 en el corte-α y un umbral de significancia de 0.05.

Comparado los resultados del algoritmo FuzzyID3 con los obtenidos por el algoritmo de Wang y Mendel se observa

Tabla 3. Resultados sobre varias bases de datos

			FuzzyID3		FuzzConRI		Wang & Mendel	
Base de datos	Atributos	Instancias	#Reglas	%Error	#Reglas	%Error	#Reglas	%Error
Glass	9	214	173	23.8%	51	28.0%	68	29.9%
Heart-statlog	13	270	199	4.1%	46	12.6%	234	4.1%
Pima	8	768	299	19.3%	42	27.9%	173	23.81
Iris	4	150	21	4.7%	8	2.7%	22	4.7%
WBC	9	683	105	0.34%	34	7.5%	254	1.0%
Sonar	60	208	91	0.0%	37	14.9%	207	0.0%
Vehicle	18	946	1844	15.0%	174	29.1%	496	28.4%

una mejora apreciable de la precisión (una tasa de error menor). Con respecto al número de reglas, en muchos casos se observa que el número de reglas generado por FuzzyID3 es mayor que el que obtiene el algoritmo de Wang y Mendel. Esto se debe a que FuzzyID3 genera reglas que cubren todo el espacio, mientras que el algoritmo de Wang y Mendel genera reglas que solo cubren los ejemplos del conjunto de datos. En estos casos, un proceso de podado de reglas que no se activen significativamente para ningún ejemplo suele reducir apreciablemente el número de reglas del sistema generado por FuzzyID3 (por ejemplo, para la base de datos Glass el número de reglas se reduce a 52, mientras que para la base de datos Vehicle se reduce a 323).

Con respecto al algoritmo FuzzConRI, como norma general se observa que el número de reglas generadas es bastante menor que el de otros algoritmos, a costa de una tasa de error más alta. Esto se debe a la elección de la precisión difusa como medida de bondad, ya que esta medida premia a las reglas con mayor número de ejemplos positivos cubiertos, incluso aunque el número de errores de clasificación sea alto.

6 CONCLUSIONES

Los algoritmos de inducción de reglas difusas se dividen en dos grupos: los que generan etiquetas difusas asociadas a cada regla (*clustering*) y los que utilizan etiquetas definidas previamente (*grid*). Estos últimos son mejores en cuanto a interpretabilidad de las etiquetas, pero suelen generar un número de reglas demasiado alto. La adaptación de algunos algoritmos de clasificación clásicos al mundo difuso nos permite obtener clasificadores difusos con reglas más compactas lo que mejora significativamente la interpretabilidad. Dos de estos algoritmos (FuzzyID3 y FuzzConRI) han sido incorporados al entorno de desarrollo Xfuzzy, lo que permite integrar estos algoritmos al ciclo de diseño desarrollado por este entorno.

Agradecimientos

Este trabajo ha sido parcialmente financiado gracias al proyecto de investigación TEC2005-04359/MIC del Ministerio de Educación y Ciencias y a los proyectos de investigación TIC2006-635, TEP2006-375 y P07-TIC-03179 de la Junta de Andalucía.

Referencias

[1] I. Baturone, F.J. Moreno Velo, A. Gersnoviez. Identifying fuzzy systems from numerical data with Xfuzzy. *Proc. 4th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2005)*, Pág 1257-1262, 2005.

[2] I. Baturone, F.J. Moreno-Velo, S. Sanchez-Solano, A. Barriga, P. Brox, A. Gersnoviez, M. Brox. Using Xfuzzy Environment for the Whole Design of Fuzzy Systems. *Proc. IEEE International Conference on Fuzzy Systems*, Pág 1-6, 2007.

[3] J.C. Bezdek, Pattern recognition with fuzzy objective function algorithms, *Plenum Press*, 1981.

[4] S. L. Chiu. A Cluster Estimation Method with Extension to Fuzzy Model Identification. *Proc.3rd IEEE Conference on Fuzzy Systems,* Pág. 1240-1245, 1994.

[5] K.J. Cios, L.M. Sztandera. Continuous ID3 algorithm with fuzzy entropy measures. *Proc. IEEE Int. Conf. on Fuzzy Systems*, Pág. 469-476, 1992.

[6] P. Clark, T. Niblett. The CN2 Induction Algorithm. *Machine Learning*, vol. 3, Pág. 261-283, 1989.

[7] P. Clark, R. Boswell. Rule induction with CN2: some recent improvements. *Proc. Fifth European Working Session on Learning (EWSL-91)*, Pág. 151-163, 1991.

[8] I. Gath, A. B. Geva. Unsupervised Optimal Fuzzy Clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, Pág. 773- 781, 1989.

[9] D. E. Gustafson, W. C. Kessel. Fuzzy Clustering with a Covariance Matrix. *Proc. IEEE Conference on Decision and Control,* Pág. 761-766, 1978.

[10] R.S. Michalski, I. Mozetic, J. Hong, N. Lavrac. The multi-purpose incremental learning system AQ15 and its testing application to three medical domains. *Proc. 5th National Conference on Artificial Intelligence (AAAI86)*, Pág. 1041-1047, 1986.

[11] D. Nauck, R. Kruse. NEFCLASS – A Neuro-Fuzzy Approach for the Classificatioin of Data. *Proc. 1995 ACM Symposium on Applied Computing*, Pág 461-465, 1995.

[12] J.R. Quinlan. Induction of decision trees. *Machine Learning*, vol. 1, Pág. 81-106, 1986.

[13] J. R. Quinlan. Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research*, vol. 4, Pág. 77-90, 1996.

[14] R. Senhadji, S. Sánchez-solano, A. Barriga, I. Baturone, F.J. Moreno-Velo. NORFREA: An algorithm for non-redundant fuzzy rule extraction. *Proc. IEEE International Conference on Systems, Man and Cybernetics*, Pág. 604-608, 2002.

[15] T. Tani, M. Sakoda, K. Tanaka, Fuzzy Modeling by ID3 algorithm and its application to prediction, *Proc IEEE International Conference on Fuzzy Systems*, Pág. 923-930, 1992.

[16] M. Umano, H. Okamoto, I. Hatono, H. Tamura, F. Kawachi, S. Umedzu, J. Kinoshita. Fuzzy Decisión Trees by Fuzzy ID3 Algorithm and Its Application to Diagnosis Systems. *Proc IEEE International Conference on Fuzzy Systems*, Pág. 2113-2118, 1994.

[17] J. van Zyl, I. Cloete. An inductive algorithm for learning conjunctive fuzzy rules. *Proc. 3rd IEEE International Conference on Machine Learning and Cybernetics*, Pág. 4181-4187, 2004.

[18] L.Wang, J.M. Mendel. Generation of Rules by Learning from Examples. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 22, Pág. 1414-1427, 1992.

[19] R. Weber. Fuzzy ID3: a class of methods for automatic knowledge acquisition. *Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks*, Pág. 265-268, 1992.

XFUZZYLIB: UNA LIBRERÍA DE MÓDULOS PARA LA SÍNTESIS HARDWARE DE SISTEMAS DE INFERENCIA DIFUSOS

S. Sánchez-Solano[1] **M. Brox**[2] **I. Baturone**[1] **A. Barriga**[1]

[1] Instituto de Microelectrónica de Sevilla. Centro Nacional de Microelectrónica. CSIC.
{santiago, lumi, barriga}@imse.cnm.es
[2] Dpto. de Arquitectura de Computadores, Electrónica y Tecnología Electrónica. Escuela Politécnica Superior.
Universidad de Córdoba. mbrox@uco.es

Resumen

Esta comunicación presenta una nueva técnica de implementación de sistemas difusos que está basada en el uso de una librería de módulos específicos, denominada *XfuzzyLib*, y cuyo flujo de diseño combina las herramientas de modelado y simulación del entorno Matlab con las de síntesis e implementación de FPGAs de Xilinx. La estrategia propuesta, que constituye la base de una nueva herramienta de síntesis hardware del entorno *Xfuzzy*, aprovecha las ventajas de flexibilidad y facilidad de configuración que brindan las diferentes herramientas de Matlab y Xilinx, permitiendo acelerar considerablemente las etapas de descripción, síntesis y verificación funcional de los sistemas bajo desarrollo.

Palabras Clave: Lógica difusa, Realizaciones hardware, Herramientas de diseño para DSP.

1 INTRODUCCIÓN

La capacidad de los sistemas difusos para emular los esquemas de almacenamiento y procesado de información empleados por el cerebro humano, así como para facilitar la descripción de sistemas complejos mediante reglas expresadas en lenguaje natural, ha sido explotada con éxito durante las últimas décadas en el desarrollo de numerosas aplicaciones de control para automatización industrial, máquinas-herramienta, robótica y productos de consumo [1]-[3]. Sin embargo, a pesar del grado de madurez alcanzado por la tecnología fuzzy, existen actualmente muy pocas herramientas o entornos de diseño capaces de transformar la descripción de alto nivel de un sistema de inferencia difuso en una implementación hardware eficiente.

La literatura especializada recoge numerosas alternativas para la implementación hardware de controladores difusos mediante ASICs o FPGAs [4]-[6]. En particular, el constante desarrollo de las tecnologías de fabricación de circuitos integrados ha provocado importantes avances en los dispositivos lógicos programables, que han afectado, no solo al incremento del número de recursos de propósito general incluidos en los dispositivos, sino también a la introducción de elementos específicos, así como a la existencia de numerosos bloques de sistema disponibles como módulos de propiedad intelectual (módulos-IP) que facilitan el desarrollo de sistemas de procesado adaptados a un determinado dominio de aplicación. Estos avances se complementan, asimismo, con la propuesta de nuevas metodologías de diseño y la disponibilidad de potentes herramientas de CAD que permiten acortar el ciclo de desarrollo de nuevos productos [5].

En esta comunicación se describe una nueva estrategia de síntesis sobre FPGAs de sistemas de inferencia basados en lógica difusa. El flujo de desarrollo empleado hace uso de la herramienta System Generator (SysGen) de Xilinx dentro del entorno de modelado y simulación de Matlab. La Sección 2 presenta brevemente el ciclo de diseño con dicha herramienta. La librería de módulos que permite la síntesis de sistemas difusos mediante esta estrategia es detallada en la Sección 3, mientras que la Sección 4 ilustra un ejemplo de aplicación en el desarrollo de un controlador para navegación de un vehículo autónomo. Las principales conclusiones del trabajo se resumen en la Sección 5.

2 EL ENTORNO SYSGEN

SysGen facilita el desarrollo de algoritmos para sistemas de procesado digital de señal (DSP) sobre las FPGAs de Xilinx [8]. Está basado en Simulink, la herramienta interactiva para el modelado, la simulación y el análisis de

Este trabajo ha sido soportado por los proyectos TEC2005-04359/MIC del Ministerio de Educación y Ciencia y TIC2006-635 y TEP2006-375 de la Junta de Andalucía.

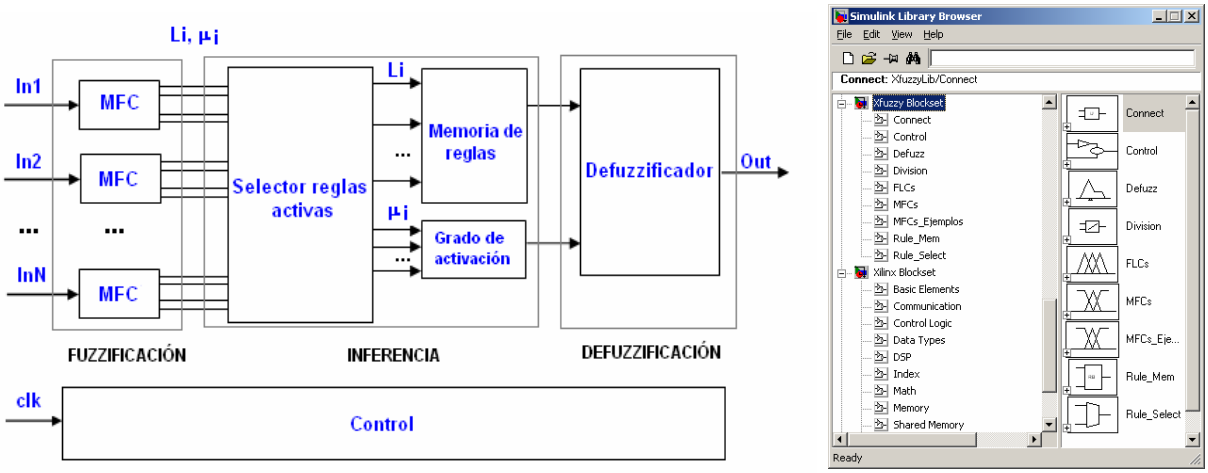


Figura 1. (a) Bloques básicos de una arquitectura de sistema difuso basada en reglas activas.
(b) Módulos de la librería *XfuzzyLib* integrados en el navegador de librerías de Simulink.

sistemas dinámicos integrada en Matlab. SysGen incluye una librería de módulos Simulink (denominada "*Xilinx Blockset*") que proporciona los elementos básicos de un sistema digital, así como el software necesario para facilitar la síntesis e implementación sobre las FPGAs de Xilinx de los algoritmos descritos y simulados en el entorno Matlab.

El ciclo de diseño típico con esta herramienta comienza con la descripción esquemática del sistema mediante un modelo Simulink. La verificación se lleva a cabo mediante simulación y utilizando las distintas facilidades de generación y visualización de señales proporcionadas por Matlab. Una vez verificada la funcionalidad del diseño, SysGen lo traslada de forma automática en una implementación hardware optimizada en términos de área y velocidad.

3. LA LIBRERÍA XFUZZYLIB

La Figura 1-a muestra los bloques que conforman la estructura general de una arquitectura eficiente para la implementación digital de sistemas difusos basada en el procesado de reglas activas [9]. Haciendo uso de la librería "*Xilinx Blockset*" y utilizando las facilidades proporcionadas por Simulink, se ha generado una librería de módulos especializados, a la que hemos denominado "*Xfuzzy Blockset*" o *XfuzzyLib*, que proporciona diferentes alternativas para implementar cada uno de los elementos básicos de esta arquitectura. La Figura 1-b muestra el navegador de librerías de Simulink con los diferentes módulos agrupados por bloques funcionales: generadores de funciones de pertenencia, selectores de reglas activas, conectivos de antecedentes, memorias de reglas, defuzzificadores y elementos de control.

A modo de ejemplo, la Figura 2 muestra el esquema de conexionado de un módulo generador de funciones de pertenencia mediante técnicas aritméticas (A_MFC). Este bloque es capaz de generar familias de funciones de pertenencia triangulares (salvo la primera y la última función que pueden ser de tipo "Z" y "S", respectivamente), con grado de solapamiento 2 y normalizadas (en cualquier punto la suma de ambos grados de pertenencia es igual a la unidad).

Los distintos módulos de la librería "*Xilinx Blockset*" admiten una serie de parámetros que definen detalles específicos de su funcionalidad, así como el tamaño o tipo de aritmética utilizada. Del mismo modo, una vez completado el diagrama de bloques de un nuevo módulo

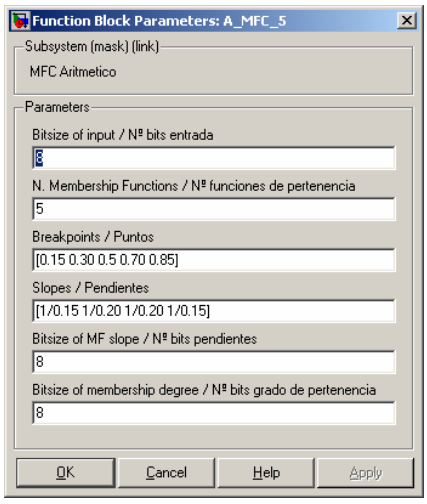


Figura 3. Parámetros que definen la funcionalidad del módulo A_MFC.

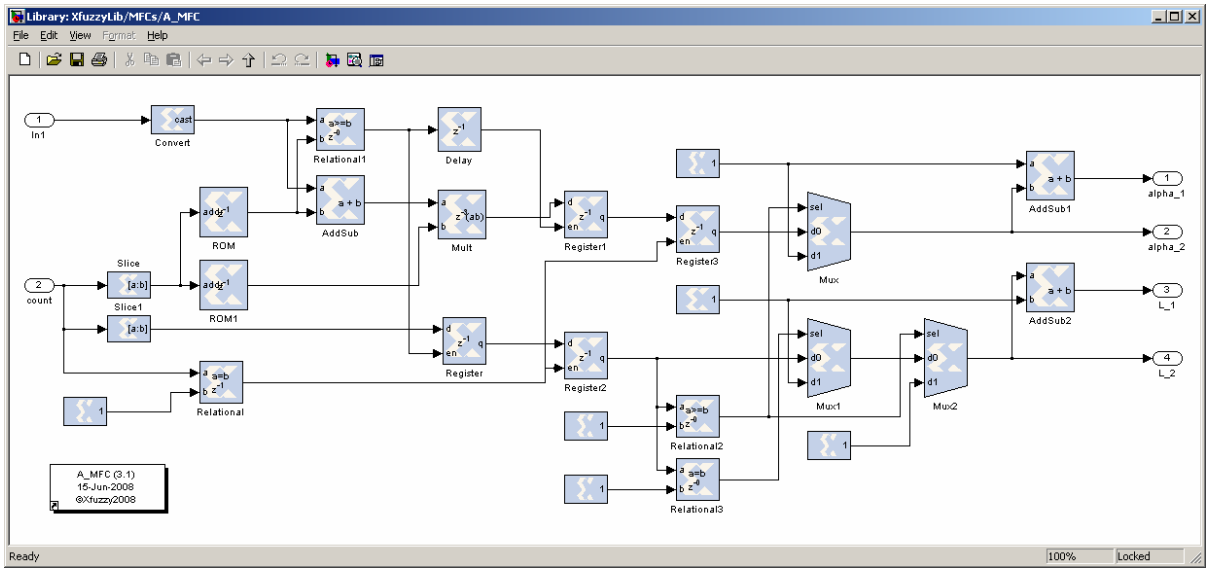


Figura 2. Modelo Simulink del módulo de generación de funciones de pertenencia (A_MFC).

de la librería *XfuzzyLib*, dicho módulo se encapsula como un "subsistema" y se añade una "máscara" para identificar los distintos parámetros que caracterizan al módulo (Figura 3). Cuando dicho subsistema es usado en un nivel jerárquico superior, estos parámetros pueden asignarse mediante valores numéricos o mediante variables a las que se les asigna un valor numérico en la ventana de comandos de Matlab o a través de un fichero ".m". Como muestra la Figura 4, La funcionalidad del diseño puede verificarse en todo momento con ayuda del simulador Simulink y sus diferentes facilidades para generar señales de excitación y para capturar y representar gráficamente los datos de salida.

Utilizando los diferentes módulos de la librería, la construcción de un sistema difuso requiere únicamente la elección, interconexión y parametrización de los bloques necesarios. Además de los bloques que constituyen los elementos constructivos básicos de un sistema difuso, la librería *XfuzzyLib* incluye bloques con descripciones de arquitecturas tipo de módulos de inferencia difusos que se diferencian entre sí en el número de entradas, el conectivo que usan para calcular el grado de activación de las reglas y el método de defuzzificación empleado. Al desarrollar un sistema difuso para una aplicación determinada, el usuario puede utilizar estas arquitecturas tipo o bien definir nuevas arquitecturas mediante la interconexión de los bloques de elementos básicos que sean necesarios. Asimismo, es posible combinar de forma jerárquica estos módulos de inferencia para definir sistemas difusos de mayor complejidad.

La Figura 5 muestra el conjunto de controladores difusos tipo (FLCs) actualmente incluidos en la librería *XfuzzyLib*. El diagrama de bloques de un FLC de dos entradas que emplea el producto como conectivo y el

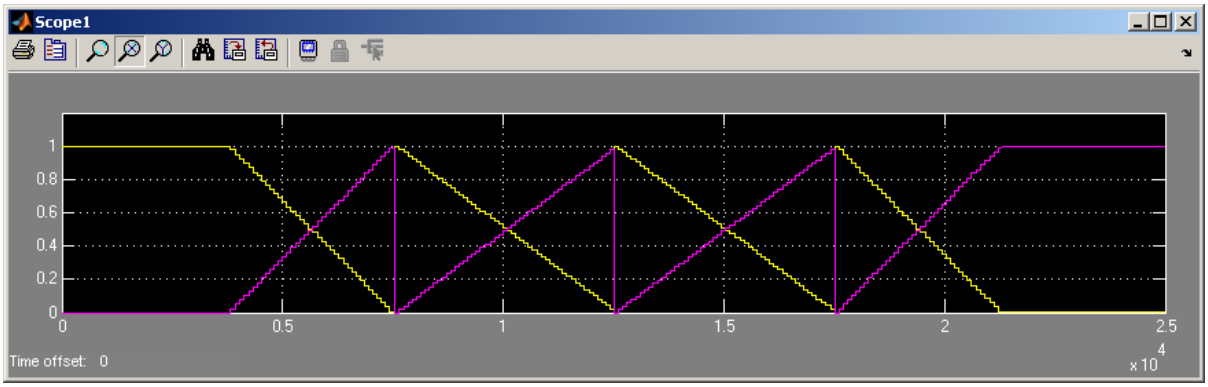


Figura. 4. Verificación de la funcionalidad del módulo A_MFC mediante Simulink.

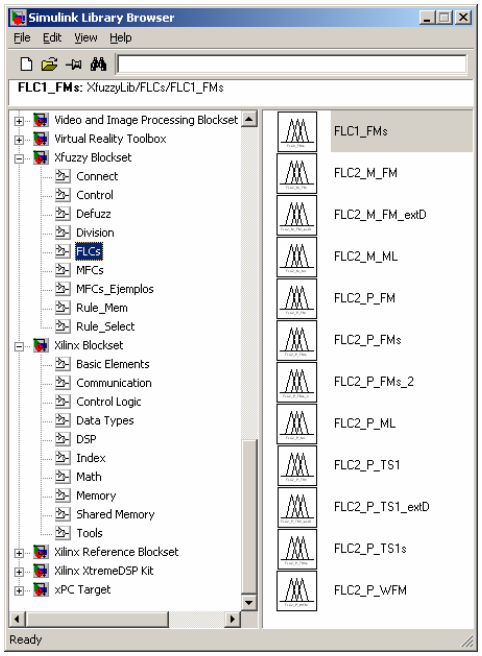


Figura 5. FLCs tipo incluidos en *XfuzzyLib*.

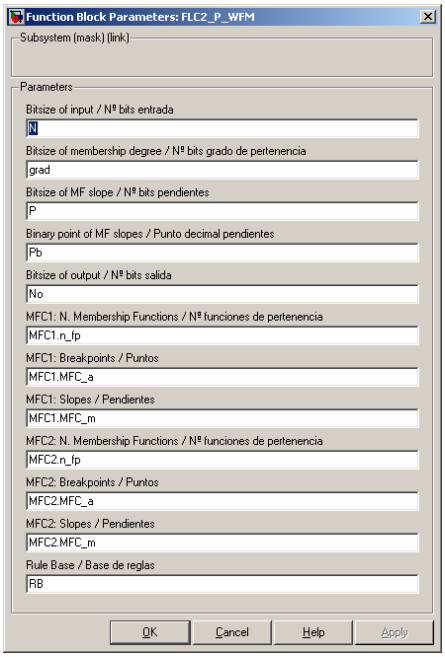


Figura 7. Parámetros del módulo FLC2_P_WFM.

método de defuzzificación WeightedFuzzyMean aparece en la Figura 6. Como ocurre con los bloques básicos de la librería, los bloques que describen arquitecturas de FLCs son totalmente parametrizables, por lo que el diseñador de sistemas difusos puede adaptar la funcionalidad del módulo a los requerimientos del problema planteado sin más que definir los parámetros correspondientes al dimensionado del sistema de inferencia y a la definición de la base de conocimiento del mismo. Análogamente a los bloques básicos, también aquí es posible aprovechar toda la funcionalidad del entorno Matlab para verificar la funcionalidad del sistema de inferencia. La Figura 7 muestra la máscara de definición de parámetros del FLC de la Figura 6. La Figura 8 ilustra cómo estos parámetros pueden definirse mediante un fichero ".m" para obtener una funcionalidad determinada.

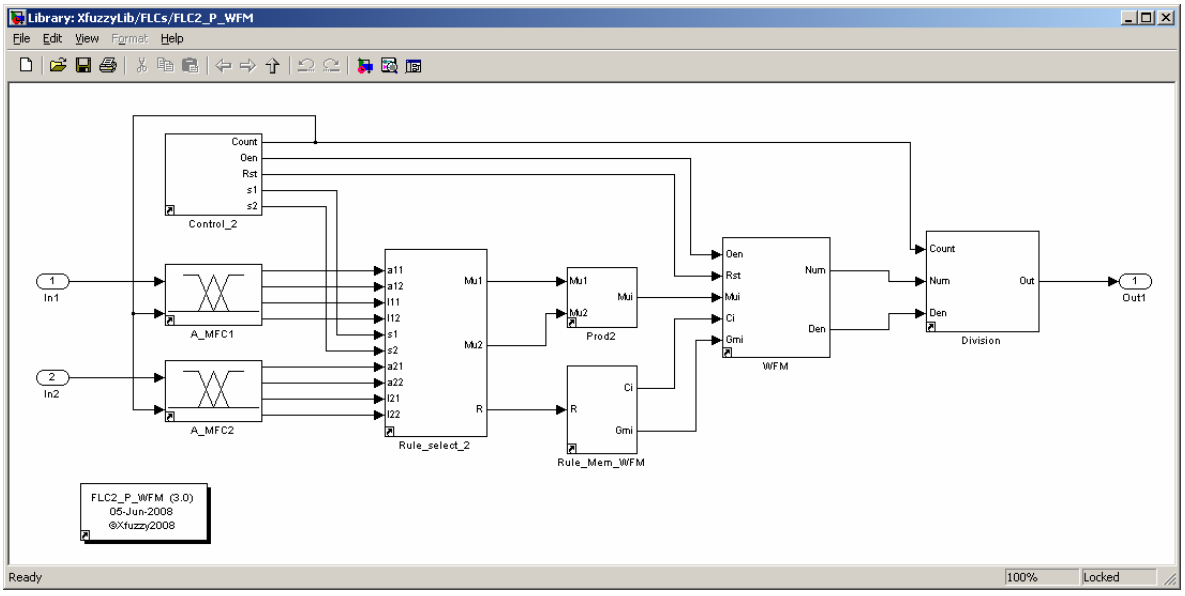
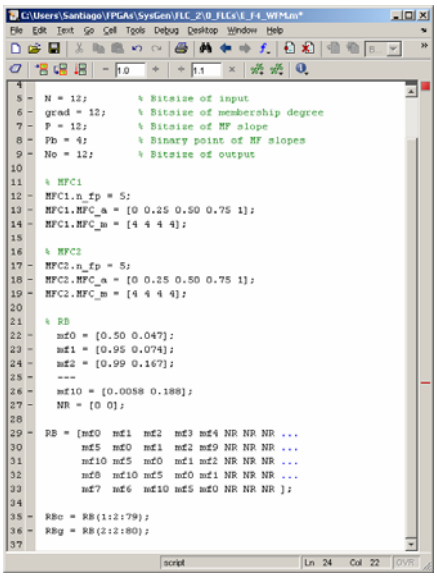
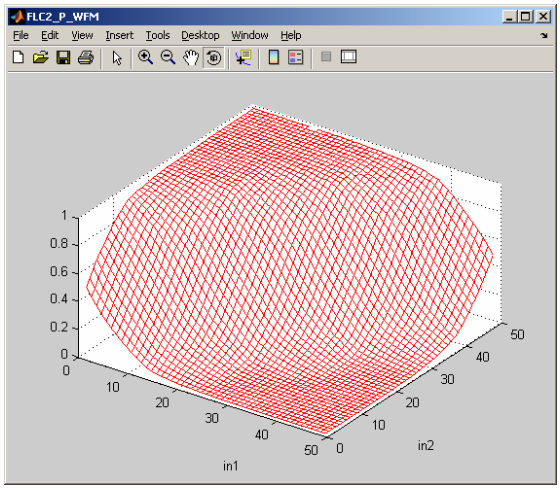


Figura 6. Diagrama de bloques de un FLC con "producto" como conectivo y WeightedFuzzyMean.



```
N = 12;          % Bitsize of input
grad = 12;       % Bitsize of membership degree
P = 12;          % Bitsize of MF slope
Pb = 4;          % Binary point of MF slopes
No = 12;         % Bitsize of output

% MFC1
MFC1.n_fp = 5;
MFC1.MFC_a = [0 0.25 0.50 0.75 1];
MFC1.MFC_m = [4 4 4 4];

% MFC2
MFC2.n_fp = 5;
MFC2.MFC_a = [0 0.25 0.50 0.75 1];
MFC2.MFC_m = [4 4 4 4];

% RB
  mf0 = [0.50 0.047];
  mf1 = [0.95 0.074];
  mf2 = [0.99 0.167];
  ---
  mf10 = [0.0058 0.188];
  NR = [0 0];

RB = [mf0  mf1  mf2  mf3 mf4 NR NR NR ...
      mf5  mf0  mf1  mf2 mf9 NR NR NR ...
      mf10 mf5  mf0  mf1 mf2 NR NR NR ...
      mf8  mf10 mf5  mf0 mf1 NR NR NR ...
      mf7  mf6  mf10 mf5 mf0 NR NR NR ];

RBc = RB(1:2:79);
RBg = RB(2:2:80);
```



(a) (b)

Figura 8. a) Definición de los parámetros del FLC a través de un fichero ".m". b) Comportamiento entrada/salida obtenido.

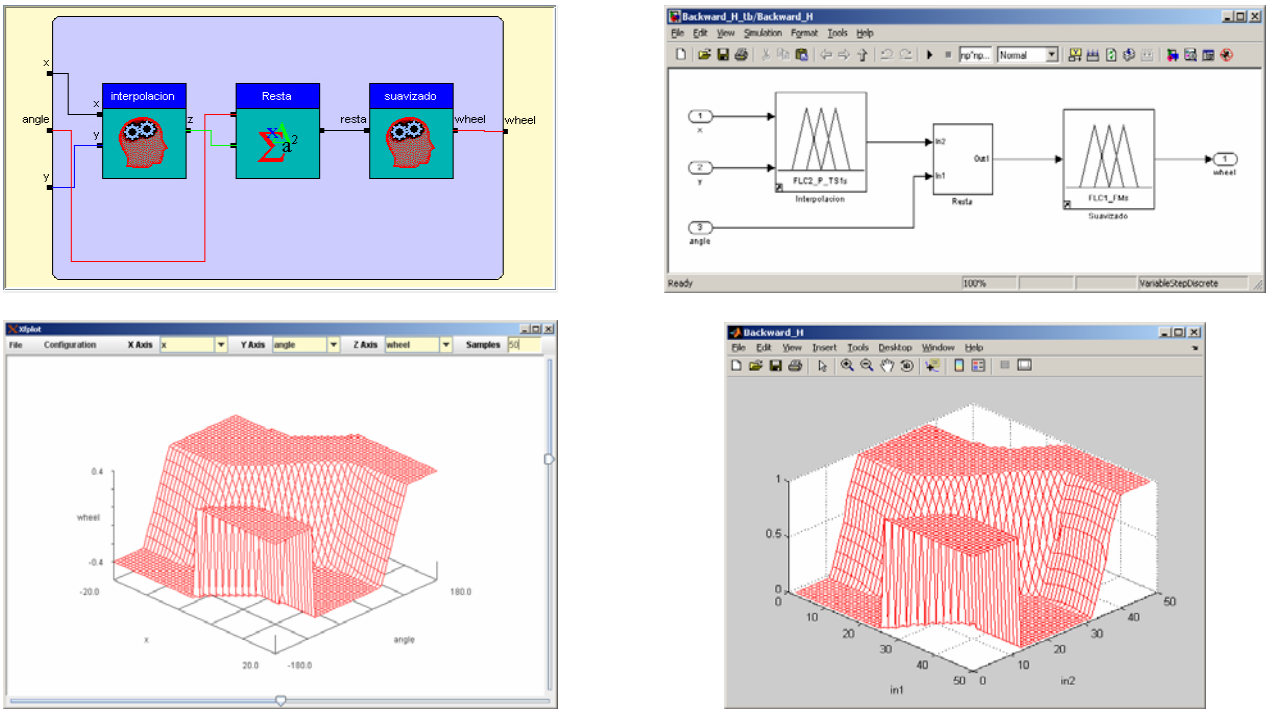
Una vez completado el modelo Simulink, la incorporación en el mismo del bloque "System Generator" facilita su síntesis hardware mediante la traslación del modelo a diferentes tipos de netlists, la generación del fichero de programación de la FPGA o, incluso, la verificación funcional del controlador en lazo cerrado a través de la cosimulación hardware/software de la implementación del sistema de control junto a un modelo Simulink de la planta.

4. UNA APLICACIÓN A LA ROBÓTICA

La técnica de implementación descrita ha sido aplicada a la realización de diferentes sistemas de control difuso para una aplicación típica de robótica: el aparcamiento de un vehículo autónomo en un espacio limitado [10]-[12]. El objeto bajo control es un vehículo eléctrico dotado de distintos sensores conectados a un procesador digital de señal (DSP) TMS-320LF de Texas Instruments que permite calcular los valores de la posición, orientación y velocidad del vehículo en cada momento [13]. La salida del controlador actúa sobre los motores de tracción y dirección del vehículo para fijar los valores adecuados de velocidad y ángulo de giro de las ruedas.

La Figura 9 muestra la estructura de uno de los sistemas de control analizados que usa una base de conocimiento jerárquica para calcular el ángulo de giro de las ruedas cuando el vehículo circula marcha atrás. El sistema utiliza dos módulo de inferencia difusos y un operador resta para definir la estrategia de control. Uno de los módulos difusos es del tipo Takagi-Sugeno de dos entradas, mientras que el otro emplea FuzzyMean como método de defuzzificación [11]. Para desarrollar esta estructura de controlador se ha hecho uso de las distintas utilidades del entorno *Xfuzzy* [14]. La traslación de la especificación de un sistema difuso en el entorno *Xfuzzy* al flujo de diseño propuesto en este trabajo requiere la identificación y la definición de los parámetros de los bloques proporcionados por la librería *XfuzzyLib*. La Figura 9-a muestra la descripción del controlador en el entorno *Xfuzzy* mientras que la Figura 9-b ilustra su representación mediante un modelo Simulink. Las Figuras ponen de manifiesto, asimismo, la similitud entre las superficies de control correspondientes a ambas descripciones.

La Figura 10 muestra el modelo Simulink de un sistema de control más complejo que emplea una base de conocimiento jerárquica que incluye módulos de toma de decisiones (para discriminar si el vehículo debe circular en un sentido u otro) y módulos de control (para calcular la velocidad y el ángulo de giro de las ruedas) [12]. Aprovechando la funcionalidad que proporciona SysGen, la verificación funcional del sistema de control puede llevarse a cabo mediante la integración en un lazo cerrado de cosimulación de un modelo software del robot descrito en Matlab y la implementación hardware del controlador sobre la FPGA. La Figura 11 muestra el modelo usado para verificar el sistema completo. Algunos de los resultados obtenidos aparecen en la Figura 12.



(a) (b)

Figura 9. Sistema de control difuso para el problema de aparcamiento de un vehículo autónomo: a) descripción del controlador en *Xfuzzy*. b) modelo Simulink con los módulos de la librería *XfuzzyLib*.

La implementación del controlador mostrado en la Figura 10, utilizando 8 bits en las entradas y salidas de las distintas bases de reglas y empleando una placa de desarrollo XUP-V2Pro (que incluye una FPGA Virtex-2 Pro XC2VP30 con 13.696 slices, 136 multiplicadores de 18 bits y 2.448Kb de memoria RAM) ocupa 1.094 slices y emplea 16 multiplicadores (lo que representa, respectivamente, el 7% y el 11% de los recursos disponibles en la FPGA). La frecuencia de operación alcanzada supera los 100 Mhz.

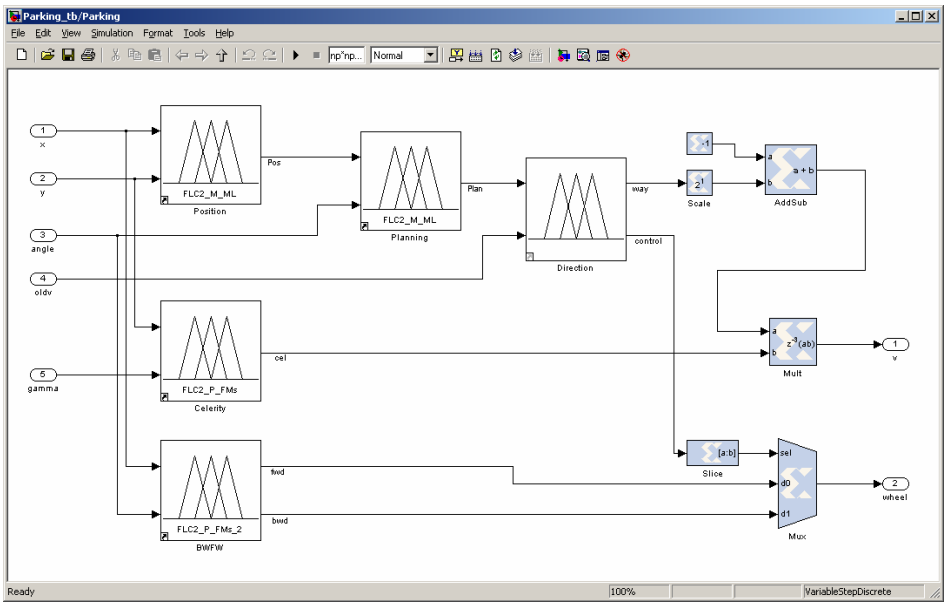


Figura 10. Modelo Simulink de un sistema de control difuso para el problema de aparcamiento.

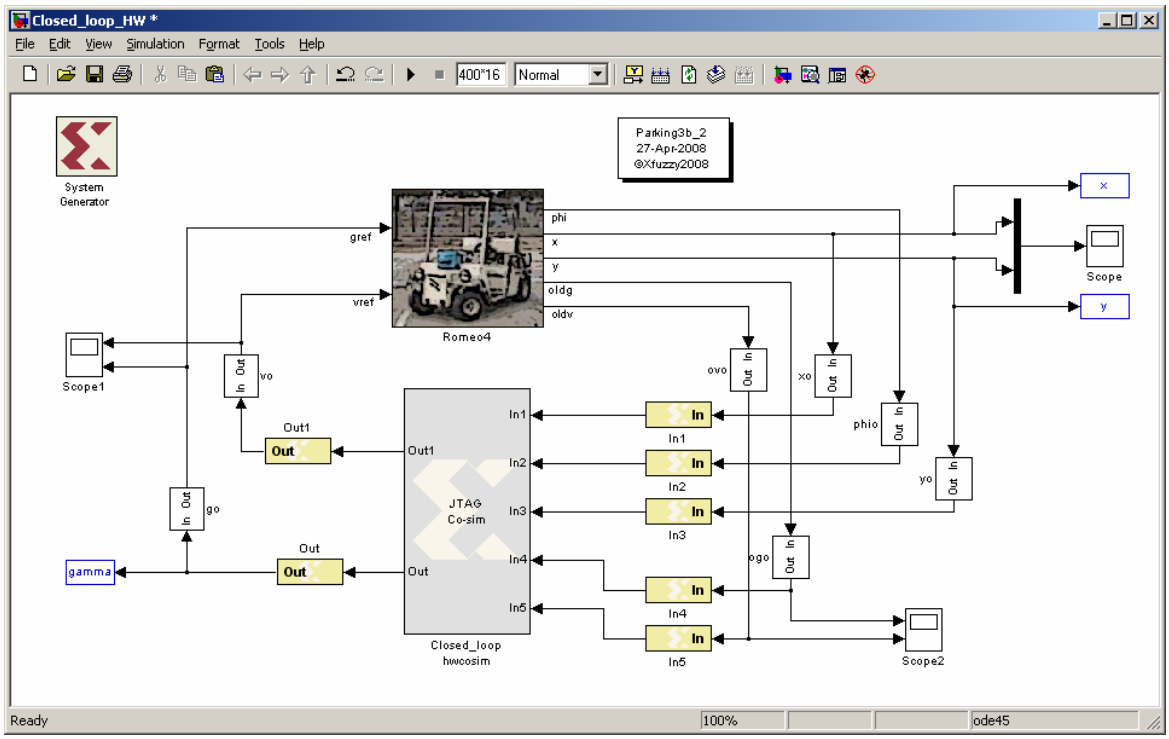


Figura 11. Verificación funcional del controlador en lazo cerrado mediante cosimulación hardware/software.

5. CONCLUSIONES

Se ha presenta una técnica de implementación de sistemas difusos sobre FPGAs basada en el uso de la herramienta de diseño de DSP de Xilinx. Al estar integrado en el entorno Matlab, el flujo de diseño propuesto presenta ventajas significativas en cuanto a flexibilidad y configurabilidad, y permite el uso de las distintas facilidades de cálculo y representación gráfica que proporciona dicho entorno. La estrategia propuesta es la base de una nueva herramienta de síntesis para la próxima versión del entorno *Xfuzzy*, aportando como principales novedades la inclusión de nuevos operadores y métodos de defuzzificación, una funcionalidad mejorada de muchos de los elementos básicos de la arquitectura y la implementación directa de sistemas difusos con estructuras jerárquicas.

Referencias

[1] J. Yen, R. Langari, L. Zadeh, (Eds.), *Industrial Applications of Fuzzy Logic and Intelligent Systems*, IEEE Press, 1995.

[2] M. J. Patyra, D. M. Mlynek, (Eds.) *Fuzzy Logic. Implementation and Applications*, Wiley & Teubner, 1996.

[3] T. J. Ross, *Fuzzy Logic with Engineering Applications*, 2nd Edition, Wiley, 2004.

[4] I. Baturone, A. Barriga, S. Sánchez-Solano, C. J. Jiménez, D. López, *Microelectronic Design of Fuzzy Logic-Based Systems*, CRC Press, 2000.

[5] D. Kim, "An Implementation of fuzzy logic controller on the reconfigurable FPGA system", *IEEE Trans. on Industrial Electronics*, Vol. 47, No. 3, pp. 703-715. Jun. 2000.

[6] M. McKenna, B. M. Wilamowski, "Implementing a fuzzy system on a field programmable gate array", *Proc. Int. Joint Conf. on Neural Networks*, Vol. 1, pp. 189-194. Jul. 2001.

[7] W. Savage, J. Chilton, R. Camposano. "IP reuse in the System on a Chip Era". *13th Int. Symp. on System Synthesis*, pp. 2-7. 2000.

[8] Xilinx. *Xilinx System Generator v8.2 for Simulink. User Guide*. Xilinx Blockset Reference Guide.

[9] A. Cabrera, S. Sánchez-Solano, C. J. Jiménez, A. Barriga, I. Baturone, "Arquitectura eficiente para la implementación hardware de sistemas de inferencia difusos", *Ingeniería Electrónica, Automática y Comunicaciones*, Vol. XXIII, No. 1, pp. 59-66, 2003.

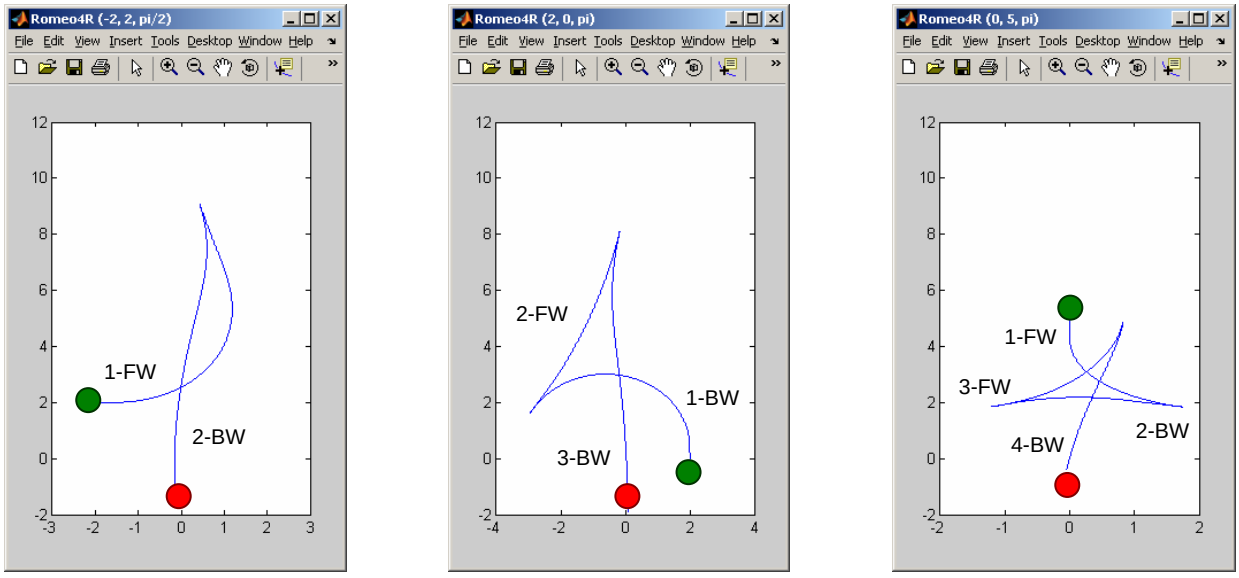


Figura 12. Trayectorias experimentales del vehículo al partir desde distintas posiciones iniciales.

[10] T.-H.S. Li, Shih-Jie Chang,Yi-Xiang Chen, "Implementation of human-like driving skills by autonomous fuzzy behavior control on an FPGA-based car-like mobile robot", *IEEE Trans. on Industrial Electronics*, Vol. 50, No. 5, pp. 867- 880. Oct. 2003.

[11] I. Baturone, F. J. Moreno-Velo, S. Sánchez-Solano, A. Ollero, "Automatic design of fuzzy controllers for car-like autonomous robots". *IEEE Trans. on Fuzzy Systems*, Vol. 12, No. 4, pp. 447-465, Aug. 2004.

[12] S. Sánchez-Solano, A. Cabrera, I. Baturone, F. J. Moreno-Velo, M. Brox, "FPGA Implementation of Embedded Fuzzy Controllers for Robotic Applications". *IEEE Trans. on Industrial Electronics*, Vol. 54, No. 4, pp. 1937-1945. Aug. 2007.

[13] J. Ferruz, V. Blanco, A. Ollero, and J. V. Acevedo. "An embedded DSP-based controller for the Romeo-4R vehicle". *Proc. IFAC Int. Symp. Intell. Compon. and Instrum. Control Appl.*, pp. 101–106, Jul. 2003

[14] Xfuzzy: Fuzzy Logic Design Tools, (http://www.imse.cnm.es/Xfuzzy)

FUZZY SETS OF QUASI-PERIODIC SIGNALS

Gonzalo Bailador[1] **Gracian Trivino**[2] **Albert van der Heide**[2]

[1] Universidad Politécnica de Madrid, gonzalo.bailador@upm.es
[2] Europen Centre for Soft Computing, {gracian.trivino, albert.vdheide}@softcomputing.es

Abstract

A quasi-periodic signal is a signal that, having a repetitive recognizable periodic pattern, has variations in period and amplitude. Examples of this type of signals are, e.g., the electrocardiogram, the accelerations produced during the human gait, the vibration of a musical instrument, etc.
Currently these signals are analyzed using, among others, statistical methods, e.g. Hidden Markov Models, or Soft Computing methods, e.g. Neural Networks.
This paper contributes to this field by introducing a method for creating a Linguistic Model of a set of these signals based on a Fuzzy Finite States Machine. The obtained model is more intuitive than the existing approaches and its computational implementation is easy to interpret. A practical example of application is provided.

Key words: Signal analysis, Quasi-periodic signal, Fuzzy finite state machines

1 INTRODUCTION

Some signals obtained sampling physical phenomena provide a similar repetitive temporal pattern. These signals are called quasi-periodic due to irregularities in period and amplitude. Popular approaches to deal with these quasi-periodic signals vary from Wavelets transform [3], classical curve fitting methodologies to Hidden Markov Models [7] and Neural Networks [2, 10]. A survey of different techniques that could be applied in these cases can be found in [1].
In this paper we present the fundamentals of a new technique for the analysis of quasi-periodical signals based on Fuzzy Finite States Machines (FFSM). We extend fuzzy logic modelling techniques [13] by defining a linguistic model of the signal which is associated with a new kind of membership function. By creating this type of linguistic model we want to contribute to the field 'Computing with words and perceptions', as defined by Zadeh [15]. The idea consists of aggregating a set of simple perceptions obtained during the temporal sampling of a signal to create a more complex perception of the process in evolution, i.e. the sequence of values of a linguistic variable that describe a signal's evolution is converted into a more general concept that we have called a Linguistic Model of the signal. Fuzzy Finite State Machines can overcome the discontinuous character of FSM [4]. FFSM allows a subset of possible states to be active to some degree at any moment in time [8] resulting in a system with continuous flow from one state into another. Because the activation of states is based upon fuzzy variables this has the additional advantage that the system can explain in human understandable terms what the cause was for current states in the model of the process. The usefulness of fuzzy automata has been demonstrated in many areas, e.g. in fault diagnosis [9], ECG signal recognition [12] and process monitoring [5]. The specific contribution of this paper is the definition of the concept of a fuzzy set of quasi-periodic signals, as well as the idea of a Linguistic Model of the common characteristics of a group of similar quasi-periodic signals. The so obtained signal description and the associated monitoring system are easier to interpret than the ones obtained using classical methods.

2 INITIAL DEFINITIONS

Fuzzy set. Let X be a space of points (objects) with a generic element of X devoted by x. A fuzzy set (class) A in X is characterized by a membership function $f_A(x)$ which associates with each point in X a real number in the interval [0, 1], with the value of $f_A(x)$ representing the "grade of membership" of x in A [14].

Quasi-periodic signal. A quasi-periodic signal is a signal that, having a repetitive recognizable periodic pattern, has variations in period and amplitude.
Fuzzy set of quasi-periodic signals. Following the definition of fuzzy set we define a fuzzy set of quasi-periodic signals U being $LMQ(u)$ a membership function that defines the degree with which a quasi-periodic signal $u(t)$ belongs to U.
We call (LMQ) the Linguistic Model of a set of quasi-periodic signals and it could be considered as an extension of the concept of Linguistic Label, first introduced by Zadeh.
Fuzzy finite state machine (FFSM). The concept of Finite State Machine can be extended using fuzzy logic [6]. In general, a FFSM is a tuple: $FFSM = \{Q, X, Y, f, g, s\}$ where:

- $Q = \{q_1, q_2, \ldots, q_n\}$ is the set of states where q_i can take values in the interval [0,1]
- X is the set of fuzzy input variables
- Y is the set of fuzzy output variables
- f is the state transition function $Q(t) = f(X, Q(t-1))$
- g is the output function $Y(t) = g(X, Q(t-1))$
- s is the initial state. ($s = q_1$ in all the examples presented in this paper)

In the rest of this paper we introduce a method, based on the concept of FFSM, to create LMQ. In this introductory work we do not develop all the possibilities of the model, for example we do not use the set of output variables Y and therefore we will not need to implement the function g.

3 MODEL OF A SET OF PERIODIC RECTANGULAR SIGNALS

In order to describe the method to create a LMQ, we will start with a simple example, creating the Linguistic Model of a set of rectangular signals. Provided $u_1(t) = \sin(\omega t)$ the rectangular signal $u_2(t)$ is defined as:

$$u_2(t) = \begin{cases} -1 & \text{if } u_1(t) \leq -0.5 \\ 0 & \text{if } -0.5 < u_1(t) \leq 0.5 \\ 1 & \text{if } u_1(t) > 0.5 \end{cases}$$

See figure 1 where the signal is represented with $\omega = 1$.

3.1 Define the states of the signal during a period

The first step of the method consists of choosing a set of states which represent the relevant aspects of the repetitive pattern. How to choose and define states is a subjective matter and depends upon the application. In the rectangular signal we have identified four states $Q = \{q_1, q_2, q_3, q_4\}$:

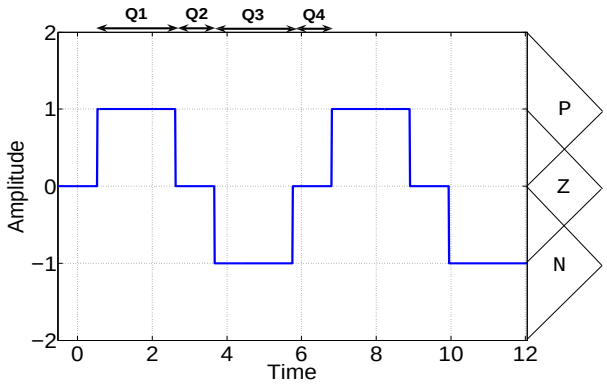


Figure 1: Rectangular signal with related fuzzy labels.

- $q_1 = High$
- $q_2 = Medium$
- $q_3 = Low$
- $q_4 = Medium2$

As can be seen in figure 1 the states correspond to clearly defined parts of the signal.

3.2 Identify the characteristics of different states

In the second step we must obtain from the signal $u(t)$ the set of fuzzy variables x_i that are needed to identify every state. This set of variables $X = \{x_1, x_2, \ldots, x_n\}$ must be defined using their associated linguistic labels A_{ij}. As we will see, depending on user requirements and on the signal complexity, the number of input variables can vary and the procedure to obtaining them from the original signal can differ.
In the example of the rectangular signal we will use a fuzzy input variable x_1 that correspond directly with the signal amplitude and we use a set of three linguistic labels $\{A_{11}, A_{12}, A_{13}\}$ to define the possible values of this variable. For simplicity we have renamed these labels to $\{P, Z, N\}$ meaning Positive, Zero, and Negative. These labels are shown on the right vertical axis of the diagram in figure 1.

3.3 Define the transitions

In this step we define the temporal order with which the signal follows the states. Figure 2 shows the diagram of a FFSM that represents the temporal evolution of the rectangular signal.

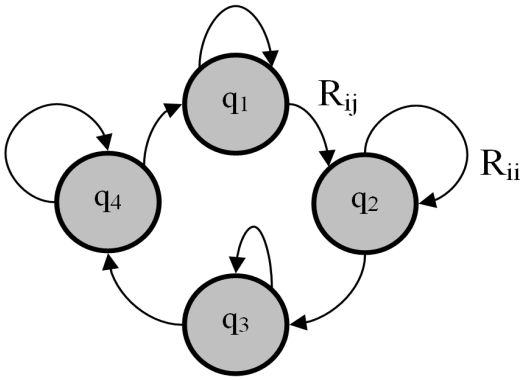


Figure 2: States of the FFSM for the rectangular signal.

3.4 Define the constraints

The next step of the method consists of using a set of fuzzy rules to define the state transition function $Q(t) = f(Q(t-1), X(t))$. We must define how the activation of states is influenced by the input variables and by the degree of activation of the states at the previous instant of time. We distinguish between rules to stay in a state i and rules to change:
R_{ii} : IF $q_i(t-1)$ AND (x_1 is A_{ii}) THEN $q_i(t)$
R_{ij} : IF $q_i(t-1)$ AND (x_1 is A_{ij}) THEN $q_j(t)$
In the example, the current state of the signal is represented by a vector of four states, $Q = \{q_1, q_2, q_3, q_4\}$, and the set of input variables is $X = \{x_1\}$.
The following is the set of rules that describes the transitions in the example of the rectangular signal:

$R_{11} : q_1(t-1) \wedge (x = P) \rightarrow q_1(t)$
$R_{22} : q_2(t-1) \wedge (x = Z) \rightarrow q_2(t)$
$R_{33} : q_3(t-1) \wedge (x = N) \rightarrow q_3(t)$
$R_{44} : q_4(t-1) \wedge (x = Z) \rightarrow q_4(t)$
$R_{12} : q_1(t-1) \wedge (x = Z) \rightarrow q_2(t)$
$R_{23} : q_2(t-1) \wedge (x = N) \rightarrow q_3(t)$
$R_{34} : q_3(t-1) \wedge (x = Z) \rightarrow q_4(t)$
$R_{41} : q_4(t-1) \wedge (x = P) \rightarrow q_1(t)$

In general, it will be needed to incorporate temporal constrains. We do this using two additional linguistic labels (See figure 3):
Time_to_stay$_i$ is the maximum time the signal is allowed to stay in a state i.
Time_to_change$_{ij}$ is the minimum time that the signal must be in state i before changing to state j.
The following set of rules includes the temporal restrictions in the model of the rectangular signal:

$R_{11} : q_1(t-1) \wedge (x = P) \wedge (d_1 = T_{stay1}) \rightarrow q_1(t)$
$R_{22} : q_2(t-1) \wedge (x = Z) \wedge (d_2 = T_{stay2}) \rightarrow q_2(t)$
$R_{33} : q_3(t-1) \wedge (x = N) \wedge (d_3 = T_{stay3}) \rightarrow q_3(t)$
$R_{44} : q_4(t-1) \wedge (x = Z) \wedge (d_4 = T_{stay4}) \rightarrow q_4(t)$
$R_{12} : q_1(t-1) \wedge (x = Z) \wedge (d_1 = T_{change12}) \rightarrow q_2(t)$
$R_{23} : q_2(t-1) \wedge (x = N) \wedge (d_2 = T_{change23}) \rightarrow q_3(t)$
$R_{34} : q_3(t-1) \wedge (x = Z) \wedge (d_3 = T_{change34}) \rightarrow q_4(t)$
$R_{41} : q_4(t-1) \wedge (x = P) \wedge (d_4 = T_{change41}) \rightarrow q_1(t)$

where d_i is the Timespan of the state q_i. A detailed explanation of the calculation of the Timespan is given later.

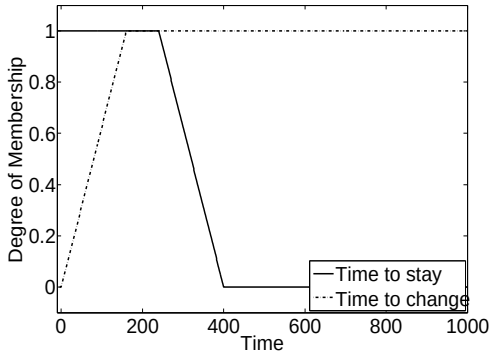


Figure 3: (a) *Time_to_change*, (b) *Time_to_stay*.

4 COMPUTATIONAL IMPLEMENTATION

The objective is to implement the membership function $LMQ(u)$ that provides the degree of matching of a signal with the Linguistic Model and therefore the degree with which a signal $u(t)$ belongs to the fuzzy set U. Figure 4 shows a diagram including the main elements of the computational implementation of this membership function.

Extraction of characteristics block consists of sensors, hardware processing, and software mechanisms that obtain from the analyzed physical phenomena the sequence of numerical data, $u(t)$, that will be used as input to the Fuzzification block.

The Fuzzification block uses the linguistic labels for describing the relevant characteristics of the signal to obtain the set of fuzzy input variables X.

The f(Q, X) block consists of the implementation of the set of rules. They are Takagi-Sugeno-Kang inference rules of zero-order with a constant output of value equal to 1 [11].
They can be expressed as:
$R_{ii} : (\hat{q}_i(t-1) = 1) \wedge (x = A_{ii}) \rightarrow q_i(t) = 1$
$R_{ij} : (\hat{q}_i(t-1) = 1) \wedge (x = A_{ij}) \rightarrow q_j(t) = 1$
And are implemented as:

$R_{ii} : q_i(t) = MIN(\hat{q}_i(t-1), A_{ii}(x))$
$R_{ij} : q_j(t) = MIN(\hat{q}_i(t-1), A_{ij}(x))$
where;
$q_i(t)$ is the activation of state i,
$A_{ii}(x)$ is the degree with which the signal matches the conditions to stay in the state i,
$A_{ij}(x)$ is the degree with which the signal matches the conditions to change from state i to j,
$\hat{q}_i(t)$ is the normalization of $q_i(t)$ that is calculated in the Feedback block.
The total degree of activation of a state q_i is calculated as:

$$q_i(t) = max_{r=1}^{N}(q_{ir}) \tag{1}$$

where;
q_{ir} is the contribution to the degree of activation of the state i delivered by the rule r. Note that defuzzification is not needed.

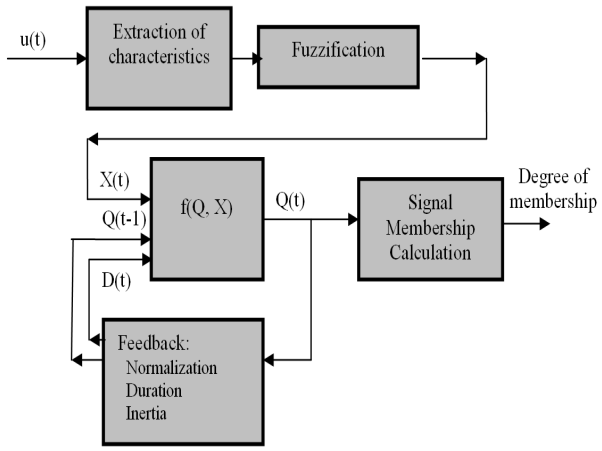


Figure 4: Computational implementation of LMQ.

The Feedback block performs three functions, namely, Normalization, Timespan calculation and Low pass filtering.
We need as input to the FFSM the current distribution of activation between all the possible states. Using the following formula we perform such normalization (redistribution) by processing the values obtained at the FFSM output.

$$\hat{q}_i(t) = \frac{q_i(t)}{\sum_{j=1}^{n} q_j(t)} \tag{2}$$

where n is the number of states.
In addition, the Feedback block calculates a vector $D = (d_1, d_2, \ldots, d_n)$ which represents the Timespan in each state; d_i is calculated as the time that $q_i > \theta$. Where θ is a parameter depending upon the specific implementation.

Finally, in order to minimize the influence of noise on the states evolution the Feedback Block implements a low-pass filter that introduces certain inertia to the state. The value of q_i is modified by the formula:

$$q_i(t) = \alpha q_i(t-1) + (1-\alpha) q_i(t) \tag{3}$$

where α is a parameter called inertia factor. The higher the value the more resistant to change.

Membership calculation block calculates to what degree the current signal $u(t)$ is a member of the set of signals U. At every instant the vector $q_1(t), q_2(t), ..., q_n(t)$ represents the degree with which the signal is matched by the model.

We define the **Instantaneous Degree of Matching** with the model of a signal $u(t)$ as:

$$IDM(u,t) = OR_1^n(q_i(t)) = max_1^n(q_i(t)) \tag{4}$$

where n is the number of states.
Based on this measure we define the **Accumulated Degree of Matching** as:

$$ADM(u,t) = \frac{1}{t}\sum_{s=1}^{t} IDM(u,s) \tag{5}$$

And finally we define the **Degree of Membership** of the signal $u(t)$ to the fuzzy set U as the Accumulated Degree of Matching after a relatively long period of time, e.g. several periods:

$$LMQ(u) = \frac{1}{T}\sum_{s=1}^{T} IDM(u,s) \tag{6}$$

5 FIRST RESULTS

As an illustration of how the model works let us to consider $u_3(t)$, which is obtained by adding a small sinusoidal signal to $u_2(t)$.

$$u_3(t) = u_2(t) + sin(100\omega t) \tag{7}$$

Figure 5 shows the curves corresponding with $IDM(u_3)$, and $ADM(u_3)$. See how $IDM(u_3)$ changes when the signal does not match the model.

The calculation of the membership of function $u_2(t)$ and $u_3(t)$ to the set of functions provides:
$LMQ(u_2) = 1$
$LMQ_(u_3) = 0.87$
Signal $u_2(t)$ is clearly a prototypical signal of our model and therefore has a higher membership value to the set of rectangular signals than u_3.

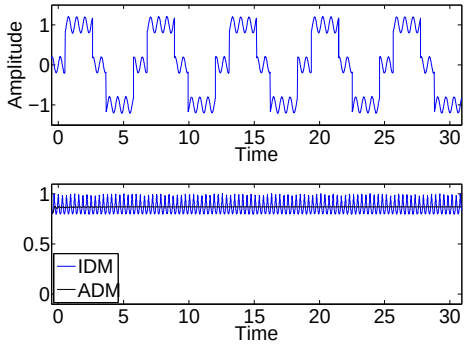


Figure 5: $IDM(u_3)$, and $ADM(u_3)$.

In addition, as an illustration of how the model works with the temporal constrains, figure 6 shows a signal $u_4(t)$ with increasing period. It shows how $IDM(u_4)$, and $ADM(u_4)$ decreases when timespan in a state is longer than the defined by $Time_to_Stay$.

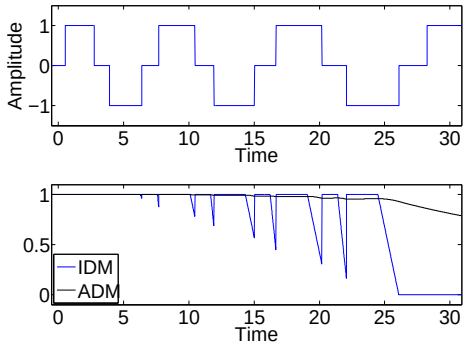


Figure 6: $IDM(u_4)$, and $ADM(u_4)$

6 A PRACTICAL EXAMPLE

In music, the note 'A4' corresponds with a sine with a frequency $440Hz$ (figure 7). However when we produce 'A4' with different musical instruments the resultant acoustic signals will be different from the perfect signal and will have a different degree of membership to the fuzzy set defined around the 'pure' signal. We have followed the steps described above to create a membership function that calculates the degree with which a signal belongs to the set 'Note-A4' of quasi-periodic signals.

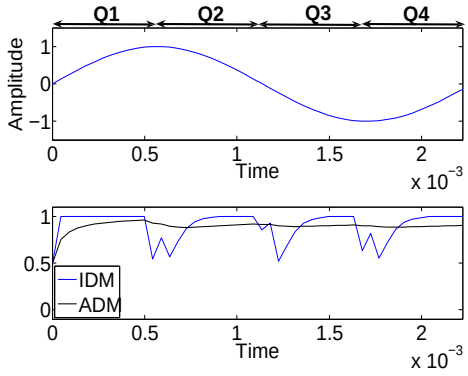


Figure 7: Sinusoidal signal

6.1 Define the states

As before the states correspond to parts in the signal period. A depiction of the states is given in figure figure 7 where the states are shown at the top. The different states are defined as:

- $q_1 = Positive_increasing$; this state matches the first part of the signal, when the amplitude is positive and increasing.
- $q_2 = Positive_decreasing$; this state matches the second part of the signal, when the signal has passed its positive peak and is decreasing towards an amplitude of zero.
- $q_3 = Negative_decreasing$; the amplitude is negative and getting more negative.
- $q_4 = Negative_increasing$; the amplitude is negative, and getting smaller (going towards zero).

6.2 Identify characteristics

In order to distinguish the states in this case, we use as an additional input variable x_2 based on the derivative of $u(t)$. The derivative makes it possible to distinguish between the increasing and decreasing parts of the signal. We will use two linguistic labels to assign fuzzy values to the signal amplitude $x_1(t)$, namely, Positive (P) and Negative (N) and we will use two linguistic labels to assign fuzzy values to x_2, namely, Decreasing (D), and Increasing (I).

6.3 Define the transitions

A basic sequential ordering of states is sufficient. In fact the state transitions will be identical to rectangular signal (figure 1).

6.4 Define the constraints

$$R_{11} : q_1(t-1) \wedge (x_1 = P) \wedge (x_2 = I) \wedge (d_1 = T_{stay}) \rightarrow q_1(t)$$
$$R_{22} : q_2(t-1) \wedge (x_1 = P) \wedge (x_2 = D) \wedge (d_2 = T_{stay}) \rightarrow q_2(t)$$
$$R_{33} : q_3(t-1) \wedge (x_1 = N) \wedge (x_2 = D) \wedge (d_3 = T_{stay}) \rightarrow q_3(t)$$
$$R_{44} : q_4(t-1) \wedge (x_1 = N) \wedge (x_2 = I) \wedge (d_4 = T_{stay}) \rightarrow q_4(t)$$
$$R_{12} : q_1(t-1) \wedge (x_1 = P) \wedge (x_2 = D) \wedge (d_1 = T_{change}) \rightarrow q_2(t)$$
$$R_{23} : q_2(t-1) \wedge (x_1 = N) \wedge (x_2 = D) \wedge (d_2 = T_{change}) \rightarrow q_3(t)$$
$$R_{34} : q_3(t-1) \wedge (x_1 = N) \wedge (x_2 = I) \wedge (d_3 = T_{change}) \rightarrow q_4(t)$$
$$R_{41} : q_4(t-1) \wedge (x_1 = P) \wedge (x_2 = I) \wedge (d_4 = T_{change}) \rightarrow q_1(t)$$

The parameters used in this experiment are: for inertia $\alpha = 0.5$ and for the timespan $\theta = 0.01$. Note that due to the symmetry of sinus function T_{change} and T_{stay} are equal for all rules.

6.5 Experimental results

We have calculated the membership of various A4-tone signals: a perfect A4-tone, an A4-tone produced by a piano, an A4-tone produced by a violin, and an A4-tone produced by a tuning fork. Additionally we calculated the membership of a A5-tone signal produced by a piano to the set of A4-tone signals.

$LMQ(Reference_{A4}) = 0.9$
$LMQ(Piano_{A4}) = 0.74$
$LMQ(Violin_{A4}) = 0.56$
$LMQ(Fork_{A4}) = 0.87$
$LMQ(Piano_{A5}) = 0$

As can be seen in the obtained results the similar signals produce high degrees of membership while the last signal obtains a low degree. In fact the last signal fails to match the first state and therefore does not synchronize with the model.

7 CONCLUSIONS

We have introduced the fundamentals for a new technique of analysis of quasi-periodical signals. The use of linguistic labels to define the set boundaries makes it easier to design and understand the model. The results so far have been promising and there are many possibilities for further improvements and research. Future investigation will focus on how to define with more precision the Linguistic Models corresponding with complex sets of signals.

References

[1] M. Berthold and D. J. Hand, editors. *Intelligent Data Analysis: An Introduction*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1999.

[2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, November 1995.

[3] S. C. Burrus, R. A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelets Transforms*. Prentice Hall, August 1997.

[4] M. Doostfatemeh and S. C. Kremer. New directions in fuzzy automata. *International Journal of Approximate Reasoning*, 38(2):175–214, February 2005.

[5] R. Kempf and J. Adamy. Sequential pattern recognition employing recurrent fuzzy systems. *Fuzzy Sets and Systems*, 146(3):451–472, September 2004.

[6] J. N. Mordeson and D. S. Malik. *Fuzzy Automata and Languages: Theory and Applications*. Chapman & Hall, CRC, London, Boca Raton, FL, 2002.

[7] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, pages 267–296, 1990.

[8] L. Reyneri. An introduction to fuzzy state automata. In *Biological and Artificial Computation: From Neuroscience to Technology*, pages 273–283. Springer Berlin / Heidelberg, 1997.

[9] G. G. Rigatos and S. G. Tzafestas. Fuzzy automata for fault diagnosis: A syntactic analysis approach. In *Methods and Applications of Artificial Intelligence*, pages 301–310. 2004.

[10] B. D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, January 1996.

[11] M. Sugeno. *Industrial applications of fuzzy control*. Elsevier Science Pub. Co., 1985.

[12] M. B. Tumer, L. A. Belfore, and K. M. Ropella. A syntactic methodology for automatic diagnosis by analysis of continuous time measurements using hierarchical signal representations. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 33(6):951–965, 2003.

[13] R. R. Yager and D. P. Filev. *Essentials of Fuzzy Modelling and Control*. John Wiley & Sons, Inc, 1994.

[14] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

[15] L. A. Zadeh. Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems*, 4(2):103–111, May 1996.

ALGORITMO DE SEGMENTACIÓN EN TIEMPO REAL SOBRE VÍDEO COMPRIMIDO H.264/AVC

C.J. Solana-Ciprés[1] **L. Rodríguez-Benítez**[2] **J. Moreno-García**[3] **G. Fernández-Escribano**[3]

[1] E.S. Informática de Ciudad Real, cayetanoj.solana@uclm.es
[2] E.U. Politécnica de Almadén, luis.rodriguez@uclm.es
[3] E.U.I.T. Industrial de Toledo, {juan.moreno, gerardo.fernandez}@uclm.es

Resumen

En este artículo se presenta un algoritmo de segmentación en tiempo real que trabaja directamente sobre la señal de vídeo comprimido en formato H.264/AVC. Se utiliza una representación lingüística de los objetos móviles de la escena basada en lógica difusa y se explotan algunas de las características de la compresión de vídeo en formato H.264/AVC, como el modo de decisión en la representación de los macrobloques. El algoritmo ha sido aplicado al seguimiento de objetos móviles en escenas de cámara estática en tiempo real con resultados satisfactorios.

Palabras Clave: segmentación, modo de decisión, H.264/AVC, sistemas de tiempo real.

1. INTRODUCCIÓN

El término segmentación en visión por computador se refiere al reconocimiento y particionamiento de los objetos presentes en una escena de vídeo. Una de las soluciones clásicas consiste en utilizar la intensidad de cada píxel de una escena de vídeo, si bien este tipo de algoritmos no son adecuados para aplicaciones de tiempo real. Para mejorar la eficiencia se ha de trabajar directamente sobre la señal de vídeo comprimido. En el presente estudio se expone una novedosa técnica de segmentación sobre vídeo comprimido en formato H.264, el cual constituye un estándar de codificación de vídeo avanzado que proporciona tasas de compresión notablemente superiores a estándares previos sin un elevado incremento de complejidad en su diseño.

El principal inconveniente ligado al procesamiento de vídeo comprimido es la escasez de información y la incertidumbre asociada a la misma. En el algoritmo de segmentación no se utiliza información de luminancia y crominancia a nivel de píxel, sino que se emplea información asociada a la compensación de movimiento entre cuadros. Este hecho provoca un incremento de la incertidumbre asociada a la misma. En este entorno ruidoso, la lógica difusa se presenta como el soporte teórico más adecuado para disminuir la influencia negativa de la incertidumbre. Así, se emplean etiquetas lingüísticas para dotar de una semántica propia a los datos utilizados y se proponen representaciones lingüísticas difusas para mejorar la interpretación de los resultados.

El artículo se estructura del modo descrito a continuación. En la Sección 2 se describen algunas características del estándar H.264/AVC. En la Sección 3 se estudia el estado del arte de la segmentación sobre vídeo comprimido. Posteriormente, en la Sección 4 se definen los conceptos lingüísticos difusos que se emplean en el algoritmo. En la Sección 5 se analiza el algoritmo de segmentación propuesto. A continuación se exponen los resultados obtenidos en la Sección 6 y en la Sección 7 se detallan las conclusiones.

2. CARACTERÍSTICAS DE H.264/AVC

H.264/AVC [1] constituye un estándar de codificación ampliamente extendido y con un prometedor futuro. En [6] se analiza en profundidad el diseño del formato H.264/AVC. Una de las características de H.264 es la mejora del proceso de estimación y compensación de movimiento, en el cual se eliminan las posibles redundancias temporales que pudieran existir entre imágenes independientes. Este proceso se realiza mediante la comparación de la imagen que está siendo codificada con otras anteriores o posteriores, en la búsqueda de un patrón que indique el tipo de movimiento que se produce entre ellas. Este patrón se representa como un vector de movimiento (Sección 2.3). Para mejorar la eficiencia de la codificación, el estándar H.264 soporta la subdivisión de los macrobloques en bloques

menores, capaces de contener y aislar el movimiento por sí mismos (Figura 1), y toma estas subdivisiones como unidades de comparación. Por tanto, el modo óptimo de codificación para un macrobloque, junto con su vector o vectores de movimiento, se empleará para predecir un macrobloque dado.

2.1. TIPOS DE IMÁGENES

En la codificación de una secuencia de vídeo puede haber imágenes (cuadros) de tres tipos. En un cuadro de **Tipo I**, la codificación se realiza teniendo en cuentra sólo dicho cuadro, es decir, explotando las redundancias espaciales. En un cuadro de **Tipo P** o **Tipo B** se emplea estimación y compensación de movimiento; en el primer caso se utiliza como referencia un cuadro anterior y en el segundo, un cuadro pasado y otro futuro.

2.2. MACROBLOQUE

El macrobloque (MB) es la unidad básica en la que se dividen las imágenes para su codificación. Contiene la información correspondiente a una región de 16x16 píxeles y en función del tipo de codificación que se le aplique se distinguen básicamente dos tipos: **Intra**, en el que se toma directamente el macrobloque de la imagen en curso que se está codificando y se aplican algoritmos de Intra-predicción según el estándar H.264, e **Inter**, en el que se emplea la estimación y la compensación de movimiento para la codificación de un MB, empleando para ello un MB de referencia previo (anterior, posterior o una combinación de ambos).

Un MB puede codificarse con diferentes tamaños de bloque (ver Figura 1). Como H.264 permite el cálculo de la estimación y compensación de movimiento con tamaño de bloque variable, se puede considerar el modo de decisión o tipo de bloque como un elemento que aporta información adicional para conocer el grado de movimiento presente en un bloque determinado de una imagen dada (ver Figura 2). Así, un mayor número de subdivisiones en un macrobloque implica un mayor grado de movimiento implícito en dicha zona. Esta información es relevante para el algoritmo de segmentación.

2.3. VECTOR DE MOVIMIENTO

Un vector de movimiento define una distancia y una dirección y tiene dos magnitudes: *right_x* y *down_x*. En H.264, un macrobloque de tipo Inter puede tener desde 0 hasta 16 vectores de movimiento, dado que cada bloque y sub-bloque tendrá asociado un vector de movimiento. Por ejemplo, un MB de tipo Intra no tendrá ningún vector de movimiento, un MB de tipo

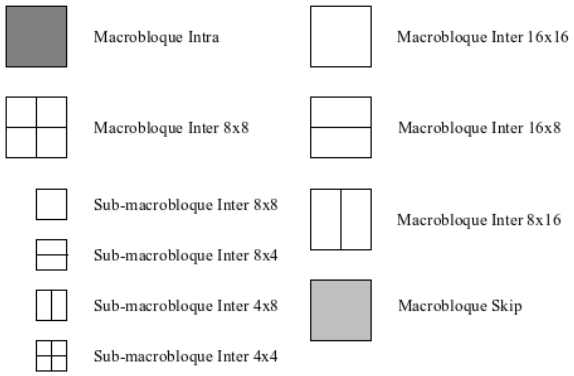


Figura 1: Tipos de macrobloque.

Figura 2: Partición de macrobloques de un cuadro.

Inter dividido en dos bloques de 16x8 tendrá dos vectores de movimiento (uno asociado a cada bloque), y un MB de tipo Inter dividido en cuatro bloques 8x8, con partición en cada bloque de 4x4 píxeles, tendrá 16 vectores de movimiento ($4 + 4 + 4 + 4 = 16$).

3. TRABAJOS PREVIOS

La lógica difusa ha sido aplicada en multitud de tareas de visión por computador ([4], [8]) dado su gran potencial y su vínculo con el lenguaje natural. En situaciones ambiguas, como las provocadas por vídeos de baja calidad o por insuficiencia de información del contexto, la lógica borrosa permite traducir la información del vídeo en representaciones lingüísticas susceptibles de ser procesadas mediante técnicas difusas.

Los algoritmos de segmentación que trabajan directamente sobre vídeo comprimido explotan dos características del macrobloque: los vectores de movimiento y los coeficientes DCT. La transformada DCT (Transformada Discreta del Coseno) se aplica sobre la información residual a codificar tras la predicción de los macrobloques y genera, para cada macrobloque, una matriz de coeficientes DCT que representan la codificación del residuo. En [12] se presenta un algoritmo de segmentación que utiliza un modelo de campo aleatorio de Markov (MRF) para la segmentación en función del campo disperso de vectores de movimiento obtenidos directamente del flujo de bits de la secuencia de

vídeo. Una aproximación más reciente ha sido propuesta en [3], donde se presenta un algoritmo de segmentación espacio-temporal en tiempo real. La segmentación espacial se basa exclusivamente en el campo de vectores movimiento de cada cuadro, mientras que para la clasificación de las regiones segmentadas se utiliza el valor del residuo de los bloques.

4. DEFINICIÓN DE CONCEPTOS LINGÜÍSTICOS DIFUSOS

Además de la falta de vectores que permitan obtener un campo de movimiento denso e uniforme en una secuencia de cuadros, los vectores de movimiento son aproximaciones imprecisas a los campos de flujo óptico reales. En cuadros de vídeo MPEG [10], las áreas que con más probabilidad pueden contener vectores de movimiento erróneos son aquellas en las que hay una escasa variación en los valores de luminancia (Figura 3(a)) y aquellas que corresponden con macrobloques frontera de objetos en la escena (Figura 3(b)). En estas situaciones, es complejo para los algoritmos de compensación de movimiento encontrar correspondencias entre macrobloques de diferentes cuadros. A partir de un conjunto de experimentos [9] se pudo observar que al estar los cuadros de tipo B más cerca de los cuadros de referencia que los cuadros de tipo P, en general, sus vectores se aproximaban más a los vectores de movimiento reales. Los campos de movimiento en los cuadros de tipo P, debido a esta distancia, serían menos fiables.

Como consecuencia de los errores introducidos por los codificadores, la mayoría de técnicas que trabajan en el dominio de vídeo comprimido aplican algún filtro al campo de vectores de movimiento para eliminar estos valores anómalos. La aplicación de este tipo de filtros requiere de la utilización de los valores de otros vectores de movimiento en el propio cuadro o en cuadros vecinos. Al utilizar mayor cantidad de datos, este tipo de operaciones suelen ser costosas desde el punto de vista computacional y, por tanto, no accesibles en algoritmos de tiempo real.

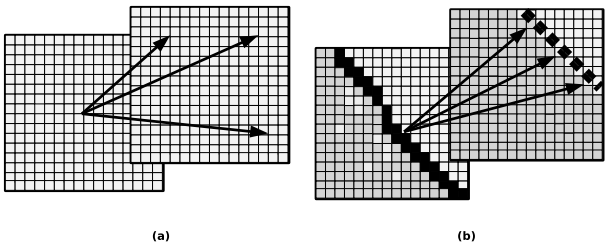


Figura 3: Incertidumbre inherente a la codificación.

Esta incertidumbre asociada a la codificación de macrobloques justifica la utilización de la lógica difusa en el ámbito del vídeo comprimido. Dado que el objetivo de este trabajo consiste en la obtención de una descripción lingüística de las regiones de la escena en las que se detecta movimiento, es necesario definir previamente el conjunto de elementos lingüísticos que se utilizan en la descripción de regiones.

4.1. VARIABLE LINGÜÍSTICA

En este estudio se utilizan cinco variables lingüísticas [11] para representar la posición horizontal (PH), la posición vertical (PV), la velocidad horizontal (VH), la velocidad vertical (VV) y el grado de movimiento (GM). El punto crítico del diseño es la definición de la función de pertenencia de cada valor del dominio a las diferentes etiquetas lingüísticas. En las Figuras 4 a 6 se pueden visualizar las etiquetas lingüísticas de tres variables utilizadas en el presente estudio, cuyos valores se pueden modificar para adaptarlos al escenario con el objetivo de obtener un diseño óptimo para cada secuencia.

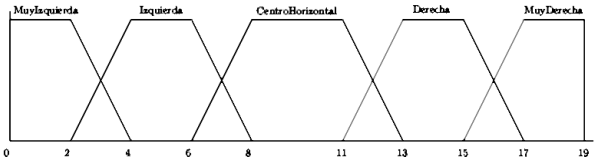


Figura 4: Variable Posición Horizontal PH.

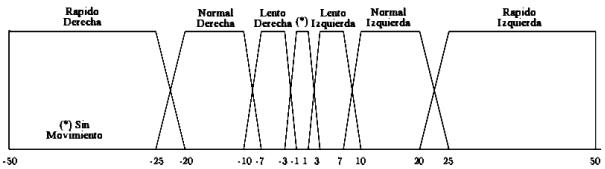


Figura 5: Variable Velocidad Horizontal VH.

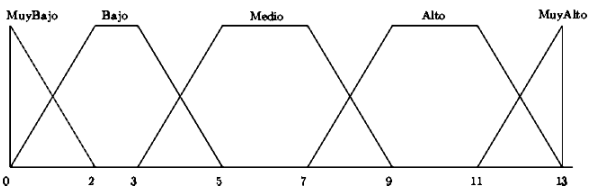


Figura 6: Variable Grado de Movimiento GM.

Las variables que representan la posición del macrobloque (PH y PV) vienen determinadas por el tamaño de los cuadros del vídeo. Así, para un vídeo con resolución 320x240 habrá que diseñar la variable PH con dominio de 0 a 19 porque en cada fila del cuadro habrá 20 macrobloques ($20 \cdot 16 = 320$) de 16x16 píxeles cada uno. En dicho ejemplo, el dominio de la variable PV irá de 0 a 14 porque en cada columna habrá 15 macrobloques de 16x16 píxeles cada uno ($15 \cdot 16 = 240$).

La variable que representa el grado de movimiento de un macrobloque GM se define sobre el dominio del conjunto de las posibles decisiones que puede tomar un codificador de vídeo H.264 a la hora de representar cada macrobloque de un cuadro. Como se ha comentado en la Sección 2.2, un macrobloque puede codificarse de diversas formas atendiendo a la división en bloques. El conjunto de decisiones se representa de forma diferente según el codificador empleado. En el Cuadro 1 se muestra la nomenclatura adoptada en la codificación de los macrobloques. Nótese cómo el identificador indica de forma creciente el grado de movimiento implícito en el macrobloque, lo cual queda reflejado en el diseño de las etiquetas lingüísticas (ver Figura 6).

Cuadro 1: Nomenclatura adoptada en la codificación.

Id	Tipo de Macrobloque
0	Skip, sin información adicional respecto al macrobloque homónimo del cuadro anterior.
1	Macrobloque codificado en 16x16 píxeles.
2	Macrobloque codificado como dos bloques de 16x8.
3	Macrobloque codificado como dos bloques de 8x16.
4	Bloque 8x8 codificado como un sub-bloque de 8x8.
5	Bloque 8x8 codificado como dos sub-bloques de 8x4.
6	Bloque 8x8 codificado como dos sub-bloques de 4x8.
7	Bloque 8x8 codificado como cuatro sub-bloques de 4x4.
8	Macrobloque codificado en cuatro bloques 8x8.
9	Macrobloque intra-frame codificado como 4x4.
10	Macrobloque intra-frame codificado como 16x16.
11	Sin uso en una codificación normal.
12	Sin uso en una codificación normal.
13	Macrobloque intra-frame codificado como 8x8.

4.2. INTERVALO LINGÜÍSTICO

Un **intervalo lingüístico ordenado** $I_X^{p,q}(a)$ es la representación difusa de un valor real a fuzzificado sobre una variable lingüística X y se define como:

$$I_X^{p,q}(a) = \begin{cases} \{A_X^p : \mu_{A_X^p}(a);\ A_X^q : 1 - \mu_{A_X^p}(a)\} & Si\ p \neq q \\ \{A_X^p : 1\} & Si\ p = q \end{cases} \quad (1)$$

donde p y q son las posiciones en X de las dos etiquetas lingüísticas que componen el intervalo y el valor de $\mu_{A_X^p}(a)$ siempre ha de ser mayor que cero. Por ejemplo, si un macrobloque está situado en la columna número siete ($c = 7$), el resultado de fuzzificar ese valor sobre la variable PH (Figura 4) es $I_{PH}^{2,3}(7) = \{Izquierda : 0,5\ ; Centro\ Horizontal : 0,5\}$.

4.3. VECTOR DE MOVIMIENTO LINGÜÍSTICO

Un **vector de movimiento lingüístico** [7] es la representación lingüística de un vector de movimiento asociado a un macrobloque:

$$VML =< NC, I_{PH}, I_{PV}, I_{VH}, I_{VV} > \quad (2)$$

donde el primer elemento identifica el número del cuadro en el que se encuentra el vector de movimiento y los cuatro elementos restantes corresponden con cuatro intervalos lingüísticos. Los intervalos I_{PH} e I_{PV} representan la columna y la fila del macrobloque al que se asocia el vector. Los intervalos I_{VH} e I_{VV} se calculan a partir de las magnitudes *right_x* y *down_x* del vector de movimiento asociado al macrobloque. En el Cuadro 2 se muestra un ejemplo de vector de movimiento lingüístico.

Cuadro 2: Vector de movimiento lingüístico.

NumeroCuadro	23
$I_{PH}(3)$	{*Muy Izquierda*: 0,5; *Izquierda*, 0,5}
$I_{PV}(3)$	{*Muy Arriba*: 0,25; *Arriba*, 0,75}
$I_{VH}(45)$	{*Rápido Izquierda*, 1}
$I_{VV}(0)$	{*Sin Movimiento*, 1}

Un **vector de movimiento lingüístico válido** es aquel vector de movimiento lingüístico que contiene información relevante acerca de la dirección y velocidad de un objeto, es decir, aquel que contiene una etiqueta distinta a *Sin Movimiento* en al menos uno de los dos intervalos lingüísticos que representan la velocidad (I_{VH} e I_{VV}). Según esto, el vector representado en el Cuadro 2 es válido porque contiene la etiqueta *Rápido Izquierda* en el intervalo I_{VH}.

4.4. MODO DE DECISIÓN LINGÜÍSTICO

Un **modo de decisión lingüístico** es la representación lingüística del modo decisión escogido por el codificador para realizar la partición del macrobloque (ver Figura 1):

$$MD =< NC, I_{PH}, I_{PV}, I_{GM} > \quad (3)$$

donde el primer elemento identifica el número del cuadro al que pertenece el macrobloque. I_{PH} e I_{PV} representan la posición horizontal y vertical del macrobloque y el intervalo I_{GM} se obtiene mediante la fuzzificación del modo de decisión (ver Cuadro 1) respecto de la variable lingüística GM (ver Figura 6). Por ejemplo, si un macrobloque se codifica como 4 bloques de tamaño 8x8, el modo de decisión del codificador será 8 (ver Cuadro 1) y, al realizar la fuzzificación del valor 8 sobre la variable GM, se obtiene como resultado el intervalo lingüístico $I_{GM}^{3,4}(8) = \{Medio : 0,5\ ; Alto : 0,5\}$.

4.5. BLOB LINGÜÍSTICO

Un **blob lingüístico** es una tupla de ocho elementos que representa de manera lingüística a una región (blob) en el cuadro que está compuesta por uno o más vectores de movimiento conceptualmente similares:

$$BL =< NC, Tam, MBs, I_{PH}, I_{PV}, I_{VH}, I_{VV}, I_{GM} > \quad (4)$$

donde NC es el número de cuadro en el que se localiza el blob, Tam es el número de vectores de movimiento lingüísticos que componen el blob, MBs es una lista con los identificadores de los macrobloques agregados al blob y los últimos cinco elementos son intervalos lingüísticos asociados a cada una de las variables lingüísticas utilizadas y que caracterizan la posición, la velocidad y el grado de movimiento del blob.

5. PROCESO DE SEGMENTACIÓN

En la Figura 7 se muestra el esquema general del sistema propuesto. El flujo de vídeo comprimido en formato H.264/AVC constituye la información inicial de entrada. La primera fase consiste en la decodificación del flujo de vídeo en H.264 con el objetivo de extraer los vectores de movimiento y los modos de decisión de acuerdo a un formato adecuado para su procesamiento durante el resto del algoritmo. En la etapa de *fuzzificación*, se transforman los vectores de movimiento del flujo de vídeo en vectores de movimiento lingüísticos y se obtiene la representación difusa del valor de los modos de decisión tal y como se anticipa en las Secciones 4.3 y 4.4. A continuación, se elimina información ruidosa mediante la utilización exclusiva de los vectores de movimiento válidos. La siguiente fase es la segmentación. Se lleva a cabo de forma diferenciada para los vectores de movimiento lingüísticos válidos y para los modos de decisión. En ambos casos se obtiene como resultado un conjunto de blobs lingüísticos. El objetivo de la duplicidad de blobs consiste en dotar de mayor consistencia y fiabilidad a la segmentación. Los dos conjuntos de blobs lingüísticos se comparan y se genera un conjunto de blobs resultado. Este último conjunto representa la salida final del proceso de segmentación. Nótese que el resultado no es un conjunto de *objetos* sino de *blobs*, que constituyen el paso previo a la identificación de los objetos presentes en la escena.

5.1. SEGMENTACIÓN DE LOS VECTORES DE MOVIMIENTO

El primer proceso de segmentación está basado en una medida de distancia euclídea D que se define en la Sección 5.1.1 y un agrupamiento o *clustering* de los vectores de movimiento lingüísticos válidos en blobs. Si existe un parecido entre un blob y un vector que no pertenece a ningún blob, entonces el vector pasa a formar parte del blob, modificándose la representación de este último en función de la del vector añadido y de su propia representación, de acuerdo a la *agregación ponderada* definida en la Sección 5.1.3.

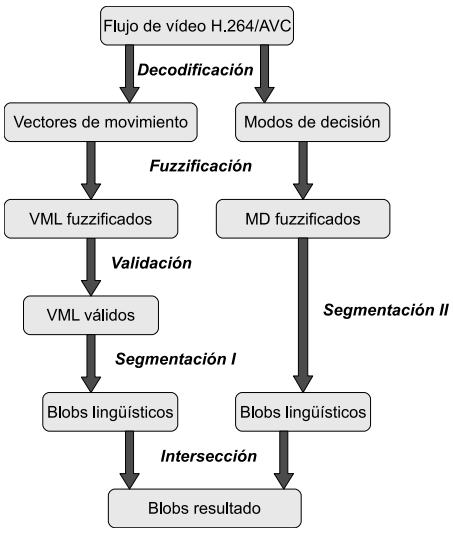


Figura 7: Esquema general del sistema.

5.1.1. Medida de distancia entre intervalos lingüísticos

Para disponer de un criterio que establezca la similitud entre un vector y un blob se propone una **medida de distancia euclídea** basada en el orden de numeración de las etiquetas definidas sobre una variable lingüística. Se define la distancia entre dos intervalos lingüísticos como:

$$D_X = D(I_X^{p,q} - I_X^{r,s}) = \left| \frac{p+q}{2} - \frac{r+s}{2} \right| \quad (5)$$

Por ejemplo, considerando la variable lingüística PH de la Figura 4, la distancia euclídea entre los intervalos:

$$I_{PH}^{2,3}(7) = \{Izquierda : 0,5 \; ; Centro\ Horizontal : 0,5\} \quad (6)$$

$$I_{PH}^{3,3}(9) = \{Centro\ Horizontal : 1,0\} \quad (7)$$

será:

$$D_{PH}(I_{PH}^{2,3}(7), I_{PH}^{3,3}(9)) = \left| \frac{2+3}{2} - \frac{3+3}{2} \right| = 0,5 \quad (8)$$

5.1.2. Distancia total entre un vector y un blob

A partir de esta distancia euclídea, se define la **distancia total** entre un vector de movimiento lingüístico válido $VMLV_x$ y un blob BL_y como el máximo de

las distancias euclídeas entre cada uno de los intervalos lingüísticos de las cuatro variables lingüísticas PH, PV, VH y VV:

$$DT(VMLV_x, BL_y) = max(D_{PH}, D_{PV}, D_{VH}, D_{VV}) \quad (9)$$

Un vector $VMLV_x$ se agregará al blob BL_y, es decir, pasará a formar parte de la misma región si se cumplen las siguientes condiciones:

1. La distancia total del vector $VMLV_x$ al blob BL_y es menor que la distancia total de cualquier otro vector de movimiento lingüístico a BL_y.
2. La distancia $DT(VMLV_x, BL_y)$ es menor que el umbral U_{vm} a partir del cual un vector se considera susceptible de pertenecer a un blob. El umbral U_{vm} es un valor que oscila entre 0 y 1, y en las pruebas realizadas, los valores óptimos se sitúan entre 0,3 y 0,6.

5.1.3. Agregación ponderada de un vector en un blob

La **agregación ponderada** [7] de $VMLV_x$ a BL_y da como resultado la representación modificada del blob BL'_y y se ejecuta en cinco fases:

1. $\Pi_{NC}(BL'_y) = \Pi_{NC}(BL_y)$, es decir, el número de cuadro al que pertenece el blob permanece invariable.
2. $\Pi_{Tam}(BL'_y) = \Pi_{Tam}(BL_y) + 1$, es decir, se incrementa en una unidad el tamaño del blob.
3. $\Pi_{MBs}(BL'_y) = \Pi_{MBs}(BL_y) + \Pi_{MB}(VMLV_x)$, es decir, se inserta en la lista de macrobloques del blob BL'_y el número de macrobloque al que se asocia el vector $VMLV_x$.
4. Se realiza la *unión ponderada* de los intervalos correspondientes a las variables lingüísticas PH, PV, VH y VV del vector en función del tamaño del blob. Por ejemplo, la unión ponderada de los intervalos (el nombre de las etiquetas se abrevia por motivos de claridad):

$$I_{PH}^{2,3}(BL_y) = \{Izq : 0,5 \; ; CenHor : 0,5\} \quad (10)$$

$$I_{PH}^{3,3}(VMLV_x) = \{CenHor : 1,0\} \quad (11)$$

considerando que el tamaño del blob BL_y es 8 será:

$$I_{PH}(BL'_y) = \{Izq : \tfrac{0,5\cdot 8}{9} + \tfrac{0}{9} \; ; CenHor : \tfrac{0,5\cdot 8}{9} + \tfrac{1,0}{9}\} = \\ = \{Izq : 0,44 \; ; CenHor : 0,55\} \quad (12)$$

5. $\Pi_{I_{GM}}(BL'_y) = 0$; el último elemento del blob lingüístico, I_{GM}, tendrá valor nulo ya que en los vectores de movimiento no se representa información del modo de decisión.

5.2. SEGMENTACIÓN DE LOS MODOS DE DECISIÓN

El tamaño de las particiones de los macrobloques en H.264 aporta información adicional para conocer el grado de movimiento presente en un macrobloque determinado, como se ha comentado en la Sección 2.2. En el Cuadro 1 y en la Figura 6 se refleja cómo la convención de nombres utilizada indica de un modo creciente el grado de movimiento implícito en dicho macrobloque. Así, un macrobloque de tipo 0 (*skip*) indica que el grado de movimiento es nulo, mientras que un macrobloque de tipo 10 (*intra-frame*) refleja un alto grado de movimiento.

En la segmentación de los modos de decisión, se trabajará con blobs en los que las componentes I_{VH} y I_{VV} tendrán valor nulo porque son intervalos lingüísticos referidos a la velocidad que toman su valor de los vectores de movimiento.

5.2.1. Medida de distancia posicional

Se define la **distancia posicional** entre un modo de decisión MD_x y un blob BL_y como el máximo de las distancias euclídeas de los intervalos pertenecientes a las variables lingüísticas de posición PH y PV:

$$DP(MD_x, BL_y) = max(D_{PH}, D_{PV}) \quad (13)$$

5.2.2. Agregación ponderada de un modo de decisión en un blob

En el Cuadro 3 se muestra una comparativa de la frecuencia con la que aparece cada modo de decisión en un conjunto de secuencias de vídeo de ejemplo [5]. Se detecta que el 70.2 % de los macrobloques de la escena se codifican como cuadros *skip* (valor 0) o como 16x16 (valor 1), los dos tipos de cuadros que implican menor grado de movimiento en el macrobloque. Con la finalidad de acelerar el proceso de segmentación, se obviarán los macrobloques codificados como 0 ó 1 porque no representan movimiento en ese área del cuadro.

Un modo de decisión lingüístico MD_x se agregará a un blob lingüístico BL_y si se cumplen tres condiciones:

1. $DP(MD_x, BL_y)$ es menor que la distancia posicional de cualquier otro modo de decisión al blob BL_y.

Cuadro 3: Frecuencia de modos de decisión.

Tipo	Macrobloques	Frecuencia(%)
0	87006	19.6
1	224552	50.6
2	14720	3.3
3	9650	2.2
4	47461	10.7
5	8370	1.9
6	6831	1.5
7	2142	0.5
8	1571	0.4
9	21091	4.8
10	3817	0.9
11	0	0.0
12	0	0.0
13	16489	3.7

2. $DP(MD_x, BL_y)$ es menor que el umbral U_{md}, donde U_{md} es un valor entre 0 y 1 que indica el límite posicional.

3. El modo de decisión del codificador es diferente de 0 y 1, por la razón expuesta anteriormente.

La agregación ponderada del modo de decisión MD_x al blob lingüístico BL_y es análoga a la agregación ponderada descrita para la segmentación de los vectores de movimiento con la diferencia de que se realiza la unión ponderada para el intervalo I_{GM} y no para I_{VH} e I_{VV}.

5.3. INTERSECCIÓN DE BLOBS LINGÜÍSTICOS

Como resultado de los dos procesos independientes de segmentación (Secciones 5.1 y 5.2) se obtienen dos conjuntos de blobs lingüísticos, de los cuales se eliminan aquellos que tengan un tamaño menor a un umbral U_{tam} predefinido. Existen dos tipos de blobs en función del proceso de segmentación del que provengan:

$$BL_{vm} =< NC, Tam, MBs, I_{PH}, I_{PV}, I_{VH}, I_{VV}, 0 > \quad (14)$$

$$BL_{md} =< NC, Tam, MBs, I_{PH}, I_{PV}, 0, 0, I_{GM} > \quad (15)$$

Cada BL_{md} se asigna al blob BL_{vm} con el que mayor coincidencia tenga, i.e., aquel en el que el número de macrobloques coincidentes sea mayor. Si a un BL_{vm} no le ha sido asignado ningún blob de modos de decisión, entonces permanece inmutable. Sin embargo, si un BL_{mv} tiene asignados uno o más BL_{md}, se eliminarán de la lista de macrobloques del blob BL_{mv} aquellos macrobloques que no aparezcan en alguno de los blobs asociados de tipo BL_{md}. De este modo, se elimina ruido en los resultados finales y se atenúa el problema de la sobresegmentación.

6. RESULTADOS EXPERIMENTALES

El algoritmo desarrollado ha sido probado con una secuencia de vídeo extraída de [5] con una resolución de 640x240. Dicha secuencia tiene bastante ruido debido a que el nivel de iluminación no es homogéneo a lo largo de la secuencia. El valor de los parámetros de configuración con el que se optimizan los resultados en esta secuencia es: $U_{vm} = 0,5$, $U_{md} = 0,3$ y $U_{tam} = 6$. El valor óptimo de los parámetros dependerá del escenario específico en cada ocasión.

En las Figuras 8 y 9 se muestra el resultado de la aplicación del algoritmo en dos cuadros de la secuencia. En el primero de ellos se observa que el algoritmo detecta una región en el cuadro, la cual corresponde a una persona. Sin embargo, en la segunda escena se puede observar cómo el algoritmo agrupa en un mismo blob a dos personas que caminan juntas y en la misma dirección. Este problema, conocido como solapamiento, puede corregirse en el seguimiento (tracking).

Figura 8: Cuadro resultado.

Figura 9: Cuadro con problema de solapamiento.

En el Cuadro 4 se muestra el tiempo de procesamiento del algoritmo propuesto. Dado que se obtienen resultados en torno a 25 ms. (equivalente a 40 cuadros por segundo) se puede asegurar que la segmentación se realiza en tiempo real.

Cuadro 4: Tiempos de procesamiento.

Proceso	**Tiempo medio**
Fuzzificación	7,63 ms.
Validación	1,38 ms.
Segmentación	13,19 ms.
Intersección	2,62 ms.
TOTAL	**24,82 ms.**

7. CONCLUSIÓN

En este estudio se ha presentado un algoritmo de segmentación espacial basado en técnicas de lógica difusa aplicable a secuencias de vídeo H.264. Una de las ventajas del sistema propuesto radica en que los resultados obtenidos generan representaciones lingüísticas conceptuales que permiten describir de modo comprensible la posición, velocidad y tamaño de las regiones detectadas en la segmentación. Por otro lado, el tamaño de la información de entrada -modos de decisión y vectores de movimiento- que utiliza el algoritmo es sensiblemente menor que la información que emplean otros algoritmos que trabajan a nivel de píxel. La segmentación espacial presenta problemas de sobresegmentación y solapamiento que podrán ser corregidos en la fase de seguimiento. Asimismo, en el seguimiento se puede corregir la detección de falsos objetos ya que no tendrán continuidad en el tiempo.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Educación y Ciencia (TIN2007-62568) y el Gobierno Regional de Castilla-la Mancha (PAC06-0141 y PBC06-0064).

Referencias

[1] ITU-T Recommendation H.264. *Advanced Video Coding for Generic Audiovisual Services*, Mayo 2003.

[2] J.S.R. Jang. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 23, Nº 3, Pág.665-685. Mayo-Junio 1993.

[3] Zhi Liu, Yu Luy y Zhaoyang Zhang. Real-time spatiotemporal segmentation of video objects in the H.264 compressed domain. *Journal of Visual Communication and Image Representation*, Nº 18, Pág.275-290, 2007.

[4] P. Matsakis, J.M. Keller, L. Wendling, J. Marjamaa y O. Sajhputera. Linguistic description of relative positions in images. *IEEE Transactions on systems, man and cybernetics*. Vol. 31, Nº 4, 2001.

[5] PETS Data Sets. *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, 2002. Disponible en ftp://ftp.pets.rdg.ac.uk/pub/ (último acceso 28 de marzo de 2008).

[6] Iain Richardson. H.264 and MPEG-4 Video Compression. Ed. John Wiley & Sons. 2003.

[7] L. Rodríguez-Benítez, J. Moreno-García, J.J. Castro, J. Albusac y L. Jiménez. An approximate reasoning technique for segmentation on compressed MPEG video. *2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007)*. Barcelona, España. Marzo 2007.

[8] O. Sajhputera, P. Matsakis, J.M. Keller y R. Bondougula. Linguistic descriptions for an object in motion. *Proceedings of the NAFIPS*, Pág.243-248. New Orleans, Louisiana, USA. 2002.

[9] R. Wang, H. Zhang, Y. Zhang. A Confidence Measure Based Moving Object Extraction System Built for Compressed Domain *IEEE International Symposium on Circuits and Systems (ISCAS 2000)*, Pág.21-24, 2000.

[10] T. Yoshida, H. Katoh, Y. Sakai. Block matching motion estimation using block integration based on reliability metric. *Proceedings of the International Conference on Image Processing (ICIP '97)*, Vol. 2, Pág.152-155, 1997.

[11] L.A. Zadeh. The Concept of Linguistic Variable and its Application to Approximate Reasoning. *Information Sciences*, Nº 8, Pág.199-249, 1975.

[12] Wei Zeng, Jun Du, Wen Gao y Qingming Huang. Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model. *Real-Time Imaging*, Nº 11, Pág.290-299, 2005.

Sistema neuro-difuso intervalo-valorado aplicado a la segmentación de imágenes de ultrasonidos

Humberto Bustince[1] Gerardo Artola[1] Miguel Pagola[1] Edurne Barrenechea[1] Hamid Tizhoosh[2]

[1] Dpto. Automática y Computación, Universidad Pública de Navarra, Pamplona, España
bustince@unavarra.es
[2] Pattern Analysis and Machine Intelligence Lab, University of Waterloo, Waterloo, Canadá
tizhoosh@uwaterloo.ca

Resumen

En este trabajo desarrollamos un sistema neuro difuso intervalo valorado para la segmentación de imágenes de ultrasonidos. El objetivo del sistema es segmentar las zonas de imágenes de mamas sospechosas de ser quiste. El sistema nos permite obtener un conjunto de reglas para clasificar cada pixel de la imágen en dos zonas diferentes. Entrenamos nuestro modelo con unas imágenes previamente segmentadas por un radiólogo. Los resultados experimentales de segmentar otras imágenes de test muestran un alto potencial de poder aplicar esta tecnica.

Palabras Clave: Imágenes de Ultrasonidos; Segmentación; Conjuntos Intervalo Valorados Difusos; Sistema Neuro-difuso.

1 Introducción

Las imágenes de ultrasonidos son utilizadas para la evaluación de anomalías en las mamas durante el diagnóstico de exámenes mamográficos. Los quistes en el pecho, pequeñas acumulaciones de uido, son la causa más importante de tumores benignos (no cancerígenos). Cuando se examina una anomalía en la mama, para determinar si es un quiste o no, el radiólogo estudia el tamaño, el área y otras características morfológicas de la anomalía. Por tanto, el desarrollo de métodos que separen automáticamente la zona perteneciente al quiste del resto de la imagen, es muy importante para un rápido y correcto diagnóstico.

La segmentación de imágenes de ultrasonidos está fuertemente in uenciada por su mala calidad. El ruido, las sombras y la atenuación hacen que la segmentación sea una tarea muy complicada. Existen varios métodos para segmentar imágenes de ultrasonidos, una revisión muy completa puede encontrarse en [10]. Algunos métodos utilizan técnicas de aprendizaje. En todos ellos se utilizan imágenes previamente segmentadas por un experto radiólogo para entrenar un sistema. En [15] se utiliza la técnica del aprendizaje por refuerzo para entrenar a un agente dedicado a la segmentación de imágenes de próstata. Zhang et. al. [17] entrenan los pesos de una transformada ponderada wavelet para detectar microcalcificaciones en imágenes mamográficas. Las redes neuronales fueron utilizadas en [11] para identificar posibles daños en imágenes del hígado y en [4] se utilizan algoritmos genéticos para entrenar los pesos de una red neuronal autoorganizada (red de Kohonen).

La metodología que proponemos está basada en el aprendizaje neuro-difuso [9]. Este técnica nos permite obtener automáticamente un conjunto de reglas difusas a partir de un conjunto de datos.

Debido a la complejidad de las imágenes en este trabajo vamos utilizar conjuntos intervalo valorados difusos. De esta forma intentamos representar la incertidumbre que está presente en estas imágenes. Mendel en [7] realiza la adaptación de los algoritmos de aprendizaje para sistemas difusos a conjuntos difusos tipo 2 y a conjuntos intervalo valorados difusos.

El objetivo de este trabajo es desarrollar un sistema de lógica difusa de tipo 2 intervalar (Interval Type 2 Fuzzy Logic System, a partir de ahora IT2FLS) para segmentar quistes en imágenes de ultrasonidos.

Como veremos más adelante un sistema intervalar es equivalente a un sistema difuso, pero los conjuntos que representan a los antecedentes y a los consecuentes del conjunto de reglas son conjuntos intervales en lugar de conjuntos difusos. En este trabajo vamos a utilizar un algoritmo de aprendizaje a partir de datos para crear el sistema intervalar. Para utilizar el algoritmo de aprendizaje debemos representar el sistema en función de parámetros. Debido a que las funciones

de pertenencia intervalares son más complejas que las funciones de pertenencia difusas, un sistema intervalar con el mismo número de reglas, variables y etiquetas ling¨uísticas, tiene ms parámetros libres que un sistema difuso simple. Muchos autores [6, 8] dicen que esta característica, tener más grados de libertad, que es una ventaja frente a los sistemas difusos clásicos. Aseguran que de esta forma se puede manejar la imprecisión existente en los datos de entrenamiento y así alcanzar una mejor representación del problema y obtener mejores resultados. Es decir, para un misma complejidad de representación del problema (etiquetas y reglas), los sistemas intervalares tienen un mayor número de parámetros que pueden adaptarse mejor a los datos de entrenamiento.

El trabajo esá organizado de la siguiente forma: primero hacemos una introducción de los sistemas de lógica difusa de tipo 2 intervalar (IT2FLS). En la sección 3 mostramos el algoritmo de Mendel [7] que permite definir un sistema a partir de datos de entrenamiento. Posteriormente, en la siguiente sección 4 presentamos el modelo que proponemos para segmentar imágenes de ultrasonidos. Finalizamos mostrando los resultados experimentales y las conclusiones y trabajos futuros.

2 Sistemas neuro-difusos intervalo-valorados

Un conjunto intervalo-valorado difuso es tal que el grado de pertenencia de cada elemento al conjunto esta dado por un subintervalo cerrado del intervalo [0,1].

El concepto de conjunto difuso tipo 2 fue introducido por Zadeh ([16]) como una generalización de un conjunto difuso ordinario. En los conjuntos difusos de tipo 2, el grado de pertenencia de cada elemento al conjunto es así mismo un conjunto difuso en [0,1].

Un conjunto difuso de tipo 2 intervalar $\overline{\overline{A}}$ en U esta definido por

$$\overline{\overline{A}} = \{(u, A(u), \mu_u(x))|u \in U, A(u) \in L([0,1])\},$$

donde $A(u) = [\underline{A}(u), \overline{A}(u)]$ es una función de pertenencia; es decir, un subintervalo cerrado de $[0,1]$, y la función $\mu_u(x)$ representa el conjunto difuso asociado con el elemento $u \in U$ obtenido cuando x engloba al intervalo $[0,1]$; $\mu_u(x)$ es dada de la siguiente forma:

$$\mu_u(x) = \begin{cases} a \text{ si } \underline{A}(u) \leq x \leq \overline{A}(u) \\ 0 \text{ en otro caso} \end{cases},$$

donde $0 \leq a \leq 1$.

En [6]–[8], se demuestra que un conjunto difuso de tipo 2 intervalar es igual que un conjunto intervalo-valorado difuso si tomamos $a = 1$.

Ejemplo: Podemos representar un conjunto intervalo valorado difuso mediante la función de pertenencia del extremo superior del intervalo y la función de pertenencia del extremo inferior del intervalo. En este trabajo todas las funciones de pertenencia las vamos a representar mediante una función gaussiana:

$$\mu_k^l(x_k) = exp\left[-\frac{1}{2}\left(\frac{x_k - m_k^l}{\sigma_k^l}\right)\right] \ \sigma_k^l \in [\sigma_{k1}^l, \sigma_{k2}^l] \quad (1)$$

En la figura 1 mostramos la función de pertenencia para el valor de $m_k^l = 0.4$. Tomamos $\sigma_{k1}^l = 0.05$ para representar el extremo inferior y $\sigma_{k2}^l = 0.1$ para el extremo superior.

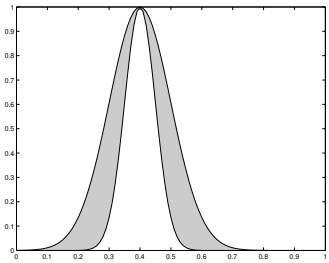


Figura 1: Función de pertenencia intervalo valorada con primaria gaussiana, $m = 0.4$, $\sigma_{k1}^l = 0.05$ y $\sigma_{k2}^l = 0.1$.

Un IT2FLS, es un sistema basado en reglas en el que los conjuntos que representan los antecedentes y los consecuentes son conjuntos difusos intervalares de tipo 2 (o conjuntos intervalo valorados difusos). En la figura 2 observamos los apartados más importantes de un IT2FLS (ver [7]).

El sistema recibe como entrada unos valores crisp. El fusificador se encarga convertir la entrada crisp en un conjunto intervalar. Después, en la parte principal del sistema utilizando las reglas y el motor de inferencia obtenemos unas conclusiones, representadas por conjuntos intervalares. Para poder utilizar estas conclusiones en un sistema real, el defusificador se encarga de convertir las conclusiones en valores crisp. Como las conclusiones son conjuntos intervalares, en los IT2FLS existe otra fase que se llama reducción de tipo. En la reducción de tipo se convierten los conjuntos intervalares en conjuntos difusos. De esta forma después se puede utilizar cualquier técnica de defusificación para pasar de un conjunto difuso a un valor crisp.

Consideramos un IT2FLS con p entradas $x_1 \in X_1, \cdots, x_p \in X_p$ y una salida $y \in Y$. Asumimos que

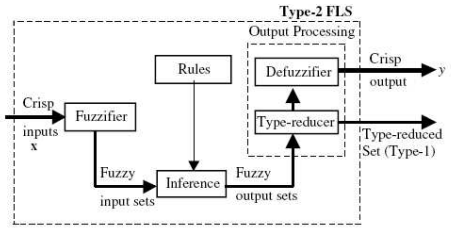


Figura 2: Sistema Intervalo valorado tipo 2

el sistema tiene M reglas de la siguiente forma:

$$R^l: \text{ SI } x_1 \text{ es } F_1^l \text{ y } \ldots \text{ y } x_p \text{ es } F_p^l, \\ \text{ENTONCES } y \text{ es } G^l \text{ con } l = 1, \ldots, M \quad (2)$$

Donde $F_1^l, \ldots, F_p^l, G^l$ son conjuntos intervalo valorados difusos. Cada regla la interpretamos como una implicación:

$$R^l : F_1^l \times F_2^l, \cdots \times F_p^l \rightarrow G^l = A^l \rightarrow G^l \text{ con } l = 1, \ldots, M \quad (3)$$

R^l esta descrita por la función de pertenencia $\mu_{R^l}(\mathbf{x}, y) = \mu(x_1, \ldots, x_p, y)$ donde,

$$\mu_{R^l}(\mathbf{x}, y) = \mu_{A^l \rightarrow G^l}(\mathbf{x}, y) = \left[\bigcap_{i=1}^{p} \mu_{F_i^l}(x_i)\right] \bigcap \mu_{G^l}(y) \quad (4)$$

Utilizando la extensión de la regla composicional de Zadeh para conjuntos intervalares el consecuente lo calculamos de la siguiente forma:

$$\mu_{B^l}(y) = \mu_{A^l \circ R^l}(y) = \bigcup_{x \in X} \left[\mu_{A_x}(\mathbf{x}) \bigcap \mu_{R^l}(\mathbf{x}, y)\right] \\ \text{con } y \in Y, l = 1, \ldots, M \quad (5)$$

Existen varios trabajos sobre la inferencia cuando tenemos reglas intervalo valoradas [1, 5], pero en este trabajo vamos a utilizar el método utilizado por Mendel en [7] (para un estudio más profundo ver [7]).

Si representamos los conjuntos intervalares utilizando una función de pertenencia para el extremo superior y otra función de pertenencia para el extremo inferior, se pueden simplificar las operaciones, por ejemplo: el grado de activación de una regla, para un vector de entrada $\mathbf{x}$ se puede calcular de la siguiente forma:

$$F^l(\mathbf{x}) = \left[\underline{f}^l(\mathbf{x}), \overline{f}^l(\mathbf{x})\right] \text{ donde}$$

$$\underline{f}^l(\mathbf{x}) = \underline{\mu}_{F_1^l}(x_1) \wedge \underline{\mu}_{F_2^l}(x_2) \cdots \wedge \underline{\mu}_{F_p^l}(x_p) \quad (6)$$

$$\overline{f}^l(\mathbf{x}) = \overline{\mu}_{F_1^l}(x_1) \wedge \overline{\mu}_{F_2^l}(x_2) \cdots \wedge \overline{\mu}_{F_p^l}(x_p) \quad (7)$$

De forma análoga en [7] se presentan las expresiones del motor de inferencia la reducción de tipo para conjuntos intervales. Estas expresiones son más sencillas que cuando utilizamos conjuntos difusos tipo 2.

3 Diseño de Sistemas Neuro-difusos intervalo-valorados a partir de datos

En este apartado mostramos como diseñar completamente el IT2FLS utilizando datos de entrenamiento. Dada una colección de N pares de datos entrada-salida $(\mathbf{x}^{(1)}, y^{(1)}), (\mathbf{x}^{(2)}, y^{(2)}), \cdots (\mathbf{x}^{(N)}, y^{(N)})$ donde $\mathbf{x}^{(t)}$ es el vector de datos de entrada e $y^{(t)}$ es el valor escalar de salida para el par de entrenamiento t-ésimo, definimos el error del IT2FLS para la entrada t-ésima como:

$$e^{(t)} = f_{s2}(\mathbf{x}^{(t)}) - y^{(t)} \quad t = 1, \ldots, N \quad (8)$$

donde f_{s2} es la salida del sistema IT2FLS. Dicho valor depende de los parámetros que definen los conjuntos intervalares de los antecedentes y consecuentes de las reglas (m_k^l, σ_{k1}^l, σ_{k2}^l, y_l^j e y_r^j, con $k = 1, \ldots, p$, para p entradas, $j = 1, \ldots, M$, para M reglas). Además y_l^j e y_r^j representan los extremos de las funciones de pertenencia del consecuente tras la reducción de tipo. Definimos el error cuadrático del IT2FLS para la entrada t-ésima como

$$E^{(t)} = \frac{1}{2}[e^{(t)}]^2 \quad t = 1, \ldots, N \quad (9)$$

y la función de performance del sistema como

$$E = \sum_{i=1}^{N} E^{(t)} \quad t = 1, \ldots, N \quad (10)$$

El entrenamiento del IT2FLS consiste en encontrar los valores de los parámetros que minimizan (9) o (10). Si minimizamos (9) (dicho método se explica en [7]) realizamos un aprendizaje en modo incremental, ya que ajustamos los parámetros del sistema en forma iterativa después de procesar cada par de entrenamiento. Si minimizamos (10), el método que hemos seguido en este trabajo, realizamos aprendizaje en modo bloque, ya que ajustamos los parámetros después de procesar todos los pares de entrenamiento. El proceso de los N pares de entrenamiento y el posterior ajuste de parámetros se denomina época de entrenamiento. En este trabajo, en cada época, modificamos los parámetros del sistema en una cantidad proporcional a los gradientes $\frac{\partial E}{\partial m_k^l}$, $\frac{\partial E}{\partial \sigma_{k1}^l}$, $\frac{\partial E}{\partial \sigma_{k2}^l}$, $\frac{\partial E}{\partial y_l^j}$ e $\frac{\partial E}{\partial y_r^j}$. Esta cantidad la calculamos utilizando el algoritmo de *resilient backpropagation* (en adelante RPROP, ver [12]). Este algoritmo ajusta automtícamente sus propios parámetros en el proceso de entrenamiento, es fácil de implementar y además se logra una gran velocidad de convergencia. Un resumen del proceso de aprendizaje, para n épocas, se describe a continuación

1. Inicializar todos los parámetros en los antecedentes y los consecuentes
2. Iniciar el contador de, épocas $ep = 1$
3. Aplicar N entradas $\mathbf{x}^{(t)}$ al sistema y calcular las correspondientes salidas $f_{s2}(\mathbf{x}^{(t)})$.
4. Calcular el Error del sistema utilizando (10).
5. Calcular los gradientes $\frac{\partial E}{\partial m_k^l}$, $\frac{\partial E}{\partial \sigma_{k1}^l}$, $\frac{\partial E}{\partial \sigma_{k2}^l}$, $\frac{\partial E}{\partial y_l^j}$ e $\frac{\partial E}{\partial y_r^j}$.
6. Calcular el valor de los nuevos parámetros según RPROP.
7. $ep = ep + 1$. Si $ep = n + 1$, PARAR, en otro caso volver al paso 3.

En el paso 4 del proceso de aprendizaje se calcula la salida del sistema neuro-difuso para el par de entrenamiento t-ésimo como

$$f_{s2}(\mathbf{x}^{(t)}) = \frac{y_l(\mathbf{x}^{(t)}) + y_r(\mathbf{x}^{(t)})}{2} \tag{11}$$

donde

$$y_l = \frac{\sum_{i=1}^{L} \overline{f}^i y_l^i + \sum_{j=L+1}^{M} \underline{f}^j y_l^j}{\sum_{i=1}^{L} \overline{f}^i + \sum_{j=L+1}^{M} \underline{f}^j} \tag{12}$$

$$y_r = \frac{\sum_{i=1}^{R} \underline{f}^i y_r^i + \sum_{j=R+1}^{M} \overline{f}^j y_l^j}{\sum_{i=1}^{R} \underline{f}^i + \sum_{j=R+1}^{M} \overline{f}^j} \tag{13}$$

donde $\underline{f}^i$ y $\overline{f}^j$ se calculan según (6) y (7) respectivamente. En tanto L y R se obtienen al resolver el cálculo del centroide (ver [7]) en la reducción de tipo de la salida del motor de inferencia.

4 Método propuesto

El objetivo de nuestro método es clasificar cada pixel de la imagen. Es decir, decidir si pertenece a la zona importante a segmentar (quiste) o si pertenece al fondo. Debido a que en las imágenes suelen aparecer varias zonas para analizar, el usuario debe seleccionar el punto central de la zona que quiere segmentar. Por tanto es un sistema parcialmente supervisado ya que el usuario aporta cierta información.

El IT2FLS que proponemos tiene 5 variables de entrada, una salida, y 20 reglas. Las variables de entrada son

1. pos_i: la distancia, en píxeles, del pixel considerado al punto central seleccionado por el usuario, en el eje horizontal.
2. pos_j: la distancia, en píxeles, del pixel considerado al punto central seleccionado por el usuario, en el eje vertical.
3. *dist*: cercanía del pixel considerado al punto central seleccionado por el usuario, utilizando un algoritmo de llenado.
4. *edg*: nivel de borde del pixel considerado.
5. *mgr*: media del nivel de gris del vecindario del pixel considerado (vecindario de 5×5 píxeles) en la imagen realzada.

En los siguientes apartados mostramos las técnicas que utilizamos para obtener los valores de cercanía, borde y realce de la imagen.

4.1 Imagen realzada

En este apartado utilizamos el algoritmo propuesto por Sabha et. al. [13] para realzar las imágenes de ultrasonidos. La idea principal es eliminar el ruido de las imágenes y realzar el nivel de gris de la zona seleccionada. Primero se elimina el ruido utilizando el filtro de la mediana (7×7 o 9×9). Después cada pixel es fusificado según su nivel de gris utilizando una función de pertenencia que tiene en cuenta el valor medio de gris alrededor del pixel central, (para mas detalles [14]).

4.2 Imagen de cercanía

En este apartado vamos a procesar la imagen de tal forma que el valor que asignamos a cada pixel representa la cercanía al pixel central de la zona seleccionada, pero esta cercanía va a tener en cuenta las diferentes zonas que tiene la imagen. Primero calculamos los bordes de la imagen realzada utilizando el algoritmo de Canny [3]. Después, partiendo del punto seleccionado por el usuario y utilizando 8 conectividad, utilizamos un algoritmo de llenado. El algoritmo de llenado va etiquetando los píxeles con la distancia al pixel central. En la primera iteración a los vecinos del pixel central se les asigna la distancia 1. En cada iteración del algoritmo la distancia aumenta una unidad. Los bordes los utilizamos como diques de tal forma que la distancia obtenida esta relacionada con las zonas de la imagen.

4.3 Imagen de falso borde

Para crear la imagen que representa el grado de borde que es cada pixel, calculamos lo que comúnmente se llama falso borde. En [2] presentamos un método para obtener falsos bordes utilizando t-normas y t-conormas. Para cada pixel de la imagen construimos

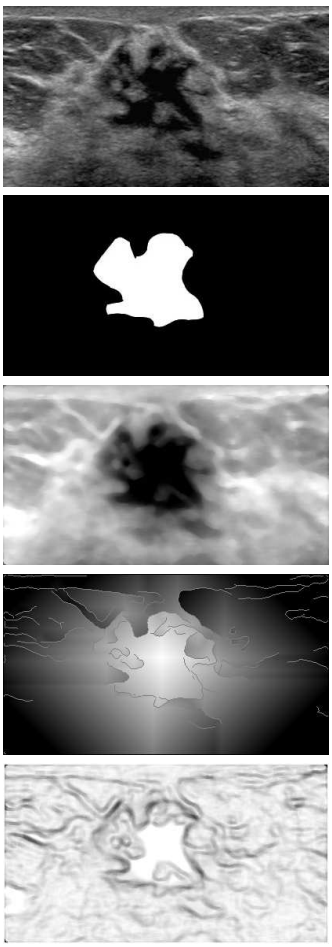


Figura 3: De arriba abajo: Imagen Original, Imagen segmentada ideal, Imagen realzada, Imagen de cercanía, Imagen de falsos bordes.

una submatriz (3×3, 5×5, etc.). Aplicando t-normas a los elementos de esa matriz obtenemos el valor inferior de un intervalo. Si aplicamos t-conormas a los elementos de la submatriz obtenemos el valor superior del intervalo. Llamamos falso borde a la longitud del intervalo, es decir, a la diferencia entre el valor superior y el inferior. En el caso de utilizar la t-norma mínimo y la t-conorma máximo el falso borde es la resta del mayor elemento y el menor elemento presentes en la submatriz.

5 Resultados experimentales

Para evaluar nuestro método utilizamos un conjunto de ocho pares de imágenes consistentes en una imagen original y la segmentación ideal generada por un experto radiólogo. Cada imagen tiene asociado un punto en el centro de la región a segmentar elegido por el experto. De estas imágenes escogemos dos para entrenamiento mientras que las restantes se utilizan para comprobar el porcentaje de acierto del sistema. De estas dos imágenes, escogemos de forma aleatoria 200 píxeles para formar los pares de entrenamiento. Estos píxeles los dividimos a su vez en un conjunto específico de entrenamiento y un conjunto de validación en una proporción de 80% y 20% respectivamente. El conjunto de entrenamiento se usa para ajustar los parámetros del sistema según se describe en la sección 3. El conjunto de validación los utilizamos para evaluar el rendimiento del sistema IT2FLS durante la fase de entrenamiento, de tal forma que escojemos como mejor sistema el definido en la época de menor error medido sobre este conjunto validación.
Para evaluar la eficacia del método de segmentación utilizamos la medida del solapamiento de áreas (S_A) entre el área de la imagen segmentada ideal y la imagen segmentada por el sistema IT2FLS.

$$S_A = \frac{|Ideal \cap IBF|}{|Ideal \cup IBF|} \tag{14}$$

Donde $Ideal$ es la imagen binaria segmentada por el experto, IBF es la imagen binaria obtenida utilizando nuestro método, $\cap$ y $\cup$ la intersección y la unión entre conjuntos crisp respectivamente.

Debido a que los parámetros del sistema los elegimos de forma aleatoria antes del entrenamiento, se obtienen diversas soluciones posibles. Después de realizar una veintena de entrenamientos, mostramos los resultados obtenidos en el mejor caso. En las figuras 4 y 5 mostramos las imágenes binarias obtenidas. En la tabla 1 mostramos los porcentajes del solapamiento de áreas para las 8 imágenes (recordamos que sólo las imágenes 1 y 2 fueron utilizadas para el entrenamiento).

Imagen	1	2	3	4	5	6	7	8
$S_A(\%)$	80	72	77	88	86	69	65	56

Tabla 1: Solapamiento del área.

6 Conclusiones y trabajos futuros

Hemos propuesto un nuevo método para segmentar imágenes de ultrasonidos utilizando un sistema difuso de tipo 2 intervalar. El sistema tiene 5 variables de entrada, una variable de salida y 20 reglas. Entrenamos el sistema utilizando unas imágenes previamente segmentadas por un experto. Alcanzamos un porcentaje medio del 70% de acierto con respecto a unas imágenes ideales.

Para mejorar el sistema podemos introducir nuevas variables de entrada, añadiendo nuevos preprocesos a

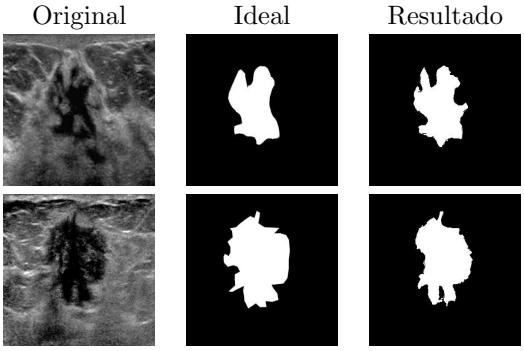


Figura 4: Comparación de las imágenes ideales con las obtenidas por el sistema de las imágenes utilizadas para el entrenamiento.

la imagen, como las transformadas wavelet que han dado muy bueno resultados en imágenes de ultrasonidos. Otra mejora para tratar la incertidumbre existente en las imágenes es no tratar los valores de entrada como crisp, sino como intervalos.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto TIN2007-65981.

Referencias

[1] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets", *International Journal of Approximate Reasoning*, vol. 23, pp. 137–209, 2000.

[2] H. Bustince, E. Barrenechea, M. Pagola, R. Orduna, Construction of interval type 2 fuzzy images to represent images in grayscale. False edges. FUZZY IEEE 2007 . London, U.K.

[3] J. Canny, A Computational approach to edge detection, IEEE Transactions on Pattern Recognition and Machine Intelligence, 8(6), 679-698 (1986).

[4] Z. Dokur and T. Olmez, Tissue segmentation in ultrasound images by using genetic algorithms, *Expert Systems with Applications*, 34(4), 2739-2746 (2008)

[5] M.B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets", *Fuzzy Sets and Systems*, vol. 21, pp. 1–17, 1987.

[6] J.M. Mendel, R. I. Bob John. Type-2 fuzzy sets made simple, *IEEE Transactions on Fuzzy Systems*, **10**(2), pages 117–127, 2002.

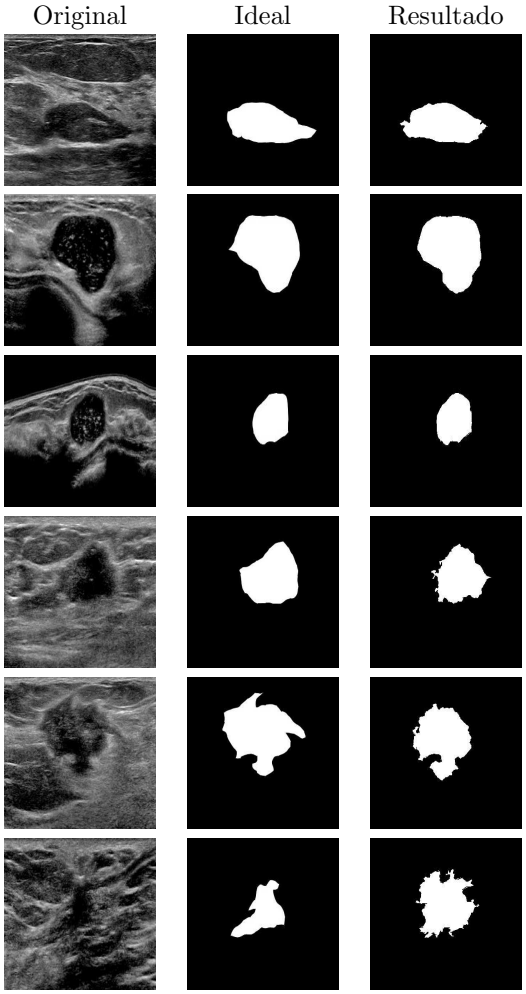


Figura 5: Comparación de las imágenes ideales con las obtenidas por el sistema de las imágenes de test.

[7] J.M. Mendel. Uncertain Rule-Based Fuzzy Logic Systems, *Prentice-Hall, Upper Saddle River*, NJ, 2001.

[8] J.M. Mendel. Advances in type-2 fuzzy sets and systems, *Information Sciences*, Volume 177(1), Pages 84-110, 2007.

[9] Mitra, S., Hayashi, Y., Neuro-fuzzy rule generation: survey in soft computing framework, IEEE Transactions on Neural Networks, 11(3), 748 - 768 May (2000).

[10] J.A. Noble and D. Boukerroui, Ultrasound image segmentation: A survey, IEEE Transactions on Medical Imaging 25 (2006), pp. 9871010.

[11] K. Ogawa, M. Fukushima, K. Kubota and N. Hisa, Computer-aided diagnostic system for diffuse liver-diseases with ultrasonography by neural networks, *IEEE Transactions on Nuclear Science*, 45 (6) 3069-3074 (1998).

[12] M. Riedmiller, H. Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco,1993

[13] F.Sahba, H.R. Tizhoosh, M.M. Salama, A coarse-to-fine approach to prostate boundary segmentation in ultrasound images, BioMedical Engineering OnLine 2005, 4:58

[14] H.R. Tizhoosh, Observer-Dependent Image Enhancement, Proceeding of FUZZ-IEEE 2001.

[15] H. R. Tizhoosh. Image thresholding using type II fuzzy sets, *Pattern recognition*, 38 (12), pages 2363-2372, 2005.

[16] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning – I, *Information Science*, **8**, 199–249, 1975.

[17] W. Zhang, H. Yoshida, R.M. Nishikawa and K. Doi, Optimally weighted wavelet transform based on supervised training for detection of microcalcifications in digital mammograms, Medical Physics 25 (6) (1998), pp. 949–956.

Umbralización de imágenes utilizando conjuntos intervalo-valorados difusos. Importancia del algoritmo de Tizhoosh

H. Bustince, J.Sanz y J.Olagoitia
Dpt. Automática y Computación
Universidad Pública de Navarra,
Campus Arrosadía s/n, 31106 Pamplona. España,
bustince@unavarra.es

Resumen

En este trabajo desarrollamos algunas re- exiones sobre el algoritmo de umbralización propuesto por Tizhoosh en [17]. El objeto de estas re exiones es completar las consideraciones publicadas recientemente en [18] y [19] sobre dicho algoritmo. También demostramos que bajo determinadas construcciones, el algoritmo de Tizhoosh permite obtener información adicional a la obtenida con los algoritmos difusos habitualmente utilizados.

Palabras Clave: Umbralización de imágenes, Conjunto difuso tipo II, Conjunto intervalo valorado difuso, Entropía intervalo valorada difusa, Entropía difusa.

1 Introducción

Desde nuestro punto de vista, la principal aportación del artículo de Tizhoosh (ver [17]): "Image thresholding using type-2 fuzzy sets" es la *utilización de conjuntos intervalo-valorados difusos en un algortimo de umbralización de imágenes*. Dichos conjuntos, permiten modelizar el desconocimiento que tiene el experto a la hora de abordar la construcción de la función de pertenencia del conjunto difuso que mejor representa a la imagen (ver [2, 10, 11]).

Creemos que la aportación anteriormente mencionada es tan importante que las siguientes consideraciones:

- El autor no relaciona el concepto de ultrafusines con el de IV entropía (ver [3]) presentado 10 años antes.
- En el título del artículo se utiliza el concepto de conjunto difuso tipo II (ver [21]). Sin embargo en dicho trabajo sólo se emplean conjuntos intervalo-valorados difusos, los cuales son un caso particular de los conjuntos difusos tipo II, (ver [12, 13]);

son detalles menores. No obstante, debemos agradecer a Vlachos y Sergiadis las apreciaciones que hacen en [19] sobre este trabajo. También debemos destacar la precisión y el rigor que utilizan para relacionar los conceptos presentados por Tizhoosh con otros ya existentes en la literatura.

Respecto a la relación entre los conjuntos intervalo-valorados difusos y los conjuntos intuicionistas de Atanassov (ver [1]), consideramos que aunque matemáticamente estos conceptos son equivalentes, al aplicarlos al procesamiento de imagen en el sentido propuesto por Tizhoosh, su interpretación es totalmente diferente. El motivo es el siguiente: Tizhoosh utiliza los conjuntos intervalo-valorados difusos para modelizar la imprecisión o desconocimiento que tiene el experto en la elección de la función de pertenencia asociada a la imagen (ver Fig. 1). Sin embargo, cuando se utilizan los conjuntos de Atanassov se habla de la función de pertenencia y de la función de no pertenencia, pero nunca de la imprecisión asociada a la elección de cada una de las funciones (pertenencia y no pertenencia). En el mejor de los casos, el índice intuicionista de Atanassov podría interpretarse como una medida de la imprecisión en la elección de ambas funciones a la vez (ver [14]).

A parte de las consideraciones anteriores, en este trabajo nos proponemos demostrar que *bajo ciertas construcciones*, el algoritmo para la umbralización de imágenes utilizando los conjuntos intervalo-valorados, permite:

1. Seleccionar el mejor umbral entre dos umbrales obtenidos mediante el algoritmo de las áreas (ver [6]), el cual es una versión particular del algoritmo de Huang and Wang propuesto en 1995 (ver [11]);
2. Representar los píxeles de la imagen para los

cuales los expertos, que utilizan el algoritmo de las áreas, no se ponen de acuerdo respecto a su pertenencia al fondo o al objeto.

2 Preliminares

Definición 1 *Un conjunto difuso $\widetilde{A}$ definido sobre un universo finito y no vacío $U = \{u_1, \cdots, u_n\}$ viene dado por (ver [20]):*

$$\widetilde{A} = \{(u_i, \mu_{\widetilde{A}}(u_i))|u_i \in U\}$$

donde $\mu_{\widetilde{A}} : U \to [0,1]$ es la función de pertenencia.

En este trabajo vamos a denotar como $\mathcal{FS}(U)$ al conjunto de todos los conjuntos difusos definidos sobre un referencial finito U.

En 1972 Deluca y Termini [9] formalizaron las propiedades de la entropía difusa a través de los siguientes axiomas:

Definición 2 *Una función real $E : \mathcal{FS}(U) \to [0,1]$ es una entropía en $\mathcal{FS}(U)$, si E tiene las siguientes propiedades:*

$(E1)$ $E(\widetilde{A}) = 0$ *si y sólo si $\widetilde{A}$ no es difuso;*

$(E2)$ $E(\widetilde{A}) = 1$ *si y sólo si* $\widetilde{A} = \{(u_i, \mu_{\widetilde{A}}(u_i) = 0.5)|u_i \in U\}$;

$(E3)$ $E(\widetilde{A}) \leq E(\widetilde{B})$ *si $\widetilde{A}$ refina $\widetilde{B}$; esto es, $\mu_{\widetilde{A}}(u_i) \leq \mu_{\widetilde{B}}(u_i)$ cuando $\mu_{\widetilde{B}}(u_i) \leq 0.5$ y $\mu_{\widetilde{A}}(u_i)$ $\mu_{\widetilde{B}}(u_i)$ cuando $\mu_{\widetilde{B}}(u_i)$ 0.5;*

$(E4)$ $E(\widetilde{A}) = E(\widetilde{A}_N)$, *donde* $\widetilde{A}_N = \{(u_i, 1 - \mu_{\widetilde{A}}(u_i))|u_i \in U\}$.

Definición 3 *Sea $L([0,1]) = \{[M_L, M_U]|(M_L, M_U) \in [0,1]^2$ y $M_L \leq M_U\}$ el conjunto de todos los subintervalos cerrados del intervalo cerrado $[0,1]$. Un conjunto intervalo-valorado difuso (IVFS) A definido sobre el universo U viene dado por:*

$$A = \{(u_i, M_A(u_i))|u_i \in U\}$$

donde la función $M_A : U \to L([0,1])$ define el grado de pertenencia de un elemento $u_i \in U$ a A.

Vamos a denotar con $\mathcal{IVFS}(U)$ el conjunto de todos los conjuntos intervalo-valorados difusos definidos sobre un referencial finito U.

En la literatura hay dos definiciones diferentes del concepto de entropía intervalo-valorada difusa (IV entropía). La primera, $\mathcal{E}_F$, fue presentada en 1996 en [3], y la segunda, $\mathcal{E}_c$, en 2001 en[16]. La diferencia entre ambos surge de que $\mathcal{E}_F$ es una medida de lo lejos que está un IVFS de un conjunto difuso, mientras que $\mathcal{E}_c$ es una medida de lo lejos que un IVFS está de un conjunto clásico. Por lo tanto, $\mathcal{E}_F$ está basada en las ideas de Sambuc (ver [15]) y $\mathcal{E}_c$ está basada en el concepto de entropía difusa de Deluca y Termini. En este trabajo sólo vamos a utilizar $\mathcal{E}_F$.

Definición 4 *Una función real $\mathcal{E}_F : \mathcal{IVFS}(U) \to \mathbb{R}^+$ es una entropía en $\mathcal{IVFS}(U)$ si tiene las siguientes propiedades:*

$(IF1)$ $\mathcal{E}_F(A) = 0$ *si y sólo si $A \in \mathcal{FS}(U)$;*

$(IF2)$ $\mathcal{E}_F(A) = 1$ *si y sólo si* $M_{AL}(u_i) = 0$ *y* $M_{AU}(u_i) = 1$ *para todo $u_i \in U$;*

$(IF3)$ $\mathcal{E}_F(A) = \mathcal{E}_F(A_N)$ *para todo $A \in \mathcal{IVFS}(U)$, donde* $A_N = \{(u_i, M_{A_N} = [1 - M_{AU}(u_i), 1 - M_{AL}(u_i)])|u_i \in U\}$;

$(IF4)$ *si $A \preceq B$; esto es, si $M_{AL}(u_i) \leq M_{BL}(u_i)$ y $M_{BU}(u_i) \leq M_{AU}(u_i)$ para todo $u_i \in U$, entonces* $\mathcal{E}_F(A)$ $\mathcal{E}_F(B)$.

En [3], estas entropías $\mathcal{E}_F$ están estudiadas en profundidad y se presenta un teorema para su construcción a partir de funciones $f : [0,1] \to [0,1]$. También se dice que la expresión más utilizada es la obtenida cuando tomamos $f(x) = x$, es decir

$$\mathcal{E}_F(A) = \frac{1}{n} \sum_{i=1}^{n} M_U(u_i) - M_L(u_i) \qquad (1)$$

para todo $A \in \mathcal{IVFS}$s(U). Es fácil comprobar que (1) es el *índice de indeterminación* introducido por Sambuc en 1975.

Vlachos y Sergiadis demuestran en [19] que el concepto de medida de ultrafusines que Tizhoosh utiliza en su algoritmo, es el concepto de entropía intervalo-valorada difusa $\mathcal{E}_F$ presentada en la definición 4. Además, la expresión utilizada en [17] para calcular el grado de ultrafusines coincide con el índice de indeterminación de Sambuc; es decir, con (1).

Debemos hacer notar que para nosotros, el mejor umbral es el umbral asociado al conjunto difuso o al conjunto intervalo-valorado difuso con menor entropía (difusa o intervalo-valorada difusa según sea el caso en estudio). Además, en los algoritmos y en las construcciones que presentamos en la próxima sección, vamos a utilizar el concepto de automorfismo (ver [5]).

Un automorfismo sobre el intervalo unidad es una función $\varphi : [0,1] \to [0,1]$ continua, estrictamente creciente tal que $\varphi(0) = 0, \varphi(1) = 1$.

3 Tres algoritmos fuzzy para umbralizar imágenes

En el marco de la teoría de conjuntos difusos los algoritmos más populares para calcular el umbral asociado a una imagen Q (dada en una escala con L niveles de gris) son aquellos que usan el concepto de entropía difusa [2, 10, 11]. En este sentido el más utilizado es el siguiente:

(a) Asignar L conjuntos difusos $\widetilde{Q}_t$ a cada imagen Q. Cada uno esta asociado a un nivel de intensidad t, $(t = 0, 1, \cdots, L-1)$, de la escala de grises L usada.

(b) Calcular la entropía difusa de cada uno de los L conjuntos difusos $\widetilde{Q}_t$ asociados a Q.

(c) Elegir, como el *mejor umbral* de nivel de gris, el t asociado al conjunto difuso con menor entropía.

Algoritmo 1

La razón por la cual en el Algoritmo 1 utilizamos la entropía difusa es la siguiente:

Umbralizar una imagen es el proceso de encontrar la intensidad $t \in \{1, \cdots, L-1\}$ tal que al representar con uno (cero) los píxeles con intensidades mayores o iguales que t y con cero (uno) los demás, obtengamos la imagen binaria que mejor separe el fondo del objeto. Por este motivo, de entre todos los conjuntos difusos construidos en el Algoritmo 1 (paso (a)), tomamos el que tiene menor entropía difusa, es decir, el que tiene valores de la función de pertenecia más próximos al cero o al uno.

El principal problema del Algoritmo 1 es el paso (a). Esto es debido a la dificultad de elegir la función de pertenencia que debemos tomar para construir el conjunto difuso que mejor represente a la imagen [17]. Este problema ha llevado a muchos autores a proponer diferentes versiones del Algoritmo 1. El *Algoritmo de las áreas* (ver [6]), que mostramos a continuación, es una de estas versiones. Dicha versión, ha sido construida utilizando las relaciones entre las medidas de información estudiadas en [7].

Evidentemente, el Algoritmo de las áreas es más sencillo y eficiente en tiempo que Algoritmo 1. Nótese que $A(\widetilde{Q}_t, \varphi_1, \varphi_2)$ crece a medida que q es más próximo a $m_b(t)$ o a $m_o(t)$. ($h(q)$ es el número de píxeles de la imagen con intensidad q).

(a) Seleccionar dos automorfismos φ_1 y φ_2.
(b) FOR $t := 0$ TO $L-1$ DO
(b-1) Calcular:

$$m_b(t) = \frac{\sum_{q=0}^{t} qh(q)}{\sum_{q=0}^{t} h(q)} \text{ y } m_o(t) = \frac{\sum_{q=t+1}^{L-1} qh(q)}{\sum_{q=t+1}^{L-1} h(q)}$$

(b-2) Calcular el área:

$$A(\widetilde{Q}_t, \varphi_1, \varphi_2) =$$
$$\sum_{q=0}^{t} h(q)\varphi_1^{-1}\left(1 - 0.5|\varphi_2(\frac{q}{L-1}) - \varphi_2(\frac{m_b(t)}{L-1})|\right) +$$
$$\sum_{q=t+1}^{L-1} h(q)\varphi_1^{-1}\left(1 - 0.5|\varphi_2(\frac{q}{L-1}) - \varphi_2(\frac{m_o(t)}{L-1})|\right).$$

ENDFOR
(c) Elegir como mejor umbral el valor de t el asociado con el mayor valor de $A(\widetilde{Q}_t, \varphi_1, \varphi_2)$.

Algoritmo 2 (Algoritmo de las áreas)

El algoritmo que presentamos a continuación es el propuesto por Tizhoosh utilizando $\mathcal{IVFS}$s:

(a) Asignar L conjuntos difusos $\widetilde{Q}_t$ (con $t \in \{0, \cdots, L-1\}$) a cada imagen Q.
(b) Coger $\alpha \in (1, \infty)$.
(c) Asociar a cada conjunto difuso $\widetilde{Q}_t$ su correspondiente IVFS Q_t construido con el siguiente método (ver [4]):

$Q_t = \{(q, M_{Q_t}(q)) | q \in \{0, \cdots, L-1\}\}$ dado por
$M_{Q_t}(q) = [\mu^{\alpha}_{\widetilde{Q}_t}(q), \mu^{\frac{1}{\alpha}}_{\widetilde{Q}_t}(q)]$, (ver Fig. 1).

(d) Para cada IVFS Q_t calcular $\mathcal{E}_F$:

$$\mathcal{E}_F(Q_t) = \frac{1}{\sum_{q=0}^{L-1} h(q)} \sum_{q=0}^{L-1} h(q)\left(\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_t}(q) - \mu^{\alpha}_{\widetilde{Q}_t}(q)\right)$$

(e) Coger como umbral la intensidad t asociada al IVFS Q_t con el menor valor de $\mathcal{E}_F$.

Algoritmo 3 (Algoritmo de Tizhoosh)

En el Algoritmo 3 (paso (a)) vamos a utilizar el siguiente método para construir los conjuntos difusos $\widetilde{Q}_t$:

(i) Seleccionar el automorfismo φ;
(i) FOR t:=0 TO L-1 DO
Calcular: $m_b(t)$ y $m_o(t)$
FOR q:=0 TO L-1 DO

$$\mu_{\widetilde{Q}_t}(q) = \begin{cases} 1 - 0.5|\varphi(\frac{q}{L-1}) - \varphi(\frac{m_b(t)}{L-1})|) \text{ if } q \leq t \\ 1 - 0.5|\varphi(\frac{q}{L-1}) - \varphi(\frac{m_o(t)}{L-1})|) \text{ if } q > t, \end{cases} \quad (2)$$

ENDFOR

ENDFOR

El motivo por el cual utilizamos esta construcción es el siguiente: Si $\mathcal{E}_F(Q_t) \to 0$, entonces $\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_t}(q) \approx \mu^{\alpha}_{\widetilde{Q}_t}(q)$ para todo $q \in \{1, \cdots, L-1\}$, teniendo en cuenta que las funciones de pertenencia construidas en (2) son tales que: $0.5 \leq \mu_{\widetilde{Q}_t}(q) \leq 1$ (ver [10]) tenemos que $\mu_{\widetilde{Q}_t}(q) \to 1$ para todo $q \in \{1, \cdots, L-1\}$, por lo tanto $q \approx m_b(t)$ o $q \approx m_o(t)$.

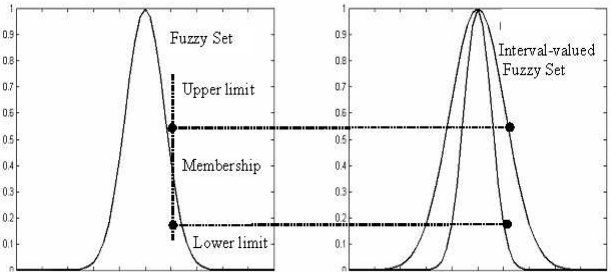


Figura 1: Método de construcción de un IVFS a partir de un conjunto difuso

4 Relación entre los algoritmos

En [6, 7] está probado que el algoritmo de las áreas es un caso particular del Algoritmo 1. Más aún, también está demostrado que el método de Otsu es un caso particular del Algoritmo de las áreas.

Respecto a la relación existente entre el Algoritmo de las áreas y el Algoritmo de Tizhoosh tenemos los siguientes resultados:

Proposición 1 *Sea* $\alpha \quad 1$*, para todo* $t \in \{1, \cdots, L-1\}$ *se cumple la siguiente desigualdad:*

$$A(\widetilde{Q}_t, x^\alpha, x) = \sum_{q=0}^{t} h(q)\left(1 - 0.5|\frac{q}{L-1} - \frac{m_b(t)}{L-1}|\right)^{\frac{1}{\alpha}} + \sum_{q=t+1}^{L-1} h(q)\left(1 - 0.5|\frac{q}{L-1} - \frac{m_o(t)}{L-1}|\right)^{\frac{1}{\alpha}}$$

$$A(\widetilde{Q}_t, x^{\frac{1}{\alpha}}, x) = \sum_{q=0}^{t} h(q)\left(1 - 0.5|\frac{q}{L-1} - \frac{m_b(t)}{L-1}|\right)^{\alpha} + \sum_{q=t+1}^{L-1} h(q)\left(1 - 0.5|\frac{q}{L-1} - \frac{m_o(t)}{L-1}|\right)^{\alpha}.$$

Dem. Directa, basta tener en cuenta que $x^{\frac{1}{\alpha}} \quad x^\alpha$ para todo $x \in [0,1]$ con $\alpha \quad 1$. ■

Teorema 1 *Si en el paso* (a) *del algoritmo de Tizhoosh utilizamos las construcciones presentadas en* (2) *con* $\varphi(x) = x$ *para todo* $x \in [0,1]$*, entonces*

$$\mathcal{E}_F(Q_t) = \frac{1}{\sum_{q=0}^{L-1} h(q)}\Big(A(\widetilde{Q}_t, x^\alpha, x) - A(\widetilde{Q}_t, x^{\frac{1}{\alpha}}, x)\Big).$$

Dem. Directa. ■

Teorema 2 *Sea* $\alpha \quad 1$ *y sea* $\varphi_2(x) = x$*. Sean* $t_{\frac{1}{\alpha}}$ *y* t_α *los umbrales calculados con el algoritmo de las áreas tomando* $\varphi_1(x) = x^\alpha$ *y* $\varphi_1(x) = x^{\frac{1}{\alpha}}$ *respectivamente. Sean los conjuntos difusos* $\widetilde{Q}_{t_{\frac{1}{\alpha}}}$ *y* $\widetilde{Q}_{t_\alpha}$ *construidos con el método* (2) *tomando* $\varphi(x) = x$*. Bajo estas condiciones:*

$$\mathcal{E}_F(Q_{t_{\frac{1}{\alpha}}}) \quad \mathcal{E}_F(Q_{t_\alpha})$$

donde

$Q_{t_{\frac{1}{\alpha}}} = \{(q, [\mu^{\alpha}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q), \mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q)]) | q \in \{0, \cdots, L-1\}\}$ *y*

$Q_{t_\alpha} = \{(q, [\mu^{\alpha}_{\widetilde{Q}_{t_\alpha}}(q), \mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_\alpha}}(q)]) | q \in \{0, \cdots, L-1\}\}.$

Dem. $t_{\frac{1}{\alpha}}$ es el umbral calculado mediante el Algoritmo de las áreas, por tanto para todo $t \in \{0, 1, \cdots, L-1\}$ tenemos:

$$A(\widetilde{Q}_{t_{\frac{1}{\alpha}}}, x^\alpha, x) = \sum_{q=0}^{L-1} h(q)\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q) \quad \sum_{q=0}^{L-1} h(q)\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_t}(q)$$

y en particular, para $t = t_\alpha$ tenemos:

$$A(\widetilde{Q}_{t_{\frac{1}{\alpha}}}, x^\alpha, x) = \sum_{q=0}^{L-1} h(q)\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q) \quad \sum_{q=0}^{L-1} h(q)\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_\alpha}}(q)$$

Por el mismo razonamiento, para t_α tenemos:

$$A(\widetilde{Q}_{t_\alpha}, x^{\frac{1}{\alpha}}, x) = \sum_{q=0}^{L-1} h(q)\mu^{\alpha}_{\widetilde{Q}_{t_\alpha}}(q) \quad \sum_{q=0}^{L-1} h(q)\mu^{\alpha}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q)$$

Por lo tanto

$$\mathcal{E}_F(Q_{t_{\frac{1}{\alpha}}}) - \mathcal{E}_F(Q_{t_\alpha}) =$$

$$\sum_{q=0}^{L-1} h(q)\Big(\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q) - \mu^{\alpha}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q)\Big) -$$

$$\sum_{q=0}^{L-1} h(q)\Big(\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_\alpha}}(q) - \mu^{\alpha}_{\widetilde{Q}_{t_\alpha}}(q)\Big) =$$

$$\sum_{q=0}^{L-1} h(q)\Big(\mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q) - \mu^{\frac{1}{\alpha}}_{\widetilde{Q}_{t_\alpha}}(q)\Big) +$$

$$\sum_{q=0}^{L-1} h(q)\Big(\mu^{\alpha}_{\widetilde{Q}_{t_\alpha}}(q) - \mu^{\alpha}_{\widetilde{Q}_{t_{\frac{1}{\alpha}}}}(q)\Big) \quad 0.\blacksquare$$

Una consecuencia del Teorema 2 es la siguiente:

Si en la ejecución del algoritmo de Tizhoosh utilizamos las construcciones (2) con $\varphi = \varphi_2$, siendo φ_2 el automorfismo utilizado en el algoritmo de las áreas para calcular $t_{\frac{1}{\alpha}}$ y t_α, entonces el algoritmo de Tizhoosh nos dice que el mejor umbral entre esos dos es el que nos permite construir un IVFS con menor IV entropía, es decir, t_α.

Corolario 1 *Bajo las condiciones del Teorema 2, si t_H es el umbral calculado con el Algoritmo de Tizhoosh, entonces:*

$$\mathcal{E}_F(Q_{t_H}) \leq \mathcal{E}_F(Q_{t_\alpha}) \leq \mathcal{E}_F(Q_{t_{\frac{1}{\alpha}}})$$

Dem. Directa. ■

Del Corolario 1 deducimos que el algoritmo de Tizhoosh siempre nos proporciona un umbral mejor que los obtenidos con el algoritmo de las áreas cuando lo aplicamos primero a los extremos inferiores y luego a los superiores de los intervalos de los conjutos IVFSs construidos con nuestros métodos.

Además, si "dibujamos con uno (cero) los píxeles cuya intensidad se encuentra entre $t_{\frac{1}{\alpha}}$ y t_α y con cero (uno) el resto", entonces queda claro que para uno de los umbrales los píxeles representados son del fondo (objeto) y para el otro umbral son del objeto (fondo); esto es, son los píxeles para los cuales el experto que ha proporcionado el umbral $t_{\frac{1}{\alpha}}$ y el que ha proporcionado el umbral t_α no se ponen de acuerdo.

Ejemplo 1 *Las imágenes de este ejemplo son una selección de las utilizadas por Tizhoosh en su artículo. Por este motivo, no mostramos la correspondiente imagen umbralizada.*

Tomamos $\alpha = 2$. En la primera fila de la Tabla 1 presentamos los nombres de las imágenes objeto de estudio. En las restantes filas mostramos los valores de $t_{\frac{1}{\alpha}}$, $\mathcal{E}_F(Q_{t_{\frac{1}{\alpha}}})$, t_α, $\mathcal{E}_F(Q_{t_\alpha})$, t_H y $\mathcal{E}_F(Q_{t_H})$ respectivamente. Teniendo en cuenta el Teorema 2 y el Corolario 1, el algoritmo de Tizhoosh nos recomienda utilizar, por ejemplo para los bloques, en primer lugar t_H, y entre $t_{\frac{1}{\alpha}}$ y t_α el umbral $t_\alpha = 43$. Siguiendo este razonamiento, para la imagen del arroz debemos tomar el umbral de 51 en primer lugar y en cualquier otro caso 52, etc.

En la figura 2 se encuentran las imágenes originales (Blocks, Rice, Wheel, Newspaper). A continuación presentamos, para cada imagen, el porcentaje de píxeles entre los correspondientes valores de $t_{\frac{1}{\alpha}}$ y t_α; es decir, el porcentaje de píxeles tales que para un umbral son del fondo y para el otro del objeto. Por último dibujamos en blanco dichos píxeles.

$IMAGE$	Blocks	Rice	Wheel	Newspaper
$t_{\frac{1}{\alpha}}$	48	55	86	199
$\mathcal{E}_\mathcal{F}(Q_{t_{\frac{1}{\alpha}}})$	0.04771	0.04218	0.04001	0.03647
t_α	43	52	81	201
$\mathcal{E}_\mathcal{F}(Q_{t_\alpha})$	0.04765	0.04210	0.03996	0.03643
t_H	42	51	79	201
$\mathcal{E}_\mathcal{F}(Q_{t_H})$	0.04764	0.04210	0.03995	0.03643

Tabla 1: Umbrales y IV entropías

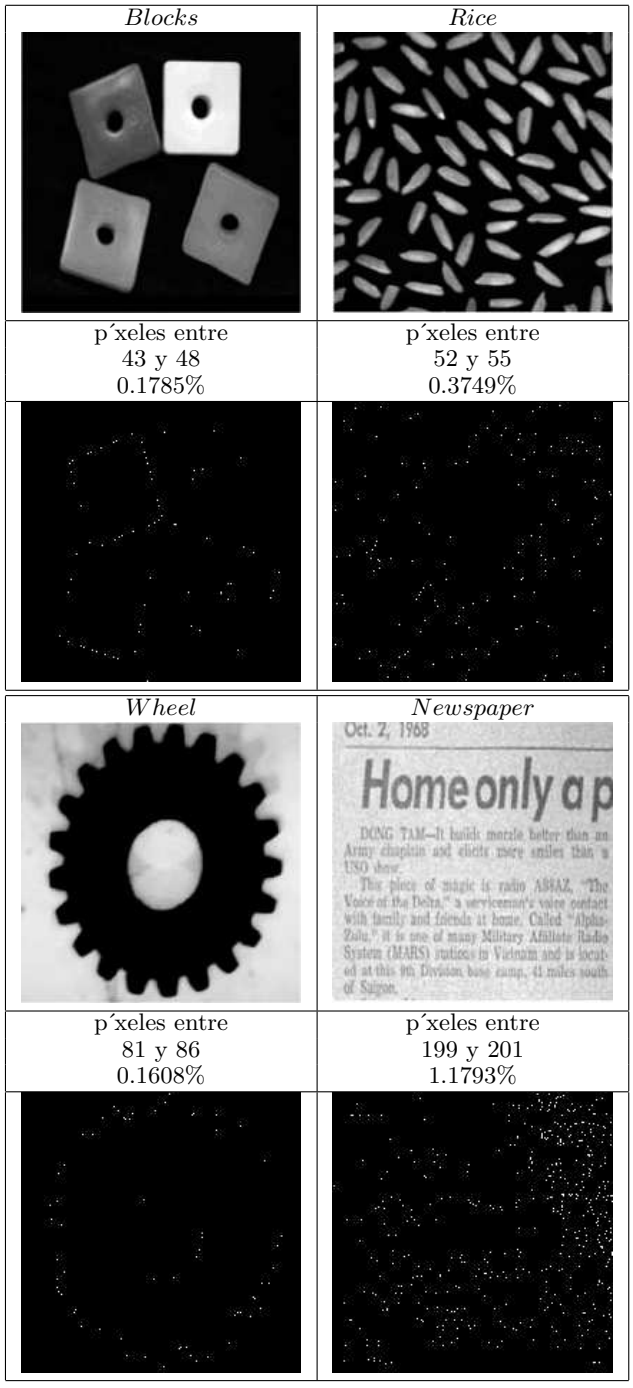


Figura 2: Imágenes originales y Píxeles Dudosos

5 Conclusiones

En este trabajo probamos que cuando utilizamos las mismas construcciones (2) para ejecutar el Algoritmo de las áreas (y por tanto el algoritmo 1) y el Algoritmo de Tizhoosh, éste último nos permite seleccionar el mejor umbral entre los calculados mediante el algoritmo de las áreas al aplicarlo a los extremos inferiores y superiores de los IVFSs construidos. Además el umbral calculado mediante el algoritmo de Tizhoosh es siempre el mejor, es decir, el de menor IV entropía.

Consideramos que estos argumentos justifican el empleo del algoritmo de Tizhoosh en procesamiento de imagen y por tanto, también justifican el empleo de los conjuntos intervalo-valorados difusos en el cálculo del umbral de una imagen.

En un futuro próximo tenemos previsto analizar la relación existente entre los algoritmos presentados en este artículo cuando utilicemos cualquier tipo de construcción, no sólo para la (2) y proponer nuevas definiciones de IV entropía fuzzy utilizando las ideas propuestas en [8].

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el proyecto de referencia: TIN2007-65981.

Referencias

[1] K. Atanassov, Intuitionistic fuzzy sets, In: VIIth ITKR Session, Deposited in the Central Science and Technology Library of the Bulgarian Academy of Sciences, Sofia, Bulgaria, Pág. 1684-1697, 1983.

[2] J.C. Bezdek, J. Keller, R. Krisnapuram, N. R. Pal, *Fuzzy Models and algorithms for pattern recognition and image processing*, in The Handbooks of Fuzzy Sets Series, Series Editors: D. Dubois and H. Prade, Kluwer Academic Publishers, Boston/London/Dordrecht, 1999.

[3] P. Burillo, H. Bustince, Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets, Fuzzy Sets and Systems 78, Pág. 305-316, 1996.

[4] P. Burillo, H. Bustince, Construction theorems for intuitionistic fuzzy sets, Fuzzy Sets and Systems 84, Pág. 271-281, 1996.

[5] H. Bustince, P. Burillo and F. Soria, Automorphisms, Negations and Implication Operators, *Fuzzy Sets and Systems* 134, Pág. 209–229, 2003.

[6] H. Bustince, E. Barrenechea, M. Pagola, Image thresholding using restricted equivalence functions and maximizing the measures of similarity, *Fuzzy Sets and Systems* 158(5), Pág. 496-516, 2003.

[7] H. Bustince, E. Barrenechea, M. Pagola, Relationship between restricted dissimilarity functions, restricted equivalence functions and normal EN-functions: Image thresholding invariant, *Pattern Recognition Letters* 29(4), Pág. 525-536, 2008.

[8] M. Delgado, M.J. Martín-Bautista, D. Sánchez, M.A. Vila, A Probabilistic Definition of a Nonconvex Fuzzy Cardinality, *Fuzzy Sets and Systems* 126(2), Pág. 177-190, 2002.

[9] A. Deluca, S. Termini, A definition of a nonprobabilistic entropy in the setting of fuzzy set theory, *Inform and Control*, **20** Pág. 301-312, 1972.

[10] M.G. Forero, *Fuzzy thresholding and histogram analysis*, in Fuzzy Filters for Image Processing, Edited by M. Nachtegael, D. Van der Weken, D. Van de Ville and E.E. Kerre, Springer, Pág. 129-152, 2003.

[11] L.K. Huang, M.J. Wang, Image thresholding by minimizing the measure of fuzziness *Pattern recognition*, **28(1)**, Pág. 41-51, 1995.

[12] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, Upper Saddle River, NJ, 1995.

[13] J.M. Mendel, Advances in type-2 fuzzy sets and systems, Information Sciences 177, Pág. 84-110, 2007.

[14] J. Montero, D. Gómez, H. Bustince, On the relevance of some families of fuzzy sets, *Fuzzy Sets and Systems* 158(22), Pág. 2429-2442,2007.

[15] R. Sambuc, Function Φ-Flous, Application a l'aide au Diagnostic en Pathologie Thyroidienne, These de Doctorat en Medicine, University of Marseille, 1975.

[16] E. Szmidt, J. Kacprzyk, Entropy for intuitionistic fuzzy sets, Fuzzy Sets and Systems 118(3), Pág. 467-477, 2001.

[17] H.R. Tizhoosh, Image thresholding using type-2 fuzzy sets, Pattern Recognition 38, Pág. 2363-2372, 2005.

[18] H.R. Tizhoosh, Interval-valued versus intuitionistic fuzzy sets: Isomorphism versus semantics, Pattern Recognition 41(5), Pág. 1812-1813, 2008.

[19] I. K. Vlachos, G. D. Sergiadis, Comment on: Image thresholding using type II fuzzy sets, Pattern Recognition 41 (5), Pág. 1810-1811, 2008.

[20] L.A. Zadeh, Fuzzy sets, Information Control 8, Pág. 338-353, 1965.

[21] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning – I, Information Sciences 8, Pág. 199-249, 1975.

SUPERPOSICIÓN CRANEOFACIAL BASADA EN ALGORITMOS GENÉTICOS Y LOCALIZACIÓN DIFUSA DE PUNTOS DE REFERENCIA CEFALOMÉTRICOS

Óscar Ibáñez [1] **Óscar Cordón** [1] **Sergio Damas** [1] **José Santamaría** [2]

[1] European Centre for Soft Computing, C/ Gonzalo Guitiérrez Quirós s.n, Mieres 33600 Asturias, {oscar.ibanez, oscar.cordon, sergio.damas} @ softcomputing.es
[2] Dpto. de Informática, Universidad de Jaén. C/ Paraje Las Lagunillas, s/n, Edificio Tecnológico y de Ingenierías A-3, Jaén 23071, jslopez@ujaen.es

Resumen

La superposición craneofacial es la segunda tarea de un proceso forense complejo, que tiene como objetivo la identificación de una persona a partir de una fotografía y del cráneo encontrado. Concretamente, se trata de encontrar la posición más adecuada del cráneo para proyectarlo sobre la fotografía. El proceso se guía por una serie de puntos de referencia identificados tanto en el cráneo (puntos craneométricos) como en la fotografía (puntos cefalométricos). En este artículo hemos ampliado nuestra anterior aproximación al problema basada en algoritmos genéticos, considerando la incertidumbre presente en la localización de los puntos cefalométricos. Los resultados que presentamos corresponden a dos casos reales, resueltos por el laboratorio de Antropología Física de la Universidad de Granada.

Palabras Clave: Superposición craneofacial, algoritmos genéticos, puntos de referencia difusos, registrado de imágenes.

1 INTRODUCCIÓN

La supra-proyección fotográfica [8] es un proceso forense en el que fotografías de una persona desaparecida son comparadas con el cráneo encontrado. Proyectando una sobre la otra (o, incluso mejor, haciendo corresponder un modelo tridimensional escaneado del cráneo, con la foto de la cara), el forense trata de determinar si se trata de la misma persona.

Para llevar a cabo este proceso, primero se necesita un modelo 3D preciso del cráneo. A continuación se deben identificar dos conjuntos de puntos radiométricos: puntos antropométricos de la cara, en la fotografía, y puntos antropométricos del cráneo, en el modelo 3D. Estos dos conjuntos serán los que guíen el proceso de emparejamiento del modelo 3D del cráneo y la fotografía [8]. Finalmente, en una etapa de toma de decisiones, se analizan los emparejamientos logrados entre los distintos pares de puntos de referencia. Algunos de ellos emparejarán perfectamente, otros lo harán de manera parcial, y otros no lo harán. Al final de este proceso, el experto forense debe decidir si el cráneo analizado corresponde o no a la persona desaparecida de la foto. Como se puede observar, el anterior procedimiento es muy costoso en tiempo. Además, no hay una metodología sistemática, cada experto aplica su propio método. Por lo tanto, hay un gran interés en diseñar métodos automáticos que faciliten la tarea a los antropólogos forenses.

En un trabajo previo [2], abordamos la segunda etapa del proceso, (la superposición craneofacial) usando un algoritmo genético (AG) con codificación real. Entonces asumimos que el proceso estaba exento de incertidumbre. Sin embargo, podemos identificar dos fuentes de incertidumbre en el problema de la superposición craneofacial.

Por un lado, llamaremos *incertidumbre en la localización* a la derivada de la extrema complejidad de la tarea de la localización de los puntos de referencia con la precisión requerida para este proceso. Por otro lado, la *incertidumbre en el emparejamiento* se referirá a la incertidumbre presente en el emparejamiento de los puntos de referencia correspondientes a los dos objetos diferentes: la fotografía y el cráneo (como ya dijimos, hay una situación de claro emparejamiento parcial).

Esta contribución tiene como objetivo extender nuestro trabajo [2] con una primera aproximación para abordar la incertidumbre en la localización mediante la definición de diferentes áreas difusas donde el experto debería localizar cada punto de referencia cefalométrico. El sistema genético difuso resultante [5] se prueba en dos problemas de superposición de casos reales de identificación ya

resueltos por el laboratorio de Antropología Física de la Universidad de Granada.

El artículo se ha estructurado de la siguiente forma. La Sección 2 repasa nuestra propuesta previa referida a la superposición craneofacial. Las fuentes de incertidumbre existentes en el problema se detallan en la Sección 3. Nuestra propuesta se describe en la Sección 4. Después, la Sección 5 presenta el estudio experimental y, por ultimo, la Sección 6 recoge las conclusiones y los trabajos futuros.

2 ALGORITMO GENÉTICO DE CODIFICACIÓN REAL PARA LA SUPERPOSICIÓN CRANEOFACIAL

En [3] analizamos el estado del arte en el problema de la superposición craneofacial, destacando que había pocos trabajos que llevaran a cabo esta tarea de manera automática. El método propuesto por Nickerson et al.'s [9] usa un AG para encontrar los parámetros óptimos de las transformaciones de perspectiva y similitud, que superpongan el modelo 3D del cráneo sobre la fotografía de la cara. Debido a la falta de espacio, no podemos describir esas transformaciones geométricas. Los lectores interesados pueden consultar [2]. Sin embargo, ninguno de los trabajos previos logró resultados aceptables con nuestros datos (ni con los sintéticos, ni con los reales). Por lo tanto, propusimos un nuevo diseño para el AG considerando la propuesta de Nickerson et al.'s cómo punto de partida.

El problema que estamos abordando, se basa en encontrar la mejor transformación perspectiva que, aplicada sobre el modelo 3D del cráneo, lo proyecta adecuadamente sobre la fotografía 2D. La búsqueda de la mejor transformación se guía por una serie de pares de puntos craneométricos y cefalométricos localizados en el cráneo y en la cara, respectivamente. Esta transformación se define mediante un conjunto de parámetros reales. Por lo tanto, nuestra primera mejora con respecto a la propuesta de Nickerson fue usar un AG con codificación real en lugar de uno de codificación binaria. En nuestro caso, los parámetros de la transformación se codificaron en un cromosoma compuesto por 12 genes reales. Tras el estudio de diferentes definiciones de la función objetivo, la que mejores resultados consiguió fue el error cuadrático medio:

$$ECM = \frac{\sum_{i=1}^{N} \left\| f(Ci) - Fi \right\|^2}{N} \quad (1)$$

donde ||·|| es la distancia Euclídea 2D, N es el número de puntos de referencia considerados (sólo cuatro en la propuesta de Nickerson), Ci se corresponde a cada punto de referencia 3D, Fi se refiere a cada punto de referencia 2D de la foto y f(Ci) representa la posición de la transformación de cada punto 3D del cráneo en el plano proyectado.

Además, hemos dotado al AG de operadores genéticos más eficientes. De esta manera, usamos la selección mediante torneo [4], el cruce BLX-α [6] y la mutación aleatoria [1]. En la selección mediante torneo, t individuos (t = 2 suele ser la elección más común) se seleccionan aleatoriamente de la población, de los que el mejor se inserta en la nueva población para más tarde ser procesado por los operadores genéticos. Este procedimiento se repite hasta rellenar la nueva población con el número de individuos necesarios. Aunque este esquema selecciona buenos individuos para la siguiente generación, no puede garantizar que la mejor solución obtenida sobreviva a lo largo del proceso de optimización. Por esta razón, también se usa el proceso llamado elitismo, donde el individuo con mejor (menor) valor de *fitness* se mantiene inalterado en la siguiente generación. El cruce BLX-α, que se aplica con una probabilidad P_c, escoge aleatoriamente nuevos individuos con valores en el intervalo [x',y'], donde x e y representan los valores de los genes de los dos progenitores para una variable determinada. Los descendientes se obtienen de la siguiente manera: $x' + r \cdot (y' - x')$ donde: $x' = x - \alpha \times (y - x)$; $y' = y + \alpha \cdot (y - x)$; r es un número aleatorio en el intervalo [0, 1], cumpliéndose que x' e y' estarán entre los extremos de la variable. La mutación aleatoria se aplica con una probabilidad P_m. y consiste en seleccionar aleatoriamente uno de los genes de un progenitor y generar un nuevo valor entre los límites de las variables con la que se corresponde ese gen.

3 INCERTIDUMBRE EN LA SUPERPOSICIÓN CRANEOFACIAL

Uno de los aspectos clave que hace tan desafiante el problema de la superposición craneofacial es la incertidumbre involucrada en el proceso así como las diferentes fuentes de las que esta proviene. Como ya dijimos, hemos identificado dos tipos de incertidumbre, a las que nos referimos como, incertidumbre en la localización e incertidumbre en el emparejamiento.

Po un lado, la incertidumbre en la localización se refiere a la gran dificultad inherente a la tarea de ubicar los puntos de referencia en una localización invariable y con la precisión requerida para esta aplicación. De hecho, cada antropólogo forense tiende a situar estos puntos en diferentes (cercanas) localizaciones. Los puntos de referencia anatómicos se especifican como ubicaciones biológicamente significativas que se pueden definir de forma inequívoca y se pueden localizar repetidamente con

un alto grado de exactitud y fidelidad [10]. Podemos apreciar como la incertidumbre está intrínsicamente ligada al punto de referencia. Incluso su definición usando lenguaje natural incluye expresiones como "un alto grado de", típicamente usadas en sistemas difusos. De hecho, el grado de precisión y fidelidad dependerá claramente de cada caso concreto. En nuestro caso, aunque los puntos de referencia se seleccionan buscando que sean lo más invariables posibles, factores como los cambios en las expresiones faciales o cambios en la edad, dificultan la localización precisa de estos puntos por el antropólogo.

Además de la dificultad inherente a la hora de localizar los puntos de referencia, la información de que disponemos no siempre está en las mejores condiciones. Por ejemplo, la ambigüedad para ubicar los puntos puede deberse en ocasiones a razones como la variación en la distribución de sombras debido a las condiciones de iluminación durante la toma de la fotografía, un enfoque inadecuado, una calidad de la imagen pobre, etc. Pues, la causa de la ambigüedad y por lo tanto su naturaleza puede variar de una imagen a otra e incluso de una característica a otra en una misma imagen [11]. Obviamente, la incertidumbre en la localización afecta a todos los puntos de referencia involucrados en el proceso de superposición craneofacial, tanto si pertenecen a la fotografía como si pertenecen al modelo del cráneo. Por simplificar, en este artículo centramos nuestro estudio en la incertidumbre relativa a los puntos cefalométricos, como primera aproximación. Consideramos que un punto de referencia difuso (ver Sección 4) representa la posición de una estructura biológica que se define de forma precisa pero que ocupa un área mayor que un simple punto en el sistema de referencia del observador. La comunidad de antropólogos físicos ya ha mostrado su interés en los puntos de referencia difusos [12].

Por otro lado, la incertidumbre en el emparejamiento se refiere a la incertidumbre involucrada en la correspondencia entre puntos de referencia asociados a diferentes objetos: la fotografía y el cráneo. Como ya dijimos, los puntos de referencia seleccionados se escogen de manera que sean lo más estables ante la presencia de cambios en las expresiones faciales, al tamaño o la edad de la persona. De cualquier manera, hay una cierta distancia entre cada punto craneométrico en el cráneo y su correspondiente punto cefalométrico en la fotografía. Esta distancia se debe a una cierta cantidad de tejido que separa los puntos de referencia del hueso del cráneo con sus correspondientes en la cara. Además, los puntos craneométricos y cefalométricos se asocian a ubicaciones biológicamente significativas. Si estudiamos el cráneo y la cara por separado, se puede observar cómo puntos significativos en el cráneo no siempre se localizan justo encima de los correspondientes puntos significativos en la cara (ver Figura 1). De este modo, nunca se dará una superposición perfecta de todos los puntos de referencia, incluso logrando la posición óptima del cráneo con respecto a la fotografía. Algunos estarán muy cerca, mientras que otros estarán separados por una pequeña pero significativa distancia. Nótese, que no todos los puntos de referencia contribuirán de igual manera a la incertidumbre del emparejamiento, cuanto mayor sea la distancia entre un punto cefalométrico y otro craneométrico, más importante será la aportación de ese par de puntos a la incertidumbre total del emparejamiento. Además, la proyección podría magnificar la distancia entre puntos de referencia alejados de la línea vertical definida por el punto medio entre los ojos y la barbilla. Por lo tanto, incluso si la distancia entre puntos cefalométricos y craneométricos es insignificante, podría magnificarse en el caso de imágenes en donde la persona desaparecida no tiene una pose completamente frontal.

Figura 1: De izquierda a derecha: puntos de referencia en el cráneo, en la cara, y superposición manual de ambas imágenes hecha por los antropólogos. Las flechas se usan para localizar un punto de referencia concreto donde se hace evidente un emparejamiento difuso.

4 LOCALIZACIÓN DIFUSA DE PUNTOS DE REFERENCIA PARA LA SUPERPOSICIÓN CRANEOFACIAL

En este artículo, el uso de puntos de referencia difusos tiene como objetivo afrontar la incertidumbre en la localización en la fotografía de una persona desaparecida. Conforme a este problema, cada experto forense podría situar cada punto de referencia en diferentes localizaciones de la imagen 2D. Por lo tanto, cuanto mayor es la incertidumbre relativa a un punto de referencia, más amplia será la zona donde los diferentes expertos forenses situarán los puntos de referencia.

Basándonos en el trabajo de Sinha [11], en lugar de usar un solo punto en la imagen para localizar los puntos cefalométricos, los hemos considerado como zonas rectangulares, a las que nos referiremos como puntos de referencia difusos. Cada punto de referencia difuso es un conjunto de píxeles que forman un rectángulo definido por esquinas diagonalmente opuestas: (x_1, y_1), (x_2, y_2) (donde $x_2 > x_1$ e $y_2 > y_1$). Por lo tanto, cuanto mayor es ese rectángulo, mayor será la incertidumbre asociada al punto de referencia.

Shina [11] introdujo estos conceptos para proporcionar mayor robustez y tolerancia a una Red Neuronal Artificial (RNA) utilizada para el emparejamiento de imágenes 2D de caras. Ghosh y Sinha [7] adaptaron la propuesta original para abordar la superposición craneofacial 2D. En estos trabajos se usaron rectángulos en lugar de puntos de referencia *crisp*, para evitar errores humanos debidos a la ambigüedad de las imágenes. Cada punto de referencia difuso se desfuzificaba temporalmente eligiendo el centroide del rectángulo como la característica objetivo. Usando los centroides para una primera adaptación de los pesos de la RNA, había una segunda etapa donde los rectángulos se adaptaban (reducían) en función de las salidas de la RNA. Aplicado al problema de la superposición craneofacial, este trabajo presenta dos limitaciones importantes, señaladas por los autores. Por un lado, nunca se consigue un ajuste correcto de una parte de la imagen del cráneo. Por otro lado, el método necesita una imagen frontal. Además, su propuesta no es completamente aplicable por su alto tiempo de computación, y por la necesidad de aplicar dos redes separadamente; una para el contorno superior del cráneo y otra para las características frontales del cráneo, donde la superposición conseguida por la primera podría alterar el resultado de la segunda.

A diferencia del trabajo de Sinha, en donde se usan puntos de referencia rectangulares para entrenar un RNA, nosotros modelamos esta incertidumbre para guiar nuestro procedimiento (basado en AGs) para realizar la superposición craneofacial. En nuestro caso, tratamos de evitar mínimos locales, priorizando unos puntos de referencia (localizados de manera más precisa) frente a otros (localizados de forma más difusa).

Por tanto, hemos modificado la definición previa de la función objetivo (ver Sección 2) como sigue:

$$Fitness = \frac{\sum_{i=1}^{N} \sqrt{\left[u_i (x'_{ci} - x_{fi})\right]^2 + \left[v_i (y'_{ci} - y_{fi})\right]^2}}{N} \quad (2)$$

donde x'_{ci} e y'_{ci} son las coordenadas de las transformaciones de los puntos 3D del cráneo Ci en el plano proyectado, x_{fi} e y_{fi} son los centroides de los puntos de referencia difusos asociado a cada punto cefalométrico de la foto 2D, y N es el número de puntos de referencia considerados. Los términos u_i, v_i se usan para representar la incertidumbre alrededor de cada punto de referencia. Cada valor depende del tamaño de su respectivo rectángulo, de la siguiente manera:

$$u_i = \frac{1}{1 + |x_2 - x_1|} \quad (3)$$

y,

$$v_i = \frac{1}{1 + |y_2 - y_1|} \quad (4)$$

donde (x_1, y_1) y (x_2, y_2) son esquinas diagonalmente opuestas. En esta formulación $(x_2 - x_1)$ y $(y_2 - y_1)$ son medidas de la incertidumbre en los ejes X e Y, respectivamente.

Según la formulación de la ecuación 2, a mayor tamaño del rectángulo que define el punto de referencia difuso (esto es, menor valor de u_i y/o v_i), menor será el peso de dicho punto de referencia en la función de *fitness*. Por lo tanto, cuanto mayor es la incertidumbre relativa a un punto de referencia, menor será la importancia de dicho punto en el proceso de optimización (minimización del error).

Por otro lado, como alternativa a la anterior formulación, en lugar de ponderar la localización de cada componente de la distancia Euclídea (Ec. 2), proponemos ponderar cada componente de la distancia como sigue:

$$Fitness_2 = \frac{\sum_{i=1}^{N} \sqrt{u_i (x'_{ci} - x_{fi})^2 + v_i (y'_{ci} - y_{fi})^2}}{N} \quad (5)$$

Nótese que $Fitness_2$ reduce ligeramente el efecto de la incertidumbre.

5 EXPERIMENTACIÓN

Tras explicar las fuentes de incertidumbre y nuestra propuesta para manejar la incertidumbre en la localización, estudiamos su comportamiento. La Sección 5.1 presenta el diseño experimental, mientras que en las Secciones 5.2 y 5.3 se analiza el comportamiento de nuestra propuesta afrontando dos casos reales.

5.1. DISEÑO EXPERIMENTAL

Teniendo en cuenta las características de la imagen (tamaño, resolución, contraste, etc.) y qué puntos de referencia son más sensibles a la incertidumbre (ver Sección 3), usamos diferentes tamaños de rectángulos para diferenciar la influencia de distintos grupos de puntos de referencia. Además, para los dos casos reales usados en este trabajo, simulamos el proceso seguido por diez antropólogos diferentes a la hora de identificar manualmente los puntos cefalométricos. Concretamente, asumimos una distribución aleatoria dentro del rectángulo que define cada punto de referencia difuso. Por lo tanto, se simula cada punto de referencia *crisp* seleccionando de manera aleatoria un punto dentro del rectángulo. A

continuación, usamos los diez conjuntos de puntos de referencia y aplicamos nuestro AG (Sección 2) para resolver los correspondientes problemas de superposición. Finalmente, comparamos estos resultados con los logrados por el método actual que ahora proponemos de superposición craneofacial basado en la localización difusa de los puntos de referencia cefalométricos.

Todas las pruebas se han hecho sobre casos reales de identificación que, han sido resueltos por el personal del laboratorio de Antropología Física de la Universidad de Granada, con la colaboración de la policía científica de España. Como se ha dicho, consideraremos dos casos de estudio, resueltos siguiendo una aproximación manual para la superposición fotográfica. Para cada caso consideraremos la fotografía 2D de la persona desaparecida y su correspondiente modelo 3D del cráneo, adquirido en el mencionado laboratorio usando su escáner láser Konica-Minolta 3D VI-910.

Todos los experimentos se llevaron a cabo con las dos variantes de función de *fitness* (ver Sección 4). En la Tabla 1 se muestran los parámetros del AG usados, seleccionados como óptimos a partir de nuestro trabajo previo [2].

Tabla 1. Conjunto de parámetros del algoritmo genético usado para la superposición craneofacial

Parámetros del AG			
Generaciones	600	Tamaño de población	600
Tamaño de ventana	2	Parámetro BLX-α	0.7
Probabilidad de cruce	0.9	Probabilidad de mutación	0.2

Para todas las pruebas, se consideran 30 ejecuciones del AG. Las Tablas 2 y 3 muestran el mínimo (m), máximo (M), media (μ) y desviación típica (σ) del ECM relativo a la distancia Euclídea entre puntos de referencia cefalométricos y los correspondientes craneométricos, una vez transformados, usando la mejor solución alcanzada por el AG. Nótese que en el caso de puntos de referencia *crisp*, la función de *fitness* es igual al ECM. Sin embargo, en el caso de puntos de referencia difusos, la función de *fitness* no es el ECM, sino una suma ponderada de las distancias entre los puntos de referencia del cráneo transformados y el centroide de sus respectivas regiones rectangulares de la fotografía. Debido a esto, los resultados numéricos no son del todo comparables. No obstante, hemos usado el ECM como la única forma de asignar un mismo error de medida para ambos casos.

5.2. CASO DE ESTUDIO 1

La fotografía de la cara de la persona desaparecida fue proporcionada por la familia, haciendo este estudio con su consentimiento. Su identidad fue confirmada usando la supra-proyección fotográfica. La imagen 2D original, es una imagen color de 1062 x 745 píxeles (ver Figura 2, izquierda). Sin embargo, los forenses centraron su estudio en una versión recortada y escalada de la imagen original con un nuevo tamaño de 290 x 371 píxeles (ver Figura 2, centro). El modelo 3D del cráneo (ver figura 2, derecha) está compuesto por 243202 puntos.

Figura. 2. Caso de estudio 1. De izquierda a derecha. Imagen original en donde se encuentra la persona desaparecida (izquierda). Fotografía de la persona desaparecida, recortada de la original y posteriormente escalada (centro). Modelo 3D del cráneo (derecha)

Los antropólogos identificaron siete puntos cefalométricos. Los puntos craneométricos correspondientes fueron extraídos manualmente del modelo 3D del cráneo. Debido a la orientación de la cara y a la calidad y resolución de la fotografía, se definieron tres tamaños de puntos de referencia difusos. En primer lugar, los cuatro puntos de referencia en la línea vertical que divide a la cara en dos partes simétricas (Nasio, Subnasal, Prostio, Gnatio) los definimos como un rectángulo de 2 x 3 píxeles. Por otro lado, para los dos puntos de referencia que se encuentran en el centro de las cavidades oculares (Ectocantions) elegimos un rectángulo de 8 x 8 píxeles. Por último, el punto de referencia próximo a la oreja derecha (Zygion) lo representamos por un rectángulo de 12x12 píxeles. Como se ha dicho anteriormente, cuanto más cerca está un punto de referencia a la mencionada línea vertical, más pequeño será el rectángulo que defina a ese punto de referencia.

La Tabla 2 muestra el ECM de las superposiciones craneofaciales, usando tanto localización difusa como precisa. Nótese que los valores se corresponden con 30 ejecuciones del AG y además, en el caso de localización precisa, los valores se corresponden a diez conjuntos de puntos de referencia distintos. Es importante reseñar que los resultados no son del todo comparables ya que el proceso de superposición en el caso de la localización difusa, minimiza una función diferente al ECM.

Tabla 2. Caso de estudio 1. Resultados de la superposición usando localización precisa y difusa.

Prueba	ECM			
	m	M	μ	σ
Puntos *crisp* (10 "forenses")	0.014	0.135	0.031	0.017
Puntos difusos (*fitness* Ec. 2)	0.020	0.321	0.129	0.101
Puntos difusos (*fitness* Ec. 5)	0.020	0.210	0.053	0.046

En la Figura 3 se presentan los resultados de la superposición de las aproximaciones *crisp* y difusa. Por un lado, el mejor y peor resultado logrado por diez

expertos forenses[1], a la izquierda y centro respectivamente de la Figura 3. Se puede observar la alta variabilidad entre ambos resultados debido a los diferentes conjuntos de puntos de referencia usados. Por otro lado, a la derecha de la Figura 3, se muestra el mejor resultado del AG usando una aproximación difusa de la localización de los puntos de referencia. Nótese que las tres superposiciones se corresponden con el mejor de 30 ejecuciones del AG.

Figura 3. Caso de estudio 1. De izquierda a derecha. Mejor y peor resultado alcanzado usando puntos de referencia *crisp*. A la derecha mejor resultado considerando puntos de referencia difusos.

5.3. CASO DE ESTUDIO 2

El segundo caso se corresponde con una vista frontal de una chica joven. El modelo 3D del cráneo esta formado por 327641 puntos y la fotografía es una imagen en color de 744 x 1147 puntos. A diferencia del caso anterior, los antropólogos pudieron localizar ocho puntos cefalométricos y sus correspondientes craneométricos fueron localizados manualmente en el modelo 3D del cráneo.

Debido a las fuentes de incertidumbre mencionadas y teniendo en cuenta que en este caso disponemos de una fotografía frontal, sólo se definieron dos tamaños diferentes de rectángulo para los puntos de referencia difusos. Los cuatro puntos en la ya mencionada línea vertical son definidos, igual que en el caso anterior, mediante un rectángulo de 2 x 3 píxeles. Los cuatro puntos en las zonas interior y exterior de las cavidades oculares (Ectocantion izquierdo y derecho, y Endocantion izquierdo y derecho) se localizaron mediante un rectángulo de 8 x 8 píxeles.

La Tabla 3 muestra el ECM de las superposiciones craneofaciales, usando tanto localización difusa como precisa. Nótese que como en al caso 1, los valores se corresponden con 30 ejecuciones del AG y además, en el caso de la localización precisa, los valores mostrados son una media de los resultados obtenidos sobre cada uno los diez conjuntos de puntos de referencia diferentes. Es importante reseñar que los resultados no son del todo comparables ya que el proceso de superposición en el caso de la localización difusa, minimiza una función diferente al ECM. Aun así, en este caso el ECM es muy similar en los resultados de ambos métodos.

[1] Como ya se ha comentado, emulamos a diez expertos forenses en la extracción de los puntos cefalométricos (Sección 5.1).

Tabla 3. Caso de estudio 2. Resultados de la superposición usando localización precisa y difusa.

Prueba	ECM			
	m	M	μ	σ
Puntos *crisp* (10 "forenses")	0.004	0.057	0.018	0.012
Puntos difusos (*fitness* Ec. 2)	0.004	0.049	0.022	0.011
Puntos difusos (*fitness* Ec. 5)	0.004	0.055	0.017	0.012

En la Figura 4 se presentan los resultados de la superposición de las aproximaciones *crisp* y difusa. Por un lado, como en la figura 3, el mejor y peor resultado logrado por diez expertos forenses[1], a la izquierda y centro respectivamente. Se puede observar nuevamente la alta variabilidad entre ambos resultados debido a los diferentes conjuntos de puntos de referencia usados. Por otro lado, a la derecha de la Figura 4, se muestra el mejor resultado del AG usando una aproximación difusa de la localización de los puntos de referencia. Nótese que las tres superposiciones se corresponden con el mejor de 30 ejecuciones del AG. Además, la superposición lograda usando puntos difusos parece ser visualmente mejor que el mejor resultado visual logrado usando puntos de referencia *crisp* ya que, en el caso de puntos difusos la orientación alcanzada por el cráneo es más frontal y por lo tanto más similar a la orientación de la cara en la fotografía.

Figura 4. Caso de estudio 2. De izquierda a derecha. Mejor y peor resultado alcanzado usando puntos de referencia *crisp*. A la derecha mejor resultado considerando puntos de referencia difusos.

6 CONCLUSIONES Y TRABAJO FUTURO

El objetivo final de nuestro proyecto es diseñar un procedimiento automático basado en *soft computing* para asistir a los antropólogos forenses en la tarea de identificación mediante supra-proyección fotográfica.

En este artículo hemos propuesto y validado el uso de puntos de referencia cefalométricos difusos para abordar la tarea de localización de los puntos de referencia usados

en la superposición craneofacial. Se han mostrado y analizado los resultados obtenidos en dos casos reales para los que hemos simulado el comportamiento de 10 antropólogos forenses, cada uno de los cuales habría extraído su conjunto de puntos de referencia de cada imagen 2D de las personas desaparecidas. Basándonos en un AG que introducimos en [2] para llevar a cabo la superposición automática, hemos propuesto dos nuevas funciones de *fitness* para el caso de la superposición basada en puntos de referencia difusos. Hemos comparado las respectivas superposiciones automáticas *crisp,* con las logradas usando puntos de referencia difusos, pudiendo concluir que se han logrado resultados prometedores, especialmente usando la función de f*itness* dada por la ecuación 5. Los resultados de las superposiciones muestran que la aproximación difusa es capaz de modelar la incertidumbre involucrada en la localización de los puntos de referencia cefalométricos. Es decir, ninguno de los diez expertos emulados ha logrado un resultado mejor que los conseguidos siguiendo una aproximación difusa. Por tanto, podemos decir que nuestra propuesta es más robusta cuando tenemos diferentes localizaciones posibles para los puntos de referencia. Sin embargo, estos resultados necesitan ser confirmados en un estudio más detallado incluyendo un mayor número de casos.

Como trabajos futuros, haremos una encuesta entre diferentes forenses antropólogos con el fin de definir la forma y tamaño más adecuados de los puntos de referencia difusos, mediante la recogida de diferentes conjuntos de puntos de referencia *crisp*. Estos últimos se usaran para generar nuevas superposiciones automáticas con las que comparar los resultados obtenidos usando puntos difusos. Además, se buscará, nuevas definiciones de la función de *fitness*, tras la modelización de la incertidumbre en la localización mediante conjuntos difusos. Así mismo, pretendemos determinar los parámetros de las transformaciones manuales logradas por los expertos forenses y usar éstas como solución óptima con la que poder comparar los resultados obtenidos. Así podríamos medir la precisión del proceso automático de superposición *crisp* y difuso. Además, planeamos introducir la incertidumbre en el emparejamiento.

Finalmente, pretendemos afrontar la tercera etapa del proceso de identificación, (la toma de decisiones) mediante lógica difusa, con el objetivo de asistir al experto forense a la hora de tomar la decisión final en la tarea de identificación.

Agradecimientos

Este trabajo está soportado por el Ministerio de Ciencia y Tecnología (ref. TIN2006-00829) y por la Consejería de Innovación y Ciencia de la Junta de Andalucía (ref. TIC1619), ambos incluyendo fondos FEDER.

Referencias

[1] Bäck T., Fogel D.B., Michalewicz Z. (Eds.), *Handbook of evolutionary computation*, IOP Publishing Ltd and Oxford University Press, 1997

[2] Ballerini L., Cordón O., Damas S., Santamaría J.: Craniofacial superimposition in Forensic identification using genetic algorithms. *Technical report ECSC AFE 2008-03. European Center for Soft Computing* (2008)

[3] Ballerini L., Cordón O., Damas S., Santamaría J., Alemán I., Botella M.: Identification by computer-aided photographic supra-projection: a survey. *Technical report ECSC AFE 2007-04. European Center for Soft Computing* (2007)

[4] Blickle T.: Tournament selection, in: T. Bäck, D. B. Fogel, Z. Michalewicz (Eds.), *Handbook of evolutionary computation*, IOP Publishing Ltd and Oxford University Press, Ch. 2.3 (1997)

[5] Cordón, O., Gomide, F., Herrera, F., Hoffmann, F., Magdalena, L.: Ten years of genetic fuzzy systems: Current framework and new trends, *Fuzzy Sets and Systems* 141 (1) 5-31 (2004)

[6] Eshelman L.J.: Real-coded genetic algorithms and interval schemata, in: L. D. Whitley (Ed.),*Foundations of Genetic Algorithms* 2, Morgan Kaufmann, San Mateo, pp. 187–202. (1993)

[7] Ghosh A., Sinha P.: An economized craniofacial identification system. *Forensic Science International* 117 (1-2), 109-119 (2001)

[8] Iscan M. Y.: Introduction to techniques for photographic comparison. In: M. Y. Iscan, R. Helmer (eds.), *Forensic Analysis of the Skull*, pp. 57-90. Wiley, (1993)

[9] Nickerson B. A., Fitzhorn P. A., Koch S. K., Charney M.: A methodology for near-optimal computational superimposition of two-dimensional digital facial photographs and three-dimensional cranial surface meshes, *Journal of Forensic Sciences* 36 (2) 480–500 (1991)

[10] Richtsmeier J., Paik C., Elfert P., Cole T.M., Dahlman F.: Precision, repeatability and validation of the localization of cranial landmarks using computed tomography scans. *Cleft Palate Craniofac* J 32:217-227 (1995)

[11] Sinha, P.: A symmetry perceiving adaptive neural network and facial image recognition. *Forensic Science International* 98 (1-2), 67-89 (1998)

[12] Valeri, C.J., Theodore, M.C., Lele, S., Richtsmeier, J.T.: Capturing Data From Three-Dimensional Surfaces Using Fuzzy Landmarks. *American Journal of Physical Anthropology*. 107, 113-124 (1998)

VISUAL COARSENESS REPRESENTATION BY MEANS OF FUZZY SETS

J. Chamorro-Martínez, P. Martínez-Jiménez
Department of Computer Science and Artificial Intelligence
University of Granada
{jesus,pmartinez}@decsai.ugr.es

Abstract

In this paper, the texture feature ”coarseness” is modelled by means of fuzzy sets, relating representative coarseness measures (our reference set) with the human perception of this type of feature. In our study, a wide variety of measures have been analyzed, defining unidemsional and bidimensional fuzzy set for different combination of measures. The fineness human perception has been collected from polls filled by human subjects, performing an aggregation of their assessments by means of OWA operators. Using as reference set a combination of some measures, the membership function corresponding to the fuzzy set is modelled as the function which provides the best fit of the collected data.

Keywords: Image features, textural features, fuzzy texture, human perception, coarseness, fineness.

1 INTRODUCTION

It is usual for humans to describe visual textures according to some vague ”textural concepts” like *coarseness/fineness*, *orientation* or *regularity* [2, 17, 22]. From all of them, the *coarseness/fineness* is the most popular one, being common to associate the presence of fineness with the presence of texture (from this point of view, texture is defined as local variations against the idea of homogeneity). In this sense, a *fine* texture is considered as small texture primitives with big gray tone differences between neighbor primitives (e.g. the image in figure 1(A)), whereas a *coarse* texture corresponds to bigger primitives formed by several pixels (e.g. the image in figure 1(I)). By considering the importance of this textural concept, in this paper we will focus our study on the ”fineness” modelling (let us remark that ”coarseness” and ”fineness” are opposite but related textural concepts).

There are many measures in the literature that, given an image, capture the fineness (or coarseness) presence in the sense that the greater the value given by the measure, the greater the perception of texture [9]. However, there is no perceptual relationship between the value given by these measures and the degree in which the humans perceive the texture. Thus, given a certain value calculated by applying a measure to an image, there is not an immediate way to decide whether there is a fine texture, a coarse texture or something intermediate (i.e. there is not a textural interpretation).

The imprecision associated to these fineness measures suggests the use of representation models that incorporate the uncertainty. Nevertheless, the majority of the approaches that can be found in the literature are crisp proposals [5, 9, 18, 22] where uncertainty is not properly taken into account. To face this problem, fuzzy logic has been recently employed for representing the imprecision related to texture. However, in many of these approaches, fuzzy logic is usually applied just during the process but the output do not habitually model the imprecision (being often a crisp one). Examples of this fact are frequently found in the literature, like those approaches that use texture to perform image segmentation or classification on the basis of fuzzy clustering [8, 20, 23], fuzzy rules [6, 10, 16], etc.

Other interesting approaches emerge from the content-based image retrieval scope, where semantic data is managed by means of fuzzy sets [12, 13]. In these proposals, a mapping from low-level statistical features to high level textural concepts is performed by defining membership functions for each textural feature. However, given a feature, these membership functions are not obtained by considering the relationship between the computational feature and the human perception of texture, so the linguis-

[1]This work has been partially supported by the Andalusian Government under the TIC1570 and TIC249

tic labels related to these membership functions do not necessarily match what a human would expect.

In this paper we propose to solve this problem by representing the concept of fineness on the basis of fuzzy sets relating representative measures of the modelled texture (usually some statistic) with its human perception of fineness. In our analysis, a wide variety of measures have been considered, including classical statistical measures, frequency domain and fractal dimension measures, etc. Thus, this paper propose a group of fuzzy sets (defined on different measures) which model the textural fineness according to the human perception. The performance of each fuzzy set is analyzed and checked with the human assessments, proposing a subgroup of them as the most adequate for modelling fineness perception in texture images.

The rest of the paper is organized as follows. In section 2 we introduce our methodology to obtain the fuzzy sets that model the fineness textural concept. In section 3 we show the results of applying the models and the main conclusions and future work are summarized in section 4.

2 FUZZY MODELLING OF FINENESS

In this paper we propose to model the fineness perception as a fuzzy set defined on the domain of a given set of measures. For this purpose, two questions will be faced: (i) what reference set should be used for the fuzzy set, and (ii) how to obtain the related membership functions.

As reference set, a vector of fineness measures will be used. Concretely, we have analyzed the $K = 17$ measures indicated in the first column of table 1, that includes classical statistical measures well known in the literature, measures in the frequency domain, fractal dimension analysis, etc. All of them are automatically computed from the texture image. From now on, we will note $\mathcal{P} = \{P_1, \ldots, P_K\}$ the set of K measures of fineness analyzed in this paper, $F \subseteq \mathcal{P}$ a subset of $K' \leq K$ measures, and $\mathcal{T}_F$ the fuzzy set defined on the domain of F[1].

To obtain the membership function[2] of the fuzzy set $\mathcal{T}_F$, a functional relationship between the subset of measures F and the presence degree of fineness related to it will be learnt. To do it, we will use a set $\mathcal{I} = \{I_1, \ldots, I_N\}$ of N images that fully represent the different degrees of fineness. Thus, for each image $I_i \in \mathcal{I}$, we will obtain (a) a vector of measures $\mathbf{M}^i = [m_1^i, \ldots, m_{K'}^i]$, where m_k^i is a value for the measure $P_k \in \mathcal{P}$ applied to the image I_i, and (b) a human assessment of the fineness degree perceived, noted as v^i, which will be collected by means of a poll with human subjects (section 2.1). Once we have a multiset of valid pairs $\Psi_F = \{(\mathbf{M}^1, v^1), \ldots, (\mathbf{M}^N, v^N)\}$, we shall estimate the membership function by fitting a suitable function to the multiset of points Ψ (section 2.2).

Figure 1: Some examples of images with different degrees of fineness

2.1 ASSESSMENT COLLECTION

In this section, the way to obtain a vector $\Gamma = [v1, \ldots, v^N]$ of the assessments of the perception degree of fineness from the image set $\mathcal{I} = \{I_1, \ldots, I_N\}$ will be described. Thus, firstly the image set $\mathcal{I}$ will be selected. After that, a poll which allows to get assessments of the perception degree of fineness will be designed. These assessments will be obtained for each image in $\mathcal{I}$, so an aggregation of the different assessments will be performed.

[1] Given a set $\mathcal{P}$ of K measures, the fuzzy set $\mathcal{T}_F$ can be defined by using all the measures ($K' = K$) or a subset of them ($K' < K$)

[2] To simplify the notation, as it is usual in the scope of fuzzy sets, we will use the same notation $\mathcal{T}_F$ for the fuzzy set and for the membership function that defines it

2.1.1 The texture image set

A set $\mathcal{I} = \{I_1, \ldots, I_N\}$ of $N = 80$ images representative of the concept of *fineness* has been selected. Figure 1 shows some images extracted from the set $\mathcal{I}$. Such set has been selected satisfying the following properties:

1. It covers the different presence degrees of fineness.
2. The number of images for each presence degree is representative enough.
3. Each image shows, as far as possible, just one presence degree of fineness.

Due to the third property, each image can viewed as "homogeneous" respect to the fineness degree represented, i.e., if we select two random windows (with a dimension which does not "break" the original texture primitives and structure), the perceived fineness will be the same for each window (and also respect to the original image). In other words, we can see each image $I_i \in \mathcal{I}$ as a set of lower dimension images (windows) with the same fineness degree of the original one. For example, images of Figure 1 are 128×128 in size, but the perceived fineness is the same for windows of size 32×32; from this point of view, we can think in $I_i \in \mathcal{I}$ as an unique image of size 128×128 or as 9216 images of size 32×32 (all of them with the same degree of fineness).

As we explained in the introduction of this section, given an image $I_i \in \mathcal{I}$, we obtain a vector of measures $\mathbf{M}^i = [m_1^i, \ldots, m_{K'}^i]$, with m_k^i being the result of applying the measure $P_k \in \mathcal{P}$ to the image I_i. However, and thanks to the third property, we really can obtain a set of vectors associated to each image $I_i \in \mathcal{I}$ (one vector for each window). Let W be the number of windows considered for each image; from now on we will note as $\mathbf{M}_w^i = [m_1^{i,w}, \ldots, m_{K'}^{i,w}]$, with $w = 1, \ldots, W$, the vector of measures for the w-th window of the image I_i, with $m_k^{i,w}$ being the result of applying the measure $P_k \in \mathcal{P}$ to the w-th window of the image I_i. Therefore, the multiset of valid pairs $\Psi_F =$ will be given by $\Psi_F = \{(\mathbf{M}_w^i, v^i), i = 1, \ldots, N; w = 1, \ldots, W\}$.

2.1.2 The poll

Given the image set $\mathcal{I}$, the next step is to obtain assessments about the perception of fineness from a set of subjects. From now on we shall note as $\Theta^i = [o_1^i, \ldots, o_L^i]$ the vector of assessments obtained from L subjects for the image I_i. To get Θ^i, subjects will be asked to assign images to classes, so that each class has associated a perception degree of fineness.

In particular, 20 subjects have participated in the poll and 9 classes have been considered. The first nine images in figure 1 show the nine representative images for each class used in this poll. It should be noticed that the images are decreasingly ordered according to the presence degree of the fineness concept. The first class (Figure 1(A)) represents a presence degree of 1 while the ninth class (Figure 1(I)), represents a presence degree of 0. The rest of the classes (Figure 1(B)-(H)) represent degrees in the interval (0,1).

As result, a vector of 20 assessments $\Theta^i = [o_1^i, \ldots, o_{20}^i]$ is obtained for each image $I_i \in \mathcal{I}$. The degree o_j^i associated to the assessment given by the subject S_j to the image I_i is computed as $o_j^i = (9 - k) * 0.125$, where $k \in \{1, \ldots, 9\}$ is the index of the class C_k to which the image is assigned by the subject.

2.1.3 Assessment aggregation

Our aim at this point is to obtain, for each image in the set $\mathcal{I}$, one assessment v^i that summarizes the assessments Θ^i given by the different subjects about the presence degree of fineness.

To aggregate opinions we have used an OWA operator guided by a quantifier [25]. Concretely, the quantifier "the most" has been employed, which allows to represent the opinion of the majority of the polled subjects. This quantifier is defined as

$$Q(r) = \begin{cases} 0 & \text{if } r < a, \\ \frac{r-a}{b-a} & \text{if } a \leq r \leq b, \\ 1 & \text{if } r > b \end{cases} \quad (1)$$

with $r \in [0, 1]$, $a = 0.3$ and $b = 0.8$. Once the quantifier Q has been chosen, the weighting vector of the OWA operator can be obtained following Yager [25] as $w_j = Q(j/L) - Q((j-1)/L), j = 1, 2, ..., L$. According to this, for each image $I_i \in \mathcal{I}$, the vector Θ^i obtained from L subjects will be aggregated into one assessment v^i as follows:

$$v^i = w_1 \hat{o}_1^i + w_2 \hat{o}_2^i + ... + w_L \hat{o}_L^i \quad (2)$$

where $[\hat{o}_1^i, \ldots, \hat{o}_L^i]$ is a vector obtained by ranking in non-increasing order the values of the vector Θ^i.

2.2 FITTING THE MEMBERSHIP FUNCTION

At this point, the aim is to obtain, for a given subset of measures $F \subseteq \mathcal{P}$, the corresponding membership function $\mathcal{T}_F$. It is defined on a vector of K' real measures, i.e.,

$$\mathcal{T}_F : \mathbb{R}^{K'} \rightarrow [0, 1] \quad (3)$$

with K' being the cardinality of F. Since we are searching for a function which associates the measure vectors ($\mathbf{M}_w^i$) and the human assessments of fineness (v^i), we propose to estimate $\mathcal{T}_F$ by fitting a suitable curve to the multiset of points $\Psi_F = \{(\mathbf{M}_w^i, v^i), i = 1, \ldots, N; w = 1, \ldots, W\}$.

The values of the measures can be affected by some factors independent of the fineness (brightness, contrast, noise), which typically causes outliers in the fitting points. For this reason, in our approach the membership function is calculated by means of a Robust Fitting of the multiset Ψ_F (this fitting could be done according to some constraints, for example, to obtain a monotonic function). The robust fitting methods avoid that these outliers move the solution far from the true fit, as it happens in least squares. For this purpose we propose one of the most used techniques in computer vision: M-estimation.

M-estimators are a generalization of the traditional maximum likelihood estimation and, therefore, of the least squares fitting. If we consider a function $\mathcal{T}_F$ defined by D parameters $p_1, \ldots, p_D$, these parameters will be obtained as follows:

$$\underset{p_1,\ldots,p_D}{\operatorname{argmin}} \sum_{i=1}^{N} \rho\left(\mathbf{r}_w^i\right) \tag{4}$$

where $\mathbf{r}_w^i$ are the fitting residuals defined as

$$\mathbf{r}_w^i = v^i - \mathcal{T}_F(\mathbf{M}_w^i; p_1 \ldots p_D) \tag{5}$$

and the function ρ gives the contribution of each residual in the model, whose objective is to reduce the effect of outliers. This function ρ should have the following properties:

- ρ should be a symmetric, positive-definite function.
- ρ should have a unique minimum at zero (when the residual is zero).
- ρ should be derivable.

The minimization problem is solved by finding the parameters $p_1, \ldots, p_D$ which are the solution of the following system of D equations:

$$\sum_{i=1}^{N} \frac{\partial \rho\left(\mathbf{r}_w^i\right)}{\partial \mathbf{r}_w^i} \frac{\partial \mathbf{r}_w^i}{\partial p_j} = 0 \qquad j = 1, \ldots, D \tag{6}$$

Instead of solving directly this problem, we define a weight function

$$\omega(x) = \frac{\partial \rho(x)/\partial x}{x} \tag{7}$$

and then, Equation 6 becomes

$$\sum_{i=1}^{N} \omega\left(\mathbf{r}_w^i\right) \mathbf{r}_w^i \frac{\partial \mathbf{r}_w^i}{\partial p_j} = 0 \qquad j = 1, \ldots, D \tag{8}$$

which is a weighted least-squares problem. The weights, however, depend on the residuals, the residuals depend on the estimated coefficients, and the estimated coefficients depend on the weights. Therefore, an iterative solution (called iteratively reweighted least-squares, IRLS) is required. This process alternates steps of calculating the weights using the current estimated coefficients and solving Equation 8 to estimate the new coefficients with the weights fixed.

The function ρ is chosen so that the weights assigned to possible outliers become smaller in each iteration. We have taken one of the most commonly used functions in literature, that achieves a robust outlier rejection: the Tukey's biweight function [3]. The definition of the objective function $\rho(\mathbf{r}_w^i)$ and the weight function $\omega(\mathbf{r}_w^i)$ is:

$$\rho(\mathbf{r}_w^i) = \begin{cases} \frac{k^2}{6}\left[1 - \left(1 - \left(\frac{\mathbf{r}_w^i}{k}\right)^2\right)^3\right] & \text{if } \left|\mathbf{r}_w^i\right| \le k \\ \frac{k^2}{6} & \text{if } \left|\mathbf{r}_w^i\right| > k \end{cases} \tag{9}$$

$$\omega(\mathbf{r}_w^i) = \begin{cases} \left(1 - \left(\frac{\mathbf{r}_w^i}{k}\right)^2\right)^2 & \text{if } \left|\mathbf{r}_w^i\right| \le k \\ 0 & \text{if } \left|\mathbf{r}_w^i\right| > k \end{cases} \tag{10}$$

where k is a parameter called the *tuning constant* and it is defined as

$$k = 4.685\sigma \qquad \text{with } \sigma = \frac{median\left|\mathbf{r}_w^i - \mu_r\right|}{0.6745} \tag{11}$$

This parameter k must be recalculated in each iteration.

It should be noticed that there is an error related to the fitting (the mean of the absolute difference between the data points and the curve). This error can also be viewed as a goodness measure of the measures used in F; according to this, different subsets of measures could be analyzed, selecting the one with the lowest error.

For selecting the K' measures which compounds the subset F, we can consider 2^K possible combinations. These possible combinations include from a vector comprised of K items to vectors with only one item. In this paper, the cases of $K' = 1$ and $K' = 2$ will be considered. The first case ($K' = 1$) will allow to perform a study of each measure separately and to compare its goodness to model the *fineness*; the second one ($K' = 2$) will allow to combine measures in order to improve the individual ones. For higher dimensions ($K' \ge 3$) the Robust Fitting is too complex, so new methods for finding $\mathcal{T}_F$ need to be considered (for example, some kind of heuristic approaches). These cases will be considered in future works.

Table 1: Fitting errors related to each measure and parameter values for the two measures with least error

Measure	Error
Correlation [9]	0.1466
Amadasun [2]	0.1515
Abbadeni [1]	0.1880
Fractal dim. (FD) [15]	0.1891
Tamura [22]	0.1994
Edge density (ED) [4]	0.2044
DGD [11]	0.2120
Local Homogeneity [9]	0.2156
Short Run Emphasis [7]	0.2211
SNE [21]	0.2365
Weszka [24]	0.2383
Newsam [14]	WSD
Entropy [9]	WSD
Uniformity[9]	WSD
FMPS [26]	WSD
Variance[9]	NR
Contrast [9]	NR

Parameters for Correlation	
a_3	-1.83828
a_2	1.23377
a_1	-1.20788
a_0	1.03395
α	0.76851
β	0.02893

Parameters for Amadasun	
a_3	-0.00006
a_2	0.00556
a_1	-0.17053
a_0	1.57472
α	16.5285
β	3.82800

To define $\mathcal{T}_F$, the following considerations will be taked into account:

- $\mathcal{T}_F$ should be a monotonic function
- The values $\mathcal{T}_F(x) = 0$ and $\mathcal{T}_F(x) = 1$ should be achieved from a certain value

2.2.1 Unidimensional fitting

Regarding the above properties, in the case of $K' = 1$ we propose to define $\mathcal{T}_k$ as a function

$$\mathcal{T}_k : \mathbb{R} \rightarrow [0, 1] \tag{12}$$

of the form[3]

$$\mathcal{T}_k(x; a_n \dots a_0, \alpha, \beta) = \begin{cases} 0 & x < \alpha, \\ poly^n(x; a_n \dots a_0) & \alpha \leq x \leq \beta, \\ 1 & x > \beta \end{cases} \tag{13}$$

with $poly^n(x; a_n \dots a_0)$ being a polynomial function

$$poly^n(x; a_n \dots a_0) = a_n x^n + \dots + a_1 x^1 + a_0 \tag{14}$$

In our proposal, the parameters $a_n \dots a_0$, α and β of the function $\mathcal{T}_k$ are calculated by carrying out a Robust Fitting on Ψ_k taking into account the restriction of obtaining a monotonic function and considering n=1,2,3 (i.e. linear, quadratic and cubic functions) to define the polynomial function. Table 1 shows for each measure $P_k \in \mathcal{P}$ the least fitting error obtained. Note that this value can be viewed as a goodness measure of the ability of the measure to represent the perception of fineness.

Table 1 has been sorted in increasing order of the errors. The parameter values of the two measures with the lowest error (Correlation and Amadasun) are also shown in Table 1. It should be noticed that we haven't carried out the fitting with six of the measures. Four of them (marked with WSD) are rejected because their values are affected by the window size, ie., they are window size dependent. The other two (marked with NR) produce a diffuse cloud of points Ψ_F which implies that these measures are not providing a representative information about the perception of fineness.

2.2.2 Bidimensional fitting

In the case of $K' = 2$ we define $\mathcal{T}_k$ as a function

$$\mathcal{T}_k : \mathbb{R}^2 \rightarrow [0, 1] \tag{15}$$

of the form

$$\mathcal{T}_k(x, y; a_{(n!+n)} \dots a_0) = \mathcal{T}_k(x, y; coef) =$$

$$= \begin{cases} 0 & poly2^n(x, y; coef) < 0, \\ poly2^n(x, y; coef) & 0 < poly2^n(x, y; coef) < 1, \\ 1 & poly2^n(x, y; coef) > 1 \end{cases} \tag{16}$$

with $poly2^n(x, y; coef)$ being a polynomial function of two variables

$$poly2^n(x, y; a_{(n!+n)} \dots a_0) = \sum_{i=0}^{n} \sum_{j=0}^{i} a_{(i!+j)} x^j y^{i-j} =$$

$$= a_{(n!+n)} x^n + a_{(n!+n-1)} x^{n-1} y^1 + \dots + a_2 x^1 + a_1 y^1 + a_0 \tag{17}$$

Like in the previous case, the parameters $a_{(n!+n)} \dots a_0$ of the function $\mathcal{T}_k$ are calculated by carrying out a Robust Estimation considering n=1,2,3 to define the polynomial function,but in this case the fitting points are located in a tridimensional space and the polynomial function has two variables. The measures are taken in pairs by combining only the first six measures in Table 1.

Table 2 shows the least fitting error obtained for each pair of measures. It can be noticed that the first pair of measures (FD + Amadasun) obtains a fitting error much smaller than the others, but moreover it is much smaller than the fitting error of each measure separately. This implies that it provides the most representative information about the perception of fineness. The parameter values of this pair of measures appear in Table 2.

[3]Note that this function is defined for measures that increase according to the perception of fineness but for those that decreases, the function needs to be changed appropriately

Table 2: Fitting errors related to each pair of measures and parameter values for the pair with least error

Measure x	Measure y	Error
FD	Amadasun	0.1095
Correlation	ED	0.1332
Correlation	Tamura	0.1339
Amadasun	Correlation	0.1354
Abbadeni	Tamura	0.1373
Amadasun	ED	0.1431
Correlation	Abbadeni	0.1433
Abbadeni	Amadasun	0.1444
Correlation	FD	0.1455
Abbadeni	FD	0.1475
Tamura	Amadasun	0.1529
FD	Tamura	0.1634
ED	Tamura	0.1636
ED	FD	0.1764
Abbadeni	ED	0.1799

Parameters for FD+Amadasun	
a_9	1.11686
a_8	-0.00076
a_7	0.00000
a_6	-0.00853
a_5	-11.3885
a_4	0.06165
a_3	0.22887
a_2	36.8053
a_1	-1.26158
a_0	-34.7570

3 Results

In this section, the membership function $\mathcal{T}_k$ with least fitting error (obtained for the pair of measures FD+Amadasun and defined by the parameter values shown in Table 2) will be applied in order to analyze the performance of the proposed model. For this purpose, some experiments using images with different associated fineness presence degrees will be performed, analyzing the goodness of the fuzzy set to represent the perception of fineness. Moreover, these results will be compared with those obtained by using directly the measure values in order to verify the improvements introduced by the model.

Let's consider Figure 2(A) corresponding to a mosaic made by several images, each one with a different increasing perception degree of fineness. The perception degree of fineness for each subimage has been calculated using the proposed model and the results are shown in Figure 2(B) where a white grey level means maximum perception of fineness, while a black one corresponds to no perception of fineness. The numeric values are also shown on each subimage: v is the value of the human assessment and f is the computed degree. It can be noticed that our model captures the evolution of the perception degrees of fineness.

Figure 2(C) shows a mapping from the original image to its "fineness" membership degree using the proposed model. For each pixel in the original image, a centered window of size 32×32 has been analyzed and its fineness membership degree has been calculated. Thus, Figure 2(C) represents the degree in which the human perceives the texture, with a white grey level meaning maximum perception of fineness, and a black one meaning no perception of fineness (i.e., maximum perception of coarseness). The histogram of this image is shown beside it. We can see four clearly differ-

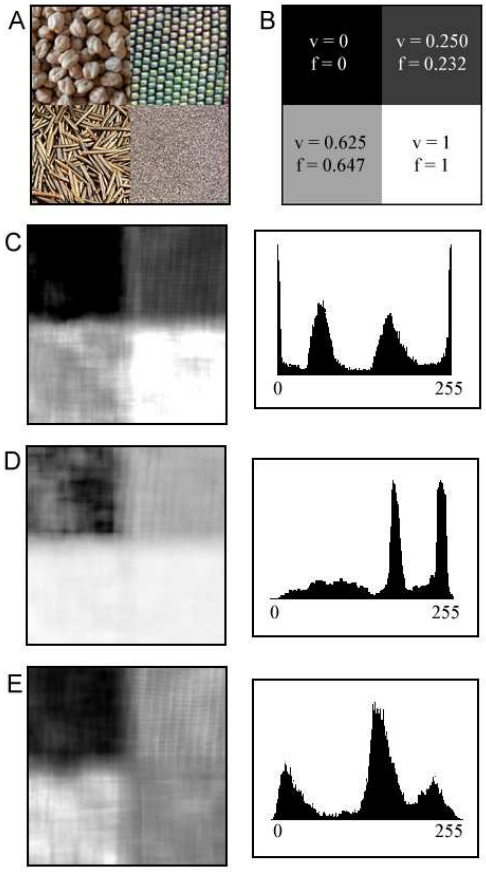


Figure 2: Results for a mosaic image. (A) Original mosaic image (B) Degree of fineness obtained with the proposed model for the pair of measures FD+Amadasun (C)(D)(E) Fineness membership degree of each pixel obtained with the proposed model and with the individual measures Amadasun and FD, respectively, and histograms of these images

enced peaks, corresponding to the four different fineness degrees in the image.

Figures 2(D) and 2(E) show a mapping from the original image using the Amadasun and FD measures directly, without any model. In this case the grey level of each pixel represents the obtained measure value instead of the perception degree of fineness so these images don't provide as much perceptual information as Figure 2(C). Furthermore, their histograms show that the four peaks can't be obtained with both measures separately.

Figure 3 presents an example where the proposed fuzzy sets have been employed for pattern recognition. In this case, the figure shows a microscopy image (Figure 3(A)) corresponding to the microstructure of a metal sample [19]. The lamellae indicates islands of eutectic, which are to be separated from the uniform light regions. The brightness values in regions of the original image are not distinct, so

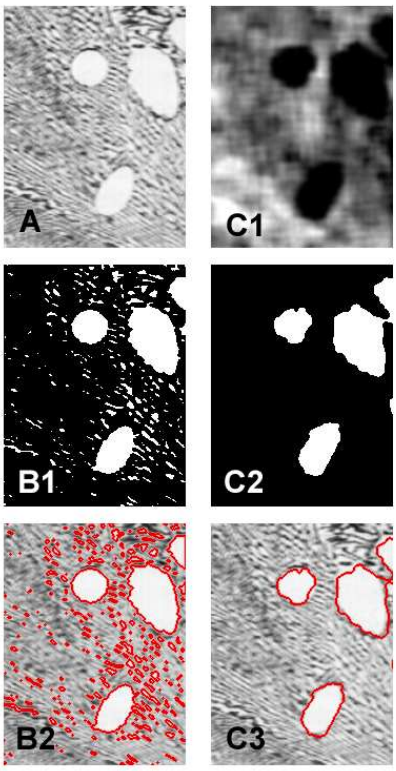


Figure 3: Example of pattern recognition using the proposed model (A) Original image (B1) Binary image obtained by thresholding the original image (B2) Region outlines of B1 superimposed on original image (C1) Fineness membership degrees obtained with our model from the original image (C2) Binary image obtained by thresholding C1 (C3) Region outlines of C2 superimposed on original image

texture information is needed for extracting the uniform areas. This fact is showed in Figure 3(B1,B2), where a thersholding on the original image is displayed (homogeneous regions cannot be separated from the textured ones as they "share" brightness values).

Figure 3(C1) shows a mapping from the original image to its "fineness" membership degree. In this case, we use a window of size 20×20. Thus, Figure 3(C1) represents the degree in which the human perceives the texture and it can be noticed that uniform regions correspond to areas with low degrees of fineness (i.e., high coarseness), so if only the pixels with fineness degree lower than 0.1 are selected (which it is equivalent to a coarseness degree upper than 0.9), the uniform light regions emerge with ease (Figure 3(C2,C3)).

4 Conclusions

In this paper, fuzzy sets for fineness perception have been defined on the domain of certain measures. The membership function associated to the fuzzy set has been learnt by finding the functional relationship between a certain measure (automatically computed over the image) and the presence degree of the textural concept. The measures have been chosen so that they collect information of the textural fineness under study whereas the presence degree has been obtained by performing a poll with humans and aggregating their assessments by means of OWA operators. Satisfactory performance has been found when using the different fuzzy sets with real images and the results show a high level of connection with the assessments given by subjects.

References

[1] N. Abbadeni, N. Ziou, and D.S. Wang. Autocovariance-based perceptual textural features corresponding to human visual perception. In *Proc. of 15th International Conference on Pattern Recognition*, volume 3, pages 901–904, 2000.

[2] M. Amadasun and R. King. Textural features corresponding to textural properties. *IEEE Transactions on Systems, Man and Cybernetics*, 19(5):1264–1274, 1989.

[3] A.E. Beaton and J.W. Tukey. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16:147–185, 1974.

[4] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[5] K.I. Chang, K.W. Bowyer, and M. Sivagurunath. Evaluation of texture segmentation algorithms. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 57, pages 294–299, 2003.

[6] X. Dai and J. Maeda. Unsupervised segmentation of textured color images using fuzzy homogeneity decision. In *Proc. 4th International Symposium on Video/Image Processing and Multimedia Communications*, pages 75–78, 2002.

[7] M.M. Galloway. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4:172–179, 1975.

[8] M. Hanmandlu, V. K. Madasu, and S. Vasikarla. A fuzzy approach to texture segmentation. In *Proc. International Conference on Information Technology:*

Coding and Computing, volume 1, pages 636–642, 2004.

[9] R.M. Haralick. Statistical and structural approaches to texture. *Proceedings IEEE*, 67(5):786–804, 1979.

[10] G. Karmakar, L. Dooley, and M. Murshed. Fuzzy rule for image segmentation incorporating texture features. In *Proc. of 2002 International Conference on Image Processing*, volume 1, pages I–797–I–800, 22-25 Sept. 2002.

[11] S.I. Kim, K.C. Choi, and D.S. Lee. Texture classification using run difference matrix. In *Proc. of IEEE 1991 Ultrasonics Symposium*, volume 2, pages 1097–1100, December 1991.

[12] S. Kulkarni and B. Verma. Fuzzy logic based texture queries for cbir. In *Proc. 5th International Conference on Computational Intelligence and Multimedia Applications*, pages 223–228, 2003.

[13] H.C. Lin, C.Y. Chiu, and S.N. Yang. Finding textures by textual descriptions, visual examples, and relevance feedbacks. *Pattern Recognition Letters*, 24(14):2255–2267, 2003.

[14] S.D. Newsam and C. Kammath. Retrieval using texture features in high resolution multi-spectral satellite imagery. In *Data Mining and Knowledge Discovery: Theory, Tools, and Technology VI, SPIE Defense and Security*, April 2004.

[15] S. Peleg, J. Naor, R. Hartley, and D. Avnir. Multiple resolution texture analysis and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):518–523, 1984.

[16] M. Potrebić. Iterative fuzzy rule base technique for image segmentation. In *Proc. of 7th seminar on Neural Network Applications in Elecrical Engineering*, pages 221–224, 23-25 Sept. 2004.

[17] A.R. Rao and G.L. Lohse. Identifying high level features of texture perception. *Graphical Models and Image Processing*, 55(3):218–233, 1993.

[18] Todd R. Reed and J.M. Hans Du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372, 1993.

[19] J.C. Russ. *The Image Processing Handbook*. CRC Press and IEEE Press, third edition, 1999.

[20] A. Shackelford. A hierachical fuzzy classification approach for high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1920–1932, 2003.

[21] C. Sun and W.G. Wee. Neighboring gray level dependence matrix for texture classification. *Computer Vision, Graphics and Image Processing*, 23:341–352, 1983.

[22] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. on Systems, Man and Cybernetics*, 8:460–473, 1978.

[23] C.B. Wang, H.B. Wang, and Q.B. Mei. Texture segmentation based on an adaptively fuzzy clustering neural network. In *Proc. of 2004 International Conference on Machine Learning and Cybernetics*, volume 2, pages 1173–1176, 2004.

[24] J.S. Weszka, C.R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Trans. on SMC*, 6:269–285, 1976.

[25] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. on SMC*, 18(1):183–190, 1988.

[26] H. Yoshida, D.D. Casalino, B. Keserci, A. Coskun, O. Ozturk, and A. Savranlar. Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images. *Physics in Medicine and Biology*, 48:3735–3753, 2003.

HISTOGRAMS ON LINGUISTIC COLOR VARIABLES BASED ON FUZZY NATURALS

J. Chamorro-Martínez, D. Sánchez

Dpt. of Computer Science and Artificial Intelligence
University of Granada
{jesus,daniel}@decsai.ugr.es

J.M. Soto-Hidalgo

Dpt. of Computer Architecture, Electronics and Electronic Technology
University of Córdoba
jmsoto@uco.es

Abstract

In this paper we introduce a new fuzzy histogram approach based on fuzzy naturals. In our proposal, histograms are functions on a fuzzy color space, where the occurrences of a given fuzzy color are counted by means of a new definition of fuzzy natural number (represented by means of a probability distribution on $\mathbb{N}$). This approach to histograms avoids the well-known disadvantages of the ordinary sigma-count. We illustrate the potential application of the proposal by applying it to the problem of dominant color selection.

Keywords: Fuzzy colors, Fuzzy histograms.

1 INTRODUCTION

In some image processing applications it is important to take into account the particular perception of colors of a certain user. However, there is a semantic gap between the human perception of colors and the data stored in a pixel, i.e., the three (or more) components corresponding to the primary colors in a given color space (e.g., red green and blue for RGB). To face this problem, several authors have proposed to employ fuzzy sets [10, 11]. In these approaches, colors in an image are linguistic labels represented as fuzzy subsets of triplets of a certain color space [19, 21]. This is important in order to determine the semantic interpretation of colors for human-machine interaction, or to take into account the similarity between colors in a color space as perceived by the user [18].

In the scope of image analysis and processing, the histogram is the basis for numerous techniques for image restoration, enhancement, segmentation, retrieval, etc [16].

[1]This work has been partially supported by the Andalusian Government under the TIC1570 and TIC249

The histogram of an image is a function $h(\mathbf{c_k}) = n_k$ where $\mathbf{c_k} = [x, y, z]$ is a color and n_k is the number of pixels in the image having the color $\mathbf{c_k}$. It is common to normalize a histogram by dividing each of its values by the total number of pixels, obtaining an estimate of the probability of occurrence of a color $\mathbf{c_k}$.

Working with fuzzy colors suggests to extend the notion of histogram to "fuzzy histogram" in order to manage the imprecision in (i) color description by means of fuzzy colors and, consequently, in (ii) fuzzy color counting. In this sense, a fuzzy histogram will give us information about the frequency of each fuzzy color.

In the literature there are several proposals which define histograms over a set of fuzzy colors [6, 9, 15]. One drawback of most of these proposals is that they work only with intensities. In addition, the counting of fuzzy colors is performed by using the sigma-count (i.e., the sum of membership degrees). However, the sigma-count is not a suitable measure of cardinality in many applications, as it has been recognized by several authors [8, 12]. Proposals based on the sigma-count summarize the counting in a single number, so they do not represent the imprecision of the count itself.

To solve the problems above, after introducing a formalization of the concept of fuzzy color space, in this paper we propose to "count" the occurrences of a given fuzzy color by means of the new definition of fuzzy natural numbers introduced in [4]. In this definition, a natural number is represented by means of a probability distribution on $\mathbb{N}$.

The rest of the paper is organized as follows. Section 2 formalize the concepts of fuzzy color and fuzzy color space. In section 3 the fuzzy natural numbers are introduced. On the basis of the fuzzy naturals, the fuzzy histogram over fuzzy colors is defined in section 4. An application to dominant color selection is presented in section 5 to illustrate the potential applications and performance of the new definition of fuzzy histogram. Some examples are showed in section 6 and, finally, the main conclusions are summarized in section 7.

2 FUZZY MODELLING OF COLORS

In this section, the notions of fuzzy color (section 2.2) and fuzzy color space (section 2.3) are introduced as an extension of the classical concepts of color and color space (section 2.1).

2.1 COLOR AND COLOR SPACES

For representing colors, several color spaces can be used. In essence, a color space is a specification of a coordinate system and a subspace within that system where each color is represented by a single point. The most commonly used color space in practice is RGB because is the one employed in hardware devices (like monitors and digital cameras). It is based on a cartesian coordinate system, where each color consists of three components corresponding to the primary colors red, green, and blue.

Nevertheless, it is well known that RGB it is not always the most adequate space for color image analysis. Furthermore, the color components of this space do not have an intuitive interpretation according to the human perception of color. Instead, other color spaces based on human perception (HSI, HSV or HSL) seem to be a better choice [16]. In these perceptual spaces, the hue component (H) represents the color tone (for example, red or blue), saturation (S) is the amount of color (for example, vivid red or pale red) and the third component (called intensity, value or lightness) is the amount of light (it allows to distinguish between a dark color and a light color). Therefore, if linguistic labels are needed for color description, this kind of color spaces are more suitable than RGB, linear combination of RGB (like CMY, YCbCr, YUV, etc.) or perceptually uniform color spaces (like CIELa*b*, CIELuv, etc.).

2.2 FUZZY COLOR

In order to manage the imprecision in color description, we introduce the following definition of fuzzy color:

Definition 2.1 *A fuzzy color $\widetilde{C}$ is a normalized fuzzy subset of colors.*

As previously explained, colors can be represented as a triplet of real numbers corresponding to coordinates in a color space. Hence, a fuzzy color can be defined as a normalized fuzzy subset of points of a color space. From now on, we shall note XYZ a generic color space with components X, Y and Z[1], and we shall assume that a color space XYZ, with domains D_X, D_Y and D_Z of the corresponding color components is employed. This leads to the following more specific definition:

Definition 2.2 *A fuzzy color $\widetilde{C}$ is a linguistic label whose semantics is represented in a color space XYZ by a normalized fuzzy subset of $D_X \times D_Y \times D_Z$.*

Notice that it is possible to represent the same fuzzy color by means of different fuzzy subsets on different color spaces, provided they have the same expressive power. For example, a fuzzy color *red* could be represented as a fuzzy subset with semantics "approximately [255,0,0]" if we employ the RGB color space and also as a fuzzy subset with semantics "approximately [0,1,128]" if we use the HSV space.

As in the case of colors, that can be described by means of three precise component values, it is possible (though not obligatory) to define a fuzzy color by means of fuzzy subsets on the domain of each component. We introduce the following definition:

Definition 2.3 *A fuzzy color component $\widetilde{C}_X$ (resp. $\widetilde{C}_Y$, $\widetilde{C}_Z$) is a linguistic label whose semantics is represented by a normalized fuzzy subset of D_X (resp. D_Y, D_Z).*

By using these concepts, a fuzzy color $\widetilde{C}$ can be defined and represented in practice by a triple $[\widetilde{C}_X, \widetilde{C}_Y, \widetilde{C}_Z]$, where $\widetilde{C}_X$, $\widetilde{C}_Y$, and $\widetilde{C}_Z$ are fuzzy color components of $\widetilde{C}$. This way, the fuzzy subset representing a fuzzy color can be obtained by combining the corresponding fuzzy color components in a suitable way. Thus, for every crisp color $[x, y, z]$, its membership degree to $\widetilde{C}$ is defined as

$$\widetilde{C}(x, y, z) = \bigwedge\{\widetilde{C}_X(x), \widetilde{C}_Y(y), \widetilde{C}_Z(z)\} \qquad (1)$$

with $\bigwedge$ being a t-norm, usually the minimum. From now on, we shall note $I_{\widetilde{C}}$ the fuzzy subset of pixels with color $\widetilde{C}$

2.3 FUZZY COLOR SPACE

For extending the concept of color space to the case of fuzzy colors, and assuming a fixed color space XYZ, with D_X, D_Y and D_Z being the domains of the corresponding color components, the following definition is introduced:

Definition 2.4 *A fuzzy color space $\widetilde{XYZ}$ is a set of fuzzy colors that define a partition of $D_X \times D_Y \times D_Z$.*

As in the case of single colors, one very convenient way of defining and representing a fuzzy color space is to employ

[1] Although we are assuming a three dimensional color space, the proposal can be easily extended to color spaces with more components.

a fuzzy space for each component. We introduce this idea in the following definition:

Definition 2.5 *A fuzzy space over the component X (resp. Y, Z) is a set of fuzzy color components that define a partition of D_X (resp. D_Y, D_Z).*

This procedure has several advantages. First, less linguistic labels have to be defined. Second, we can represent and work with every component of the fuzzy color individually (this is specially interesting in perceptual color spaces like HSI where, for example, we could query for colors with red hue). Finally, the linguistic labels associated to fuzzy colors can be obtained by combining the corresponding linguistic labels of each component (for example, in the case of HSI color space, the color label "red-saturated-bright" is obtained by combining hue, saturation and intensity labels).

3 FUZZY NATURALS

In order to obtain an histogram, which give us information about the frequency of each color in the image, we need to count the number of pixels having a particular color. In our case, we must count pixels having fuzzy colors. For this purpose we employ the concepts of fuzzy natural.

In this section we explain the definition of fuzzy naturals introduced in [4] (section 3.2). For defining fuzzy naturals the definition of cardinality of a fuzzy set will be used (section 3.1)

3.1 FUZZY CARDINALITY

The definition of fuzzy natural can be obtained directly from the definition of cardinality of a fuzzy set. In the crisp case, as it is well known, naturals are the possible cardinalities of crisp sets. The usual (in the sense of most employed) way to extend cardinality to the fuzzy case is the scalar sigma-count, that can defined for any fuzzy set F with membership function $F : X \to [0, 1]$ as

$$sc(F) = \sum_{x \in X} F(x) \tag{2}$$

However, the sigma-count is not a suitable measure of cardinality in many applications, as it has been recognized by several authors [8, 12]. In the particular problem of defining fuzzy naturals, sigma-count is counterintuitive, since sc is not a natural number in general. Even if the result is a natural number, we can obtain the same result in very different situations, thus loosing information. This is the case if we consider two situations like having 100 pixels compatible with a color $\widetilde{C}$ to a degree 0.1, and having 10 pixels compatible with $\widetilde{C}$ to a degree 1; in both cases, the result of the sigma count applied to the fuzzy set of pixels compatible with $\widetilde{C}$ is 10. On the basis of these problems, it is widely accepted that the best way to represent the cardinality of a fuzzy set is by means of a fuzzy subset of the naturals [3, 7].

Most of the approaches consider that such fuzzy subsets of the naturals must be convex. However, in [3] we showed that in some cases this is counterintuitive. Consider for example the fuzzy set given by $A = 1/x_1 + 0.5/x_2 + 0.5/x_3$. The cardinality of A could be one (because x_1 belongs to A for sure) or, if we relax our criterion to accept elements in A, the cardinality could be three (accepting x_2 and x_3 belong to A as well). However, the cardinality cannot be two, since if $x_2 \in A$ then $x_3 \in A$ and vice versa. This way, the cardinality is not convex. In addition, this example illustrates that the sigma-count is not always a good measure since $sc(A) = 2$. In summary, sc is not a good cardinality approach for the definition of fuzzy natural numbers since it is a real number and can give counterintuitive results.

Several authors [3, 13, 14] have pointed out that the possible cardinalities of a fuzzy set are the cardinalities of its α-cuts, since these are the possible crisp representatives of the fuzzy set. In our previous example the possible cardinalities of A are 1 or 3 since its possible α-cuts are $\{x_1\}$ and $\{x_1, x_2, x_3\}$. In general, this is in accordance with the principle that there is a single membership scale for fuzzy sets, i.e., $F(x) = F(x')$ means that x is F just like x', so we cannot consider $x \in F$ and $x' \notin F$ at the same time, and either both or none of them contribute to the cardinality of F.

As we have just seen, the sigma-count does not necessarily correspond to the cardinality of an α-cut ($sc(A) = 2$). On the contrary, there are two fuzzy cardinalities that comply with this idea, calculating the possible cardinalities of a fuzzy set as the cardinalities of its α-cuts, and assigning to each cardinality a degree of representativity, as we detail in the following subsections:

3.1.1 Zadeh's first fuzzy cardinality

Zadeh introduced in 1979 a non-convex fuzzy measure for the cardinality of a fuzzy set F to be [20]

$$\mu_{CARD(F)}(k) = sup\{\alpha \mid |F_\alpha| = k\} \tag{3}$$

For example, for $A = 1/x_1 + 0.5/x_2 + 0.5/x_3$ we have $\mu_{CARD(A)} = 1/1 + 0.5/3$. In general, $\mu_{CARD(\cdot)}$ is a fuzzy subset of $\mathbb{N}$ verifying it is normal, its core consists of one value only, and it is strictly decreasing in its support.

3.1.2 The Fuzzy Cardinality ED

In [3] the fuzzy cardinality of a fuzzy set F is defined as

$$ED(F) = \sum_{\alpha_i \in \Lambda(F)} (\alpha_i - \alpha_{i+1}) \, / \, |F_{\alpha_i}| \qquad (4)$$

where $\Lambda(F) = \{\alpha_1, \ldots, \alpha_q\} = \{F(x_i) \,|\, x_i \in supp(F)\} \cup \{1\}$ is the set of representative α-cuts of F, with $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_q > \alpha_{q+1} = 0$. For example, $\Lambda(A) = \{1, 0.5\}$ and $ED(A) = 0.5/1 + 0.5/3$.

The measure ED is a probability distribution on $\mathbb{N}$. It is possible to show that the converse is also true, i.e., any probability distribution on $\mathbb{N}$ is the cardinality of some fuzzy set as given by ED. The proposition and the corresponding proof can be found in [4].

In [3] it is shown that ED is the basic probability assignment of the possibility distribution $\mu_{CARD(\cdot)}$, and hence both cardinalities are isomorphic. However, ED has been shown to be more appropriate than $\mu_{CARD(\cdot)}$ when using the fuzzy cardinality in practice, for example in the evaluation of quantified sentences [5].

3.2 FUZZY NATURAL DEFINITION

The idea that cardinality and natural numbers are basically identical was employed in [4] to define fuzzy natural numbers as follows:

Definition 3.1 *A fuzzy natural N is any probability distribution on $\mathbb{N}$ of the form*[2]

$$N = \sum_{n_i \in \mathbb{N}} p_i / n_i \qquad (5)$$

with $\sum_{n_i \in N} p_i = 1$.

As an example, let us consider the fuzzy natural

$$N = 0.2/0 + 0.6/2 + 0.2/4 \qquad (6)$$

Having N objects means that the number of objects is 0 with probability 0.2, 2 with probability 0.6, and 4 with probability 0.2.

In the context of histograms, this definition allows us to count the number of pixels that comply with a certain fuzzy color $\widetilde{C}$ as a fuzzy natural, corresponding to the cardinality of the fuzzy subset of pixels whose color is $\widetilde{C}$. This fuzzy natural can be understood as a (fuzzy) estimation of the counting of $\widetilde{C}$ in the image.

[2]Notice that we are using in the definition of N the usual notation $\sum$ for fuzzy sets on a discrete referential, while in $\sum_{n_i \in N} p_i = 1$, $\sum$ means addition.

4 FUZZY HISTOGRAMS

We define a fuzzy histogram as follows:

Definition 4.1 *Let $\widetilde{XYZ}$ be a fuzzy color space. A fuzzy histogram is a function h that assigns a fuzzy natural to every fuzzy color in $\widetilde{XYZ}$, i.e.,*

$$h(\widetilde{C}) = N_{\widetilde{C}} \qquad (7)$$

with $N_{\widetilde{C}}$ a fuzzy natural.

Notice that this definition extends to the fuzzy case the idea of histogram as a function that assign a count, given by a natural number, to every color. In the fuzzy context, we have fuzzy colors (instead crisp ones) and fuzzy naturals for imprecise counting.

Fuzzy histograms are descriptions of the fuzzy counting for a pixel having a color $\widetilde{C}$, for every color in a fuzzy color space. However, if this is the final information to be provided to a user, it is convenient to summarize it so that it is easier to understand. Of course, any possible summary causes a loss of information. We can consider several options as a summary of the fuzzy frequency of a fuzzy color:

- Linguistic summary of the fuzzy naturals. In this option, a triangular fuzzy number (in the classical sense of "possibility distribution on the set or real numbers") $\mathcal{R}_{\widetilde{C}}$ is calculated as an approximation of the fuzzy natural $N_{\widetilde{C}}$. The number is obtained by approximation techniques as the number that yields a maximum accomplishment degree with $N_{\widetilde{C}}$ with minimum entropy. This technique is described in [1].

- The interval of natural numbers $[a, b]$ with $a = \min\{i \in support(N_{\widetilde{C}})\}$ and $b = \max\{i \in support(N_{\widetilde{C}})\}$

- The natural number with higher probability in the probability distribution $N_{\widetilde{C}}$. Notice that this number, being an scalar, is obtained from the cardinality of an α-cut of $I_{\widetilde{C}}$, and in this sense is coherent with the ideas explained at the beginning of section 3.

- The natural number l obtained as the cardinality of the 0.5-cut of the fuzzy subset of pixels $I_{\widetilde{C}}$. This scalar value is the definition of approximation of the cardinality of a fuzzy set proposed by Ralescu [13]. In addition, if the fuzzy color space is a partition in the sense of Ruspini, the 0.5-cut corresponds to a crisp partition of the color space, and the values obtained for all the fuzzy colors correspond to a crisp histogram based on this crisp partition.

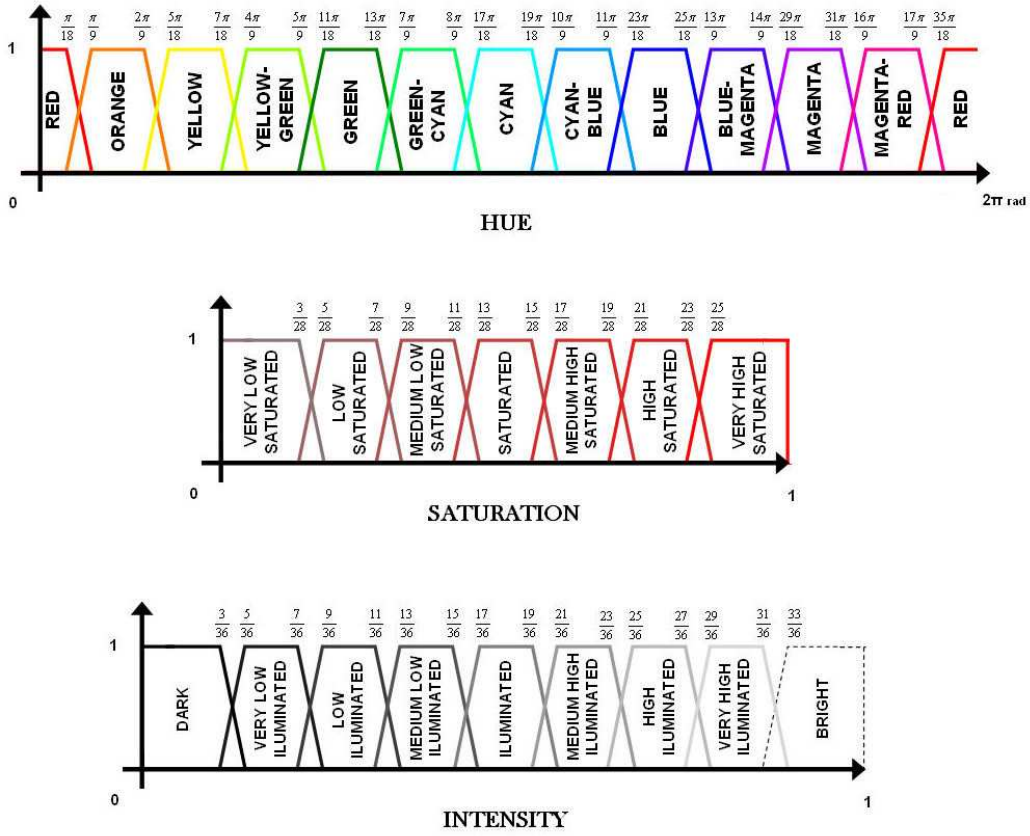


Figure 1: The fuzzy HSI color space used in the experiments.

Let us remark once more that the latter are different ways to obtain a summary that can be useful for a better understanding of the fuzzy histogram by an user, at the cost of loosing information. However, if the histogram is just an intermediate step in a calculation, we can extend the operations to the fuzzy case by operating on α-cuts and then summarizing the final result [4]. This option preserves all the information until the final step. The option of extending the operations to the fuzzy case by operating on one of the previously suggested summaries discards many information, and it is reasonable in our opinion only when time restrictions force us to obtain a less accurate, but faster, calculation.

Notice that we are not suggesting that the calculations performed by α-cuts are necessarily slow; the idea is that performing a certain operation in every α-cut takes more time than performing the same operation on a single scalar approximation. Depending on the time expended by each operation and the number of α-cuts employed (the latter can be fixed in order to obtain an approximate, but more accurate, result; a typical value is 20), we may choose one or another alternative.

5 APPLICATION TO DOMINANT COLOR SELECTION

In this section we illustrate the potential applications and performance of the new definition of fuzzy histogram by applying it to select a set of dominant fuzzy colors. Dominant colors arise as a powerful tool for describing the representative colors in an image. In fact, they are a very efficient tool for retrieving images in large databases.

Intuitively, a color is dominant to the extent it appears frequently in a given image. Therefore, the use of fuzzy histograms comes up as a suitable tool for finding the dominant fuzzy colors.

Let h be a fuzzy histogram for a color image on a fuzzy color space $\widetilde{XYZ}$. Given a color $\widetilde{C}$, its fuzzy frequency is given by $h(\widetilde{C}) = N_{\widetilde{C}}$ with $N_{\widetilde{C}} = \sum_{n_i \in \mathbb{N}} p_i/q_i$ being a fuzzy natural number in the sense of definition 3.1. We define the degree of dominance of the fuzzy color $\widetilde{C}$ as

$$Dom(\widetilde{C}) = \sum_{q_i > T} p_i \tag{8}$$

with $T \in (0, 1]$ being a threshold for selecting significant frequencies.

As set of dominant fuzzy colors we will select those ones with higher dominant degree (we can fix a number of dominant colors or select those ones which dominant degrees are greater than a threshold)

6 RESULTS

In this section we show the results obtained by applying the proposed technique to several examples. Firstly, the fuzzy histogram will be calculated over several images. Secondly, these histograms will be used to select dominant fuzzy colors.

In our experiments, the HSI color space have been used. To define the fuzzy color space $\widetilde{HSI}$ we have employed the fuzzy spaces for hue, saturation, and intensity proposed in [2] that are shown in figure 1. This proposal used as reference the Munsell color space [17] which divided in 12, 7 and 9 intervals the hue, saturation, and intensity respectively. Each one of these intervals are fuzzified using a trapezoidal function to define the memberships functions, obtaining a fuzzy partition in the sense of Ruspini.[3]

Figure 2 shows two images, the first one containing eight colors, each one compatible with degree 1 with a fuzzy color obtained by combining the fuzzy saturation label "High saturated" and the fuzzy intensity label "Illuminated", with eight different fuzzy hue labels, specifically all the hue labels except "Yellow-Green", "Cyan", "Blue-Magenta", and "Red" in figure 1. In the second one we have four colors compatible with the same saturation and intensity labels to degree 1. Each one of these four colors is compatible to a degree 0.5 with two of the eight fuzzy hue labels mentioned before. In both images, pixels are equidistributed among colors.

In both images, the set of fuzzy colors with probability greater than 0 is the set of eight fuzzy colors mentioned above. It is easy to see that the sigma-count histogram is the same for both images, specifically the probability is 1/8 for each fuzzy color. However, the perception of the colors that appear in the image and the corresponding frequencies is very different in both images.

On the contrary, the histogram based on our definition of fuzzy frequencies yield a different result in each case. In the first one, since we are in a crisp case, the result is the same provided by the sigma-count histogram, i.e., probability 1/8 (given as a probability distribution 1/(1/8)) for each fuzzy color. On the contrary, our approach yields for the second image a fuzzy histogram where, for each one of the eight color, the fuzzy probability is represented by the probability distribution $0.5/0 + 0.5/(1/4)$. This is in accordance with our intuition since each fuzzy color is compatible to a degree 0.5 with one of the colors that appear in the image and, as a consequence, we have a probability 0.5 that the fuzzy label appears in the image, and the same probability that it does not appear. Hence, the same probabilities are assigned to the value 1/4 in the first case, and 0 in the second case.

[3]The problem of non-representability in semichromatic (undefined hue) and achromatic (undefined hue and saturation) is taken into account in the sense proposed in [2]

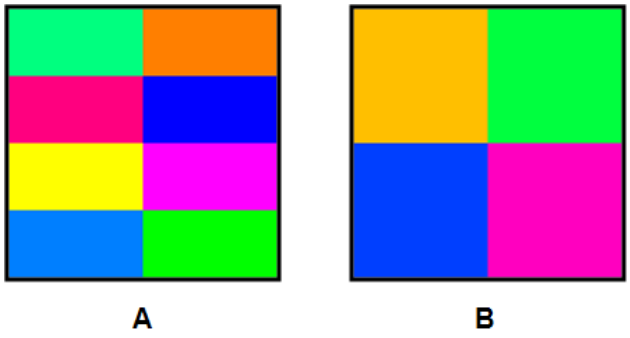


Figure 2: Multicolored images. (A) Image containing eight colors, each one compatible with degree 1 with a fuzzy color (B) Image containing four colors, each one compatible with degree 0.5 with a fuzzy color

Figures 3(B) to 3(I) show the fuzzy histogram for image 3(A). For plotting the histogram, a graph has been drawn for each intensity label. The X axis of these graphs shows the color labels ordered by hue and, within a given hue, by saturation. The Y axis shows the support of $h(\widetilde{C})$ normalized by the number of pixels n. More specifically, the graph plots pairs $(\widetilde{C}, n_i/n)$ where n_i are the crisp naturals that appear in the probability distribution $N_{\widetilde{C}}$. In addition, a dotted line is drawn between the maximum and minimum values a and b of the support, $a = \min\{i \in support(N_{\widetilde{C}})\}$ and $b = \max\{i \in support(N_{\widetilde{C}})\}$. This dotted line shows the interval $[a, b]$ that corresponds to one of the ways of summarizing $h(\widetilde{C})$, as explained in section 4. The values $N_{\widetilde{C}}(n_i)$ are not displayed.

We have compared the results obtained in this and other images with the crisp histogram for a crisp partition obtained as the set of 0.5-cuts of the fuzzy colors. Though the latter is a summary and many information is lost, we have found that higher (resp. lower) crisp frequencies in the crisp histogram correspond to higher (resp. lower) values in the support of fuzzy frequencies and, in this sense, the results are coherent with what is expected. However, the fuzzy histogram contains much more information since, contrary to the crisp one, it reflects the counting imprecision.

The fuzzy histogram shows greater frequency values in the support of the most frequent fuzzy colors in the image. For example, the highest values correspond to the fuzzy colors Black (Dark intensity) and the brown-like colors [Orange,Medium-Low-Saturated,Illuminated] and [Orange,Saturated,Illuminated], these three fuzzy colors corresponding to the background. The next colors in terms

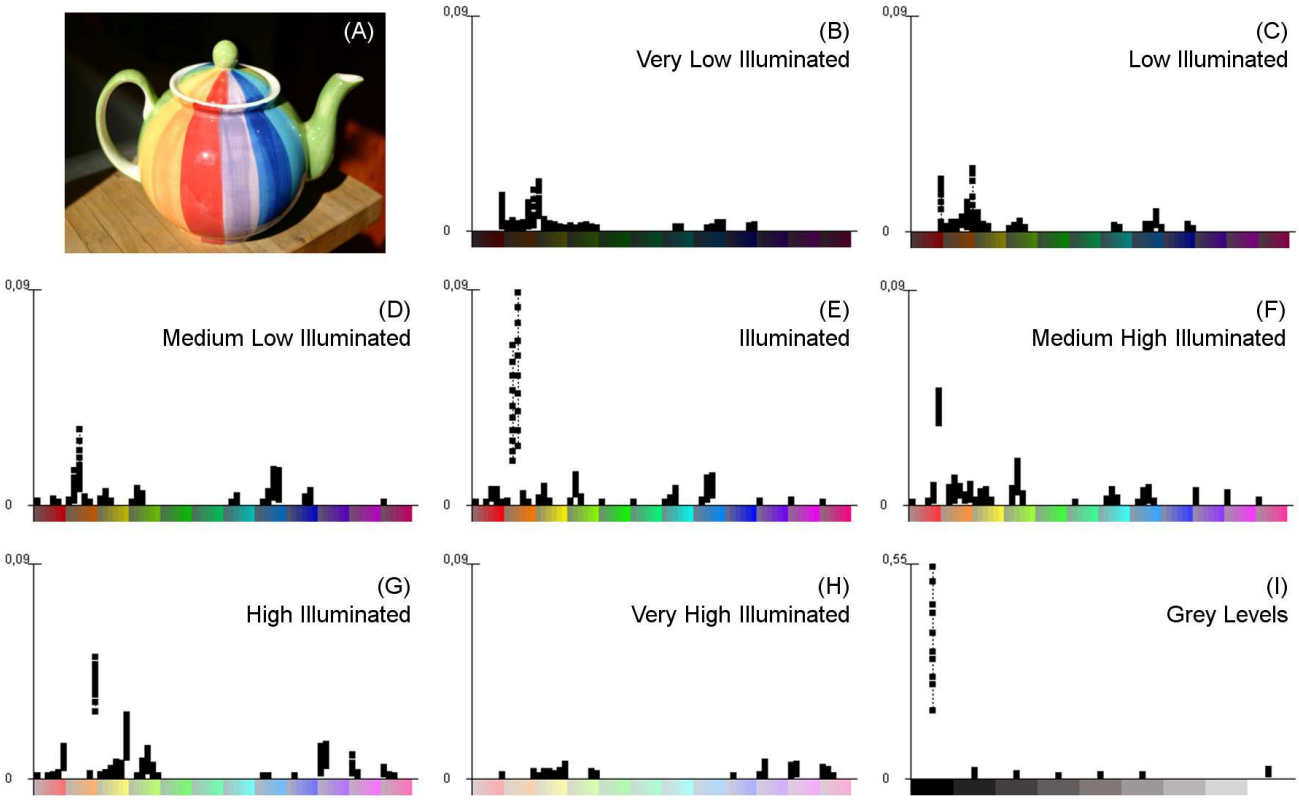


Figure 3: (A): color image. (B)-(H): fuzzy histograms for colors with different intensity labels except “Dark” and “Bright”. In each histogram, the 12x6 colors (“Very low saturated” colors, i.e., fuzzy grey levels, are not considered) are ordered first by hue and then by saturation. (I): Fuzzy colors corresponding to fuzzy grey levels.

of high frequency are those of the teapot.

Figure 4 shows the degree of dominance of those colors with degree of dominance greater than 0 for the image in Figure 3(A). The dominance has been calculated by using a threshold $T = 0.015$. The results are coherent with the information in the fuzzy histogram. With respect to our perception of dominance in the image, we must take into account that the fuzzy color space employed in the chromatic area of the HSI color space is a equidistributed fuzzy partition. However, as pointed out by several authors, we tend to employ a non-equidistributed partition in which a single fuzzy color covers approximately the support of the fuzzy colors with hue “Cyan”, “Cyan-Blue” and “Blue”. Something similar happens with the fuzzy colors with hue labels around “Green”. The consequence in this example is that the colors we see as blue are distributed between the fuzzy colors with hue “Cyan”, “Cyan-Blue” and “Blue”, so the degree of dominance of what we see as blue is maybe less than expected. On the contrary, if we consider a fuzzy color space with a label “Blue” covering the support of the three above mentioned blue-based colors, the dominance of the new label would be similar to other fuzzy colors with “Red”, “Orange” or “Yellow” fuzzy hue.

7 CONCLUSIONS

In this paper a novel definition of fuzzy histogram has been proposed. Firstly, a formalization of the concept of fuzzy color and fuzzy color space has been presented. Secondly, the occurrences of a given fuzzy color have been counted by means of a new definition of fuzzy natural number.

The proposal improves the classical sigma-count based histogram. In fact, our fuzzy histogram represents the imprecision of the count itself, contrary to sigma-count that summarizes the counting in a single number.

The performance of the proposal has been shown with several images, introducing some examples where histograms based on sigma-count are not a suitable representation of the amount of pixels painted with each color. In addition, we have illustrated the potential applications of the fuzzy histogram by applying it to the problem of dominant fuzzy colors selection.

References

[1] I. Blanco, D. Sánchez, J.M. Serrano, and M.A. Vila. A new proposal of aggregation functions: The linguistic summary. In T. Bilgiç, B. De Baets,

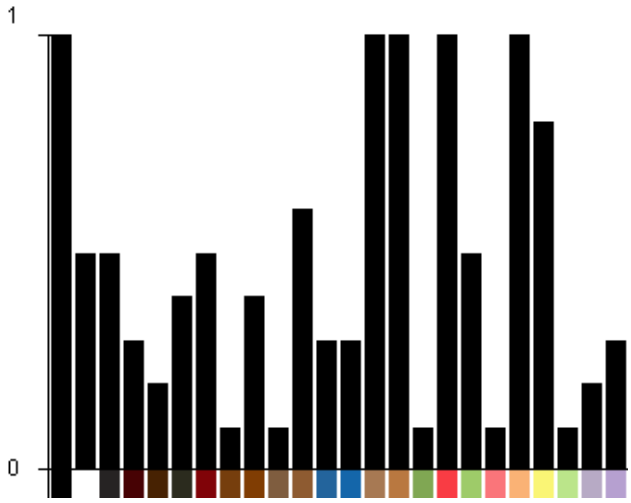


Figure 4: Degree of dominance of fuzzy colors in the image of figure 3(A). Only fuzzy colors with non-zero dominance are shown.

and O. Kaynak, editors, *Fuzzy Sets and Systems - Proceedings IFSA'03, LNAI 2715*, pages 127–134. Springer, 2003.

[2] J. Chamorro-Martínez, J.M. Medina, C. Barranco, E. Galán-Perales, and J.M. Soto-Hidalgo. Retrieving images in fuzzy object-relational databases using dominant color descriptors. *Fuzzy Sets and Systems*, 158(3):312–324, February 2007.

[3] M. Delgado, M.J. Martín-Bautista, D. Sánchez, and M.A. Vila. A probabilistic definition of a nonconvex fuzzy cardinality. *Fuzzy Sets and Systems*, 126(2):41–54, 2002.

[4] M. Delgado, M.J. Martín-Bautista, D. Sánchez, and M.A. Vila. Fuzzy integers: Representation and arithmetic. In *Proceedings of IFSA'05*, 2005.

[5] M. Delgado, D. Sánchez, and M.A. Vila. Fuzzy cardinality based evaluation of quantified sentences. *International Journal of Approximate Reasoning*, 23:23–66, 2000.

[6] A. Doulamis and N. Doulamis. Fuzzy histograms for efficient visual content representation: application to content-based image retrieval. In *IEEE International Conference on Multimedia and Expo*, pages 893–896, Aug 2001.

[7] D. Dubois and H. Prade. Fuzzy cardinality and the modeling of imprecise quantification. *Fuzzy Sets and Systems*, 16:199–230, 1985.

[8] D. Dubois, H. Prade, and T. Sudkamp. A discussion of indices for the evaluation of fuzzy associations in relational databases. In T. Bilgic, B. De Baets, and O. Kaynak, editors, *Fuzzy Sets and Systems - Proceedings IFSA'03, LNAI 2715*, pages 111–118. Springer, 2003.

[9] J. Han and Kai-Kuang. Fuzzy color histogram and its use in color image retrieval. *IEEE Transactions on Image Processing*, 11(8):944–952, 2002.

[10] L. Hildebrand and M. Fathi. Knowledge-based fuzzy color processing. *IEEE. Tran. on Systems, Man and Cybernetics. Part C*, 34(4):499–505, November 2004.

[11] G. Louverdis, I. Andreadis, and Ph. Tsalides. New fuzzy model for morphological colour image processing. In *IEEE Proc. Vis. Image Signal Processing*, volume 149, June 2002.

[12] M.J. Martín-Bautista, D. Sánchez, M.A. Vila, and H.L. Larsen. Measuring effectiveness in fuzzy information retrieval. In H.L. Larsen, J. Kacprzyk, S. Zadrozny, T. Andreasen, and H. Christiansen, editors, *Proceedings FQAS 2000, Advances in Soft Computing Series*, pages 396–402. Springer-Verlag, 2000.

[13] D. Ralescu. Cardinality, quantifiers and the aggregation of fuzzy criteria. *Fuzzy Sets and Systems*, 69:355–365, 1995.

[14] D. Rocacher and P. Bosc. The set of fuzzy rational numbers and flexible querying. *Fuzzy Sets and Systems*, 155(3):317–339, 2005.

[15] S. Romani, P. Sobrebilla, and E. Montseny. Obtaining the relevant colors of an image through stability-based fuzzy color histograms. In *IEEE International Conference on Fuzzy Systems*, volume 2, pages 914–919. St. Louis, Missouri (USA), May 2003.

[16] J.C. Russ. *The Image Processing Handbook*. CRC Press and IEEE Press, third edition, 1999.

[17] S.J. Sangwine and R.E.N. Horne. *The colour image processing handbook*. Chapman and Hall, 1998.

[18] M. Seaborn, L. Hepplewhite, and J. Stonham. Fuzzy colour category map for the measurement of colour similarity and dissimilarity. *Pattern Recognition*, 38(4):165–177, 2005.

[19] N. Sugano. Color-naming system using fuzzy set theorical approach. In *IEEE Int. Conference on Fuzzy Systems*, pages 81–84, 2001.

[20] L. A. Zadeh. A theory of approximate reasoning. *Machine Intelligence*, 9:149–194, 1979.

[21] H. Zhu, H. Zhang, and Y. Yu. Deep into color names: Matching color descriptions by their fuzzy semantics. *LNAI 4183*, pages 138–149, 2006.

Aplicación de un algoritmo de extracción de reglas difusas para minería de uso web

Cristóbal José Carmona del Jesus **Pedro González García**
Víctor Manuel Rivas Santos **María José del Jesus Díaz**

Depto. Informática. Universidad de Jaén, {ccarmona, pglez, vrivas, mjjesus}@ujaen.es

Resumen

El gran impulso de las nuevas tecnologías hace que la World Wide Web sea actualmente una de las fuentes de información más importantes en el momento. Dentro del campo de la minería de uso web, que analiza la información de la WWW, se aplican diferentes tipos de técnicas. En nuestra aproximación se procesa la base de datos MSNBC mediante una técnica evolutiva capaz de extraer conocimiento de interés sobre el uso de la web y expresarlo en forma de reglas difusas de descripción de subgrupos. Los resultados obtenidos en esta primera aproximación se caracterizan por la obtención de reglas que cubren la mayoría de los ejemplos con alto nivel de relevancia y confianza.

Palabras Clave: Minería de Uso Web, Patrones Frecuentes, MSNBC, Descubrimiento de Subgrupos, Sistemas Evolutivos Difusos.

1 Introducción

El uso de las nuevas tecnologías y la explosión de información que ha surgido a raíz de la World Wide Web, provoca que sea una de las fuentes de datos más importantes que existen en la actualidad. Por este motivo se ha convertido en una de las áreas de investigación más importantes dentro del descubrimiento de información en grandes bases de datos.

En minería web se distinguen tres dominios de extracción de conocimiento de acuerdo con la naturaleza de los datos [6][31]: *minería de contenido web* que se centra en la extracción de conocimiento del contenido de los documentos o de sus descripciones, *minería de estructura web* que intenta descubrir conocimiento a partir de la forma en que se enlazan (o se incluyen) unos documentos con otros y la *minería de uso web* que obtiene el conocimiento de los ficheros de registro de los servidores.

En este trabajo nos centraremos en el dominio de la *minería de uso web*, aplicando una técnica de minería de datos descriptiva como el descubrimiento de subgrupos. Nuestra implementación utiliza técnicas de *computación flexible*, como los *algoritmos evolutivos* [16][18] y la *lógica difusa* [43][44], que facilitan la interpretación por los expertos del conocimiento extraído y mejoran los resultados obtenidos por los algoritmos clásicos en resolución de bases de datos con una alta dimensionalidad. Aplicaremos esta técnica sobre la base de datos del *MSNBC Anonymous Web Data*, que es una base de datos que se caracteriza por su alta dimensionalidad y que describe las páginas visitadas por los usuarios que accedieron a la web del MSNBC[1] el 28 de septiembre del 1999.

El artículo se organiza de la siguiente forma: En la sección 2 se definen los aspectos más importantes de la minería de uso web. En la sección 3 se introduce el problema que pretendemos resolver, en la sección 4 se define el descubrimiento de subgrupos, a continuación en la sección 5 se describe la aproximación del algoritmo SDIGA. En la sección 6 se encuentra la experimentación llevada a cabo sobre la base de datos *MSNBC Anonymous Web Data*. Para concluir, en la sección 7 se detallan las conclusiones y posibles trabajos futuros a realizar.

2 Minería de uso web

La minería de uso web extrae conocimiento de datos de utilización de la web o *web logs*. Un *web log* es una lista de datos de referencias de páginas. En estos *logs* se pueden examinar, desde una perspectiva del cliente o desde una perspectiva de servidor (la más destacada

[1] http://www.msnbc.msn.com

de ellas es la del lado del servidor), las páginas alojadas en cada uno de los sitios, en la que el proceso de minería de uso web descubre información sobre el uso y el acceso que los usuarios hacen del sitio.

La minería de uso web es una de las áreas de investigación que más desarrollo está teniendo en la actualidad para descubrir conocimiento oculto dentro de la Web [38], y consta de tres fases distintas y separadas que identificamos de forma secuencial: Preparación de los datos, Minería de datos y Evaluación e Interpretación.

La actividad más habitual llevada a cabo dentro de la minería de uso web es la mejora del diseño del sitio, aunque según Facca y Lanzi [11], los ficheros *web log* se pueden aplicar para mejorar en distintas áreas de aplicación de la minería de uso web. Para la resolución de estos problemas se pueden aplicar distintos tipos de técnicas de minería de datos. Estas áreas y técnicas son las siguientes:

- Personalización, en la que analizando la secuencia de páginas a las que accede un usuario se pueden desarrollar un perfil sobre el mismo. Dentro de este área se pueden encontrar entre otras, una técnica de *fuzzy clustering* que obtiene perfiles de usuario y los usa para generar una web de interés para el usuario automáticamente [21], una técnica de reglas de asociación, donde el soporte su ajusta automáticamente [29], o también, una técnica basada en la agregación de perfiles de usuario, con *clustering* y reglas de asociación [34]. Entre otras técnicas relacionadas con la lógica difusa también se pueden encontrar distintos trabajos como por ejemplo el de Delgado[9], que mediante reglas de asociación difusas intenta obtener el comportamiento de los usuarios en la web, o trabajos como [32] que trata los perfiles de usuario en la Web. En [33] se puede encontrar una revisión de las técnicas utilizadas dentro de este campo.
- *Caching*, donde se estudian técnicas de clasificación para saber qué páginas visitará el usuario y cargarlas con anterioridad. Dentro de este área se puede encontrar un modelo de predicción para mejorar los sistemas de Internet basados en técnicas de reglas de asociación y clasificadores [26], además de predicción de caching basada en *n-gram* que es un algoritmo de reglas de asociación basado en la filosofía GSDF [42], entre otras.
- Diseño, donde se estudia la calidad y efecitividad global de las páginas del sitio, Podemos encontrar una técnica para descubrir patrones de secuencias y clasificadores para mejorar el diseño de un sitio web [5], técnicas mediante clasificadores para obtener un sistema web adaptativo para facilitar los accesos de los usuarios [12], técnicas que miden la calidad de navegación, del servicio o de la página mediante técnicas de patrones de secuencia [39], o, técnicas que se basan en el *backtracking* de los usuarios para descubrir comportamiento para el diseño, buscando patrones de secuencia en las páginas [40].
- e-comercio, donde podemos encontrar entre otras, una técnica de clustering para medir las relaciones secuenciales en las páginas web, basándose en lógica difusa y algoritmos genéticos [17].

Considerando el problema que pretendemos resolver, donde se procesa la base de datos del MSNBC, que describe las páginas visitadas por los usuarios que accedieron a la web el 28 de septiembre del 1999, se aplicará la técnica de descubrimiento de subgrupos para obtener reglas que sean capaces de mostrar a los expertos los patrones de navegación para este sitio más relevantes dentro de la base de datos. El descubrimiento de subgrupos es una técnica de inducción descriptiva, cuya aplicación dentro del campo de la minería de uso web es novedosa y hasta el momento desconocida dentro de este área de aplicación.

3 Problema a resolver

El problema que se pretende resolver consiste en la extracción de patrones frecuentes en la base de datos *MSNBC Anonymous Web Data*, disponible en el repositorio UCI[2] [2]. Esta base de datos, obtenida de los ficheros *log* de los servidores del portal MSNBC y preprocesada y almacenada como un conjunto de sesiones de usuario, describe las páginas visitadas por los usuarios que accedieron a la web del MSNBC el 28 de septiembre del 1999. Las características principales de esta base de datos se pueden observar en la tabla 1.

Tabla 1: Características principales de MSNBC

Información Relevante	
Número de usuarios	989818
Número medio de visitas por usuario	5,7
Longitud mínima de peticiones por sesión	1
Longitud máxima de peticiones por sesión	500
Número de URLs por categoría	10 a 5000

La base de datos MSNBC, consiste en un conjunto de datos de longitud variable con información sobre 989818 usuarios, donde para cada usuario se almacena la información de una *sesión*, es decir, las páginas visitadas por el usuario en un periodo de 24 horas, por orden de petición. En esta base de datos, cada una

[2] http://www.ics.uci.edu/∼mlearn/MLRepository.html

de las páginas se asocia a una categoría (cat) a la que se asigna un código (cód); así la categoría frontpage se codifica con el número 1, la categoría news con el número 2 y así sucesivamente hasta las 17 categorías existentes, tal y como muestra la tabla 2. Dentro de cada categoría, de las 17 que hay, se pueden encontrar asociadas un número de páginas web que pueden oscilar entre 10 y 5000, y el número de peticiones realizadas por un usuario en una sesión, pueden ir desde 1 como mínimo hasta 500 como máximo, es decir, una sesión puede tener un tamaño que oscila entre 1 y 500.

Tabla 2: Códigos para las categorías del MSNBC

cat	cód	cat	cód	cat	cód
frontpage	1	misc	7	summary	13
news	2	weather	8	bbs	14
tech	3	health	9	travel	15
local	4	living	10	msn-news	16
opinion	5	business	11	msn-sports	17
on-air	6	sports	12		

Una de las ventajas de usar la base de datos MSNBC, es la gran difusión que tiene dentro de la literatura actual con la aplicación de distintas técnicas para la resolución de problemas de minería de uso web, como por ejemplo:

- Búsqueda de secuencias de patrones mediante el algoritmo FS-Miner [10] que aplica técnicas de reglas de asociación para la aplicación de *caching* y *pre-fetching* sobre los *web logs*.
- Descubrimiento de órdenes parciales en las secuencias de las sesiones de usuario mediante estructuras [15].
- Búsqueda de patrones frecuentes [19][20] que presenta tres algoritmos capaces de obtener patrones, secuencias y representaciones gráficas, respectivamente.
- Técnicas de *rough clustering hierarchy* de datos secuenciales [25].
- Hiperenlaces adaptativos mediante las secuencias de acceso [30] que es una aproximación mediante teoría de grafos para mejorar la navegación.
- Técnicas de *clustering* basadas en medidas de probabilidad para la obtención de resultados [35].
- Descubrir patrones para describir el comportamiento de los usuarios mediante técnicas de *clustering* [37].

Como el objetivo es la obtención de patrones frecuentes mediante reglas interpretables tras la aplicación de descubrimiento de subgrupos, ha sido necesario realizar una preparación de los datos, para convertir la base de datos de longitud variable a una base de datos con tamaño fijo. De esta forma obtenemos un conjunto de datos de 17 variables (categorías) y 989818 usuarios (sesiones), donde cada categoría almacena el número de peticiones realizadas por el usuario para cada sesión.

El descubrimiento de subgrupos tiene la finalidad de obtener subgrupos de la población, donde dando el conjunto de datos y una propiedad de esos datos en la que el usuario esté interesado se buscan subgrupos que sean interesantes para el usuario en el sentido de que tengan una distribución estadística inusual respecto a la propiedad resaltada por el usuario. Esta filosofía se puede asemejar en cierto modo a la extracción de patrones frecuentes, tal y como se puede observar en los trabajos de Iváncsy [19][20]. En estos trabajos, se obtienen patrones frecuentes partiendo de unos determinados niveles de umbral y confianza, consiguiendo así los mejores patrones dentro de la base de datos. Por contra, en nuestra aproximación se obtienen reglas basándose exclusivamente en un umbral de soporte, el cual detallaremos más adelante.

Como se detalla en la sección 5, la aproximación se basa en técnicas de *soft-computing* para la resolución del problema. Dentro de este campo, SDIGA utiliza las técnicas de lógica difusa y algoritmos genéticos. Mediante estas técnicas se obtiene un procesamiento más flexible para la obtención de resultados, ya que permiten tratar los problemas asemejándose más a los problemas que existen en la realidad. Tal y como se puede observar en [36], las técnicas de *soft-computing* aplicadas a los problemas de minería web para el reconocimiento de patrones son muy relevantes. No solo la lógica difusa y los algoritmos genéticos, sino también técnicas como las redes neuronales o los conjuntos rugosos [1].

4 Descubrimiento de Subgrupos

Las técnicas de descubrimiento de subgrupos generan modelos basados en reglas cuya finalidad es descriptiva, empleando una perspectiva predictiva para obtenerlos [28]. Las reglas utilizadas en la tarea de descubrimiento de subgrupos tienen la forma $Cond \rightarrow Clase$, donde la propiedad de interés para el descubrimiento de subgrupos es el valor de la Clase que aparece en el consecuente de la regla [13][27] y el antecedente es una conjunción de variables (parejas atributo-valor) seleccionadas entre las variables del conjunto de datos.

El concepto de descubrimiento de subgrupos fue formulado inicialmente por Klösgen [23] y Wrobel [41], como "dado un conjunto de datos y una propiedad de

esos datos que sea de interés para el usuario, buscar subgrupos que sean de mayor interés para el usuario". En este sentido, se dice que un subgrupo es interesante cuando tiene una distribución estadística inusual respecto a la propiedad en la que estamos interesados. El objetivo es descubrir propiedades características de subgrupos construyendo reglas individuales sencillas (con una estructura comprensible y con pocas variables), altamente significativas y con soporte alto (que cubran muchas instancias de la clase objetivo).

Entre los distintos modelos más interesantes que obtienen descripciones de subgrupos representados de diferentes formas, tenemos: *AprioriSD* [22], el cual adopta un algoritmo de aprendizaje de reglas de asociación al descubrimiento de subgrupos; *CN2-SD* [28], cuyo propósito es inducir subgrupos en forma de reglas utilizando como medida de calidad la relación entre el ratio de verdaderos positivos y el ratio de falsos positivos; y *SD-MAP* [3], que es un algoritmo exhaustivo de extracción de reglas de descubrimiento de subgrupos basado en el algoritmo *FP-Growth* de reglas de asociación.

Para llevar a cabo una tarea de descubrimiento de subgrupos, es necesario definir cuatro elementos fundamentales [4]:

- El tipo de variable objetivo. Puede ser binaria, nominal o numérica. Dependiendo del número de variable objetivo son posibles diferentes cuestiones analíticas.
- El lenguaje de descripción que especifique los individuos de la población de referencia que pertenecen al subgrupo.
- La función de calidad que mide el interés de los subgrupos. Se han propuesto gran variedad de funciones de calidad [14][24][23]. Las funciones de calidad aplicables en cada caso vienen determinadas por el tipo de variable ojetivo y el problema analítico.
- La estrategia de búsqueda. Este aspecto es muy importante, puesto que el espacio de búsqueda crece exponencialmente con el número de posibles expresiones que pueden formar parte de una descripción de un subgrupo.

Uno de los aspectos más importantes de cualquier enfoque de inducción de reglas que describan subgrupos es la elección de las medidas de calidad a utilizar, tanto para seleccionar las reglas, como para valorar los resultados obtenidos en el proceso.

5 SDIGA: Subgroup Discovery Iterative Genetic Algorithm

El algoritmo SDIGA es un modelo evolutivo para la extracción de reglas difusas para la tarea de descubrimiento de subgrupos. En [8] puede encontrarse una descripción detallada del mismo.

Este algoritmo puede trabajar con reglas difusas o nítidas y con distintas estructuras de regla dependiendo del problema que deseemos resolver (DNF o canónicas). En este trabajo utilizaremos reglas difusas canónicas en las que el antecedente está compuesto por una conjunción de pares variable-valor.

El modelo evolutivo propuesto sigue el enfoque IRL, en el que el que cada cromosoma representa una regla, la solución del AG es el mejor individuo, y la solución global está formada por los mejores individuos de una serie de ejecuciones sucesivas [7]. Así, el núcleo de SDIGA es un AG que utiliza una etapa de post-procesamiento basada en una búsqueda local que incrementa la generalidad de la regla obtenida. Este AG híbrido se incluye en un proceso iterativo para la extracción de un conjunto de reglas difusas de descripción de subgrupos soportado por zonas distintas del espacio de búsqueda, que extrae reglas mientras se cubran nuevos ejemplos y las reglas generadas superen el umbral de confianza. Este proceso iterativo permite la extracción de reglas diferentes en sucesivas ejecuciones del algoritmo.

La función de adaptación (*fitness*) se calcula de acuerdo a la ecuación 1, en la que se utilizan el soporte (Sop_c) como medida de generalidad, la confianza ($Conf$) como medida de precisión y el interés (Int) de la regla como medida de novedad. Las variables ω_1, ω_2 y ω_3 son los pesos asignados a cada una de las medidas.

$$fitness(c) = \frac{\omega_1 \times Sop_c(c) + \omega_2 \times Conf(c) + \omega_3 \times Int(c)}{\omega_1 + \omega_2 + \omega_3} \quad (1)$$

Tal y como se puede ver en [8], estas medidas se definen de la siguiente forma:

- **Soporte:** Es una medida del grado de cobertura que la regla ofrece a los ejemplos de la clase. Para promover la obtención de diferentes reglas en sucesivas ejecuciones del algoritmo, en su cálculo solo se consideran los ejemplos no cubiertos por las reglas previamente generadas.
- **Confianza:** La confianza de una regla es una medida estándar que determina la frecuencia relativa de los ejemplos que satisfacen tanto el antecedente

como el consecuente de una regla entre aquellos que satisfacen sólo el antecedente.

- **Interés:** Para el cálculo del interés, en la práctica es adecuado utilizar criterios objetivos como medidas que permitan seleccionar reglas potencialmente interesantes y criterios subjetivos para que el usuario final determine reglas realmente interesantes.

La lógica difusa mejora la interpretabilidad de las reglas extraídas debido a la utilización de una representación del conocimiento más cercana al experto, permitiendo además la utilización de variables numéricas sin necesidad de realizar una discretización previa. Los conjuntos difusos que definen las etiquetas lingüísticas se definen mediante las correspondientes funciones de pertenencia especificadas por el usuario o definidas mediante una partición uniforme si no está disponible el conocimiento del experto. Como se puede observar en la figura 1, se utilizan particiones uniformes con funciones de pertenencia triangulares (en este caso en concreto con 5 etiquetas).

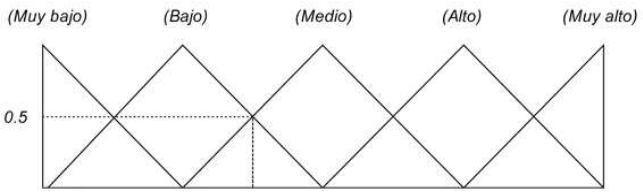


Figura 1: Conjunto difuso utilizado por SDIGA para 5 etiquetas

6 Experimentación realizada

La experimentación se ha llevado a cabo sobre la base de datos MSNBC. Considerando que en los algoritmos de descubrimiento de subgrupos hay que indicar cuál es el atributo objetivo, se necesitaría realizar un estudio para cada una de las categorías de la base de datos (como la variable objetivo tiene 17 valores posibles, habría que realizar una experimentación para cada uno de éstos, obteniendo una base de datos distinta para cada categoría). Sin embargo, en la literatura actual podemos encontrar distintos estudios realizados sobre MSNBC [19], que indican como categorías relevantes para la extracción de patrones frecuentes las categorías *frontpage*, *news* y *on-air*. Por esta razón, se ha realizado la experimentación sólo sobre estas tres categorías.

El objetivo principal que se pretende para esta primera aproximación de aplicación de descubrimiento de subgrupos en el área de la minería de uso web es obtener un conjunto de patrones de MSNBC para cada una de las categorías seleccionadas como variable objetivo y descubrir conocimiento sobre ellas.

Debido al gran tamaño de la base de datos (989818 sesiones descritas mediante 17 variables), es necesario realizar una reducción de instancias para reducir el tiempo de procesamiento y almacenamiento. En la literatura actual se pueden encontrar distintas técnicas de reducción de instancias aplicadas a la base de datos MSNBC [10][20][25]. En esta experimentación, hemos aplicado una selección de instancias aleatoria del 10% sobre la base de datos, obteniendo una nueva base de datos con 98982 usuarios.

Se ha realizado una *5-Validación Cruzada*, generándose cinco pares distintos de entrenamiento y prueba con un 20% de patrones para el entrenamiento y un 80% para el conjunto de prueba. Las particiones se han realizado en esta proporción ya que debido al gran tamaño de la base de datos, el porcentaje de entrenamiento es suficiente para que el modelo generado sea capaz de describir información a partir del conjunto de prueba.

Se ha desarrollado una experimentación sobre estas bases de ejemplos con los algoritmos clásicos de descubrimiento de subgrupos, Apriori-SD [22] y CN2-SD [28], pero no se obtienen resultados para poder realizar una comparación con SDIGA, ya que la necesidad de aplicar una discretización sobre la base de datos para el procesamiento de éstos, produce una gran pérdida de información que provoca la obtención de reglas vacías. Este problema no se presenta con SDIGA, puesto que no necesita discretización.

Dado el carácter no determinista de SDIGA, se han realizado tres ejecuciones independientes para cada uno de estos pares de conjuntos entrenamiento-prueba y los resultados mostrados en la tabla 3 son la media de los resultados. Como parámetros de la experimentación de SDIGA, se ha utilizado un umbral de confianza del 80%, un tamaño de la población del AG de 100, un máximo de 10000 evaluaciones de individuos para el AG, y una distribución de pesos de 0.4, 0.3 y 0.1 para el soporte, confianza e interés, respectivamente.

Para realizar la experimentación, se ha realizado la comparación de resultados considerando distinto número de etiquetas para los conjuntos difusos. De esta forma tenemos los algoritmos *SDIGA-3*, *SDIGA-5* y *SDIGA-7*, que utilizan 3,5 y 7 etiquetas respectivamente para la descripción de las variables continuas.

Los resultados se pueden observar en la tabla 3, donde *Objetivo* es la categoría utilizada como variable objetivo, *Alg* es el nombre del Algoritmo utilizado, *NReg* es el número de reglas obtenidas por el algoritmo,

Tabla 3: Resultados obtenidos de los experimentos realizados.

Objetivo	*Alg*	*NReg*	*Var*	$Sop_c(\%)$	$Conf(\%)$	*Int*	*Rel*
frontpage	*SDIGA-3*	**2**	**2**	92.90	**50.59**	**1.02**	**153.80**
	SDIGA-5	3	**2**	**95.09**	**50.59**	1.00	128.90
	SDIGA-7	**2**	**2**	94.01	50.29	1.00	135.23
news	*SDIGA-3*	**2**	**2**	99.13	64.52	**1.04**	**582.93**
	SDIGA-5	3.13	2.01	**99.17**	**64.74**	1.03	402.85
	SDIGA-7	**2**	**2**	98.70	64.17	1.03	326.48
on-air	*SDIGA-3*	**2**	**2**	**90.18**	50.17	**1.02**	120.21
	SDIGA-5	3	**2**	86.26	50.09	**1.02**	183.06
	SDIGA-7	**2**	**2**	87.75	**50.53**	1.01	**226.37**

Var es el número de variables que contiene cada regla, Sop_c, *Conf* e *Int* son el Soporte, la Confianza y el Interés, y *Rel* es la relevancia, medida muy utilizada en tareas de descubrimiento de subgrupos que se calcula en términos de su razón de verosimilitud (*likelihood ratio*) normalizada con la razón de verosimilitud del umbral de relevancia (99%).

Analizando los resultados obtenidos de la experimentación, se puede observar que el número de etiquetas con el que se obtiene mejores resultados depende de la categoría a estudiar. Así, para las categorías *frontpage* y *news* son mejores los conjuntos de 5 etiquetas, mientras que para la categoría *on-air* se obtienen mejores resultados con el conjunto de 7 etiquetas.

Se obtienen conjuntos con pocas reglas, con reglas sencillas interpretables, por el uso de reglas difusas lingüísticas, y en todos los casos con valores de soporte y confianza muy similares. Son reglas que abarcan casi la totalidad del conjunto de ejemplos pero con una confianza ligeramente baja. Además, en el caso de las clases o grupos on-air y frontpage son más difíciles de representar de forma precisa con el conjunto de instancias seleccionadas y el método de selección de instancias aleatorio utilizado.

7 Conclusiones y Trabajos futuros

Se ha presentado una primera aproximación a la aplicación de un algoritmo de descubrimiento de subgrupos con reglas difusas para un problema de minería de uso web. Hasta donde conocemos, no se han aplicado algoritmos de descubrimiento de subgrupos con lógica difusa a este problema. Las aproximaciones más cercanas son [17][21], que se centran en técnicas de agrupamiento difuso. Los resultados obtenidos muestran que es posible extraer conocimiento descriptivo en este tipo de problemas pero que es necesario hacer un análisis detallado del preprocesamiento, es decir de la selección de instancias.

En este trabajo se ha utilizado un método de selección de instancias aleatorio con una reducción muy importante (90% por la dimensionalidad de la base de datos original). Es conveniente estudiar el efecto de otros métodos de selección de instancias y la influencia del sesgo del mismo en el método posterior. La integración de un método de selección de instancias dentro del propio método de minería de datos, podría dar buenos resultados.

Como trabajos futuros se pretende realizar un completo estudio estadístico sobre técnicas de preprocesamiento para esta base de datos. El objetivo es por un lado, demostrar que es posible realizar una selección de instancias sobre MSNBC perdiendo la mínima información posible, y por otro, realizar un análisis de la influencia del método de preprocesamiento utilizado en el conocimiento extraído tras aplicar un método de minería de datos descriptiva posterior.

Además, se pretende obtener mediante descubrimiento de subgrupos las secuencias que existen en la base de datos, intentando ir más allá de la mera extracción de patrones frecuentes e indicando las secuencias que los usuarios realizan para llegar a una categoría objetivo.

Referencias

[1] D. Arotaritei and S. Mitra. Web mining: a survey in the fuzzy framework. *Fuzzy Sets and Systems*, 148(1):5–19, 2004.

[2] A. Asuncion and D.J. Nweman. UCI Machine Learning Repository. *University of California*, 2007.

[3] M. Atzmüller and F. Puppe. Sd-map - a fast algorithm for exhaustive subgroup discovery. In Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors, *PKDD*, volume 4213 of *Lecture Notes in Computer Science*, pages 6–17. Springer, 2006.

[4] M. Atzmüller, F. Puppe, and Hans-Peter Buscher. Towards Knowledge-Intensive Subgroup Discovery. In Andreas Abecker, Steffen Bickel, Ulf Brefeld, Isabel Drost, Nicola Henze, Olaf Herden, Mirjam Minor, Tobias Scheffer, Ljiljana Stojanovic, and Stephan Weibelzahl, editors, *LWA*, pages 111–117. Humbold-Universität Berlin, 2004.

[5] B. Berendet. Using site semantics to analyze, visualize, and support navigation. *Data Mining and Knowledge Discovery*, 6(1):37–59, 2002.

[6] R. Cooley, B. Mobasher, and J. Srivastava. Web Mining: Information and Pattern Discovery on the World Wide Web. *On Tools with Artificial Intelligence*, pages 558–567, 1997.

[7] O. Cordón, M.J. del Jesus, F. Herrera, and M. Lozano. MOGUL: A Methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach. *Internation Journal of Intelligent Systems*, 14:1123–1153, 1999.

[8] M.J. del Jesus, P. González, F. Herrera, and M. Mesonero. Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A case study in marketing. *IEEE Transactions on Fuzzy Systems*, 12(3):296–308, 2007.

[9] M. Delgado, N. Marín, D. Sánchez, and M.A. Vila. Fuzzy Asociation Rules: General Model and Applications. *IEEE Transactions on Fuzzy Systems*, 11:214–225, 2003.

[10] M. El-Sayed, E. Rudensteiner, and C. Ruiz. FS-Miner: Efficient and Incremental Mining of Frequent Sequence Patterns. 2004.

[11] F.M. Facca and P.L. Lanzi. Mining interesting knowledge from weblogs: a survey. In *Data & Knowledge Engineering*, volume 53, pages 225–241. Elsevier, 2005.

[12] Y. Fu, M. Creado, and C. Ju. Reorganizing web sites based on user access patterns. In ACM Press, editor, *Tenth International Conference on Information and Knowledge Managemen*, pages 583–585, 2001.

[13] D. Gamberger and N. Lavrac. Expert-Guided Subgroup Discovery: Methodology and Application. *J. Artif. Intell. Res. (JAIR)*, 17:501–527, 2002.

[14] D. Gamberger and N. Lavrac. Active subgroup mining: a case study in coronary heart disease risk group detection. *Artificial Intelligence in Medicine*, 28(1):27–57, 2003.

[15] G. Garriga, P. Díaz-López, and J.L. Balcázar. A lattice-based method for structural analysis. 2005.

[16] D.E. Golberg. Genetic Algorithms in search, optimization and machine learning. *Addison-Wesley*, 1989.

[17] B. Hay, G. Wets, and K. Vanhoof. Clustering navigation patterns on a website using a sequence alignment method.

[18] J.H. Holland. Adaptation in natural and artificial systems. *University of Michiga Press*, 1975.

[19] R. Iváncsy and I. Vajk. Different Aspects of Web Log Mining. 2005.

[20] R. Iváncsy and I. Vajk. Frequent Pattern Mining in Web Log Data. *Acta Polytechnica Hungarian*, 3(1), 2006.

[21] T. Kamdar. *Creating adaptive web servers using incremental web log mining*. PhD thesis, Computer Science Department, University of Maryland, Baltimore Country, 2001.

[22] B. Kavsek and N. Lavrac. APRIORI-SD: Adapting association rule learning to subgroup discovery. *Applied Artificial Intelligence*, 20:543–583, 2006.

[23] W. Klösgen. Explora: A Multipattern and Multistrategy Discovery Assistant. In *Advances in Knowledge Discovery and Data Mining*, pages 249–271. 1996.

[24] W. Klösgen and J. Zytkow. *Handbook of Data Mining and Knowledge Discovery*. Oxford, 2002.

[25] P. Kumar, P.R. Krishna, R.S. Bapi, and S. Kumar De. Rough Clustering of Sequential Data. *Data Knowl. Eng.*, 63(2):183–199, 2007.

[26] B. Lan, S. Bressan, B.C. Ooi, and K.-L. Tan. Rule-assisted prefetching in web-server caching. In ACM Press, editor, *Ninth International Conference on Information and Knowledge Management (CIKM 2000)*, pages 504–511, 2000.

[27] N. Lavrac, B. Cestnik, D. Gamberger, and P.A. Flach. Decision Support Through Subgroup Discovery: Three Case Studies and the Lessons Learned. *Machine Learning*, 57(1-2):115–143, 2004.

[28] N. Lavrac, B. Kavsek, P.A. Flach, and L. Todorovski. Subgroup Discovery with CN2-SD. *Journal of Machine Learning Research*, 5:153–188, 2004.

[29] W. Lin, S.A. Alvarez, and C.Ruiz. Efficient adaptive support association rule mining for recommender systems. *Data Mining and Knowledge Discovery*, 6(1):83–105, 2002.

[30] P. Lingras and R. Lingras. Adaptative hyperlinks using page access sequences and minimum spanning trees. 2006.

[31] Z. Markov and D.T. Larose. *Data Mining The Web. Uncovering patterns in Web Content, Structure and Usage*. Wiley, 2007.

[32] M.J. Martín-Bautista, D.H. Kraft, M.A. Vila Miranda, J. Chen, and J. Cruz. User profiles and fuzzy logic for web retrieval issues. *Soft Comput.*, 6(5):365–372, 2002.

[33] B. Mobasher. *Web Usage Mining and Personalization*. CRC Press LLC, 2005.

[34] B. Mobasher, H. Dai, and M. Tao. Discovery and evaluation of aggregate usage profiles for web personalization. *Data Mining and Knowledge Discovery*, 6:61–82, 2002.

[35] V. Nikulin. Universal Clustering with Family of Power Loss Functions in Probabilistic Space. In Marcus Gallagher, James M. Hogan, and Frédéric Maire, editors, *IDEAL*, volume 3578 of *Lecture Notes in Computer Science*, pages 311–318. Springer, 2005.

[36] S.K. Pal and S. Mitra. *Neuro-Fuzzy Pattern Recognition: Methods in Soft-Computing*. Wiley, New York, 1999.

[37] G. Pallis, F. Angelis, and A. Vakali. Model-bases cluster analysis for web user sessions. 3488:219–227, 2005.

[38] M. Spiliopoulou. The laborious way from data mining to web mining. *International Journal of Computer System*, 14:113–126, 1999.

[39] M. Spiliopoulou and C. Pohle. *Data Mining for Measuring and Improving the Success of Web Sites*, volume 5, pages 85–114. Kluwer Academic Publishers, 2001.

[40] R. Srikant and Y. Yang. Mining web logs to improve website organization. In *WWW*, pages 430–437, 2001.

[41] S. Wrobel. An Algorithm for Multi-relational Discovery of Subgroups. In Henryk Jan Komorowski and Jan M. Zytkow, editors, *PKDD*, volume 1263 of *Lecture Notes in Computer Science*, pages 78–87. Springer, 1997.

[42] Q. Yang and H.H. Zhang. Web-log mining for predictive web caching. *IEEE Transactions on Knowledge and Data Engineering*, 15(4):1050–1054, 2003.

[43] L.A. Zadeh. Information Control. *Fuzzy sets*, 8:338–353, 1965.

[44] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning, Parts I, II, III. *Information Science*, 8-9:199–249, 301–357, 43–80, 1975.

EXTRACCIÓN DE ESTRUCTURAS DE CONOCIMIENTO EN CAMPOS TEXTUALES

María J. Martín-Bautista[1] **Sandro Martínez-Folgoso**[2] **María-Amparo Vila**[1]

[1] Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, {mbautis,vila}@decsai.ugr.es
[2] Dpto. of Computer Science, University of Camagüey, 74650 - Camagüey, sandro.martinez@reduc.edu.cu

Resumen

En este trabajo, proponemos una posible solución al manejo de campos textuales sin estructura en almacenamiento de datos estructurado. La idea general consiste en encontrar una estructura de conocimiento para poder manejar los atributos de tipo texto de la misma forma que el resto de los campos de la base de datos. Para ello, proponemos una representación basada en unas estructuras matemáticas denominadas conjuntos-AP y sus operaciones inspiradas en los conjuntos de itemsets frecuentes derivados del algoritmo Apriori que normalmente se utiliza en minería. Para obtener dicha representación, proponemos una arquitectura del sistema y una metodología a partir de las cuales podemos obtener el Tipo Abstracto de Dato del atributo textual junto con sus métodos.

Palabras Clave: Textos cortos, conjuntos-AP, bases de datos, minería de texto.

1 INTRODUCCIÓN

La falta de estructura de los datos textuales hace difícil su procesamiento automático para manejarlos de forma masiva. Esto sucede no solo en repositorios de textos, sino también en atributos de tipo texto en bases de datos relacionales, como las de tipo médico [10] [8], encuestas, bases de datos de correos electrónicos, etc. Existen diversas áreas de investigación tales como el Procesamiento de Lenguaje Natural, la Minería de Texto, la Recuperación de Información, etc., que tratan con este problema e intentan resolverlo desde diferentes puntos de vista. La mayoría de las soluciones vienen de encontrar algún tipo de estructura en el texto y/o transformarlo en lo que es llamado forma intermedia [5]. Como se mencionó anteriormente, en este trabajo nos planteamos como objetivo primordial el resolver esta problemática mediante la transformación del atributo de tipo texto en una estructura intermedia [2]. Dicha estructura está basada en el concepto de itemsets frecuentes y sus propiedades, llamada conjunto-AP.

Para resolver esta problemática, se comienza en la sección 2 del presente trabajo definiendo las estructuras en las que se transforman los atributos textuales y sus operaciones. Se introducen los conceptos de conjunto-AP y estructura-AP como estructuras básicas que componen el modelo de representación que se ha propuesto. En la sección 3, se introducen un conjunto de operaciones y medidas que permitirán realizar búsquedas sobre las estructuras de conocimiento asociadas a los textos libres, y que además nos permitirán en su momento:

1. Sugerirle al usuario conjuntos de términos para refinar su búsqueda.
2. Decirle la cantidad de términos que aparecen representados de todos los que está buscando.
3. Si su frase aparece completamente incluida en la estructura de conocimiento, etc.

En la sección 4, se describe el sistema explicando su arquitectura. La metodología para obtener el Tipo Abstracto de Dato (TDA) a partir del atributo textual original queda recogida en la sección 5, donde además se introducen parte de los resultados de un ejemplo experimental sobre un conjunto de datos real. Finalmente, se incluyen las conclusiones en la sección 6.

2 Definición formal de las estructuras matemáticas

En esta sección se comienzan a formalizar las ideas presentadas anteriormente, definiendo las estructuras matemáticas que serán la base para la representación formal de los datos. Se establecerán las definiciones y propiedades de los conjuntos que tienen la propiedad ' *APriori* ' (conjuntos-AP) [2]. A continuación se da la definición formal y las propiedades de las estructuras subyacentes en los textos, que capturan la semántica que los mismos encierran y que como se verá, está compuesta por un conjunto de conjuntos-AP. Dichas estructuras se denominan estructuras-AP. Cabe destacar que en este trabajo se muestran sólo las definiciones y propiedades más importantes de estas representaciones. Una explicación más detallada de las mismas puede ser consultada en [7].

2.1 Definición y propiedades de los conjuntos-AP

2.1.1 Definición de Conjunto-AP

Sean $X = \{x_1...x_n\}$ un conjunto referencial de items y $R \subseteq P(X)$ un conjunto de itemsets frecuentes, diremos que R es un conjunto-AP si y solo si:

1. $\forall Z \in R \Rightarrow P(Z) \subseteq R$
2. $\exists Y \in R$ tal que :
   (a) $card(Y) = max_{Z \in R}(card(Z))$ y no exista $Y' \in R$ tal que $card(Y') = card(Y)$
   (b) $\forall Z \in R; Z \subseteq Y$

El conjunto Y de máxima cardinalidad caracteriza al conjunto-AP y este será llamado *conjunto generador de R*. Se denotará $R = g(Y)$, y significa que $g(Y)$ será el conjunto-AP con conjunto generador Y. Obsérvese que $g(Y)$ es el conjunto de las partes de Y.

Se denotará *grado de* $g(Y)$ al cardinal de Y, o sea, a la cantidad de elementos del conjunto generador Y. Obviamente, los conjuntos-AP de grado 1 son los elementos de X, y además consideraremos el conjunto vacío $\emptyset$ como el conjunto-AP de grado cero.

2.1.2 Ejemplo 1

Sea $X = \{1, 2, 3, ..., 10\}$ y
$R = \{\{1\}, \{3\}, \{5\}, \{1, 3\}, \{1, 5\}, \{3, 5\}, \{1, 3, 5\}\}$, entonces el conjunto generador es $Y = \{1, 3, 5\}$

Como se observa en el ejemplo anterior y teniendo en cuenta la definición 1, el conjunto generador $Y = \{1, 3, 5\}$ se corresponde con el conjunto-AP de mayor cardinalidad y que incluye a su vez, todas las combinaciones presentes en R.

Se debe destacar que la definición 1 implica que cualquier conjunto-AP $g(Y)$ es de hecho el retículo de $P(Y)$. Esto se puede afirmar ya que la definición 1, además de verificar que Y sea el conjunto de máxima cardinalidad y que no exista ningún otro conjunto con su misma cardinalidad, garantiza que todas las combinaciones del conjunto generados estén presentes. De esta forma con dichas combinaciones es construido el retículo del conjunto-AP.

2.2 Definición y propiedades de la estructura-AP

Una vez que ya se han establecido los conceptos de conjunto-AP, se usarán estos para definir las estructuras de información que aparecen cuando son obtenidos los itemsets frecuentes. Debe considerarse que dichas estructuras son obtenidas de forma constructiva, por la generación inicialmente de itemsets con cardinalidad igual a 1; seguidamente, estos son combinados para obtener los de cardinalidad igual a 2, y así sucesivamente hasta obtener los itemsets de máxima cardinalidad, fijando para ello un soporte mínimo. Por consiguiente, la estructura final será un conjunto de conjuntos-AP, la cual será definida formalmente como sigue.

2.2.1 Definición de Estructura-AP

Sea $X = \{x_1...x_n\}$ un conjunto referencial de items y $S = \{A, B, ...\} \subseteq P(X)$ un conjunto de itemsets frecuentes, tal que:

$$\forall A, B \in S\,;\, A \not\subseteq B\,,\, B \not\subseteq A$$

Llamaremos *estructura-AP del generador S*, $T = g(A, B, ...)$, al conjunto de conjuntos-AP cuyos conjuntos generadores son $A, B, ...$

De la definición anterior queda claro entonces, el hecho de que se mencionó anteriormente: la estructura-AP no es más que una colección de conjuntos-AP. Tal como se definió, los conjuntos generadores de la estructura-AP $A, B, ...$ no pueden estar contenidos unos en otros, utilizando para esta interpretación la definición de conjunto-AP incluido que se dio anteriormente. Entonces, la estructura-AP quedará constituida por todos los conjuntos generadores que se obtengan de las combinaciones de X presentes, dentro de todas las posibles ($P(X)$).

Debemos destacar también, que cualquier estructura-AP es un retículo de subconjuntos, cuyos extremos

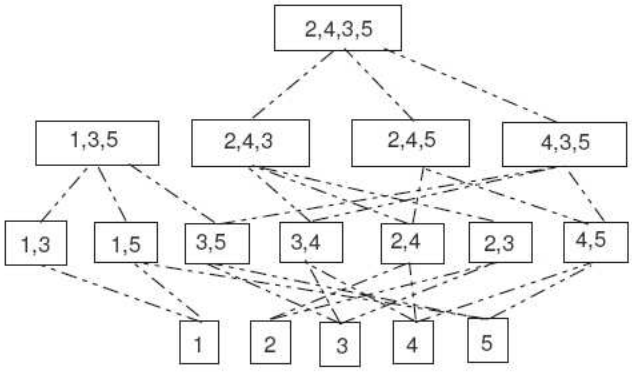


Figura 1: **Retículo de la estructura-AP.**

superiores son sus conjuntos generadores. La figura 1 muestra un ejemplo del retículo subyacente en $g(\{1,3,5\},\{2,4,3,5\})$. Como se observa en la figura, los conjuntos $\{1,3,5\}$ y $\{2,4,3,5\}$ son los conjuntos generadores. Es de destacar que aunque no se encuentran uno completamente incluido en el otro, comparten elementos comunes y se encuentran en la raíz del árbol que forma el retículo. Como también se aprecia, en los conjuntos-AP de cardinalidad dos, ambos conjuntos generadores ya comparten un conjunto-AP común, el $\{3,5\}$, lo mismo que en las hojas del árbol donde ya tienen en común los conjuntos-AP de cardinalidad uno ($\{3\}$ y $\{5\}$).

A continuación se introduce una de las operaciones más importantes que se pueden definir sobre esta estructura-AP, la operación subestructura-AP inducida, que no será más que la estructura-AP resultante de intersecar una estructura-AP cualquiera, con un conjunto dado. Esta operación reviste especial importancia, por el hecho de que será la que permita encontrar la representación del tipo de dato abstracto, que corresponde a una tupla dada de la base de datos.

2.2.2 Definición de Subestructura-AP inducida

Sea la estructura-AP $T = g(A_1, A_2, ..., A_n)$ con conjunto referencial de items X y $Y \subseteq X$. Definiremos la *estructura-AP de T inducida por Y* como:

$$T' = T \bigwedge Y = g(B_1, B_2, ..., B_m)$$

donde

$$\forall B_i \in \{B_1, ..., B_m\} \Rightarrow \exists A_j \in \{A_1, A_2, ..., A_n\} \text{ tal que } B_i = A_j \bigcap Y$$
$$\forall A_j \in \{A_1, ..., A_n\} \Rightarrow \exists B_i \in \{B_1, B_2, ..., B_m\} \text{ tal que } A_j \bigcap Y \subseteq B_i$$

Según esta definición, queda claro que T' es la estructura-AP generada por las intersecciones de Y con los conjuntos generadores de T. Para ello, se interseca cada elemento de Y con todos los elementos $(A_1, A_2, ..., A_n)$ de T; a los elementos B_i de T' irán solo las intersecciones que no estén completamente incluidas en otro elemento B_i, o sea, las intersecciones resultantes entre los elementos de T y Y que ya se encuentren incluidas en otro conjunto B_i serán eliminadas. De esta forma, se garantiza que la estructura-AP obtenida solo esté formada por conjuntos generadores maximales, tal como plantea el concepto de estructura-AP introducido anteriormente. El siguiente ejemplo clarifica estas ideas.

2.2.3 Ejemplo 2

Sea $X = \{1, 2, ..., 9\}$,

$T = g(\{1,2,3\},\{2,3,4\},\{4,5\},\{6\})$,

$Y = \{2,3,4,5\}$ entonces se tiene

$T \bigwedge Y = g(\{2,3,4\},\{4,5\})$.

Del ejemplo anterior se debe destacar que el conjunto $\{2,3\}$ no está incluido en $T \bigwedge Y$, mientras que este conjunto sí pertenece a la intersección de Y y el conjunto generador de T. La razón de la no inclusión es, como se explicó anteriormente, que el conjunto $\{2,3\}$ ya está incluido en uno de los conjuntos de salida de $T \bigwedge Y$, el $\{2,3,4\}$.

3 Consultando la base de datos: intersección de conjuntos con estructuras-AP

En esta sección, estableceremos las definiciones necesarias para consultar la base de datos. Teniendo en cuenta que la estructura-AP es obtenida a partir de los términos relevantes del atributo textual que se esté procesando, se puede decir entonces que contendrá la mayoría de los términos relevantes que aparecen en dicho atributo. De aquí que se pueda afirmar entonces que la estructura-AP es el dominio activo del atributo del que fue obtenida. La idea es que el usuario expresará sus requerimientos como conjuntos de términos que se contrastarán con las estructuras-AP almacenadas en la base de datos como valor de atributo. Para evaluar dichas comparaciones, necesitamos operaciones de acoplamiento definidas a continuación.

3.1 Definición de Acoplamiento fuerte

Sea la estructura-AP $T = g(A_1, A_2, ..., A_n)$ con conjunto referencial de items X y $Y \subseteq X$. Se define el *acoplamiento fuerte entre Y y T* como la operación

lógica:

$$Y \bigodot T = \begin{cases} \text{verdadero si} & \exists A_i \in \{A_1, A_2, ..., A_n\} \\ & / Y \subseteq A_i \\ \text{falso} & \text{en otro caso} \end{cases}$$

En la definición anterior, se define el tipo de acoplamiento fuerte, que será el que tendrá el conjunto Y con la estructura-AP T cuando dicho conjunto aparezca completamente incluido en alguno de los conjuntos generadores de T. En caso contrario se dice que Y no tiene un acoplamiento fuerte con la estructura-AP.

De forma análoga, se puede definir la operación que verifica si el conjunto Y está parcialmente incluido en T.

3.2 Definición de Acoplamiento débil

Sea la estructura-AP $T = g(A_1, A_2, ..., A_n)$ con conjunto referencial de items X y $Y \subseteq X$. Se define el *acoplamiento débil entre Y y T* como la operación lógica:

$$Y \bigoplus T = \begin{cases} \text{verdadero si} & \exists A_i \in \{A_1, A_2, ..., A_n\} \\ & / Y \bigcap A_i \neq \emptyset \\ \text{falso} & \text{en otro caso} \end{cases}$$

Como se observa, en esta definición se hace uso de la operación intersección entre conjuntos ($\bigcap$) para determinar si la intersección entre Y y los conjuntos generadores de la estructura-AP T no es vacía. Esto significa que dicho conjunto se acopla, al menos parcialmente, con alguno de los A_i generadores de T.

Estas definiciones se pueden complementar dando alguna medida o índice que cuantifique estos acoplamientos. La idea es considerar que el acoplamiento de un conjunto largo de términos tendrá un índice mayor que uno con un menor número de términos. Adicionalmente, si algún conjunto de términos se acopla con más de un conjunto generador, éste tendrá un índice mayor que el de otro que solo se acopla con un solo conjunto.

3.3 Índice de acoplamiento fuerte(débil)

Sea la estructura-AP $T = g(A_1, A_2, ..., A_n)$ con conjunto referencial de items X y $Y \subseteq X$. Se define el *índice de acoplamiento fuerte (débil) entre Y y T* como sigue:

$\forall A_i \in \{A_1, A_2, ..., A_n\}$ se denota $m_i(Y) = card(Y \bigcap A_i)/card(A_i)$, $S = \{i \in \{1, ..., n\} | Y \subseteq A_i\}$, $W = \{i \in \{1, ..., n\} | Y \bigcap A_i \neq \emptyset\}$.

Entonces se define el *índice de acoplamiento fuerte y débil entre Y y T* como sigue:

$$\textbf{Índice fuerte} = S(Y|T) = \sum_{i \in S} m_i(Y)/n$$

$$\textbf{Índice débil} = W(Y|T) = \sum_{i \in W} m_i(Y)/n$$

De estas dos definiciones se desprende que:

$$\forall Y \text{ y } T \ , \ S(Y|T) \in [0,1] \ , \ W(Y|T) \in [0,1] \text{ y}$$

$$W(Y|T) \geq S(Y|T)$$

Lo que significa que para todo conjunto Y y estructura-AP T, ambos índices de acoplamiento tomarán valor entre 0 y 1; de igual forma, siempre el índice de acoplamiento débil será mayor o igual que el índice de acoplamiento fuerte.

4 Descripción del sistema

El sistema desarrollado trata con textos cortos en ciertos dominios. Estos pueden estar no solo en repositorios de textos si no también en bases de datos (Relacionales, Relacional Orientado a Objetos, etc.). En este último caso, estos pueden ser atributos de tipo texto compuestos por algunas palabras o varias frases. Las operaciones sobre estos atributos son las clásicas en las bases de datos para atributos de tipo texto, tales como, preguntar sobre el contenido del atributo, o buscar si el atributo contiene determinadas palabras. Sin embargo, la falta de un dominio definido y una estructura en estos atributos, les impide que sean tratados como el resto de los atributos de la base de datos. El problema puede ser peor, cuando procesos tales como Datawarehousing, Minería de Datos, o simplemente una consulta semántica tiene que ser realizado sobre estos campos de tipo texto. Además contar el número de palabras repetidas en el atributo de tipo texto que cumplan cierta condición, no es calculable con la representación y operadores tradicionales de las bases de datos relacionales.

Nuestra hipótesis se basa en el hecho de que existe una semántica subyacente en determinados atributos textuales de la base de datos. A pesar de que el dominio de estos no es finito y es muy complicado de manejar, puesto que está compuesto por elementos del lenguaje natural, la semántica e incluso el vocabulario en los atributos, es en alguna medida restringido. De esta forma, podemos encontrar conjuntos de términos relacionados semánticamente que aparecen repetidamente en el atributo analizado. Para este propósito, nosotros usamos una forma intermedia basada en itemsets frecuentes obtenidos vía el algoritmo Apriori con la propiedad Apriori[1]. Dicha forma intermedia es

llamada conjuntos-AP. Con todos los conjuntos-AP obtenidos en la base de datos, es formada una estructura global intermedia nombrada estructura-AP. Esta estructura refleja la semántica del atributo textual en la base de datos.

4.1 Arquitectura del sistema

El sistema está compuesto básicamente por una base de datos inicial, un módulo de manejo de textos cortos, la base de datos modificada y un cliente de consulta. Dado que el principal objetivo de esta arquitectura es el manejo de atributos textuales como el resto de los atributos de la base de datos, podemos tener un Servidor OLAP que se aproveche de las ventajas del Sistema, como se muestra en la figura 2.

Como ya se mencionó, la solución propuesta parte de una base de datos inicial en PostgreSQL, en ella se almacena el Data Warehouse que contiene la información sensible a ser consultada por el cliente final. Para soportar la consulta y construcción de cubos multidimensionales en los que puedan participar dimensiones de textos libres, la base de datos inicial es sometida a un proceso de transformación, donde para los atributos de textos que sean de interés en la toma de decisiones, se les aplica un proceso de transformación para lograr su representación estructurada que pueda posibilitar su posterior consulta. Este proceso será realizado por una aplicación que tomando como entrada todos los atributos de textos seleccionados, obtendrá una base de datos inicial modificada soportada por el modelo objeto-relacional de bases de datos y además una base de datos con conocimiento subyacente.

Dado que la forma de representación intermedia escogida para la representación de los textos libres, son las reglas de asociación utilizando el algoritmo A priori, este proceso se encargará de obtener los conjuntos-AP generadores y las estructuras-AP resultantes del proceso. Como resultado se obtiene además, un diccionario que está compuesto por todos los términos relevantes correspondientes a cada atributo procesado, con su correspondiente frecuencia de aparición y la lista de Item-Sets frecuentes con su soporte. También se almacenan los conjuntos-AP detallados con su soporte. Ambas estructuras pueden ser utilizadas en el proceso de actualización de la base de datos inicial modificada, teniendo en cuenta la frecuencia de aparición de los términos con respecto a un umbral determinado y/o para realizar consultas sobre la estructura general, buscando posibles asociaciones entre términos para sugerir otras posibles respuestas de interés al cliente. Tanto la base de datos inicial modificada como el diccionario y la base de datos que contiene los conjuntos-AP detallados, serán almacenados en PostgreSQL.

Por otra parte, la arquitectura cuenta con dos herramientas de consultas que permitirán al usuario construirse información muy variada, y de una forma flexible será asistido por las aplicaciones. La primera de estas será el cliente de consulta, aquí el usuario podrá indagar sobre información almacenada en la base de datos, haciendo marcado hincapié en la información almacenada en el diccionario y los conjuntos-AP detallados, podrá realizar preguntas para cualquiera de los atributos de textos previamente procesados, y además recibirá sugerencias sobre los términos consultados. La segunda será el Servidor OLAP de libre disposición Wonder OLAP Server, construido cien por ciento utilizando software libre. Aquí el usuario podrá construir cubos multidimensionales donde pueden formar parte de este, dimensiones que respondan a atributos de texto libre, se podrán realizar las operaciones clásicas del modelo multidimensional (Roll up, Drill Down, Slice y Dice) sobre dichas estructuras, obtener reportes y gráficos muy flexibles para facilitar la interpretación de los resultados brindados por el sistema.

De forma resumida, a nuestros efectos, los principales componentes de esta arquitectura serían:

BD inicial: La base de datos almacena diferentes tipos de datos, incluyendo atributos textuales. Aquellos atributos textuales con relevancia en el proceso de toma de decisiones pueden ser transformados, a fin de realizar consultas simples incluyendo estos atributos. A pesar de que nosotros estamos manejando una base de datos Relacional Orientada a Objetos, cualquier otro tipo de base de datos puede ser considerada.

Módulo de manejo de textos cortos: En este módulo, los atributos textuales de la base de datos son procesados a fin de obtener la forma intermedia. La nueva estructura de conocimiento nos permitirá consultar todos los atributos de la base de datos y llevar a cabo cualquier proceso de sumarización, Datawarehousing, etc.

BD modificada: Esta base de datos almacena las nuevas estructuras de conocimiento correspondientes a los atributos textuales de la base de datos inicial. Estas estructuras están basadas en los conjuntos de términos obtenidos por el algoritmo Apriori. Esta es la razón por la que son llamadas conjuntos-AP y estructuras-AP. Diferentes datos intermedios son obtenidos en el proceso de obtener dichos conjuntos. El diccionario está compuesto por todos los términos relevantes de cada atributo textual (una vez removidas las stop-words, preprocesados los acrónimos, etc) y por los itemsets frecuentes y su soporte. El diccionario y los conjuntos-AP expandidos con su soporte son almacenados en la base de datos con conocimiento. Esta base de datos puede ser usada para actualizar la

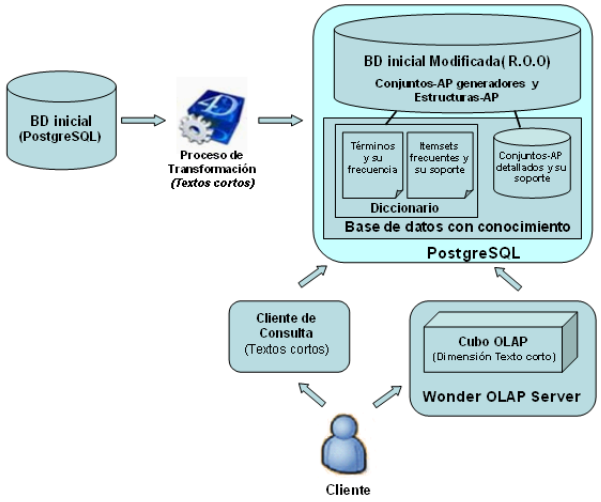


Figura 2: Arquitectura del Sistema.

base de datos Modificada, así como para completar las respuestas del sistema al Cliente de Consulta, sugerir un nuevo término por el que se puede consultar, etc.

Cliente de Consulta: Mediante este módulo, el cliente realiza consultas sobre la BCD, con operaciones específicas para los atributos textuales, tal como, consultar las estructuras de la base de datos con conocimiento.

Servidor OLAP: El usuario puede construir cubos multidimensionales donde una o más dimensiones pueden estar relacionadas a los atributos textuales de la base de datos. La nueva representación de estos atributos, nos permite realizar las operaciones usuales del modelo multidimensional (Roll up, Drill Down, Slice y Dice). Concretamente, el Wonder Servidor OLAP incluido en la figura 2 es un desarrollo propio sobre el cuál se puede encontrar información en [9].

5 Desde un atributo textual hasta la estructura-AP: el mecanismo para obtener el TDA

La metodología presentada es aplicada a atributos textuales de la base de datos, con la meta de obtener una nueva representación estructurada que mantenga la semántica de dichos atributos [5]. Esta representación no debe ser estática, sino ser un modelo matemático con operaciones sobre los datos, esto es, ser un Tipo Abstracto de Datos (TDA). De esta forma, podemos implementar las estructuras en un Sistema de Gestión de Base de Datos Objeto Relacional (SGBDOR), así como sus operaciones, en orden de consultar la base de datos por términos y realizar consultas semánticas, minería, y procesos como OLAP y Datawarehousing.

Un mecanismo es utilizado para obtener los conjuntos-AP desde los datos originales y formar la estructura-AP y el TDA. Sea R una relación con atributos $\{A_1, A_2, A_T, ..., A_n\}$, donde A_T representa un atributo que es de tipo texto y sin una estructura predecible. Por la heterogeneidad y estructura en sí que puede tener A_T, el conjunto de valores que puede tomar, sobre el dominio en el que se esté trabajando, se hace impredecible, por lo que se hace necesario obtener una forma de representación intermedia, que nos permita realizar consultas más poderosas que las tradicionales verificaciones de si el atributo es igual a un determinado valor, o si sigue algún patrón especificado en una máscara. El mecanismo seguido, de forma general, para tratar con A_T desde su representación original hasta realizar consultas que traten con su semántica, es el siguiente:

1. Realizar el proceso de limpieza sobre A_T, teniendo en cuenta la sustitución de sinónimos y acrónimos, además de tener en cuenta, la posible separación de frases diferentes en una misma tupla con uno o varios caracteres separadores (Ejemplo: signo +, punto y coma, coma).

2. Eliminar las stops-words y crear el diccionario con los términos relevantes que posee el A_T y se obtiene además la base de datos transaccional donde los atributos son los diferentes términos del diccionario y cada tupla corresponde a un record. El valor de un determinado atributo en una tupla será 1, si el término del atributo aparece en el record original, en otro caso será 0.

3. Obtener los Itemsets frecuentes a partir de la base de datos transaccionales siguiendo el algoritmo Apriori.

4. A partir de los Itemsets frecuentes se calculan los Itemset Maximales que son considerados los conjuntos-AP generadores, con ellos se forman las estructuras-AP que en su conjunto serán el dominio considerado para el TDA obtenido para cada tupla de la base de datos fuente.

5. Para obtener el valor del TDA para cada fila en el atributo A_T, se utilizará el fichero de Metadatos que contiene para cada fila la representación de A_T removiendo las stops-words. Cada valor de las filas de dicho fichero se intersecará con la estructura-AP y el resultado final será el que se escriba en otra columna A_T como valor de TDA para AT .

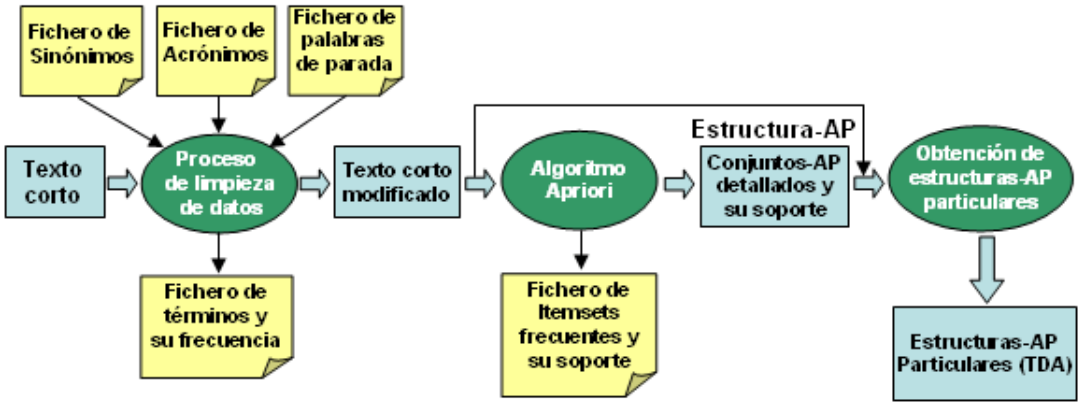


Figura 3: Proceso de transformación de un texto corto a TDA

6. Una vez que se tiene el atributo A_T que contiene el valor del TDA para el texto libre, se está en condiciones de realizar operaciones de consultas más complejas que incluso encierren resultados semánticos. La implementación de estas operaciones de consulta serán posibles utilizando las especificaciones de tipos abstractos de datos que posee SQL-99, para ello se creará el TDA Tconjunto-AP con sus métodos asociados que será el tipo de dato que tome la columna A_T. Además se creará la clase Testructura-AP para representar el dominio sobre el que puede tomar valores el TDA, esta clase tendrá los métodos asociados a las consultas que se pueden realizar sobre A_T. Finalmente se creará la clase TQuery que será la que se utilice directamente por el usuario final para tramitar sus consultas sobre la BD, la clase TQuery implementará los métodos necesarios para determinar qué funcionalidad de las que implementa TAP-Structure es necesario desencadenar para responder a una consulta determinada, así como contendrá un método *Suggest* para sugerir al usuario posibles resultados o términos a consultar.

Las tareas antes mencionadas, para su mejor estudio, y desarrollo a mayor profundidad, las agrupamos de la forma siguiente:

1. Limpieza de Datos.
2. Obtención de la estructura-AP global de conocimiento.
3. Obtención de la estructura-AP particular para cada tupla de la base de datos.

Para la realización de la fase experimental, utilizaremos un ejemplo de datos médicos, tanto para clarificar las ideas como para mostrar como el proceso va siendo llevado a cabo en un caso real. Este ejemplo experimental es descrito brevemente de la forma siguiente:

Conjunto de datos experimental: El conjunto de datos está compuesto por atributos textuales de la base de datos de Operaciones Quirúrgicas del Hospital Clínico "San Cecilio" de Granada, España. De 25527 registros de datos, se extrajeron los atributos textuales correspondientes a la Cirugía Propuesta y el Diagnóstico Médico. El contenido de estos atributos es corto (desde uno a cincuenta términos), y ellos pueden incluir una o más frases.

En la tabla 1 se puede observar un extracto del TDA resultante correspondiente al atributo Cirugía de dicho conjunto de datos. Como se puede observar, inicialmente ellos están en lenguaje natural, incluyendo signos de puntuación, abreviaturas y hasta símbolos como " + ". Por consiguiente, la tarea de preprocesamiento y limpieza de datos es imprescindible para el caso de datos médicos.

6 Conclusiones

La metodología presentada en este trabajo nos da el procedimiento a seguir para manejar atributos textuales en bases de datos (Relacionales, Relacional Orientado a Objetos, etc.). Con un preprocesamiento del texto y la obtención de los itemsets frecuentes, siguiendo el algoritmo Apriori, puede ser construida la estructura global de conocimiento que encierra la

Tabla 1: Ejemplos de textos originales y su TDA correspondiente.

Dato de tipo Texto	**TDA asociado**
CURA DE ABCESO PERIANAL	{CURA, ABSCESO, PERIANAL}
HISTERECTOMIA ABDOMINAL	{ABDOMINAL, HISTERECTOMIA}
EMBOLECTOMIA BYPASS AXILO-FEMORAL IZDO	{AXILO-FEMORAL, BY-PASS, EMBOLECTOMIA, IZQUIERDO}
EECC + L.I.O. O.I.(20.O)	{EECC}, {LIO, OI}
ARTROPLASTIA DE RODILLA DCHA	{ARTROPLASTIA, RODILLA},{ DERECHA }

semántica del atributo textual. Dicha estructura tiene un modelo matemático que nos permite considerarla como un TDA fácilmente implementable. Hemos aplicado esta metodología a un caso real y los resultados refuerzan nuestra hipótesis inicial.

En el futuro, pretendemos introducir incertidumbre y soporte en las estructuras del modelo y sus operaciones. Así mismo, pretendemos crear el modelo multidimensional que incluya la nueva representación obtenida para los atributos textuales.

Referencias

[1] R. Agrawal, T. Imielinski, A. Swami (1993). Mining Association rules between set of items in large databases. In *Proceedings of ACM SIGMOD Conference*, May, 1993, Washington DC, USA.

[2] R. Agrawal, R. Srikant (1994). Fast Algorithms for mining Association rules. In *Proceedings of VLDB*, Sept, 1994, Santiago, Chile.

[3] F. Berzal, J.C. Cubero, N. Marín, J.M. Serrano (2001). TBAR: An efficient method for association rule mining in relational databases. *Data and Knowledge Engineering*, Volume 37, Number 1, pp. 47-64.

[4] J. Celko (2005). *SQL for smarties. Advanced SQL Programming*. Morgan Kaufmann.

[5] C. Justicia, M.J. Martín-Bautista, D. Sánchez, M.A. Vila (2005). Text Mining: Intermediate Forms for Knowledge Representation. In *Proceedings of the conference Eusflat'2005*, pages 1082-1087, Barcelona, Spain, September 2005.

[6] N. Marín, M.J. Martín-Bautista, M. Prados, M.A. Vila (2006). Enhancing short text retrieval in databases. In Proceedings of the Conference on Flexible Query-Answering Systems (FQAS), Milan, June 2006.

[7] M.J. Martín-Bautista, M. Prados, M.A. Vila, S. Martínez-Folgoso (2006). A knowledge representation for short texts based on frequent itemsets. In Proceedings of the 11th Conference of Information Processing and Management of Uncertainty (IPMU), Paris, July 2006.

[8] R. Mack, M. Hehenberger (2002). Text-based knowledge discovery: search and mining of life-sciences documents. *Drug Discovery Today*, volume 7, number 11 (Suppl.)

[9] S. Martínez-Folgoso, J. Recio, E. Rodríguez. "Wonder Olap Server". Solución OLAP de libre disposición. UXI. N.9 (Diciembre 2007). Revista de Software Libre de la UCI (Universidad de las Ciencias Informáticas).

[10] M. Prados, C. Peña, B. Prados-Suárez, M.A. Vila (2006). Generation and use of one ontology for intelligent information retrieval from electronic health medical record. In *8th International Conference on Enterprise Information Systems*, May 2006, Cyprus (to appear).

[11] G. Salton, M.J. McGill (1983). *Introduction to Modern Information Retrieval*. McGraw-Hill.

TEXT KNOWLEDGE MINING: AN APPROACH TO TEXT MINING

C. Justicia de la Torre

Servicio de Informática
Universidad de Granada
consuelo@ugr.es

M.J. Martín-Bautista, D. Sánchez[1], I. Blanco

Depto. Ciencias de la Computación e I.A., Universidad de Granada
E.T.S.I.I.T., C/Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
{mbautis,daniel,iblanco}@decsai.ugr.es

Abstract

In this paper we review several alternative views of text mining (TM) as proposed by different authors, and we introduce our own view. We propose a classification of TM techniques on two main groups: techniques based on *inductive inference*, that we call *text data mining* (TDM, comprising most of the existing proposals in the literature), and techniques based on *deductive* or *abductive inference*, that we call *text knowledge mining* (TKM). To our knowledge, the TKM view of TM is new though, as we shall show, several existing techniques could be considered in this group. We discuss about the possibilities and challenges of TKM techniques. In addition, we sketch a new TM technique we proposed in a previous work, that can be fully considered in the TKM group. We also discuss about the application of existing theories, like fuzzy sets, in possible future research in this field.

Keywords: Text Mining, Text Knowledge Mining, Inference, Fuzzy Sets.

1 INTRODUCTION

Text mining (TM) refers to the discovery of non-trivial, previously unknown, and potentially useful knowledge from a collection of texts. Since its origin, text mining has been considered an analog of data mining (DM, interpreted as Knowledge Discovery in Databases, or KDD) applied to text repositories. TM is very important since nowadays, around 80% of the information stored in computers (not considering audio, video, and images) consist of text.

What seems to give text mining a clear distinction from data mining is that the latter deals with structured data, whereas text presents special characteristics and its explicit appearance is basically unstructured. Following this vision of TM as DM on unstructured data, most of the approaches to TM have been mainly concerned with obtaining structured datasets from text, called *intermediate forms* (IF), on which usual DM techniques are applied. Table 1 shows the most usual IF that have been proposed and the minimal text units (atomic data) they are comprised of. For example, the *bag of words* represents a piece of text by means of a set of weighted terms, like that employed in Information Retrieval (IR). For a review of IF see for example [5]. DM techniques employed on the IF forms of table 1 include clustering, dependence analysis, and association rules among others.

Text Unit	**Intermediate Form**
Word	Bag-of-Words
	N-grams
Concept	Concept Hierarchy
	Conceptual Graph
	Semantic Graph
	Conceptual Dependence
Phrase	N-phrases
	Multi-term text phrases
	Trends
Paragraph	Paragraph
	N-phrases
	Multi-term text phrases
	Trends
Document	Document

Table 1: Minimal text units and their related Intermediate forms

However, textual data is not inherently unstructured.

[1]Corresponding author.

On the contrary, text is characterized by a very complex implicit structure that has defeated many representation attempts, with a very rich semantics. On this basis, several authors disagree with the most usual vision of TM as applying DM techniques, because these are designed to work on very simple data structures with a very limited expressive power. By employing these data structures as IF, we are loosing most of the semantics of text; in addition, when using DM techniques on richer semantic structures, we are exploiting a very small part of their expressive power.

Following this discussion, Hearst distinguishes in [8] two types of TM; first, the aforementioned vision of TM as DM techniques providing simple patterns and regularities. Second, what she calls *real TM*, that obtains new and useful knowledge pieces with semantics as richer as that of text. As an example of the latter, she cites the work of Swanson and Smalheiser [21] in the Arrowsmith project. They provide a software tool able to employ the "causal" relationships that appear in the medical literature, as appearing in titles of papers in the MEDLINE database, in order to generate new, previously unknown hypothesis. The following example is a quote from Hearst [8]:

> "For example, when investigating causes of migraine headaches, he [Swanson] extracted various pieces of evidence from titles of articles in the biomedical literature. Some of these clues can be paraphrased as follows:
>
> - stress is associated with migraines
> - stress can lead to loss of magnesium
> - calcium channel blockers prevent some migraines
> - magnesium is a natural calcium channel blocker
> - spreading cortical depression (SCD) is implicated in some migraines
> - high levels of magnesium inhibit SCD
> - migraine patients have high platelet aggregability
> - magnesium can suppress platelet aggregability
>
> These clues suggest that magnesium deficiency may play a role in some kinds of migraine headache; a hypothesis which did not exist in the literature at the time Swanson found these links."

The discussion about what is and what is not TM has been also considered by other authors, together with reflections about the importance of knowledge and its representation in TM. In this paper we offer our own view of TM. In our view, TM techniques can be classified into two main groups: techniques based on *inductive inference*, that we call *text data mining* (TDM, comprising most of the existing proposals in the literature), and techniques based on *deductive* or *abductive inference*, that we call *text knowledge mining* (TKM). Usual DM techniques, based on induction, fall in the first group. The technique for generating new hypothesis proposed by Swanson and Smalheiser can be considered, as we shall see, as based on abduction; in this sense, is an example of technique of the second group.

To our knowledge, the TKM view of TM is new though, as we have just seen, some existing techniques could be considered in this group. We discuss about the possibilities and challenges of TKM techniques. In addition, we sketch a new TM technique we proposed in a previous work, that can be fully considered in the TKM group. We also discuss about the application of existing theories, like fuzzy sets, in possible future research in this field

2 TEXT KNOWLEDGE MINING

2.1 MOTIVATION AND DEFINITION

DM techniques for KDD are based on inductive inference, i.e., starting from a set of particular cases described by some (structured) data model, they obtain general patterns consistent with the particular examples. These patterns are subject to objective and/or subjective assessment; the former is based on statistics in general, while the latter takes into account the background knowledge (BK) in the domain in order to assess the novelty, usefulness, and non-triviality of the patterns, in accordance with the objectives of KDD. Inductive inference is the natural way to obtain new knowledge from databases since they contain particular examples, simple facts, arranged in data structures.

As we have seen, the usual view of TM consists of obtaining a structured dataset from text and then applying on it the usual inductive DM techniques. However, in our opinion, this strong association between *mining* and *inductive learning* that seems to dominate the research in TM, comes from the fact that DM was the first "mining" paradigm. Remarkably, the word *induction* does not appear in the definition of TM. TM is concerned with obtaining new, non-trivial, and potentially useful knowledge from text repositories stored in computers. Though inductive approaches have been shown to be very useful, they do not seem to be the "natural" choice for TM as they are for DM.

In our view, the main difference between data and text is that the former is represented by means of data mod-

els whilst the latter is represented by natural language. And these representations have very different characteristics. Data models have a limited expressive power, mainly restricted to the structured representation of simple facts. On the contrary, natural language has a far richer expressive power, and it is able to represent not only simple facts like those stored in a database, but general knowledge of almost any kind, from relations like rules to complex procedures. Unfortunately, natural language has some well-known disadvantages like lack of a simple, computationally easy to manage structure, vagueness, ambiguity, etc.

Summarizing, and leaving apart the issue of structure and manageability, the main difference between data and text can be stated as follows:

- Data represents a collection of simple facts, particular cases, from where knowledge can be obtained in a natural way by means of inductive inference.
- On the contrary, text expressed in natural language is a representation of general knowledge, a knowledge base [8], whose semantics is usually much richer than simple facts, on which the natural way of obtaining new knowledge is by means of deductive and abductive inference.

Let us point out that

- the idea of using deductive and/or abductive inference does not contradict the definition of TM, since there is no restriction about what kind of techniques should be employed to obtain the new knowledge, and
- though we consider deductive and abductive inference to be the natural choice for TM, this does not means that TM based on inductive inference is *not* TM. Almost all the TM approaches existing in the literature, that have been shown to be very useful in practice, are based on induction.

That is, by proposing TM based on non-inductive inference, we propose a new direction in TM that, as we shall show, leads to techniques able to provide new knowledge from text that cannot be obtained by approaches based on induction. In summary, we propose to classify TM techniques into two groups:

- **Text Data Mining (TDM) techniques** obtain new knowledge by applying inductive inference on a structured data representation obtained from text. Most existing TM techniques fall in this category.
- **Text Knowledge Mining (TKM) techniques** obtain new knowledge by applying deductive and/or abductive inference on a structured knowledge representation of the semantic content of text.

Notice that we employ TKM meaning "mining by reasoning using knowledge contained in text", i.e., the process of obtaining new knowledge taking the knowledge contained in the texts as the starting point. We understand TKM as a particular case of what we call *knowledge mining* (KM), the latter defined as "obtaining non-trivial, previously unknown, and potentially useful knowledge from knowledge repositories". We think that TKM techniques are examples of what Hearst call "real text mining" in [8] because i) the real knowledge contained in text is employed instead of text metadata or IF with a poorer semantic content, and ii) new, non trivial knowledge is derived from data, with the same expressive power.

The phases of the TKM process are exactly the same as those of TDM, i.e., first text refining for obtaining a computationally manageable IF representing the text, a mining procedure to obtain new knowledge, and a final step for assessing and filtering the knowledge obtained. However, there are substantial differences between TKM and TDM in the way these procedures are performed and the challenges they pose. We shall go into details in section 3.

2.2 KM: WHAT WE DO NOT MEAN

KM has been employed by some authors in a different sense from that introduced in the previous section. In [17], KM is employed to specify that the IF employed contains knowledge richer than simple data structures (though inductive DM techniques are used on that IF). In [10], knowledge mining is employed to indicate that BK and user knowledge is incorporated in the mining process in order to ensure that the IF and the final patterns contain only those concepts interesting for the user; again, inductive DM techniques are used. In [19] knowledge mining refer to using BK to evaluate the novel and interesting patterns after an inductive process.

Another term appearing in the literature with a different meaning with respect to our proposal is "deductive mining". We have not found a proper definition, but references to the work [11]. In this paper, deductive mining refer to a group of text mining techniques, where the better known example is said to be the Information Extraction (IE). IE is frequently referred as "the mapping of natural language texts into predefined, structured representation, or templates, which, when filled, represent an extract of key information

from the original text"; or alternatively "the process of filling the fields and records of a database from unstructured text".

IE appears in relation to text mining in many papers [9, 15, 16], but it is concerned mainly with the task of obtaining structured intermediate representations from text, that is, with translating the knowledge existing in text into another, structured, representation; however, no new knowledge is derived. The term "deductive" is applicable to Information Extraction since it departs from a knowledge about the kind of structures to be filled and rules about how to find instances of values in text to fill them.

However, the knowledge employed for the filling task is not the knowledge contained in the text, but *metaknowledge* about text and how to translate it into a predefined structure, and the deductive inference performed is just to translate from one representation to another, but no new knowledge is discovered. This fact is pointed out by Hearst [8] in relation to other techniques coming from computational linguistics and applied to particular problems like automatic augmentation of Wordnet relations. In [7], metaknowledge is employed for these purposes for so-called KM. In our view, IE performs TDM but not TKM; however, it can play a very important role in obtaining intermediate representations for TKM.

The use of deduction in combination with data mining has been proposed by some authors referring to the incorporation of deductive capabilities to mining tools in order to improve the assessment of the knowledge obtained [20, 6]. In [13], the term "deductive data mining" is employed to indicate that the knowledge should be obtained mathematically and taking into account the hidden assumptions about data. Finally, in [18], a "deductive data mining" paradigm is proposed, where the term "deductive" refer to the fact that the discovery process is performed on the result of user queries (seen as the result of a deductive process) that limit the possibility to work on corrupt data. In all these cases, the generation of new knowledge is purely inductive.

2.3 RELATED WORK

Related to TKM, an interesting summary and discussion about [8] is provided in [12]. In this paper, the authors distinguish between *non-novel investigation*, like IR, in which no new knowledge is generated; *semi-novel investigation*, like standard text mining and KDD, in which patterns/trends that already exist in data are discovered, and *novel investigation* or *knowledge creation*, in which new knowledge, telling something about the world outside the data collection itself, is created. The latter includes solutions to problems like, "how can the linguistic features of text be used to create knowledge about the outside world?", "does a new theme reflect reality?", and "which business decisions are implied?".

For achieving these goals of knowledge creation, the authors propose in [12] what they call *Intelligent Text Mining*, i.e., the use of i) interaction between users and mining tools, and ii) AI techniques (no particular technique is specified). However, these techniques are proposed for the assessment of knowledge only. No indication about the kind of inference employed for obtaining the new knowledge is provided.

Some advanced TM techniques that can be considered TKM are related to literature-based discovery [3, 21, 14, 24]. The objective of these techniques is to provide hypothesis in the form of possible relations between concepts appearing in the literature about a specific topic. In this sense, they can be seen as abduction-based TKM techniques, though some authors consider that their work is not about discovery but about assisting human experts in formulating new hypothesis by means of an interactive process [14, 24]. The process is performed usually on the title and abstract of papers in a bibliographic database. The main work has been made in the medical field by using the MEDLINE database [21, 14, 24], plus some additional BK like the UMLS thesaurus in some cases [24]; however, there are results in other areas [3].

The basic idea of literature-based discovery in [3, 21, 24] is the following: let $\{A_1, \ldots, A_n, B_1, \ldots, B_m, C\}$ be concepts and suppose we are interested in finding hypothesis involving possible relations between C and some of the A_i. Suppose that no A_i and C appear simultaneously in the collection, but they can appear simultaneously with some of the B_i. Then, if the set of B_j that appear simultaneously with A_i and C (obviously, in different documents) is large enough, a relation between A_i and C can be hypothesized. In [14], several statistical measures are employed. This procedure has provided new hypothesis that have been later confirmed by experimental study, as that in the introduction or that in [24]. Let us remark that C and restrictions about what concepts can play the role A or B are to be specified in advance.

A key point in this approach is the identification of concepts to obtain the IF of documents. The main problem with this point is that a lot of BK about the text is necessary, and varies a lot from one domain to another. In addition, let us remark that i) the relation between concepts is based in most of the cases on their simultaneous appearance in the same title or abstract only, and ii) the generation of hypothesis does

not follow a formal procedure. However, it is inspired on one of the ways humans generate hypothesis. As stated in [24], using a better semantical analysis seems to be promising. Remarkably, some formal model for conjectures, consequences and hypothesis have been published recently [23], potentially providing a formal tool for literature-based discovery and TKM.

A TKM technique was proposed in [4] for discovering contradictions in a collection of texts. The starting point is a set of text documents $T = \{t_1, \ldots, t_n\}$ of arbitrary length. The objective is to discover possible contradictions that may be deduced from the content of these texts. For this purpose, a first order logic representation in clausal normal form of the content of text is employed as IF, and a deductive inference procedure based on resolution is performed.

The IF is obtained by a semi-automatic procedure in two steps. First, an interlingua representation based on United Nation's standard UNL [1, 4] is obtained from text. This has the advantages that there are UNL encoders available for different languages, hence eliminating the multilinguism problem typical of TM, and that UNL is a computationally manageable representation. From the UNL representation, the first order clausal form is obtained by means of IE techniques. BK in the form of either text or clauses can be incorporated to the process as an additional document.

The mining process is performed by using a resolution-based deductive inference procedure. This is applied to every subset of texts looking for contradictions. The procedure is similar to the levelwise + candidate generation exploration in frequent itemsets discovery, but replacing the minimum support criteria by non-contradiction in the KB obtained by the union of the IF of texts. First, individual text are evaluated to discard the possibility that they contain a contradiction. This corresponds to the exploration of the first level of the lattice of $\mathcal{P}(T)$ with respect to inclusion. Texts containing contradictions are not considered in the following level. Candidate generation for the next level is performed in the same way as in frequent itemset mining in the successive levels. The final result is a collection of subsets of texts in the collection that generate a contradiction. In addition, intermediate clauses obtained by the resolution process are also new pieces of knowledge obtained by this TKM technique.

Looking for contradictions is very useful. It can be employed to assess the consistency of the text collection, or to assess the validity of a new text to be incorporated, in terms of the knowledge already contained in the collection. In addition, if we consider collections in which contradictions are known to be (like for example repositories of papers containing opinions about topics), the search for contradictions can be employed in order to separate different groups of texts and, consequently, of authors in terms of consistency of their ideas. Finally, let us remark that reasoning with ontologies can be reduced in the end to checking the consistency (i.e., non contradiction) of KB, hence looking for contradictions can be seen as a basic procedure for performing different types of reasoning.

3 CHALLENGES OF TKM

As DM owe a lot to machine learning techniques, statistics, and other previously existing areas, so TKM will benefit of the large amount of existing results in areas like knowledge representation (KR), reasoning algorithms for performing deductive and abductive inference, and knowledge-based systems (KBS), as well as NLP techniques. As in the case of DM, these techniques must be adapted for TKM; new specific research results are needed as well. In the following we detail some of the challenges that TKM must address in several aspects.

3.1 OBJECTIVES

KBS and TKM systems have very different objectives that affect the way techniques coming from KR and KBS are adapted to TKM:

- In a KBS, a KB is built containing the knowledge needed by the KBS to solve a specific problem. In TKM, the knowledge managed by the systems is collected from a collection of texts, each of one can be regarded as a particular KB; however, these KB in texts are not generated for the specific purpose of building an effective KBS, but generally as historical or normative reports.
- KBS intends to give answer to *every possible question*, or to solve any possible problem posed to the system. On the contrary, the objective of TKM is just to provide new pieces of non-trivial, previously unknown, and potentially useful knowledge derived from the collection of text. A TKM is not required to obtain *every such piece of knowledge*; in particular, such set of knowledge pieces is unknown by definition.
- A KBS performs a reasoning process as a result of a query, stated by the user of a consultive system, or by the necessity to perform an action in a certain situation in the case of intelligent agents. In this sense, the inference procedures employed by KBS are in general oriented and optimized to give an answer to specific questions. On the contrary, a TKM system is most generally asked to

find new knowledge without specifying a specific question. Of course, the TKM process can also be focused by means of specific queries (e.g. "find new knowledge involving X", with X being and object, a kind of action, or anything else).

3.2 KNOWLEDGE REPRESENTATION

TM requires to translate text into a computationally manageable IF. We have introduced some of the most employed IF for TDM in table 1. Techniques coming form NLP and IR have been employed for obtaining IF from text. This step of the TM process is crucial and poses several challenges, common to TDM and TKM, like obtaining IF from multilingual text collections, defining structures for IE, identifying semantic links between new concepts, relating different identifiers, etc.

It is important to remark that the IF and inference procedures employed are obviously linked to each other, e.g., in order to perform logical deduction, a logic-based representation of the knowledge in texts is needed. Usually, abductive and (specially) deductive inference require a more expressive, semantically richer IF that inductive inference, that are more difficult to obtain from texts in general. However, let us remark that many TDM techniques perform inductive inference on IF that are also suitable for TKM, like semantic graphs and conceptual dependencies in table 1, that allow certain kind of non-inductive inference.

A key problem with obtaining IF for TKM is that the current state-of-the-art techniques for translating text into those IF are mainly semiautomatic, i.e., they require the participation of human experts in order to verify that the translation is correct. On the other hand, several ongoing projects deal with the possibility of expressing all or at least key parts of the knowledge in text by means of KR models directly. This is the case for example of work on the Semantic Web, where ontologies represented by description logics or languages with the same expressive power are employed in order to represent the semantics of concepts and their relations [2]. TKM in the Semantic Web seems a promising direction since efficient deductive inference techniques are also available.

The objective though of an IF for TKM is not to represent every possible aspect of the semantic content of text, but those related to the kind of inference we are interested in, that are on its turn determined by the kind of knowledge we want to discover. Even the lack of a text analyzer able to translate without loss from text to the IF of our choice has no more negative effect that preventing some pieces of knowledge from being discovered. However, as we discussed in the previous section, TKM is not expected to be exhaustive in obtaining new knowledge and, in this sense, this is not a major problem. Of course, better analyzers provide more effective TKM systems.

On the other hand, if the analyzer obtain an inexact representation of text in a IF, this can affect the knowledge discovered, leading to false discoveries. Minimizing such misleading discoveries (that, in DM and TDM, use to be the larger part of the patterns discovered, together with the uninteresting ones) is an important challenge. In this sense, not only the misfunctioning of text analyzers, but also the reliability of text itself and the adequacy of the IF introduce new problems in TKM. The knowledge in a KBS is assumed to be reliable and consistent, but the same cannot be assumed in general in a collection of texts, even before the IF of each one is obtained. Coping with the reliability of knowledge in the inference process is a major challenge in TKM. In this point, Soft Computing techniques like fuzzy or probabilistic representation of opinions, importance, reliability and confidence of knowledge may be very useful.

The possible contribution of Soft Computing is far larger. Several extensions to TDM in order to cope with imprecise data, and fuzzy extensions of IF, have been proposed [5]. Other Soft Computing techniques are being employed currently in DM and can be adapted to TDM. However, Soft Computing can contribute a lot also to TKM. Since text use to describe changing world, non-monotonic logics are needed for certain TKM tasks. Many of the most important thinkers in the Soft Computing field claim that one of the most promising applications of the theory of fuzzy sets and related areas is the representation of, and reasoning with, the semantics of natural language. It is well known the suitability of fuzzy sets for this purpose due to the use of imprecise terms and connectives in human discourse. Logics related to fuzzy sets allow us to capture the semantics of imprecise concepts and to model the subtleties of the discourse relative to conjunction, disjunction, negation and implication. An example of this that, in our opinion, is very close to (if not fully) a TKM process has been presented in [22].

3.3 BACKGROUND KNOWLEDGE

The necessity of incorporating BK to TM is widely recognized. However, it is specially crucial in TKM, since usually text does not contain commonsense knowledge and specific domain knowledge that is necessary in order to perform TKM. For example, a text contaning "A is father of B" and "B is father of A" is not recognized as contradictory unless commonsense knowledge is considered. The recopilation and representation by IF of BK is a challenging problem.

BK can contribute to TKM in some other very interesting aspects. First, it allows us to "create" new knowledge that is not fully contained in the collection of text, but can be derived from a combination of text and BK. This was a requirement of *intelligent text mining* in [12]. Another interesting application of BK is in the assessment of knowledge, see section 3.5.

3.4 REASONING AND COMPLEXITY

Complexity is one of the important aspects of automatic inference in KBS. Any KR paradigm, from which maybe description logics is one of the most visible nowadays, must find a tradeoff between expresiveness and complexity of reasoning. Obviously, TKM systems are also affected by these facts. However, the problem is somehow less hard than in the case of KBS or ontological reasoning.

First, TM and DM are inherently hard problems in terms of complexity. Researchers in these fields are used to cope with exponential algorithms, and many strategies for designing efficient mining algorithms are available. Second, as we have pointed out previously, TKM is not intended to be exhaustive in the derivation of new knowledge. Hence, time and/or space limitations can be stated in the algorithms so that the knowledge obtained is just the knowledge that has been possible to obtain with those resources.

3.5 ASSESSING THE RESULTS

In DM, objective measures intend to assess the statistical significance of results as a first filter to discard uninteresting patterns; then, subjective assessments about the novelty and usefulness of the patterns obtained are employed as a final filter. Usually, this process is performed directly by human experts, though some methodologies for subjective pattern evaluation have been developed. We consider the following aspects regarding the assessements of results in TKM:

- The reliability of the results in TKM comes from several aspects. First, the validity and reliability of the texts; second, the inference procedures employed. Formal deductive procedures are mathematically sound and consistent, for instance; there are also formal models of abductive inference [23]. Uncertainty in the IF as a result of the IE process may be propagated to the results as well.
- Non-triviality of the results is to be evaluated by the expert user. However, notice that deductive and abductive inference are not trivial reasoning procedures in general.
- Novelty of the results may be evaluated in terms of the BK employed: the new knowledge obtained is novel if it was not previously contained neither in the text collection nor in the BK.
- Usefulness is to be evaluated by the experts. However, a possible way to look for useful knowledge is by leaving the expert to specify the topic (object, action, etc.) he is interested in obtaining knowledge about, prior to the mining process.

4 CONCLUSIONS

This paper introduces the idea of text knowledge mining (TKM) as a particular case of *knowledge mining*, the latter defined as "obtaining non-trivial, previously unknown, and potentially useful knowledge from knowledge repositories". To our knowledge, this idea does not exists in the literature, though the term "knowledge mining" has been employed by different authors in different senses.

Knowledge mining differs from data mining in that the former deals with explicit knowledge, whilst the latter deals with data in which the knowledge is implicit. As a consequence, the natural way to obtain new knowledge in knowledge mining is by means of deductive and abductive inference; on the contrary, inductive inference is the natural way to proceed in data mining.

Most of the existing text mining techniques transform text into data and/or perform inductive inference on the intermediate representation obtained in order to make explicit the implicit knowledge in data. This *text data mining* approach has proved to be very useful in many applications. However, though datasets can be obtained from it, text is primarily an explicit representation of knowledge in natural language. As a consequence, in this paper we have proposed text knowledge mining as an alternative and natural way to text mining that obtains new knowledge from the explicit knowledge expressed in the text, via deductive and/or abductive inference.

This new proposal opens a wide range of possibilities in text mining. The application of non-inductive inference techniques that, remarkably, take into account a richer representation of the semantic content of text, leads to new types of text mining applications that can take advantage of the extensive available literature in natural language processing, knowledge representation, and reasoning. The adaptation to different objectives, the translation of natural language into manageable knowledge representation models, the inclusion of background knowledge, complexity issues of reasoning algorithms, and assessment of the inferred knowledge are some of the exciting challenges to be addressed in the future.

References

[1] *Web of the Institute of Advanced Studies, The United Nations University, Spanish Language Center- Universidad Politécnica de Madrid. http://www.unl.fi.upm.es/homepage.htm.*

[2] G. Antoniou and F. Van Harmelen. *A Semantic Web Primer*. The MIT Press, Cambridge, MA, 2004.

[3] K.A. Cory. Discovering hidden analogies in an online humanities database. *Computers and the Humanities*, 31:1–12, 1997.

[4] C. Justicia de la Torre, M. J. Martín-Bautista, D. Sánchez, and I. Blanco. Un enfoque deductivo para la minería de textos. In *Actas de Estylf 2006*, pages 241–246, 2006.

[5] C. Justicia de la Torre, M. J. Martín-Bautista, D. Sánchez, and M.A. Vila. Text mining: Intermediate forms for knowledge representation. In *Joint Eusflat-Lfa 2005 Conference*, pages 1082–1087, 2005.

[6] S. Greco, E. Masciari, and L. Pontieri. Combining inductive and deductive tools for data analysis. *AI Communications*, 14:69–82, 2001.

[7] U. Hahn and K. Schnattinger. Knowledge mining from textual sources. In F.Golshani and K.Makki, editors, *CIKM'97 - Proceedings of the 6th Intl. Conference on Information and Knowledge Management.*, pages 83–90, ACM, NY, USA, 1997.

[8] M.A. Hearst. Untangling text data mining. In *Proc. 37 Annual Meeting of the Association for Computational Linguistics*, pages 20–26, 1999. http://citeseer.ist.psu.edu/367814.html.

[9] H. Karanikas, C. Tjortjis, and B. Theodoulidis. An approach to text mining using information extraction. In *Proc. KMTA 2000*, 2000.

[10] K. Kaufman and R.S. Michalski. From data mining to knowledge mining. In C.R. Rao, J.L. Solka, and E.J. Wegman, editors, *Handbook in Statistics, Vol. 24: Data Mining and Data Visualization*, pages 47–75. Elsevier/North Holland, 2005.

[11] Y. Kodratoff. Rating the interest of rules induced from data and within texts. In *Proc. 12th Int. DEXA Workshop*, pages 265–269. 2001.

[12] J.H. Kroeze, M.C. Matthee, and T.J.D. Bothma. Differentiating data- and text-mining terminology. In J. Eloff, A. Engelbrecht, P. Kotzé, and M. Eloff, editors, *Proceedings of the 2003 Annual Research Conference of the SAICSIT*, pages 93–101. SAICSIT, Republic of South Africa, 2003.

[13] T.Y. Lin. Deductive data mining: Mathematical foundation of database mining. In G. Wang, Q. Liu, Y. Yao, and A. Skowron, editors, *Proc. 9th International Conference, RSFDGrC. LNAI 2639.*, pages 403–404. Springer, 2003.

[14] R.K. Lindsay and M.D. Gordon. Literature-based discovery by lexical statistics. *JASIS*, 50:574–587, 1999.

[15] A. McCallum and D. Jensen. A note on the unification of information extraction and data mining using conditional-probability, relational models. In *IJCAI03 Workshop on Learning Statistical Models from Relational Data*, 2003. http://citeseer.ist.psu.edu/mccallum03note.html.

[16] R.J. Mooney and R. Bunescu. Mining knowledge from text using information extraction. *SIGKDD Explorations*, 7(1):3–10, 2005.

[17] T. Nasukawa and T. Nagano. Text analysis and knowledge mining system. *IBM Systems Journal*, 40(4):967–984, 2001.

[18] L. Russell. Deductive data mining: Uncertainty measures. In A. Borgida, V.K. Chaudhri, and M. Staudt, editors, *Proceedings of the 5th KRDB Workshop*, pages 15.1 – 15.5, 1998.

[19] M.S. Siadaty and W.A. Knaus. Locating previously unknown patterns in data-mining results: A dual data- and knowledge-mining method. *BMC Med Inform Decis Mak.*, 6(13), 2006. Available online doi: 10.1186/1472-6947-6-13.

[20] E. Simoudis, B. Livezey, and R. Kerber. Integrating inductive and deductive reasoning for data mining. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors, *Advances in Knowledge Discovery and Data Mining*, pages 353 – 373. AAAI Press, Menlo Park, CA, 1996.

[21] D.R. Swanson and N.R. Smalheiser. Implicit text linkages betwen medline records : Using arrowsmith as an aid to scientific discovery. *Library trends*, 48(1):48–59, 1999.

[22] E. Trillas. Lógica borrosa y narrativa: Un párrafo de Vila-Matas. In *Actas de ESTYLF 2004*, pages 15–20, 2004.

[23] E. Trillas, S. Cubillo, and E. Castiñeira. On conjectures in orthocomplemented lattices. *Artificial Intelligence*, 117:255–275, 2000.

[24] M. Weeber, H. Klein, L.T.W. de Jong-Van Den Berg, and R.Vos. Using concepts in literature-based discovery: Simulating Swansons Raynaud-fish oil and migraine-magnesium discoveries. *JASIST*, 52(7):548–557, 2001.

UN MODELO LÓGICO PARA LA REPRESENTACIÓN DE REGLAS DE ASOCIACIÓN DIFUSAS MEDIANTE NIVELES DE RESTRICCIÓN*

Miguel Delgado[1], M. Dolores Ruiz[1], M. Amparo Vila[1]

[1] Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada
ETSIIT, C/Periodista Daniel Saucedo Aranda s/n, 18071, Granada
{mdelgado}@ugr.es, {mdruiz vila}@decsai.ugr.es

Resumen

Este trabajo pretende dar un modelo lógico con buenas propiedades para la representación y evaluación de reglas de asociación difusas. Para ello se van a combinar dos propuestas. La primera de ellas es un modelo para la representación y evaluación de reglas de asociación crisp. La segunda es una nueva propuesta para la representación de propiedades imprecisas (en particular para conjuntos difusos) mediante niveles de restricción que verifica todas las equivalencias lógicas del caso crisp [10]. Combinando ambas conseguiremos un modelo sólido, con unas buenas características para la representación y evaluación de reglas de asociación difusas que además generaliza de forma natural la concepción de regla de asociación crisp.

Palabras Clave: Reglas de asociación difusas, modelo lógico, niveles de restricción.

1. INTRODUCCIÓN

Las reglas de asociación difusas desde su aparición a finales de los años noventa han sufrido un gran desarrollo y se han aplicado en numerosas situaciones. Los primeros trabajos hacían uso de la Teoría de Subconjuntos Difusos propuesta por L.A. Zadeh [11] para agrupar valores imprecisos con una semántica clara y más tarde también para representar los diferentes tipos de imprecisión de la información almacenada en la base de datos. Las ventajas del uso de conjuntos difusos es por un lado el suavizado de las fronteras y por otro dotar de una representación del conocimiento que sea semánticamente significativa para el usuario. En [7], [6] podemos encontrar descritos algunos de los trabajos más importantes en la extracción de reglas de asociación difusas.

La existencia de diferentes tipos de reglas difusas sugiere que en cada momento, según las necesidades del usuario, se debería guiar el proceso para obtener el conocimiento que interese extraer. Por tanto, parece comprensible que la evaluación del cumplimiento de la asociación no debería ser independiente de su interpretación. De esta forma queda reflejado que debería ser decisión del usuario (dependiendo del conocimiento que se quiera extraer) usar un tipo de evaluación u otro que se ajuste a sus exigencias. Más aún, puesto que las medidas de evaluación pueden ser generalizadas de diversas formas, sería necesario buscar un soporte lógico o una base axiomática que refleje la elección de la medida específica. Este trabajo pretende ir en esta línea y establecer un modelo lógico que generalice y conserve de manera natural todas las características de las reglas de asociación crisp, así como las distintas medidas de evaluación.

El resto del trabajo sigue con un breve resumen de un modelo propuesto en [8] para la representación de reglas de asociación crisp que fue introducido por primera vez en [4]. Después se verán los conceptos que se usarán de una propuesta para

*Este trabajo ha sido parcialmente sufragado por los proyectos TIN2006-15041-C04-01 y TIN2006-07262.

la representación de propiedades imprecisas mediante niveles de restricción [10]. Por último se presentará una propuesta para la representación y evaluación de reglas de asociación difusas usando una combinación de lo presentado en las anteriores secciones. Terminaremos con las conclusiones y posibles líneas futuras de investigación.

2. UN MODELO LÓGICO PARA REGLAS DE ASOCIACIÓN

El modelo lógico que vamos a tratar está basado en un método desarrollado en los años sesenta por Háyek et al. [4]. Este método se llama GUHA (General Unary Hypotheses Automaton) y tiene una buena base lógica y estadística que contribuye a una mejor comprensión de dos aspectos de las reglas de asociación: su naturaleza y las propiedades de las medidas utilizadas para su evaluación. Recientemente, algunos autores han implementado un buen y rápido algoritmo [9] basado en una representación de la base de datos usando cadenas de bits. Este algoritmo se utiliza para extraer reglas de asociación en base al modelo lógico que vamos a presentar a continuación.

Sea M una matriz de datos donde sus filas $O_1, \ldots, O_n$ están asociadas a un conjunto de objetos observados y sus columnas $A_1, \ldots, A_K$ son los atributos asociados que describen los objetos. De esta forma, la entrada (i, j) de M será 1 cuando el objeto O_i presente o cumpla el atributo A_j y 0 en otro caso. Dentro del marco de la extracción de reglas de asociación, cada matriz M representará una transacción y el conjunto de todas las matrices (transacciones) la representaremos mediante una base de datos a la que llamaremos D.

D	$< O_1, A_1 >$	$\ldots$	$< O_n, A_K >$
t_1	1	$\ldots$	0
t_2	0	$\ldots$	1
$\vdots$	$\vdots$	$\ddots$	$\vdots$
t_n	1	$\ldots$	1

Para el modelo lógico que estamos presentando, un *ítem* será un par de la forma $< O_1, A_1 >$, y un *itemset* será una agregación de items usando los conectores lógicos: $\wedge, \vee, \neg$. Una regla de asociación en el modelo propuesto por [5] es una expresión del tipo $\varphi \approx \psi$ donde φ y ψ son itemsets (en el sentido anterior) derivados de una base de datos D, y el símbolo $\approx$, llamado *cuantificador*, es una evaluación o condición para el cumplimiento de la regla de asociación que dependerá de la medida usada y de la tabla de cuatro huecos (*four fold table*, 4ft) asociada a los itemsets φ y ψ. A dicha tabla la notaremos por $\mathcal{M} = 4ft(\varphi, \psi, D) = \langle a, b, c, d \rangle$ donde a, b, c y d serán enteros no negativos cumpliendo que a es el número de objetos (i.e. de filas de D) que contienen a la vez a φ y ψ, b el número de objetos que contienen φ y no ψ, y de forma análoga para c y d. Es obvio que siempre se cumple $a+b+c+d > 0$. De manera gráfica lo representamos como sigue:

$\mathcal{M}$	ψ	$\neg\psi$
φ	a	b
$\neg\varphi$	c	d

Cuando el cumplimiento de la condición impuesta por un cuantificador $\approx$ proviene de una tabla como la anterior, diremos que $\approx$ es un *cuantificador* $4ft$. La regla de asociación $\varphi \approx \psi$ será cierta en la base de datos D (o en la matriz $\mathcal{M}$) si y sólo si la condición asociada al cuantificador $4ft$, $\approx$, se satisface para la tabla $4ft(\varphi, \psi, D)$.

Según el tipo de cuantificador $4ft$ podremos expresar diferentes tipos de asociación entre los itemsets φ y ψ. En [5] y [8] podemos encontrar numerosos ejemplos. La propuesta clásica que usa las medidas del soporte y la confianza puede modelarse por el cuantificador llamado *implicación fundada*, $\Rightarrow_{minconf,minsupp}$, como sigue:

$$\frac{a}{a+b} \geq minconf \wedge \frac{a}{n} \geq minsupp \qquad (1)$$

donde $0 < minconf, minsupp < 1$ son los umbrales de mínima confianza y mínimo soporte respectivamente, y $n = a+b+c+d$ es el número total de transacciones en la base de datos D. En [3] se explica la relación entre los cuantificadores $4ft$ y las medidas de interés usadas en la evaluación y validación de reglas de asociación.

3. REPRESENTACIÓN POR NIVELES DE RESTRICCIÓN

Una propiedad imprecisa en un universo X puede ser representada por una colección de realizaciones crisp. La propuesta de [10] para representar

propiedades imprecisas mediante niveles de restricción extiende las operaciones usuales del caso crisp al difuso cumpliendo las equivalencias lógicas entre dichas operaciones.

Definición 1. [10] Un *RL-set* Λ es un conjunto finito de niveles de restricción $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verificando $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$, $m \geq 1$.

En general, el *RL*-set de una propiedad atómica representada mediante un conjunto difuso A está definido como sigue:

Definición 2. [10] Sea A un conjunto difuso definido sobre el referencial X. Entonces el *RL*-set para A viene dado por:

$$\Lambda_A = \{A(x) \,|\, x \in X\} \cup \{1\} \tag{2}$$

donde $A(x)$ es el grado de pertenencia de x al conjunto difuso A.

El *RL*-set empleado para representar una propiedad imprecisa se obtiene como la unión de los *RL*-sets de las propiedades atómicas que definan dicha propiedad.

Para representar una propiedad imprecisa en X mediante niveles de restricción usaremos una *RL-representation* definida como un par (Λ, ρ) donde Λ es un *RL*-set y $\rho : \Lambda \to \mathcal{P}(X)$ es una función que aplica cada nivel de restricción en su realización crisp que representa la propiedad imprecisa en dicho nivel. Como ejemplo, la *RL*-representation de una propiedad atómica imprecisa definida por un conjunto difuso A será el par (Λ_A, ρ_A), donde Λ_A viene dada por la ecuación (2) y $\rho_A(\alpha) = A_\alpha = \{x \in X \,|\, A(x) \geq \alpha\}$ para todo $\alpha \in \Lambda_A$.

Definición 3. [10] Sea (Λ, ρ) una *RL*-representación con $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verificando que $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$. Sea $\alpha \in (0, 1]$ y $\alpha_i, \alpha_{i+1} \in \Lambda$ cumpliendo que $\alpha_i > \alpha > \alpha_{i+1}$. Entonces definiremos

$$\rho(\alpha) = \rho(\alpha_i). \tag{3}$$

Si nos fijamos en esta definición, esta extensión para valores que no estén en el *RL*-set de la función ρ, es la natural si pensamos en un conjunto difuso A y en sus α-cortes. Usando esta definición se definirá la equivalencia entre dos *RL*-representaciones.

Definición 4. [10] Sean (Λ, ρ) y (Λ', ρ') dos *RL*-representaciones sobre X. Diremos que ambas representaciones son *equivalentes* y lo notaremos por $(\Lambda, \rho) \equiv (\Lambda', \rho')$, si y sólo si, $\forall \alpha \in (0, 1]$

$$\rho(\alpha) = \rho(\alpha'). \tag{4}$$

3.1. OPERACIONES LÓGICAS

En esta sección veremos las operaciones lógicas de disyunción, conjunción y negación, que nos harán falta para la generalización del modelo lógico al caso de reglas de asociación difusas. Las ideas básicas de cómo se definen dichas operaciones se encuentran en [10].

Definición 5. Sean P, Q dos propiedades imprecisas con *RL*-representaciones (Λ_P, ρ_P), (Λ_Q, ρ_Q). Entonces, $P \wedge Q$, $P \vee Q$ y $\neg P$ son propiedades imprecisas representadas por $(\Lambda_{P\wedge Q}, \rho_{P\wedge Q})$, $(\Lambda_{P\vee Q}, \rho_{P\vee Q})$ y $(\Lambda_{\neg P}, \rho_{\neg P})$ respectivamente, donde

$$\begin{aligned} \Lambda_{P\wedge Q} = \Lambda_{P\vee Q} &= \Lambda_P \cup \Lambda_Q \\ \Lambda_{\neg P} &= \Lambda_P \end{aligned} \tag{5}$$

y, para todo $\alpha \in (0, 1]$,

$$\begin{aligned} \rho_{P\wedge Q}(\alpha) &= \rho_P(\alpha) \cap \rho_Q(\alpha), \\ \rho_{P\vee Q}(\alpha) &= \rho_P(\alpha) \cup \rho_Q(\alpha), \\ \rho_{\neg P}(\alpha) &= \overline{\rho_P(\alpha)}, \end{aligned} \tag{6}$$

donde $\overline{Y}$ es el complemento usual de un conjunto crisp Y.

Proposición 1. [10] Las operaciones $\wedge, \vee, \neg$ entre *RL*-representaciones verifican las propiedades ordinarias de equivalencia lógica como $\neg\neg A \equiv A$, las leyes de Morgan (una de ellas $\neg(A \wedge B) \equiv (\neg A \vee \neg B)$) y la ley del tercero excluido que puede ser expresada como $A \wedge \neg A \equiv \bot$ ó $A \vee \neg A \equiv \top$, donde $\top$ y $\bot$ son las propiedades atómicas que representan respectivamente la tautología (cuya *RL*-representación se obtiene del referencial X) y la contradicción (cuya representación se obtiene de $\emptyset$).

4. UN MODELO LÓGICO PARA REGLAS DE ASOCIACIÓN DIFUSAS

En [1] se generaliza el concepto de transacción y regla de asociación para el caso en el que éstas

sean difusas.

Definición 6. [1] Sea $I = \{i_1, \ldots, i_m\}$ un conjunto finito de items. Una transacción difusa es un subconjunto difuso no vacío $\tilde{\tau} \subseteq I$.

Para cada item $i \in I$ y cada transacción $\tilde{\tau}$ tendremos que i pertenece a $\tilde{\tau}$ con grado[1] $\tilde{\tau}(i)$, donde $\tilde{\tau}(i)$ es un número real en el intervalo $[0, 1]$.

Sea A un itemset de I, es decir, un subconjunto de items, en una transacción difusa $\tilde{\tau}$. Se define el grado de pertenencia de $A \subseteq I$ a la transacción difusa $\tilde{\tau}$ como $\tilde{\tau}(A) = \min_{i \in A} \tilde{\tau}(i)$.

Usando la definición 6 una transacción será un caso especial de transacción difusa donde todos los items tienen valor 0 ó 1 según se encuentren o no en la transacción.

Ejemplo 1. Consideremos el conjunto de items $I = \{i_1, i_2, i_3, i_4\}$ y el conjunto de transacciones difusas dada en la tabla 1.

Tabla 1: Conjunto de transacciones difusas $\tilde{D}_1$

	i_1	i_2	i_3	i_4
$\tilde{\tau}_1$	0	0.6	0.7	0.9
$\tilde{\tau}_2$	0	1	0	1
$\tilde{\tau}_3$	1	0.5	0.75	1
$\tilde{\tau}_4$	1	0	0.1	1
$\tilde{\tau}_5$	0.5	1	0	1
$\tilde{\tau}_6$	1	0	0.75	1

En particular, podemos observar que $\tilde{\tau}_2$ es una transacción crisp. Algunos grados de inclusión podrían ser: $\tilde{\tau}_1(\{i_3, i_4\})$ =0.7, $\tilde{\tau}_1(\{i_2, i_3, i_4\})$ =0.6 y $\tilde{\tau}_4(\{i_1, i_4\})$ =1.

Definición 7. [1] Sea I un conjunto de items, $\tilde{D}$ un conjunto de transacciones difusas y $A, B \in I$ dos itemsets disjuntos, es decir, $A \cap B = \emptyset$. Una regla de asociación difusa $A \rightarrow B$ se cumple en $\tilde{D}$ si y sólo si, $\tilde{\tau}(A) \leq \tilde{\tau}(B)$ para todo $\tilde{\tau} \in \tilde{D}$, esto es, el grado de inclusión de B es mayor que el grado de inclusión de A para todas las transacciones difusas $\tilde{\tau}$ en $\tilde{D}$.

Esta definición preserva el significado de las reglas de asociación crisp ya que si necesitamos que $A \subseteq \tilde{\tau}$ en algún sentido, necesitaremos que también $B \subseteq \tilde{\tau}$ se cumpla, en nuestro caso sería $\tilde{\tau}(A) \leq \tilde{\tau}(B)$. De esta forma, como una transacción crisp es un caso especial de transacción difusa, una regla de asociación será un caso especial de regla de asociación difusa.

Consideremos los conjuntos difusos $\tilde{\Gamma}_A$ y $\tilde{\Gamma}_B$ definidos en $\tilde{D}$ como $\tilde{\Gamma}_A(\tilde{\tau}) = \tilde{\tau}(A)$ y $\tilde{\Gamma}_B(\tilde{\tau}) = \tilde{\tau}(B)$ respectivamente. Puesto que son dos conjuntos difusos, podemos obtener sus respectivas RL-representaciones a las que notaremos[2] por (Λ_A, ρ_A), (Λ_B, ρ_B), definidas de la forma descrita en las ecuaciones (2) y (3).

Los conjuntos de items siguientes: $A \wedge B, A \wedge \neg B, \neg A \wedge B, \neg A \wedge \neg B$ forman una partición de la base de datos $\tilde{D}$ (al igual que lo hacían $\varphi \wedge \psi, \varphi \wedge \neg\psi, \neg\varphi \wedge \psi$ y $\neg\varphi \wedge \neg\psi$ para el caso crisp). Podemos considerar los conjuntos difusos asociados definidos en $\tilde{D}$ con sus respectivas RL-representaciones a las que llamaremos $(\Lambda_{A\wedge B}, \rho_{A\wedge B})$, $(\Lambda_{A\wedge\neg B}, \rho_{A\wedge\neg B})$, etc. Observemos que los RL-sets que aparecen contendrán los mismos niveles de restricción, es decir, todos ellos serán igual a $\Lambda_A \cup \Lambda_B$.

Puesto que para cada $\alpha \in \Lambda_Y$, $\rho_Y(\alpha)$ es un conjunto crisp, podremos calcular su cardinal como viene siendo habitual, y lo notaremos por $|\rho_Y(\alpha)|$.

De esta forma para cada nivel de restricción $\alpha_i \in \Lambda_A \cup \Lambda_B$ definimos la tabla de cuatro huecos asociada $\mathcal{M}_{\alpha_i} = 4ft(\tilde{\Gamma}_A, \tilde{\Gamma}_B, \tilde{D}, \alpha_i)$ como sigue:

$\mathcal{M}_{\alpha_i}$	$\tilde{\Gamma}_B$	$\tilde{\Gamma}_{\neg B}$
$\tilde{\Gamma}_A$	a_i	b_i
$\tilde{\Gamma}_{\neg A}$	c_i	d_i

donde a_i, b_i, c_i y d_i serán enteros no negativos tales que $a_i = |\rho_{A\wedge B}(\alpha_i)|$, $b_i = |\rho_{A\wedge\neg B}(\alpha_i)|$ y análogamente con c_i y d_i.

Observemos que $\forall \alpha_i \in \Lambda_A \cup \Lambda_B$ se tiene que

$$a_i + b_i + c_i + d_i = n = |\tilde{D}|. \qquad (7)$$

Veamos que esta definición es una buena generalización del caso crisp. Supongamos que A y B son dos itemsets de una base de datos crisp D,

[1] Observemos que $\tilde{\tau}(i)$ no es más que $\mu_{\tilde{\tau}}(i)$ donde $\mu_{\tilde{\tau}} : I \longrightarrow [0, 1]$ es la función de pertenencia del conjunto difuso $\tilde{\tau}$ sobre el referencial $I =$ { conjunto de items }.

[2] A y B son conjuntos crisp de items disjuntos y $\tilde{\Gamma}_A, \tilde{\Gamma}_B$ son conjuntos difusos definidos en $\tilde{D}$.

entonces en este caso, $\tilde{\Gamma}_A(\tilde{\tau}) = \tilde{\tau}(A) \in [0,1]$ equivaldría a $t(A) \in \{0,1\}$ donde $t(A)$ viene dada por una función indicadora de la forma:

$$t(A) = \begin{cases} 1 & \text{si } A \in t \\ 0 & \text{si } A \notin t \end{cases} \tag{8}$$

con t una transacción crisp de D. Los RL-sets asociados tanto a $\tilde{\Gamma}_A$ como a $\tilde{\Gamma}_B$ serán $\Lambda_A = \Lambda_B = \{1\}$. Y las RL-representaciones asociadas serán: (Λ_A, ρ_A) y (Λ_B, ρ_B) con

$$\begin{aligned} \rho_A(1) = A_1 = \{t \in D \,|\, t(A) \geq 1\} = \{t \in D \,|\, A \in t\} \\ \rho_B(1) = B_1 = \{t \in D \,|\, t(B) \geq 1\} = \{t \in D \,|\, B \in t\} \end{aligned} \tag{9}$$

De forma análoga podemos calcular las RL-representaciones para los conjuntos $\tilde{\Gamma}_A \wedge \tilde{\Gamma}_B$, $\tilde{\Gamma}_A \wedge \neg\tilde{\Gamma}_B$, $\neg\tilde{\Gamma}_A \wedge \tilde{\Gamma}_B$, $\neg\tilde{\Gamma}_A \wedge \neg\tilde{\Gamma}_B$. Por tanto, la tabla $4ft$ para el nivel de restricción $\alpha = 1$ coincide con la tabla $4ft$ para los itemsets A y B vista en la sección 2:

$\mathcal{M}_1 \equiv \mathcal{M}$	B	$\neg B$
A	a_1	b_1
$\neg A$	c_1	d_1

Usando esta representación es fácil generalizar cualquier tipo de medida de interés usada para el caso crisp, en particular, cualquier cuantificador $4ft$. Veámoslo para los casos concretos de soporte y confianza.

Definición 8. (Soporte de un itemset) Sea $A \subseteq I$ un itemset y (Λ_A, ρ_A) la RL-representación asociada al conjunto difuso $\tilde{\Gamma}_A$ en $\tilde{D}$. Entonces, el soporte de A en el conjunto de transacciones difusas $\tilde{D}$ se define como sigue

$$sop(A) = \sum_{\alpha_i \in \Lambda_A} (\alpha_i - \alpha_{i+1}) \left(\frac{|\rho_A(\alpha_i)|}{|\tilde{D}|} \right). \tag{10}$$

Si consideramos la parte derecha de la formula (10) y usando la tabla $4ft$ asociada a los itemsets A y B para el nivel de restricción α_i es fácil ver que:

$$\frac{|\rho_A(\alpha_i)|}{|\tilde{D}|} = \frac{a_i + b_i}{a_i + b_i + c_i + d_i} \tag{11}$$

que es justamente lo que conocemos como soporte de un itemset cuando la base de datos es crisp. Además como hicimos notar anteriormente, $a_i + b_i + c_i + d_i$ permanece constante para cualquier nivel de restricción y es justamente el número de transacciones (en este caso difusas) de $\tilde{D}$ que notaremos como viene siendo habitual por n.

De forma parecida definiremos el soporte y la confianza para la regla de asociación difusa $A \to B$.

Definición 9. Sean $A, B \subseteq I$ dos itemsets disjuntos y $(\Lambda_A, \rho_A), (\Lambda_B, \rho_B)$ las RL-representaciones asociadas a los conjuntos difusos $\tilde{\Gamma}_A$ y $\tilde{\Gamma}_B$ en $\tilde{D}$. Entonces, el soporte de la regla difusa $A \to B$ en $\tilde{D}$ se define como

$$Sop(A \to B) = \sum_{\alpha_i \in \Lambda_A \cup \Lambda_B} (\alpha_i - \alpha_{i+1}) \left(\frac{|\rho_{A \wedge B}(\alpha_i)|}{|\tilde{D}|} \right) \tag{12}$$

Definición 10. Sean $A, B \subseteq I$ dos itemsets disjuntos y $(\Lambda_A, \rho_A), (\Lambda_B, \rho_B)$ las RL-representaciones asociadas a los conjuntos difusos $\tilde{\Gamma}_A$ y $\tilde{\Gamma}_B$ en $\tilde{D}$. Entonces, la confianza de la regla difusa $A \to B$ es

$$Conf(A \to B) = \sum_{\alpha_i \in \Lambda_A \cup \Lambda_B} (\alpha_i - \alpha_{i+1}) \left(\frac{|\rho_{A \wedge B}(\alpha_i)|}{|\rho_A(\alpha_i)|} \right) \tag{13}$$

En el caso de encontrarnos en la última parte de la fórmula (13) la indeterminación "$\frac{0}{0}$" tomaremos el valor 1 para preservar la definición 7 de regla de asociación difusa, ya que dicha indeterminación viene a decir que hay cero transacciones que cumplen el antecedente y el consecuente a la vez de entre cero transacciones que sólo cumplen el antecedente. De nuevo, es fácil ver que para cada nivel de restricción α_i la última parte de las fórmulas (12) y (13) coincide con el soporte y la confianza definidas para el caso crisp. Queremos hacer notar también que las anteriores definiciones extienden perfectamente las medidas de soporte y confianza del caso crisp al difuso. Es inmediato comprobarlo ya que dichas medidas coinciden con la generalización presentada en [1] en la que se utiliza una aproximación semántica basada en la evaluación de sentencias cuantificadas (ver también [2]) usando el método GD y el cuantificador relativo difuso la mayoría $Q_M(x) = x$ para evaluar las sentencias. En dichos trabajos se prueba que usando el método GD y el cuantificador Q_M el soporte y la confianza definidas en (12)-(13) son el soporte y la confianza ordinaria del itemset A y de la regla $A \to B$ respectivamente.

Tabla 2: RL-representaciones asociadas a $\tilde{\Gamma}_A$,$\tilde{\Gamma}_B$, $\neg\tilde{\Gamma}_A$, $\neg\tilde{\Gamma}_B$ y las posibles conjunciones entre ellos.

α_i	ρ_A	ρ_B	$\rho_{\neg A}$	$\rho_{\neg B}$	$\rho_{A\wedge B}$	$\rho_{A\wedge\neg B}$	$\rho_{\neg A\wedge B}$	$\rho_{\neg A\wedge\neg B}$
1	$\emptyset$	$\{\tilde{\tau}_2,\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_5,\tilde{\tau}_6\}$	$\tilde{D}$	$\{\tilde{\tau}_1\}$	$\emptyset$	$\emptyset$	$\{\tilde{\tau}_2,\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_5,\tilde{\tau}_6\}$	$\{\tilde{\tau}_1\}$
0.9	$\emptyset$	$\tilde{D}$	$\tilde{D}$	$\emptyset$	$\emptyset$	$\emptyset$	$\tilde{D}$	$\emptyset$
0.75	$\{\tilde{\tau}_3,\tilde{\tau}_6\}$	$\tilde{D}$	$\{\tilde{\tau}_1,\tilde{\tau}_2,\tilde{\tau}_4,\tilde{\tau}_5\}$	$\emptyset$	$\{\tilde{\tau}_3,\tilde{\tau}_6\}$	$\emptyset$	$\{\tilde{\tau}_1,\tilde{\tau}_2,\tilde{\tau}_4,\tilde{\tau}_5\}$	$\emptyset$
0.1	$\{\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_6\}$	$\tilde{D}$	$\{\tilde{\tau}_1,\tilde{\tau}_2,\tilde{\tau}_5\}$	$\emptyset$	$\{\tilde{\tau}_3,\tilde{\tau}_4,\tilde{\tau}_6\}$	$\emptyset$	$\{\tilde{\tau}_1,\tilde{\tau}_2,\tilde{\tau}_5\}$	$\emptyset$

Tabla 3: $4ft(\mathcal{M}_{\alpha_i},A,B,\tilde{D})$ con $\alpha_i\in\Lambda_A\cup\Lambda_B$

	a_i	b_i	c_i	d_i
$\mathcal{M}_1$	0	0	5	1
$\mathcal{M}_{0.9}$	0	0	6	0
$\mathcal{M}_{0.75}$	0	0	6	0
$\mathcal{M}_{0.1}$	3	0	3	0

Además, usando el modelo lógico propuesto que se basa en la representación mediante las tablas $4ft$ para cada nivel de restricción es inmediato probar que si A y B son itemsets de un conjunto de transacciones crisp, las definiciones para el soporte y la confianza (12)-(13) coinciden con las definiciones ordinarias de soporte y confianza.

Ejemplo 2. Sean I y $\tilde{D}_1$ los definidos en el ejemplo anterior. Consideremos los itemsets $A=\{i_1,i_3\}$ y $B=\{i_4\}$. Sus conjuntos difusos asociados $\tilde{\Gamma}_A = 0/\tilde{\tau}_1 + 0/\tilde{\tau}_2 + 0.75/\tilde{\tau}_3 + 0.1/\tilde{\tau}_4 + 0/\tilde{\tau}_5 + 0.75/\tilde{\tau}_6$ y $\tilde{\Gamma}_B = 0.9/\tilde{\tau}_1 + 1/\tilde{\tau}_2 + 1/\tilde{\tau}_3 + 1/\tilde{\tau}_4 + 1/\tilde{\tau}_5 + 1/\tilde{\tau}_6$ tendrán RL-representaciones (Λ_A,ρ_A) y (Λ_B,ρ_B). En la tabla 2 vemos sus RL-representaciones y las representaciones resultantes de aplicar la negación a ambos conjuntos difusos y la conjunción entre ellos.

Usando las anteriores RL-representaciones podemos ver cuáles son las tablas $4ft(\mathcal{M}_{\alpha_i},\tilde{\Gamma}_A,\tilde{\Gamma}_B,\tilde{D})$ para cada nivel de restricción en $\Lambda_A\cup\Lambda_B$ (tabla 3). De esta forma, podemos comprobar que el soporte de la regla $A\rightarrow B$ es 0.267 usando la ecuación (12) y su confianza es 1. Esto último es fácilmente comprobable puesto que $\tilde{\tau}_i(A)\leq\tilde{\tau}_i(B)$ para $i=1,2,3,4$ siguiendo la definición 7. Si usamos la fórmula para la confianza dada en (13) también llegamos al mismo resultado pero teniendo en cuenta que cada vez que aparezca la indeterminación "$\frac{0}{0}$" tendremos que considerarla como 1 como ya se comentó después de la definición 10.

5. FUTUROS TRABAJOS Y CONCLUSIONES

Hemos presentado un modelo sencillo con buenas propiedades para la representación y evaluación de reglas de asociación difusas que generaliza de forma natural al caso crisp. Además, planeamos usar el algoritmo presentado en [9] para la extracción de reglas difusas ya que puede adaptarse fácilmente gracias al modelo propuesto.

Referencias

[1] M. Delgado, N. Marín, D. Sánchez, and M.A. Vila. Fuzzy association rules: General model and applications. *IEEE Transactions on Fuzzy Systems*, 11(2):214–225, 2003.

[2] M. Delgado, M.D. Ruiz, and D. Sánchez. Reglas de asociación difusas: Nuevos retos. In *ESTYLF'08*, 2008, submitted.

[3] M. Delgado, M.D. Ruiz, and D. Sánchez. Studying interest measures for association rules through a logical model. *IJUFKS*, Submitted, 2007.

[4] P. Hájek, I. Havel, and M. Chytil. The guha method of automatic hypotheses determination. *Computing*, 1:293–308, 1966.

[5] P. Hájek and T. Havránek. *Mechanising Hypothesis Formation-Mathematical Foundations for a General Theory*. Springer-Verlag: Berlin, Heidelberg and New York, 1978.

[6] E. Hüllermeier. Fuzzy sets in machine learning and data mining. *Applied Soft Computing, to appear*, 2008.

[7] T-P Hong and Y-C Lee. *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, chapter An Overview of Mining Fuzzy Association Rules, pages 397–410. Springer, 2008.

[8] J. Rauch. Logic of association rules. *Applied Intelligence*, 22:9–28, 2005.

[9] J. Rauch and M. Šimunek. An alternative approach to mining association rules. *Studies in Computational Intelligence (SCI)*, 6:211–231, 2005.

[10] D. Sánchez, M. Delgado, and M.A. Vila. A restriction level approach to the representation of imprecise properties. In *IPMU'08*, Málaga, Spain, 2008, To appear.

[11] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

Obtaining fuzzy sequential patterns for Ambient Intelligence Problem Solving

Miguel Delgado[1] **María Ros**[2] **M. Amparo Vila**[3]

[1] Department of Computer Science and Artificial Intelligence. University of Granada, mdelgado@ugr.es
[2] Department of Computer Science and Artificial Intelligence. University of Granada, marosiz@decsai.ugr.es
[3] Department of Computer Science and Artificial Intelligence. University of Granada, vila@decsai.ugr.es

Abstract

Many traditional environment applications base their operations in using sensors, for this reason sensors are becoming more and more common in daily life. One of that applications is the Tagged World which uses the information from sensors to identify the user behavior and to provide any service. In this paper, we present a method to use the sensor information in order to extract the behavior pattern on time performed by the user. The method uses frequent itemsets to represent the common actions that are realized by user. After, it obtains the sequences patterns which defines the specific behavior from the extracted frequent itemsets.
However, the user behavior has random imprecision by definition. Thus, we have designed a method to manage this imprecision from the user behavior, establishing a temporal constraint: a Fuzzy Temporal Window.

Keywords: Tagged World, Frequent Itemsets, Sequence Patterns, α-cut.

1 INTRODUCTION

The ubiquitous systems is one of the today's technology reefs. Sensors are able to collect the human activity. That permits to develop a system which is able to detect the user's actions from sensor information. Therefore, the goal of such systems is to find a mechanism that would identify the different actions with the specific activity carried out by the user.
A Tagged World [8] is defined as a smart area that serves to recognize user's behavior using information about their daily activity. This information is collected by sensors placed (embedded) in the environment.
The sensor information is continuous on time, i.e., it could arrive at the system at any time. This information can be processed with different techniques of Data Mining, or Stream Mining specifically, in order to extract some knowledge pattern.
When the information is continuos, we need to use some techniques to realize a representative sampling of the data [3]. There exists several techniques to reduce the input data [7]: sampling, sliding window, histogram,sketches, etc.
Sensor information should be processed to be used in real applications. In [5] it is proposed a system to identify correct behavior using Data Mining Techniques. The system is divided into two fundamental parts: inductive learning mechanism, whose final output will be a behavior database using Frequent Itemsets; and a system for the recognition of sequences performed by the users by Regular Grammar. The second stage will be built on the behavior database obtained in the previous process.
That system uses the Frequent Itemset concept to process the sensor information, without keeping in mind the time when the action happens. That is an important problem because this time determines some system features. There are many proposal about time constraints for sequence mining. In [6] Fiot et al. present a soften temporal constraints used for generalized sequential pattern mining. The temporal constraints are established between the itemset that are in the sequence patterns, pointing out the minimum and maximum distance among them. The result is fuzzy sequences patterns.
There exists many papers which propose some algorithm to obtain the sequences patterns. AprioriAll and AprioriSome, which were proposed by Rakesh Agrawal et al. [Rakesh Agrawal et al.1995], are the first method we can find in the bibliography. Henceforth, many variants have been appeared, where we could emphasize the algorithm PrefixSpan [Pei et al. 2001].which explores prefix projection in sequential pattern min-

ing. PrefixSpan mines the complete set of patterns but greatly reduces the efforts of candidate subsequence generation. The method to obtain the sequence of actions that defines an user behavior. However, these method solves crips problem. But, we have mentioned that the system have to manage random imprecision. There are other methods to work with fuzzy sequences patterns, but the most common approach is to extract the fuzzy sequences patterns to quantitative attributes. For example, [4] propose Fuzzy Grids Based Sequential Pattern Mining Algorithm (FGBSPMA) to generate all fuzzy sequential patterns from relational database, where each quantitative attribute is viewed as a linguistic variable, and can be divided into many candidate I-dim fuzzy grids. Although that, there are not many applications of fuzzy sequential patterns to discrete attributes.
There are many kind of sensors, between them we could emphasize the RFID. RFID technology is based on radio-frequency communication and permits the univocal identification of the elements in the system. RFID systems [12] consists of tags (chips that contain identifying and often other data) and reading devices that convey information from the tags to computers. RFID Devices are beginning to replace magnetic-stripe security cards for unlocking doors and granting access to secured areas, especially at facilities with special security needs, such as military installations. Such systems are already in limited use and are being tested widely for applications involving the tracking of inventory from manufacturers to stores. However, nowadays new uses are appearing for RFID systems such that identifying object in a Tagged World [9]. Philipose and et al. [9] propose a system to inferring Activities of Daily Living in Tagged World. They present a new paradigm for ADL inferencing leverages radio-frequency-identification technology, data mining and a probabilistic inference engine to recognize ADLs based on the objects people use. As sensor, they use RFID technology with other sensor streams to fill in the gaps. The system represents activities as linear sequences of activity stages, and annotate each stage with the objects involved and the probability of their involvement.

This paper is organized into five main sections. In section 2 we present the formal problem to solve with the designed method. In the next section, the crisp model is presented explaining the method to obtain the knowledge for a specific behavior. In section 4 we present the Fuzzy Temporal Window concept and the method that permits to work with the uncertainty of the problem. As well, we present an illustrative example of operation model. Finally, the conclusions and future works are reported in section 5.

2 FORMAL PROBLEM

This section defines the formal problem we try to solve: to obtain the pattern sequences that identify the user behavior in a specific domain and context. Thus, we have to define some basic concepts which will be the systems' aims.
When we observe an user, the most essential concept that he could do is the action.

Definition 2.1 (Action) *An action is defined as the occurrence of a fact over a specific object.*

However, it results very interested to know when the action happened on time.

Definition 2.2 (Action on time) *An action is the occurrence of a fact over a specific object in a given time and it denotes as a pattern $i_j = (h_j, l_j)$, where h_j is a fact and l_j is a temporal label defines the time of occurrence of the action.*

These definitions give us the basic element to work, but our aim is to find the different actions make a behavior up.

Definition 2.3 (Behavior) *Let $A = \{a_1, ..., a_n\}$ be the set of possible user's actions in some situation or domain. A user behavior is a finite set of actions:*

$$\beta = \{\alpha_1, \alpha_2, , ...\alpha_{p(\beta)}\}$$

with $\alpha_j \in A \forall j$, and where α_j is performed before α_k iif $j \leq k$.

In this definition, we have not taken in mind the time, although, in general, a behavior happen in a specific moment on time.

Definition 2.4 (Behavior on time) *Let $A = \{a_1, ..., a_n\}$ be the set of possible user's actions in some situation or domain. Let τ be the temporal line, then A user behavior on time is a finite set of pairs such as:*

$$b_I = \{(\alpha_1, t_1), ...(\alpha_{n(b)}, t_{n(b)})\}$$

where $\alpha_j \in A \forall j$, where α_j is performed before α_k iif $j \leq k$, $t_j \in [0, \tau]$ $\forall t_i < t_j$ si $i \geq j$, con $I = [t_1, t_{n(b)}]$

The system takes, as its basis, a database about the user action. This database has been obtained from the observation of the user, i.e., from the actions the user has done. In the rest of the document, this database will be named by Observation Data Base (ODB).
The selected representation for ODB is a transactional database T, where each fils is a observation over the user, normally his activity all day long; and each column is the possible actions in the A set, $A = \{a_1, ..., a_n\}$.

3 CRISP MODEL

In this section, we present a solution to the problem has been explained in section 2 but only to the crisp problem. This version of the problem consider each action as an occurrence, without any time. The process followed to obtain the pattern sequences is explained in [5]. That process has two stages: to extract the common actions and to obtain the correct sequences.

3.1 OBTAINING COMMON ACTIONS

Before to explain the method, it is convenient to understand what the common actions are and what they represent on the method.
When we talk about common actions, we refer the actions what person performs generally to do a specific behavior.

Example 3.1 (Common actions) *Let us considerer the* `to go out home` *behavior which contains always two actions,* `to take the keys` *and* `to close the door`. *However, if the user* `to take a bag` *or* `to take a scarf`, *it will not be a common action in the behavior.*

As we have mentioned above in this paper, the system observes the user and collects his activity in the ODB, which could be able to represent as a transactional database T. Thus, the common actions can be identified with the concept of frequent itemset extracted from the transactional database. Thus, the common actions will be defined as the sequence of events that occur more often in the observed knowledge. Each frequent itemset corresponds to a particular common behavior [5]. The extraction of the common actions could make by any algorithm based in Apriori Algorithm [1].
After applying the Apriori Algorithm, itemsets have been obtained. However, a itemset is a set and lacks order among its elements. That results a problem, because a behavior has an order among its actions.

Example 3.2 (Valid order) *Let us considerer a* `to leave home` *behavior. The itemset defines that behavior will be* `door key mobile_phone`. *To this actions, it could think in two possible ordinations:* `mobile_phone keys door` *or* `keys mobile_phone door`. *Nevertheless, we could not considered any order where the door will be before the mobile_phone or the keys.*

3.2 OBTAINING CORRECT SEQUENCES

When we have obtained the common actions, we have to obtain the possible order of these actions to build the sequence patterns. To obtain them, we will use the permutation of a set to obtain all possible sequences. But, not all permutation will be accepted as a pattern sequence. Only we accept the permutation whose order is in the ODB.

Definition 3.1 (Correct sequences) *Let T be the transactional database which represents the ODB and let I be an itemset obtained from it for behavior* b. *Then a correct sequence p is a permutation of I where the order among its elements are defined in the ODB.*

Example 3.3 (Correct sequences)
Let us consider an example of `leave home` *behavior. The ODB is showed in the Table 1. This is represented in the transactional database T. If we apply the Apriori Algorithm over T, with support = 0.9, we obtain the itemset {Keys, ODoor, MobilePhone}. From this itemset, we could obtain the permutations showed in table 2. However, the order of all sequences are not valid in the ODB. For example, the sequences s_3 and s_4 never could happen. After removing all invalid sequences, we accept as valid sequences s_2 and s_6.*

Table 1: Leave Home ODB

TRANS	ELEMENTS
t1	Clothes, Shoes, MobilePhone, Bag, Keys, ODoor, Elevator
t2	Shoes, Bag, Keys, MobilePhone,ODoor, Keys2, ODoor2, Elevator
t3	Tap, Towel, BathroomDoor, BedroomDoor, Clothes,Shoes, Bag, Keys, MobilePhone, ODoor, Keys2,ODoor2, Elevator
t4	Shoes, Keys, MobilePhone, Bag, ODoor, Elevator, Keys2, ODoor2
t5	Shoes, Bag, MobilePhone, Keys, ODoor, Keys2, ODoor2, Elevator

Table 2: Possible permutations

s_1	{Keys, ODoor, MobilePhone}
s_2	{Keys, MobilePhone, ODoor}
s_3	{ODoor, Keys, MobilePhone}
s_4	{ODoor, MobilePhone, Keys}
s_5	{MobilePhone, ODoor, Keys}
s_6	{MobilePhone, Keys, ODoor}

So, we use the knowledge of the ODB to determine if a sequence is valid or not. When a sequence is valid,

we name it as *correct pattern*. The algorithm 2.1 is designed to detect if we have found a *correct pattern* or not.

Algorithm 3.1 (Valid sequences)

For each $i \in$ behavior sequence
  For each $t \in$ ODB
    Check if all elements in the sequence i respect the t elements order
      Get the position of two t elements
      Check the positions
        If this positions are incorrect in the sequence i
          Get other two t elements
  If the sequence does not find in neither of ODB, delete the sequence.

As we explained in the subsection 1, there exists many algorithm to obtain the sequences patterns. However, we are not looking for this type of sequence patterns but sequence patterns which represent all valid ordinations for elementos of the frequent itemset.
In addition, the traditional algorithms do not solve the main problem of the frequent itemset: different order between their elements. The sequences has been obtained by these type of method are statics, i.e., the order is to set up when their build the sequences. These sequences will only be valid if their order is frequent in the data base.

4 INTRODUCTION OF A FUZZY WINDOWS

In the previous section, we have presented a model crisp, but the human behavior is not crisp. The human activity has random imprecise by definition, because the human has free thinking. Then the system has to work with it and to adapt itself to the user habits.

4.1 MOTIVATIONS

In the section 3 we have presented the process to obtain the correct patterns sequences of actions from user behavior. This method extracts from a ODB particular the common action. The ODB contains tuples which are only composed by the action to a specific behavior. However, a general ODB contains actions that have been realized by the user during whole day. For this reason, we do not know where the behavior starts and ends.
In [5], the method propose that an expert would indicate when a behavior performs. However, this proposal is not applicable to a real system, where each person has particular activities and habits. Therefore, it has to try to find a mechanism to identify behavior patterns from the sequence data collected by the sensors.
Traditional Stream Mining techniques focus their activity on the study of a range of information [3]. It defines a Sliding Windows that establishes the range and is moved according to the time increases. Generally, there exists some behavior that are realized by the user at the same time every day.

Example 4.1 *We want to control the behavior* `Luis go out home`*. But, we do not know when the behavior starts or ends, we only know that Luis go out home at 8:30 o'clock.*

Then, we could use this knowledge to situate an interval on the temporal line. The result will be a subset from the ODB particular to the studied behavior.
However nobody does the action at the same moment each day, so he does the action roughly at the same moment. This raises a new problem: how to fix the window's ranges to detect a particular activity. Ranges should consider the random imprecise of the situations.

Example 4.2 *As we know* Luis go out home at 8:30 o'clock*, then we could control the actions happen about 8:30.*

This interval, which has been defined on the temporal line above a specific time, will be named as *Temporal Window*. It permits to get a subset from each ODB tuple (See figure 1). In this subset, the actions have not the same degree of importance, since the actions are about the specific time are more important than the actions far from the interval center. So, we could assign a degree of importance at each action, in function of its situation in the interval. Then the interval will be defined as a fuzzy set and name as *Fuzzy Temporal Window* (See figure 2).

4.2 MODEL FORMULATION

In this section we present a formal representation to the *Temporal Window* and *Fuzzy Temporal Window*.

Definition 4.1 (Temporal Window) *Let τ be a interval from the temporal line τ, ODB the Observation Data Base and $t \in ODB$ a tuple from ODB. Let i_j be an action $i_j = (h_j, l_j)$, then a Temporal Window W to a specific behavior, which ocurrs in the interval τ, is defined as a subset of t where*

$$\forall i_j \in W(t) \text{ then } i_j \in t \text{ and } l_j \in \tau$$

The next stage is to include some random imprecision in the specific Temporal Window.

Definition 4.2 (Fuzzy Temporal Window) *Let τ be a interval from the temporal line τ, ODB the Obser-*

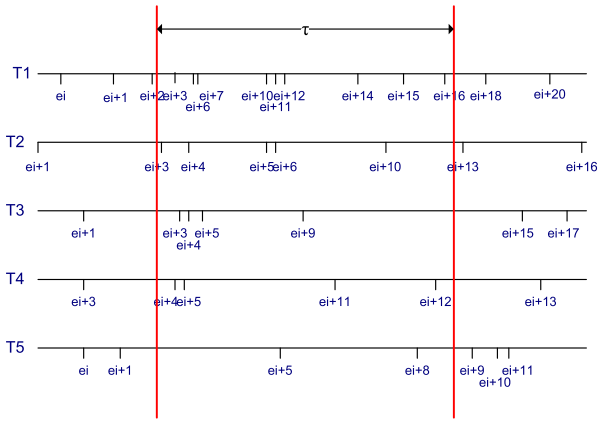


Figure 1: Temporal Window over the ODB.

vation Data Base and $t \in ODB$ *a tuple from ODB. Let* i_j *be an action* $i_j = (h_j, l_j)$ *and W a Temporal Window to a specific behavior. Let fs be a fuzzy set over* τ*, then it defines the FW* Fuzzy Temporal Window *as a Temporal Window where*

$$\forall i_j \in W(t)\ \mu_{FW}(i_j) = \mu_{fs}(l_j)$$

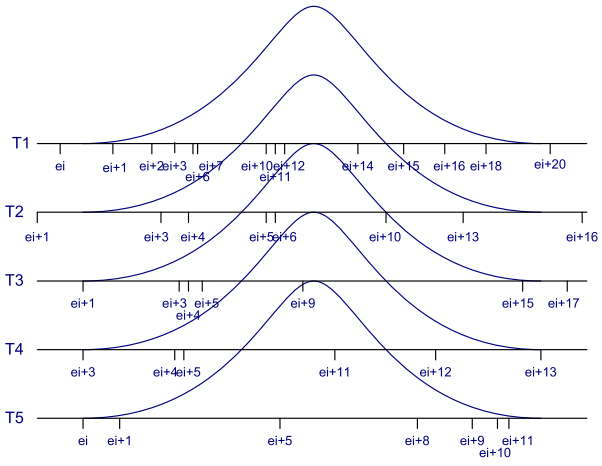


Figure 2: Fuzzy Temporal Window over the ODB.

Applying a Fuzzy Temporal Window on the ODB, we obtain a image of it where we assign to each item its membership degree. We have represented the ODB as a Transactional Database, thus we should obtain the image of Transactional Database too.

Definition 4.3 (Fuzzy ODB applying W) *Let ODB be a Observation Data Base and W a Fuzzy Temporal Window, it defines* $O\tilde{D}B$ *as a Fuzzy Observation Data Base constructs as W(ODB) such as*

$$\forall t \in ODB, W(t) \in O\tilde{D}B$$

Definition 4.4 (Fuzzy Transactional Data Base) *Let T be a Transactional Data Base and W a Fuzzy Temporal Window, it defines* $\tilde{T}$ *as a Fuzzy Transactional Data Base constructs as W(t)* $\forall t \in T$*.*

We represent the Fuzzy Transactional DataBase as a table where for each fils and column we have the membership degree corresponding to the Fuzzy Temporal Window.
Once we have the data, we should get the common actions and valid sequences for each behavior. However, we can not apply the method explained in the section 3 directly, because we have to manage the membership degree corresponding to the Fuzzy Temporal Window. The following subsection explains models' modifications to manage the random imprecision.
Therefore, the aim is developed a system that obtains the stimulation from the user who realizes some actions sequentially. So the system has to find the action sequence which corresponds with a specific behavior that is defined by a Fuzzy Temporal Window.

4.3 OBTAINING PATTERNS SEQUENCES BY ALPHA-CUTS

Here, we present the techniques used to extract the sequences patterns when we have a specific Fuzzy Temporal Window W and the Fuzzy ODB $O\tilde{D}B$, which is represented by the Fuzzy Transactional Database $\tilde{T}$. The ideal situation will be the sequences patterns were fuzzy. However, our method is designed to crisp problem. Thus, we have to work in first time with a crisp problem and, after, transform it to a fuzzy problem again. To do that, we are using the α-cut concept and the Representation Theorem.

4.3.1 Description

The problem consist of extracting the sequences patterns to specific behavior when we have a Fuzzy ODB $O\tilde{D}B$, represented as $\tilde{T}$, and a Fuzzy Temporal Window W defined from the user knowledge.
In the first stage, we have to transform a fuzzy problem to a crisp problem with the aim of applying the method explained in the section 2. To do that, we could use the α-cut set [10].

Definition 4.5 (α-cut set of A)

$$A^{\alpha} = x \in X | \mu_A(x) \geq \alpha \qquad (1)$$

By the α-cut, we would have a new image of $\tilde{T}$, T^{α}, crisp where each value is in $\{0, 1\}$. Then, we can apply the classical method to T^{α}, obtaining frequent itemsets and sequences patterns to specific α value (I^{α} and

P^{α}, respectively). After extracting all sequences patterns, we could create a fuzzy set **P** by the sequences patterns which have obtained to each α-value applying the Representation Theorem.

Theorem 4.1 (Representation Theorem) *Let A be a fuzzy set then we can verify*

$$A = \bigcup_{\alpha \in [0,1]} A^{\alpha} \quad (2)$$

This theorem has an important restriction, the *consistent restriction*: If $\alpha_1 > \alpha_2$ then $A^{\alpha_1} \subseteq A^{\alpha_2}$.
As we use the frequent itemset, we need to represent the frequent itemsets extracted as an unique fuzzy set. So we have to ensure the *consistent restriction* among each α frequent itemset I^{α_i}

Proposition 4.1 *Let $\tilde{T}$ be a fuzzy transactional data base, T^{α_1} and T^{α_2} the result to use the α-cut on $\tilde{T}$ with $\alpha = \alpha_1, \alpha_2$, $\alpha_1, \alpha_2 \in [0,1]$, $\alpha_1 \geq \alpha_2$. And let I^{α_1}, I^{α_2} be the set of frequent itemsets extracted from T^{α_1}, T^{α_2} respectively. Then, it is declared what $I^{\alpha_1} \subseteq I^{\alpha_2}$, ie, if I is frequent to level α_1 then it is frequent to level α_2.*

Demonstration
We suppose we have $\tilde{T}$ a fuzzy transactional data base, where each $t \in \tilde{T}$ represents a fuzzy set of T. If we apply the α-cut on $\tilde{T}$, we obtain T^{α}. Then, if we use the values $\alpha_1, \alpha_2 \in [0,1]$ with $\alpha_1 \geq \alpha_2$ we get T^{α_1}, T^{α_2} respectively. So, $T^{\alpha_1} \subseteq T^{\alpha_2}$, by consistent restriction. *Let I^{α_1}, I^{α_2} be the frequent itemsets from T^{α_1}, T^{α_2} respectively. Then $\forall i \in I^{\alpha_1}, \exists t \in T^{\alpha_1} / i \in t$; with $T^{\alpha_1} \subseteq T^{\alpha_2} \Longrightarrow t \in T^{\alpha_2}$*
Let I^{α_1} be the frequent itemset to $T^{\alpha_1} \Longrightarrow \forall i \in I^{\alpha_1}$, $supp(i) \geq minsup$ to $T^{\alpha_1} \Longrightarrow supp(i) \geq minsup$ to T^{α_2}, ie, $i \in I^{\alpha_2} \Longrightarrow I^{\alpha_1} \subseteq I^{\alpha_2}$

In our method, we use the frequent itemset obtained to create the sequences patterns. However, now we have the frequent itemset to specific α value, so we will get sequences patterns to that specific α value. To represent the patterns sequences as a fuzzy set we use the Representation Theorem, for this fact we need to ensure the *consistent restriction.*

Proposition 4.2 *Let P^{α_i}, P^{α_j} be sets of sequences patterns where $\alpha_i \geq \alpha_j$ then $P^{\alpha_i} \subseteq P^{\alpha_j}$.*
Demonstration *Let I^{α_i}, I^{α_j} be the set of frequent itemset obtained from T^{α_i} and T^{α_j}, respectively, with $\alpha_i \geq \alpha_j$. Let P^{α_i}, P^{α_j} be the pattern valid set of I^{α_i} and I^{α_j}, respectively.*
We define the function $perm(I)$ such that:

$$perm(I) = \{p / p \text{ is a } I \text{ permutation and } p \in T\} \quad (3)$$

where T is a transactional data base and I is a frequent itemset from T.
Then,
$\forall I \in I^{\alpha_i}$ $perm(I) \in P^{\alpha_i}$. As $I^{\alpha_i} \subseteq I^{\alpha_j} \Longrightarrow I \in I^{\alpha_j}$.
So $perm(I) \in P^{I^{\alpha_i}} \Longrightarrow P^{I^{\alpha_i}} \subseteq P^{I^{\alpha_j}}$

4.3.2 An illustrative example

In this example we want to show the operation method. We start from ODB which is collected from the touched object by the user in the daily activity. We do not show the Data Base due to the lack of space. Let W a Fuzzy Temporal Window defines by the following fuzzy set over temporal line τ:

$$W = \left\{ \frac{0}{8:20}, \frac{0{,}2}{8:21}, \frac{0{,}4}{8:22}, \frac{0{,}6}{8:23}, \frac{0{,}8}{8:24}, \frac{1}{8:25}, \frac{1}{8:26}, \frac{1}{8:27}, \frac{1}{8:28}, \frac{1}{8:29}, \frac{1}{8:30}, \frac{1}{8:31}, \frac{1}{8:32}, \frac{1}{8:33}, \frac{1}{8:34}, \frac{1}{8:35}, \frac{0{,}8}{8:36}, \frac{0{,}6}{8:37}, \frac{0{,}4}{8:38}, \frac{0{,}2}{8:39}, \frac{0}{8:40} \right\} \quad (4)$$

Then, we apply the W over the T and obtain a Fuzzy T $\tilde{T}$ where each time of happenes are replaced with the membership degree value of the W. We apply the α-cut for α values $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\alpha_3 = 0.8$, $\alpha_4 = 1.0$. We obtain for each α values the T^{α}, the result of apply the α-cut in $\tilde{T}$.

After these operations, we have transformed the fuzzy problem to a crisp problem. We can extract frequent itemsets and sequences patterns with the model explained in section 3.

We have executed the Apriori Algorithm with two support values: *minsup = 0.8 and minsup = 0.9.* The results are showed in tables 3 and table 4.

Table 3: Frequent itemset to minsup = 0.8

I^{α_1}	{Shoes, Bag, Keys, ODoor, MobilePhone} {Keys, Keys2, ODoor, ODoor2, MobilePhone}
I^{α_2}	{Shoes, Bag, Keys, ODoor, MobilePhone} {Keys, Keys2, ODoor, ODoor2, MobilePhone}
I^{α_3}	{Shoes, Bag, Keys, ODoor, MobilePhone}
I^{α_4}	{Shoes, Bag, Keys, ODoor, MobilePhone}

Table 4: Frequent itemset to minsup = 0.9

I^{α_1}	{Keys, ODoor, MobilePhone}
I^{α_2}	{Keys, ODoor, MobilePhone}
I^{α_3}	{Keys, ODoor}
I^{α_4}	{ODoor}

Since we have the common actions to specific behavior, the next stage consists of obtaining the valid sequences patterns and the final representation as a fuzzy set applying the Identity Principle. The valid patterns are

showed in table 5 and table 6. And the final pattern representation in equation 5 to minsup = 0.8 and equation 6 to minsup = 0.9.

Table 5: Valid patterns to minsup = 0.8

P^{α_1}	{Shoes, Bag, Keys, MobilePhone, ODoor}	p_1
	{Shoes, Keys, MobilePhone, Bag, ODoor}	p_2
	{Shoes, MobilePhone, Bag, Keys, ODoor}	p_3
	{Keys, MobilePhone, ODoor, Keys2, ODoor2}	p_4
	{MobilePhone, Keys, ODoor, Keys2, ODoor2}	p_5
P^{α_2}	IDEM	
P^{α_3}	{Shoes, Bag, Keys, MobilePhone, ODoor}	p_1
	{Shoes, Keys, MobilePhone, Bag, ODoor}	p_2
	{Shoes, MobilePhone, Bag, Keys, ODoor}	p_3
P^{α_4}	IDEM	

$$\tilde{P} = \left\{ \frac{p_1}{\alpha_4}, \frac{p_2}{\alpha_4}, \frac{p_3}{\alpha_4}, \frac{p_4}{\alpha_2}, \frac{p_5}{\alpha_2} \right\} = \left\{ \frac{p_1}{1}, \frac{p_2}{0.4}, \frac{p_3}{1}, \frac{p_4}{1}, \frac{p_5}{0.6} \right\} \quad (5)$$

Table 6: Valid patterns to minsup = 0.9

P^{α_1}	{Keys, MobilePhone, ODoor}	p_1
	{MobilePhone, Keys, ODoor}	p_2
P^{α_2}	IDEM	
P^{α_3}	{Keys, ODoor}	p_3
P^{α_4}	{ODoor}	p_4

$$\tilde{P} = \left\{ \frac{p_1}{\alpha_2}, \frac{p_2}{\alpha_2}, \frac{p_3}{\alpha_3}, \frac{p_4}{\alpha_4} \right\} = \left\{ \frac{p_1}{0.6}, \frac{p_2}{0.6}, \frac{p_3}{0.8}, \frac{p_4}{1} \right\} \quad (6)$$

5 CONCLUSIONS

In this paper we have proposed a method to detect the user behavior. This problem is random imprecise and uncertain by definition.

- random imprecise, because we do not know the interval of the time in which the action happen.
- uncertain, because it could exists a mistake in the sensor reading.

We have studied the random imprecision of the problema. This imprecision appears due of a behavior is not static on a context, domain and for an user. Thus, we have designed a method which uses fuzzy logic to limit the events in a specific behavior using the Fuzzy Temporal Window. With the Window, we can assign a membership degree for each event to the studied behavior.

The system result is the fuzzy patterns sequences which define the behavior. To get it, we use the α-cuts and the Representation Theorem to obtain a fuzzy result applying the crisp method. That let us, we could demonstrate that the method obtain a fuzzy pattern sequence set.

In future work, we would like to extend the Fuzzy Temporal Window concept, to adjust the interval using the knowledge from user we have. This adjustment will be made with quantified sentences. With this adjustment we will solve the possibility of a person changes his habits.

As well, we have to present a system which uses the knowledge organized in that stage to take decision and to send the correct alarms.

When we work with sensors we have to control their operations. There are many reason to provoke the mistake in the sensor reading: the sensors or the reader could fail, the environment affects to the sensors, the sensor damage, etc. The uncertainty from this situation will be studied and controlled in future works too.

References

[1] R. Agrawal; R. Srikan. Fast Algorithms for Mining Association Rules. *IBM Research Report RJ9839*, Junio 1994.

[2] R. Agrawal; R. Srikant. Mining Sequetial Pattern. In *11th International Conference on Data Engineering* pages 3-14, Taipei, Taiwan, 1995. IEEE Computer Society Press.

[3] Charu C. Aggarwal. Data Streams: Models and Algorithms. Hardcover, Advances in Database Systems Ser. 2003

[4] Ruey-Shun Chen , Gwo-Hshiung Tzengb, C. C. Chen and Yi-Chung Hu. Discovery of fuzzy sequential patterns for fuzzy partitions in quantitative attributes. *ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*, 2001, pp. 144-150.

[5] Miguel Delgado, María Ros, Amparo Vila. Correct Behavior Identification System in a Tagged World. *1th Jornadas Cientficas sobre RFID*, 2007.

[6] Cline Fiot, Anne Laurent, Maguelonne Teisseire. Extended Time Constraints for Sequence Mining. *14th International Symposium on Temporal Representation and Reasoning TIME'07*,2007

[7] Mohamed Medhat Gaber, Arkady Zaslavsky, Shonali Krishnaswamy.Mining Data Streams: A Review.SIGMOD Record 2005. Vol(34), 2

[8] Kyohei Koyama, Kouichi Nakagawa, and Hiromitsu Shimakawa, Embedded Action Detector to Enhance Freedom from Care, *Proc. of 11th International Conference on Human-Computer Interaction*, 8 pages, Las Vegas, Jul., 2005

[9] Matthai Philipose, Kenneth P Fishkin, Mike Perkowitz, Donald J. Patterson, Dieter Fox, Henry Kautz, Dirk Hhnel (2004). Inferring Activities from Interactions with Objects. *Context-Aware Computing. Pervasive computing* 3(50-57)

[10] Hung T. Nguyen, Elbert A. Walker, *A first Course in Fuzzy Logic*, Chapman & Hall/CRC Taoylr & Frances Group. 2006

[11] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal and M-C. Hsu. PrefixSpan: Mining Sequential Patterns Efficiently by PrefixProjected Pattern Growth. In. Proc. 2001 Int. Conf. Data Engineering (ICDE'01), pages 215-224, Heidelberg, Germany, April 2001. http://citeseer.ist.psu.edu/pei01prefixspan.html

[12] Want R. RFID: A key to automating everything. *Scientific American Magazine*, page 10, January 2004.

Or gate Bayesian networks for text classification: A discriminative alternative approach to multinomial naive Bayes

Luis M. de Campos Juan M. Fernández-Luna Juan F. Huete Alfonso E. Romero

Departamento de Ciencias de la Computación e Inteligencia Artificial
E.T.S.I. Informática y de Telecomunicación, Universidad de Granada,
18071 – Granada, Spain
{lci,jmfluna,jhg,aeromero}@decsai.ugr.es

Abstract

We propose a simple Bayesian network-based text classifier, which may be considered as a discriminative counterpart of the generative multinomial naive Bayes classifier. The method relies on the use of a fixed network topology with the arcs going form term nodes to class nodes, and also on a network parametrization based on noisy or gates. Comparative experiments of the proposed method with naive Bayes and Rocchio algorithms are carried out using three standard document collections.

Keywords: Bayesian network, noisy or gate, multinomial naive Bayes, text classification

1 Introduction: Probabilistic Methods for Text Classification

The classical approach to probabilistic text classification may be stated as follows: We have a class variable C taking values in the set $\{c_1, c_2, \ldots, c_n\}$ and, given a document d_j to be classified (described by a set of attribute variables, which usually are the terms appearing in the document), the posterior probability of each class, $p(c_i|d_j)$, is computed in some way, and the document is assigned to the class having the greatest posterior probability. Learning methods for probabilistic classifiers are often characterized as being generative or discriminative. Generative methods estimate the joint probabilities of all the variables, $p(c_i, d_j)$, and $p(c_i|d_j)$ is computed according to the Bayes formula:

$$p(c_i|d_j) = \frac{p(c_i, d_j)}{p(d_j)} = \frac{p(c_i)p(d_j|c_i)}{p(d_j)} \propto p(c_i)p(d_j|c_i)\,. \tag{1}$$

The problem in this case is how to estimate the probabilities $p(c_i)$ and $p(d_j|c_i)$. In contrast, discriminative probabilistic classifiers model the posterior probabilities $p(c_i|d_j)$ directly.

The naive Bayes classifier is the simplest generative probabilistic classification model that, despite its strong and often unrealistic assumptions, performs frequently surprisingly well. It assumes that all the attribute variables are conditionally independent on each other given the class variable. In fact, the naive Bayes classifier can be considered as a Bayesian network-based classifier [1], where the network structure is fixed and contains only arcs from the class variable to the attribute variables, as shown in Figure 1.

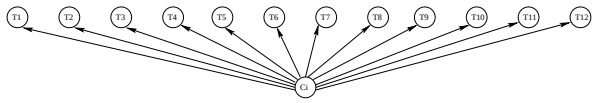


Figure 1: Network structure of the naive Bayes classifier.

In this paper we are going to propose another simple Bayesian network-based classifier, which may be considered as a discriminative counterpart of naive Bayes, in the following senses: (1) it is based on a type of Bayesian network similar to that of naive Bayes, but with the arcs in the network going in the opposite direction; (2) it requires the same set of simple sufficient statistics than naive Bayes, so that the complexity of the training step in both methods is the same, namely linear with the number of attribute variables; the complexity of the classification step is also identical.

The rest of the paper is organized in the following way: in Sections 2 and 3 we describe the two probabilistic text classifiers we are considering, naive Bayes and the proposed new model, respectively. Section 4 is focused on the experimental results. Finally, Section 5 contains the concluding remarks and some proposals for future work.

2 The Multinomial Naive Bayes Classifier

In the context of text classification, there exist two different models called naive Bayes, the multivariate Bernoulli naive Bayes model [2, 4, 8] and the multinomial naive Bayes model [5, 6]. In this paper we shall only consider the multinomial model. In this model a document is an ordered sequence of words or terms drawn from the same vocabulary, and the naive Bayes assumption here means that the occurrences of the terms in a document are conditionally independent given the class, and the *positions* of these terms in the document are also independent given the class[1]. Thus, each document d_j is drawn from a multinomial distribution of words with as many independent trials as the length of d_j. Then,

$$p(d_j|c_i) = p(|d_j|)\frac{|d_j|!}{\prod_{t_k\in d_j} n_{jk}!}\prod_{t_k\in d_j} p(t_k|c_i)^{n_{jk}}\,, \quad (2)$$

where t_k are the distinct words in d_j, n_{jk} is the number of times the word t_k appears in the document d_j and $|d_j| = \sum_{t_k\in d_j} n_{jk}$ is the number of words in d_j. As $p(|d_j|)\frac{|d_j|!}{\prod_{t_k\in d_j} n_{jk}!}$ does not depend on the class, we can omit it from the computations, so that we only need to calculate

$$p(d_j|c_i) \propto \prod_{t_k\in d_j} p(t_k|c_i)^{n_{jk}}\,. \quad (3)$$

The estimation of the term probabilities given the class, $\hat{p}(t_k|c_i)$, is usually carried out by means of the Laplace estimation:

$$\hat{p}(t_k|c_i) = \frac{N_{ik}+1}{N_{i\bullet}+M}\,, \quad (4)$$

where N_{ik} is the number of times the term t_k appears in documents of class c_i, $N_{i\bullet}$ is the total number of words in documents of class c_i, i.e. $N_{i\bullet} = \sum_{t_k} N_{ik}$, and M is the size of the vocabulary (the number of distinct words in the documents of the training set).

The estimation of the prior probabilities of the classes, $\hat{p}(c_i)$, is usually done by maximum likelihood, i.e.:

$$\hat{p}(c_i) = \frac{N_{i,doc}}{N_{doc}}\,, \quad (5)$$

where N_{doc} is the number of documents in the training set and $N_{i,doc}$ is the number of documents in the training set which are assigned to class c_i.

The multinomial naive Bayes model can also be used in another way: instead of considering only one class variable C having n values, we can decompose the problem using n binary class variables C_i taking its values in the sets $\{c_i, \overline{c}_i\}$. This is a quite common transformation in text classification [9], especially for multilabel problems, where a document may be associated to several classes. In this case n naive Bayes classifiers are built, each one giving a posterior probability $p_i(c_i|d_j)$ for each document. In the case that each document may be assigned to only one class (single-label problems), the class $c^*(d_j)$ such that $c^*(d_j) = \arg\max_{c_i}\{p_i(c_i|d_j)\}$ is selected. Notice that in this case, as the term $p_i(d_j)$ in the expression $p_i(c_i|d_j) = p_i(d_j|c_i)p_i(c_i)/p_i(d_j)$ is not necessarily the same for all the class values, we need to compute it explicitly through

$$p_i(d_j) = p_i(d_j|c_i)p_i(c_i) + p_i(d_j|\overline{c}_i)(1-p_i(c_i))\,.$$

This means that we have also to compute $p_i(d_j|\overline{c}_i)$. This value is estimated using the corresponding counterparts of eqs. (3) and (4), where

$$\hat{p}(t_k|\overline{c}_i) = \frac{N_{\bullet k} - N_{ik} + 1}{N - N_{i\bullet} + M}\,. \quad (6)$$

$N_{\bullet k}$ is the numbers of times that the term t_k appears in the training documents, i.e. $N_{\bullet k} = \sum_{c_i} N_{ik}$, and N is the total number of words in the training documents.

3 The OR Gate Bayesian Network Classifier

The document classification method that we are going to propose is based on another restricted type of Bayesian network with the following topology: Each term t_k appearing in the training documents (or a subset of these terms in the case of using some method for feature selection) is associated to a binary variable T_k taking its values in the set $\{t_k, \overline{t}_k\}$, which in turn is represented in the network by the corresponding node. There are also n binary variables C_i taking its values in the sets $\{c_i, \overline{c}_i\}$ (as in the previous binary version of the naive Bayes model) and the corresponding class nodes. The network structure is fixed, having an arc going from each term node T_k to the class node C_i if the term t_k appears in training documents which are of class c_i. Let nt_i be the number of different terms appearing in documents of class c_i. In this way we have a network topology with two layers, where the term nodes are the "causes" and the class nodes are the "effects", having a total of $\sum_{i=1}^{n} nt_i$ arcs. An example of this network topology is displayed in Figure 2. It should be noticed that the proposed topology, with arcs going from attribute nodes to class nodes, is the opposite of the one associated to the naive Bayes model.

[1] The length of the documents is also assumed to be independent on the class.

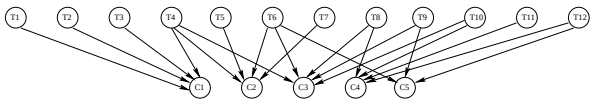


Figure 2: The OR gate classifier.

It should also be noticed that this network topology explicitly requires modeling the "discriminative" conditional probabilities $p(c_i|pa(C_i))$, where $Pa(C_i)$ is the set of parents of node C_i in the network (the set of terms appearing in documents of class c_i) and $pa(C_i)$ is any configuration of the parent set (any assignment of values to the variables in this set). As the number of configurations is exponential with the size of the parent set[2], we use a canonical model to define these probabilities, which reduce the number of required numerical values from exponential to linear size. More precisely, we use a noisy OR Gate model [7].

The conditional probabilities in a noisy OR gate are defined in the following way:

$$p(c_i|pa(C_i)) = 1 - \prod_{T_k \in R(pa(C_i))} (1 - w(T_k, C_i)) \ , \quad (7)$$

$$p(\overline{c}_i|pa(C_i)) = 1 - p(c_i|pa(C_i)) \, , \quad (8)$$

where $R(pa(C_i)) = \{T_k \in Pa(C_i) \,|\, t_k \in pa(C_i)\}$, i.e. $R(pa(C_i))$ is the subset of parents of C_i which are instantiated to its t_k value in the configuration $pa(C_i)$. $w(T_k, C_i)$ is a weight representing the probability that the occurrence of the "cause" T_k alone (T_k being instantiated to t_k and all the other parents T_h instantiated to $\overline{t}_h$) makes the "effect" true (i.e., forces class c_i to occur).

3.1 Classification as Inference

Once the weights $w(T_k, C_i)$ have been estimated, and given a document d_j to be classified, we instantiate in the network each of the variables T_k corresponding to the terms appearing in d_j to the value t_k (i.e. $p(t_k|d_j) = 1$ if $t_k \in d_j$), and all the other variables T_h (those associated to terms that do not appear in d_j) to the value $\overline{t}_h$ (i.e. $p(t_h|d_j) = 0 \ \forall t_h \notin d_j$). Then, we compute for each class node C_i the posterior probabilities $p(c_i|d_j)$. As in the case of the naive Bayes model, we would assign to d_j the class (or classes) having the greatest posterior probability.

The combination of network topology and numerical values represented by OR gates allows us to compute very efficiently and in an exact way the posterior probabilities:

$$\begin{aligned} p(c_i|d_j) &= 1 - \prod_{T_k \in Pa(C_i)} (1 - w(T_k, C_i) \times p(t_k|d_j)) \\ &= 1 - \prod_{T_k \in Pa(C_i) \cap d_j} (1 - w(T_k, C_i)) \ . \end{aligned} \quad (9)$$

In order to take into account the number of times a word t_k occurs in a document d_j, n_{jk}, we can replicate each node T_k n_{jk} times, so that the posterior probabilities then become

$$p(c_i|d_j) = 1 - \prod_{T_k \in Pa(C_i) \cap d_j} (1 - w(T_k, C_i))^{n_{jk}} \ . \quad (10)$$

3.2 Training as Weight Estimation

The estimation of the weights in the OR gates, $w(T_k, C_i)$, can be done in several ways. The simplest one is to compute $w(T_k, C_i)$ as $\hat{p}(c_i|t_k)$, the estimated conditional probability of class c_i given that the term t_k is present. We can do it by maximum likelihood:

$$w(T_k, C_i) = \frac{N_{ik}}{N_{\bullet k}} \, , \quad (11)$$

or using Laplace:

$$w(T_k, C_i) = \frac{N_{ik} + 1}{N_{\bullet k} + 2} \, . \quad (12)$$

Another, more accurate way of estimating $w(T_k, C_i)$ is directly as $\hat{p}(c_i|t_k, \overline{t}_h \, \forall T_h \in Pa(C_i), \, T_h \neq T_k)$. However, this probability cannot be reliably estimated, so that we are going to compute an approximation in the following way:

$$\hat{p}(c_i|t_k, \overline{t}_h \, \forall h \neq k) = p(c_i|t_k) \prod_{h \neq k} \frac{p(c_i|\overline{t}_h)}{p(c_i)} \quad (13)$$

This approximation results from assuming a conditional independence statement similar to that of the naive Bayes classifier, namely

$$p(t_k, \overline{t}_h \, \forall h \neq k|c_i) = p(t_k|c_i) \prod_{h \neq k} p(\overline{t}_h|c_i). \quad (14)$$

In that case

$$\begin{aligned} p(c_i|t_k, \overline{t}_h \, \forall h \neq k) &= \frac{p(t_k, \overline{t}_h \, \forall h \neq k|c_i) p(c_i)}{p(t_k, \overline{t}_h \, \forall h \neq k)} \\ &= \frac{p(t_k|c_i) \left(\prod_{h \neq k} p(\overline{t}_h|c_i) \right) p(c_i)}{p(t_k) \prod_{h \neq k} p(\overline{t}_h \, \forall h \neq k)} \\ &= p(c_i|t_k) \prod_{h \neq k} \frac{p(c_i|\overline{t}_h)}{p(c_i)} \end{aligned}$$

[2] Notice that $|Pa(C_i)| = nt_i$.

The values of $p(c_i|t_k)$ and $p(c_i|\overline{t}_h)/p(c_i)$ in eq. (13) are also estimated using maximum likelihood. Then, the weights $w(T_k, C_i)$ are in this case:

$$w(T_k, C_i) = \frac{N_{ik}}{N_{\bullet k}} \times \prod_{h \neq k} \frac{(N_{i\bullet} - N_{ih})N}{(N - N_{\bullet h})N_{i\bullet}} \qquad (15)$$

Another option is to relax the independence assumption in the following way:

$$p(t_k, \overline{t}_h \, \forall h \neq k|c_i) = \frac{p(t_k|c_i)}{nt_i} \prod_{h \neq k} p(\overline{t}_h|c_i). \qquad (16)$$

We are assuming that the joint probability of the events $[t_k, \overline{t}_h \, \forall h \neq k]$ is smaller than the pure independence assumption would dictate. The weights $w(T_k, C_i)$ would be in this case

$$w(T_k, C_i) = \frac{N_{ik}}{nt_i N_{\bullet k}} \times \prod_{h \neq k} \frac{(N_{i\bullet} - N_{ih})N}{(N - N_{\bullet h})N_{i\bullet}} \qquad (17)$$

In any case, the set of sufficient statistics necessary to compute the weights are N_{ik} $\forall t_k$, $\forall c_i$, i.e. the number of times each term t_k appears in documents of each class c_i, the same required by multinomial naive Bayes.

4 Experimentation

For the evaluation of the proposed model we have used three document test collections: Reuters-21578, Ohsumed and 20 Newsgroups. Reuters-21578 (ModApte split) contains 12,902 documents (9603 for training and 3299 for testing) and 90 categories (with at least 1 training and 1 test documents). Ohsumed, including 20000 medical abstracts from the MeSH categories (10000 for training and 10000 for testing) of the year 1991, and 23 categories. 20 Newsgroups corpus contains 19997 articles for 20 categories taken from the Usenet newsgroups collection, where only the subject and the body of each message were used. Note that there is no fixed literature split for this collection. All the three collections were preprocessed in the same way using stemming (Porter's algorithm) and stopword removal (SMART's system 571 stopword list). No term selection was carried out.

The evaluation takes into account that the classification process will generate an ordered list of possible categories, in decreasing order of probability[3], instead of a definite assignment of categories to each document. Then, as performance measures, we have firstly selected the typical measures used in multi-label categorization problems (as they are Reuters-21578 and Ohsumed): *breakeven* point[4] (BEP) and the *average 11-point precision*[5] (Av-11). Another measure commonly used is F_1[6]. However F_1 requires a precise assignment of classes to each document, so that we shall use instead F_1 at one (F_1@1) and also F_1 at three and five (F_1@3, F_1@5) document level: the F_1 value obtained by assuming that the system assigns to each document either the first or the first three or five most probable classes. Both breakeven and F_1 values will be computed in micro-average (micr.) and macro-average (macr.). In all the measures, a higher value means a better performance of the model.

We have executed experiments using naive Bayes and the proposed OR gate classifier, although we have also included another non probabilistic classifier in the comparison, namely the well-known Rocchio method [3], used as a perspective. Tables 1, 2 and 3 display the values of the performance measures obtained. We do not display the results of the OR gate classifier using all the different parameter estimation methods commented in Subsection 3.2. Instead, some preliminary experimentation showed that the best performing methods were those based on equation 12 (for estimation based on $\hat{p}(c_i|t_k)$) and equation 17 (for estimation based on $\hat{p}(c_i|t_k, \overline{t}_h \, \forall h \neq k)$).

Several conclusions can be drawn from these experiments: the proposed OR gate model is quite competitive, frequently outperforming Rocchio and naive Bayes. Particularly, this new model seems to perform quite well in terms of macro averages: it gives a more balanced treatment to all the classes, and this is especially evident in those problems where the class distribution is quite unbalanced, as Reuters and Ohsumed. At the same time, the OR gate model (the one based on eq. 17) performs generally well also in terms of micro averages.

5 Concluding Remarks

We have described a new approach for document classification, the so called "OR Gate classifier", with different variants based on several parameter estimation methods. It is based on a Bayesian network representation which is, in some sense, the opposite that the

[3] We are therefore using an instance of the so-called category-ranking classifiers [9].

[4] The point where precision equals recall, by moving a threshold.

[5] The precision values are interpolated at 11 points at which the recall values are 0.0, 0.1,. . ., 1.0, and then averaged.

[6] The harmonic mean of precision and recall.

	micr.BEP	macr.BEP	Av-11
Reuters			
NBayes	0.73485	0.26407	0.84501
Rocchio	0.47183	0.42185	0.84501
OR-eq12	0.66649	0.**55917**	0.81736
OR-eq17	**0.76555**	0.54370	**0.89725**
Ohsumed			
NBayes	**0.58643**	0.49830	**0.76601**
Rocchio	0.42315	0.44791	0.68194
OR-eq12	0.48017	**0.58792**	0.64739
OR-eq17	0.53122	0.56450	0.72925
20Newsgroups			
NBayes	0.71778	0.73629	**0.88834**
Rocchio	0.60940	0.63875	0.86583
OR-eq12	**0.80732**	**0.81333**	0.87889
OR-eq17	0.77689	0.79779	0.86208

Table 1: Micro and macro breakeven points and average 11-point precision.

	micr.F1@1	micr.F1@3	micr.F1@5
Reuters			
NBayes	**0.75931**	0.49195	0.35170
Rocchio	0.70047	0.47399	0.34979
OR-eq12	0.67297	0.45352	0.33493
OR-eq17	0.75369	**0.51461**	**0.36825**
Ohsumed			
NBayes	**0.53553**	**0.53979**	0.42718
Rocchio	0.46064	0.46313	0.40912
OR-eq12	0.41676	0.46329	0.40292
OR-eq17	0.49048	0.50883	**0.42773**
20Newsgroups			
NBayes	0.80881	**0.48705**	**0.33212**
Rocchio	0.77233	0.48323	0.33153
OR-eq12	**0.81000**	0.47266	0.32614
OR-eq17	0.77858	0.47284	0.32695

Table 2: Micro F_1 values.

one associated to the naive Bayes classifier. The complexity of the training and classification steps for the proposed model is equivalent to that of the naive Bayes too. In fact we can think of the OR gate model as a kind of discriminative version of naive Bayes, which is not a discriminative but a generative method.

According to the results of the experimental comparison carried out using several standard text collections, we found that the new model can compete with naive Bayes, especially in terms of macro averages and in those cases where the class distribution is unbalanced.

We believe that the OR gate model could greatly benefit from a term/feature selection preprocessing, and we plan to test this assertion in the future.

	macr.F1@1	macr.F1@3	macr.F1@5
Reuters			
NBayes	0.39148	0.28935	0.23116
Rocchio	0.39148	0.28935	0.23116
OR-eq12	0.11263	0.20092	0.26404
OR-eq17	**0.45762**	**0.39169**	**0.30959**
Ohsumed			
NBayes	0.40627	0.47260	0.40732
Rocchio	**0.45421**	0.50604	0.44945
OR-eq12	0.19980	0.42017	0.45870
OR-eq17	0.43602	**0.53615**	**0.50103**
20Newsgroups			
NBayes	**0.80985**	0.54983	0.39048
Rocchio	0.77095	0.54086	0.39113
OR-eq12	0.80880	0.58083	0.44502
OR-eq17	0.78682	**0.59099**	**0.44722**

Table 3: Macro F_1 values.

Acknowledgments. This work has been jointly supported by the Spanish Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía, Ministerio de Educación y Ciencia and the research programme Consolider Ingenio 2010, under projects TIC-276, TIN2005-02516 and CSD2007-00018, respectively.

References

[1] S. Acid, L.M. de Campos, J.G. Castellano. Learning Bayesian network classifiers: searching in a space of acyclic partially directed graphs. *Machine Learning* 59(3), 213-235 (2005).

[2] D. Koller, M. Sahami. Hierarchically classifying documents using very few words. In: *Proceedings of the Fourteenth International Conference on Machine Learning*, pp. 170–178, (1997).

[3] T. Joachims. A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization. In *Proceedings of ICML-97, 14th International Conference on Machine Learning*, pp. 143–151, (1997).

[4] L.S. Larkey, W.B. Croft. Combining classifiers in text categorization. In: *Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 289–297 (1996).

[5] D. Lewis, W. Gale. A sequential algorithm for training text classifiers. In: *Proceedings of the 17th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 3–12 (1994).

[6] A. McCallum, K, Nigam. A Comparison of event models for Naive Bayes text classification. In: *AAAI/ICML Workshop on Learning for Text Categorization*, pp. 137–142, AAAI Press (1998).

[7] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan and Kaufmann, San Mateo (1988).

[8] S.E. Robertson, K. Spärck-Jones. Relevance weighting of search terms. *Journal of the American Society for Information Science* 27, 129–146 (1976).

[9] F. Sebastiani. Machine Learning in automated text categorization. *ACM Computing Surveys* 34, 1–47 (2002).

EL VALOR EDUCATIVO DE LAS IDEAS DE LA LÓGICA BORROSA EN LA ENSEÑANZA DE LAS MATEMÁTICAS

Claudi Alsina Catalá
Sección Matemáticas. ETSAB. Univ. Politécnica de Catalunya.
Diagonal 649, 08028 Barcelona
claudio.alsina@upc.edu

Resumen

En este trabajo se analiza el valor educativo que podrían tener ideas clave de la lógica borrosa en los procesos de enseñanza y de aprendizaje de las matemáticas ampliando las posibilidades del razonamiento plausible descrito por Pólya y los procesos de matematización de Freudenthal.

Palabras Clave: Lógica Borrosa, Educación Matemática, Razonamiento, Modelización, Aplicaciones.

1 INTRODUCCIÓN

Uno de los problemas curriculares de la educación matemática actual, en sus diversos niveles, es el de la actualización constante de los contenidos. En un enfoque tradicional de programación, los contenidos clásicos agotan, por su riqueza y variedad, las posibilidades efectivas de implementación docente. Sin embargo el desarrollo de la matemática, y de las disciplinas científicas y tecnológicas inherentes, abre el debate de que tipo de implicaciones deberían tener las nuevas ideas en los procesos de enseñanza y aprendizaje. El tema es complejo pues toda actualización curricular pasa necesariamente por una minuciosa y previa actualización en el desarrollo profesional del profesorado y la necesidad de seleccionar temas esenciales exige la cautela de una elección muy estricta de todo aquello que tenga realmente sentido desde un punto de vista formativo.

En este contexto, el objetivo de este trabajo es hacer unas consideraciones sobre las aportaciones que la Lógica Borrosa [14] podría hacer a la formación matemática (véase [4]).

2 LÓGICA BORROSA Y COMPETENCIAS MATEMÁTICAS

La tendencia actual es superar los listados tradicionales de contenidos y fijar en cada caso los objetivos competenciales. Siguiendo el programa DeSeCo (OCDE, 2003) la *competencia* es la capacidad de afrontar con éxito demandas complejas en contextos particulares lo que en el caso de matemáticas ha llevado [9] a que se consideren esenciales las competencias de:

(i) Pensamiento y razonamiento matemático;

(ii) Planteamiento y resolución de problemas y de modelizaciones;

(iii) Comunicación de ideas matemáticas;

(iv) Uso de instrumentos, representaciones y formalismos.

Es precisamente en este marco teórico de la formación competencial donde ideas básicas de lógica borrosa tienen posibilidades formativas de interés. En el nivel universitario pueden ofrecerse cursos de lógica borrosa pero en el nivel pre-universitario creemos que no se trata de incluir capítulos de lógica borrosa (como hace décadas se incluyó teoría de conjuntos) pero sí de aprovechar *ideas, ejemplos* y *aplicaciones* que ayuden precisamente a abordar la complejidad del mundo, viendo como los nuevos instrumentos permiten abordar dicho tratamiento. Haciendo un simil, podríamos considerar que la situación es parecida a la de otras teoría no-tradicionales como la teoría de grafos o los fractales. Ningún currículo ([7]) de primaria o secundaria puede dedicar capítulos a estas teorías, sin embargo, la belleza de las imágenes fractales, su dibujo por ordenador, los esquemas de grafos aplicados (organigramas, mapas, etc.) si que tienen sentido y deben

tenerse en cuenta. No es tanto describir la esponja de Menger como entender el crecimiento fractal de las plantas o del sistema arterial humano.

3 LA COMPETENCIA EN LÓGICA BORROSA

La lógica borrosa aporta ideas interesantes, para la formacin, en tres niveles:

Nivel 1. *Modelización de fenómenos y del lenguaje natural a través de valoraciones numéricas que amplían el modelo clásico booleano.*

La insistencia en el modelo booleano de verdad-falsedad que solo se corresponde con proposiciones muy estrictas susceptibles de valoración 1-0 puede complementarse con las valoraciones en [0,1] de los conjuntos borrosos haciendo ver como la borrosidad lleva a interesantes consideraciones multi-valuadas, con especial referencia a las clasificaciones. Estas es una idea educativa muy rica.

Ejemplo 1. Las escalas usadas en evaluación escolar constituyen un claro ejemplo donde las etiquetas lingísticas (suspenso, aprobado, notable,... etc.) se corresponden con valoraciones numéricas.

Un interesante uso de conjuntos borrosos en evaluación ha sido desarrollado recientemente por Nykanen [10], Garriga [6] y Montero [8].

Ejemplo 2. *Escalas técnicas.*

En sismografía, huracanes y otros fenómenos de la naturaleza se han introducido interesantes escalas de evaluación (Richter, Mercalli, Fujita,...) de naturaleza borrosa. Algunas incluyen aparatos que miden determinados parámetros pero en otras e incluyen estadísticas de evaluaciones subjetivas de impactos. Una interesante colección de escalas a explotar didácticamente.

Nivel 2. *Ejemplos y aplicaciones de la lógica borrosa en problemas diversos, con especial atención a problemas de la vida cotidiana.*

Es especialmente interesante mostrar como con imágenes, problemas de control, automatismos, etc., la lógica borrosa aporta modelos interesantes en ingeniería, computación e inteligencia artificial [1,2,12]. Clips mostrando esta presencia real de la borrosidad darían interesantes informaciones (lavadoras, máquinas de fotografía, aparatos de presión,...).

Ejemplo 3. *Índice de pobreza humana*

El índice de pobreza humana HPI-1 introducido por la ONU se define mediante la media

$$HPI-1=\left[\left(P_1^3+P_2^3+P_3^3\right)/3\right]^{1/3},$$

donde P_1 es la probabilidad al nacimiento de no sobrevivir 40 años (multiplicada por 100), P_2 es la ratio de analfabetismo adulto y P_3 es la media de la población sin acceso sostenible a agua potable y la población infantil con peso inferior al normal. Normalizando este índice al [0,1] resulta un interesante índice borroso de pobreza en un páis.

Ejemplo 4. *Clasificaciones de formas*

En [13] J. Wagensberg plantea una definición borrosa de las formas que ha dado lugar a diversos módulos del museo Cosmocaixa (Madrid, Barcelona). En cada clase (espirales, ondas, fractales, ángulos, etc.) se agrupan formas naturales y de diseño y pueden introducirse de forma elemental.

Nivel 3. *Reconocimiento de la borrosidad presente en la propia formulación de muchos conceptos matemáticos.*

Muchos son los conceptos matemáticos (probabilidad, frecuencias relativas, errores relativos, proporciones,...) que puden ser reinterpretados o construidos con valores en [0,1], lo que permite hacer ver que más allá del subjetivismo que puede recogerse en el nivel 1 también la idea de borrosidad es atractiva en modelizaciones paramétricas.

Ejemplo 5. *El grado de aproximación de una figura plana a una circunferencia*

Siguiendo la idea de Pólya ([11]), dado que en cualquier figura plana con área A, limitada por una curva cerrada rectificable de longitud L, vale la desigualdad isoperimétrica $L^2 \leq 4\pi A$ con igualdad en el caso de la circunferencia, el cociente $L^2/4\pi A$ en [0,1] evalua el grado en que la figura se aproxima a un círculo.

Ejemplo 6. *El grado de perpendicularidad*

La propia definición trigonométrica de la función seno en triángulos rectángulos da una clara medida de la perpendicularidad: sen $A = 1$ solo si $A = \pi/2$. La función cos A da el grado de dependencia y resulta dual de sen A mediante la negación circular $n(x) = \sqrt{1-x^2}$.

Especificados estos tres niveles posibles de tratamiento de conceptos borrosos tiene sentido formular el desarrollo de la *competencia en borrosidad* como la ca-

pacidad de matematizar situaciones mediante la construcción de valores borrosos, cuantificando situaciones complejas y comparando modelos alternativos de descripción.

4 BORROSIDAD Y RAZONAMIENTO

Desarrollar el razonamiento deductivo clásico de la matemática es un proceso complejo desde el punto de vista del aprendizaje. A este desarrollo contribuye el ejercicio de la argumentación y tener ideas intuitivas, es decir, en el sentido de Pólya, ir construyendo el razonamiento plausible ([11]). Así pues debería entenderse que el propio desarrollo del razonamiento deductivo pasa por sucesivas etapas de borrosidad, desde un nivel puramente experimentativo y de observación hasta un nivel ya más abstracto y susceptible de formalización. Algo parecido a lo que en el aprendizaje de la geometría ha sido descrito por los niveles de los Van Hiele.

El caso más evidente de borrosidad puede apreciarse en el proceso de resolución de problemas descrito por Miguel de Guzmán ([5]) donde la elección inicial de estrategias posibles para abordar una posible solución requiere un tanteo que tiene muy poco de booleano y una fuerte componente de borrosidad (véase [3] en resolución gráfica de demostraciones).

Una de las dificultades de aprendizaje de las matemáticas en el nivel secundario obligatorio que ha detectado el estudio PISA de la OCDE es precisamente el problema de comprensión de los enunciados verbales de problemas y la aplicación de los recursos matemáticos curriculares a situaciones reales. En el fondo es la dificultad de introducir precisión (matemáticas) en contextos cargados de borrosidad (realidad).

5 A MODO DE CONCLUSIÓN

Por todo lo expuesto parece razonable considerar que ideas de la lógica borrosa podrían formar parte del aprendizaje matemático, especialmente si desea trabajarse la modelización y aplicaciones como motor educativo. Pero esta consideración lleva necesariamente a divulgar estas buenas ideas borrosas y ofrecer situaciones didácticas adecuadas. Pólya acaba [11] diciendo *"...buena parte de nuestra confianza en los razonamientos demostrativos puede proceder del razonamiento plausible"*. La borrosidad es una característica de nuestra realidad y de nuestras formas de pensar que puede tener valor formativo.

Referencias

[1] C. Alsina. (2004) *Contar bien para vivir mejor*, Rubes, Barcelona.

[2] C. Alsina. (2008) *Vitaminas matemáticas*, Ariel, Barcelona.

[3] C. Alsina y R.B. Nelsen. (2008) *Math Made Visual. Creating Images for Understanding Mathematics*, MAA, Washington.

[4] C. Alsina y E. Trillas. Fuzzy Sets and Mathematics Education, For the learning of Mathematics, Univ. Concordia Montreal, V. 1. N. 3, 16-19, (1991).

[5] M. de Guzmán. (1994) *Para pensar mejor,* Labor, Barcelona.

[6] G. Garriga. (2005) *A new approach to the synthesis of fuzzy systems form input-outpput data.* PhD. tesis, Engiyeria i Arquitectura La Salle. Universitat Ramon Llull.

[7] Ministerio de Educación y Ciencia. Evaluación PISA 2003. Resumen de los resultados en España, 2003. http://www.ince.mec.es/pub/pisa2003resumenespana.pdf (octubre 2007)

[8] J.A. Montero, R.M. Alsina, J.A. Morán y M. Cid. Fuzzy logic system for student's evaluacion. *Lecture Notes in Computer Science,* 3512: 1246-1253, (2005).

[9] M. Niss. (2003) Mathematical competencies and the learning of mathematics, en A.A. Gnatsi et al. (eds). 3r. Med. Conf. Math. Zd., Math. Soc., Atenas, 115-124.

[10] O. Nykanen. Inducing fuzzy models for student classificaiton. *Educational Technology and society*, 9(2): 223-234, (2006).

[11] G. Pólya. (1966) *Matemática y Razonamiento Plausible*, Tecnos, Madrid.

[12] E. Trillas. (1998) *Inteligencia Artificial. Máquinas y personas.* Debate, Madrid.

[13] J. Wagensberg. (2004) *La Rebelión de las Formas. O como perseverar cuando la incertidumbre aprieta.* Metatemas 84, Tusquets Editors, Barcelona.

[14] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338-353, (1965).

DiffuCalc: A simple computational tool for Fuzzy-Logic based modelling

Luis Argüelles Méndez [1]
Carbonar S.A., R+D Department
Cardenal Cienfuegos, 8., 33007 - Oviedo, Spain
[1] arguelles@carbonar.es

Abstract

Commercial software toolboxes for creating, modifying and visualizing fuzzy-sets and fuzzy-logic based systems are available from several vendors throughout the world. This work presents *Diffucalc*, a free, world wide web downloadable computing environment for building and exploring simple to medium size scientific and technical models based on Fuzzy-Logic.

Keywords: Spreadsheet, fuzzy-logic toolbox, fuzzy modelling, control systems.

1 Introduction

Running under the actual powerful computers, spreadsheets allow to calculate not only large financial models, but also to solve complex technical problems. In fact, simple to medium size fuzzy-logic based models can be designed in the framework of electronic spreadsheets if some simplifications are assumed. *DiffuCalc* accepts linguistic variables based on trapezoidal or triangular normal fuzzy-sets. From an adequate arrangement of the a_i values on the support of the fuzzy-sets at play, a suitable set of membership functions can be modelled, as can be seen in figures 1-4. *DiffuCalc* supports models with two linguistic variables as antecedents and one output variable as a consequent, formed by Sugeno's Singletons[**3**] in order to allow a simple and efficient defuzzyfication process. Later, and by means of linking several spreadsheets, more complex topologies can be implemented. As we shall see, this type of fuzzy model admits a matrix-based formalism that can be easily implemented in the cells of any spreadsheet.

In this work, we shall first present the mathematical basis of *DiffuCalc* and later, an aircraft landing control model will be shown as an application example before reaching the final conclusions.

2 Mathematical Formalism

2.1. Fuzzyfication

A simple Fuzzy-Logic based system can be formed from a set of knowledge rules relating two input linguistic variables and one output linguistic variable. The input linguistic variables M_1 and M_2, containing n and m fuzzy-sets respectively, can be written in matrix form in the following way:

$$M_1 = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ . & . & . & . \\ a_{n,1} & a_{n,2} & a_{n,3} & a_{n,4} \end{bmatrix}$$

$$M_2 = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ . & . & . & . \\ b_{m,1} & b_{m,2} & b_{m,3} & b_{m,4} \end{bmatrix}$$

The output variable S can be written as a set of Sugeno's Singletons:

$$S = [s_1, s_2, ..., s_t]$$

From two input values p, q, belonging respectively to the universes of discourse where M_1 and M_2 are defined, we can build two vectors whose elements are the membership values to every fuzzy-set contained on the linguistic variables M_1, M_2:

$$V_1 = [\mu_1, \mu_2, ..., \mu_n]$$

$$V_2 = [\mu_1', \mu_2', ..., \mu_m']$$

Every μ_i membership degree is given by the expression:

$$\mu_i = \max(a, b, c)$$

where, observing figures 1-4, we can see that:

$$if\,(p > a_{i,1}) \wedge (p < a_{i2}) \rightarrow a = \frac{p - a_{i1}}{a_{i2} - a_{i1}}$$
$$otherwise \rightarrow a = 0$$

$$if\,(p \geq a_{i2}) \wedge (p \leq a_{i3}) \rightarrow b = 1.0$$
$$otherwise \rightarrow b = 0$$

$$if\,(p > a_{i,3}) \wedge (p < a_{i4}) \rightarrow c = \frac{a_{i4} - p}{a_{i4} - a_{i3}}$$
$$otherwise \rightarrow c = 0$$

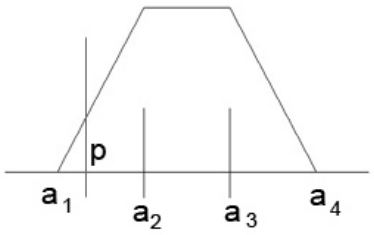


Figure #1: A symmetrical trapezoidal Fuzzy-Set

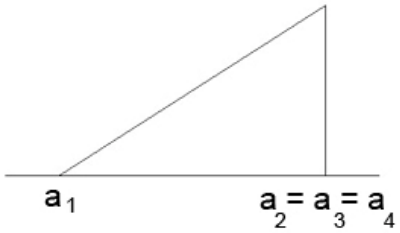


Figure #2: A triangular non-symmetrical Fuzzy-Set

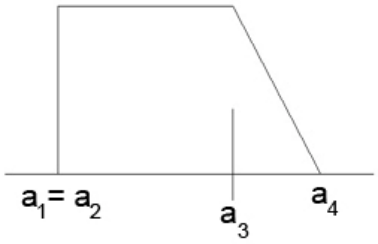


Figure #3: A non-symmetrical trapezoidal Fuzzy-Set

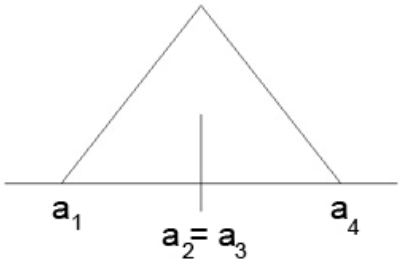


Figure #4: A triangular symmetrical Fuzzy-Set

It easy to see that μ'_i values are calculated in a similar way.

2.2. Rules, inference and defuzzyfication

As it is well known, a common way of representing human knowledge is to build expressions of natural language of the type:

IF antecedent$_1$ AND/OR antecedent$_2$ THEN consequent

Under a computational point of view, every fuzzy rule can be seen as a relationship between real numbers, so we can translate this expression into

IF μ_i AND/OR μ'_i THEN $\lambda_i s_i$

Where λ_i are real numbers obtained after applying either a t-norm or t-conorm[1] on the membership degrees μ_i and μ'_i, always remembering that s_i are singleton values. For convenience, from vectors V_1 and V_2, we can define a compositional matrix Q expressed as:

$$Q = \begin{bmatrix} \mu_1 & \mu'_1 \\ \mu_1 & \mu'_2 \\ . & . \\ \mu_1 & \mu'_m \\ \mu_2 & \mu'_1 \\ \mu_2 & \mu'_2 \\ . & . \\ \mu_2 & \mu'_n \\ . & . \\ \mu_n & \mu'_m \end{bmatrix}$$

From this matrix, we can build a second one adding some columns. The third one is the result of computing the λ_i values from the sets of memberships expressed by V_1 and V_2. For simplicity, we shall use in this work an algebraic product $\mu.\mu'$

The fourth column in the matrix is represented by the set S of Sugeno's singletons from the output's universe of discourse, and the fifth column is the result of multiplying the third and fourth columns, that is, the result of calculating $\lambda_i s_i$. The resulting matrix R is shown as follows:

$$R = \begin{bmatrix} \mu_1 & \mu_1' & \mu_1\mu_1' & s_1 & \mu_1\mu_1' s_1 \\ \mu_1 & \mu_2' & \mu_1\mu_2' & s_2 & \mu_1\mu_2' s_2 \\ . & . & .. & . & ... \\ . & . & .. & . & ... \\ \mu_1 & \mu_m' & \mu_1\mu_m' & s_{i-1} & \mu_1\mu_m' s_{i-1} \\ \mu_2 & \mu_1' & \mu_2\mu_1' & s_i & \mu_2\mu_1' s_i \\ \mu_2 & \mu_2' & \mu_2\mu_2' & s_{i+1} & \mu_2\mu_2' s_{i+1} \\ . & . & .. & . & ... \\ . & . & .. & . & ... \\ \mu_n & \mu_m' & \mu_n\mu_m' & s_t & \mu_n\mu_m' s_t \end{bmatrix}$$

The summation of the third and fifth columns produces two real numbers A and B such as:

$$A = \sum \mu_i \mu_i'$$

$$B = \sum \mu_i \mu_i' s_i$$

And finally, the deffuzified output Z of the fuzzy-logic based model exposed in this work is given by B/A, that is:

$$Z = \frac{\sum \mu_i \mu_i' s_i}{\sum \mu_i \mu_i'}$$

3 An application example

The following application must be understood as a simple example of an aircraft landing control model adapted from[**2**] in order to show the implementation made by *DiffuCalc*, and not a finished model of a real aircraft landing control system.

The model describes the final descent and landing approach of an aircraft, where as the height (altitude) of the aircraft diminishes, the desired downwards velocity must be smaller and smaller. In this way the aircraft is able to promptly descend from altitude but touches down gently at the airport.

Figure #5: Schematic diagram for an aircraft landing control model

As can be seen in figures 5-8, the model has two input variables named "Height" and "Vertical velocity" and an output variable named "Control force" representing a force than, when applied to the aircraft, alters its height and velocity. The rules for controlling the system are given in Table 1.

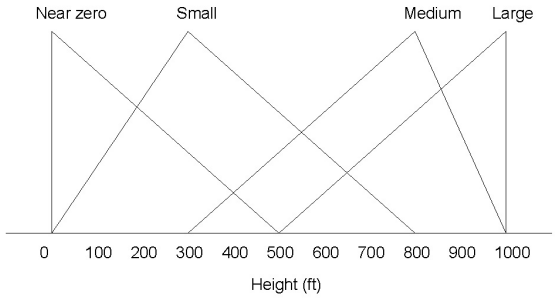


Figure #6: Linguistic variable for "Height" (input)

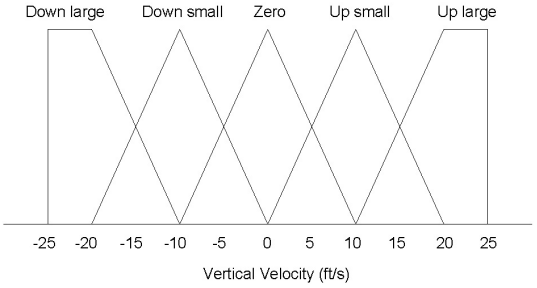


Figure #7: Linguistic variable for "Vertical velocity" (input)

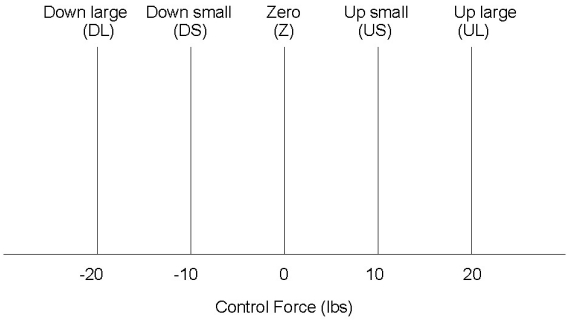


Figure #8: Linguistic variable for "Control force" as Singletons (output)

Table 1: Model's rules

	Velocity				
Height	**DL**	**DS**	**Zero**	**US**	**UL**
L	Z	DS	DL	DL	DL
M	US	Z	DS	DL	DL
S	UL	US	Z	DS	DL
NZ	UL	UL	Z	DS	DS

Given two input values Height = 600 ft, Vertical velocity = -11 ft/s and having into account the material exposed on the previous section, we can write the following matrixes:

$$M_1 = \begin{bmatrix} 0 & 0 & 0 & 500 \\ 0 & 300 & 300 & 800 \\ 300 & 800 & 800 & 1000 \\ 500 & 1000 & 1000 & 1000 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} -25 & -25 & -20 & -10 \\ -20 & -10 & -10 & 0 \\ -10 & 0 & 0 & 10 \\ 0 & 10 & 10 & 20 \\ 10 & 20 & 25 & 25 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0.1 & 0 & 20 & 0 \\ 0 & 0.9 & 0 & 20 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -10 & 0 \\ 0 & 0 & 0 & -10 & 0 \\ 0.4 & 0.1 & 0.04 & 20 & 0.8 \\ 0.4 & 0.9 & 0.36 & 10 & 3.6 \\ 0.4 & 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & -10 & 0 \\ 0.4 & 0 & 0 & -10 & 0 \\ 0.6 & 0.1 & 0.06 & 10 & 0.6 \\ 0.6 & 0.9 & 0.54 & 0 & 0 \\ 0.6 & 0 & 0 & -10 & 0 \\ 0.6 & 0 & 0 & -20 & 0 \\ 0.6 & 0 & 0 & -20 & 0 \\ 0.2 & 0.1 & 0.02 & 0 & 0 \\ 0.2 & 0.9 & 0.18 & -10 & -1.8 \\ 0.2 & 0 & 0 & -20 & 0 \\ 0.2 & 0 & 0 & -20 & 0 \\ 0.2 & 0 & 0 & -20 & 0 \end{bmatrix}$$

then, after performing the appropriate summations, the output Z (control force) will be:

$$Z = \frac{3.20}{1.20} = 2.67 \text{ lb.}$$

Since *DiffuCalc* is a predesigned data and algorithm structure, implementing fuzzy-based models on it is a straightforward task, and users only need to copy and paste cells with only a bit of editing. In fact, the exposed aircraft landing control model can be written in *DiffuCalc* in only a few minutes.

	A	B	C	D	E
1	**Linguistic Variable #1**				
2	Variable's Name: Height (ft)				
3	Near Zero	0	0	0	500
4	Small	0	300	300	800
5	Medium	300	800	800	1000
6	Large	500	1000	1000	1000
7	**Data:**	600			
8					
9					
10	**Linguistic Variable #2**				
11	Variable's Name: Vertical Velocity (ft/s)				
12	Down large	-25	-25	-20	-10
13	Down small	-20	-10	-10	0
14	Zero	-10	0	0	10
15	Up small	0	10	10	20
16	Up large	10	20	25	25
17	**Data:**	-11			

Figure #9: Matrixes M_1 and M_2 written in spreadsheet form

22	**Memberships to "Height"**		**Memberships to "Vertical Velocity"**	
23	**Near Zero**	**0,000**	**Down large**	**0,100**
24	**Small**	**0,400**	**Down small**	**0,900**
25	**Medium**	**0,600**	**Zero**	**0,000**
26	**Large**	**0,200**	**Up small**	**0,000**
27			**Up large**	**0,000**

Figure #10: Vectors V_1 and V_2 written in spreadsheet form

31	If Height is	Near Zero	**0,0**	and	Vel. Is	Down large	**0,1**	Force is	Up large	**0,00**	20	0,00
32	If Height is	Near Zero	**0,0**	and	Vel. Is	Down small	**0,9**	Force is	Up large	**0,00**	20	0,00
33	If Height is	Near Zero	**0,0**	and	Vel. Is	Zero	**0,0**	Force is	Zero	**0,00**	0	0,00
34	If Height is	Near Zero	**0,0**	and	Vel. Is	Up small	**0,0**	Force is	Down small	**0,00**	-10	0,00
35	If Height is	Near Zero	**0,0**	and	Vel. Is	Up large	**0,0**	Force is	Down small	**0,00**	-10	0,00
36												
37	If Height is	Small	**0,4**	and	Vel. Is	Down large	**0,1**	Force is	Up large	**0,04**	20	0,80
38	If Height is	Small	**0,4**	and	Vel. Is	Down small	**0,9**	Force is	Up small	**0,36**	10	3,60
39	If Height is	Small	**0,4**	and	Vel. Is	Zero	**0,0**	Force is	Zero	**0,00**	0	0,00
40	If Height is	Small	**0,4**	and	Vel. Is	Up small	**0,0**	Force is	Down small	**0,00**	-10	0,00
41	If Height is	Small	**0,4**	and	Vel. Is	Up large	**0,0**	Force is	Down small	**0,00**	-10	0,00
42												
43	If Height is	Medium	**0,6**	and	Vel. Is	Down large	**0,1**	Force is	Up small	**0,06**	10	0,60
44	If Height is	Medium	**0,6**	and	Vel. Is	Down small	**0,9**	Force is	Zero	**0,54**	0	0,00
45	If Height is	Medium	**0,6**	and	Vel. Is	Zero	**0,0**	Force is	Down small	**0,00**	-10	0,00
46	If Height is	Medium	**0,6**	and	Vel. Is	Up small	**0,0**	Force is	Down large	**0,00**	-20	0,00
47	If Height is	Medium	**0,6**	and	Vel. Is	Up large	**0,0**	Force is	Down large	**0,00**	-20	0,00
48												
49	If Height is	Large	**0,2**	and	Vel. Is	Down large	**0,1**	Force is	Zero	**0,02**	0	0,00
50	If Height is	Large	**0,2**	and	Vel. Is	Down small	**0,9**	Force is	Down small	**0,18**	-10	-1,80
51	If Height is	Large	**0,2**	and	Vel. Is	Zero	**0,0**	Force is	Down large	**0,00**	-20	0,00
52	If Height is	Large	**0,2**	and	Vel. Is	Up small	**0,0**	Force is	Down large	**0,00**	-20	0,00
53	If Height is	Large	**0,2**	and	Vel. Is	Up large	**0,0**	Force is	Down large	**0,00**	-20	0,00

Figure #11: Matrix R written in spreadsheet form

Figures 9-11 show how the set of matrixes described in this section is organized in a commercially available spreadsheet. Since the mathematical computations are predefined, implementing a fuzzy-based model is only a question of "filling in the blanks" operation. Especially

interesting is the structure shown in Figure 11, where a complete implementation of the matrix R is shown. As can be appreciated, not only the five columns in the matrix are present, but also a set of linguistic labels that help users to friendly implement the fuzzy rules.

4 Conclusions

Aside being freely available from www.diffucalc.org, *DiffuCalc* is virtually independent from any operating system or computing environment, since it can run on any spreadsheet supporting the most basic logical formulae and even on handheld computers, as can be seen in Figure 12.

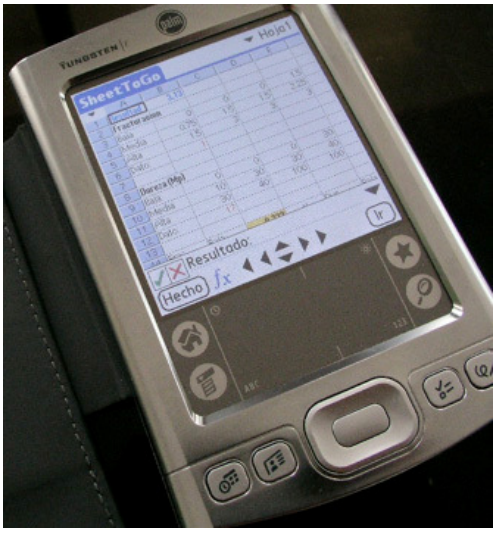


Figure 12: *Diffucalc* runnig on a handheld computer

Its simplicity and easy to use interface makes it an excellent tool not only for fuzzy-logic prototyping and quick developing, but also for teaching and learning the foundations of fuzzy-logic.

References

[1] Klir, G., Yuan, B., Fuzzy Sets and Fuzzy Logic, Prentice-Hall, 1995.

[2] Ross, T., Fuzzy Logic with Engineering Applications, McGraw-Hill, 1995

[3] Sugeno, M., Industrial applications of fuzzy control, Elsevier Science Pub. Co., 1985

LA HOJA DE CÁLCULO: UN ENTORNO PARA LA ENSEÑANZA Y ESTUDIO DE RELACIONES BORROSAS

Jordi Casabó [1] **Joan Jacas** [1] **Jordi Recasens** [2]

[1] Sección de Matemáticas y Informática, ETSAB Universidad Politécnica de Cataluña
Diagonal 649, 08028 Barcelona
{Jorge.Casabo,Joan.Jacas}@upc.edu
[2] Sección de Matemáticas y Informática, ETSAV Universidad Politécnica de Cataluña,
Pere Serra, 1-15, 08190 Sant Cugat de Vallès
j.recasens@upc.edu

Resumen

En este trabajo se estudia la posibilidad de introducir conceptos de teoría de conjuntos borrosos en los currículos correspondientes a distintos niveles de enseñanza. Se hace especial hincapié en la enseñanza de las relaciones borrosas presentando un entorno Excel© como soporte docente y de experimentación

Palabras clave: Relaciones borrosas, software educativo, hoja de cálculo, enseñanza de conjuntos borrosos.

1 Introducción

En el mercado existen diversos programas diseñados expresamente para la manipulación de relaciones borrosas enfocadas a la modelización de sistemas y a problemas de clasificación automática. por ejemplo:
http://www.mathworks.com/products/fuzzylogic/.
Sin embargo, todos los programas consultados están diseñados para desarrollos industriales o de investigación.
[ver por ejemplo: companies supplying fuzzy tools in
http://www.cs.cmu.edu/groups/AI/html/ai/fuzzy].
En el ámbito educativo, si bien existe una extensa literatura utilizable en cursos de segundo ciclo universitario o postgrado [12, 13, 15], hay escasos ejemplos de software educativo [8] que no estén dirigidos a los usuarios mencionados anteriormente y requieren, en general, la compra de determinadas licencias o la compilación de archivos. Dentro del ámbito específico de la enseñanza y manipulación de relaciones borrosas, no hemos encontrado ninguna aplicación específica. Echábamos en falta un entorno interactivo y amable que permitiera, con un coste bajo de aprendizaje, la construcción, simulación y visualización de las propiedades más importantes de las relaciones borrosas, así como de algunos de los métodos de clasificación borrosa. Por ello, hemos iniciado el desarrollo, bajo el formato de hoja de cálculo, de un software educativo para la generación y estudio de relaciones borrosas.

2 Lógica borrosa y currículos docentes

Si hacemos un repaso de la situación actual de los currículos docentes a nivel europeo [14], nos encontramos que todos los cursos están dirigidos a alumnos universitarios de segundo ciclo o de postgrado. No hemos constatado la existencia de estos temas en currículos de enseñanza primaria o media. Sin embargo, a nivel de Ciencias de la Educación, hemos encontrado distintos artículos en donde se propone la lógica borrosa (o difusa) como nueva herramienta [1] para superar la lógica clásica a fin de afrontar el paradigma de la complejidad formulado por Morin [11]. En otros, se propone la introducción de conceptos de teoría de conjuntos borrosos en currículos de matemáticas de enseñanzas secundarias [10].
La viabilidad de esta propuesta requiere por una parte, la formación del profesorado y por otra, la cuidadosa elección de los conceptos a introducir puesto que se puede caer en el mismo error que se cometió con la introducción a nivel de enseñanza primaria de la teoría de conjuntos "a lo Papy".
La introducción tendría más sentido si se hiciera bajo el paradigma docente de enseñanza basada en problemas.

3 Software educativo sobre conjuntos borrosos

En el ámbito de ingenierías existe un creciente interés por la introducción de conceptos de lógica borrosa, no sólo como conceptos de aplicación, sino también como herramientas de carácter pedagógico [6,7].
De hecho, todo el software de carácter educativo existente presenta entornos para la modelización y desarrollo de

sistemas de control, sistemas expertos, de soporte a la toma de decisiones. Un ejemplo lo constituye el "fuzzy toolbox" para Matlab®. Sin embargo, existen muy pocos ejemplos de software educativo enfocado al estudio y aprendizaje de los conceptos básicos de teoría de conjuntos borrosos y lógica borrosa. Software que permita trabajar la aritmética borrosa, comparar las distintas operaciones conjuntistas, estudiar las conectivas, etc.
Por todo ello, creemos que es interesante el desarrollo de un entorno para la enseñanza de lógica borrosa a nivel de último curso de secundaria o primer ciclo de universidad.
En este artículo presentamos un ejemplo de desarrollo de un material para la enseñanza, el estudio, generación y utilización de relaciones borrosas.

4 Un entorno Excel© para relaciones borrosas

Ante la necesidad de elegir un entorno informático, planteamos una serie de requerimientos que éste debiera cumplir:

1. Ser interactivo en el sentido de que el usuario pueda experimentar fácilmente con él.
2. Ser fácilmente automatizable y/o programable para obtener aplicaciones de soporte que visualicen o pongan de manifiesto determinados conceptos y propiedades.
3. Tener la suficiente potencia de cálculo que nos permita hacer simulaciones de situaciones reales o problemas verosímiles.
4. Por último, y no por ello menos importante, disponer de un entorno fácilmente accesible para todos los usuarios.

El entorno que mejor se adapta a los requerimientos establecidos es la hoja de cálculo en general y en concreto la propuesta que presentamos: los libros de Excel© [16].

Algunas de las ventajas que ofrecen estos entornos, muy extendidos entre los usuarios, las podemos resumir en:

1. El entorno de las hojas de cálculo es de fácil y rápido aprendizaje.

2. Son muy interactivas, son lo más parecido al lápiz y papel tradicional, ya que todo queda a la vista.

3. Permiten un alto grado de simulación puesto que una vez establecido el modelo y el algoritmo que lo implementa es muy fácil experimentar con situaciones análogas.

4. Con ellos es muy fácil implementar modelos recursivos, (copiar y pegar), que en otros entornos son difíciles de controlar (Figura 1).

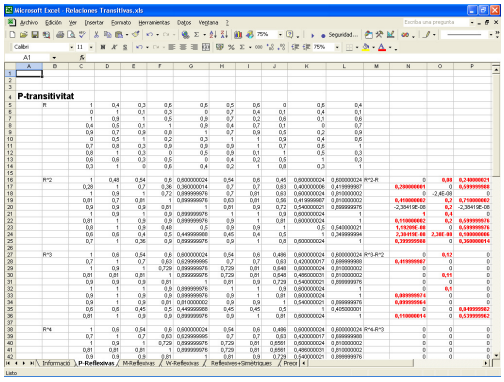


Figura 1: cálculo recursivo de la clausura P-transitiva

En el entorno Excel© podemos añadir otras ventajas:

1. Dispone de un lenguaje de alto nivel, *Visual Basic for Applications* (**VBA**), con el que se pueden diseñar aquellas aplicaciones que el uso interactivo del entorno haría impracticables. Este sería el caso de la aplicación que hemos diseñado para la visualización de t-normas, aplicación que también puede utilizarse para la representación de buen número de funciones de dos variables (figura 2).

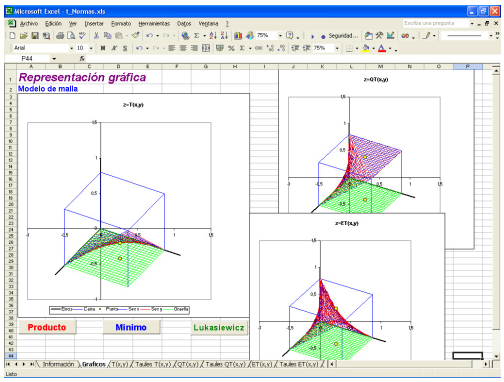


Figura 2: Representación de funciones con Excel©

2. Nos permite crear nuestras propias funciones de usuario con lo cual el uso interactivo de este entorno se hace mucho más fácil.

3. De cara a su uso didáctico, permite fácilmente la fragmentación de los temas ("libros") en sub-temas ("hojas").

En nuestra implementación, hemos optado por crear pequeños entornos que ejemplifican los distintos aspectos del tema, que van de la representación de las t-normas más utilizadas hasta la simulación numérica de los

teoremas de representación para preordenes y relaciones de T-indistinguibilidad.

2.1 Descripción de los distintos entornos

1. ***Herramientas.xls***.
   Contiene los procedimientos y funciones de usuario usadas en los distintos entornos.

2. ***T-normas.xls***.
   Entorno de demostración en que se visualizan la t-norma Producto, Mínimo y de Lukasiewicz, así como sus residuaciones y birresiduaciones.

3. ***Relaciones Fuzzy.xls***.[3]
   Este libro constituye en núcleo central del entorno que presentamos.
   En las hojas "*P-Reflexivas", "M-Reflexivas"* y "W-*Reflexivas"* se construyen las clausuras transitivas de una relación reflexiva respecto a las t-normas producto, mínimo y de Lukasiewicz .
   La hoja "*Reflexiva+Simetrica"* se dedica a la clausura transitiva de una relación reflexiva y simétrica.
   La construcción de preordenes i T-indistinguibilades unidimensionales se trata en las hojas "*Preordenes"* y "*T-indistinguibilidades*".
   El teorema de representación para varios conjuntos borrosos se trata de forma manual y automatizada en "*TR-Manual"* y "*TR-Auto"*.
   Finalmente, el tema de las T-indistinguibilidades descomponibles se trata en "*Descomponibles"*.
   En este entorno, se dispone además de un modulo con procedimientos de generación aleatoria de matrices, con las que se pueden generar relaciones reflexivas y/o simétricas ("*GenRang", "GenReflex"* y "*GenReflexSimet"*).
   Así mismo, se ha diseñado una función definida de usuario para el teorema de representación, "*TR()"*, aplicable hasta un máximo de 29 relaciones.

4. ***OWA.xls***.
   Construcción de indistinguibilidades mediante agregación [9].
   En "*OWA Demo"* se construye, tanto manualmente como de forma automática, las "OWA" de 5 relaciones, tanto para la t-norma Producto como para la de Lukasiewicz.
   En este entorno, se dispone de un módulo con dos funciones definidas de usuario, "*OWAProd"* y *"OWALuk",* que generan "OWA's" generalizadas para un máximo de 29 relaciones.

5. ***Fuzzy c_Means***.

   Si bien el problema de clasificación borrosa no es un tema directamente ligado con la enseñanza de relaciones borrosas, hemos incluido este libro como ejemplo del potencial que tiene la hoja de cálculo para la implementación y docencia de procesos recursivos.
   En este libro, disponemos de dos versiones, una manual (interactiva) y otra automatizada del programa de clasificación Fuzzy c-Means [2], en las que se obtiene la clasificación borrosa de un número no superior a 125 puntos u "objetos", cada uno de ellos con un número determinado de características o "eventos". La limitación del número de objetos es debida al hecho de que las hojas de cálculo admiten un máximo 256 columnas. Este entorno es eminentemente visual, en cada una de sus hojas se visualiza un ejemplo concreto con ayuda de los gráficos integrados que ofrece el programa
   Se dispone de un modulo con un procedimiento, "*Fuzzy_c-Means*", de clasificación automática. Asi mismo disponemos de tres procedimientos más: "*ComprobarPuntos*" analiza si los datos referentes a los "eventos" de los "objetos" sean realmente numéricos. "*ComprobarMatriz*" verifica si se ha escrito correctamente la matriz de pertenencia. "*CrearMatriz*" escribe de forma automática la matriz de pertenencia.

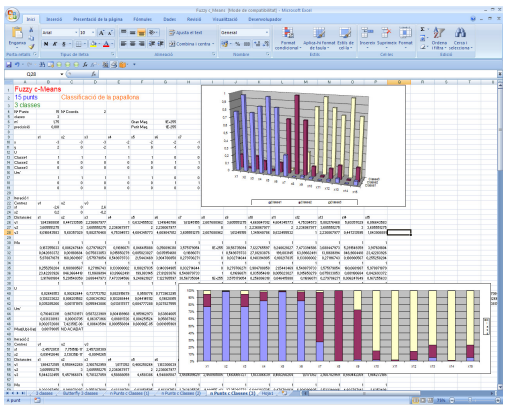


Figura 3. Clasificación de la base de datos "butterfly"

6. ***SGBD Fuzzy***.
   Versión de "***Fuzzy c_Means"*** que vence la limitación anterior pudiéndose clasificar un número de objetos superior a 256. En este entorno el procedimiento de clasificación es "*FuzzyBD*". Los datos originales deben encontrarse en una hoja de un libro abierto, en donde cada fila contiene los distintos eventos de cada objeto, el número máximo de objetos a clasificar es 1024. El procedimiento genera un nuevo libro que recoge el proceso iterativo de clasificación. Además

en el libro original se añade una hoja, de nombre "*Mat. Pertenencia*", en la queda escrita la matriz de pertenencia resultante.

5 Conclusiones

La enseñanza de conceptos de lógica borrosa, fuera del ámbito universitario de las ingenierías, no se ha hecho efectivo hasta ahora por distintas razones:

1 La necesidad de preparación específica del profesorado.
2 Resistencia por parte del profesorado universitario que enseña matemáticas.
3 Falta de material didáctico experimental.

Sin embargo, creemos que existe una creciente penetración de los conceptos borrosos en profesores del área de Ciencias de la Educación tanto a nivel de utilización del paradigma "fuzzy" para evaluación como su utilización en docencia para estudio de fenómenos reales. Por ello creemos interesante la creación de un material de nivel polivalente para la enseñanza de las bases de la teoría de conjuntos borrosos y lógica borrosa que pueda ser utilizada tanto en los primeros ciclos universitarios como en cursos de enseñanza secundaria.

Referencias

[1] L. Ballester, Antonio J. Colom. Lógica difusa: una nueva epistemología para las Ciencias de la Educación, Revista Educación vol 340 Mayo-Agosto 2006 pp 995-1008.

[2] J. C. Bezdek. (1981) Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York.

[3] D. Boixader, J. Jacas, J. Recasens. (2000) Fuzzy equivalence relations, advanced material in: Fundamentals of Fuzzy Sets D. Dubois H. Prade Eds. Kluwer , Boston.

[4] J. Casabó, et al. Temes de càlcul numèric amb el full de càlcul. Convocatoria UPC 2003/2004 de ayudas para proyectos de mejora docente.

[5] J. Casabó et al. Automatització i programació dels fulls de càlcul, Convocatoria UPC 2005/2006 de ayudas para proyectos de mejora docente.

[6] K. Degtiarev, R. Nizamiev, "Program SFDT: An Educational Software Tool for Fuzzy System Model Design", International Conference on Computational Intelligence (ICCI), pp. 193 - 199, Nicosia (2004).

[7] E. Inelmen, E. Inelmen, A. Ibrahim. (2003) A new approach to teaching fuzzy logic system design in: LNAI 2715 T. Bilgiç et al. (Eds.) Springer-Verlag Berlin, pp 79-86.

[8] InrecoLAN Fuzzy Math for OpenOffice .org Calc in: http://www.openfuzzymath.org .

[9] J. Jacas, J. Recasens. Aggregation of T- Transitive relations. Int. Journal of Intelligent Systems, vol. 18, nº 12, pp. 1193-1214 (2003).

[10] B. López , La enseñanza de los conjuntos difusos. Facetas: epistémica, cognitiva e instruccional. Universitas 2000, vol. 29 nº 3-4 pp. 1-16 (2005).

[11] E. Morin. Introducción al pensamiento complejo. 1995 Barcelona, Gedisa.

[12] H. T. Nguyen, E. A. Walker. (2000) A first course in fuzzy Logic, Chapman & Hall/CRC BocaRaton.

[13] T. J. Ross. (1995) Fuzzy Logic with Engineering Applications, Mc Graw-Hill.

[14] Teaching Activities in Fuzzy Logic and Technologies in: http://www.eusflat.org/research/teaching .htm .

[15] E. Trillas, C. Alsina., J.M.Terricabras. (1995) Introducción a la Lógica Borrosa, Ariel Barcelona.

[16] J. Walkenbach. (2000) Programación en Excel 2000 con VBA. Ed. Anaya Multimedia.

Pirámides y Cópulas

Claudio Moraga

European Centre for Soft Computing
33600, Mieres, Asturias
claudio.moraga@softcomputing.es

Resumen

Se presenta una familia de funciones que en el cubo unitario tienen forma de pirámides. Se muestra que mediante manipulaciones elementales con geometría Euclidiana es posible calcular valores y analizar algunas propiedades de estas funciones. Desde el punto de vista de la lógica borrosa, estas funciones son todavía "extrañas", por cuanto pueden no ser ni asociativas ni conmutativas. La razón de esto es que son cópulas.

1 INTRODUCCIÓN

Las cópulas tienen su origen y su "habitat natural" en el mundo de las probabilidades [1], [2]. En el de la lógica borrosa son tratadas "con respeto", pero aun no han encontrado demanda a la hora de las aplicaciones. Esta situación puede cambiar: hay trabajos que abogan seriamente por un *diseño* de las reglas si–entonces de la lógica borrosa (ver, por ejemplo, [3]), proceso en el cual no solo hay que considerar que los conjuntos borrosos elegidos para representar formalmente las variables lingüísticas usadas en las premisas deben ser un reflejo fiel del uso de tales conceptos en el lenguaje, sino también que los conectivos utilizados lleven a conclusiones consistentes con las expectaciones del usuario o con los datos empíricos que estén disponibles. Otros trabajos han señalado situaciones en las cuales el conectivo "y" establece diferentes requisitos al operador que vaya a realizarlo [4] o que no tengan representación adecuada con una t-norma [5], cuando frecuentemente una t-norma suele ser considerada como la "opción obvia". Cabe destacar [6], trabajo en el cual se hace mención explícita a las cópulas y al papel que pueden ocupar en la lógica borrosa.

El resto del trabajo está organizado de la siguiente manera: La próxima sección abarca los formalismos necesarios, los cuales van a ser presentados de una manera didáctica, para hacerlos accesibles a quienes no han tenido experiencia previa con cópulas. Las secciones siguientes cubrirán las conclusiones, los agradecimientos y la bibliografía.

2 DEFINICIONES, TEOREMAS Y PROPIEDADES

Según la sabiduría popular, "*un cuadro dice mas que 1000 palabras*", razón por la cual se mostrará en la figura 1 primeramente el dibujo de una pirámide, representativa de las que se estudiarán en este trabajo, antes de presentar la definición formal.

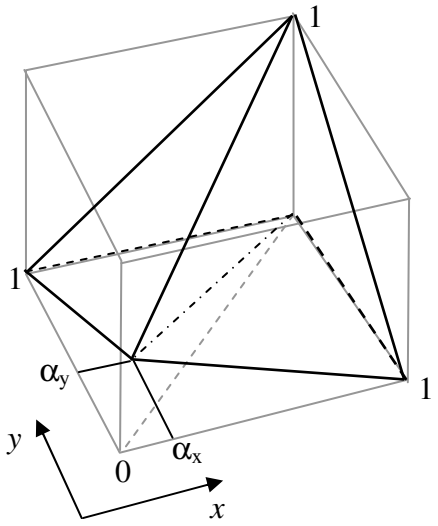


Figura 1: La función pirámide

En el resto del trabajo se utilizarán paréntesis angulares para señalar las coordenadas de un punto en el cuadrado unitario, para evitar confusión con intervalos abiertos.

Definición 1:

El parámetro característico de una pirámide es el punto $\langle \alpha_x , \alpha_y \rangle$, con $\alpha_x + \alpha_y < 1$ y $0 < \alpha_x \cdot \alpha_y$. Si $\alpha_x = \alpha_y$, la pirámide es simétrica y si $\alpha_x \neq \alpha_y$, es asimétrica. El triángulo cuyos vértices son $\langle 1 , 1 \rangle$, $\langle 0 , 1 \rangle$ y $\langle \alpha_x , \alpha_y \rangle$, se denominará proyección triangular superior (de la pirámide) y aquél cuyos vértices son $\langle 1 , 1 \rangle$, $\langle \alpha_x , \alpha_y \rangle$ y $\langle 1 , 0 \rangle$, proyección triangular inferior.

Para calcular la altura de la pirámide en un punto interior dado, se puede utilizar el siguiente procedimiento geométrico, ilustrado en la figura 2, suponiendo que el punto se encuentra en la proyección triangular inferior.

1. Determinar la posición del punto $\langle x, y \rangle$

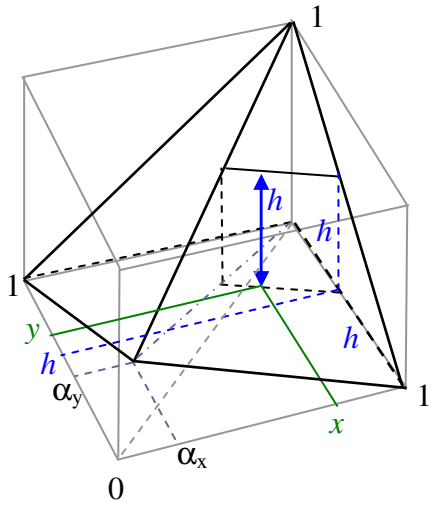


Figura 2: Obtención del valor de la función en un punto dado

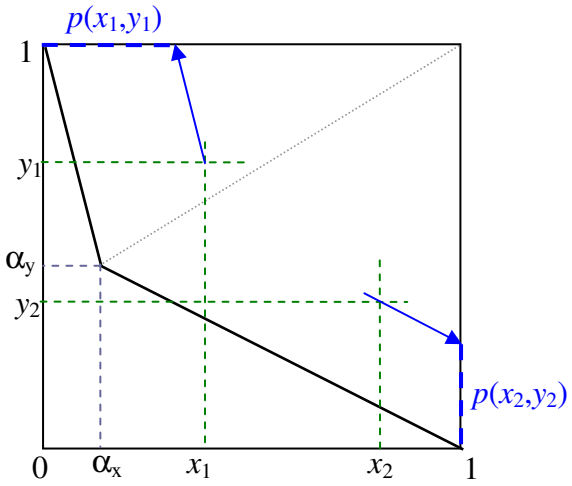


Figura 3: método simplificado para calcular el valor de la función en un punto dado

2. Construir un plano vertical que pase por el punto $\langle x, y\rangle$ y su orientación sea paralela a la arista frontal de la proyección triangular inferior. Este plano intercepta la cara frontal de la pirámide a la altura h, que es la altura de la pirámide en el punto $\langle x, y\rangle$
3. El plano construido en el punto anterior intercepta la cara lateral de la pirámide. En la intersección, la altura es obviamente h. Recordando que –(por construcción)– la cara lateral de la pirámide es un triángulo isósceles, la distancia desde la intersección del plano con la cara lateral de la pirámide al punto $\langle 1, 0\rangle$ a lo largo de la línea $x = 1$, es también h.
4. Proceder de manera análoga si el punto se encuentra en la proyección triangular superior.

Dado el análisis anterior, resulta evidente que si un punto está fuera de la pirámide (pero dentro del cuadrado unitario) tiene altura 0. Además, es posible calcular la altura de la pirámide en puntos dentro de su base, solamente utilizando la proyección bidimensional, como se muestra en la figura 3, donde la función se denomina p (tomado de *p*irámide). La relación entre las figuras 2 y 3 es que el valor de h en un punto $\langle x_i, y_j\rangle$ de la pirámide es igual a $p(x_i, y_j)$.

Sean $\beta_x = \alpha_y/(1-\alpha_x)$ y $\beta_y = \alpha_x/(1-\alpha_y)$. (1)

La ecuación de la línea que pasa por los puntos $\langle 1, 0\rangle$ y $\langle \alpha_x, \alpha_y\rangle$ es

$$y = \beta_x (1-x) \quad (2)$$

De modo similar, la ecuación de la línea a través de los puntos $\langle 0, 1\rangle$ y $\langle \alpha_x, \alpha_y\rangle$ es

$$x = \beta_y (1-y) \quad (3)$$

o bien

$$y = 1 - (x/\beta_y) \quad (4)$$

Tomando en cuenta la ecuación (2), la ecuación de la recta paralela por el punto $\langle x_2, y_2\rangle$ es

$$y_2 = \beta_x (1-x_2) + p(x_2, y_2) \quad (5)$$

de donde

$$p(x_2, y_2) = y_2 - \beta_x (1-x_2) \quad (6)$$

De la misma manera, tomando en cuenta la ecuación (3), la ecuación de la recta paralela por el punto $\langle x_1, y_1\rangle$ es

$$x_1 = \beta_y (1-y_1) + p(x_1, y_1) \quad (7)$$

de donde

$$p(x_1, y_1) = x_1 - \beta_y (1-y_1) \quad (8)$$

Definición 2:

Una función $p : [0,1] \times [0,1] \rightarrow [0,1]$ que satisface las condiciones siguientes es una *pirámide asimétrica*:

i) $p(1, y) = y$; $p(x, 1) = x$
(*i.e.* 1 es el elemento neutral o identidad)

ii) Con $\alpha_x \neq \alpha_y$ y β_x, β_y tal como en (1),
si $x \leq \beta_y(1-y)$ o si $y \leq \beta_x(1-x)$ -(vale decir, si $\langle x, y\rangle$ está fuera de las regiones de proyección triangular)- entonces $p(x, y) = 0$
(Nótese el caso particular $p(0, y) = p(x, 0) = 0$
i.e. 0 es un atractor o elemento dominante)

iii) Si $\langle x, y\rangle$ está en la proyección triangular inferior,
entonces $p(x, y) = y - \beta_x (1-x)$

iv) Si $\langle x, y\rangle$ está en la proyección triangular superior,
entonces $p(x, y) = x - \beta_y(1-y)$

Los puntos iii) y iv) especifican que la pirámide asimétrica no es conmutativa.

Permítase en la definición anterior que $\alpha_x = \alpha_y$ con lo cual $\beta_x = \beta_y$. Entonces la definición corresponde a una pirámide simétrica, la cual es conmutativa, porque

de iii) se tiene $p(x, y) = y - \beta_x(1-x)$
y de iv) se tiene $p(x, y) = x - \beta_x(1-y)$
de lo cual se sigue que $p(x, y) = p(y, x)$

Lemma 1:
Una pirámide asimétrica no es asociativa.

Demostración:
Supóngase que una pirámide asimétrica fuera asociativa. Se presentará un contraejemplo para obtener una prueba por contradicción. Para esto se utilizará un análisis geométrico ilustrado en la figura 4. Resulta obvio que $p(p(u,v), z) > 0$, mientras que $p(u, p(v,z)) = 0$. Como esto contradice la suposición inicial, se concluye que una pirámide asimétrica no es asociativa.

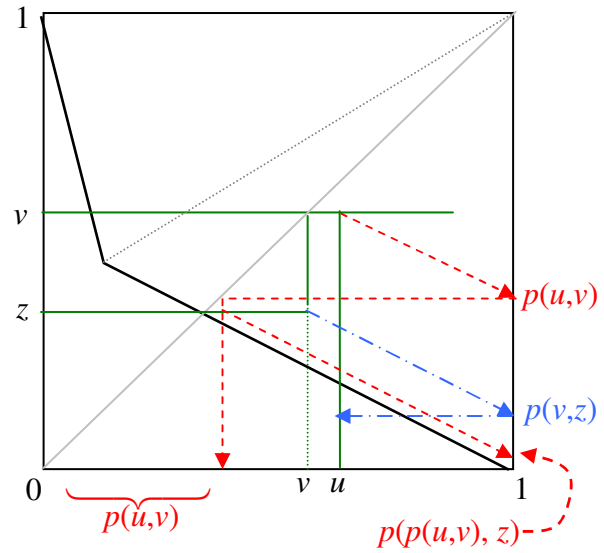


Figura 4: Contraejemplo a una supuesta asociatividad de una pirámide asimétrica

Análisis:
Para calcular $p(p(u,v), z)$, primero se determina el punto $\langle u, v\rangle$ y se dibuja una paralela al borde inferior de la pirámide por dicho punto, la cual intersecta la línea $x = 1$ dando $p(u,v)$. Usando como espejo la línea $x = y$ (en el cuadrado unitario), $p(u,v)$ se convierte en abscisa para poder posicionar el punto $\langle p(u,v), z\rangle$. Finalmente usando el mismo procedimiento antes señalado, se obtiene $p(p(u,v), z) > 0$.

Para calcular $p(u, p(v, z))$, primero hay que localizar el punto $\langle v, z\rangle$. Como v tiene que ser abscisa, nuevamente se usa la diagonal del cuadrado unitario para realizar la transformación requerida y con la ordenada z encontrar el punto deseado. Una paralela al borde inferior de la pirámide por dicho punto, intersecta la línea $x = 1$ dando $p(v, z)$. El punto $\langle u, p(v, z)\rangle$ se obtiene con la intersección de la abscisa u y de la ordenada $p(v, z)$. Es claramente visible que este punto cae fuera de la pirámide; por lo tanto $p(u, p(v, z)) = 0$.

Se sigue que $p(p(u,v), z) \neq p(u, p(v,z))$ *i.e.* p no es asociativa.

Corolario 1.1:
Una pirámide simétrica tampoco es asociativa.

Demostración:
El contraejemplo utilizado en el caso asimétrico también es válido para éste, por cuanto todos los puntos utilizados en los cálculos yacen al mismo lado de la diagonal del cuadrado unitario. Sin embargo se realizará una demostración analítica la cual permitirá realzar un aspecto importante de esta pirámide.

Sean $\alpha_x = \alpha_y = \alpha$ y $\beta = \alpha/(1-\alpha)$. Además, sean u, v y z tal como en el caso anterior.

Según el punto iii) de la definición 3 (para el caso simétrico),

$$p(u, v) = v - \beta(1-u)$$

entonces:

$$p(p(u, v), z) = z - \beta(1 - p(u, v)) =$$

$$= z - \beta(1 - (v - \beta(1-u))) = z - \beta(1-v) - \beta^2(1-u)$$

y $$p(u, p(v, z)) = p(v, z) - \beta(1-u) =$$

$$= z - \beta(1-v) - \beta(1-u)$$

Por lo tanto

$$p(p(u, v), z) \neq p(u, p(v, z))$$

Nótese que desde el punto de vista numérico, la asociatividad sería válida si $\beta = \beta^2$; pero esto solo es posible si $\beta = 0$ o 1 lo cual requiere que $\alpha = 0$ o $½$, valores que están justo al borde, pero fuera del dominio de α. Si α tomara el valor 0, entonces la pirámide sería igual a la t-norma mínimo y is α tomara el valor ½ entonces la pirámide sería igual a la t-norma W de Łukasiewicz, (como puede deducirse de la figura 1). Esto lleva a la siguiente *conjetura*:

Existe una familia infinita de pirámides entre la t-norma minimum y la t-norma W de Łukasiewicz. Si p es una pirámide, entonces

$$W < p < mínimo \qquad (9)$$

Antes de abordar el Lemma principal, se necesitan algunas especificaciones formales adicionales.

La ecuación de la diagonal de la (base de la) pirámide, *i.e.* la línea que une los puntos $\langle 1, 1\rangle$ y $\langle \alpha_x, \alpha_y\rangle$ está dada por

$$\frac{1-\alpha_y}{1-\alpha_x} = \frac{y-\alpha_y}{x-\alpha_x}$$

$$y = \left(\frac{1-\alpha_y}{1-\alpha_x}\right)x - \frac{\alpha_x - \alpha_y}{1-\alpha_x} \qquad (10)$$

pero como

$$\frac{\alpha_x - \alpha_y}{1-\alpha_x} = \frac{-1+\alpha_x+1-\alpha_y}{1-\alpha_x} =$$
$$= \frac{-(1-\alpha_x)+1-\alpha_y}{1-\alpha_x} = -1+\frac{1-\alpha_y}{1-\alpha_x}$$

entonces la ecuación (10) puede escribirse como sigue:

$$y = \left(\frac{1-\alpha_y}{1-\alpha_x}\right)(x-1)+1 \qquad (11)$$

Nótese además que

$$\frac{\beta_x - 1}{\beta_y - 1} = \frac{\frac{\alpha_y}{1-\alpha_x}-1}{\frac{\alpha_x}{1-\alpha_y}-1} = \frac{\frac{\alpha_y-1+\alpha_x}{1-\alpha_x}}{\frac{\alpha_x-1+\alpha_y}{1-\alpha_y}} = \frac{1-\alpha_y}{1-\alpha_x} \qquad (12)$$

Reemplazando (12) en (11) se obtiene

$$y = \left(\frac{\beta_x - 1}{\beta_y - 1}\right)(x-1)+1 \qquad (13)$$

Definición 3: [1], [2]

Una cópula bidimensional es una función $C : [0,1] \times [0,1] \rightarrow [0,1]$ la cual satisface las condiciones de contorno $C(0,0) = 0$ y $C(1,1) = 1$, a la vez que la condición de monotonía

$$C(x_1, y_1) - C(x_2, y_1) - C(x_1, y_2) + C(x_2, y_2) \geq 0 \qquad (14)$$

toda vez que $x_1 \leq x_2, y_1 \leq y_2$.

Lemma 2:

Una pirámide es una cópula.

Demostración:

i) Según los puntos i) y ii) de la definición 2,

$$p(0,0) = 0 \quad \text{y} \quad p(1,1) = 1$$

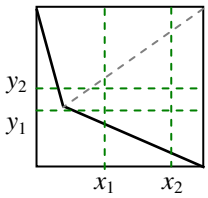
Por lo tanto p satisface las condiciones de contorno especificadas en la definición 3.

ii) Para el análisis de la inecuación de monotonía (14) se considerarán cuatro casos representativos con $x_1 \leq x_2$ e $y_1 \leq y_2$

ii.1)

Sea $p(x_i, y_j) > 0$, $i,j \in \{1, 2\}$ con los cuatro puntos dentro de una misma proyección triangular. Sin pérdida de generalidad se eligirá la región inferior.

$$p(x_1, y_1) = y_1 - \beta_x(1 - x_1)$$
$$p(x_2, y_1) = y_1 - \beta_x(1 - x_2)$$
$$p(x_1, y_2) = y_2 - \beta_x(1 - x_1)$$
$$p(x_2, y_2) = y_2 - \beta_x(1 - x_2)$$

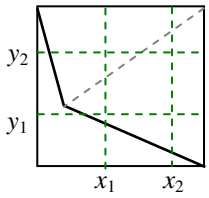


$$p(x_1, y_1) - p(x_2, y_1) - p(x_1, y_2) + p(x_2, y_2) =$$
$$= y_1 - \beta(1 - x_1) - [\, y_1 - \beta(1 - x_2)] -$$
$$- [\, y_2 - \beta(1 - x_1)] + [y_2 - \beta(1 - x_2)] = 0 \qquad (15)$$

ii.2)

Sean $x_2 > y_2 > x_1 > y_1$ tal que todos los puntos estén dentro de las proyecciones triangulares: $\langle x_1, y_2\rangle$ en la superior y los demás en la inferior.

$$p(x_1, y_1) = y_1 - \beta_x(1 - x_1)$$
$$p(x_2, y_1) = y_1 - \beta_x(1 - x_2)$$
$$p(x_1, y_2) = x_1 - \beta_y(1 - y_2)$$
$$p(x_2, y_2) = y_2 - \beta_x(1 - x_2)$$



$$p(x_1, y_1) - p(x_2, y_1) - p(x_1, y_2) + p(x_2, y_2) =$$
$$= y_1 - \beta_x(1 - x_1) - [\, y_1 - \beta_x(1 - x_2)] -$$
$$- [x_1 - \beta_y(1 - y_2)] + [y_2 - \beta_x(1 - x_2)] =$$
$$= y_1 - \beta_x(1 - x_1) - y_1 + \beta_x(1 - x_2) - x_1 + \beta_y(1 - y_2)$$
$$+ y_2 - \beta_x(1 - x_2)$$
$$= - \beta_x(1 - x_1) - x_1 + \beta_y(1 - y_2) + y_2 =$$
$$= - (\beta_x - \beta_y) + x_1(\beta_x - 1) - y_2(\beta_y - 1) = \Gamma \qquad (16)$$

Dado que $\langle x_1, y_2\rangle$ está sobre la diagonal y tomando en cuenta la ecuación (14), se tiene:

$$y_2 > \left(\frac{\beta_x - 1}{\beta_y - 1}\right)(x_1 - 1)+1,$$

lo cual se puede expresar como

$$y_2 = \left(\frac{\beta_x - 1}{\beta_y - 1}\right)(x_1 - 1)+1+\delta \quad \text{con} \quad 0 < \delta \,^{1} \qquad (17)$$

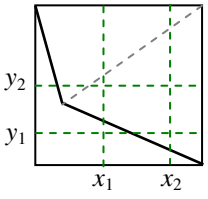
Introduciendo (17) en (16) se obtiene:

$$\Gamma = - (\beta_x - \beta_y) + x_1(\beta_x - 1) - (\beta_x - 1)(x_1 - 1) -$$
$$- (\beta_y - 1)(1 + \delta) =$$
$$= - \beta_x + \beta_y + x_1(\beta_x - 1) - x_1(\beta_x - 1) + \beta_x - 1 +$$
$$+ (1 - \beta_y)(1 + \delta) =$$
$$= (1 - \beta_y)\,\delta > 0 \qquad (18)$$

ii.3)

Sean $x_2 > y_2 > x_1 > y_1$ tal que $\langle x_1, y_1\rangle$ esté fuera de la proyección basal de la pirámide y los restantes puntos en la proyección triangular inferior.

$$p(x_1, y_1) = 0$$
$$p(x_2, y_1) = y_1 - \beta_x(1 - x_2)$$
$$p(x_1, y_2) = y_2 - \beta_x(1 - x_1)$$
$$p(x_2, y_2) = y_2 - \beta_x(1 - x_2)$$



[1] δ también tiene una cota superior para preservar que $1 \geq y_2$, pero no es necesaria para la demostración que se está realizando.

$$p(x_1, y_1) - p(x_2, y_1) - p(x_1, y_2) + p(x_2, y_2) =$$

$$= 0 - [\, y_1 - \beta_x(1 - x_2)] - [\, y_2 - \beta_x(1 - x_1)] + \\ + [y_2 - \beta_x(1 - x_2)] = - y_1 + \beta_x(1 - x_1)$$

Si se calcula la intersección de la arista inferior de la pirámide (ecuación (2)) con la vertical $x = x_1$ se obtiene la ordenada $y' = \beta_x(1 - x_1)$, la cual obviamente es mayor que y_1. Insertando esto en la expresión anterior, se obtiene:

$$- y_1 + \beta_x(1 - x_1) = - y_1 + y' > 0 \qquad (19)$$

ii.4)

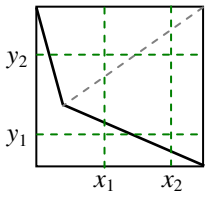
Considérese $\langle x_1, y_1 \rangle$ fuera de la proyección basal; $\langle x_1, y_2 \rangle$ en la proyección triangular superior y los restantes dos puntos en la inferior

$$p(x_1, y_1) = 0$$

$$p(x_2, y_1) = y_1 - \beta_x(1 - x_2)$$

$$p(x_1, y_2) = x_1 - \beta_y(1 - y_2)$$

$$p(x_2, y_2) = y_2 - \beta_x(1 - x_2)$$



$$p(x_1, y_1) - p(x_2, y_1) - p(x_1, y_2) + p(x_2, y_2) =$$

$$= 0 - [\, y_1 - \beta_x(1 - x_2)] - [x_1 - \beta_y(1 - y_2)] + \\ + [y_2 - \beta_x(1 - x_2)] =$$

$$= - y_1 + \beta_x(1 - x_2) - x_1 + \beta_y(1 - y_2) + y_2 \\ - \beta_x(1 - x_2) =$$

$$= - y_1 - x_1 + \beta_y(1 - y_2) + y_2 = - y_1 - x_1 + \beta_y + \\ + y_2(1 - \beta_y)$$

Tal como en el caso ii.2 se tomará

$$y_2 = \left(\frac{\beta_x - 1}{\beta_y - 1} \right)(x_1 - 1) + 1 + \delta \ \text{ con } 0 < \delta \qquad (20)$$

con lo cual la expresión anterior se convierte en:

$$- y_1 - x_1 + \beta_y + (1 - \beta_x)(\, x_1 - 1) + (1 - \beta_y)(1 + \delta) =$$

$$= - y_1 - x_1 + \beta_y + (1 - \beta_y)(1 + \delta) + (\, x_1 - 1) - \\ - \beta_x(\, x_1 - 1)$$

$$= - y_1 + \beta_x(1 - x_1) - x_1 + \beta_y + (1 - \beta_y)(1 + \delta) \\ + (\, x_1 - 1)$$

$$= [- y_1 + \beta_x(1 - x_1)] - 1 + \beta_y + (1 - \beta_y) + \delta(1 - \beta_y)$$

$$= [- y_1 + \beta_x(1 - x_1)] + \delta(1 - \beta_y)$$

y utilizando la ecuación (19) finalmente se obtiene

$$[- y_1 + \beta_x(1 - x_1)] + \delta(1 - \beta_y) = [y' - y_1] + \delta(1 - \beta_y) > 0$$

En todos los casos se cumple la condición de monotonía requerida en la definición 1, por lo tanto p es una cópula.

Lemma 3: [2]

Si C es una cópula, entonces cuando $x_1 \le x_2,\ y_1 \le y_2$

$$0 \le C(x_2, y_2) - C(x_1, y_1) \le x_2 - x_1 + y_2 - y_1$$

Por lo tanto, toda cópula es no-decreciente en cada argumento y tiene la continuidad de Lipschitz. Además cualquier cópula C satisface el entorno

$$W \le C \le min \qquad (21)$$

Resulta evidente que este Lemma confirma la conjetura planteada en la ecuación (9). Los márgenes estrictos especificados en (9) fueron elegidos así explícitamente con restricciones en $\langle \alpha_x , \alpha_y \rangle$ – (recordar la definición 1)– con el fin de destacar las cópulas no asociativas (y mayormente no conmutativas).

3 COMENTARIO RETROSPECTIVO

Los lectores familiarizados con el mundo de las cópulas habrán reconocido en este trabajo el ejercicio 3.8 propuesto en la página 65 de [1] y en las ecuaciones (9) y (21), las cotas de Fréchet-Hoeffding.

4 AGRADECIMIENTOS

Al Profesor Enric Trillas, por haberme guiado con mucha paciencia por los caminos de la Lógica Borrosa y haberme brindado su amistad.

BIBLIOGRAFÍA

[1] Nelsen R.: *An introduction to Cópulas.* (2nd. Edition), Springer, Heidelberg, 2006

[2] Alsina C., Frank M.J., Schweizer B.: *Associative functions on intervals: A primer on triangular norms.* World Scientific, Hackensack, 2005

[3] Trillas E.: On some features of fuzzy logic. *Proc. 27th Annual Meeting of the North American Fuzzy Information Processing Society NAFIPS.* 2008

[4] Guadarrama S.: Contribución al estudio de la computación con palabras/percepciones. Tesis Doctoral, Universidad Politécnica de Madrid, 2007

[5] Moraga C., Temme K.-H.: Functional equivalence between S-neural networks and fuzzy models. In: *Technologies for Constructing Intelligent Systems 2: Tools.* (B. Bouchon-Meunier, J. Gutiérrez Ríos, L. Magdalena and R.R. Yager, Eds.), 355-363, Physica-Verlag, Heidelberg, 2002

[6] Trillas E., Alsina C., Pradera A.: On a class of fuzzy sets theories. *Proc. FuzzIEEE-2007*, 120 – 124. IEEE-CS-Press, 2007

COMBINACIÓN DE CREENCIAS Y LÓGICA BORROSA

Julián Velarde Lombraña
Departamento de Filosofía, Universidad de Oviedo, velarde@uniovi.es

Resumen

La combinación de creencias y de evidencia puede ser expresada y formalizada dentro de la teoría de la evidencia de Dempster-Shafer. En este artículo, a partir de la crítica de Zadeh a la regla de combinación de Dempster, describimos y comparamos los métodos propuestos en la teoría de la posibilidad para combinar creencias tomando en cuenta la información acerca de la fiabilidad de las fuentes y del grado de conflicto entre ellas.

Palabras Clave: Combinación de creencias, evidencia, regla de Dempster, distribución de posibilidad.

1. INTRODUCCIÓN

El cuerpo de *creencias* de un sujeto en interacción con el mundo cambiante y dinámico está destinado a evolucionar ante la *información* que le llega al sujeto con diferentes grados de *certeza, fiabilidad, evidencia*; nociones todas ellas ligadas en el ámbito de la filosofía del conocimiento o epistemología. Y la adquisición de nuevas evidencias puede corroborar, aumentar o entrar en conflicto con anteriores creencias, obligando a revisar (reestructurar) el cuerpo de creencias. Varios son los modelos lógico-matemáticos que se han desarrollado para la formalización y tratamiento de las creencias en sus múltiples aspectos. En tanto que las creencias están relacionadas con la evidencia, la *teoría de la evidencia* de Dempster - Shafer ha alcanzado gran predicamento. El modelo formal allí propuesto busca cómo computar las funciones de creencia simples y la combinación de funciones, lo que permite cuantificar la amplitud de evidencia que conlleva un mensaje (información) complejo y no susceptible de descomposición en componentes independientes. Sin embargo, el problema central con el que se encuentra esta teoría —al igual que la teoría bayesiana y todas sus alternativas— es la combinación de evidencias y creencias en conflicto, junto con sus aplicaciones. Por otro lado, la *teoría de la posibilidad* desarrollada por Zadeh proporciona un modelo formal para la cuantificación de la posibilidad subjetiva o de un estado de creencia y para el razonamiento aproximado en el ámbito de la incertidumbre. El problema principal con el que se encuentra esta teoría es el de la normalización de las distribuciones de posibilidad. A partir de la crítica de Zadeh a la *regla de combinación* de evidencia de Dempster – Shafer, se analizan los problemas que presentan ambas teorías y las propuestas de solución que se vienen ofreciendo.

2. TEORÍA DE LA EVIDENCIA

En los años 60 Dempster puso las bases de esta teoría, y luego fue desarrollada por Shafer, para tratar cierto tipo de incertidumbre (la llamada "incertidumbre epistémica", de distinta naturaleza que la "incertidumbre aleatoria", tratada por la teoría clásica de la probabilidad. En la incertidumbre aleatoria el grado numérico (comprendido entre 0 y 1) expresa la probabilidad frecuencialista u objetiva de un suceso; el grado numérico de creencia (igualmente comprendido entre 0 y 1) representa el nivel de credibilidad que se concede a la realización de un suceso o a una proposición, apoyándose en la evidencia resultante de la información de la que se dispone. Así pues, en tanto que los grados de probabilidad se asignan a sucesos en un proceso "objetivo" azaroso, los grados de credibilidad concedidos a una hipótesis (o proposición) son resultado de un acto de juicio subjetivo.
Las tres funciones básicas utilizadas en la teoría de la evidencia son: (1) las *funciones de masa* (en ingles: "basic belief mass functions" o "basic probability assignment functions", en terminología de Shafer). (2) Las *funciones de credibilidad* (en inglés: "belief functions"). (3) Las *funciones de plausibilidad* (en inglés: "plausibility functions").

Las funciones de masa

La teoría de Dempster–Shafer está normalmente basada sobre un conjunto Θ no vacío, exclusivo, exhaustivo y supuestamente finito de hipótesis, estados o acontecimientos:

$$\Theta = \{H_1, H_2, \ldots, H_n\}$$

A partir de Θ se obtiene el conjunto 2^Θ, llamado *marco de discernimiento*, que comprende las 2^n proposiciones (estados, hipótesis) A de Θ :

$$2^\Theta = \{A / A \subseteq \Theta\} = \{\varnothing, \{H_1\}, \{H_2\}, \ldots, \{H_1, H_2\}, \ldots, \Theta\}$$

A cada una de estas proposiciones A de 2^Θ se asigna una distribución de masa de evidencia *m(A)*, asociada a la creencia y la plausibilidad. La función de masa *m(A)* representa la confianza prestada estrictamente a *A*, sin que pueda ser repartida entre las hipótesis que componen *A* (los X ⊂ A). Las masas representan los grados numéricos de creencia asignados a los subconjuntos de 2^Θ. La cantidad *m(A)* es llamada *masa de A*, y representa en qué medida la evidencia soporta la hipótesis A.

Dado un marco de discernimiento 2^Θ, una función $m : 2^\Theta \rightarrow [0, 1]$ es llamada una *función de masa*, si y sólo si,

1. $m(\varnothing) = 0$
2. $\sum_{A \subseteq \Theta} m(A) = 1$

Los subconjuntos A ⊆ Θ tales que $m(A) > 0$ son llamados *elementos focales*, y el par $\langle F , m \rangle$, en donde *F* denota el conjunto de todos los elementos focales inducidos por *m* es llamado *cuerpo de evidencia.*

A partir de las funciones de masa son definidas las funciones de *credibilidad* y de *plausibilidad*. La credibilidad de un conjunto de elementos A en un marco de discernimiento 2^Θ es definida como la credibilidad total de A; es decir, como la evidencia total que soporta A, lo cual representa la suma de todas las masas asignadas a los elementos contenidos en A más la masa atribuida a la propia A, obteniendo así la función:

$$Cred(A) = \sum_{X \subseteq A} m(X)$$

La función *Cred*: $2^\Theta \rightarrow [0, 1]$, caracterizada por la función de masa $m : 2^\Theta \rightarrow [0, 1]$, es llamada *función de credibilidad* sobre 2^Θ; y *Cred*(A) es llamada *grado de credibilidad* de *A*.

La plausibilidad de un conjunto de elementos *A*, en un marco de discernimiento 2^Θ es definida como la credibilidad maximal posible que podría ser atribuida a *A* en caso de que toda evidencia futura apoyase a *A*; esto es, mide toda la evidencia que no contradice a *A*; en términos formales:

$$Pl(A) = \sum_{X \cap A \neq \varnothing} m(X)$$

Una función $Pl : 2^\Theta \rightarrow [0, 1]$, caracterizada por la función de masa $m : 2^\Theta \rightarrow [0, 1]$, es llamada *función de plausibilidad* sobre 2^Θ; y *Pl*(A) es llamada *grado de plausibilidad* de *A*. Y en virtud de las relaciones de *Cred* con *m*, la fórmula anterior deviene:

$$Pl(A) = 1 - Cred(A')$$

3. REGLA DE COMBIANCIÓN DE FUNCIONES DE CREDIBILIDAD (REGLA DE DEMPSTER)

La regla de combinación de Dempster proporciona un método para cambiar las opiniones (creencias) anteriores a la luz de nueva evidencia. Se trata de una regla que permite calcular, a partir de dos funciones de credibilidad sobre el mismo marco de discernimiento 2^Θ, una nueva función de credibilidad ⊕, llamada su *suma ortogonal*, basada sobre la evidencia combinada. El operador ⊕ cumple las propiedades conmutativa y asociativa.

Supongamos que m_1 sea una función de masa sobre un marco de discernimiento 2^Θ con A_i elementos focales; y que m_2 sea otra función de masa sobre el mismo marco de discernimiento con B_j elementos focales; entonces de la combinación de las masas de evidencia m_1 y m_2 obtenemos nueva evidencia sobre los elementos de 2^Θ. Esto es, la nueva evidencia que apoya a C es la suma de la evidencia de todas las intersecciones entre los conjuntos A_i y B_j que den como resultado C. Así:

$$[m_1 \oplus m_2](C) = \sum_{A_i \cap B_j = C} m_1(A_i) \cdot m_2(B_j)$$

El problema de este esquema es que puede atribuir una masa de evidencia no nula al conjunto vacío, lo que contradice las reglas de base para las funciones de credibilidad; en efecto, dado que A_i y B_j son elementos focales: $m_1(A_i) > 0$ y $m_2(B_j) > 0$, su producto resulta ser no nulo; y así mismo cuando $A_i \cap B_j = \varnothing$, esto es, cuando su realización simultánea es imposible. La masa $m(\varnothing)$ asociada al conjunto vacío refleja el conflicto entre las dos fuentes de evidencia. Los conflictos surgen normalmente al combinar funciones de masa que representan (parcialmente) evidencia contradictoria. Para resolver este problema es necesario descartar los productos $m_1(A_i) \cdot m_2(B_j)$ tales que $A_i \cap B_j = \varnothing$, redistribuyendo $m(\varnothing) = K$ entre los restantes productos $m_1(A_i) \cdot m_2(B_j)$ del marco de discernimiento tales que $A_i \cap B_j = \varnothing$; esto es, cada uno de los otros valores-masa es dividido por $1 - K$ (de manera equivalente: multiplicado por $(1 - K)^{-1}$), siendo

$$K = \sum_{A_i \cap B_j = \varnothing} m_1(A_i) \cdot m_2(B_j) > 0$$

Se trata de inflar la masa de evidencia de los demás subconjuntos en forma proporcional a la evidencia asignada al conjunto vacío (para que, según las reglas

de base, su suma sea igual a 1); tales subconjuntos quedan así *normalizados*. La regla de combinación de Dempster deviene con la normalización:

$$[m_1 \oplus m_2](C) = \sum_{Ai \cap Bj=C} \frac{m_1(A_i) \cdot m_2(B_j)}{1-K}$$

El coeficiente *K* traduce el conflicto existente entre las dos fuentes m_1 y m_2 de evidencia. Cuando K = 1, las fuentes están en conflicto total, y no pueden ser combinadas. Y a la inversa, cuando K = 0, las fuentes están en acuerdo perfecto. El elemento neutro del operador es la masa de la ignorancia total: $m(\Theta) = 1$; y los elementos absorbentes son las funciones que afectan a la totalidad de la masa de evidencia sobre una de las hipótesis elementales: $m(\{H_i\}) = 1$.

Ejemplo: dado Θ = {a , b , c}; el marco de discernimiento: 2^{Θ} = { ∅ ,{a}, {b},{c},{a , b},{a , c},{b , c},{a , b ,c}, {Θ}}; y dos fuentes de evidencia (dos funciones de masa), m_1 y m_2, sobre el mismo marco de discernimiento 2^{Θ}, con valores y sus productos $m_1(A_i) \cdot m_2(B_j)$ tal como se refleja en la tabla 1:

Tabla 1: combinación de m_1 y m_2

	m_1{ a } = 0,40	m_1{ a, b }= 0,20	m_1{ c } = 0,40
m_2{ a } = 0,7	0,28	0,14	0,28
m_2{ Θ} = 0,3	0,12	0,06	0,12

Y ahora normalizando, esto es, dividiendo cada uno de los productos $m_1(A_i) \cdot m_2(B_j)$ tales que $A_i \cap B_j \neq \emptyset$ por 1 – K, esto es, por 1 – 0,28 = 0,72, tenemos:

$$m_1(A_i) \oplus m_2(B_j)\{a\} = \frac{0,28+0,14+0,12}{0,72} = 0,75$$

$$m_1(A_i) \oplus m_2(B_j)\{a , b\} = \frac{0,06}{0,72} = 0,08$$

$$m_1(A_i) \oplus m_2(B_j)\{c\} = \frac{0,12}{0,72} = 0,16$$

Características importantes de esta función de combinación son:

(1) Ítems de evidencia concordantes se refuerzan entre sí. Así, la combinación de masas de evidencia en {a} es 0,54 antes de la normalización; y 0,75 después de la normalización, cuando todas las masas de evidencia son divididas por 1 – K.

(2) Ítems de evidencia conflictivos se destruyen entre sí. Y cuando la masa en el conjunto nulo es muy amplia (hay amplio conflicto entre las fuentes de evidencia), la reasignación puede causar problemas.

(3) El uso de esta regla sólo es posible si m_1 y m_2 no están totalmente en conflicto, i. e., si existen al menos dos elementos focales A y B de m_1 y m_2 tales que $A \cap B \neq 0$.

Problemas de la normalización

El factor de normalización, por una parte, produce convergencia hacia la opinión dominante; y por otra parte tiene el efecto de ignorar completamente el conflicto, atribuyendo cualquier masa de evidencia en conflicto al conjunto nulo. Zadeh [9], en su reseña del libro de Shafer, *A Mathematical Theory of Evidence*, aduce un ejemplo en el que la regla de combinación de Dempster produce resultados considerados usualmente insatisfactorios o contra-intuitivos. El ejemplo de Zadeh en el marco de la teoría de la evidencia es el siguiente:

Sea Θ = {M , C , T} con las posibles enfermedades de un paciente (M = *Meningitis*; C = *Contusión cerebral*; T = *Tumor cerebral*); con dos masas de evidencia (dos opiniones) de dos expertos (médicos), m_1 y m_2. El primer médico cree, en grado 0,99, que el paciente tiene meningitis, o bien un tumor cerebral (en grado 0,01). El segundo médico cree, en grado 0,99, que el paciente tiene una contusión cerebral, o bien un tumor cerebral, en grado 0,01. Aplicando la regla de combinación de Dempster, tal como aparece en la tabla 2.

Tabla 2: Combinación de m_1 y m_2

	m_1{M} = 0,99	m_1{T} = 0,01
m_2{C} = 0,99	0,9801	0,0099
m_2{T} = 0,01	0,0099	0,0001

Por tanto:

$$m_1 \oplus m_2(A) \begin{cases} = 0,0001 \ldots\ldots para\ A = \{T\} \\ = 0,9999 \ldots\ldots para\ A = \emptyset \\ = 0 \ldots\ldots para\ todo\ otro\ A \subset \Theta \end{cases}$$

Y normalizando:

$$m_1 \oplus m_2\ (\mathrm{A}) \begin{cases} -1 \ldots\ldots para\ A = \{T\} \\ = 0 \ldots\ldots para\ A = \{M\} \\ = 0 \ldots\ldots para\ A = \{C\} \end{cases}$$

En el ejemplo cabe resaltar: (1) que las fuentes de evidencia m_1 y m_2 son altamente conflictivas: al ser $m_1\{C\} = 0$ y $m_2\{M\} = 0$, quedan excluidas completamente M y C como posibles respuestas; (2) que la conclusión parece ser contra-intuitiva, ya que la enfermedad T, que ha sido casi completamente descartada por cada uno de los dos médicos, resulta ser, tras la aplicación de la regla, la única opción verdadera.

Métodos de combinación alternativos

Muchos autores han utilizado el ejemplo de Zadeh, bien para criticar la teoría de Dempster-Shafer en su conjunto, bien para proponer reglas de combinación alternativas que eviten los resultados contra-intuitivos producidos por la regla de Dempster en caso de evidencia conflictiva. Las propuestas alternativas pueden agruparse en dos familias: en la primera se reagrupan las reglas de combinación que reposan sobre el postulado de la *fiabilidad* de las fuentes a combinar; por ejemplo, la combinación propuesta por Smets [7]. En la segunda, aquellas en las que una (aunque ignorando cuál) de las fuentes es fiable; aquí se ubican las propuestas de Yager [8] y de Dubois y Prade [2].

La combinación propuesta por Smets concede masa al conjunto nulo, eliminando así el factor de normalización: Así, siendo $m_1 \oplus m_2 = m_3$:

$$m_3\ (\mathrm{A}) \begin{cases} = m_3(A) \ldots\ldots \forall A \subseteq \Theta \\ - m_3(\emptyset) - K \end{cases}$$

En la regla propuesta por Yager se supone que una de las fuentes de evidencia que intervienen en la combinación es fiable. La solución, por tanto, está en el referencial; pero al no saber qué fuente proporciona la verdadera solución, Yager propone atribuir la masa conflictiva al conjunto Θ. Así:

$$m_3\ (\mathrm{A}) \begin{cases} = m_3(A) \ldots\ldots \forall A \subset \Theta; A \neq \Theta; A \neq \emptyset \\ = m_3(\Theta) + m_3(\emptyset) \ldots\ldots\ldots si\ A = \Theta \\ = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots si\ A = \emptyset \end{cases}$$

La regla de Yager puede ser considerada como una interpretación epistemológica honesta de la evidencia, por cuanto que no cambia la evidencia mediante la normalización del conflicto: la masa asociada al conflicto es atribuida al conjunto universal, ampliando el grado de ignorancia.

La combinación propuesta por Dubois y Prade consiste en lo siguiente: Sean $m_1(B)$ y $m_2(C)$: cuando B y C son contradictorias y no sabiendo cuál es la fuente fiable, el principio de mínimum de especificidad impone redistribuir la masa asociada a esa contradicción $m_1(B) \,.\, m_2(C)$ sobre la unión de las proposiciones $B \cup C$:

$$\forall A \subseteq \Theta$$

$$m_3(A) = m_3(A) + \sum_{\substack{B \cup C = A \\ B \cap C = \emptyset}} m_1(B) \cdot m_2(C)$$

La unión no genera conflicto alguno y no rechaza ninguna de las informaciones proporcionadas por las fuentes; no se requiere, por tanto, proceso de normalización. Esta regla de combinación se adapta mejor y es más específica que la de Yager en la asignación de la masa conflictiva.

Evaluación de las alternativas a la normalización

Todas estas propuestas alternativas de combinación evitan los resultados contra-intuitivos generados por la regla de Dempster cuando se presenta evidencia muy conflictiva. Pero cada una de ellas acarrea, por su parte, otras desventajas. Así, el método que no normaliza, asigna una masa de evidencia decreciente a los elementos focales en las sucesivas combinaciones: al usar el operador intersección sin normalización para representar un límite inferior en la creencia, resulta una creencia que se aproxima a cero a medida que crece el número de combinaciones. Y esto unido a un límite superior proporcionado por el operador unión da como resultado una brecha constantemente creciente entre creencia y plausibilidad a medida que continúan las combinaciones. Se produce así la paradoja de que a medida que se acumula evidencia crece la ignorancia: a medida que crece el número de combinaciones, la creencia asignada a la ignorancia se aproxima a 1.

4. CONFLICTO Y FIABILIDAD

En la discusión de la regla de combinación de Dempster aparecen varias nociones —*evidencia, creencia, fiabilidad, fuente de información*— y relaciones entre ellas que no siempre adquieren la misma significación en los diversos modelos formales

propuestos para su tratamiento. Si, como refleja la bibliografía, lo que se busca es un modelo adecuado para la combinación de evidencias (creencias) conflictivas, entonces el tipo, la amplitud y la exactitud de la evidencia (creencia), así como el tipo y la fiabilidad de las fuentes y sus interdependencias, pueden ser interpretados como rasgos pertinentes en la valoración del conflicto. Y luego de la valoración del conflicto, la consideración más importante es la relevancia del conflicto presenten el contexto dado. La modelización de creencias está sujeta a ciertas limitaciones, y los modelos tienen validez dentro de cierto rango; de ahí que al combinar la información proveniente de múltiples fuentes es preciso tomar en cuenta el rango y las limitaciones del modelo de creencia utilizado para cada fuente. El problema de la fiabilidad de las fuentes está ligado al problema del conflicto. En efecto, la existencia de conflicto indica la existencia de al menos una fuente no fiable. Pero, al mismo tiempo, fuentes no fiables pueden concordar y la ausencia de conflicto no garantiza la fiabilidad de las fuentes. Los coeficientes de fiabilidad controlan la influencia de las respectivas fuentes sobre los resultados de la combinación; y dependen, no sólo del modelo seleccionado, sino también de las características del entorno y de las peculiaridades de los inputs. La elección de una regla de combinación depende, pues, del contexto, de la fiabilidad de las fuentes, y no existe teoría alguna capaz de prescribir un único modelo de combinar fuentes de información en conflicto.

5. LOS MODELOS POSIBILISTAS Y DIFUSOS

Dentro del esquema posibilista (Dubois y Prade, 1992) la información obtenida de la fuente (experto) *i* viene representada por la función de distribución de posibilidad π (análoga a la función de masa *m* en la teoría de la evidencia) sobre un espacio de conocimiento $\Omega = \{\omega_1 , \omega_2 , \ldots, \omega_n \}$

$\pi(\omega) : \Omega \rightarrow [0 , 1] \qquad \forall\omega\in\Omega$

$\pi(\omega) > \pi(\omega')$ significa que ω es más plausible que ω'

$\pi(\omega) = 0$ significa que ω es imposible.

$\pi(\omega) = 1$ no significa que ω es el mundo real (como sí ocurre en el esquema probabilista), sino simplemente que nada impide que sea así; en Ω caben muchos mundos con $\pi(\omega) = 1$. Por lo que no vige la propiedad aditiva.

En la combinación de evidencias (creencias) se emplea la regla disyuntiva cuando se sabe con seguridad que *al menos una* fuente de datos es fiable; pero no se sabe cuál:

$\pi^* \cup = \cup_i \pi_i(\omega) = \text{máx.}_i \ \pi_i(\omega) \qquad \forall\omega\in\Omega$

Y para combinar fuentes igualmente fiables se emplea la regla conjuntiva siguiente:

$\pi^* \cap = \cap_i \pi_i(\omega) = \text{mín.}_i \ \pi_i(\omega) \qquad \forall\omega\in\Omega$

El modelo posibilista ofrece varias reglas de combinación que toman en cuenta de manera explícita el conocimiento disponible acerca del valor de los coeficientes de fiabilidad. Sobre la base del conocimiento disponible acerca de la fiabilidad de las fuentes se presentan varias posibilidades (Dubois y Prade, [3]):

(1) Asignar un grado numérico (absoluto o relativo) de fiabilidad a cada fuente.

(2) Establecer un orden de fiabilidad de las fuentes aun sin conocer los valores precisos de cada una de ellas.

(3) Un subconjunto de fuentes es fiable, pero ignorando cuál de ellos.

Y ante esas situaciones caben dos estrategias en la modelización de creencias que toma en cuenta la fiabilidad de las mismas: (a) utilizar de manera explícita la fiabilidad de las fuentes. Esta vía se bifurca: (a_1) incluir la fiabilidad de cada fuente *antes* de la combinación, para compensar su diferente fiabilidad y hacerlas así totalmente (al menos igualmente) fiables antes de la combinación; (a_2) o bien, si cada fuente no puede ser transformada por separado, incluir coeficientes de fiabilidad que modifiquen el operador de combinación. Y (b) estrategias para identificar la calidad de los datos de entrada en los procesos de combinación, eliminando los datos de escasa fiabilidad. En la situación (a_1), si se conoce el grado de certeza R de que la fuente *i* es fiable, entonces cabe aplicar (Dubois y Prade [2], [3]) la siguiente regla de ajuste que cambia la distribución de posibilidad π_i, proporcionada por la fuente *i*, en π'_i:

$$\pi'_i = \text{máx} (\pi_i , 1 - R_i)$$

En donde R_i es el grado de certeza de que la fuente *i* es fiable. Si $R_i = 1$ (esto es, la fuente *i* es completamente fiable), entonces $\pi_i = \pi$. Cuando $R_i = 0$ (esto es, la fuente *i* no es nada fiable), entonces $\pi_i = 1$, lo que significa ignorancia total. Las fuentes, una vez reajustadas de acuerdo con sus niveles de fiabilidad, son tomadas como absolutamente fiables y son combinadas mediante la regla conjuntiva. El inconveniente de esta regla es que cuando $R_i = 0$, entonces el soporte de la distribución reajustada deviene el universo entero.

Con todo, la modelización de creencias basada en los sistemas posibilistas y difusos ha resultado ser de gran utilidad en su aplicación en múltiples campos: procesamiento de imágenes médicas (Bloch [1]), reconocimiento de rostros (Mirhosseini, Yan, Lam y

Pham [6]), reconocimiento de crisis (Larsen y Yager [5]).

6. CONCLUSIONES

Cuando se presenta evidencia conflictiva, la regla de Dempster para la combinación de creencias genera resultados que no reflejan adecuadamente la actual distribución de creencias. Las reglas alternativas propuestas conducentes a evitar tales resultados presentan, a su vez, ventajas e inconvenientes, dependiendo en gran medida del tipo de creencia, de la fiabilidad de las fuentes y de su interdependencia. De ahí que toda modelización de creencias esté sujeta a ciertas limitaciones y obligada a especificar su rango de aplicación. La especificación más pertinente y usualmente propuesta consiste en establecer la *fiabilidad* de las creencias, siguiendo varias estrategias. Una de ellas, la basada en los modelos posibilistas y difusos, permite establecer reglas de combinación de creencias con aplicaciones útiles en múltiples dominios.

Referencias

[1] I. Bloch. Information combination operators for data fusion: a comparative review with classification. *IEEE Transactions on Systemas, Man and Cybernetics, Part A,* 26(1), Pág. 52 – 67, 1996.

[2] D. Dubois y H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence* 4, Pág. 244 – 264, 1988.

[3] D. Dubois y H. Prade. Combination of fuzzy information in the framework of possibility theory. *Data Fusion in Robotics and Machine Intelligence*. M. Abidi y R. González (eds.), Pág. 481 -505, Academic Press, 1992.

[4] G. Jaeger y U. Benz. Supervised fuzzy classification of SAR data using multiple sources. *Proceedings of IEEE 1999 International Geoscience and Remote Sensing Symposium, 3*, Pág. 1603 – 1605, 1999.

[5] H. Larsen y R. Yager. A framework for fuzzy recognition technology. *IEEE Transactions on Systems, Man, and Cybernetics, Part C* , 30(1), Pág. 65 – 76, 2000.

[6] A. R. Mirhosseini y H. Yan y K.-M. Lam y T. Pham. Human face image recognition: an evidence aggregation approach. *Computer Vision and Image Understanding*, 7(2), Pág. 213 – 230, 1998.

[7] Ph. Smets. The combination of evidence in transferable belief model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2), Pág. 447 – 458, 1990.

[8] R. Yager. On the Dempster-Shafer framework and new combination rueles. *Information Sciences*, 41, Pág. 93 – 138, 1987.

[9] L. A. Zadeh. Review of Mathematical Theory of Evidence, by G. Shafer. *AI Magazine*, 5(3), Pág. 81 – 83, 1984.

LA REGLA COMPOSICIONAL DE ZADEH,
UNA LECCIÓN PARA PRINCIPIANTES

Enric Trillas[1] **Itziar García-Honrado**[2]

[1] European Center for Soft-Computing. Mieres, Asturias (España), enric.trillas@softcomputing.es
[2] European Center for Soft-Computing. Mieres, Asturias (España), itziar.garcia@softcomputing.es

Resumen

La regla composicional de inferencia borrosa es el instrumento básico para el control de los sistemas descritos por reglas imprecisas. Se efectúa una presentación teórica de tal regla de inferencia y se obtienen, como casos particulares, los llamados métodos de Mamdani, Larsen y Takagi-Sugeno.

Palabras Clave: CRI, Mamdani-Larsen, Takagi-Sugeno.

1 Introducción

A la regla composicional de inferencia borrosa, introducida por Lotfi A. Zadeh, y en inglés abreviada por CRI (Compositional Rule of Inference), se le deben muchos de los éxitos conseguidos por las tecnologías borrosas. También se le debe el resultado teórico conocido como *teorema de la aproximación universal* de una función continua, por medio de un sistema de reglas en un intervalo cerrado de $\mathcal{R}$ que describa lingüísticamente tal función.
Pese a su indudable relevancia para las aplicaciones de la lógica borrosa, en general, en los cursos de introducción la CRI se explica de una forma muy deficiente. Una forma que puede llevar a creer que es una regla que aparece bien como "sacada de la manga", bien como una regla arbitraria, cuando no es sino una condición suficiente para obtener una "buena" conclusión del Modus Ponens General,

$$\frac{\begin{array}{c}\text{Si x es P, entoces y es Q}\\ \text{x es } P^*\end{array}}{\text{y es } Q^*.}$$

Es decir, se trata de encontrar una función μ_{Q^*} que represente la conclusión "y es Q^*".

Además, cuando con ella se avanza hacia el control borroso de un sistema dinámico, puede dar la impresión de que los diversos métodos, como el de Takagi-Sugeno, tienen una lejana relación con la CRI cuando, en realidad, no son sino casos particulares de la misma. En el presente artículo/lección se estudiará de dónde viene la CRI, de forma asequible para los alumnos de un curso de introducción a la lógica borrosa y suponiendo que han asimilado cuanto antecede a ésta lección.

2 Conceptos básicos

2.1

La descripción lingüística y reglada de un sistema dinámico con dos variables (x, y), que toman valores respectivamente en los conjuntos X, Y , se efectúa en lenguaje natural por medio de reglas del estilo

$$r_i\text{: Si x es } P_i\text{, entonces y es } Q_i, \quad 1 \leq i \leq n,$$

con P_i y Q_i predicados lingüísticos en X e Y, respectivamente. La representación de cada r_i se obtiene mediante un T-condicional J_i, i.e., una función $J_i : [0,1] \times [0,1] \to [0,1]$, para la que existe una t-norma continua T_0 verificando la desigualdad, $T_0(a, J_i(a,b)) \leq b$, para todos los a, b en $[0,1]$. Con ella la representación de r_i es la función

$$J_i(\mu_{P_i}(x), \mu_{Q_i}(y)), \;\; x \in X, \; y \in Y,$$

una vez los predicados P_i y Q_i han sido representados por sendos conjuntos borrosos $\mu_{P_i} \in [0,1]^X$ y $\mu_{Q_i} \in [0,1]^Y$. Nótese que r_i es un enunciado en lenguaje natural que, por tanto, tiene un significado; ese significado, y no otro , es el que debe ser capturado por J_i. Por ejemplo, si r_i debiese ser interpretada como

$$No\ (x \text{ es } P_i)\ \acute{}\text{o}\ (y \text{ es } Q_i),$$

entonces J_i deberá responder a esa interpretación, es decir, será del tipo $J_i(\mu_{P_i}(x), \;\mu_{Q_i}(y)) =$

$S_i(N_i(\mu_{P_i}), \mu_{Q_i}(y))$, y aún deberán determinarse la t-conorma S_i y y la negación fuerte N_i. Si r_i debiera ser interpretada como

$$(x \text{ es } P_i) \ y \ (y \text{ es } Q_i),$$

J_i deberá responder al modelo $J_i(\mu_{P_i}(x), \mu_{Q_i}(y)) = T_i(\mu_{P_i}(x), \mu_{Q_i}(y))$, y aún deberá determinarse la más adecuada t-norma continua T_i para el caso en cuestión.

2.2

Las funciones J_i deben satisfacer una desigualdad $T_0(a, J_i(a,b)) \leq b$, llamada MP[1]. La razón es la siguiente. Dada la regla r_i, y dado también su antecedente "x es P_i", se exige que el consecuente "y es Q_i" sea una consecuencia lógica de esas dos premisas , es decir,

De "Si x es P_i, entonces y es Q_i" *y de* "x es P_i", *se sigue lógicamente* "y es Q_i". (Modus Ponens).

Ello, traducido al lenguaje borroso, significa que debe existir una t-norma continua T_0, tal que

$$T_0(\mu_{P_i}(x), J_i(\mu_{P_i}(x), \mu_{Q_i}(y))) \leq \mu_{Q_i}(y),$$

para todos los x en X, y en Y. En consecuencia, para que J_i pueda ser utilizado para cualesquiera conjuntos borrosos, debe verificar $T_0(a, J_i(a,b)) \leq b$, para todos los a, b en $[0,1]$, con una adecuada t-norma T_0.
Por ejemplo, si J es la S-implicación de Kleene-Dienes $J(a,b) = \max(1-a, b)$, de $T_0(a, \max(1-a,b)) \leq b$, con $b = 0$ se sigue necesariamente, $T_0(a, 1-a) = 0$; es decir, existe la t-norma continua $T_0 = W$, con la que, en efecto, es $W(a, \max(1-a,b)) = \max(0, a + \max(1-a,b) - 1) = \max(0, a+b-1) = W(a,b) \leq b$. Obsérvese que ni $T_0 = \min$ ni $T_0 = prod$, pueden servir; por ejemplo, es $\min(0.4, \max(1-0.4, 0.3)) = \min(0.4, 0.6) = 0.4 > 0.3$, y $0.2 \times \max(1-0.2, 0.1) = 0.2 \times 0.8 = 0.16 > 0.1$.

[1]En lógica clásica, el Modus Ponens (MP) '$a, a \to b : b$' se traduce por la desigualdad $a \cdot (a \to b) \leq b$ que, en un álgebra de Boole, es equivalente a la desigualdad $a \to b \leq a' + b$ (intersección=$\cdot$, unión=$+$, complemento=$'$) y, por tanto el condicional $a \to b = a' + b$ es el mayor de todos ellos.Con esa desigualdad, b es una consecuencia lógica del conjunto de premisas $\{a, a \to b\}$. Además, también en un álgebra de Boole B, resulta

$$\sup\{z \in B; a \cdot z \leq b\} = \sup\{z \in B; z \leq a' + b\} = a' + b.$$

Sin embargo, éstas propiedades sólo valen en las álgebras de Boole. En un retículo ortomodular, por ejemplo, no existe un condicional máximo, sino algunos que son maximales como el de Sasaki $a \to b = a' + a \cdot b$, y el de Dishkant $a \to b = b + a' \cdot b'$.

En el caso $J_i(a,b) = T_i(a,b)$, es $T_0(a, T_i(a,b)) \leq \min(a, \min(a,b)) = \min(\min(a,a), b) = \min(a,b) \leq b$, es decir, T_0 puede ser cualquier t-norma. Sin embargo, la que más acercará el término $T_0(a, J(a,b))$ al término b, será $T_0 = \min$, la mayor t-norma, que es la que se toma en esos casos.

2.3

Es un resultado matemático conocido que la desigualdad MP, $T_0(a, J(a,b)) \leq b$, es equivalente a la desigualdad

$$J(a,b) \leq J_{T_0}(a,b) := \sup\{z \in [0,1]; T_0(z,a) \leq b\},$$

donde J_{T_0} recibe el nombre de R-implicación (R por residuada). Por lo tanto, la R-implicación J_{T_0} es el mayor de todos los T_0-condicionales. Por ejemplo, como

$$J_W(a,b) = \sup\{z \in [0,1]; W(a,z) = \max(0, z+a-1) \leq b\} = \min(1, 1-a+b),$$

resulta $\max(1-a, b) \leq \min(1, 1-a+b)$. Análogamente, es

$$J_{prod}(a,b) = \begin{cases} 1, & \text{si } a \leq b \\ \frac{b}{a}, & \text{si } a > b, \end{cases}$$

y como sea que

$$J(a,b) = \begin{cases} 1, & \text{si } a \neq 0, b = 1 \\ \min(1, \frac{b}{a}, \frac{1-a}{1-b}), & \text{en otro caso,} \end{cases}$$

es un prod-condicional, también resulta $J(a,b) \leq \begin{cases} 1, & \text{si } a \leq b \\ \frac{b}{a}, & \text{si } a > b, \end{cases}$ desigualdad que, que por lo demás es obvia.

2.4

Para terminar este párrafo, veamos un ejemplo con el que se pretende ilustrar la importancia de elegir correctamente el T-condicional J que deberá representar a la regla.
Dada la regla, 'Si x es pequeño, entonces y es grande', con x, y en $[0,1]$ y $\mu_P(x) = 1-x$, $\mu_G(y) = y$, de elegir el W-condicional $J(a,b) = \max(1-a, b)$, resulta que la regla viene representada por

$$J(\mu_P(x), \mu_G(y)) = \max(1-\mu_P(x), \mu_G(y)) = \max(x, y),$$

que cabe interpretar como $\max(\mu_G(x), \mu_G(y))$, es decir, "x es grande ó y es grande". De la misma manera, de elegir el W-condicional $J_W(a,b) = \min(1, 1-a+b)$, resulta

$$J_W(\mu_P(x), \mu_G(y)) = \min(1, 1-\mu_P(x) + \mu_Q(y)) = \min(1, x+y) = W^*(x,y),$$

otra forma de representar 'x es grande ó y es grande'. Ambos resultados pueden resultar sorprendentes.

3 La regla composicional de inferencia borrosa

De momento, nos limitaremos a considerar un sistema ideal (x,y) descrito por una única regla,

$$\text{r: Si } x \text{ es } P, \text{ entonces } y \text{ es } Q,$$

con $x \in X$, $y \in Y$. De ordinario, en las aplicaciones los conjuntos X e Y son intervalos cerrados de la recta real. Supondremos capturados los significados de P, Q y r por medio de funciones de pertenencia $\mu_P \in [0,1]^X, \mu_Q \in [0,1]^Y$ y un T-condicional $J : [0,1] \times [0,1] \to [0,1]$.
Como por hipótesis el sistema (x, y) es dinámico, al observar la variable x difícilmente cabrá afirmar 'x es P', sino que las observaciones serán del tipo 'x es P^*', para cierto predicado P^* en X y, con ello, el consecuente correspondiente será del tipo 'y es Q^*' para un cierto predicado Q^* en Y. Se trata de resolver el problema de hallar qué conjunto borroso $\mu_{Q^*} \in [0,1]^Y$ corresponde al $\mu_{P^*} \in [0,1]^X$. La formalización requiere que se cumpla,
De $\mu_{P^*}(x)$ y de $J(\mu_P(x), \mu_Q(y))$, debe seguirse lógicamente $\mu_{Q^*}(y)$, para todo $x \in X$ y $y \in Y$.

Es decir,

$$T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))) \leq \mu_{Q^*}, \quad \forall x \in X, \forall y \in Y, \tag{1}$$

con la condición de que si $P^* = P$, debe ser $Q^* = Q$. Por lo tanto la t-conorma continua T_0 debe hacer de J un T_0-condicional, es decir, debe verificar $T_0(a, J(a,b)) \leq b$ para todos los $a, b \in [0,1]$.

De la anterior desigualdad se sigue, sin más,

$$\sup_{x \in X} T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))) \leq \mu_{Q^*}, \quad \forall y \in Y,$$

que da una buena cota inferior de $\mu_{Q^*}(y)$. *La regla composicional de inferencia* de Zadeh *consiste en tomar esa cota inferior como salida* del sistema. Es decir, en *definir la salida* por la fórmula,

$$\mu_{Q^*}(y) = \begin{cases} \mu_Q(y), & \text{si } \mu_{P^*} = \mu_P \\ \sup_{x \in X} T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))), & \text{en otro caso,} \end{cases}$$

con la t-norma T_0 para la que J es un T_0-condicional. Como usualmente, μ_{P^*} no es μ_P, se usará sólo la última línea de la fórmula anterior.

En el caso de que X sea finito, $X = \{x_1, x_2, ..., x_n\}$, el supremo se alcanza y, por lo tanto podemos hablar de máximo,

$$\sup_{x \in X} T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))) = \max_{x \in X} T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))).$$

Si además Y también es finito, $Y = \{y_1, y_2, ...y_m\}$, el cálculo de μ_{Q^*} puede realizarse mediante el producto de matrices $Max - T_0$, del siguiente modo,

$$\mu_{Q^*} = \begin{pmatrix} \mu_{Q^*}(y_1) & ... & \mu_{Q^*}(y_m) \end{pmatrix} =$$

$$\begin{pmatrix} \mu_{P^*}(x_1) & ... & \mu_{P^*}(x_n) \end{pmatrix} \bullet$$

$$\begin{pmatrix} J(\mu_P(x_1), \mu_Q(y_1)) & ... & J(\mu_P(x_1), \mu_Q(y_m)) \\ ... & ... & ... \\ J(\mu_P(x_n), \mu_Q(y_1)) & ... & J(\mu_P(x_n), \mu_Q(y_m)) \end{pmatrix},$$

donde $\bullet$ representa el producto $Max - T_0$ calculado como,

$$\mu_{Q^*}(y_j) = \max_{i \in \{1,2,...,n\}} T_0(\mu_{P^*}(x_i), J(\mu_P(x_i), \mu_Q(y_j))).$$

3.1 Ejemplos

1. Si la observación es numérica, es decir, del tipo $x = x_0$, entonces es $P^* =$ 'x_0' y, por tanto, 'x es P^*' no es sino 'x es x_0'. Con ello, es $\mu_{P^*} = \mu_{\{x_0\}}$, es decir,

   $$\mu_{P^*}(x) = \begin{cases} 1, & \text{si } x = x_0 \\ 0, & \text{si } x \neq x_0. \end{cases}$$

   Por tanto, resulta

   $$\mu_{Q^*}(y) = \sup_{x \in X} T_0(\mu_{\{x_0\}}(x), J(\mu_P(x), \mu_Q(y))) =$$

   $$J(\mu_P(x_0), \mu_Q(y)), \quad \forall y \in Y.$$

2. Si, en el ejemplo anterior, es $J(a,b) = \min(a,b)$, resulta

   $$\mu_{Q^*}(y) = \min(\mu_P(x_0), \mu_Q(y)), \quad \forall y \in Y.$$

3. Si es $J(a,b) = a \cdot b$, resulta

   $$\mu_{Q^*}(y) = \mu_P(x_0) \cdot \mu_Q(y), \quad \forall y \in Y$$

4. Si es $J(a,b) = S(N(a), b)$, resulta

   $$\mu_{Q^*}(y) = S(N(\mu_P(x_0)), \mu_Q(y)), \quad \forall y \in Y$$

5. Si en (2) y (3), el consecuente de la regla también es numérico, '$y = y_0$', en los dos casos resulta

   $$\mu_{Q^*}(y) = \begin{cases} \mu_P(x_0), & \text{si } y = y_0 \\ 0, & \text{si } y \neq y_0. \end{cases}$$

4 Métodos de Mamdani, de Larsen y de Takagi-Sugeno (de orden cero)

4.1 Método de Mamdani

El llamado método de Mamdani no consiste en otra cosa que representar la regla por $J(a,b) = \min(a,b)$, con lo que es $T_0 = \min$. En ese caso particular se obtiene la salida

$$\mu_{Q^*}(y) = \sup_{x\in X} \min(\mu_{P^*}(x), \min(\mu_P(x), \mu_Q(y))) =$$

$$\sup_{x\in X} \min(\min(\mu_{P^*}(x), \mu_P(x)), \mu_Q(y)),$$

a causa de la asociatividad de la operación min.

4.2 Ejemplo

La regla: 'Si x es pequeño, entonces y es grande', con $x \in [0,1]$, $y \in [0,10]$, y $\mu_{peq}(x) = 1-x, \mu_{gran}(y) = \frac{y}{10}$, con la observación '$x$ es muy pequeño' ($\mu_{P^*} = (1-x)^2$), lleva a la salida

$$\mu_{Q^*}(y) = \min(\sup_{x\in[0,1]} \min((1-x)^2, 1-x), \frac{y}{10})$$

$$= \min(\sup_{x\in[0,1]} (1-x)^2, \frac{y}{10}) = \min(1, \frac{y}{10}) = \frac{y}{10}.$$

4.3 Método de Larsen

El llamado método de Larsen no consiste en otra cosa que tomar $J(a,b) = a{\cdot}b$, con lo que es $T_0 = \min$ y se obtiene la salida

$$\mu_{Q^*}(y) = \sup_{x\in X} \min(\mu_{P^*}(x), \mu_P(x){\cdot}\mu_Q(y)), \quad \forall y \in Y.$$

Con los datos del ejemplo anterior, resulta

$$\mu_{Q^*}(y) = \sup_{x\in[0,1]} \min((1-x)^2, (1-x){\cdot}\frac{y}{10}) =$$

$$\sup_{x\in[0,1]} ((1-x){\cdot}\min(1-x, \frac{y}{10})) = \frac{y}{10}.$$

4.4 Método de Takagi-Sugeno de orden cero

El llamado método de Takagi-Sugeno de orden cero[2], no es sino el de Larsen con observación numérica $x = x_0$, y consecuente numérico $y = y_0$, es decir, $\mu_{P^*} = \mu_{\{x_0\}}$ y $\mu_Q = \mu_{\{y_0\}}$. Con ello, como se ha visto en el ejemplo 5 de 3.1, resulta la función salto,

$$\mu_{Q^*}(y) = \mu_P(x_0){\cdot}\mu_{\{y_0\}}(y) = \begin{cases} 0, & \text{si } y \neq y_0 \\ \mu_P(x_0), & \text{si } y = y_0. \end{cases}$$

[2]Es llamado de orden cero porque la salida obtenida es una constante, es decir un polinomio de orden cero. En el caso de que la salida fuera un polinomio de orden n, se llama método de Takagi-Sugeno de orden n.

4.5 Ejemplo

Con la regla 'Si x es pequeño, entonces $y = 5$' y la observación $x = 0.3$, con $x \in [0,1], y \in [0,1]$ y $\mu_{peq}(x) = 1-x$, resulta $\mu_{peq}(0.3) = 0.7$ y

$$\mu_{Q^*}(y) = \begin{cases} 0, & \text{si } y \neq 5 \\ 0.7, & \text{si } y = y_0. \end{cases}$$

4.6

El predicado P^* de la observación no tiene por qué ser únicamente del tipo $x = x_0$; la observación puede indicar que el valor está en un intervalo del tipo $[x_0 - \varepsilon, x_0 + \varepsilon]$. En ese caso, la CRI también facilita una salida del sistema.

4.7 Ejemplos

a. Regla, 'Si x es grande, entonces $y = 0.8$, con x, y en $[0,1]$. Observación, $x \in [0.4, 0.6]$. Por tanto,

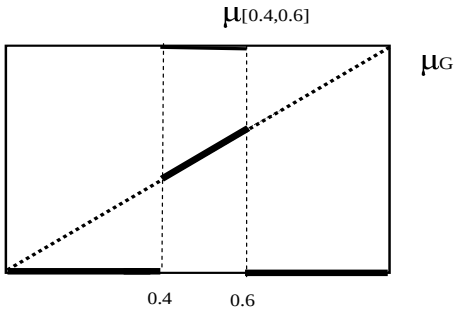
$$J(\mu_G(x), \mu_{\{0.8\}}(y)) = x{\cdot}\mu_{\{0.8\}}(y),$$

$$\mu_{Q^*}(y) = \sup_{x\in[0,1]} \min(\mu_{[0.4,0.6]}(x), x{\cdot}\mu_{\{0.8\}}(y)).$$

Con ello,

$$\mu_{Q^*}(y) = \begin{cases} 0, & \text{si } y \neq 0.8 \\ 0.6, & \text{si } y = 0.8, \end{cases}$$

ya que



b. Con la misma observación anterior, y la regla 'Si x es grande, entonces y es pequeño', con $\mu_G(x) = x, \mu_P(y) = 1-y, J(a,b) = \min(a, 1-b)$ (Mamdani), se obtiene

$$\mu_{Q^*}(y) = \sup_{x\in[0,1]} \min(\mu_{[0.4,0.6]}(x), \min(x, 1-y)) =$$

$$\sup_{x\in[0,1]} \min(\min(\mu_{[0.4,0.6]}(x), x), 1-y) =$$

$$\sup_{x\in[0.4,0.6]} \min(x, 1-y) = \min(0.6, 1-y).$$

c. Sean $X = \{1, 2, 3\}$, $Y = \{6, 7\}$ dos conjuntos finitos , y la regla 'Si x está alrededor de 2, entonces $y = 6$', con $\mu_{\mathrm{A2}}(x) = \begin{cases} 0.5, & \text{si } x = 1 \\ 1, & \text{si } x = 2 \\ 0.5, & \text{si } x = 3, \end{cases}$.

Con la observación $\mu_{P^*} = 0.6/1 + 0.9/2 + 0.7/3$, y utilizando $J(a, b) = a{\cdot}b$ y $T = \min$ (Larsen), se calcula μ_{Q^*} con el producto min-max de matrices, del siguiente modo,

$$\mu_{Q^*} = \begin{pmatrix} \mu_{Q^*}(6) & \mu_{Q^*}(7) \end{pmatrix} =$$

$$\begin{pmatrix} 0.6 & 0.9 & 0.7 \end{pmatrix} \bullet \begin{pmatrix} J(0.5, 1) & J(0.5, 0) \\ J(1, 1) & J(0.5, 0) \\ J(0.5, 1) & J(0.5, 0) \end{pmatrix}$$

$$= \begin{pmatrix} \max(0.5, 0.9, 0.5) & \max(0, 0, 0) \end{pmatrix} = \begin{pmatrix} 0.9 & 0 \end{pmatrix},$$

y resulta $\mu_{Q^*} = 0.9/6 + 0/7$.

5 Caso de varias reglas

Realmente no hay sistemas físicos descritos por una única regla , los sistemas siempre vienen descritos por varias reglas r_i, $1 \le i \le n$. Dada una observación, y n reglas

r_1 : Si x es P_1, entonces y es Q_1

r_2 : Si x es P_2, entonces y es Q_2

.....

r_n : Si x es P_n, entonces y es Q_n,

¿cómo encontrar la salida común μ_{Q^*}? Obviamente, cada regla da una salida $\mu_{Q_i^*}$; es decir, de dispone de n salidas $\mu_{Q_1^*}, \mu_{Q_2^*}, ...\mu_{Q_n^*}$, cada una facilitada por cada una de las reglas r_i.
Suponiendo que se *dispara* cada vez una de esas reglas, es decir, "r_1 ó r_2 ó ... ó r_n", entonces cabe definir una disyunción de las n-salidas como la salida común del sistema; es decir, tomar $\mu_{Q^*}(y) = S(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y), ..., \mu_{Q_n^*}(y))$, para todo $y \in Y$, y una t-conorma continua S. En general, como sea que max es la menor t-conorma, se toma

$$\mu_{Q^*}(y) = \max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y), ..., \mu_{Q_n^*}(y)), \quad \forall y \in Y.$$

5.1 Ejemplo

Un sistema con dos variables continuas $x, y \in [0, 10]$, está descrito por las 3 reglas

r_1 : Si x es grande, entonces $y = 2$

r_2 : Si x es pequeño, entonces $y = 8$

r_3 : Si x está alrededor de 4, entonces $y = 6$.

¿Cuál es la salida correspondiente a la observación x=4.5?
La solución se obtiene siguiendo los cuatro pasos siguientes.

- <u>1^{er} *Paso* :</u> Representar los predicados G='grande', P='pequeño', y A4='alrededor de 4'. Supondremos,

$$\mu_G(x) = \frac{x}{10}, \mu_P(x) = 1 - \frac{x}{10}$$

$$\mu_{A4}(y) = \begin{cases} 0, & \text{si } 0 \le x \le 3, 5 \le x \le 10 \\ x - 3, & \text{si } 3 \le x \le 4 \\ 5 - x, & \text{si } 4 \le x \le 5. \end{cases}$$

- <u>2^o *Paso* :</u> Representar cada regla por un T-condicional J. Supondremos, como es usual en el control borroso que $J_1(a, b) = J_2(a, b) = J_3(a, b) = a.b$. Con ello,
  - r_1 se representará por $J_1(\mu_G(x), \mu_{\{2\}}(y)) = \begin{cases} \frac{x}{10}, & \text{si } y = 2 \\ 0, & \text{si } y \ne 2. \end{cases}$
  - r_2 se representará por $J_2(\mu_P(x), \mu_{\{8\}}(y)) = \begin{cases} 1 - \frac{x}{10}, & \text{si } y = 8 \\ 0, & \text{si } y \ne 8. \end{cases}$
  - r_3 se representará por $J_3(\mu_{A4}(x), \mu_{\{6\}}(y)) = \begin{cases} \mu_{A4}(x), & \text{si } y = 6 \\ 0, & \text{si } y \ne 6. \end{cases}$
- <u>3^{er} *Paso* :</u> Obtener las salidas de cada regla. Se aplica la regla composicional de inferencia

$$\mu_{Q_i^*}(y) = \sup_{x \in [0,10]} \min(\mu_{\{4.5\}}(x), J_i(\mu_{P_i}(4.5), \mu_{Q_i}(y))),$$

para todo y. Resulta:

$$\mu_{Q_1^*}(y) = \begin{cases} 0.45, & \text{si } y = 2 \\ 0, & \text{si } y \ne 2, \end{cases}$$

$$\mu_{Q_2^*}(y) = \begin{cases} 0.55, & \text{si } y = 8 \\ 0, & \text{si } y \ne 8, \end{cases}$$

$$\mu_{Q_3^*}(y) = \begin{cases} \mu_{A4}(4.5) = 0.5, & \text{si } y = 6 \\ 0, & \text{si } y \ne 6, \end{cases}$$

- <u>4^o *Paso* :</u> Obtener la salida común $\mu_{Q^*} = \max(\mu_{Q_1^*}, \mu_{Q_2^*}, \mu_{Q_3^*})$. Resulta la función discontinua,

$$\mu_{Q^*}(y) = \begin{cases} 0.45, & \text{si } y = 2 \\ 0.5, & \text{si } y = 6 \\ 0.55, & \text{si } y = 8 \\ 0, & \text{si } y \in [0, 10] \setminus \{2, 6, 8\}. \end{cases}$$

5.2 Ejemplo

Con x, y en $[0,1]$, se consideran las dos reglas

- r_1: Si x es grande, entonces y es pequeño, representada por $J_1(a,b) = a \cdot b$
- r_2: Si x es muy pequeño, entonces y es muy grande, representada por $J_2(a,b) = \min(a,b)$,

y la observación $x = 0.4$. Si los predicados G=grande, MG=muy grande, P=pequeño, MP=muy pequeño, están representados por

$$\mu_G(x) = x,\ \mu_{MG}(y) = y^2,$$

$$\mu_P(y) = 1 - y,\ \mu_{MP}(x) = (1-x)^2,$$

entonces

$$J_1(\mu_G(x), \mu_P(y)) = J_1(x, 1-y) = x \cdot (1-y),$$

$$J_2(\mu_{MP}(x), \mu_{MG}(y)) = \min((1-x)^2, y^2) =$$

$$(\min(1-x, y))^2,$$

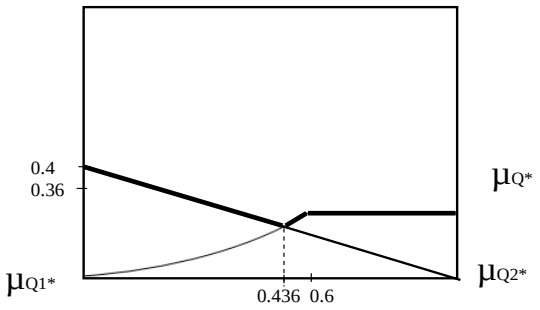
con lo cual, resulta:

$$\mu_{Q_1^*}(y) = J_1(0.4, 1-y) = 0.4(1-y),$$

$$\mu_{Q_2^*}(y) = J_2((1-0.4)^2, y^2) = (\min(0.6, y))^2 =$$

$$\begin{cases} 0.36, & \text{si } y \geq 0.6 \\ y^2, & \text{si } y < 0.6. \end{cases}$$

Gráficamente, como $0.4(1-y) = y^2 \Leftrightarrow y = 0.463$, resulta



Con todo ello,

$$\mu_{Q^*}(y) = \max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) =$$

$$\begin{cases} 0.4(1-y), & \text{si } 0 \leq y \leq 0.436 \\ y^2, & \text{si } 0.436 \leq y \leq 0.6 \\ 0.36, & \text{si } 0.6 \leq y \leq 1. \end{cases}$$

6 Conclusión

Se ha intentado explicar de un modo simple qué es la regla composicional de inferencia borrosa de Zadeh, una regla sin la cual no se habría podido desarrollar ni la inferencia aproximada, ni los métodos para el control borroso de sistemas dinámicos que son, todos ellos, inmediatamente derivados de aquella regla; casos particulares de la misma en realidad.
Nada se ha dicho aquí de los métodos de desborrosificación para compactar la salida funcional $\mu_{Q^*} \in [0,1]^Y$ en un único valor $Des(\mu_{Q^*}) \in Y$. Son métodos que han quedado fuera del objetivo del artículo/lección aunque son esenciales tanto para el control borroso, como para probar que el comportamiento inferido del sistema, es decir, la función que, dadas las reglas, le hace corresponder a cada $x \in X$ el valor $Des(\mu_{Q^*}) \in Y$, difiere en tan poco como se quiera de la función de X en Y real continua, que representa el comportamiento del sistema dinámico dado, el llamado *teorema de la aproximación universal*. Es el caso, por ejemplo, del método de Takagi-Sugeno de orden cero, en el que a cada $x_0 \in X$ le corresponde, como de ha visto, una función discreta del tipo

$$\mu_{Q^*}(y) = \begin{cases} \alpha_1, & \text{si } y = y_1 \\ \alpha_2, & \text{si } y = y_2 \\ ... & \\ \alpha_n, & \text{si } y = y_n \\ 0, & \text{en otro caso,} \end{cases}$$

con la que, obviamente,

$$Des(\mu_{Q^*}) = \frac{\alpha_1 y_1 + \alpha_2 y_2 + ...\alpha_n y_n}{\alpha_1 + \alpha_2 + ...\alpha_n}.$$

Representar un predicado por una curva, una regla por una superficie, y la salida de un sistema de reglas imprecisas más una observación de su antecedente por la CRI, son tres de las más importantes características de la lógica borrosa.

Referencias

[1] H. T. Nguyen and E. A. Walker. *A First course in Fuzzy Logic*. Chapman & Hall/Crc, 2000.

[2] K. Tanaka. *An introduction to Fuzzy Logic for practical applications*. Springer-Verlag, New York, 1996.

[3] G. J. Klir and B. Yuan. *Fuzzy sets and fuzzy logic*. Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[4] E. Trillas. On Logic and Fuzzy Logic. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 1(2):107–137, 1993.

REPRESENTATION OF IMPRECISE PROPERTIES VIA RESTRICTION LEVELS: AN ALTERNATIVE TO FUZZY NEGATION

Daniel Sánchez, Miguel Delgado, and María-Amparo Vila

Depto. Ciencias de la Computación e I.A., Universidad de Granada
E.T.S.I.I.T., C/Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
{daniel,vila}@decsai.ugr.es, mdelgado@ugr.es

Abstract

In this paper we discuss about the alternative representation of imprecise properties via restriction levels. The representation is equivalent to that of fuzzy sets as collections of alpha-cuts for properties that are either atomic (i.e., not defined in terms of logical operations) or defined in terms of conjunction and disjunction only. Operations are defined on levels, leading to conjunction and disjunction based on minimum and maximum, respectively. However, negation is also performed on levels, and hence the representation obtained does not always correspond to the alpha-cuts of a fuzzy set. We discuss about the properties of this representation.

Keywords: Restriction Levels, Imprecise Properties, Fuzzy Sets, Negation.

1 INTRODUCTION

There is a virtually infinite number of logics for reasoning with imprecise properties represented by fuzzy sets. Different triplets of dual t-norms and t-conorms via a fuzzy negation can be employed for this purpose. However, many properties of these and other operations in the crisp case do not hold or must be redefined in different ways.

This problem arises frequently when the operations and definitions we want to extend involve negations (eq. complement of the representation). As an example, the excluded middle axioms do not hold for most of the combinations of t-norm, t-conorm, and negation, and specifically for the most employed ones, like the triplet comprised of maximum, minimum, and the standard negation. This fact is argued usually as desirable because it allows for the representation of imprecise boundaries.

Let us consider the case of set difference. Given crisp sets R and S, $R \backslash S = R \cap \bar{S} = T$ so that $T \cup (R \cap S) = R$. However, for many of the most employed t-norms, t-conorms and negations, the previous properties do not hold. In particular, if F is a fuzzy set defined on a crisp universe X, $X \backslash F = X \cap \bar{F} = \bar{F}$ but $F \cup \bar{F} = X$ is not always true since the excluded middle axioms do not hold in general. Another problem involving negation is the extension of implication as $\neg F \vee G$ and the corresponding set interpretation as a degree of inclusion (in particular, the original definitions of fuzzy set inclusion and equality are crisp).

In this paper we recall the restriction-level representation of imprecise properties introduced in [2] and the alternatives it offers for dealing with the negation of logical expressions. The representation is equivalent to that of fuzzy sets as collections of alpha-cuts for atomic properties. However, fuzzy sets are just a particular case of restriction-level representation in general, since the crisp sets representing the property in each level are not forced to be nested in the usual way. Operations are defined on levels, leading in the particular case of fuzzy sets to conjunction and disjunction based on minimum and maximum, respectively. However, since negation is performed on levels as well, the representation obtained does not always correspond to the alpha-cuts of a fuzzy set. We show that the new representation and operations allow for the representation of imprecision while maintaining properties related to the negation.

2 RESTRICTION-LEVEL REPRESENTATION

In [2] we introduced the restriction-level representation (RL-representation for short) of imprecise properties as an extension of the usual alpha-cut representation of fuzzy sets.

2.1 REPRESENTATION

The RL-representation of an imprecise property defined on a set X is a collection of crisp subsets of X, each one corresponding to a crisp realization of the property under a *restriction* level. The notion of restriction level of representation is very close to that of alpha-cut of a fuzzy set, i.e., a crisp set associated to a degree $\alpha \in [0, 1)$; in fact, the representation of a fuzzy set via the set of its alpha-cuts can be seen as a restriction-level representation of the imprecise property represented by the fuzzy set. Every value $\alpha \in [0, 1)$ defines a restriction level. In the particular case of fuzzy sets, in the restriction level α, the alpha-cut of a fuzzy set μ_P can be seen as a (crisp) approximation of the imprecise property P obtained by relaxing our criterion to decide whether an object verifies the property; specifically, in the restriction level α we assume that x verifies P iff $\mu_P(x) \geq \alpha$.

The main difference between a RL-representation and the representation via a set of alpha-cuts of a fuzzy set is that the crisp sets of a RL-representation are not required to verify the usual nested structure of alpha-cuts. Hence, a restriction-level representation does not always represent a fuzzy set. This idea may look strange to those familiar with fuzzy set theory, in particular if we try to make an interpretation of the representation as we use to do with alpha-cuts. We shall discuss on this later.

We distinguish between *atomic* and *derived* properties. Atomic properties are those that cannot be defined in terms of other properties in our problem. Derived properties are defined by operations on other properties. Obviously, the kind of derived properties we can represent depends on the operations we define. In this paper, we shall employ the definition of conjunction, disjunction, and negation introduced in [2] on the basis of the corresponding set operations of intersection, union, and complement.

In [2] we consider that atomic imprecise properties are represented by fuzzy sets, and hence restrictions are of the form *degree* $\geq \alpha$ with $\alpha \in (0, 1]$, and restriction levels are associated to values $\alpha \in (0, 1]$. In the same case, the crisp realization of an atomic imprecise property represented by a fuzzy set A in the restriction level α corresponds to the α-cut A_α. However, let us remark that this is not the only possible way to obtain a RL-representation of an atomic property, but one that allows us to employ a familiar mechanism to represent imprecision, fuzzy sets.

We assume that, though theoretically an infinite number of restriction levels can be considered, there is a finite set of restriction levels necessary in order to represent a certain property. Specifically, if we consider the levels ordered by degree α in decreasing order, we shall consider any finite set of restriction levels including those where there is a change in the crisp representation with respect to the previous level. This consideration is not a practical limitation since humans are able to distinguish a limited number of restriction or precision levels and, in practice, the limit in precision and storage of computers allows us to work with a finite number of objects and degrees (and consequently, of levels) only. We call such finite subsets of $[0, 1)$ *RL-sets*.

Definition 2.1 *A* RL-set Λ *is a finite set of restriction levels* $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ *verifying* $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$, $m \geq 1$.

In practice, the RL-set for an atomic property represented by a fuzzy set A on an universe X is

$$\Lambda_A = \{A(x) \mid x \in \mathit{support}(A)\} \cup \{1\} \tag{1}$$

However, notice that any finite subset of $\Lambda \subset [0, 1)$ with $\Lambda_A \subseteq \Lambda$ is a suitable RL-set for the representation of A. Following this idea, the RL-set employed to represent a derived property is obtained as the union of the RL-sets of the atomic properties in terms of which the property is defined. The idea is that in order to represent the result of an operation between representations, we need to consider at least the minimum restriction levels for those representations involved.

The RL-representation of an imprecise property on X is defined in [2] as follows:

Definition 2.2 ([2]) *A RL-representation on* X *is a pair* (Λ, ρ) *where* Λ *is a RL-set and* ρ *is a function*

$$\rho : \Lambda \to \mathcal{P}(X) \tag{2}$$

The function ρ indicates the crisp realization that represents the imprecise property for each restriction level. As an example, the RL-representation for an atomic imprecise property defined by a fuzzy set A is the pair (Λ_A, ρ_A), where Λ_A is obtained using equation (1), and $\rho_A(\alpha) = A_\alpha \ \forall \alpha \in \Lambda_A$.

From now on, we shall consider that the set X on which the imprecise property is defined is implicit.

Definition 2.3 *Given an imprecise property* P *represented by* (Λ_P, ρ_P), *we define the set of crisp representatives of* P, Ω_P, *as*

$$\Omega_P = \{\rho_P(\alpha) \mid \alpha \in \Lambda_P\} \tag{3}$$

For an atomic property A, the set of crisp representatives Ω_A is the set of significant α-cuts of A, as we have seen. However, notice that in definition 2.2 there is no restriction about the possible crisp representatives for non-atomic properties. In particular, as a consequence of operations, they don't need to be nested, so the final RL-representation of a derived property is not always equivalent to the α-cut representation of fuzzy sets.

In order to define properties by operations, it is convenient to extend the function ρ to any RL $\alpha \in (0,1]$.

Definition 2.4 *Let (Λ, ρ) be a RL-representation with $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ verifying $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$. Let $\alpha \in (0,1]$ and $\alpha_i, \alpha_{i+1} \in \Lambda$ such that $\alpha_i \geq \alpha > \alpha_{i+1}$. Then*

$$\rho(\alpha) = \rho(\alpha_i) \tag{4}$$

Equation 4 remarks a very important idea discussed before: though Λ defines a minimum set of restriction levels employed in order to represent a certain imprecise property, the representation exists for the rest of possible restriction levels. Indeed, to specify Λ for every property has just the practical purpose of limiting the space and time needed to store and operate with RL-representations.

Definition 2.5 *Two RL-representations (Λ, ρ) and (Λ', ρ') are* equivalent*, noted $(\Lambda, \rho) \equiv (\Lambda', \rho')$, iff $\forall \alpha \in (0,1]$*

$$\rho(\alpha) = \rho'(\alpha).$$

Two representations are equivalent if they are representing equivalent properties, in the sense that every object in X verify both properties in the same way. Let us remark that this is an equivalence relation. In particular, if $\Lambda \subseteq \Lambda'$ then $(\Lambda, \rho) \equiv (\Lambda', \rho)$, i.e., if we add restriction levels to the RL-set and we define the crisp representatives in the new levels by using the extension in definition 2.5, we obtain an equivalent representation of the same property.

Definition 2.6 *Two RL-representations (Λ, ρ) and (Λ', ρ') are* equal*, noted $(\Lambda, \rho) = (\Lambda', \rho')$, iff $\Lambda = \Lambda'$ and $\rho = \rho'$.*

2.2 INTERPRETATION IN TERMS OF EVIDENCE

Given a RL-representation (Λ_A, ρ_A) for an atomic property A, the values of Λ_A can be interpreted as values of possibility of a possibility measure defined $\forall \rho_A(\alpha_i) \in \Omega_A$ as

$$Pos(\rho_A(\alpha_i)) = \alpha_i. \tag{5}$$

Following this interpretation we define a basic probability assignment in the usual way:

Definition 2.7 *Let (Λ, ρ) be a RL-representation with crisp representatives Ω. The associated probability distribution $m : \Omega \to [0,1]$ is*

$$m(Y) = \sum_{\alpha_i \mid Y = \rho(\alpha_i)} \alpha_i - \alpha_{i+1}. \tag{6}$$

The basic probability assignment m_F gives us information about how representative of the property F is each crisp set in Ω_F. For each $Y \in \Omega_F$, the value $m_F(Y)$ represents the proportion to which the available evidence supports the claim that the property F is represented by Y. From this point of view, a RL-representation can be seen as a basic probability assignment in the sense of the theory of evidence, *plus a structure indicating dependencies between the possible representations of different properties.* We'll be back to this point in the next section.

From m_F it is possible to obtain an object-centered summary μ_F of the information available about the imprecise property m_F as a fuzzy set, as follows:

Definition 2.8 *Let $F = (\Lambda_F, \rho_F)$ be a RL-representation on X. The evidence that $x \in X$ verifies the property F is*

$$\mu_F(x) = \sum_{Y \in \Omega_F \mid x \in Y} m_F(Y) \tag{7}$$

By definition, μ_F is a fuzzy set that summarizes the evidence that x verifies F for every $x \in X$ as the addition of the basic probability assignments of those representatives where x appears. This is useful since it is easier to understand for those familiar with fuzzy sets, though information about how the representatives are distributed on restriction levels (and hence, crucial information about dependencies between representations) is lost. In other words, both representations are not equivalent in general. However, as a particular case, for any atomic property A and $\forall x \in X$

$$\mu_A(x) = A(x). \tag{8}$$

i.e., for atomic properties represented by means of fuzzy sets, both RL-representations and fuzzy sets are equivalent.

2.3 OPERATIONS

Operations on RL-representations are based on the following ideas:

1. There is a single membership scale, i.e., let F and G be fuzzy subsets of X, then $F(x) = G(x')$ means that x is F just like x' is G. Hence, it makes sense to consider the representation of a knowledge base in a certain restriction level as a knowledge base consisting of the representation in that RL of all the properties and concepts that form the KB.

2. Crisp operations are extended to imprecise properties by operating in each restriction level. In general, let $f : \mathcal{P}(X)^n \to \mathcal{P}(X)$ be a crisp operation. Then, f is extended to RL-representations as follows: let $(P_1, \ldots, P_n)$ be imprecise properties defined on X with P_i represented by a RL-representation $(\Lambda_{P_i}, \rho_{P_i})$. Then, $f(P_1, \ldots, P_n)$ is represented by $(\Lambda_{f(P_1,\ldots,P_n)}, \rho_{f(P_1,\ldots,P_n)})$ where

$$\Lambda_{f(P_1,\ldots,P_n)} = \bigcup_{1 \le i \le n} \Lambda_{P_i} \tag{9}$$

and, $\forall \alpha \in \Lambda_{f(P_1,\ldots,P_n)}$,

$$\rho_{f(P_1,\ldots,P_n)}(\alpha) = f(\rho_{P_1}(\alpha), \ldots, \rho_{P_n}(\alpha)) \tag{10}$$

The representation in a certain RL of the extension of a certain crisp operation is the result of that operation on the representatives in the same RL of the arguments involved.

On this basis, the logical operations between imprecise properties represented by RL-representations can be defined by extending the corresponding set operators in their extensive representations. Conjunction and disjunction are extended as follows: let P, Q be imprecise properties represented by (Λ_P, ρ_P), (Λ_Q, ρ_Q). Then, $P \wedge Q$ and $P \vee Q$ are imprecise properties represented by $(\Lambda_{P\wedge Q}, \rho_{P\wedge Q})$ and $(\Lambda_{P\vee Q}, \rho_{P\vee Q})$, respectively, where

$$\Lambda_{P\wedge Q} = \Lambda_{P\vee Q} = \Lambda_P \cup \Lambda_Q \tag{11}$$

with, $\forall \alpha \in (0, 1]$,

$$\rho_{P\wedge Q}(\alpha) = \rho_P(\alpha) \cap \rho_Q(\alpha) \tag{12}$$

and

$$\rho_{P\vee Q}(\alpha) = \rho_P(\alpha) \cup \rho_Q(\alpha) \tag{13}$$

Imprecise properties derived from atomic properties by using $\wedge$ and $\vee$ only are compatible with fuzzy sets when these operations are performed via minimum and maximum, in the sense that the corresponding RL-representation yields the usual nested α-cut representation of fuzzy sets.

However, this is not true in general when negation is employed. We define the negation operator as follows:

Definition 2.9 *Let P be an imprecise property represented by (Λ_P, ρ_P). Then, $\neg P$ is an imprecise property represented by $(\Lambda_{\neg P}, \rho_{\neg P})$, where*

$$\Lambda_{\neg P} = \Lambda_P \tag{14}$$

and, $\forall \alpha \in (0, 1]$,

$$\rho_{\neg P}(\alpha) = \overline{\rho_P(\alpha)} \tag{15}$$

where $\overline{Y}$ is the usual set complement of a crisp set Y.

The rationale behind this definition is that if $\rho_P(\alpha)$ is the representative of P in the restriction level α (i.e., under the restriction α), then it is reasonable to consider that, under the same restriction, the representative of $\neg P$ is $\overline{\rho_P(\alpha)}$. As an example, consider the atomic properties A and B defined by the following fuzzy sets:

$$A = 1/x_1 + 0.8/x_2 + 0.5/x_3 + 0.4/x_5$$

$$B = 0.9/x_1 + 0.6/x_3 + 0.5/x_4$$

Then we have $\Lambda_A = \Lambda_{\neg A} = \{1, 0.8, 0.5, 0.4\}$ and $\Lambda_B = \Lambda_{\neg B} = \{1, 0.9, 0.6, 0.5\}$. The RL-set for any operation between A and B is the union $\Lambda_A \cup \Lambda_B$. In table 1 we can see several properties derived from A and B using $\wedge$, $\vee$, and negation. Let us remark that:

- As we can see, the representation of the negation of an atomic property such as $\neg A$ (and $\neg B$) is not the set of α-cuts of the complement of the corresponding fuzzy set (e.g. the fuzzy set $\overline{A}(x) = 1 - A(x)$). However, for every atomic (or compatible with fuzzy sets) property A,

$$\mu_{\neg A}(x) = \overline{A}(x) = 1 - A(x) \tag{16}$$

  i.e., the summary of the RL-representation of $\neg A$ is the fuzzy complement of the fuzzy set A by means of the standard negation $n(x) = 1 - x$. This means that, from the point of view of operations on degrees, both are equivalent. However, the structure of the RL-representation of $\neg A$ and the RL-representation obtained from the fuzzy set $\overline{A}$ are not the same.

- We have an example of a property, $A \wedge \neg B$, in which the different crisp realizations corresponding to RL's are not nested, i.e., the representation does not correspond to the representation of (i.e., compatible with) a fuzzy set.

α	$\rho_A(\alpha)$	$\rho_{\neg A}(\alpha)$	$\rho_B(\alpha)$	$\rho_{\neg B}(\alpha)$	$\rho_{A\wedge\neg B}(\alpha)$	$\rho_{B\wedge\neg A}(\alpha)$
1	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\emptyset$	X	$\{x_1\}$	$\emptyset$
0.9	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\emptyset$	$\emptyset$
0.8	$\{x_1,x_2\}$	$\{x_3,x_4,x_5\}$	$\{x_1\}$	$\{x_2,x_3,x_4,x_5\}$	$\{x_2\}$	$\emptyset$
0.6	$\{x_1,x_2,x_3\}$	$\{x_4,x_5\}$	$\{x_1,x_3\}$	$\{x_2,x_4,x_5\}$	$\{x_2\}$	$\emptyset$
0.5	$\{x_1,x_2,x_3\}$	$\{x_4,x_5\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_5\}$	$\{x_2\}$	$\{x_4\}$
0.4	$\{x_1,x_2,x_3,x_5\}$	$\{x_4\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_5\}$	$\{x_2,x_5\}$	$\{x_4\}$

Table 1: RL-representation of several properties derived from two atomic properties A and B.

Let us remark that by following these ideas, we are not combining directly the evidences given by the basic probability assignments of the properties involved, but we are combining RL-representations. As we mentioned before, evidences are just part of the information contained in a RL-representation; the other part is a structure, given by the restriction levels, that indicates how the possible sets that are representing the properties correspond to each other, and hence how they can be combined. In other words, they introduce a structure of dependence between the possible representations. This idea is different from that of the theory of evidence, in which the usual way to combine the basic probability assignments is by considering them as probability distributions under the assumption of independence.

2.4 PROPERTIES

We introduce the atomic properties $\top$ and $\bot$ represented by any RL-representation equivalent to the following:

$$\Lambda_\top = \{1\} \quad (17)$$
$$\rho_\top(1) = X \quad (18)$$
$$\Lambda_\bot = \{1\} \quad (19)$$
$$\rho_\bot(1) = \emptyset \quad (20)$$
$$(21)$$

It is possible to show that operations between RL-representations verify all the usual properties, since the operations are performed via the corresponding set operations on crisp sets in each restriction level and the equivalence of RL-representations is defined in terms of equality of the representatives in all levels. In particular, let us point out the following properties:

1. Excluded Middle axioms: $A \wedge \neg A \equiv \bot$ and $A \vee \neg A \equiv \top$.
2. Double negation: $\neg\neg A \equiv A$.
3. De Morgan's laws: $\neg(\neg A \wedge \neg B) \equiv (A \vee B)$ and $\neg(\neg A \vee \neg B) \equiv (A \wedge B)$.

2.5 TRUTH ASSESSMENT AND IMPLICATION

We introduce the following definition:

Definition 2.10 *A RL-truth assessment $\mathcal{A}$ is a RL-representation (Λ,τ) on $\{0,1\}$ such that $\tau(\alpha) \in \{0,1\}$ $\forall\alpha \in \Lambda$.*

RL-truth assessments are the result of evaluating the truth of a logical expression involving properties represented by means of RL-representations. The values 0 and 1 mean *false* and *true*, respectively, in the usual way. Let us remark that the requirement $\tau(\alpha) \in \{0,1\}$ in definition 2.10 means that there is no uncertainty in the evaluation of the expression in every restriction level. From now on, given a well formed formula (wff) ϕ, we shall denote by (Λ_ϕ,τ_ϕ) the corresponding truth assessment. We shall denote by $\mathcal{T}$ and $\mathcal{F}$ the crisp truth assessments 1 and 0, respectively, i.e., $\tau_\mathcal{T}(\alpha) = 1$ and $\tau_\mathcal{F}(\alpha) = 0$ $\forall\alpha \in (0,1]$.

Let ϕ, ψ be two wff's. Then, we define the logical operations in the usual way:

1. $\Lambda_{\neg\phi} = \Lambda_\phi$, $\tau_{\neg\phi}(\alpha) = 1 - \tau_\phi(\alpha)$ $\forall\alpha \in \Lambda_{\neg\phi}$.
2. $\Lambda_{\phi\wedge\psi} = \Lambda_\phi \cup \Lambda_\psi$, $\tau_{\phi\wedge\psi}(\alpha) = \min\{\tau_\phi(\alpha),\tau_\psi(\alpha)\}$ $\forall\alpha \in \Lambda_{\phi\wedge\psi}$.
3. $\Lambda_{\phi\vee\psi} = \Lambda_\phi \cup \Lambda_\psi$, $\tau_{\phi\vee\psi}(\alpha) = \max\{\tau_\phi(\alpha),\tau_\psi(\alpha)\}$ $\forall\alpha \in \Lambda_{\phi\vee\psi}$.
4. $\Lambda_{\phi\to\psi} = \Lambda_\phi \cup \Lambda_\psi$, $\tau_{\phi\to\psi}(\alpha) = \tau_{\neg\phi\vee\psi}(\alpha)$ $\forall\alpha \in \Lambda_{\phi\to\psi}$.

It is easy to show that all the properties of logical equivalence hold. Let us remark the following:

- $\mathcal{A}_\phi = \mathcal{A}_{\neg\neg\phi}$.
- $\mathcal{A}_{\phi\wedge\neg\phi} = \mathcal{F}$.
- $\mathcal{A}_{\phi\vee\neg\phi} = \mathcal{T}$.
- Let A, B be imprecise properties on X. Then $\mathcal{A}_{A\to B} = \mathcal{T}$ iff $\forall\alpha \in (0,1]$, $\tau_{A\to B}(\alpha) = 1$ iff $\forall\alpha \in (0,1]$ $\rho_A(\alpha) \subseteq \rho_B(\alpha)$ iff $\neg A \vee B \equiv \top$.

3 CONCLUSIONS

The representation proposed in this paper is a generalization of fuzzy sets. More specifically, is a generalization of the usual representation of fuzzy sets on the basis of a collection of alpha-cuts. Fuzzy sets are employed for representing *atomic* properties, i.e., those properties that are not defined in terms of logical operations on other properties. Properties defined as logical expressions involving other properties are called *derived* properties. The kind of derived properties we can obtain depend on the logical operations employed.

Properties derived by using only conjunction and disjunction are also represented by a collection of alpha-cuts, and correspond to the usual operations on fuzzy sets via minimum and maximum, respectively. However, in general, the crisp sets that form the representation are not forced to verify the usual inclusion property of alpha-cuts with respect to the values of alpha. This difference gives to this representation the possibility to represent imprecise properties and, at the same time, to keep properties involving negation that are not verified by fuzzy sets when using maximum and minimum.

An inconvenience of RL-representations is that it may be difficult to interpret the information contained in the representation when it does not correspond to the alpha-cut representation of a fuzzy set. We consider that this representation is specially useful when the properties are intermediate results in a process, and not the final information to be provided to a human expert. As an example, in [3], RL-representations of imprecise numbers are proposed (they have been proposed also in [1] under the name *gradual numbers*), with the remarkable property that their imprecision does not necessarily increases when arithmetic operations are performed; the maximum imprecision is determined by the number of restriction levels considered and its associated degree. A RL-number associates a single number to each restriction level. Though RL-numbers (gradual numbers) solve a classical problem of fuzzy arithmetic, they are more difficult to interpret. On the contrary, fuzzy numbers are easier to interpret, but they are well suited to represent imprecise intervals rather than imprecise values [1, 3].

In order to solve the problem of interpretation, we have proposed to summarize the evidence contained in a RL-representation F as a fuzzy set μ_F. In [3] we propose to approximate μ_F by a linguistic label with a clear semantic in order to obtain a linguistic, representative summary of the information of F. However, F and μ_F are not equivalent, and the operations on the former cannot be translated to operations on the latter in general. The summary μ_F is obtained from a basic probability assignment m_F defined on the crisp sets that form the representation by considering that the values of α associated to each restriction level form a possibility distribution under the theory of evidence. This is another way of summarizing the information in F. Again, m_F is not equivalent to F in general since, as it was the case with μ_F, the structure formed by the restriction levels, that defines a set of dependencies that indicate what crisp representatives of each argument are to be operated, is lost.

An interesting feature of RL-representations (resp. gradual elements in a fuzzy set) is that imprecision and uncertainty are clearly separated. The imprecision is represented in the number of different crisp representatives of the property that appear in the representation. In particular, a property represented by the same crisp set in all the restriction levels is a property that can be represented by a single restriction level and, hence, is a crisp property. On its turn, uncertainty can be represented in each restriction level by a probability distribution.

This discussion is suggested in [1] in terms of fuzziness (given by the set of levels) and imprecision (in each level). In addition, the negation considered in [1] is coherent with fuzzy negation, i.e., gradual representations are completely equivalent to fuzzy sets both for atomic and derived properties. The fuzzy set corresponding to a "gradual representation" is obtained by a max-sup operation on the set of representatives (called *gradual elements*).

As a final remark, t-norms and t-conorms other than maximum and minimum are not considered in order to perform logical operations, but can be used as aggregation operators in order to obtain the fuzzy sets describing atomic properties.

At this moment we are working on applying RL-representations to different problems where keeping the classical properties involving negation while working with maximum and minimum is important.

References

[1] D. Dubois and H. Prade. Gradual elements in a fuzzy set. *Soft Computing*, 12:165–175, 2008.

[2] D. Sánchez, M. Delgado, and M.A. Vila. A restriction level approach to the representation of imprecise properties. In *IPMU 2008*, Accepted for publication.

[3] D. Sánchez, M. Delgado, and M.A. Vila. RL-numbers: An alternative to fuzzy numbers for the representation of imprecise quantities. In *WCCI 2008*, Accepted for publication.

TRES PREGUNTAS DE RESPUESTA INCIERTA SOBRE LA LÓGICA FUZZY

Alejandro Sobrino Cerdeiriña[1], Gracián Triviño Barros[2]
[1] Dpto. de Lógica y Filosofía Moral. Universidad de Santiago de Compostela. alejandro.sobrino@usc.es
[2] European Centre for Soft Computing. Mieres, Asturias. gracian.trivino@softcomputing.es

Resumen

La lógica borrosa es un saber que se está desarrollando, en algunos ámbitos, como alternativa a la lógica clásica. Como tal, suscita preguntas acerca de su naturaleza y de su estatus. La contestación a esas cuestiones suele ser materia de debate, a la espera de progresos categóricos que permitan pasar de la opinión a la definición. Este trabajo pretende contribuir a esa discusión.

Palabras clave: lógica borrosa, computación borrosa, objetos vagos.

1ª PREGUNTA

Algunos teóricos prefieren decir que la lógica fuzzy no es tanto una lógica como una tecno-lógica; esto es, una lógica que alcanza su legitimación en el éxito de sus aplicaciones. ¿Debería ganar 'estatuto lógico' la lógica fuzzy? ¿Es una lógica eminentemente aplicada?

En el trabajo "Local and fuzzy logics", Bellman y Zadeh [1] afirman que la lógica borrosa tiene una vocación fundamentalmente aplicada y que la demostración de propiedades metalógicas es una cuestión más bien periférica que central a la misma. Es un hecho histórico que la lógica fuzzy fue razonablemente bien admitida desde el principio en las Escuelas de Ingeniería, pero en general mal considerada en las Facultades de Matemáticas.

Los ingenieros suelen decir que, con ella, modelan situaciones de manera adecuada que anteriormente no podían tratar o cuya gestión involucraba un gasto o esfuerzo considerable. Los matemáticos, que la lógica fuzzy es, en buena medida, una lógica construída *ad hoc* para resolver problemas concretos, pero no para capturar los principios de algún tipo de racionalidad; ni siquiera de aquella que dice aprehender, la razón aproximada. No es una lógica categórica, sino variable, contextual; vale decir, nada definitiva.

Los ingenieros usan la lógica borrosa atendiendo, al menos, a tres criterios:

- Al contrario de lo que ocurre cuando se modela la incertidumbre con métodos probabilísiticos, la lógica fuzzy permite representar situaciones complejas usando pocas reglas. A menudo, basta con representar algunas heurísticas para obtener resultados bastante exitosos.
- Las soluciones, aún no siendo precisas o bastante precisas, como ocurre con el razonamiento probabilístico, aportan información suficiente para tomar una decisión razonable. Están en consonancia con el *dictum* de K. Popper de que "aunque la claridad es un valor en si misma, no sucede lo mismo con la exactitud y la precision…En ciencia no se debe ser más preciso de lo que la situación requiere" [2].
- Las soluciones, por ser aproximadas, suelen ser robustas: hay un margen de tolerancia entre lo que es y lo que no es, si bien, p. ej., es difícil demostrar a priori la estabilidad de un controlador fuzzy.

En general, estos criterios responden a la máxima que Zadeh enunció como el principio de incompatibilidad: *"as the complexity of a system increases, our ability to make a precise and yet significant statement about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics"*

Los matemáticos, en cambio, velan por la salud interna de la herramienta y por la precisión y estabilidad de sus soluciones. Situados en la tradición clásica de la preocupación por las consecuencias que se siguen de no distinguir lenguaje de metalenguaje o de los emparejamientos de cadenas de caracteres que no terminan nunca, observan con preocupación soluciones que no se sometan a los test, reales o imaginarios, más exigentes de la racionalidad precisa. Las discusiones sobre conjuntos que son al mismo tiempo elemento o sobre bucles que se llaman a si mismos, tienen un aspecto impecable, aunque no suelen mostrar incidencia en problemas reales, donde es difícil encontrarse con el conjunto máximamente grande o con procedimientos recursivos sin condición límite.

Pero la lógica no surgió del ejercicio racional puro, sino para solucionar cuestiones humanas, como dirimir los conflictos que se producían en el ágora o dotar a la ciencia de un sistema de justificación de sus proposiciones. En este cometido, Aristóteles, su creador, la concibe como una disciplina deductiva, donde las conclusiones se derivan con necesidad de las premisas. En su obra se pueden rastrear indicios de que esta ciencia estaría formada por un conjunto de sentencias Σ que satisfacen, entre otros, los siguientes postulados:

1. Postulado de realidad: Las sentencias de Σ lo son acerca de algún ámbito de entidades reales
2. Postulado de verdad. Toda sentencia de Σ es verdadera
3. Postulado deductivo. Todo lo que puede ser deducido de un subconjunto de Σ puede ser deducido también de Σ

-Σ es cerrado bajo consecuencia lógica-

Como materia de definición, estos postulados son claros, pero suscitan dudas si se pretenden aplicar a ámbitos reales, como el que cubre la ciencia: en numerosas ocasiones no se verifica que oraciones acerca de entidades reales (1) sean verdades incontestables (2). Las verdades autoevidentes pertenecen a ámbitos extraterrenales, como el de la matemática, pero no casan bien con las proprosiciones científicas que, por serlo, como diría Popper, llevan fecha de caducidad, y mucho menos con oraciones del lenguaje ordinario que deescriben escenarios comunes, banales o no. Desde la lógica clásica se obvia esta situación imponiendo un proceso de regimentación: si el lenguaje no es preciso, hagámoslo y, si no podemos, esperemos a hacerlo; mientras tanto, excrétese a la vaguedad del lenguaje, científico o no, y de su reconstrucción lógica. Para el lenguaje de la matemática, evitar la vaguedad es legítimo, pero Aristóteles no estaba pensando en esta ciencia cuando propuso su silogística.

Son los matemáticos y no los ingenieros los que suelen reclamar estatuto lógico a cualquier propuesta de justificación racional de un saber. Instalados en dominios o contextos muy generales, velan porque la soluciones particulares sean casos de las soluciones más generales. ¿Qué determina el estatuto lógico de un saber? Desde una perspectiva clásica, determinar la verdad o falsedad de sus proposiciones y transmitir la verdad en las cadenas de razonamiento.

En el caso de la lógica o de la matemática, la verdad de las proposiciones surge de su forma, de su evidencia. Así, $p \vee \neg p$ es un enunciado incontestablemente verdadero o tautológico, mientras que $p \vee q$ puede ser verdadero o no. Desde luego, $p \wedge \neg p$ es siempre falso. En la Física, la verdad suele provenir de una hipótesis recional confirmada por algún experimento. P. ej. la primera ley de Newton que dice que 'en la ausencia de fuerzas exteriores, todo cuerpo continúa en su estado de reposo o de movimiento rectilíneo uniforme a menos que actúe sobre él una fuerza que le obligue a cambiar dicho estado'. Pero en el lenguaje natural hay proposiciones que no son ni verdaderas ni falsas; no hay un criterio de determinación que permita confirmarlas o falsarlas de una vez. Así, ocurre que a veces, no se puede determinar con precisión si un coche va rápido o lento o si un individuo es inteligente o lerdo. En ese caso, para determinar la verdad de las proposiciones no vale una semántica definicionista (al estilo de la matemática) o verificacionista (similar a la de la física), sino posibilista, de valoración contextual, de elección entre las alternativas racionales que se nos aparecen a nuestra capacidad normal de evaluar problemas; vale decir, no hay una evaluación absoluta, sino relativa, gradual. Se han ofrecido al menos tres maneras de justificar la representación de una proposición de significado impreciso con un grado: (a) atendiendo a su parecido con una proposición prototípica, acerca de la cual no hay dudas sobre su valoración precisa -grado de similitud-; (b) atendiedo a la confianza o riesgo con la que se utilizaría para hacer una elección -grado de preferencia-, (c) atendiendo a las dudas o incertidumbre que generaría tomarla como un representante preciso -grado de incertidumbre-. La primera sería propia de una semántica del prototipo; la segunda, de una semántica del voto; la tercera, sería propia de una semántica posibilista.

Una lógica lo es, en el sentido clásico, cuando transmite verdad en las cadenas de razonamiento. Tarski aprehendió este concepto con la noción de consecuencia lógica, a la que caracterizó con tres propiedades: inclusión, monotonicidad y una suerte de corte o punto fijo. Chakraborty [3] generalizó esta noción para el caso graduado, con las propiedades de α-reflexividad, monotonicidad gradual y corte gradual y Castro et al. [4] mostraron que la propuesta de Chakraborty no cubría todos los métodos de razonamiento monotónicos en lógica polivalente, proponiendo definiciones alternativas de consecuencia fuzzy. En el marco de la regla composicional de inferencia, sugerida por Zadeh para capturar el Modus Ponens aproximado, Trillas [5] probó que, si se usa el condicional de Reichenbach para representar la regla que es premisa, la única T-norma que asegura consecuencias, en el sentido de Tarski, es la acotada. Estos son algunos resultados metateóricos que deben hacer creer a los lógicos más puristas que la lógica fuzzy está ganando estatuto lógico. De momento, quizás de modo no tan notable como en el caso clásico, pero se debe esperar al menos tanto como se ha aguardado por los resultados metalógicos de la lógica clásica para sacar alguna conclusión acerca de una posible deficiencia sustancial de la lógica fuzzy. Y es un hecho que, si la carrera para ganar estatuto lógico ha sido, en el caso clásico, gradual (no se han probado todos los metateoremas de una vez), no es razonable exigirle otra cosa a la lógica fuzzy. En esta espera, un elemento tranquilizador para la lógica borrosa, como para las lógicas polivalentes en general, es que cuando tratan con proposiciones precisas se comportan como una lógica bivalente; es decir, la incluyen como un caso particular suyo (tesis en la que difieren, sino en la descripción, sí en el foco, los lógicos clásicos).

En la tarea de determinar el valor de verdad de una proposición vaga, parece que una condición necesaria para representar adecuadamente el significado de un predicado borroso P es determinar el uso que se hace del mismo en un contexto determinado. El contexto permite delimitar el universo de discurso y el uso marcar, de modo razonable, para cada x, el grado en el que 'x es P'. Respecto a las proposiciones moleculares, Trillas [5] generalizó la definición de complemento o negación como una función n: [0,1] → [0,1] que es estrictamente

decreciente (n(a)>n(b) si a<b), continua y que satisface las condiciones clásicas de límite (n(1)=0; n(0)=1) e involución (n(n(a))=a. y probó este interesante resultado metateórico: $T_0(a, b) \leq T(a, b) \leq T_2(a, b)$ y $S_2(a, b) \leq S(a, b) \leq S_0(a, b)$, siendo T_0 y S_0 el producto y la suma acotada y T_2 y S_2 el min y el max; es decir, que todas las formas de agregar valores graduados están entre la t-norma mayor (max) y la t-conorma menor (min). Respecto al condicional, se conocen en la actualidad más de 50 funciones que tratan de modelar la relación entre antecedente y consecuente. Algunas cuestiones que se suscitan son: primero, como ordenarlas o clasificarlas para poder manejarlas según que propósito. En segundo lugar, si todas ellas tienen correlato en el mundo, porque en el caso fuzzy sucede lo contrario que en el mundo clásico: en la bivalencia, hay una única definición para un lenguaje vario, que hay que regimentar si se sale del corsé. En particular, ha sido muy discutido el carácter paradójico de la definición del condicional material. En lógica borrosa, en cambio, como se puede observar en Sobrino y Fernández [6] es difícil buscar un correlato claro en el lenguaje común para algunas de las definiciones de condicional fuzzy.

Volvamos con los ingenieros. Contemplan a la lógica fuzzy, no como las leyes posibles del pensamiento impreciso, sino como una herramienta que proporciona soluciones a problemas que involucran imprecisión. Esto le ha llevado a veces a hacer propuestas demasiado *ad hoc*, cuya explicación, cuando se le pide, es de sabor estrictamente funcional. Es cierto que la explicación funcional no tiene estatuto alguno en las ciencias deductivas, pero sí en ciencias experimentales, como la biología. A grandes rasgos, una explicación funcional es aquella que justifica un hecho o suceso en términos de su efecto [7]. Puede que los ingenieros conciban así a su tarea y el éxito de una aplicación sirva de justificación a la herramienta que utilizan.

Los ingenieros tuvieron una parte importante de responsabilidad en el auge de la lógica fuzzy. Nacida como una lógica cuestionada, aunque atractiva, por dar cuenta del razonamiento aproximado de manera diferente a como lo hacían otras ramas de la matemática (probabilidad, estadística, etc.), debía intentar ganar crédito de manera inmediata para mantenerse. Dado que el teórico le era negado por los matemáticos fundadores de la I.A. (MCarthy, Minsky,…), tenía que mostrar su fortaleza en las aplicaciones. Plantas de cemento, potabilizadoras de agua, sistemas de frenada en trenes de y, sobre todo, la electrónica de consumo doméstico (tensiómetros, estabilizadores de imagen en cámaras de video,…), fueron aplicaciones conocidas y celebradas. Estos logros, unidos a la novedad de lo fuzzy y a lo sugerente de su denominación, hicieron que la lógica borrosa se ganara un hueco en las escuelas de ingeniería. La eclosión de productos comerciales 'fuzzy powered' llegó, pero con ellos también alguna perplejidad: la t-norma min, común a algunas lógicas polivalentes, se mostró como la de más éxito en el modelado de reglas borrosas empleadas en el control de procesos. En cambio, otras maneras más complejas, pero igualmente razonables de definir el razonamiento aproximado, no han encontrado aún realizaciones acordes con la racionalidad que parecen expresar. Este dato ha sido utilizado para mostrar la supuesta 'debilidad' teórica de algunos de los ámbitos de aplicación más celebrados y exitosos.

En esta encrucijada, recordamos lo que dijo Carnap [8]: *"In logic there are no morals. Everyone is at liberty to build up his own logic, i. e., his own formal language, as he wishes. All that is requiered of him is that, if he wishes to discuss it he must state his methods clearly and give rise syntactical rules instead of philosophical arguments"*. La primera parte de su aseveración parece atinada, pero no la segunda. La tendencia hacia la formalización será una consecuencia de la madurez de la lógica fuzzy, pero eso vendrá soportado o provocado por discusiones filosóficas. Con la formalización, la lógica fuzzy será cada vez menos borrosa, intentará ser una lógica precisa de lo impreciso; una lógica que explicará la tartamudez sin tartamudear. Pero eso, de lograrse absolutamente, sería una contradicción en términos. Si lo impreciso tuviese una lógica precisa, la imprecisión ya no sería tal, al ser desvelada como precisa. Esta tensión sólo se ve aliviada por el debate acerca de qué aspectos del lenguaje y sus significados quedan fuera de las fórmulas, más allá de la precisión. La lógica fuzzy llegará a formalizarse más y más en la medida en que se vayan definiendo sus límites mediante la descripción de restricciones que la acoten. Pero otra cosa es el lenguaje natural. En él, no es posible introducir límites sin desvirtuarlo.

2ª PREGUNTA

¿Hay un modo de computar genuinamente fuzzy o simplemente se consiguen simular con máquinas de Turing computaciones imprecisas? ¿Dice esto algo a favor de la increíble capacidad de las Máquinas de Turing para simular lo que está alejado incluso de su propia naturaleza, como la imprecisión?

Uno de los resultados más famosos de la informática moderna dice que una función es computable si es Turing-computable (tesis de Church-Turing). Desde un punto de vista intuitivo, se dice que un problema es computable si es algorítmico. La MT procesa cualquier problema que se pueda formular en términos algorítmicos; esto es, como un conjunto de instrucciones finitas, precisas y que alcanzan una solución en un tiempo limitado.

Para cada problema algorítmico existe una MT que lo resuelve. Para ello, la MT adopta la siguiente configuración: el problema debe ser formulado en términos de un alfabeto finito, la máquina puede situarse

en una serie de estados y en función del símbolo que lea y del estado en el que se encuentre, emite una respuesta. Hay dos características a destacar: (a) la cinta, con infinitas celdas, admite símbolos del alfabeto (más el blanco) de uno en uno; no hay solape entre ellos; (b) el conjunto de estados es finito y discreto. Esto configura una computación lineal y precisa. El problema a computar, si tiene solución, hará que la MT se detenga en un estado final. Si la MT se para en un estado que no es final, el problema puede carecer de solución. Si no se detiene y, no obstante, computa una y otra vez, el problema es indecidible. Cada problema concreto reclamará su algoritmo de solución; esto es, una MT que se detenga en un estado de aceptación.

Todas las MT pueden ser simuladas por otra MT: la MT binaria, que representa un símbolo o un estado por un dígito o un conjunto de dígitos binarios. Así, una instrucción como: $10 \rightarrow 1000 1i$, indicaría que la máquina se encuentra en el estado interno 1, lee 0 en la cinta, pasa al estado interno 1000 (8), escribe 1 y se mueve hacia la izquierda un paso (la cinta se mueve hacia la derecha). Con la MT binaria es posible simular cualquier proceso efectuado por una MT no binaria. Esta es una característica señalada: cualquier MT puede ser simulada por una más general.

Pero además, la MT puede simular problemas para los que de entrada no está diseñada. Así, se pueden procesar problemas que no son secuenciales, sino paralelos, poniendo varias MT a computar. También permite simular computaciones imprecisas, emitiendo soluciones que parecen aproximadas en vez de precisas; esto es, outputs discretos que, agrupadas en una secuencia temporal ínfima, simulan una solución continua. El efecto es parecido al de los 'dibujos animados': láminas estáticas que pasan con tal rapidez que, de no saberlo, diríamos que representan a figuras que se mueven. Un aspecto destacable de las MT es su gran capacidad de simulación.

La pregunta que sobreviene ahora es la siguiente: ¿hay computaciones genuinamente fuzzy o simplemente simulan lo borroso? A tenor de lo dicho antes, la respuesta podría ser, sencillamente, que no hay un modo estrictamente fuzzy de computar, que, como argumenta Moraga, las computaciones que realiza una MT fuzzy las puede hacer una MT clásica [9]. Pero la cuestión merece algo más de debate. Consideremos dos contextos, uno teórico y otro de aplicaciones. Desde un punto de vista teórico, las computaciones fuzzy existentes son simulaciones hechas por MT clásicas para imitar los aspectos de vaguedad o imprecisión considerados. Hay dos formas de conseguir esto: (a) ponderando el alfabeto que entra/sale en la cinta o (b) dando preferencias a los estados. Ponderar o dar pesos quiere decir asociar grados. En ese caso, la MT debe disponer de dos módulos: uno de computación clásica y otro de cálculo con números reales, que dará la parte 'aproximada' de la computación. Desde esta perspectiva, la computación fuzzy es una computación clásica con cálculo de grados.

Pero desde un punto de vista aplicado estamos interesados (como en los dibujos animados), no en saber cuantas láminas se superponen y a qué velocidad para conseguir una sensación determinada (una animación cualquiera), sino en que esa sensación se aparezca ante nosotros con las connotaciones más realistas posibles. Desde esta perspectiva, no requerimos analizar cómo se computa, sino si el resultado simula la manera en que, como humanos, resolveríamos el problema y acerca del cual concluiríamos una solución que, de basarnos en información parcial o vaga, debería tener una apariencia aproximada. Así visto el tema, la computación es genuinamente fuzzy porque simula bien una situación fuzzy, independientemente de si la resuelve o no de una manera 'naturalmente' fuzzy. Un avión vuela, aunque no es un pájaro; si bien volar es una categoría que, modo natural, sólo se le puede atribuir pájaro, pues el avión es un producto manufacturado. La pregunta sobre si una simulación perfecta es tan natural como la propiedad que simula nos retrotrae a la vieja y honorable pregunta filosófica de qué es el ser. Esta cuestión excede el ámbito de este trabajo y no abundaremos en ella.

¿Cómo sería una computación fuzzy no artificial, sino natural? Es difícil aventurarlo. De momento, no se entrevé un modelo teórico de computar diferente al de la MT, aunque hay al menos dos aproximaciones, no circunscritas a lo fuzzy, que pretenden ser diferenciadoras.

Una proviene de la mecánica cuántica, en la que los estados no se suceden de uno en uno, sino que pueden ser activados conjuntamente, dándose de manera natural un paralelismo computacional. El modelo cuántico es muy sugerente como tratamiento alternativo de la vaguedad, ya que en la física cuántica, un tipo especial de vaguedad –la indeterminación- es consustancial a esa teoría. Pero la vaguedad lingüística y la vaguedad cuántica parecen tener diferente naturaleza: así, mientras en lógica borrosa la unidad de explicación fundamental es el *fuzzy set*, del que se conoce su cardinalidad, aunque para la mayoría de sus elementos no es posible determinar si están o no en el conjunto; en la mecánica cuántica es el *poset*, del que se conoce, para cada elemento, si está o no está en el conjunto, aunque nunca la cardinalidad precisa del mismo. En consecuencia, está por ver si la computación cuántica ofrece un marco apropiado no ya para verificar si hay un método alternativo de computar, sino si ese método es apropiado para procesar instrucciones o algoritmos imprecisos. Otra alternativa resulta de considerar la computación de la MT en un espacio-tiempo relativista, no newtoniano. Así, Hogart [10] demostró que es posible llevar a cabo supertareas computacionales –esto es, algoritmos que involucran un número infinito de pasos- en un tiempo finito si se ejecutan en un espacio-tiempo relativista al que denominó de Malament-Hogart.

3ª PREGUNTA

¿Sólo el lenguaje es vago –como decía Russell- o hay también objetos vagos?

En 1923, Russell [11] publicó un trabajo seminal titulado 'Vagueness' y dejó un sentencia que se ha convertido con los años, a fuerza de repetirla, en una especie de truismo: "Sólo el lenguaje es vago, no los objetos". No obstante, esta frase no es una verdad obvia y merece discusión. La pregunta por los objetos nos acerca a un viejo problema metafísico, cual es el de si hay entes y qué naturaleza tienen. Pero esto nos retrotraería a los inicios de la filosofía. No es es este un objetivo de este trabajo, así que diremos, a continuación, desde qué perspectiva lo vamos a abordar. El tema de los objetos vagos puede ser considerado, al menos, desde tres puntos de vista diferentes:

- Hay objetos vagos porque, aún teniendo una naturaleza definida, no hay información suficiente o conocimiento adecuado para nombrarlos de manera precisa.
- Hay objetos vagos porque, aún pudiendo reducirlos a una caracterización precisa, ello no revelaría su verdadero ser.
- Hay objetos vagos porque tienen una naturaleza inerradicablente vaga, esencialmente imprecisa.

Mientras los dos primeros ámbitos atienden a la representación del objeto, con el tercero se alude a la naturaleza inexacta del objeto mismo a la luz del conocimiento científico actual; mientras los dos primeros refieren a maneras de verbalizar o representar la realidad, el tercero denota la pretendida naturaleza de esa realidad. Es decir, los dos primeros se estudian en el ámbito del lenguaje y de los modelos que, para diferentes propósitos, se construyen. El tercero pertenece al ámbito de la física y, en particular, al de la mecánica cuántica.

Veamos algo de este último. Un área donde se examina la vaguedad objetual es a propósito de los objetos físicos. Dos ejemplos manidos provienen de la geografía: la delimitación de una nube es asunto impreciso y decidir dónde empieza o término una montaña es imposible. Pero el ámbito en el que se ha discutido más este aspecto es en el mundo subatómico, donde algunas cosas son indeterminadas por naturaleza. Uno de los ejemplos más nombrados es el gato de Schröedinger. Imaginemos una caja cerrada donde está un gato, una botella con gas mortal y un mineral radiactivo que tiene el 50% de probabilidades de desintegrarse. Si lo hace, la botella se romperá, el gas se liberará y el gato morirá. Si no se activa, el gato seguirá vivo. Mientras no se abra la caja, sólo se puede decir que el gato está vivo y muerto a la vez. Esta descripción, que al normal juicio parece contradictoria, es una proposición verdadera de la mecánica cuántica y consecuencia de un juicio racional.

Incluso admitiendo que el conocimiento es representacional, la cuestión a plantear en el caso de la física cuántica es si las características vagas de nuestras representaciones son, en última instancia, eliminables o no. Sólo tenemos una manera de acceder al mundo subatómico, que es a través de instrumentos. Pero estos distorsionan la realidad que permiten observar, haciendo, respecto a alguna magnitud, indeterminado lo que esa realidad sea. En la mecánica cuántica, la vaguedad se da en los propios objetos (que no son sino su descripción). P. ej., consideremos la propiedad 'tener una dirección de spin x', que puede ser adscrito con significado a una colección de electrones. El predicado está bien definido, ya que los físicos conocen perfectamente qué requisitos debe satisfacer un electrón para tener una dirección de spin x. Haciéndo uso de instrumental adecuado, pueden encontrar un cierto número, digamos n, de electrones que satisfacen esa propiedad, y se obtiene el mismo número si el experimento se repite. Sin embargo, no hay forma de asegurar qué electrones de esa colección tienen esa propiedad o qué los electrones de la primera medición son los mismos que los del segundo experimento. La noción de conjunto estándar no vale agrupar esos casos, ya que, como diría Cantor, no constituyen una reunión de entidades distintas en nuestra intuición o en nuestro pensamiento. M. L. Dalla Chiara y T. di Francia han propuesto que los objetos cuánticos se agrupen en quasi-sets. De modo informal, un quasi-set es una colección de objetos acerca de los cuales se sabe si pertenecen o no pertenecen a la colección, pero no cuáles están en la colección.

Situemonos ahora en el ámbito de la representación formal del lenguaje. En Filosofía, la cuestión de si puede haber objetos vagos la planteó G. Evans [12] en un famoso artículo de sólo media página que tituló así. Pero bajo esa pregunta escondía otra más específica: si puede haber vaguedad en la relacion de identidad cuando ésta se atribuye a designadores rígidos que, según Kripke, son términos que *"designan siempre al objeto dondequiera que el objeto exista…Los nombres son designadores rígidos...."*. Bajo este considerando, Evans se plantea si puede ser *inderminado* (∇) que un designador rígido sea idéntico a si mismo. Y responde negativamente con la siguiente prueba:

(1)∇ (a=b)	*a=b* es indeterminado	(supuesto)
(2) $\lambda x[\nabla(a=x)]b$	*b* es indeterminadamente igual a *a*	(de 1)
(3) $\neg \nabla$ (a=a)	*a=a* es determinado	Axioma
(4) $\neg\lambda x[\nabla$ (x=b)]a	*a* no es indeterminadamente igual a *b*	(de 3)
(5) a≠b	*a* no es igual a *b*	(de 2 y 4)

El argumento de Evans parece mostrar que no puede haber identidades vagas: si es indeterminado si *a* es igual a *b*, se sigue son diferentes. Pero más bien parecen extraerse consecuencias en el sentido contrario -tal y como ha visto D. Lewis-: si bien Evans parte del supuesto de que se puede atribuir vaguedad a una identidad de designadores rígidos (∇ (a=b)), la derivación prueba que el supuesto es falso (a≠b); por tanto, que el punto de partida es inadecuado. Este contraargumento muestra, de modo paradigmático, lo inapropiado que es juzgar de

modo preciso situaciones imprecisas e intentar arrojar resultados concluyentes. La imprecisión debe ser considerada seriamente y con todas las consecuencias (usualmente, aproximadas, no exactas).

Volvamos al plano representacional de la vaguedad y, ahora, en concreto, a su reconstrucción lógica. La lógica, desde Aristóteles, reclamó el estatuto de ciencia. Si esa ciencia lo es sobre objetos ideales, puede ser una ciencia deductiva. En ese caso, el laboratorio está en el cerebro. Pero si es sobre objetos reales, la lógica, si aspira a ser una ciencia, debería aplicar el método científico. Aunque hay distintas variedades de proceder científicamente, el método hipotético-deductivo constituye uno de ellos. La lógica borrosa trata con predicados vagos y estos están emparentados con el lenguaje común, con el que razona la gente corriente. Por ello, su objeto de estudio no son entes ideales, sino asuntos terrenales, reales. Si la lógica borrosa aspira a ser una ciencia, debe tratar de conjugar hipótesis, contrastación, inferencia (aproximada) y predicción. En esto no debería diferenciarse de la estructura que presenta cualquier teoría científica.

En la reconstrucción neopositivista de la actividad científica, lo preferente no es la fidelidad de la representación, sino la capacidad de deducción, para hacer predicciones precisas. Se simplifica para ser preciso, y lo que no se puede precisar, se deja a un lado. Las teorías borrosas ponen más énfasis en la representación, en tratar de ser fieles a aquello que se estudia, que en la exactitud de lo que se infiere. Desaparecido el filtro regimentativo que convierte el lenguaje de observación impreciso en preciso y las reglas de correspondencia que proyectan un lenguaje natural en un cálculo artificial preciso, la lógica fuzzy apuest por un equilibrio entre lo que se representa y lo que supuestamente 'es', entre lo que idealmente se debería predecir y lo que se es capaz de decir, usualmente ajustado a las necesidades de una solución concreta.

Hay objetos de naturaleza más o menos imprecisos de los que, una caracterización precisa, no sería informativa. El Sorites proporciona un buen ejemplo. Se puede reducir, de modo simplicado, la calvicie a un asunto de pelos y decir que si alguien tien 0 pelos, es calvoy si tiene 10^{10}, no lo es. Pero hay un rango entre ambas cantidades que que puede hacer dudar; algunas personas dirían que sí y otras que no. Trillas ha ideado un brillante método para encontrar el punto medio donde una persona es calva y no calva a la vez; es decir, un punto a partir del cual un individuo es más calvo que no calvo o al revés. Pero el problema de querer solucionar la vaguedad con una medida precisa es que enseguida aparece una vaguedad de segundo orden que suscita la pregunta de por qué no elegir el número que representa la solución inmediatemante superior o inferior en la granularidad considerada.

En la lógica fuzzy, la vaguedad se hace recaer en los predicados, en las propiedades que se le atribuyen a los objetos, no en los objetos mismos. Y el problema es ser fiel en esa representación, como la física trata de serlo con la naturaleza, proponiendo modelos para aprehender el significado de la manera más oportuna para el problema considerado, tratando de generalizar las soluciones siempre que sea posible, haciendo hipótesis sobre su carácter explicativo y contrastando las concluisones proveídas con aquellas esperadas. Este proceso dialéctico correponde a la manera de proceder de las ciencias experimentales y, a nuestro juicio, la lógica fuzzy debe seguirlo. Como la lógica clásica, debe aspirar a verdades eternas, pero como la física, la realidad debe mostrarle que sólo las logra en un grado.

Agradecimientos

Financiación a cargo del proyecto HUM 2007-66607-C04-02 del MEC.

Referencias

[1] R.E. Bellman and L.A. Zadeh, "Local and fuzzy logic", in: G. Epstein, (ed.), *Modern Uses of Multiple-Valued Logic*. Reidel, Dordrecht, 1977, pp. 103-165.

[2] K. R. Popper, *Conjeturas y refutaciones. El desarrollo del conocimiento objetivo.* Paidós. 1994, Barcelona.

[3] M. K. Chakraborty, "Use of Fuzzy Set Theory in introducing graded consequence in Multiple-Valued Logic" in *Fuzzy Logic in Knowledge-Based Systems, Decision and Control*. M. M. Gupta & T. Yamakawa (eds.), North-Holland, 1988, pp. 247-257.

[4] J. L. Castro et al. "On consequence in approximate reasoning", *J. of Applied Non-Classical Logics*, 1994, 4, 91-103.

[5] E. Trillas et al., *Introducción a la lógica borrosa*. Ariel, 1995.

[6] A. Sobrino & S. Fernández "Hacia una lógica fuzzy experimental", *Actas del XII Congreso español sobre Tecnologías y Lógica Fuzzy*. Jaén, 2004, 651-656.

[7] J. Elster, *El cambio tecnológico. Investigaciones sobre la realidad y la transformación social*. Gedisa, 1990

[8] R. Carnap, *Logical Syntax of language*, Routledge and Kegan Paul, London, 1949.

[9] C. Moraga, "Towards a Fuzzy Computability?", *Mathware & Soft Computing*, 6 (1999), 163-172.

[10] M. Hogart, "Non-Turing Computers and Non-Turing Computability", *Philosophy of Science Association 1994*, vol 1, pp 126-138.

[11] B. Russell, "Vagueness", *Australian Journal of Philosophy and Psychology*, Vol. 1 (1923), 84–92.

[12] G. Evans, "Can There Be Vague Objects?", *Analysis*, 38, 1978, p. 208.

UNA REVISIÓN DEL CONCEPTO DE INTENCIONALIDAD EN UN MARCO APROXIMADO

Martín Pereira Fariña

Licenciado en Filosofía
Universidad de Santiago de Compostela
Miembro de la Red PLIR
España
mpereiraf@msn.com

Resumen

La intencionalidad es una de las características fundamentales de las acciones humanas, pero a su vez, una de las más oscuras. En ámbitos de conocimiento vago e impreciso, como el de la lógica borrosa (LB), donde el contexto tiene una función crucial, la intencionalidad de los diseñadores, en tanto que elemento estructural del mismo, queda plasmada en la mayoría de los modelos que se usan para representar sistemas borrosos.

Palabras Clave: lógica borrosa, contexto, intencionalidad, comprensión, racionalidad.

1 INTRODUCCIÓN

Desde sus orígenes, la LB ha sido una herramienta fundamentalmente orientada a la acción, como así lo demuestran sus múltiples aplicaciones industriales. Esto ya supone una diferencia de raíz respecto de la lógica clásica, puesto que esta está principalmente orientada a la deducción de leyes universales y necesarias.

La mayor parte de las acciones que diariamente realizan los humanos son calificadas como acciones intencionales, pues están sustentadas, motivadas por una intención particular. A pesar de ser esta una cuestión ampliamente aceptada, la naturaleza de la intencionalidad y el vínculo que mantiene con la acción es un tema ampliamente debatido, pues la intención permanece siempre oculta en el interior del agente. La única vía de acceso para un observador externo a la intención de un sujeto es la observación de las acciones y de la conducta del mismo, pero es una vía indirecta y no siempre totalmente transparente. El único que puede tener un acceso directo o primario a la intención que hay detrás de una acción es el agente mismo.

Partiendo de la idea de ese carácter orientado a la acción que tiene la LB, el objetivo de este trabajo es analizar de qué modo la intencionalidad de los agentes se manifiesta en dicho ámbito, de cómo influye en la gestión de la vaguedad de los enunciados imprecisos, de cómo repercute en el significado de las expresiones y de cómo esta afecta al diseño efectivo de cualquier sistema de LB.

El punto de partida es un análisis del amplio tratamiento filosófico que dicho concepto ha tenido, asumiendo los enfoques de von Wright. Para las cuestiones relacionadas con los enunciados vagos y los aspectos del contexto se ha recurrido al tratamiento formal de Eikmeyer y Rieser, al análisis pragmático de Van Dijk y a la filosofía del lenguaje del segundo Wittgenstein. Por último, las cuestiones relativas a los conceptos de comprensión y consenso se han fundamentado en las ideas de Gadamer y von Wright.

2 ENUNCIADOS VAGOS Y ENUNCIADOS PRECISOS

La mayor parte de los enunciados que usamos en nuestra vida cotidiana se caracterizan por tener un significado con un cierto grado de imprecisión, vaguedad, indeterminación… Esto está directamente relacionado con el significado de dichos enunciados o expresiones: su aplicación a un conjunto de objetos deja una zona de indeterminación o “penumbra” en la que no puede ser aplicado ni el predicado ni su negación [4]. Por lo tanto, concretizar esa zona de indeterminación requiere el uso de otros elementos, que pueden ser tanto lingüísticos como extralingüísticos. Esto es precisamente lo que convierte a los enunciados o predicados vagos en dependientes del contexto [7]. En los enunciados precisos, sin embargo, el significado no deja ningún hueco, ninguna zona de indeterminación que deba ser rellenada con elementos que

no estén incluidos en el propio significado. Esto hace que los enunciados precisos sean libres de contexto.

A primera vista, la ambigüedad, la imprecisión, la vaguedad son características del lenguaje que deben tender a evitarse, pues pueden ser una fuente de conflictos o problemas en la comunicación. Por lo tanto, lo deseable es un lenguaje depurado de todos estos elementos. Esta fue, precisamente, una de las pretensiones de la filosofía neopositivista en la primera mitad del siglo XX: alcanzar una forma de expresión completamente precisa y libre de ambigüedades. Así, Waismann formuló el Principio de Verificación (1932) según el cual, el significado de una proposición está determinado por las experiencias que permiten determinar de un modo conclusivo si la proposición es verdadera o falsa. Por lo tanto, todo enunciado que no satisficiera dicho principio, era calificado como un pseudo-enunciado y carente de significado. La mayoría de los enunciados científicos, correctamente formulados, cumplen este principio. Sin embargo, otros enunciados, como los de la filosofía o muchos otros que empleamos cotidianamente, no satisfacen esta exigencia; por lo tanto, deberían ser considerados como carentes de significado.

Sin embargo, a pesar de este hecho, los enunciados vagos son usados diariamente con eficacia; su empleo no ocasiona grandes dificultades a aquel que los utiliza y, lo más importante, permiten a los individuos desenvolverse apropiadamente en situaciones en las que no es posible usar enunciados precisos. P.ej.: un monitor de autoescuela es capaz de enseñar a un alumno a aparcar el coche sin tener que saber la medida exacta de la plaza de aparcamiento, ni el ángulo exacto de giro de las ruedas… sino que es suficiente con que aprenda unas pocas reglas aproximadas. Así pues, ¿carecen de significado los enunciados vagos?, ¿deberían de ser desechados por lo problemático que resulta su uso? La respuesta a ambas preguntas es claramente negativa.

Carnap, en años posteriores, suavizó la rigidez del Principio de Verificación sustituyéndolo por el de Confirmabilidad: "*We can only confirm a sentence more and more. Therefore we shall speak of the problem of* confirmation *rather than problem of verification*"[3]. Es claramente menos restrictivo que el Principio de Verificación, pero está basado en la noción probabilidad, por lo que los enunciados vagos siguen siendo problemáticos para él. La vaguedad de los enunciados no depende de su probabilidad, sino más bien de su posibilidad y de su referencia a elementos contextuales. P.ej.: es posible que alguien piense en tomarse media docena de huevos para cenar, pero es muy poco probable que ese alguien lo haga por los efectos nocivos que tiene para su salud.

Los enunciados vagos no sólo suponen un problema para el análisis del significado, sino también para su tratamiento mediante la lógica clásica. La eficiencia de sus reglas está suficientemente demostrada desde su consolidación matemática, pero cuando los enunciados no son precisos, esta se demuestra deficitaria en dos aspectos:

a) Sus estrategias de formalización no son válidas para representar adecuadamente enunciados del tipo "El cielo está bastante nublado", "Si Pedro está bastante calvo entonces tiene que cortarse poco el pelo" o "La mayoría de los suecos son muy rubios", pues, precisamente, no permiten formalizar los elementos contextuales que hacen disminuir la vaguedad o imprecisión que contienen dichos enunciados.
b) Las reglas básicas de la lógica no permiten resolver razonamientos en los que no haya una coincidencia exacta entre la regla y el antecedente de la proposición. Si sólo existiera una coincidencia aproximada, la regla no se podría ejecutar.

3 EL CONTEXTO: ELEMENTO QUE LIMITA LA VAGUEDAD Y LA IMPRECISIÓN

El contexto es el elemento que va a determinar el grado de imprecisión de un enunciado vago. Así, un mismo enunciado vago puede ser muy impreciso en un contexto o muy poco impreciso en otro.

El contexto es una unidad de tipo holístico constituida por elementos de muy diversa naturaleza, que puede ser tanto lingüística como no lingüística, y que forman parte del significado de un enunciado. Esos elementos son todos los objetos que los sujetos son capaces de percibir, producir, recordar, tener creencias acerca de ellos,…

La pragmática se ocupa de analizar específicamente el contexto. En ese marco, estas son algunas definiciones:

- Conjunto de premisas que se usan en la interpretación de un enunciado [5].
- Saber de fondo compartido por el hablante y el oyente que contribuya a la interpretación por parte del oyente de lo que quiere decir el hablante con un enunciado determinado [8].

Ambas definiciones destacan la importancia del contexto para que un oyente pueda comprender adecuadamente un enunciado determinado. En definitiva, semántica y pragmática deben complementarse mutuamente.

Van Dijk [11], dando un paso más, ofrece una descripción de los elementos estructurales del contexto. Así, un contexto consta de:

- El par "aquí-ahora" como indicadores de un punto en el tiempo y un punto de lugar.

- El conjunto de personas o agentes participantes.
- El conjunto de actos comunicativos, que incluye la función de hablar (y el hablante real) y la función de oír (el oyente real).
- Varios conjuntos de conocimiento, creencias, necesidades/deseos, intenciones pertinentes reales.
- Conjunto de convenciones comunicativas de la comunidad de habla.

Algunas de estas estructuras se pueden definir de una manera adecuada empleando la teoría de los mundos posibles de Leibniz [9]. Así, como se analiza en [4], y prescindiendo de los aspectos de teoría de comunicación y relativos a los actos comunicativos, en una situación comunicativa ideal, se da el conjunto de los hablantes, el conjunto de todos los contextos posibles, y el conjunto de todos los *backgrounds* (conjunto de conocimientos, creencias, necesidades, deseos, intenciones…) posibles de los sujetos, siendo posible definir, por ejemplo, las siguientes relaciones:

- El subconjunto de aquellos contextos sobre los cuales el hablante puede ejercer su influencia directa.
- El subconjunto de contextos accesibles para el conjunto de los sujetos. De modo análogo al caso anterior, se designa sólo a aquellos contextos sobre los que los hablantes tienen influencia.
- El subconjunto de los *backgrounds* de cada uno de los sujetos en un contexto particular. Designa a los distintos conjuntos de creencias, intenciones, deseos,… que un hablante puede tener acerca de un determinado contexto.
- El subconjunto de los *backgrounds* de los sujetos en el conjunto de los contextos. Designa a la diversidad de creencias, intenciones, deseos,… que todos los hablantes pueden tener acerca de los contextos en los cuales pueden actuar.

Dados estos componentes, se pueden establecer las siguientes meta-relaciones o interacciones entre ellos:

1) El conjunto de los contextos es igual que la unión de los contextos a los que accede cada sujeto.
2) El conjunto de los *backgrounds* es igual que la unión de todos los *backgrounds* de cada uno de los sujetos.
3) Existe al menos un contexto al que todos los sujetos pueden acceder.

En una situación comunicativa real, por otra parte, no sería posible detallar completamente ni el conjunto de todos los contextos posibles ni accesibles, ni el conjunto de todos los *bakgrounds* posibles ni accesibles. Sin embargo, mediante una aproximación heurística llevada a cabo en ese marco de comunicación real, es posible llevar a cabo una descripción relativamente amplia y completa de dichos conjuntos.

La tercera interacción propuesta es la más importante, pues es la que garantizaría la existencia de un contexto compartido entre todos los hablantes, condición indispensable para que se pueda establecer una buena comunicación entre ellos. Este marco comunicativo es uno de los espacios en los que se explicitan las intenciones de los agentes participantes, quedando así expuestas a la consideración de los otros. Sin embargo, es posible que en dicha intersección exista más de un contexto, lo que puede provocar ambigüedad o vaguedad, puesto que abre la posibilidad de múltiples interpretaciones de los enunciados proferidos en dicha situación comunicativa.

Otra cuestión importante es el hecho de que no hay ninguna restricción relativa a la relación entre contextos y *backgrounds*. Esto garantiza que, en cualquier situación comunicativa, cada sujeto puede tener su propio *background* particular respecto del contexto compartido.

Por último, en cualquier situación comunicativa concreta y efectiva, se da un par ordenado formado por el contexto compartido por todos los sujetos (lo que indica la tercera interacción) y un *background* apropiado de cada sujeto para dicho contexto. Este par es lo que se denomina un mundo posible. Así, existen tantos mundos posibles distintos como *backgrounds* diferentes y apropiados tenga cada sujeto para asociar a dicho contexto compartido, puesto que no hay ninguna restricción que impida que cada contexto tenga asociados varios *backgrounds* diferentes.

En conclusión, si consideramos una situación comunicativa concreta en la que se emiten unos enunciados particulares como un mundo posible, el significado de las proposiciones emitidas sólo puede ser comprendido en ese mundo posible. Ahora bien, ¿qué grado de influencia tiene esto en el significado de los enunciados precisos y de los enunciados vagos?

En los enunciados precisos, puesto que son libres de contexto, la selección de un mundo posible u otro no va a alterar su precisión, o dicho de otro modo, son enunciados precisos en cualquier mundo posible. Esto es debido a que en los ámbitos en los que son empleados se exige una precisión máxima que permita garantizar sus pretensiones de validez, universalidad y necesidad. P.ej. un enunciado del tipo “2 es par” es verdad independientemente del universo de números que se escoja (reales, enteros, positivos, naturales,…), y de las creencias o inferencias que cualquier individuo pueda llevar a cabo. El significado de los enunciados precisos también es completo, pues, sin necesidad de recurrir a elementos que

no están contenidos en la proposición, se puede conocer exactamente su significado.

Sin embargo, si consideramos un enunciado impreciso o vago como "Este es un coche grande" o "Este es un coche caro" comprobar su valor de verdad no es una cuestión inmediata como en el caso anterior. Determinar que un coche es "grande" sólo se puede dilucidar mediante la referencia a ese mundo posible:

- Del contexto o universo de coches que se escoja. Supongamos que ese coche mide 3m. de largo, será grande si lo comparamos con la clase de los utilitarios y será pequeño si lo comparamos con la clase de las limusinas.
- Del *background* de los agentes que mantengan la conversación. Una persona considerará que un coche de 3m. es grande si va a optar por hacer un uso urbano del mismo, si quiere que le consuma poco, si quiere que se pueda aparcar fácilmente… y demás creencias que tenga dicho individuo. Por otra parte, un individuo que tenga la intención de hacer viajes largos, de pasar mucho tiempo conduciendo, que prime la comodidad y la seguridad del mismo, que no le dé demasiada importancia al consumo… no aceptará que un coche de 3m. sea calificado como grande.

En conclusión, el significado de las proposiciones anteriores no se reduce al significado individual de sus términos, sino que varía según el mundo posible en el que se use. La selección de un mundo posible u otro puede hacer aumentar o disminuir la vaguedad del enunciado, por lo tanto, no es igual de vago o impreciso en todos los mundos posibles. Wittgenstein ya dio cuenta de este fenómeno en la siguiente definición de significado "El significado de una palabra es su uso en el lenguaje" [13].

En general, para los enunciados vagos, la LB se muestra como una herramienta de análisis muy apropiada, puesto que sus estrategias de formalización permiten recoger los elementos necesarios del contexto, del mundo posible que delimita el uso de un enunciado para gestionar apropiadamente su vaguedad o imprecisión.

3.1 LÓGICA BORROSA: LÓGICA SENSIBLE AL CONTEXTO

El momento crítico para conseguir un buen análisis de cualquier problema aplicando LB es el de establecer las funciones de pertenencia y el conjunto de reglas necesarias. Seleccionar el tipo de función de pertenencia corresponde al diseñador del sistema de LB, mientras que la distribución de esas funciones en un dominio y rango y la elaboración del conjunto de reglas dependen del conocimiento de un experto en el campo.

En el diseño de un sistema de LB se pueden concretar las siguientes relaciones entre dicha tarea y la selección de un mundo posible. Aquí sólo se refieren las más específicas, pues las más genéricas ya se presuponen:

1) Selección del contexto real o contexto compartido entre todos los agentes participantes: el conocimiento de las particularidades concretas del problema objeto del análisis pertenece a los expertos en el área, pero es una información que debe ser compartida con los diseñadores del sistema de LB para que puedan hacer un buen sistema.
2) El conjunto de todos los agentes participantes: este conjunto está formado por los expertos en el campo, los diseñadores del sistema de LB y de los potenciales usuarios, que será determinado por los expertos. Por lo tanto, aunque este último conjunto es importante, los que efectivamente participan en el proceso son los dos primeros.
3) *Backgrounds*: En tanto que sujetos humanos, tanto los expertos como los diseñadores del sistema tienen unas intenciones, unos deseos… En este caso, no sólo es importante compartir un contexto sino que también es fundamental armonizar los *backgrounds* de todos los agentes participantes, pues sólo así se puede seleccionar el mejor de los mundos posibles. Esto ya es un índice de calidad *a priori*, pues en un sistema donde no se aspire al mejor de los mundos posibles y los *backgrounds* de los agentes participantes sean muy divergentes, los problemas están asegurados y surge lo que comúnmente se llama "falta de entendimiento entre las partes" que puede terminar en "diferencias irreconciliables" y en fracaso. Tanto la selección del tipo de funciones de pertenencia (triangulares, trapezoidales, lineales,…), como la utilización de unas reglas borrosas u otras (de entre las múltiples t-normas, operadores de negación, de implicación… que existen) o el tipo de inferencia aproximada que se utilice (de ser el caso) también dependen de ese mundo posible, pues estas no vienen impuestas a priori sino por la naturaleza del problema.

Como ya ha sido indicado, la intencionalidad es un elemento estructural del *background*, y por lo tanto, fundamental en la selección del mejor de los mundos posibles. Este, a su vez, es fundamental en la modelización de cualquier problema mediante la LB. A continuación, se propone una revisión de dicho concepto desde esta perspectiva.

4 REVISIÓN DE LA INTENCIONALIDAD DESDE LA LÓGICA BORROSA

La intencionalidad es una cuestión clásica y ampliamente tratada en la filosofía. Se ha abordado fundamentalmente desde dos perspectivas: una ética y otra epistemológica. En este apartado sólo va a ser presentada desde la perspectiva epistemológica, pues es la que está directamente vinculada con la LB.

Se suele calificar a la intencionalidad como el conjunto de intenciones, objetivos, metas… que cada individuo se propone como fin y con el grado de conocimientos prácticos que dicho individuo tiene para poder alcanzar eso que se ha propuesto [12]. Por lo tanto, la intención exige de una materialización en una acción, que será de tipo intencional y de acuerdo con esos conocimientos, con la que se pretende alcanzar dicha meta. Asumiendo un modelo explicativo de tipo teleológico, más que causalista, la descripción de una acción intencional debe poner de manifiesto la racionalidad de la misma. Esta concepción muestra la interdependencia existente entre la intención de un determinado sujeto y el conocimiento que ese sujeto debe tener para llevar a buen puerto dicha intención mediante la ejecución de las acciones necesarias.

Si analizamos un sistema de LB en el cual se selecciona el mejor mundo posible para obtener óptimos resultados, la intencionalidad de los individuos, como elemento estructuralmente constituyente de ese mundo posible, es determinante para el diseño del mismo.

El punto de partida del proceso es la selección de una descripción apropiada para el análisis, pues de una misma situación o problema se pueden dar muy distintas descripciones válidas, pero conducentes a identificar intenciones, metas muy distintas. Así, la máxima que debe guiar esa selección es la de buscar la descripción en la que mejor se muestre lo que el agente procura hacer, de modo que el vínculo entre los requisitos causales (fases anteriores) y las consecuencias de la acción (fases posteriores) sea lo más claro posible. P.ej., crear un sistema de LB es una acción que se puede describir como diseñar dicho sistema para un problema concreto.

Una vez descrita la acción, para reconocer la intención del creador o creadores de dicho sistema, hay que identificar cuál es el objeto de su intención, es decir, sólo es posible clarificar esa intencionalidad si se identifica el objeto al que es dirigida. En este caso, el objeto del sistema de LB es el tipo de uso que se va a hacer de él: aproximar funciones, crear controladores, toma de decisiones,…

Otro punto de análisis que ofrece la perspectiva de Wright es la relación entre la intención de un sujeto y sus conocimientos acerca de todo lo necesario para su consecución, es decir, su saber acerca de cómo satisfacer el objeto de su intención. Esto es lo que comúnmente se denomina “pericia del experto” y consiste en la selección acertada del tipo de función de pertenencia y de las t-normas, t-conormas, funciones de negación e implicaciones borrosas de entre el amplísimo abanico que existe de las mismas así como el tipo de inferencia aproximada más apropiada para cada problema. Así, la combinación de la intención con este “saber-hacer”, o sea, la pericia del agente que diseña, culminará con la concreción de un diseño efectivo de dicho sistema de LB.

Por otra parte, y dado que una de las virtudes de la LB es su carácter compartido, contextual y local (garantiza su flexibilidad y un amplio rango de aplicaciones), alcanzar un consenso entre los distintos agentes participantes en el proceso es crucial. En un problema analizado mediante lógica clásica puede haber pequeñas disensiones en la formalización del problema, pero no puede haber ninguno en la selección de un Modus Ponens, pues sólo hay uno. Sin embargo, en la LB esos puntos de disenso son mucho más abundantes e importantes. La fuente de los mismos reside en la posibilidad de que cada agente participante pueda asociar un *background* particular a ese contexto compartido, surgiendo así una multitud de mundos posibles no siempre compatibles entre sí. Lo deseable, por lo tanto, es la selección de un único mundo posible que, en aras de la optimización, debe de ser el mejor de esos mundos posibles. ¿Cómo hacerlo? Mediante el consenso y la creación de un sistema comprensivo.

El consenso consiste en la superación de la tensión existente entre las opiniones de aquellos individuos que pretenden fijar una interpretación común del mundo acerca de un ámbito de la realidad del que sólo se puede tener un conocimiento no metódico [4]. Este tipo de conocimiento sólo puede ser adquirido mediante la experiencia vital, por el día a día y las vivencias cotidianas. Se opone al conocimiento metódico, propio de la ciencia, que está basado en la verificación experimental. El conocimiento modelado por la LB es de naturaleza no metódica (de no ser así, podría ser modelado por la lógica clásica); así, las reglas de los expertos sólo son aproximadas, propuestas a partir de su experiencia y que ofrecen una flexibilidad al entorno que las reglas exactas no pueden dar. Es más, en muchas ocasiones proporcionar reglas exactas ni es posible ni es lo mejor.

El consenso sólo se puede adquirir mediante el diálogo, mecanismo mediante el cuál la opinión de uno no se impone sobre la de los demás, sino que se transforma la opinión de todos los participantes, mediante un acto de propia voluntad, en una opinión única, pactada y compartida [6]. En este momento, la divergencia existente al principio entre los *backgrounds* de los agentes se combina en un único conjunto que, unido a ese contexto compartido, es el mejor de los mundos posibles para el problema a tratar. Seleccionar el mejor de los mundos

posibles es apostar por un funcionamiento óptimo del sistema.

Si ese *background* compartido y unificado se extiende y cada vez es más ampliamente aceptado, llegará a adquirir el carácter de "acuerdo tácito" [6], es decir, no precisa de una discusión previa para ser asumido. Por lo tanto, dicho acuerdo pasará a formar parte del conjunto de todos los contextos posibles como un elemento de los mismos para una situación y viceversa; si un elemento del conjunto de los contextos se discute, dejará de ser contextual y pasará a formar parte del conjunto de los *backgrounds* de los sujetos.

Otra cuestión es la "comprensividad" de un sistema de LB, es decir, lo que permite que otros agentes puedan descubrir, ver, entender el significado del mismo. Esto facilita la participación de otros individuos y su colaboración para posibles mejoras. La "revisabilidad" de lo hecho es otro punto fuerte de la LB.

Una vez aclaradas estas cuestiones, se puede comprender mejor la relevancia de la intencionalidad en la LB, pues ya se han analizado todos los elementos adyacentes. La intencionalidad es la que, en último término, proporciona ese hilo de racionalidad y coherencia a todo el diseño propuesto. Es la respuesta a la pregunta crítica: "¿Por qué?" [1]. Si después de evaluar un sistema construido mediante LB un sujeto (bien sea externo o bien sea el propio diseñador o diseñadores) es capaz de responder a ese "¿Por qué?" está indicando la transparencia, coherencia, cohesión y la construcción consensuada y compartida de dicho sistema, además de situarlo en el conjunto de objetivos y creencias de sus diseñadores. Esto garantiza su constante revisión y evaluación. Sin embargo, una respuesta negativa a ese "¿Por qué?" denota un sistema opaco, poco comprensible, falto de coherencia y de consenso que lo aleja bastante de toda posible revisión o mejora.

En este sentido, una buena intención, entendida como una intención dirigida a un buen análisis del trabajo permite garantizar un alto grado de racionalidad en las propuestas presentadas, tanto en el establecimiento de las funciones de pertenencia como de la definición de las reglas. Esto es una garantía de obtener resultados racionales y razonables, en definitiva, un sistema de calidad.

5 CONCLUSIONES Y ASPECTOS ABIERTOS

En este trabajo se ha pretendido dejar claro que la naturaleza misma de la LB establece un vínculo estructural con los elementos contextuales del problema, tanto a nivel de selección de tipo de funciones de pertenencia, como de reglas y de tipos de razonamiento.

La intencionalidad, uno de los constituyentes estructurales del contexto, se muestra como uno de los elementos determinantes en el diseño de cualquier sistema LB debido a la peculiar conexión que establece entre tres componentes fundamentales en el análisis de cualquier problema: el objetivo, el saber-cómo y la acción misma.
El esquema de von Wright acerca de la intencionalidad es especialmente clarificador para el vínculo anterior, pues denota la racionalidad que debe tener el mismo: garantiza la coherencia de la propuesta, elimina la arbitrariedad (una de las críticas más habituales de la LB) y es una condición *sine qua non* para crear consenso entre los agentes participantes y hacer comprensivo dicho sistema. Un sistema de LB al margen de un análisis intencional y sin motivaciones se convertiría en un sinsentido, lo que va en un claro detrimento de la calidad del sistema.

Por último, una cuestión abierta para su estudio es la evaluación de los distintos tipos de reglas borrosas y de los diferentes tipos de razonamientos aproximados para crear un especie de "catálogo" que, en función de los diversos tipos de intenciones con las que los diseñadores y expertos analicen un problema, les ayude decidir cuál de ellos es el que mejor se adapta a sus necesidades.

Agradecimientos
Este artículo fue realizado con el apoyo de la Dirección Xeral de Ordenación e Calidade do Sistema Universitario de Galicia, de la Consellería de Educación e Ordenación Universitaria-Xunta de Galicia
Mi agradecimiento al profesor Alejandro Sobrino por sus constantes sugerencias y correcciones.

Referencias

[1] G.E.M. Anscombe. *Intención*. Paidós. Barcelona. 1991.

[2] R.E. Bellman & L.A. Zadeh. "Local and fuzzy logic" en J.M. Dunn and G. Epstein (eds.). Modern *Uses of Multiple-Valued Logic*. Dordrecht. Reidel Publishing Co. 1977. pp. 103-165.

[3] R. Carnap. "Testeability and Meaning" en *Philosophy of Science*. Baltimore. 3. 1936. pp.419-471, 4. 1937. pp. 1-40.

[4] H-J. Eikmeyer & H. Rieser. "Meanings, Intensions, and Stereotypes. A New Approach to Linguistic Semantics" en Eikmeyer & H. Rieser (eds.). *Words, Worlds an Contexts*. Walter de Gryter. Berlín. 1981.

[5] Mª V. Escandell. *Introducción a la pragmática*. Ariel. Barcelona. 1996.

[6] G. Gadamer. *Verdad y método II*. Sígueme. Salamanca. 1970.

[7] S. Gross. “Vagueness in Context” en Lila Gleitman and Avarind Joshi (eds.). *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Scociety*, Lawrence Erlbaum. 2000.

[8] G.N. Leech. *Principles of pragmatics*. Longman. 1986[2].

[9] G. Leibniz. *Escritos filosóficos*. Mínimo Tránsito. Madrid. 2003.

[10] S. Shapiro. *Vagueness in Context*. Oxford University Press. New York. 2006

[11] T.A. Van Dijk. *Text and Context: explorations in the semantics and pragmatics of discourse*. Longman. London. 1989.

[12] G.H. von Wright. *Explicación y comprensión.* Alianza. Madrid. 1980.

[13] L. Wittgenstein. *Investigaciones filosóficas*. Instituto de Investigaciones Filosóficas UNAM/Crítica. Barcelona. 1988.

DECIR, HACER Y COMPUTAR CON PALABRAS: ENTRE LA SEMANTICA Y LA LOGICA FUZZY

Luis A. Urtubey[1]
[1] Universidad Nacional de Córdoba (Argentina)
urtubey@ffyh.unc.edu.ar

Resumen

Siguiendo un enfoque semántico recientemente expuesto por R. Brandom, este trabajo se plantea considerar la lógica fuzzy desde la perspectiva de un vocabulario elaborado a partir de una relación semántica pragmáticamente mediada, el cual puede ser explicativo de prácticas que son, o bien necesarias o suficientes para el manejo autónomo de otros vocabularios.

Palabras Clave: Semántica, Pragmática, Significado, uso, lógica fuzzy.

1 LA RELACION SIGNIFICADO-USO: SEMANTICA VS. PRAGMATICA.

1.1

El denominado "Problema del Significado" ha interesado y afectado en casi toda época a la filosofía, aunque especialmente y en forma más explícita, en el siglo XX. Sin duda, la respuesta a la pregunta filosófica acerca de qué es el significado resulta crucial como dice Horwich [2], para una gran variedad de cuestiones filosóficas. Así, por ejemplo, entre otras áreas de la problemática filosófica para las cuales una respuesta a esta pregunta resulta relevante, se encuentran la lógica, la filosofía de la ciencia, la filosofía de la mente, la ética, además, por cierto, de la misma filosofía del lenguaje, donde se arraiga la cuestión.

Expresado en su forma más básica, el problema del significado es el problema acerca de la forma en que se relacionan el lenguaje, el pensamiento y el mundo. Esta relación, filosóficamente problematizada, conduce a una maraña de cuestiones difíciles de abarcar o resolver en una única teoría. Con todo, tradicionalmente ha sido tarea de la semántica, la de teorizar acerca de los fundamentos de la relación entre el lenguaje y el mundo. Tarea que consiste fundamentalmente, en explicar por qué algunos sonidos se consideran provistos de significado y otros no. Lo cual lleva a su vez a la necesidad de dar razón de qué es lo que se denomina significado.

Buena parte de la semántica formal en el siglo XX transitó dentro de lo que constituye la llamada "tradición analítica", en la cual prácticamente tiene su origen sistemático, la denominada Filosofía del Lenguaje. El problema del lenguaje y del significado cobró sin lugar a dudas una importancia primordial para reconocidos autores dentro de esta tradición, como Frege, Russell, Carnap, Wittgenstein, Quine, Putnam y Kripke, por nombrar sólo algunos de los más relevantes.

Como suele señalarse, Wittgenstein constituye un punto de inflexión importante dentro de esta tradición. Punto de inflexión o ruptura reflejado en el propio desarrollo de sus ideas filosóficas. Con la conocida como "segunda etapa" de la filosofía de Wittgenstein, se inicia el desafío pragmatista al proyecto analítico inicial. La preocupación por el significado es así desplazada del centro de la atención filosófica en favor de la noción de "uso"; reemplazando el interés en la semántica por el interés en la pragmática. Sin lugar a dudas, la autonomía de la semántica respecto al estudio filosófico del lenguaje, sufre en ese punto un golpe, que la conmociona definitivamente. A partir de ese momento, filósofos como Sellars, Quine, Davidson y Dummett formulan diversos argumentos que cuestionan algunas tesis fundamentales de la tradición semántica, poniendo al descubierto los nuevos dogmas.

En sus últimos trabajos, R. Brandom ha interpretado de manera particular estos avatares del proyecto analítico. Brandom [1] entiende, que el proyecto ligado a la filosofía analítica del siglo XX se interesa, sobre todo, en relaciones semánticas entre diferentes "vocabularios". La pregunta típica del análisis lógico, indaga acerca de si es posible y de qué manera se puede dar cuenta del significado expresado por un tipo de palabras, en términos de los significados expresados por otro tipo de palabras. Un caso paradigmático lo constituyen, por ejemplo, los proyectos destinados a mostrar que todo lo expresable en el vocabulario de la teoría de números, así como mediante el uso de descripciones definidas, es ya expresable en el lenguaje de la lógica cuantificacional de primer orden con identidad.

La naturaleza de la relación semántica entre estos diversos vocabularios, ha recibido variadas caracterizaciones, tales como análisis o reducciones de diversa índole. Entre la clase de vocabularios, cuyas relaciones semánticas se consideró importante investigar en este período, se hallan para Brandon, las identificadas con dos programas en particular, a saber, el empirismo y el naturalismo. De este modo, el proyecto general del análisis filosófico se puede entender como una interpretación o reducción de otros vocabularios, a estos dos vocabularios en cierto modo privilegiados.

Por su parte, el vocabulario lógico, tiene un rol especial en la especificación de estas relaciones semánticas. Se considera al menos lícito, apelar al vocabulario lógico a fin de elaborar la relación entre *analysandum* y *analysans*, es decir entre el vocabulario a reducir y el vocabulario base.

Este proyecto, tal como lo entiende Brandom, se vio confrontado a mediados del siglo pasado con un desafío importante, el que antes denominamos ya "desafío pragmático". En términos generales, el desafío pragmático conlleva un desplazamiento como centro de la atención filosófica, de la noción de significado a la de uso. No obstante, Brandom pone el acento en una formulación algo diferente y más general de esta transición, remarcando el hecho de que se apela a las prácticas vinculadas al desarrollo de diversos vocabularios, en lugar de tomar sólo en cuenta el significado que ellos expresan.

Bajo cierta interpretación, se puede entender que Wittgenstein radicaliza estas críticas pragmatistas a la semántica, advirtiendo que "no se considere el significado, sino el uso". No habría ya ningún punto privilegiado en el lenguaje desde el cual establecer principios válidos en todo el resto. El lenguaje –dice Wittgenstein- no tiene centro. Sólo un "aire de familia" reúne las prácticas lingüísticas, sin que haya ningún principio especificable en otros vocabularios. Se ve en seguida que, llevada a su extremo, esta línea de pensamiento conduce a una negación directa de toda posibilidad de análisis semántico en el sentido clásico. En lugar de la semántica queda ahora la pragmática entendida como el estudio del uso de las expresiones, sólo en virtud del cual tienen significado. A la tradición formal, matemáticamente inspirada –dice Brandom- de Frege, Russell y Carnap, se opone una investigación antropológica, histórico-natural, socio-práctica, orientada a desmitificar las pretensiones filosóficas sobre este discurso. Puesto en estos términos, todo filósofo se vio compelido a tomar partido por uno u otro bando.

1.2

Un aspecto remarcable del análisis de Brandom, es que considera una tercera alternativa, que consiste en la posibilidad de integrar ambas perspectivas, tratando de beneficiarse con el aporte de ambas, considerándolas no opuestas sino complementarias. En este sentido, la pragmática puede servir para profundizar la semántica, en lugar de excluirla.

Un aspecto que merece destacarse en las ideas de Brandom antes expuestas, es que permiten también ver en una nueva perspectiva, la aplicación al lenguaje de teorías o modelos formales o matemáticos, permitiendo reexaminar las dificultades teóricas producidas por la irrupción de elementos pragmáticos, como sucede por ejemplo en el caso de las teorías que se ocupan de la imprecisión y la vaguedad en el lenguaje, como es el caso de lógica fuzzy. En este sentido, una de las objeciones filosóficas corrientes a la aplicación de la lógica fuzzy al análisis de los lenguajes vagos o imprecisos tiene relación con lo que Settimo Termini [4] propone llamar "vagaries of fuzziness" , entendiendo por tal el hecho de que la lógica fuzzy intenta fundamentar en un nivel más general la noción de pertenencia, pero haciendo uso al mismo tiempo de nociones que necesitan teorías clásicas o nítidas para su formulación. El problema no tiene que ver con la eficacia de la teoría ni con su desarrollo matemático, sino con el sustento que ofrece a la fundamentación teórica de una noción innovadora. Se remarca como algo extraño por ello, que tal formalización sirva como *explicatum* de una noción que parece excederla en su carácter innovador y general.

Una posible línea superadora de esta clase de dificultades podría resultar entonces ampliando el proyecto de la semántica, desde su interés exclusivo en las relaciones de significado, a un interés que abarque también las relaciones entre significado y uso, dejando así abierta la posibilidad de esclarecer el uso de un vocabulario, aceptando que expresa alguna suerte de significado o contenido. Esta tesis, que puede denominarse siguiendo a Brandom, pragmatismo semántico, sostiene que la única vía para explicar que un significado dado se halla asociado con un vocabulario, es el uso del mismo: las prácticas –dice Brandom- por las cuales es conferido o las habilidades cuyo ejercicio constituye el despliegue de un vocabulario con este significado.

En los aspectos que son aquí relevantes, el aparato conceptual que Brandom introduce a fin de elaborar esta relación fundamental entre uso y significado, entre habilidades prácticas y vocabularios, comprende una serie de relaciones básicas. Por un lado, se debe tener en cuenta lo que es usar palabras o alocuciones como expresiones de significados, esto es, qué es lo que uno debe hacer para que eso cuente como decir algo que el vocabulario permite expresar. Esta relación se denominará de "suficiencia práctica-vocabulario" (Suficiencia P-V) y se da cuando el manejo de un conjunto específico de prácticas o el ejercicio de ciertas habilidades resulta suficiente para que se considere que quien las posee despliega un vocabulario específico. Por otro lado, hablar de prácticas o habilidades tiene sentido en la medida que sea relativa al vocabulario en que se especifican tales prácticas o habilidades. De este modo, aparece una

segunda relación básica entre uso y significado, que se denomina de "suficiencia vocabulario-práctica", "suficiencia V-P". Se trata de la relación que se da entre un vocabulario y un conjunto de prácticas o habilidades cuando este vocabulario es suficiente para especificar estas prácticas o habilidades.

2 SIGNIFICADO, USO Y LOGICA FUZZY

2.1

En algunos de sus últimos trabajos, entre otros Trillas [9], Enric Trillas ha expuesto una interpretación de la lógica fuzzy en línea, se podría decir, con una teoría pragmática del significado de los predicados imprecisos del lenguaje natural. Una exposición también muy elaborada de esta propuesta puede verse con gran provecho en la reciente tesis de Sergio Guadarrama [4]. Desde una perspectiva pragmática, un predicado P es caracterizado desde este punto de vista, como ejerciendo una "acción" sobre un universo de discurso X. Tal acción se percibe a su vez, a través de los enunciados "x es P" y consiste en formar un "colectivo" sobre X. Refinando este procedimiento básico, se puede alcanzar también una caracterización de predicados imprecisos, tratando de reflejar los elementos constitutivos del uso de estos predicados. A este fin, bastaría con diferenciar tres nociones básicas o primitivas, a saber, "x es P", "x es tan P como y" y "x es menos P que y". Las cuales permiten introducir dos relaciones básicas respecto al uso de P en X. La última, permite a su vez definir un concepto de "grado" para P, mediante una medida no necesariamente numérica, toda vez que esta relación define un preorden sobre X. De este modo, la representación del colectivo originado por el uso de P en X, a partir de cada grado vinculado con este uso particular, resulta en un conjunto fuzzy, tal como lo definiera Zadeh, cuando tal graduación toma sus valores en el intervalo unidad.

Esta forma de ver los conjuntos fuzzy se halla en sintonía con la sugerencias de Zadeh [11] [12], con respecto a una nueva perspectiva sobre la lógica fuzzy, que la ubica en el campo más amplio de la computación con palabras (CW), una metodología, a decir de Zadeh, en la cual se usan palabras en lugar de números para computar y razonar. En CW, las palabras son consideradas como rótulos sobre los elementos básicos denominados "gránulos", definidos como conjuntos fuzzy de puntos reunidos por una relación de similitud. La denotación de una palabra, señala Zadeh, puede verse como un predicado de orden superior, como en la gramática de Montague. De esta manera se tiene también una interpretación del significado de los predicados en términos de conjuntos fuzzy. En particular esto se puede ver en el hecho de que una de las cuestiones centrales que se plantea en CW, es precisamente la representación de las restricciones fuzzy. El problema consiste en cómo explicitar las restricciones fuzzy que se hallan implícitas en las proposiciones expresadas en el lenguaje natural. La noción de gránulo recoge esta idea al estar definida como una restricción fuzzy sobre una variable. Por su parte, una proposición es vista como una caracterización lingüística de una percepción, la cual siempre será inevitablemente fuzzy.

En ambas perspectivas brevemente expuestas, se puede apreciar un interés particular en cuanto a la formalización del significado de las expresiones del lenguaje natural. Para CW, el significado de una palabra en el lenguaje natural puede representarse como una restricción sobre una variable implícita. El uso de las palabras, a su vez, puede verse como una forma de granulación, la cual a su turno, puede considerarse como un modo de manejarse con cantidades en forma fuzzy.

De este modo, resulta que la lógica fuzzy tiene un rol fundamental que jugar en la cuestión del significado y por ello en la semántica. La idea que este trabajo quiere plantear es que este rol destacado de la lógica fuzzy puede fundamentarse adecuadamente en base al esquema de Brandom antes considerado. Ubicada en este esquema, la lógica fuzzy puede interpretarse desde la perspectiva de un vocabulario elaborado a partir de una relación semántica pragmáticamente mediada, el cual puede ser explicativo de prácticas que son o bien, PV-necesarias o suficientes para el manejo autónomo de otro vocabulario.

2.2

La noción de significado es la noción central de la semántica. Asimismo, se considera hoy que la semántica es una disciplina empírica. No obstante, puede sostenerse con cierto fundamento, como dice Enric Trillas, que no hay aún un enfoque de la noción de significado basada en modelos matemáticos y una experimentación controlada. Es así que se está lejos de contar con una buena formalización que resulte de utilidad en las ciencias de la computación, donde sólo la lógica fuzzy se esfuerza por alcanzar una especificación del significado a través de una representación con suficiente poder expresivo. No obstante, se debe reconocer que la noción de significado es a su vez una noción extremadamente compleja y presenta muchas más dificultades respecto al estatus de su estudio científico que los objetos de estudio de otras disciplinas empíricas. El desarrollo mismo de la semántica, bastante reciente en tanto su inicio no va mas allá de los años 70, muestra una variedad de enfoques en los cuales no se dejan de apreciar fuertes presupuestos filosóficos, como ha hecho notar M. Stokhof [6][7]. El significado es como objeto de estudio bastante particular. Tiene muchas facetas, deriva de diferentes fuentes, juega roles diferentes e interactúa de distintos modos con otras funciones cognitivas y no-cognitivas. Una razón aún más fundamental es que el significado tiene tanto rasgos mecánicos como conscientes. La mayor parte del tiempo es producido y procesado de una manera completamente mecánica, sin ningún esfuerzo consciente. No obstante, en

otros casos, es usado y puesto bajo cuidadoso escrutinio de manera deliberada y consciente. Todo esto pone de relieve que el significado tiene dos aspectos, individual y social. Este último se pone de manifiesto, por ejemplo, en su carácter contextual, donde aparece claramente el uso consciente del lenguaje. El mundo al cual el lenguaje se aplica no está todo dado desde un comienzo, servido ante cada uno, sino que es también una construcción social, lejos de una imagen mecanicista. La gente crea el lenguaje y también es creada por este.

De este modo, entre los aspectos del significado que plantean un desafío particular, se encuentra sobre todo, su componente normativo. El reconocimiento de este aspecto es el elemento fundamental en el cambio de la filosofía de Wittgenstein y es lo que ha llevado a derribar en cierto modo la separación entre semántica y pragmática. El desplazamiento del significado al uso, el significado como uso, remarca justamente este aspecto normativo. El uso esta regido por reglas, que encuentran su explicación última en razones y no en causas. Razones que necesitan de un marco particular para que tengan sentido. Para responder al cuestionamiento sobre por qué usamos una expresión del modo que lo hacemos, debemos recurrir a una justificación, no a una explicación causal. Justificación que sólo puede darse dentro de un determinado marco. Las prácticas en última instancia que rigen el significado son prácticas normativas. Como hace notar también Stokhof [6] , sólo una vez que los agentes son conscientes del hecho de que se comportan de un cierto modo, pero podrían hacerlo de un modo diferente, aparece la necesidad imperativa de determinar normativamente las conductas. Este es el momento en que las regularidades se transforman en reglas, los patrones de conducta devienen prácticas institucionales.

De todo esto resulta que una teoría del significado adecuado deberá incorporar necesariamente este aspecto social o contextual que le da una dimensión definitivamente pragmática. Deberá superar en cierto modo un enfoque individualista o internalista que ha marcado con fuerza el desarrollo de la semántica. Parece necesario entonces prestar atención a fenómenos típicamente colectivos o grupales que envuelven al significado, al nivel de grupos de usuarios de un lenguaje. Para ello resultará también apropiado recurrir a otras herramientas formales orientadas al análisis de la conducta cooperativa o diversos tipos de intercambio de información o comunicación.

La sensibilidad de la lógica fuzzy hacia estos fenómenos es bien remarcada por Trillas [9], cuando señala que al dar lugar a que haya un número incontable de teorías de conjuntos fuzzy, la lógica fuzzy reconoce que no hay un único modo de usar la conjunción “y”, la disyunción “o” y la negación “no” , y que las diferentes representaciones de estas conectivas lingüísticas se basan en las propiedades particulares de sus significados en el momento, en un contexto dado y con un fin particular. Claramente de aquí se desprende que los significados no pueden ser considerados como algo dado, cuya existencia es independientes de la intención de un sujeto y de toda comunicación lingüística. Por el contrario, no pueden caracterizarse sino en una relación concreta con los hablantes del lenguaje.

2.3

Como ya fue dicho, el esquema de Brandom puede servir para introducir también en un plano conceptual, una forma de considerar la lógica fuzzy en relación con los modelos formales aplicados al lenguaje natural. La lógica fuzzy brinda en este sentido un vocabulario particular de gran ductilidad para interpretar aquellos aspectos del lenguaje en donde la imprecisión y la vaguedad tienen un rol fundamental. Tal como sucede por ejemplo con los predicados que dependen de la percepción, como los términos que designan colores. Es conocida aquí la paradoja del sorites que surge con la aplicación de estos predicados. Muchas teorías semánticas optan por no dar cuenta de ella sino a través de su exclusión. No obstante, el hecho de que bajo ciertas circunstancia se pierda el rumbo al aplicar este tipo de predicados, viéndose uno forzado a aplicarlos de modo inconsistente desde un punto de vista lógico, es algo que debe quedar incorporado en su significado y tiene que contarse como parte de su uso. Un enfoque en este sentido a través de la lógica fuzzy se trata de alcanzar en Trillas y Urtubey [10].

Vale dar aquí algunas precisiones finalmente en la dirección del esquema de Brandom. ¿Qué rol jugaría en este caso la lógica fuzzy? Para responder, cabe dar una breve indicación sobre el tipo de prácticas y vocabularios involucrados. A tal fin se puede entender aquí el lenguaje de conjuntos fuzzy como un sistema de restricciones en el sentido de Zadeh. Cierta ventaja de considerar un lenguaje de restricciones, es que nos aparta de aquel marco individualista en que los significados son vistos como entidades que operan por si mismas; de este modo el significado resulta de una interacción dada por conjuntos de restricciones, que actúan de un modo no determinista, no como funciones, sobre elementos perceptivos.

En este sentido, el vocabulario de la lógica fuzzy puede presentarse bajo una elaboración algorítmica quizás, como explicativo de prácticas que son PV-necesarias para el desarrollo de otro vocabulario. Cuáles sean estas habilidades, es algo que se desprende de las consideraciones anteriores. Zadeh [11] señala en particular que CW tiene su motivación y resulta necesaria precisamente por la presencia de habilidades humanas imposibles de considerar de otro modo. Hay dos imperativos mayores – dice Zadeh- para CW. Primero, CW es una necesidad cuando la información disponible es muy imprecisa para justificar el uso de números, y segundo, cuando hay una tolerancia para la imprecisión, que puede explotarse para alcanzar mejores resultados. De este modo, CW se inspira en una destacable capacidad humana de desempeñar una

amplia variedad de tareas físicas y mentales sin mediciones o cómputos. Se puede agregar que también son capacidades observables en todos los animales y tal vez en todos los seres vivos en general. Devlin [2] refiere el ejemplo del perro, que puede determinar la trayectoria parabólica de un objeto lanzado al aire y establecer matemáticamente el punto donde podrá interceptarlo, procediendo de modo correcto y guiado sólo por su pobre percepción visual y su poderoso olfato. Debajo de esta capacidad se encuentra la habilidad del cerebro para manipular percepciones –percepciones de distancia, tamaño, peso, color, velocidad, dirección, fuerza y también verdad y posibilidad entre otras características de los objetos físicos y mentales.

Una diversidad de vocabularios se basan en estas habilidades. Vocabularios que sirven para manejar o gestionar la imprecisión en las ciencias naturales, la tecnología y también las ciencias sociales. Brandon concluye sosteniendo que el vocabulario lógico se puede elaborar algorítimicamente a partir de ciertas prácticas a las cuales también explica. Prácticas que son PV-necesarias para el desarrollo autónomo de cualquier vocabulario. Algo similar es lo que se puede entender que sucede respecto a la lógica fuzzy y su relación con otros vocabularios.

3. CONCLUSION

Finalmente, a modo de conclusión y para dar un poco más de precisión a esta idea, vale la pena considerar uno de estos vocabularios que puede constituir un caso particular de lo antes dicho. Servirá para estos fines, sólo a modo de ejemplo, el caso de la definición de la causalidad, según algunas consideraciones al respecto de Zadeh [13].

Zadeh dice que el intento de definir causalidad dentro de la estructura conceptual de la lógica clásica y la teoría de la probabilidad, no tiene chance de ser exitoso. La definibilidad se asocia con un lenguaje que se usa para definir un concepto, por ejemplo como señala Zadeh, un lenguaje natural o la lógica de primer orden. El lenguaje más general y más expresivo a este fin es el denominado PNL, un subconjunto del lenguaje natural, que se vuelve preciso mediante la traducción a GCL (lenguaje de restricciones generalizadas). Un concepto es indefinible o amorfo, según Zadeh, si no es definible mediante PNL. La definición de causalidad se ubica justamente más allá del alcance de PNL. De este modo, no se contaría con un vocabulario en el cual decir efectivamente lo que se hace cuando se dice algo sobre causalidad. El uso del concepto escapa a tal caracterización.

No obstante, en un modelo que trate lo amorfo en este sentido como una medida de la complejidad, como el considerado en [3], la lógica fuzzy da más flexibilidad para tratar con la extrema complejidad de un concepto, dando cabida a la posibilidad de expresar las prácticas que se ligan con su uso y constituyen su significado.

Agradecimientos

A Enric Trillas y a las Autoridades del ECSC por su apoyo para poder participar en ESTYLF 2008.

Referencias

[1] R. Brandom. Between Saying and Doing: Towards an Analytic Pragmatism. *The John Locke Lectures*, Oxford University, 2006.

[2] K. Devlin. *The Math Instinct: Why You're a Mathematical Genius (along with Lobsters, Birds, Cats, and Dogs)*, Avalon Publishing: Thunder's Mouth Press 2005

[3] M. Ghomshei, J. Meech y R. Naderi. Fuzzy Logic in a Posmodern Era. Presentado al *BISC Special Event in Honor of Prof. Lofti A. Zadeh, Forging New Frontiers: 40 Years of Fuzzy Pioneers (1965-2005).* BISCSE´05 – University of California, Berkeley. Nov. 2-5 2005.

[4] S. Guadarrama. A contribution to Computing with Words/Perceptions. Tesis Doctoral, Universidad Politécnica de Madrid, Facultad de Informática. Archivo Digital UPM, http://oa.upm.es/

[5] P. Horwich. *Meaning*. Clarendon Press. Oxford, 1998.

[6] M. Stokhof. Meaning, Interpretation and Semantics. En D. Baker-Plummer, D.Beaver, J. Van Benthem y P. Scotto di Luzio (eds), *Words, Proofs and Diagrams*, Stanford, CSLI Press, Pág. 217-240.

[7] M. Stokhof. Hand or Hammer. Presentado en *100 Years On Denoting*. June 24 2005. Utrecht

[8] S. Termini. On Some Vagaries of Vagueness and Information. *Annals of Mathematics and Artificial Intelligence* 35, 2002, Pág. 343–355

[9] E. Trillas. On the Use of Words and Fuzzy Sets. *Information Sciences*, 176 (2006) Pág. 1463-1487.

[10] E.Trillas y L. Urtubey. Toward the Dissolution of the Sorites Paradox. A publicarse en *Applied Soft Computing*

[11] L. Zadeh. From Computing with Numbers to Computing with Words –From Manipulation of Measurements to Manipulation of Perceptions. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 45, 1, Pág. 105-119.

[12] L. Zadeh. Fuzzy Logic = Computing with Words. En L. A. Zadeh, J. Kacprzyk (eds.) *Computing with Words in Information/Intelligent Systems 1*, Pág. 3-23.

[13] L. Zadeh. Causality is Indefinible –Toward a Theory of Hierarchical Definibility. En *Intelligence in a Materials World*, *Selected Papers form IPMM-2001*, Pág. 237, CRC Press, Nueva York.

A CONSENSUS MODEL FOR GROUP DECISION MAKING IN UNBALANCED FUZZY LINGUISTIC CONTEXTS

F.J. Cabrerizo[1] **S. Alonso**[2] **I.J. Pérez**[1] **E. Herrera-Viedma**[1]

[1] Department of Computer Science and A.I., University of Granada, {cabrerizo, ijperez, viedma}@decsai.ugr.es
[2] Department of Software Engineering, University of Granada, zerjioi@ugr.es

Abstract

The aim of this paper is to present a consensus model for group decision making problems with unbalanced linguistic information, i.e., assuming that the preferences are assessed on linguistic term sets whose terms are not symmetrically and uniformly distributed. This consensus model is based on both a fuzzy linguistic methodology to deal with unbalanced linguistic term sets and consensus criteria. Additionally, it presents a feedback mechanism to help experts for reaching a high consensus in decision making processes.

Keywords: fuzzy linguistic modelling, group decision making, consensus.

1 INTRODUCTION

In Group Decision Making (GDM) problems there are a set of alternatives in order to solve a problem and a group of experts trying to achieve a common solution. To solve these problems, the experts are faced by applying two processes before obtaining a final solution [4, 8]: *the consensus process* and *the selection process*. The consensus process is defined as a dynamic and iterative group discussion process, coordinated by a moderator helping experts to bring their opinions closer. If the consensus level is lower than a specified threshold the moderator would urge experts to discuss their opinions further in an effort to bring them closer. Otherwise, the moderator would apply the selection process which consists in obtaining the final solution to the problem from the opinions expressed by the experts. In this framework, an important question is how to substitute the actions of the moderator in order to automatically model the whole consensus process [4].

Figure 1: Example of an unbalanced linguistic term set of 8 labels

On the other hand, many GDM problems based on fuzzy linguistic approaches use symmetrically and uniformly distributed linguistic term sets [1, 3, 4, 10], i.e., assuming the same discrimination levels on both sides of mid linguistic term. However, there exist problems that need to assess their variables with linguistic term sets that are not uniformly and symmetrically distributed [2, 7, 9]. This type of linguistic term sets are called [2, 7] *unbalanced fuzzy linguistic term sets* and its use requires to define new methods of management of linguistic information (see Figure 1).

The aim of this paper is to present a consensus model for GDM problems defined in an unbalanced fuzzy linguistic context. In [2, 7] a methodology to manage unbalanced fuzzy linguistic information was presented which used hierarchical linguistic contexts [6] based on the linguistic 2-tuple computational model [5]. However, this methodology can only represent unbalanced linguistic term sets when there exists a level with an adequate granularity to represent the subset of linguistic terms on the left of the mid linguistic term and a level with an adequate granularity to represent the subset of linguistic terms on the right of the mid linguistic term. Thus, we present a new fuzzy linguistic methodology that can represent unbalanced fuzzy linguistic information when the above conditions are not satisfied. As part of the consensus model, a feedback mechanism substituting the figure of the moderator is given to help experts change their opinions on the alternatives in order to obtain the highest degree of consensus possible. It consists of simple and easy rules generating recommendations in the discussion process. Moreover, this model is based on two

types of consensus criteria, *consensus degrees* evaluating the agreement of all the experts, and *proximity measures* evaluating the distance between experts' individual opinions and the group or collective opinion which is also used in the feedback mechanism to guide the direction of the changes in experts' opinions in order to increase the consensus degrees.

To do so, the paper is set out as follows. In Section 2, we present the new fuzzy linguistic methodology to manage unbalanced fuzzy linguistic information. Section 3 introduces the new consensus model for GDM problems in an unbalanced fuzzy linguistic context. In Section 4, a practical example is given to illustrate the application of the consensus model. Finally, some concluding remarks are pointed out in Section 5.

2 A NEW FUZZY LINGUISTIC METHODOLOGY TO MANAGE UNBALANCED FUZZY LINGUISTIC INFORMATION

In this section, we make a review of the 2-tuple fuzzy linguistic representation model [5] and the concept of hierarchical linguistic contexts [6] in order to present the new methodology to manage unbalanced fuzzy linguistic information.

2.1 THE 2-TUPLE FUZZY LINGUISTIC REPRESENTATION MODEL

The 2-tuple fuzzy linguistic representation model [5] is based on the concept of *symbolic translation* and represents the linguistic information by means of a pair of values, (s, α), where s is a linguistic label and α is a numerical value that represents the value of the symbolic translation.

Definition 1. *[5] Let β be the result of an aggregation of the indexes of a set of labels assessed in a linguistic term set $S = \{s_0, s_1, \ldots, s_{g-1}, s_g\}$, i.e., the result of a symbolic aggregation operation. $\beta \in [0, g]$, being $g + 1$ the cardinality of S. Let $i = round(\beta)$ and $\alpha = \beta - i$ be two values, such that, $i \in [0, g]$ and $\alpha \in [-0.5, 0.5)$, then α is called a symbolic translation.*

This model defines a set of transformation functions to manage the linguistic information expressed by linguistic 2–tuples.

Definition 2. *Let S be a linguistic term set and $\beta \in [0, g]$ a value supporting the result of a symbolic aggregation operation, then the 2–tuple that expresses the equivalent information to β is obtained with the following function $\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5)$:*

$$\begin{aligned} &\Delta(\beta) = (s_i, \alpha) \\ &i = round(\beta) \\ &\alpha = \beta - i \end{aligned} \qquad (1)$$

where "round" is the usual round operation, s_i has the closest index label to "β" and "α" is the value of the symbolic translation.

Finally, for all Δ there exists Δ^{-1}, defined as $\Delta^{-1}(s_i, \alpha) = i + \alpha = \beta$.

2.2 HIERARCHICAL LINGUISTIC CONTEXTS

A *Linguistic Hierarchy* is a set of levels, where each level represents a linguistic term set with different granularity from the remaining levels of the hierarchy [6]. Each level is denoted as $l(t, n(t))$, where t is a number indicating the level of the hierarchy, and $n(t)$ is the granularity of the linguistic term set of t. A graphical example of a linguistic hierarchy is shown in Figure 2.

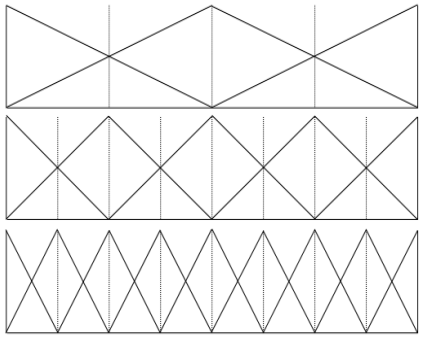


Figure 2: Linguistic hierarchy of 3, 5 and 9 labels

The levels belonging to a linguistic hierarchy are ordered according to their granularity, i.e., for two consecutive levels t and $t + 1$, $n(t + 1) > n(t)$. Then, a linguistic hierarchy LH can be defined as the union of all levels t: $LH = \bigcup_t l(t, n(t))$.

Given a LH, we denote as $S^{n(t)}$ the linguistic term set of LH corresponding to the level t of LH characterized by a granularity of uncertainty $n(t)$: $S^{n(t)} = \{s_0^{n(t)}, \ldots, s_{n(t)-1}^{n(t)}\}$. Furthermore, the linguistic term set of the level $t + 1$ is obtained from its predecessor as: $l(t, n(t)) \rightarrow l(t + 1, 2 \cdot n(t) - 1)$.

Transformation functions between labels from different levels to accomplish processes of computing with words in multigranular linguistic information contexts without loss of information were defined in [6].

Definition 3. [6] *Let $LH = \bigcup_t l(t, n(t))$ be a linguistic hierarchy whose linguistic term sets are denoted as $S^{n(t)} = \{s_0^{n(t)}, \ldots, s_{n(t)-1}^{n(t)}\}$, and let us consider the 2-tuple fuzzy linguistic representation. The transformation function from a linguistic label in level t to a label in level t' is defined as $TF_{t'}^t : l(t, n(t)) \longrightarrow l(t', n(t'))$ such that*

$$TF_{t'}^t(s_i^{n(t)}, \alpha^{n(t)}) = \Delta_{t'}\left(\frac{\Delta_t^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t') - 1)}{n(t) - 1}\right). \quad (2)$$

2.2.1 An Unbalanced Fuzzy Linguistic Representation Model

The procedure to represent unbalanced fuzzy linguistic information defined in [2, 7] works as follows:

1. Find a level t^- of LH to represent the subset of linguistic terms S_{un}^L on the left of the mid linguistic term of S_{un}.
2. Find a level t^+ of LH to represent the subset of linguistic terms S_{un}^R on the right of the mid linguistic term of S_{un}.
3. Represent the mid term of S_{un} using the mid terms of the levels t^- and t^+.

The problem appears when there does not exist a level t^- or t^+ in LH to represent S_{un}^L or S_{un}^R, respectively. Then, we propose to overcome this problem by applying the following algorithm, which is defined assuming that there does not exist t^-, as it happens with the unbalanced fuzzy linguistic term set given in Figure 1:

1. Represent S_{un}^L:
   (a) Identify the mid term of S_{un}^L, called S_{mid}^L.
   (b) Find a level t_2^- of the left sets of LH^L to represent the left term subset of S_{un}^L, where LH^L represents the left part of LH.
   (c) Find a level t_2^+ of the right sets of LH^L to represent the right term subset of S_{un}^L.
   (d) Represent the mid term S_{mid}^L using the levels t_2^- and t_2^+.
2. Find a level t^+ of LH to represent the subset of linguistic terms S_{un}^R
3. Represent the mid term of S_{un} using the levels t^+ and t_2^+.

For example, applying this algorithm the representation of the unbalanced fuzzy linguistic term set $S_{un} = \{N, VL, L, M, H, QH, VH, T\}$ shown in Figure 1 with the linguistic hierarchy LH shown in Figure 2 would be as it is shown in Figure 3. In this example,

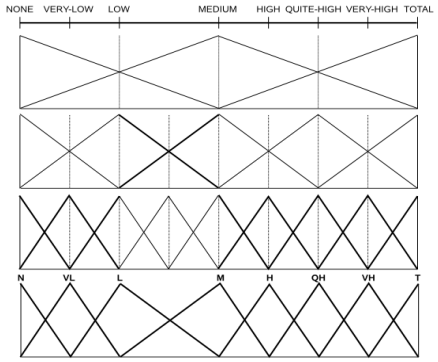


Figure 3: Representation for an unbalanced term set of 8 labels

- $S_{un}^L = \{N, VL, L\}$,
- $S_{mid}^L = L$,
- $LH^L = \{s_0^{n(1)}\} \bigcup \{s_0^{n(2)}, s_1^{n(2)}\} \bigcup \{s_0^{n(3)}, s_1^{n(3)}, s_2^{n(3)}, s_3^{n(3)}\}$.

Thus, we have that $t_2^- = 3$, $t_2^+ = 2$, $S_{mid}^L = L$ is represented using both levels, 3 and 2, and the mid term of S_{un} is represented using the levels 2 and 3.

2.2.2 An Unbalanced Fuzzy Linguistic Computational Model

In any fuzzy linguistic approach we need to define a computational model to manage and aggregate linguistic information. As in [3, 5], we have to define three types of computation operators to deal with unbalanced fuzzy linguistic information: comparison operator, negation operator and aggregation operator. In a unbalanced linguistic context, previously to carry out any computation task of unbalanced fuzzy linguistic information we have to choose a level $t' \in \{t^-, t_2^-, t^+, t_2^+\}$, such that $n(t') = max\{n(t^-), n(t_2^-), n(t^+), n(t_2^+)\}$:

1. *An unbalanced linguistic comparison operator:* The comparison of linguistic information represented by two unbalanced linguistic 2-tuples $(s_k^{n(t)}, \alpha_1)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$, and $(s_l^{n(t)}, \alpha_2)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$ is similar to the usual comparison of two 2-tuples but acting on the values $TF_{t'}^t(s_k^{n(t)}, \alpha_1) = (s_v^{n(t')}, \beta_1)$ and $TF_{t'}^t(s_l^{n(t)}, \alpha_2) = (s_w^{n(t')}, \beta_2)$. Then, we have:
   - if $v < w$ then $(s_v^{n(t')}, \beta_1)$ is smaller than $(s_w^{n(t')}, \beta_2)$.
   - if $v = w$ then

(a) if $\beta_1 = \beta_2$ then $(s_v^{n(t')}, \beta_1)$, $(s_w^{n(t')}, \beta_2)$ represent the same information.
(b) if $\beta_1 < \beta_2$ then $(s_v^{n(t')}, \beta_1)$ is smaller than $(s_w^{n(t')}, \beta_2)$.
(c) if $\beta_1 > \beta_2$ then $(s_v^{n(t')}, \beta_1)$ is bigger than $(s_w^{n(t')}, \beta_2)$.

2. *An unbalanced linguistic 2-tuple negation operator.* Let $(s_k^{n(t)}, \alpha)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$ be an unbalanced linguistic 2-tuple, then:

$$NEG(s_k^{n(t)}, \alpha) = Neg(TF_{t''}^t(s_k^{n(t)}, \alpha)), \quad (3)$$

where $t'' \in \{t^-, t_2^-, t^+, t_2^+\}$ and $Neg(s_i, \alpha) = \Delta(g - \Delta^{-1}(s_i, \alpha))$.

3. *An unbalanced linguistic aggregation operator.* To aggregate unbalanced fuzzy linguistic information by means of its representation in a LH, we use the $LOWA_{un}$ operator, which is an extension of the Linguistic Ordered Weighted Averaging operator [3], and is defined as follows:

Definition 4. *Let $\{(a_1, \alpha_1), \ldots, (a_m, \alpha_m)\}$ be a set of unbalanced assessments to aggregate, then the $LOWA_{un}$ operator ϕ_{un} is defined as:*

$$\phi_{un}\{(a_1, \alpha_1), \ldots, (a_m, \alpha_m)\} = W \cdot B^T =$$

$$C_{un}^m\{w_k, b_k,\ k = 1, \ldots, m\} =$$

$$w_1 \otimes b_1 \oplus (1 - w_1) \otimes C_{un}^{m-1}\{\beta_h, b_h,\ h = 2, \ldots, m\}$$

where $b_i = (a_i, \alpha_i) \in (S_{un} \times [-0.5, 0.5))$, $W = [w_1, \ldots, w_m]$, is a weighting vector, such that, $w_i \in [0, 1]$ and $\sum_i w_i = 1$, $\beta_h = \frac{w_h}{\sum_2^m w_k}$, $h = 2, \ldots, m$, and B is the associated ordered unbalanced 2-tuple vector. Each element $b_i \in B$ is the i-th largest unbalanced 2-tuple in the collection $\{(a_1, \alpha_1), \ldots, (a_m, \alpha_m)\}$, and C_{un}^m is the convex combination operator of m unbalanced 2-tuples. If $w_j = 1$ and $w_i = 0$ with $i \neq j$ $\forall i, j$ the convex combination is defined as: $C_{un}^m\{w_i, b_i,\ i = 1, \ldots, m\} = b_j$. And if $m = 2$ then it is defined as:

$$C_{un}^2\{w_l, b_l,\ l = 1, 2\} = w_1 \otimes b_j \oplus (1 - w_1) \otimes b_i =$$

$$= TF_t^{t'}(s_k^{n(t')}, \alpha)$$

where $(s_k^{n(t')}, \alpha) = \Delta(\lambda)$ and $\lambda = \Delta^{-1}(TF_{t'}^t(b_i)) + w_1 \cdot (\Delta^{-1}(TF_{t'}^t(b_j)) - \Delta^{-1}(TF_{t'}^t(b_i)))$, $b_j, b_i \in (S_{un} \times [-0.5, 0.5))$, $(b_j \geq b_i)$, $\lambda \in [0, n(t') - 1]$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$.

In [11] it was defined an expression to obtain W by means of a fuzzy linguistic non-decreasing quantifier Q [12]:

$$w_i = Q(i/m) - Q((i-1)/m), \quad i = 1, \ldots, m. \quad (4)$$

3 A CONSENSUS MODEL FOR GDM PROBLEMS DEFINED IN UNBALANCED FUZZY LINGUISTIC CONTEXTS

In this section we present a consensus model defined for GDM problems defined in unbalanced fuzzy linguistic contexts. A GDM problem based on preference relations is classically defined as a decision situation where there are a set of experts, $E = \{e_1, \ldots, e_m\}$ $(m \geq 2)$, and a finite set of alternatives, $X = \{x_1, \ldots, x_n\}$ $(n \geq 2)$, and each expert e_i provides his/her preferences about X by means of a preference relation, $P_{e_i} \subset X \times X$, where the value $\mu_{P_{e_i}}(x_l, x_k) = p_i^{lk}$ is interpreted as the preference degree of the alternative x_l over x_k for e_i. In this paper, we deal with GDM problems defined in unbalanced fuzzy linguistic contexts, i.e., GDM problems where the experts e_i express their preferences relations $P_{e_i} = (p_i^{lk})$ on the set of alternatives X using a linguistic term set that is not uniformly and symmetrically distributed, S_{un}.

This consensus model presents the following main characteristics:

1. It is designed to guide the consensus process of unbalanced fuzzy linguistic GDM problems.

2. It uses a new methodology to manage unbalanced fuzzy linguistic information.

3. It is based on two consensus criteria: *consensus degrees* and *proximity measures*. The first ones are used to measure the agreement amongst all the experts, while the second ones are used to learn how close the collective and individual expert's preference are. Both consensus criteria are calculated at three different levels: pair of alternatives, alternatives and relation. It will allow us to know the current state of consensus from different viewpoints, and therefore, to guide more correctly the consensus reaching processes.

4. A *feedback mechanism* is defined using the above consensus criteria. It substitutes the moderator's actions, avoiding the possible subjectivity that he/she can introduce, and gives advice to the experts to find out the changes they need to make in their opinions in order to obtain the highest degree of consensus possible.

In particular, our consensus model develops its activity in three phases (Figure 4): *computing consensus degrees, controlling the consensus state* and *feedback mechanism*.

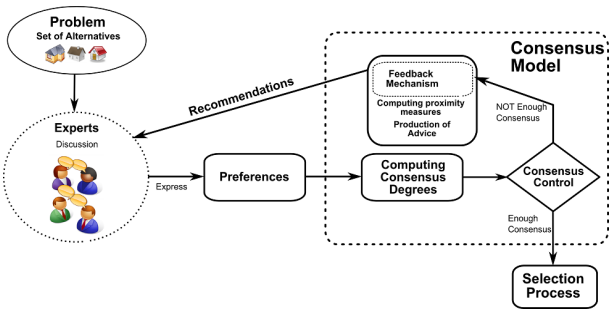


Figure 4: Consensus model

3.1 COMPUTING CONSENSUS DEGREES

The consensus degrees are used to measure the current level of consensus in the decision process. They are given at three different levels: pairs of alternatives, alternatives and relations. The computation of the consensus degrees is carried out as follows:

1. For each pair of experts, e_i, e_j $(i < j)$, a similarity matrix, $SM_{ij} = (sm_{ij}^{lk})$, is defined where

$$sm_{ij}^{lk} = 1 - \frac{|\Delta_{t'}^{-1}(TF_{t'}^{t}(p_i^{lk})) - \Delta_{t'}^{-1}(TF_{t'}^{t}(p_j^{lk}))|}{n(t') - 1}. \tag{5}$$

   being $p_i^{lk} = (s_v^{n(t)}, \alpha_1)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$ and $p_j^{lk} = (s_w^{n(t)}, \alpha_2)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$.

2. A consensus matrix, CM, is calculated by aggregating all the similarity matrices using the arithmetic mean as the aggregation function ϕ:

$$cm^{lk} = \phi(sm_{ij}^{lk}). \tag{6}$$

3. Once the consensus matrix, CM, is computed, we proceed to calculate the consensus degrees at the three different levels:

   (a) **Level 1.** *Consensus degree on pairs of alternatives,* cp^{lk}. It measures the agreement on the pair of alternatives (x_l, x_k) amongst all the experts.

$$cp^{lk} = cm^{lk}. \tag{7}$$

   (b) **Level 2.** *Consensus degree on alternatives,* ca^l. It measures the agreement on an alternative x_l amongst all the experts.

$$ca^l = \frac{\sum_{k=1, k\neq l}^{n} cp^{lk}}{n-1}. \tag{8}$$

   (c) **Level 3.** *Consensus degree on the relation,* cr. It measures the global consensus degree amongst the experts' opinions.

$$cr = \frac{\sum_{l=1}^{n} ca^l}{n}. \tag{9}$$

3.2 CONTROLLING THE CONSENSUS STATE

The consensus state control process involves deciding when the consensus process should be finished. To do so, a minimum consensus threshold, $\gamma \in [0, 1]$, is fixed before applying the consensus model. When the consensus measure, cr, satisfies this value, γ, the consensus model finishes and the selection process is applied to obtain the solution. Otherwise, the feedback mechanism is applied. To avoid that the consensus process does not converge, a maximum number of consensus rounds, *MaxRounds*, is incorporated.

3.3 FEEDBACK MECHANISM

The feedback mechanism provides recommendations to support the experts in changing their opinions. It consists on two steps: *computation of proximity values* and *production of advice*, which are explained in detail in the following subsections.

3.3.1 Computation of Proximity Measures

These measures evaluate the agreement between the individual experts' opinions and the group opinion. To compute them for each expert, we need to obtain the collective unbalanced fuzzy linguistic preference relation, $P_{e_c} = (p_c^{lk})$, calculated by means of the aggregation of the set of individual unbalanced fuzzy linguistic preference relations $\{P_{e_1}, \ldots, P_{e_m}\}$ as follows

$$p_c^{lk} = \phi_{un}(p_1^{lk}, \ldots, p_m^{lk}). \tag{10}$$

with ϕ_{un} the $LOWA_{un}$ operator defined in section 2.2.2.

Once P_{e_c} is obtain, we can compute the proximity measures carrying out the following two steps:

1. For each expert, e_i, a proximity matrix, $PM_i = (pm_i^{lk})$, is obtained where

$$pm_i^{lk} = 1 - \frac{\left|\Delta_{t'}^{-1}(TF_{t'}^{t}(p_i^{lk})) - \Delta_{t'}^{-1}(TF_{t'}^{t}(p_c^{lk}))\right|}{n(t') - 1}. \tag{11}$$

   being $p_i^{lk} = (s_v^{n(t)}, \alpha_1)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$ and $p_c^{lk} = (s_w^{n(t)}, \alpha_2)$, $t \in \{t^-, t_2^-, t^+, t_2^+\}$.

2. Computation of proximity measures at three different levels:

(a) **Level 1.** *Proximity measure on pairs of alternatives,* pp_i^{lk}. It measures the proximity between the preferences, on each pair of alternatives , of the expert, e_i, and the group.

$$pp_i^{lk} = pm_i^{lk}. \tag{12}$$

(b) **Level 2.** *Proximity measure on alternatives,* pa_i^l. It measures the proximity between the preferences, on each alternative, x_l, of the expert, e_i, and the group.

$$pa_i^l = \frac{\sum_{k=1,k\neq l}^{n} pp_i^{lk}}{n-1}. \tag{13}$$

(c) **Level 3.** *Proximity measure on the relation,* pr_i. It measures the global proximity between the preferences of each expert, e_i, and the group.

$$pr_i = \frac{\sum_{l=1}^{n} pa_i^l}{n}. \tag{14}$$

3.3.2 Production of Advice

The production of advice to achieve a solution with the highest degree of consensus possible is carried out in two steps: *Identification rules* and *Direction rules*.

1. **Identification rules (IR).** We must identify the experts, alternatives and pairs of alternatives that are contributing less to reach a high degree of consensus and, therefore, should participate in the change process.

   (a) *Identification rule of experts (IR.1).* It identifies the set of experts, $EXPCH$, that should receive advice on how to change some of their preference values.

   $$EXPCH = \{e_i \mid pr_i < \gamma\} \tag{15}$$

   (b) *Identification rule of alternatives (IR.2).* It identifies the alternatives whose associated assessments should be taken into account by the above experts in the change process of their preferences.

   $$ALT = \{x_l \in X \mid ca^l < \gamma\} \tag{16}$$

   (c) *Identification rule of pairs of alternatives (IR.3).* It identifies the particular pairs of alternatives (x_l, x_k) whose respective associated assessments p_i^{lk} the expert e_i should change.

   $$PALT_i = \{(x_l, x_k) \mid x_l \in ALT \wedge e_i \in EXPCH \wedge pp_i^{lk} < \gamma\} \tag{17}$$

2. **Direction rules (DR).** We must find out the direction of change to be applied to the preference assessment p_i^{lk}, with $(x_l, x_k) \in PALT_i$. To do this, we define the following two direction rules.

   (a) *DR.1.* If $p_i^{lk} > p_c^{lk}$, the expert e_i should decrease the assessment associated to the pair of alternatives (x_l, x_k), i.e., p_i^{lk}.

   (b) *DR.2.* If $p_i^{lk} < p_c^{lk}$, the expert e_i should increase the assessment associated to the pair of alternatives (x_l, x_k), i.e., p_i^{lk}.

4 Example of Application

Let us suppose that three different experts $E = \{e_1, e_2, e_3\}$ provide the following unbalanced fuzzy linguistic preference relations over a set of four alternatives $X = \{x_1, x_2, x_3, x_4\}$ using the unbalanced linguistic term set $S_{un} = \{N, VL, L, M, H, QH, VH, T\}$ (see Figure 1 and Figure 3):

$$P_{e_1} = \begin{pmatrix} - & (H,0) & (QH,0) & (L,0) \\ (VL,0) & - & (M,0) & (H,0) \\ (L,0) & (M,0) & - & (L,0) \\ (VH,0) & (VL,0) & (VH,0) & - \end{pmatrix}$$

$$P_{e_2} = \begin{pmatrix} - & (VL,0) & (L,0) & (VH,0) \\ (H,0) & - & (QH,0) & (T,0) \\ (QH,0) & (L,0) & - & (H,0) \\ (VL,0) & (N,0) & (L,0) & - \end{pmatrix}$$

$$P_{e_3} = \begin{pmatrix} - & (VL,0) & (M,0) & (VH,0) \\ (H,0) & - & (QH,0) & (L,0) \\ (M,0) & (L,0) & - & (T,0) \\ (L,0) & (H,0) & (N,0) & - \end{pmatrix}$$

FIRST ROUND

In the following, we show how to apply each step of the consensus model.

1. **Computing consensus degrees:**

   (a) *Similarity matrices:*

$$SM_{12} = \begin{pmatrix} - & 0.50 & 0.50 & 0.37 \\ 0.50 & - & 0.75 & 0.62 \\ 0.50 & 0.75 & - & 0.62 \\ 0.25 & 0.87 & 0.37 & - \end{pmatrix}$$

$$SM_{13} = \begin{pmatrix} - & 0.50 & 0.75 & 0.37 \\ 0.50 & - & 0.75 & 0.62 \\ 0.75 & 0.75 & - & 0.25 \\ 0.37 & 0.50 & 0.12 & - \end{pmatrix}$$

$$SM_{23} = \begin{pmatrix} - & 1.00 & 0.75 & 1.00 \\ 1.00 & - & 1.00 & 0.25 \\ 0.75 & 1.00 & - & 0.62 \\ 0.87 & 0.37 & 0.75 & - \end{pmatrix}$$

(b) *Consensus matrix:*

$$CM = \begin{pmatrix} - & 0.66 & 0.66 & 0.58 \\ 0.66 & - & 0.83 & 0.49 \\ 0.66 & 0.83 & - & 0.49 \\ 0.49 & 0.58 & 0.41 & - \end{pmatrix}$$

(c) *Consensus degrees on pairs of alternatives.* The element (l, k) of CM represents the consensus degrees on the pair of alternatives (x_l, x_k).

(d) *Consensus on alternatives:*

$$ca^1 = 0.63 \quad ca^2 = 0.66 \quad ca^3 = 0.66 \quad ca^4 = 0.49$$

(e) *Consensus on the relation:*

$$cr = 0.61$$

2. **Controlling the consensus state:** In this example, we have decided to use the value, $\gamma = 0.75$. Because $cr < \gamma$, then it is concluded that there is no consensus amongst the experts, and consequently, the consensus model computes the proximity measures to support the experts on the necessary changes in their preferences in order to increase cr.

3. **Feedback mechanism:** To calculate P_{e_c}, the systems uses the $LOWA_{un}$ operator and the linguistic quantifier *most of* defined as $Q(r) = r^{1/2}$, which applying (4), generates the following weighting vector $W = \{0.58, 0.24, 0.18\}$.

$$P_{e_c} = \begin{pmatrix} - & (M, -0.34) & (H, -0.20) & (QH, 0.10) \\ (M, 0.28) & - & (QH, -0.37) & (QH, 0.20) \\ (H, -0.20) & (M, -0.42) & - & (QH, 0.20) \\ (H, -0.28) & (M, -0.43) & (H, -0.46) & - \end{pmatrix}$$

(a) **Computation of proximity measures.**

i. *Proximity matrices:*

$$PM_1 = \begin{pmatrix} - & 0.83 & 0.85 & 0.48 \\ 0.59 & - & 0.79 & 0.85 \\ 0.65 & 0.95 & - & 0.47 \\ 0.71 & 0.68 & 0.69 & - \end{pmatrix}$$

$$PM_2 = \begin{pmatrix} - & 0.67 & 0.65 & 0.89 \\ 0.91 & - & 0.95 & 0.77 \\ 0.85 & 0.80 & - & 0.85 \\ 0.53 & 0.55 & 0.68 & - \end{pmatrix}$$

$$PM_3 = \begin{pmatrix} - & 0.66 & 0.90 & 0.89 \\ 0.91 & - & 0.95 & 0.47 \\ 0.90 & 0.80 & - & 0.77 \\ 0.66 & 0.82 & 0.43 & - \end{pmatrix}$$

ii. *Proximity on pairs of alternatives for expert e_i are given in PM_i.*

iii. *Proximity on alternatives (See Table 1):*

Table 1: Proximity measures on alternatives

x_1	x_2	x_3	x_4
$pa_1^1 = 0.72$	$pa_1^2 = 0.74$	$pa_1^3 = 0.69$	$pa_1^4 = 0.69$
$pa_2^1 = 0.73$	$pa_2^2 = 0.87$	$pa_2^3 = 0.83$	$pa_2^4 = 0.58$
$pa_3^1 = 0.82$	$pa_3^2 = 0.78$	$pa_3^3 = 0.82$	$pa_3^4 = 0.64$

iv. *Proximity on the relation:*

$$pr_1 = 0.71 \quad pr_2 = 0.75 \quad pr_3 = 0.76$$

(b) **Production of advice.**

i. *Identification rules.*
(IR.1) Set of experts to change their preferences, $EXPCH$:

$$EXPCH = \{e_i \mid pr_i < 0.75\} = \{e_1\}$$

(IR.2) Set of alternatives whose assessments should be considered in the change process, ALT:

$$ALT = \{x_l \in X \mid ca^l < 0.75\} = \{x_1, x_2, x_3, x_4\}$$

(IR.3) Set of pairs of alternatives whose associated assessments should change, $PALT_i$:

$$PALT_1 = \{(x_1, x_4), (x_2, x_1), (x_3, x_1), (x_3, x_4), (x_4, x_1), (x_4, x_2), (x_4, x_3)\}$$

which gives the following list of preference values:

$$p_1^{14} \quad p_1^{21} \quad p_1^{31} \quad p_1^{34} \quad p_1^{41} \quad p_1^{42} \quad p_1^{43}$$

ii. *Direction rules.*
Because $p_1^{14} < p_c^{14}$, $p_1^{21} < p_c^{21}$, $p_1^{31} < p_c^{31}$, $p_1^{34} < p_c^{34}$, $p_1^{42} < p_c^{42}$ and $p_1^{41} > p_c^{41}$, $p_1^{43} > p_c^{43}$ expert e_1 is advised to increase the assessment of the first five preference values and decrease the assessment of the last two preference values.

SECOND ROUND

1. **Providing new preferences:** In this example, we suppose that expert e_1 follows the advice given, and thus, his/her new preferences is as follows:

$$P_{e_1} = \begin{pmatrix} - & (H, 0) & (QH, 0) & (\mathbf{VH}, \mathbf{0}) \\ (\mathbf{M}, \mathbf{0}) & - & (M, 0) & (H, 0) \\ (\mathbf{M}, \mathbf{0}) & (M, 0) & - & (\mathbf{H}, \mathbf{0}) \\ (\mathbf{VL}, \mathbf{0}) & (\mathbf{M}, \mathbf{0}) & (\mathbf{L}, \mathbf{0}) & - \end{pmatrix}$$

2. **Computing consensus degrees.**

(a) *Similarity matrices:*

$$SM_{12} = \begin{pmatrix} - & 0.50 & 0.50 & 1.00 \\ 0.87 & - & 0.75 & 0.62 \\ 0.75 & 0.75 & - & 0.62 \\ 1.00 & 0.50 & 0.50 & - \end{pmatrix}$$

$$SM_{13} = \begin{pmatrix} - & 0.50 & 0.75 & 1.00 \\ 0.87 & - & 0.75 & 0.62 \\ 1.00 & 0.75 & - & 0.62 \\ 0.87 & 0.87 & 0.25 & - \end{pmatrix}$$

$$SM_{23} = \begin{pmatrix} - & 1.00 & 0.75 & 1.00 \\ 1.00 & - & 1.00 & 0.25 \\ 0.75 & 1.00 & - & 0.62 \\ 0.87 & 0.37 & 0.75 & - \end{pmatrix}$$

(b) *Consensus matrix:*

$$CM = \begin{pmatrix} - & 0.66 & 0.66 & 1.00 \\ 0.91 & - & 0.83 & 0.50 \\ 0.83 & 0.83 & - & 0.62 \\ 0.91 & 0.58 & 0.83 & - \end{pmatrix}$$

(c) *Consensus degrees on pairs of alternatives.* The element (l, k) of CM represents the consensus degrees on the pair of alternatives (x_l, x_k).

(d) *Consensus on alternatives:*

$$ca^1 = 0.77 \quad ca^2 = 0.75 \quad ca^3 = 0.76 \quad ca^4 = 0.77$$

(e) *Consensus on the relation:*

$$cr = 0.76$$

3. **Controlling the consensus state:** As we can observe, the changes in the preference values introduced result in an increasing of the global consensus from 0.61 to 0.76. The minimum consensus threshold is reached, $cr = 0.76 > \gamma = 0.75$, and, therefore, the consensus model would stop and the selection process would be applied to obtain the final solution of consensus.

5 CONCLUDING REMARKS

In this paper we have proposed a consensus model for GDM problems with unbalanced fuzzy linguistic information which allows to manage consensus processes based on similarity measures among preferences and to build recommendation systems of preferences to support the consensus process automatically, without moderator.

Acknowledgements

This work has been supported by the Research Projects TIN2007-61079 and SAINFOWEB-PAI00602.

References

[1] J.L. García-Lapresta. A general class of simple majority decision rules based on linguistic opinions. *Information Sciences* **176:4** (2006) 352-365.

[2] F. Herrera, E. Herrera-Viedma, L. Martinez. A Hierarchical Ordinal Model for Managing Unbalanced Linguistic Term Sets based on the Linguistic 2-tuple Model. *Eurofuse Workshop on Preference Modelling and Applications*, Granada (Spain), April 2001, 201-206.

[3] F. Herrera, E. Herrera-Viedma and J.L. Verdegay. Direct approach processes in group decision making using linguistic OWA operators. *Fuzzy Sets and Systems* **79** (1996) 175-190.

[4] F. Herrera, E. Herrera-Viedma and J.L. Verdegay: A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems* **78** (1996) 73-87.

[5] F. Herrera and L. Martínez. A 2–tuple fuzzy linguistic representation model for computing with words. *IEEE Transaction on Fuzzy Systems* **8:6** (2000) 746-752.

[6] F. Herrera and L. Martínez. A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making. *IEEE Trans. on Sys., Man and Cyb. Part B: Cyb.* **31:2** (2001) 227-234.

[7] E. Herrera-Viedma and A.G. López-Herrera. A model of information retrieval system with unbalanced fuzzy linguistic information. *International Journal of Intelligent Systems* **22:11** (2007) 1197-1214.

[8] J. Kacprzyk, H. Nurmi and M. Fedrizzi. *Consensus under fuzziness* (Boston: Kluwer Academic Publishers, 1997).

[9] V. Torra. Aggregation of linguistic labels when semantics is based on antonyms. *International Journal of Intelligent Systems* **16** (2001) 513-524.

[10] Z.S. Xu. A method for multiple attribute decision making with incomplete weight information in linguistic setting. *Knowledge-Based Systems* **20:8** (2007) 719-725.

[11] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. on Sys., Man and Cyb. Part B: Cyb.* **10:1** (1988) 183-190.

[12] L.A. Zadeh. A computational approach to fuzzy quantifiers in natural language. *Computation & Mathematics with Applications* **9** (1983) 149-184.

Un procedimiento de consenso para la ordenación de alternativas en problemas de decisión en grupo con información imprecisa.

Ignacio Contreras[1] **Amparo M. Mármol**[2] **Miguel A. Hinojosa**[1]

[1] Departamento de Economía, Mét. Cuantitativos e H. E. Universidad Pablo de Olavide, {iconrub, mahinram}@upo.es
[2] Departamento de Economía Aplicada III. Universidad de Sevilla, amarmol@us.es

Resumen

En el presente trabajo proponemos un procedimiento para la resolución problemas de decisión en grupo con múltiples criterios en los que cada agente decisor ofrece información imprecisa sobre la importancia relativa de cada criterio. La información parcial que proporciona cada agente sobre los pesos que debe asignarse a los criterios es transformada en información ordinal sobre las alternativas, determinando una ordenación de las mismas. A partir de dichas ordenaciones, se obtiene una solución de consenso mediante la aplicación de modelos basados en distancias.

Palabras Clave: Decisión en grupo, consenso, información parcial.

1 INTRODUCCIÓN

Tratamos problemas de decisión en grupo en situaciones en las que un conjunto de agentes ha de seleccionar la mejor de entre un conjunto de alternativas, o bien proporcionar una ordenación completa de éstas. En nuestro contexto, el conjunto de alternativas debe ser evaluado con respecto a múltiples criterios. En estas situaciones, cada miembro del grupo puede ofrecer su particular opinión sobre la importancia relativa que dichos criterios deben jugar en el proceso de toma de decisiones y dichas preferencias u opiniones de los individuos deben ser un elemento determinante del resultado final del proceso de decisión.

Algunos de los procedimientos desarrollados recientemente para el tratamiento de los problemas de decisión en grupo incorporan la posibilidad de operar con información incompleta o imprecisa, pudiendo ser aplicados en situaciones en las que los miembros del grupo no son capaces de ofrecer una representación precisa de sus preferencias con respecto a los elementos del problema. Véase por ejemplo [3], donde se proponen índices compuestos con elección flexible de pesos o [10] donde se presenta un procedimiento de agregación basado en la estimación de intervalos de utilidades. Una revisión de los trabajos en los que se desarrollan procedimientos de decisión multicriterio con información parcial puede verse, por ejemplo, en [6]. Otras contribuciones recientes en esta misma línea son, entre otras, [4], [9] y [2].

En el presente trabajo proponemos un procedimiento para la decisión en grupo con múltiples criterios para aquellos problemas en los que el objetivo es la obtención de una ordenación completa del conjunto de alternativas. Este procedimiento esta especialmente indicado para aquellos casos en los que los individuos solo son capaces de proporcionar información incompleta sobre sus preferencias respecto a los criterios.

El resto del trabajo se estructura de la siguiente forma. En la Sección 2 se describe el problema de decisión en grupo y se introduce el procedimiento propuesto para la agregación de preferencias individuales. La Sección 3 incluye un ejemplo que ilustra el procedimiento propuesto. En la última Sección se recogen algunas conclusiones del trabajo.

2 PROCEDIMIENTO DE DECISIÓN EN GRUPO

Consideremos un problema de decisión en grupo con múltiples criterios en el que M alternativas, $X = \{x_1, \ldots, x_M\}$, han sido evaluadas con respecto a N criterios. Las evaluaciones que cada alternativa recibe con respecto a cada criterio se representan mediante una matriz $A \in \mathcal{R}^{N \times M}$ tal que a_{ij} con $i = 1, \ldots, N$, $j = 1, \ldots, M$ denota el valor que la alternativa x_j recibe con respecto al citerio $i-$ésimo.

Consideramos que K decisores (DMs) proporcionan

información sobre sus preferencias con respecto a la importancia relativa de cada criterio. Además, asumimos que las preferencias pueden representarse mediante una función de valor aditiva. Esto es, el valor que el k-ésimo DM asocia a la alternativa x_h se obtiene como

$$V^k(x_h) = \sum_{i=1}^{N} w_i^k a_{ih}, \tag{1}$$

donde w^k denota el vector de pesos que representa la importancia relativa de los criterios para el agente k.

Consideramos el caso en el que los parámetros del problema no están determinados de manera precisa. La información que proporcionan los agentes está expresada mediante conjunto de información parcial sobre los vectores de pesos w que denotamos por $\Phi^k \subseteq \mathcal{R}^N$, $k = 1, \ldots, K$.

Un conjunto de información parcial consiste en aquellos vectores de pesos que un determinado decisor acepta como razonables para representar sus preferencias respecto a la importancia relativa de los criterios. Por convenio, consideramos que dichos vectores están normalizados de manera que sus componentes suman la unidad. Esto es, $\Phi^k \subseteq \{w^k \in \mathcal{R}^N, \sum_{i=1}^{N} w_i^k = 1, w_i^k \geq 0, i = 1, \ldots, N\}$, para cada $k = 1, \ldots, K$. En particular, consideramos aquellos casos en los que las preferencias individuales vienen dadas a través de relaciones lineales entre las componentes de dichos vectores. En tal caso, los conjuntos Φ^K son poliédricos y quedan descritos por relaciones lineales de los pesos.

En esta situación, el objetivo final es obtener una ordenación de las alternativas que tenga en cuenta, en la medida de lo posible, las preferencias individuales de todos los agentes. La idea básica del procedimiento propuesto es la determinación de una solución de compromiso a partir de los rankings u ordenaciones que pueden inducirse a partir de los conjuntos de información de cada agente. De manera implícita en este problema, se supone la existencia de una función de distancia sobre el conjunto de ordenaciones de alternativas, que verifica un conjunto de propiedades deseables en el contexto de la elección social. Un estudio detallado de modelos basados en funciones de distancia puede verse, entre otros, en [5].

Sean R_h y $R_k \in \mathcal{R}^N$ dos vectores de prioridad que representan los rankings de alternativas de los agentes h y k. Por r_{hj} denotamos la posición asignada a la alternativa x_j por el DM h. Por convenio, suponemos que la primera posición se asigna a la alternativa más valorada y que, en caso de empates, se asigna a las alternativas la media de los valores que corresponderían a las alternativas empatadas. Por $\mathcal{S}$ denotamos el conjunto de vectores de prioridad que pueden representar preferencias ordinales para el conjunto de alternativas X.

Para medir el desacuerdo entre las preferencias individuales del agente k y la solución de grupo consideramos funciones basadas en distancias. Pueden plantearse funciones de desacuerdo basadas en cualquier métrica l_p, aunque los esfuerzos en el contexto de la elección social se han centrado en las métricas l_1 y l_2. En ambos casos, la solución que se obtiene puede interpretarse en términos estadísticos, como la mediana y media respectivamente. Además, la obtención de las soluciones consiste en la resolución de un problema de asignación en el caso de la métrica l_1 y un problema de mínima varianza para la l_2. Con la propuesta que aparece en el trabajo de González-Pachón y Romero [7], el tratamiento de los los modelos basados en las métricas l_1 y l_∞ puede hacerse utilizando la Programación por Metas.

En el presente trabajo queremos enfatizar el que el desacuerdo de cada individuo tenga en cuenta la posición de todas las alternativas, y no de una única, como resulta de los modelos basados en la métrica l_∞, por lo que limitamos nuestro análisis a una función de compromiso basada en la métrica l_1.

Consideramos la siguiente función de distancia basada en la métrica l_1,

$$d(R_k, R_s) = \sum_{j=1}^{M} |r_{kj} - r_{sj}|. \tag{2}$$

En este contexto, definimos un vector de consenso R_G como el vector de prioridad que resuelve el siguiente problema

$$\begin{array}{ll} min & \sum_{k=1}^{K} d(R_k, R_G) \\ s.t. & R_G \in \mathcal{S} \end{array}. \tag{3}$$

Es importante destacar que la utilización de conjuntos de información parcial para representar las preferencias individuales implica que cada agente aceptaría una evaluación de la alternativa x_h tal que $V^k(x_h) = \sum_{i=1}^{N} w_i^k a_{hi}$, siempre que $w^k \in \Phi^k$. Consecuentemente, cada ordenación de alternativas que pueda obtenerse a partir de estos valores es aceptado por el agente k-ésimo. Es decir, a partir de las preferencias individuales de cada decisor, pueden inducirse diferentes rankings (al menos uno, ya que en otro caso $\Phi^k = \emptyset$). El procedimiento de decisión ha de incluir un criterio que seleccione cuál de los rankings individuales será considerado para la determinación de la solución de grupo. En nuestro caso, cada individuo se representa por aquel ranking más próximo a la solución de grupo, aquel que minimiza la función de desacuerdo (2).

A partir de (1) puede derivarse una relación de dominancia entre alternativas. Para un vector de pesos dado $w^k \in \Phi^k$, puede afirmarse que la alternativa x_j domina estrictamente a la alternativa x_h cuando se considera la estructura de preferencia del agente k, si $V^k(x_j) - V^k(x_h) > 0$. El ranking de grupo se determinará comparando los valores agregados de las alternativas para diferentes valores de los vectores $w^k \in \Phi^k$. Por tanto, se hace necesario un paso intermedio consistente en la determinación de un ranking individual para cada agente. Para tal fin consideramos el siguiente grupo de restricciones en las que δ_{ij} y γ_{ij} son variables binarias:

$$\begin{array}{l} V^k(x_h) - V^k(x_j) + \delta^k_{hj} B \geq 0, \quad \forall\, h \neq j, \\ V^k(x_h) - V^k(x_j) + \gamma^k_{hj} B \geq \varepsilon, \quad \forall\, h \neq j; \end{array} \tag{4}$$

donde B denota un número suficientemente grande y ε es un factor de discriminación entre alternativas, que se incluye con un propósito computacional. Esto es, puede afirmarse que la alternativa x_h domina de manera estricta a x_j si $V^k(x_h) - V^k(x_j) \geq \varepsilon$.

Las variables δ^k_{hj} deben tomar valor 1 cada vez que x_h no domine a x_j, es decir siempre que $V^k(x_j) > V^k(x_h)$. Por otro lado, γ^k_{hj} tomará valor 1 si $V^k(x_j) \geq V^k(x_h)$, es decir, siempre que x_h no sea preferida estrictamente a x_j con las preferencias del agente k-ésimo. Esta variable se incluye en el modelo para detectar los empates entre alternativas.

En la siguiente tabla se resumen los valores que tomarán las anteriores variables en cada caso,

Tabla 1: Variables binarias

	δ^k_{hj}	δ^k_{jh}	γ^k_{hj}	γ^k_{jh}
$V^k(x_h) > V^k(x_j)$	0	1	0	1
$V^k(x_h) = V^k(x_j)$	0	0	1	1
$V^k(x_h) < V^k(x_j)$	1	0	1	0

Para asegurar que la asignación de los valores se hace de manera adecuada incluimos los siguientes conjuntos de restricciones,

$$\delta^k_{hj} + \delta^k_{jh} \leq 1, \forall\, h \neq j; \tag{5}$$

$$\delta^k_{hj} + \gamma^k_{hj} + \gamma^k_{jh} \leq 2, \forall\, h \neq j; \tag{6}$$

$$\delta^k_{hj} + \gamma^k_{hj} + \gamma^k_{jh} \geq 1, \forall\, h \neq j; \tag{7}$$

$$\delta^k_{hj} + \gamma^k_{jh} = 1, \forall\, h \neq j; \tag{8}$$

Las restricciones (5)-(8) son necesarias para inducir el ranking de alternativas a partir de los valores $V^k(x_i)$. Estas relaciones garantizan que la asignación de valores a las variables δ^k_{hj} y γ^k_{kj} es correcta en el sentido de que solo si x_j domina estrictamente a x_h se tiene que $\delta^k_{hj} = 1$, y consecuentemente $\delta^k_{jh} = 0$. Los valores de γ^k_{hj} dependerán también de los valores asignados a δ^k_{hj} y permitirán representar empates entre alternativas. Es interesante hacer notar que en cada restricción es necesario, no solo considerar la relación de dominancia de x_h sobre x_j, sino también la de x_j sobre x_h.

La posición que la alternativa x_h ocupa en el vector de prioridad R_k se obtiene considerando tanto el número de alternativas que dominan a x_h como el número de alternativas empatadas con x_h. Consideramos la siguiente expresión

$$r_{kh} = \sum_{h \neq j} \frac{\delta^k_{hj} + \gamma^k_{hj}}{2} + 1, \forall\, h = 1, \ldots, M. \tag{9}$$

En ella se contabiliza una unidad por cada alternativa que domine de manera estricta a x_h, en cuyo caso tanto δ^k_{hj} como γ^k_{hj} son iguales a uno para el correspondiente valor de j, mientras que en el caso de empates se computará 0.5, ya que en tal caso γ^k_{hj} es igual a 1 y δ^k_{hj} es igual a 0 para el correspondiente valor de j.

Para imponer que los vectores R_k representen ranking de alternativas en la forma referida, incluimos la siguiente condición,

$$\sum_{j=1}^{M} r_{kj} = \frac{M(M+1)}{2}, \forall\, k = 1, \ldots, K. \tag{10}$$

Junto a las restricciones (5)-(8), la condición (10) asegura que el vector R_k está contenido en el conjunto $\mathcal{S}$, o, lo que es lo mismo, representa un vector de prioridad.

Para obtener la solución compromiso han de considerarse todas las variables incluidas en el modelo para el grupo: un vector de pesos, valores agregados y rankings inducidos a partir de éstos. Para representar esta situación, incluimos un *nuevo agente* en el conjunto de individuos con el subíndice G que representa los valores de grupo[1].

Para determinar la solución de grupo, el modelo opera con la función definida en (2). La minimización de esa función no lineal puede reducirse a un problema lineal considerando el cambio de variables propuesto en [7],

$$\begin{array}{l} \alpha_{kj} = \frac{1}{2}\left[|r_{kj} - r_{Gj}| + (r_{kj} - r_{Gj})\right] \\ \beta_{kj} = \frac{1}{2}\left[|r_{kj} - r_{Gj}| + (r_{kj} + r_{Gj})\right]. \end{array} \tag{11}$$

Incluimos en el modelo el siguiente conjunto de restricciones que permite medir la distancia entre los vectores

[1]Esto supone que el índice k que representa a los DMs variará desde 1 hasta G, esto es, $k = 1 \ldots, K, G$.

individuales y de grupo,

$$r_{kj} - r_{Gj} + \alpha_{kj} - \beta_{kj} = 0, \forall\, k = 1, \ldots, K; j = 1, \ldots, N. \tag{12}$$

De esta forma, la distancia entre R_k y R_G puede expresarse mediante la siguiente expresión,

$$d(R_k, R_G) = \sum_{j=1}^{M} (\alpha_{kj} + \beta_{kj}). \tag{13}$$

Considerando todo lo expuesto, el modelo a resolver es,

$$\begin{array}{lll}
min & \sum_{k=1}^{K} \sum_{j=1}^{M} (\alpha_{kj} + \beta_{kj}) & \\
s.t. & V^k(x_h) = \sum_{i=1}^{N} w_i^k a_{ih}, & \forall\, h, k \\
& V^k(x_h) - V^k(x_j) + \delta_{hj}^k B \geq 0, & \forall\, h \neq j, \forall\, k \\
& V^k(x_h) - V^k(x_j) + \gamma_{hj}^k B \geq \varepsilon, & \forall\, h \neq j, \forall\, k \\
& \delta_{hj}^k + \delta_{jh}^k \leq 1, & \forall\, h \neq j, \forall\, k \\
& \delta_{hj}^k + \gamma_{hj}^k + \gamma_{jh}^k \leq 2, & \forall\, h \neq j, \forall\, k \\
& \delta_{hj}^k + \gamma_{hj}^k + \gamma_{jh}^k \geq 1, & \forall\, h \neq j, \forall\, k \\
& \delta_{hj}^k + \gamma_{jh}^k = 1, & \forall\, h \neq j, \forall\, k \\
& r_{kh} = \sum_{h \neq j} \frac{\delta_{hj}^k + \gamma_{hj}^k}{2} + 1, & \forall\, h, k \\
& \sum_{h=1}^{M} r_{kh} = \frac{M(M+1)}{2} & \forall\, k \\
& r_{kh} - r_{Gh} + \alpha_{kh} - \beta_{kh} = 0, & \forall\, h, k \neq G \\
& w^k \in \Phi^k, & \forall\, k \\
& \delta_{hj}^k, \gamma_{hj}^k \in \{0, 1\}, & \forall\, h, j, k \\
& B, \varepsilon \geq 0. &
\end{array} \tag{14}$$

Deben destacarse algunos aspectos del modelo. En primer lugar, cada DM se representa por uno de los rankings que pueden inducirse con sus preferencias. En particular, el modelo selecciona aquella ordenación más favorable en términos de desacuerdo, es decir, aquel que minimiza el desacuerdo total del grupo (que no necesariamente coincide con el que minimizaría el desacuerdo individual de dicho agente). Este ranking se genera a partir de los valores agregados $V^k(x_i)$, por lo que puede considerarse que el modelo selecciona, para cada DM, aquel vector de pesos que minimiza el desacuerdo.

Los valores agregados para el grupo, $V^G(x_h)$, no tienen significado desde un punto de vista de la valoración de las alternativas, únicamente se incluyen para crear el ranking de alternativas en el formato establecido.

Además, el procedimiento propuesto cumple ciertas propiedades en el contexto de elección social como son la factibilidad, anonimato, monotonía, neutralidad y carácter no dictatorial, que se siguen directamente de la forma en la que se ha construido el procedimiento. Esto implica que siempre puede obtenerse una solución de grupo, que todos los miembros del mismo y alternativas son tratados de igual forma y que no existe un único decisor cuyas preferencias determinan la solución de grupo.

En algunos casos, un DM puede entender que la solución obtenida no refleja apropiadamente sus preferencias individuales. En tales casos, el agente puede proponer modificaciones de su conjunto de información Φ^k, para ofrecer información más refinada sobre sus preferencias.

Un ejemplo es el caso en que el k-ésimo DM prefiere en cualquier caso la alternativa x_h sobre x_j. Esto implica que la alternativa x_h debe situarse por encima de x_j en cualquiera de los rankings inducidos a partir del conjunto de información de dicho agente. Consecuentemente, debe incluirse en el conjunto Φ^k la relación $\sum_{i=1}^{N} w_i^k a_{ih} - \sum_{i=1}^{N} w_i^k a_{ij} > 0$.

Cuando el decisor proporciona a posteriori información adicional sobre sus preferencias, una etapa previa a la resolución del nuevo problema consiste en comprobar que la nueva información es consistente con la incluida en Φ^k (véase [8]). Cuando esto produce una reducción efectiva de los conjuntos de información, las nuevas relaciones deben incluirse y el modelo (14) debe ser recalculado. La inclusión de nueva información tiene como consecuencia una reducción en el conjunto de posibles rankings que el DM considera como aceptables y, posiblemente, una modificación de la solución de consenso que se obtiene de (14).

2.1 DISCRIMINANDO ENTRE SOLUCIONES MÚLTIPLES

La solución que se obtiene al resolver (14) puede no ser única. Esto puede causar dificultades en la representación de las preferencias del grupo, dado que una misma alternativa puede estar situada en posiciones diferentes en cada una de las soluciones óptimas.

Para solucionar dicha dificultad proponemos el siguiente procedimiento de refinamiento de la solución. El procedimiento esta basado en ideas análogas a las presentadas los modelos del Análisis Envolvente de Datos DEA [1] y propone evaluar a cada alternativa con la mejor valoración que ésta puede alcanzar, considerando el nivel de desacuerdo como dado. En otras palabras, se da a cada alternativa la posibilidad de ser ordenada en la mejor posición posible considerando un conjunto de restricciones comunes y al nivel de desacuerdo mínimo.

Por $\hat{D}$ representamos el valor óptimo obtenido de (14). Para cada alternativa, procedemos a determinar la mejor posición que puede ocupar preservando el nivel de desacuerdo en su valor mínimo. Para cada alternativa x_j resolvemos el siguiente problema,

$$\begin{array}{ll} min & r_{Gj} \\ s.t. & \sum_{k=1}^{K}\sum_{h=1}^{M}(\alpha_{kh}+\beta_{kh}) \leq \hat{D} \\ & \text{Restricciones de (14).} \end{array} \quad (15)$$

De esta manera se asigna la mejor posición a cada alternativa y se ordenan en función de estos valores. En caso de empates, situamos a las alternativas en la misma posición. Es importante destacar como este procedimiento proporciona una solución única.

3 EJEMPLO

Consideramos un problema en el que tres agentes han de decidir sobre siete alternativas, que denotamos por $\{x_1, \ldots, x_7\}$, y que han sido evaluadas con respecto a cinco criterios. En la Tabla 2 se muestran las evaluaciones obtenidas por las alternativas con respecto a cada uno de los criterios medidas en una escala común.

Tabla 2: Matriz de utilidades

Alt.	Crit. 1	Crit. 2	Crit. 3	Crit. 4	Crit. 5
x_1	0.298	0.056	0.024	0.074	0.041
x_2	0.106	0.096	0.017	0.095	0.272
x_3	0.059	0.056	0.155	0.101	0.088
x_4	0.037	0.101	0.224	0.081	0.097
x_5	0.016	0.022	0.079	0.108	0.198
x_6	0.027	0.079	0.231	0.061	0.032
x_7	0.138	0.096	0.117	0.074	0.069

Los agentes expresan sus preferencias con respecto a la importancia relativa de los criterios. Por ejemplo, el primer individuo considera que la importancia de los criterios está ordenada de manera decreciente según sus subíndices y que la importancia del criterio 1 no ha de ser menor al 25% del valor total. Si denotamos por $\Lambda^+ = \{w \in \mathcal{R}^5, \sum_{k=1}^{5} w_k = 1, w_k \geq 0, k = 1, \ldots, 5\}$, el conjunto de información parcial del primer agente puede formalizarse como

$$\Phi^1 = \left\{w \in \Lambda^+,\, w_i \geq w_{i+1},\, i = 1, \ldots, 4;\, w_1 \geq 0.25\right\}.$$

La información de los otros dos agentes se representa a través de los siguientes conjuntos de información parcial,

$$\Phi^2 = \{w \in \Lambda^+,\, w_i \leq w_{i+1},\, i = 1, \ldots, 4\}$$
$$\Phi^3 = \left\{w \in \Lambda^+,\, w_3 \geq \sum_{i \neq 3} w_i\right\}.$$

Esto es, el agente 2 considera la importancia de los criterios ordenada en orden inverso al agente 1 y el agente 3 considera el criterio 3 al menos tan importante como todos los demás conjuntamente.

El resultado de (14) es igual a 12 unidades pero la solución de grupo no es única. En tal caso aplicamos el procedimiento descrito en la sección anterior. Los resultados obtenidos se resumen en la Tabla 3, en la que cada columna representa un ranking de alternativas, aquél en el que dicha unidad optimiza su posición, todos ellos con el mismo nivel de desacuerdo total (si bien no necesariamente de desacuerdo individual de cada agente). El ranking final se obtiene a partir de las mejores posiciones de cada alternativa, valores que se sitúan en la diagonal principal y que destacamos en negrita.

Tabla 3: Resultados individuales de cada alternativa

	Mejor posición de x_i						
Alt.	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	**1**	1	1	2	1	2	2
x_2	4	**3**	6	6	6	6	6
x_3	5	5	**5**	5	5	5	5
x_4	2	2	2	**1**	2	1	1
x_5	7	7	7	7	**7**	7	7
x_6	6	6	4	4	4	**4**	4
x_7	3	4	3	3	3	3	**3**

Para analizar la solución consideremos los resultados obtenidos cuando se optimiza la posición de la alternativa x_1 y que se recogen en la Tabla 4. Los vectores de pesos obtenidos para cada agente son, respectivamente,

$$w^1 = (0.250, 0.211, 0.211, 0.211, 0.117),$$
$$w^2 = (0.180, 0.205, 0.205, 0.205, 0.205),$$
$$w^3 = (0.407, 0.000, 0.500, 0.000, 0.093).$$

Tabla 4: Solución óptima para la alternativa x_1.

	DM 1		**DM 2**		**DM 3**	
Alt.	$V^1(x_j)$	**Pos.**	$V^2(x_j)$	**Pos.**	$V^3(x_j)$	**Pos.**
x_1	0.112	1	0.094	4	0.137	1
x_2	0.102	4	0.118	1	0.077	6
x_3	0.091	5	0.093	5	0.110	5
x_4	0.106	2	0.110	2	0.136	2
x_5	0.071	7	0.086	7	0.064	7
x_6	0.089	6	0.087	6	0.129	3
x_7	0.103	3	0.098	3	0.121	4

Es interesante ver cómo el desacuerdo individual de cada agente puede variar de una a otra solución, puesto que solo se fija el desacuerdo total en su nivel mínimo pero no se hace ningún supuesto respecto al desacuerdo de cada agente. En este caso, los valores de desacuerdo son $d(R_1, R_G) = 0$, $d(R_2, R_G) = 6$ y $d(R_3, R_G) = 6$.

A partir de los valores de la Tabla 3 obtenemos el orden para el grupo. Las alternativas x_1 y x_4 empatan en la mejor posición, en tanto ambas podrían llegar a dicha posición cuando el desacuerdo del grupo es de 12 unidades. La solución completa es

$$\begin{pmatrix} x_1, x_4 \\ x_2, x_7 \\ x_6 \\ x_3 \\ x_5 \end{pmatrix}$$

4 CONCLUSIONES

En el presente trabajo proponemos un procedimiento para la resolución de problemas de decisión en grupo con múltiples criterios en los que los agentes únicamente ofrecen información parcial sobre sus preferencias respecto a la importancia relativa de los criterios. Se establece como supuesto que no existe intercambio de información entre los agentes por lo que no existe un incentivo a ofrecer preferencias no reales para manipular el resultado de grupo.

El procedimiento se basa en la transformación de información ordinal sobre los criterios que facilitan los agentes en información ordinal sobre las alternativas, expresada en forma de ordenaciones. Como consecuencia de que la información ofrecida es incompleta, la posición en la que cada alternativa puede situarse dentro de la ordenación puede no ser única. Por ello construimos un modelo que permite determinar una ordenación de consenso para el grupo, generando, para cada agente, una ordenación individual a partir de sus preferencias. Además proponemos un procedimiento que permite seleccionar una única ordenación de entre las múltiples soluciones que minimizan el desacuerdo total, basado en asignar a cada alternativa su mejor posición, manteniendo el nivel de desacuerdo fijo.

El trabajo se centra en el procedimiento que permite generar un ranking de alternativas a partir de cada conjunto de información parcial para, a partir de éstos, obtener una solución de consenso. Por ello, hemos considerado una única función de distancia entre vectores de prioridad. No obstante, el procedimiento propuesto puede extenderse para considerar nuevas medidas de desacuerdo basadas no sólo en distancias espaciales, sino en distancias de similitud (distancias de Kendall-Tau o de Hamming).

Referencias

[1] Charnes, A., Cooper, W.W., Rhodes, E. Measuring the efficiency of decision making units. European Journal of Operational Research 2 , Pág. 429444, 1978.

[2] Climaco, J.N., Dias, L.C. Negotiation Processes with Imprecise Information Multicriteria Additive Models. Group Decision and Negotiation, 15, Pág. 171184, 2006.

[3] Contreras, I., Hinojosa, M.A., Mármol, A.M. A class of flexible weight indices for ranking alternatives. IMA Journal of Management Mathematics 16, Pág. 71-85, 2005.

[4] Contreras, I., Mármol, A.M. A lexicographical compromise method for multiple criteria group decision problems with imprecise information. European Journal of Operational Research, 181 (3), Pág. 1530-1539, 2007.

[5] Cook, W.D., Kress, M. An extreme-point approach for obtaining weighted ratings in qualitative multicriteria decision making. Naval Research Logistic, 43, Pág. 519-531, 1996.

[6] Dias, L.C., Climaco, J.N. Dealing with imprecise information in group multicriteria decisions: A methodology and a GDSS architecture. European Journal of Operational Research, 160 (2), Pág. 291-307, 2005.

[7] González-Pachón, J., Romero, C. Distanced based consensus methods: a goal programming approach. Omega, 27, Pág. 341-347, 1999.

[8] Mármol, A.M., Puerto J., Fernández F.R. Sequential incorporation of imprecise information in multiple criteria decision processes. European Journal of Operational Research, 137, Pág. 123-133, 2002.

[9] Valadares Tavares, L. A model to support the search for consensus with conflicting rankings: Multitrident. International Transactions in Operational Research, 11, Pág. 107-115, 2004.

[10] Wang, Y.M., Yang, J.B., Xu, D.L. A preference aggregation method through the estimation of utility interval. Computers and Operations Research, 32; Pág. 2027-2049, 2005.

The Measurement of Consensus in Social Choice

Jorge Alcalde-Unzu[1] **Marc Vorsatz**[2]

[1]Departamento de Economía Aplicada IV, Universidad del País Vasco. E-mail: jorge.alcalde@ehu.es
[2]Fundación de Estudios de Economía Applicada (FEDEA). E-mail: mvorsatz@fedea.es

Abstract

We propose two ways of how to measure the consensus of opinions within a given group of individuals. According to the first rule, the consensus is determined by calculating first, for every pair of alternatives, the differences in the support of one alternative over the other and, in a second step, by averaging the differences in the support over all possible pairs of alternatives. Larger differences obviously indicate less consensus. According to the second rule, first the alikeness of any two preference relations is determined using the Kemeny [5] measure and, afterwards, the average of all Kemeny [5] measures is calculated. Our main results are axiomatic characterizations of these rules.

1 INTRODUCTION

Consensus Measures In many decision making problems, a group of individuals has to determine the social ranking of a given set of alternatives. Examples include, among others, political elections, the provision of public goods, and claims problems. The classic question is then how to aggregate individual preferences socially. The objective of this work is to analyze axiomatically the related but different problem of how to measure the *consensus* within a group of individuals, where consensus can be understood as the alikeness of individual preferences on the set of objects.

The first and probably best–known consensus measure is due to Kendall [4]. This measure, usually referred to as Kendall's tau (τ), is restricted to the case of two individuals and simply calculates the proportion of pair–wise comparisons of alternatives that both individuals agree upon. Thus, it can be reinterpreted as 1 minus the Kemeny [5] distance between the preference orderings of the two individuals. An extension to the case of any arbitrary number of individuals has been proposed by Hays [3], who suggests to calculate the *average tau* ($\bar{\tau}$), which is obtained by averaging Kendall's measure over all possible pairs of individuals. Since the limits of the average tau for different sets of individuals and alternatives may differ, he also suggests the possibility of normalizing the measure, leading to the *normalized average tau* ($\tilde{\tau}$).

Contribution We are looking for consensus measures that are consistently applicable in any *situation*; that is, for any sets of individuals and alternatives. Since we also assume that individuals have linear preferences on the set of alternatives, a consensus measure assigns, for any situation, a real number from the unit interval to every preference profile (with the implicit assumption that higher values imply more consensus).

Our paper will consider two different types of consensus measures that are intuitive. One of them is the Kendall–Hays proposal and the other, that we will call *average sigma* ($\bar{\sigma}$) and *normalized average sigma* ($\tilde{\sigma}$), consists in the following: for any pair of alternatives, take the absolute value of the difference in the number of individuals that prefer one alternative over the other and calculate the average over all possible pairs of alternatives. The measure up to this point is denoted *average sigma*. The *normalized average sigma* measure is then obtained in the same way as *normalized average tau* is constructed from *average tau*.

We first analyze whether the two normalized measures are different. We find that (i) the two measures are identical as long as the number of individuals is at most three, (ii) if the number of alternatives is two, then the two measures are ordinally equivalent meaning that they rank all preference profiles according to their consensus in the same way, and (iii) if there are at least four individuals and three alternatives, then the two measures are even in ordinal terms different from each other.

After this initial comparative study, we provide an ax-

iomatic analysis proposing a set of properties a consensus measure *should* naturally satisfy. The axioms can be classified in three three different groups. The first class consists of axioms that are linked to the classic monotonicity property in social choice theory. In the current context, the basic axiom studies the impact on the consensus measure if one individual changes her/his preferences relation by reversing two alternatives x and y that are *contiguous* to each other in the sense that if xP_iy, then there does not exist a third alternative z such that zP_ix and yP_iz. If, after reversing the preference between x and y, it is still the case that a majority of the individuals prefers the alternative that is made worse off, then the consensus should decrease. In concordance with the social choice literature, this property will be called Monotonicity. In course of our study, we also will consider variations of this central idea.

The second group of axioms refers to independence properties. that each of two changes of contiguous pairs of alternatives with some common characteristics should have the same impact on the change in the consensus. Some versions of this idea are reflected in the axioms of this group.

The third group of properties includes axioms refered to the relation between the consensus measures for different sets of alternatives and individuals. For example, the Full Range property requires that there are at least two preference profiles with consensus 0 and 1, respectively.

With these axioms, we present some characterization results, that include characterizations of *normalized average tau*, *average sigma* and *normalized average sigma*.

Remainder We proceed as follows: in the next section, we introduce basic notation and definitions. Section 3 is devoted to present some results about the relations between the normalized measures. The formal axioms and the characterizations are presented afterwards. Finally, we conclude. The proofs of the results are not incorporated to this reduced version of the paper, but are available upon request.

2 NOTATION AND DEFINITIONS

Consider an infinite set of alternatives X and an infinite set of individuals $\mathcal{N}$. Our objective is to evaluate the consensus in finite groups of individuals $N \subset \mathcal{N}$ and finite sets of alternatives $K \subset X$. Thus, we evaluate the consensus in *situations* (N, K). The cardinalities of N and K are denoted by n and k, respectively. We assume that $n \geq 2$ and $k \geq 2$. Elements of X are usually denoted by x, y and z, generic individuals are indexed by i and j.

Let P_i be the strict preference relation of individual i on $\mathcal{K}$. In particular, we assume that P_i is complete, transitive, and antisymmetric. Given an individual i with preferences P_i and a set of alternatives K, we denote by P_i^K the preference relation obtained by restricting P_i to K; that is, for all $x, y \in K$, $xP_i^K y$ if, and only if, xP_iy. Obviously, P_i^K is complete, transitive, and antisymmetric. The set of all preference relations on K is denoted by $\mathcal{P}_K$. A preference profile on K for society N, $P_N^K = (P_i)_{i\in N} \in (\mathcal{P}_K)^N$, is an n–tuple of individual preference relations on K. Given a set of individuals $N \subset \mathcal{N}$ and a set of alternatives $K \subset X$, a *consensus measure for situation* (N, K) is a function $M_{N,K} : (\mathcal{P}_K)^N \to [0, 1]$ that assigns to every preference profile $P_N^K \in (\mathcal{P}_K)^N$ a real number $M_{N,K}(P_N^K)$ from the unit interval. For notational purposes, we will write $M(P_N^K)$ instead of $M_{N,K}(P_N^K)$. A consensus measure is then a family of functions $M : \{(\mathcal{P}_K)^N \to [0, 1]\}_{K\subset X, N\subset\mathcal{N}}$ such that for all possible situations (N, K), $M_{N,K}$ is a consensus measure for that situation.

For the restricted case of two individuals ($N = \{i, j\}$), Kendall [4] proposed a consensus measure that has received particular attention. To define it, take any preference profile P_N^K belonging to the situation $(\{i, j\}, K)$. Then, for any pair of alternatives (x, y), $\tau_{x,y}^{i,j}(P_N^K) = 1$ if individual i and j agree on the binary ordering between x and y ($xP_i^K y \Leftrightarrow xP_j^K y$), otherwise $\tau_{x,y}^{i,j}(P_N^K) = 0$. The consensus in this society, $\tau_{i,j}(P_N^K)$, is then defined as the percentage of pair–wise comparisons that both individuals agree upon; that is, $\tau_{i,j}(P_N^K) = \frac{2}{k(k-1)} \sum_{x,y\in K: y\neq x} \tau_{x,y}^{i,j}(P_N^K)$. One proposal to study the general case of an arbitrary set of individuals is to naturally extend Kendall's τ by calculating the average of all $\tau_{i,j}(P_N^K)$ for all pairs of individuals in the society (see, Hays [3]).

Definition 1 The consensus measure M is called *average tau* ($\bar{\tau}$) if, and only if, for all situations (N, K) and all preference profiles $P_N^K \in (\mathcal{P}_K)^N$,

$$M(P_N^K) = \frac{2}{n(n-1)} \sum_{i,j\in N: j\neq i} \tau_{i,j}(P_N^K).$$

While this measure considers pair–wise comparisons on the individual level as the basic ingredient, an alternative approach consists in calculating difference in the support between pairs of alternatives according to the Majority Rule. Formally, given the preference profile $P_N^K \in (\mathcal{P}_K)^N$ and the pair of alternatives (x, y), let $\#(xP_N^K y)$ be the number of individuals that prefer x to y at P_N^K. Then, $n_{x,y}(P_N^K) = \#(xP_N^K y) - \#(yP_N^K x)$ denotes the difference between the number of individ-

uals that prefer x to y and the ones that prefer y to x, always at $P_N^K \in (\mathcal{P}_K)^N$.

In line with the construction of Kendall's tau, consider first the case of two alternatives ($K = \{x, y\}$). Then, given any preference profile P_N^K belonging to the situation $(N, \{x, y\})$, the consensus would be $\sigma_{x,y}(P_N^K) = \frac{|n_{x,y}(P_N^K)|}{n}$. This measure can also be naturally extended to the general case of an arbitrary number of alternatives by calculating the average of all $\sigma_{x,y}(P_N^K)$ for all pairs of alternatives. Formally, we obtain the following definition.

Definition 2 The consensus measure M is called *average sigma* ($\bar{\sigma}$) if, and only if, for all situations (N, K) and all preference profiles $P_N^K \in (\mathcal{P}_K)^N$,

$$M(P_N^K) = \frac{2}{k \cdot (k-1)} \sum_{x,y \in K: y \neq x} \sigma_{x,y}(P_N^K).$$

The measures presented so far are defined on the unit interval, but, nevertheless, the range of these measures turns out to be distinct for different situations. We also propose the normalized versions, re–defined on the unit interval. In particular, denote by $min\{M_{N,K}\}$ the minimal consensus M takes for a given situation (N, K). Define $max\{M_{N,K}\}$ accordingly.

Definition 3 Given the consensus measure M, the consensus measure $\tilde{M}$ is said to be *normalized with respect to* M if, and only if, for all situations (N, K) and all preference profiles $P \in (\mathcal{P}_K)^N$,

$$\tilde{M}(P_N^K) = \frac{M(P_N^K) - min\{M_{N,K}\}}{max\{M_{N,K}\} - min\{M_{N,K}\}}.$$

We denote the normalized average tau and the normalized average sigma by $\tilde{\tau}$ and $\tilde{\sigma}$, respectively.

3 CONSTRUCTION AND COMPARISON OF $\tilde{\sigma}$ AND $\tilde{\tau}$

Our first aim is to calculate the minimal and maximal values of $\bar{\tau}$ and $\bar{\sigma}$ in order to construct the normalized measures $\tilde{\tau}$ and $\tilde{\sigma}$. It is worth noticing that, for both measures, limits are independent of the set of alternatives K and only depend on the size of the society N.

Lemma 1 *Given a situation* (N, K),

$$min\{\bar{\tau}_{N,K}\} = \begin{cases} \frac{n-2}{2(n-1)} & \text{if } n \text{ is even} \\ \frac{n-1}{2n} & \text{if } n \text{ is odd}, \end{cases}$$

$$min\{\bar{\sigma}_{N,K}\} = \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{1}{n} & \text{if } n \text{ is odd}, \end{cases}$$

and

$$max\{\bar{\tau}_{N,K}\} = max\{\bar{\sigma}_{N,K}\} = 1.$$

With this result at hand, we are able to compare the two normalized measures for different situations. We find first that the two measures are identical to each other whenever $n \leq 3$.

Proposition 1 *For any situation* (N, K) *such that* $n \leq 3$ *and for all preference profiles* $P_N^K \in (\mathcal{P}_K)^N$, $\tilde{\tau}(P_N^K) = \tilde{\sigma}(P_N^K)$.

Limiting the number of alternatives does not help to find an exact equivalence between the two consensus measures. However, if $k = 2$, then it is the case that the two measures are ordinally equivalent meaning that if the consensus at a certain preference profile is higher than the consensus at another profile according to one measure, then the same is true also for the other measure. Formally, two consensus measures M and M' are said to be *ordinally equivalent for situation* (N, K) whenever for all preference profiles $P_N^K, \bar{P}_N^K \in (\mathcal{P}_K)^N$ such that $M(P_N^K) \geq M(\bar{P}_N^K)$ it is also the case that $M'(P_N^K) \geq M'(\bar{P}_N^K)$

Proposition 2 *For any situation* (N, K) *such that* $k = 2$, $\tilde{\sigma}$ *and* $\tilde{\tau}$ *are ordinally equivalent.*

Our final result in this section establishes that the measures are totally different from each other as soon as $n \geq 4$ and $k \geq 3$.

Proposition 3 *For any situation* (N, K) *such that* $n \geq 4$ *and* $k \geq 3$, *there are two preference profiles* $P_N^K, \bar{P}_N^K \in (\mathcal{P}_K)^N$ *such that* $\tilde{\sigma}(P_N^K) \leq \tilde{\sigma}(\bar{P}_N^K)$ *and* $\tilde{\tau}(P_N^K) > \tilde{\tau}(\bar{P}_N^K)$.

4 PROPERTIES AND CHARACTERIZATIONS

In this section, we introduce the axioms and present our main results. Let us start with some notation. We will say that the pair of alternatives (x, y) is a *contiguous pair for individual* $i \in N$ *at* P_i^K if xP_iy and there does not exist an alternative $z \in K$ such that xP_izP_iy. Then, the preference relation $\bar{P}_i^K$ is (x, y)*–different with respect to* the relation P_i^K, where $x, y \in K$, if (y, x) is a contiguous pair for i at P_i, (x, y) is a contiguous pair for i at $\bar{P}_i$, and $zP_i^K w \Leftrightarrow z\bar{P}_i^K w$ for all $(z, w) \neq (x, y)$. Finally, in situation (N, K) the preference profile $\bar{P}_N^K$ is said to be (x, y)–different with respect to P_N^K whenever $\bar{P}_i^K$ is (x, y)–different from P_i^K for some individual $i \in N$ and $\bar{P}_j^K = P_j^K$ for all $j \neq i$. Intuitively, two preference profiles are (x, y)–different from each other if one preference profile can

be derived from the other by interchanging for one individual one contiguous pair of alternatives.

4.1 MONOTONICITY PROPERTIES

The first group of axioms states the consequences for a consensus measure of a change of a contiguous pair in an individual preference.

WEAK MONOTONICITY: The consensus measure M is *weakly monotone* if for all situations (N, K) and all preference profiles P_N^K and $\bar{P}_N^K$ such that $\bar{P}_N^K$ is (x, y)–different from P_N^K,

$$n_{y,x}(P_N^K) = 1 \Rightarrow M(\hat{P}_N^K) = M(P_N^K).$$
$$n_{x,y}(P_N^K) \geq 0 \Rightarrow M(\hat{P}_N^K) > M(P_N^K).$$

Weak Monotonicity considers the situation when one single individual interchanges a contiguous pair of alternatives. If, in this situation, it is the case that the number of individuals that prefer one alternative over the other is interchanged as well, then the consensus should not vary across the two profiles. If, however, there were more individuals that prefer the alternative that is made better off with the change, then the consensus should increase. In the following, we strengthen this property.

MONOTONICITY: The consensus measure M is *monotone* if it is weakly monotone and for all situations (N, K) and all preference profiles $P_N^K, \hat{P}_N^K, \bar{P}_N^K \in (\mathcal{P}_K)^N$ such that $\hat{P}_N^K$ is (x, y)–different from P_N^K, $\bar{P}_N^K$ is (w, z)–different from P_N^K, and $0 \leq n_{x,y}(P_N^K) < n_{z,w}(P_N^K)$,

$$M(\bar{P}_N^K) > M(\hat{P}_N^K) > M(P_N^K).$$

This version is stronger, because it additionally establishes that the increase in the consensus should be larger if the number of individuals that preferred the alternative that is made better off is also larger. The last monotonicity is even stronger, because it establishes that the increase in the consensus changes proportionally with the number of individuals that prefer the alternative that is made better off.

STRONG MONOTONICITY: The consensus measure M is *monotone* if it is weakly monotone and for all situations (N, K) and all preference profiles $P_N^K, \hat{P}_N^K, \bar{P}_N^K \in (\mathcal{P}_K)^N$ such that $\hat{P}_N^K$ is (x, y)–different from P_N^K, $\bar{P}_N^K$ is (w, z)–different from P_N^K, and $0 \leq n_{x,y}(P_N^K) < n_{z,w}(P_N^K)$,

$$\frac{M(\bar{P}_N^K)}{M(\hat{P}_N^K)} = \frac{\#(zPw)}{\#(xPy)}.$$

4.2 INDEPENDENCE PROPERTIES

This subsection presents three independence properties. To clarify the notation, we should refer to the properties as Independence-1, Independence 2 and Strong Independence to reflect that the last one is stronger than the first two, but there is no logical relation between the first pair.

The first property, Independence-1, states that two changes of contiguous pairs varies the consensus in the same quantity when the number of individuals that prefer the alternative that is made better off with each change is the same.

INDEPENDENCE-1: The consensus measure M is *independent-1* if for all situations (N, K) and all preference profiles $P_N^K, \bar{P}_N^K, \hat{P}_N^K, \tilde{P}_N^K \in (\mathcal{P}_K)^N$ such that $\bar{P}_N^K$ is (x, y)–different from P_N^K, $\tilde{P}_N^K$ is (z, w)–different from $\hat{P}_N^K$, and $n_{x,y}(P_N^K) = n_{z,w}(P_N^K)$,

$$M(\bar{P}_N^K) - M(P_N^K) = M(\tilde{P}_N^K) - M(\hat{P}_N^K).$$

The second property reflects that the difference in the increases of two consecutive changes of contiguous pairs is independent of the profile where the changes took place.

INDEPENDENCE-2: The consensus measure M is *independent-2* if for all situations (N, K) and all preference profiles $P_N^K, \bar{P}_N^K, \check{P}_N^K, \hat{P}_N^K, \tilde{P}_N^K, \breve{P}_N^K$ such that $\bar{P}_N^K$ is (x, y)-different from P_N^K, $\tilde{P}_N^K$ is (z, w)–different from $\hat{P}_N^K$, $\check{P}_N^K$ is (x, y)–different from $\bar{P}_k^K$, $\breve{P}_N^K$ is (z, w)–different from $\tilde{P}_N^K$, and both $n_{x,y}(P_N^K)$ and $n_{z,w}(P_N^K)$ are strictly positive,

$$[M(P_N^K) - M(\bar{P}_N^K)] - [M(\bar{P}_N^K) - M(\check{P}_N^K)] =$$
$$[M(\hat{P}_N^K) - M(\tilde{P}_N^K)] - [M(\tilde{P}_N^K) - M(\breve{P}_N^K)].$$

The last property simply states that two changes of contiguous pairs of alternatives makes the consensus change equally.

STRONG INDEPENDENCE: The consensus measure M is *strongly independent* if for all situation (N, K) and all preference profiles $P_N^K, \bar{P}_N^K, \hat{P}_N^K, \tilde{P}_N^K \in (\mathcal{P}_K)^N$ such that $\bar{P}_N^K$ is (x, y)–different from P_N^K, $\tilde{P}_N^K$ is (z, w)–different from $\hat{P}_N^K$, and both $n_{x,y}(P_N^K)$ and $n_{z,w}(P_N^K)$ are non–negative,

$$M(\bar{P}_N^K) - M(P_N^K) = M(\tilde{P}_N^K) - M(\hat{P}_N^K).$$

4.3 PROPERTIES ACROSS DIFFERENT SITUATIONS

In this section, we will introduce four properties that reflect some relations of the values of a consensus measure across different sets of individuals and alternatives.

The first property, Neutrality, is a classical property in social choice. In this framework, it states that the name of the alternatives do not matter in the calculus of the consensus.

NEUTRALITY: The consensus measure M is *neutral* if for all situations (N, K) and (N, K') such that $k = k'$, and all preferences profiles $P_N^K \in (\mathcal{P}_K)^N$ and $\hat{P}_N^{K'} \in (\mathcal{P}_{K'})^N$ such that there exists a bijection $\mu : K \to K'$ with the property that $xP_i^K y \Leftrightarrow \mu(x)\hat{P}_i^{K'}\mu(y)$ for all $i \in N$ and all $x, y \in K$,

$$M(P_N^K) = M(\hat{P}_N^{K'}).$$

The second property is the Dalton principle, a very important property in the literature of income inequality measurement. Adopted to our framework it simply states that the consensus is invariant to replications of the society.

DALTON PRINCIPLE: The consensus measure M satisfies the *Dalton principle* if for all situations (N, K) and (N', K) with the properties that $N \subset N'$, $n' = m \cdot n$, with m being an integer, and all preferences profiles $P_N^K \in (\mathcal{P}_K)^N$ and $\hat{P}_{N'}^K \in (\mathcal{P}_K)^{N'}$ such that there exists a partition of N', $\{N_1, \ldots, N_m\}$ that satisfies $P_N^K = \hat{P}_{N_j}^K$ for all $j \in \{1, \ldots, m\}$,

$$M(P_N^K) = M(\hat{P}_{N'}^K).$$

The next property states that a consensus measure should take the theoretical minimum and maximum (0 and 1) in at least one profile each.

FULL RANGE: The consensus measure M satisfies *full range* if there exist situations (N, K) and (N', K') such that there are two preference profiles $P_N^K \in (\mathcal{P}_K)^N$ and $P_{N'}^{K'} \in (\mathcal{P}_{K'})^{N'}$ that satisfy $M(P_N^K) = 0$ and $M(P_{N'}^{K'}) = 1$.

The last property is even stronger since it requires that the theoretical extreme values are taken by some profiles in every situation.

STRONG FULL RANGE: The consensus measure M satisfies *strong full range* if for all situations (N, K), there exists two preference profiles $P_N^K, \hat{P}_N^K \in (\mathcal{P}_K)^N$ that satisfy $M(P_N^K) = 0$ and $M(\hat{P}_N^K) = 1$.

4.4 THE RESULTS

In this section, we present our main results. The first theorem explains the consequences of assuming the weakest version of monotonicity and one of the weakest versions of independence.

Theorem 1 *The following two statements are equivalent:*

- *The consensus measure satisfies Weak Monotonicity and Independence-1.*
- *For all $N \subset \mathcal{N}$ and all $K \subset X$, there exists a strictly increasing function $f_{N,K}$ such that for all $P_N^K \in (\mathcal{P}_K)^N$,*

$$M(P_N^K) = \sum_{(x,y)\in K: x\neq y} f_{N,K}\left(n_{x,y}(P_N^K)\right).$$

Using Theorem 1 as our basic result, it is easy to see which are the implications of other properties. The next two results show the implications of strength Weak Monotonicity (to Monotonicity) and Independence-1 (to Independence).

Theorem 2 *The following two statements are equivalent:*

- *The consensus measure satisfies Monotonicity and Independence-1.*
- *For all $N \subset \mathcal{N}$ and all $K \subset X$, there exists a strictly increasing and strictly convex function $f_{N,K}$ such that for all $P_N^K \in (\mathcal{P}_K)^N$,*

$$M(P_N^K) = \sum_{(x,y)\in K: x\neq y} f_{N,K}\left(n_{x,y}(P_N^K)\right).$$

Theorem 3 *The following two statements are equivalent:*

- *The consensus measure satisfies Weak monotonicity and Strong Independence.*
- *For all $N \subset \mathcal{N}$ and all $K \subset X$, there exists $a_{N,K}, b_{N,K} \in [0, 1)$ such that $b_{N,K} \leq \frac{2-2a_{N,K}}{nk(k-1)}$ and that for all $P_N^K \in (\mathcal{P}_K)^N$,*

$$M(P_N^K) = a_{N,K} + b_{N,K} \cdot \left(\sum_{(x,y)\in K: x\neq y} n_{x,y}(P_N^K) \right).$$

We also include the characterizations of some of the measures presented in the previous sections.

Theorem 4 *A consensus measure satisfies Weak Monotonicity, Strong Independence, Neutrality, Dalton principle and Full Range if and only if it is average sigma ($\bar{\sigma}$).*

Theorem 5 *A consensus measure satisfies Weak Monotonicity, Strong Independence and Strong Full Range if and only if it is normalized average sigma ($\tilde{\sigma}$).*

Theorem 6 *A consensus measure satisfies Strong Monotonicity, Independence-1, Independence-2 and Strong Full Range if and only if it is normalized average tau ($\tilde{\tau}$).*

In Theorem 6, if we weaken Strong Full Range and we add Neutrality and Dalton Principle, we do not obtain a characterization of *average tau*. The reason is that $\bar{\tau}$ does not satisfy the Dalton Principle. This is surprising, since as we have seen this change in the axioms in presence of Weak Monotonicity and Strong Independence leads to the change from a characterization of *normalized average sigma* to one of *average sigma* (see Theorems 4 and 5).

Definition 4 The consensus measure M is called *corrected average tau* ($\hat{\sigma}$) if, and only if, for all situations (N, K) and all preference profiles $P_N^K \in (\mathcal{P}_K)^N$,

$$M(P_N^K) = \begin{cases} \tilde{\tau} & \text{if } n \text{ is even} \\ \frac{1}{n^2} + \tilde{\tau} \cdot \frac{n^2-1}{n^2} & \text{if } n \text{ is odd} . \end{cases}$$

Theorem 7 *A consensus measure satisfies Strong Monotonicity, Independence-1, Independence-2, Neutrality, Dalton Principle and Full Range if and only if it is corrected average tau* ($\hat{\tau}$).

We can also show that the properties used in the characterization results are independent.

Proposition 4 *The axioms used in the characterization results are independent.*

5 CONCLUSIONS

We have presented some possible ways to measure the consensus and we have some characterizations of these measures. The measures in which we have focused in this paper take as the basic ingredient the binary comparisons between alternatives, as in the Condorcet tradition of social choice. As a consequence, our measures have some relation with the Kemeny [5] distance between two preference orderings. Other possibilities should be to start with other information of the individual orderings different from the pairwise comparisons. For example, if we construct a consensus measure taking the information of the positional comparisons as in the Borda tradition of social choice, it is probably that the measures at which we arrive should have some relation with the distance function proposed by Cook and Seiford [1, 2]. If, however, we take as the basic ingredient of comparison between preference orderings the different choice functions that they generate, the consensus measures that we construct should be in relation with the distance function proposed by Klamler [6].

Acknowledgements

Preliminar version. We are very grateful to the seminar audiences in Bilbao, Soria, and Vigo. Jorge Alcalde-Unzu acknowledges the financial support by the Spanish Ministry of Education and Science, through the Juan de la Cierva program and project SEJ2006-11510. Marc Vorsatz acknowledges the financial support through the Ramón y Cajal program of the Spanish Ministry of Education and Science. This work was initiated while the authors worked at Universidad Pública de Navarra and Maastricht University, respectively.

References

[1] W. Cook y L. Seiford. Priority Ranking and Consensus Formation. *Management Science* 24, Pág. 1721-1732, 1978.

[2] W. Cook y L. Seiford. On the Borda-Kendall Consensus Method for Priority Ranking Problems. *Management Science* 28, Pág. 621-637, 1982.

[3] W. Hays. A Note on Average Tau as a Measure of Concordance. *Journal of the American Statistical Association* 55, Pág. 331-341, 1960.

[4] M. Kendall. Rank Correlation Methods, 3er edicion, Hafner Publishing Company, New York, 1962.

[5] J. Kemeny. Mathematics Without Numbers, *Daedalus* 88, Pág. 577-591, 1959.

[6] C. Klamler. A distance measure for choice functions. *Social Choice and Welfare* 30, Pág. 419-425, 2008.

APPROACHES TO CONSENSUS: THE STATE OF THE ART

Miguel Martínez-Panero

Dep. de Economía Aplicada, PRESAD Research Group
Universidad de Valladolid, Valladolid, 47011, Spain
panero@eco.uva.es

Abstract

This paper surveys the notion of consensus from several points of view, ranging from philosophical aspects and characterizations of several quantification measures within the Social Choice framework, paying also attention to aspects of judgment aggregation as well as fuzzy or linguistic approaches, to practical applications in Decision Making and Biomathematics, among others.

Keywords: Consensus, consent, agreement.

1 INTRODUCTION

> **Consensus** [L., fr. consensus, pp. of *consentire*] **1**: group solidarity in sentiment and belief. **2 a**: general agreement, unanimity. **2 b**: the judgement arrived at by most of those concerned.
>
> **Consent 1**: compliance in or approval of what is done or proposed by another, acquiescence. **2**: agreement as to action or opinion; *specif*: voluntary agreement by a people to organize a civil society and give authority to the government.
>
> *Webster's New Collegiate Dictionary*

Consensus is a multi-facet concept. Emerson [23] refers to "general agreement" versus "majority view" considered as two widespread opposite senses of the word[1]. On his part, Emerson [22] asserts that "the general consensus on any one issue is that which is perceived to be the agreed opinion of an overwhelming number, though there may be some who dissent". Thus, this author points out an evident fact, i.e., that consensus is not endowed with an absolute meaning, but a relative one, and nowadays it is usual to deal with levels or measures of consensus, as several of the authors appearing next do.

This paper, in its second section, presents some philosophical aspects of consensus essentially focused on the doctrine that men are joined together within a society by a contract with explicit or hidden agreements, as Rousseau believed. Then we outline some further developments and connections, such as the link between Rousseau and Condorcet through his well-known "jury theorem". We also distinguish between the notion of consent and the more technical and recent concept of consensus as appearing in modern Political Science and Sociology. In the third section, we deal with several formal approaches to consensus mainly from the Social Choice framework (some of them being in fact works in progress), and recent advances from distance-based, fuzzy or linguistic points of view are presented. In addition, we point out some aspects of an emergent research field focused on "judgment aggregation". The fourth section includes some examples as signs of the power of consensus-based methods in practice, and we refer to the idea of penalizing disagreement which is the basis of a new voting system tested in 2007 (the "Majority Judgement"), to consensus as a determining condition for publishing in Wikipedia, and about how this idea naturally[2] appears in Molecular Biology and Biomathematics. Finally, we expose some concluding remarks.

[1] It is meaningful that Emerson [23] entitled his book after Tocqueville argument against the "majority tyranny" appearing in *Democracy in America*.

[2] Taking into account the common structure or pattern between what they call "bioconsensus" and Social Choice, Day and McMorris [18] advocate for the existence of an Arrow's paradigm in the Kuhn's sense.

2 PHILOSOPHICAL BACKGROUND

There is a long tradition, from Greek and Roman thinkers[3] to Locke, Hobbes, Paine and mainly Rousseau [66], supporting the idea that political authority relies on the consent of the governed through a *social contract*. Thus, the legitimacy of the government should be based on the *general will*[4]. It is interesting to point out how this concept is related to the idea of consensus as appearing in Condorcet's [14] "jury theorem". This fact has been explained by Grofman and Feld [32], who point out the following three elements of Rousseau theory then formalized by Condorcet:

1. There is a common good.
2. Citizens are not always accurate in their judgements about what is in the common good.
3. When citizens strive to identify this common good and vote in accordance with their perceptions of it, the vote of the Assembly of the People can be taken to be the most reliable means for ascertaining the common good.

An overview of other recent authors such as Habermas, Rawls, etc. about the subject of consensus-based legitimacy can be found in Knight and Johnson [52]. And it is worth to mention that Buchanan and Tullock consider their well-known work, *The Calculus of Consent*, as a contractual theory of the State (see Buchanan and Tullock [11], appendix 1).

We have used both the terms "consent" and "consensus". According to Partridge ([63], p. 71), "within modern Social Theory *consent* has been mainly a term of Political Philosophy, *consensus* mainly a term in Sociology". Indeed, Jiménez de Parga [42] points out that it was Comte who introduced in the XIXth century in the vocabulary of Social Sciences the notion of "consensus", in its Latin version, taken from the writings by Caesar, Tully and Livy. Nevertheless, according to the referenced authors, it is well known that Comte founded Sociology from the basis of Biology, and in this particular case the coined term was inspired by Medicine of his age, where consensus meant solidarity among distinct parts of the human body. Thus, in recent years, consensus refers to agreements to provide the conditions of political and social cohesion (see Lipsitz [53]).

[3] "The doctrine that society is itself a voluntary association was not unheard of in Greek and Roman times" (see Partridge [63], p. 15).

[4] See Negro Pavón [61], Saccamano [69] and Graham Jr. [31].

3 FORMAL APPROACHES TO CONSENSUS

In what follows we present several aspects of consensus and how they have been developed within different frameworks.

3.1 Consensus and Social Choice

In Theory of Democracy as understood by van Mill [57] there are two extremely different currents to deal with conflicts in a legitimate way: "theories of democratic discourse" advocated mainly by Habermas [33], which try to solve them through deliberation and consensus, and "disequilibrium theories of social choice" as proposed by Black [7], Arrow [2] and Riker [65], aggregating individual preferences into a social result (see also Britos [10], who slightly modify this approach speaking about the "argumentative discourse" versus "social choice", and tries to find some ways of conciliation).

However, Social Choice is nor merely a tool to decide in absence of consensus, and some authors have formalized this notion in the Social Choice context. In this way, Samet and Schmeidler [70] have introduced and then axiomatized "consent rules, that incorporate aspects of majoritarianism and liberalism". On their part, Bosch [9] and Alcalde and Vorsatz [1] have proposed several consensus measures and they have also characterized them. On the other hand, García-Lapresta [27] and García-Lapresta and Pérez-Román [28] have devised voting methods penalizing the disagreement, the key idea being the "marginal contribution to consensus" for each agent, Mata, Martínez and Martínez [56] have analyzed strategic aspects in consensus processes, and Ballester and García-Lapresta [6] have analyzed a recursive consensus process for selecting qualified individuals of a group.

It is also worth to mention consensus models to decision making in committees which has nothing to do with voting or coalitions, for example that appearing in Eklund, Rusinowska and De Swart [21]. These authors have performed a method where a chairman suggests the decision makers to adjust or modify their evaluation of some alternatives attending several criteria in order to improve the agreement, following an idea which recurrently appears in Consensus Theory. In this way, Carlsson et al. [13] have also considered "an advising monitor which tries to contract the decision makers into a mutual decision through soft enforcement", where convergence to consensus is ana-

lyzed by means of a "topological approach" which can measure distances between decision makers and allow these authors to model the trade-off between "degree of consensus" and "strength of majority".

Such metric techniques lead us to the next section, where the notion of distance becomes crucial.

3.2 Distance-based approaches to consensus

The idea of considering distances among preferences (already suggested in Condorcet's writings) has revealed to be very fruitful[5]. In this way, consensus can be somehow understood as closeness to unanimity, and several attempts has been done to minimize (from distance-based assumptions) the aggregate disagreement.

Hornik and Meyer [43] considers that such "optimization approach" goes back to Règnier [64]. Nonetheless, it is fair to acknowledge the pioneer analysis of so-called "compromise solutions" (those minimizing the group regret through appropriate distance functions) by Yu [73] and Freimer and Yu [26]. On the other hand, in order to determine the closest ranking to all decision-makers, Cook and Seiford [15, 16] considered the so-called metrics 1 and 2, the last one related to the Borda rule[6] independently obtained by Kendall [49] in a statistical estimation context (see González-Pachón and Romero [29] for a detailed exposition and further developments).

The last mentioned authors, González-Pachón and Romero [29], have considered the aforementioned problem of minimization as a Goal Programming model and have recently obtained consensus weights when the information taken into account by the decision maker is somehow inconsistent, i.e., does not verify some rational assumptions (see González-Pachón and Romero [30]).

3.3 Judgement Aggregation

According to List [54], Judgement Aggregation is concerned in "how can a group of individuals aggregate the group members' individual judgments on some interconnected propositions into corresponding collective judgments on these propositions", and such analysis may be done from different collective decision making contexts. This research field runs parallel to Social Choice. In fact there is a "doctrinal paradox" (or "discursive dilemma") whose homologue in Social Choice would be the "Condorcet's paradox", and there is also an Imposibility Theorem and further developments to mitigate the resulting impass, as happening with Arrow's Theorem and its variants (see List [54] and the references appearing there, all of them available online, as well as Dietrich [19], and Dietrich and List [20]).

Within this framework, Hartmann and Sprenger [35] and Hartmann, Pigozzi and Sprenger [34] are developing models of consensus appearance in judgement aggregation situations in order to achieve compromise-based decisions.

3.4 Fuzzy/Linguistic consensus models

An updated overview on the use of fuzzy preferences and majorities in obtaining group decision making outcomes and levels of consensus has been recently presented by Kacprzyk, Zadrozny, Fedrizzi and Nurmi (see [47] and the references appearing there). As for fuzzy and linguistic approaches to consensus models for group decision making it is worth to mention Herrera-Viedma, Alonso, Chiclana and Herrera [38], Herrera-Viedma, Herrera and Chiclana [39], Herrera, Herrera-Viedma and Verdegay [36, 37], Herrera-Viedma, Martínez, Alonso [3], Mata [55], Mata and Chiclana [40], and Cabrerizo, Alonso, Pérez and Herrera-Viedma [12], among others. In these papers it is usual the existence of a moderator in the process of obtaining a high level of consensus or agreement among the experts to decide in a group decision making problem (see also Fedrizzi, Kacprzyk and Zadrozny [25], Kacprzyk and Fedrizzi [44], Kacprzyk, Fedrizzi and Nurmi [45] and Kacprzyk, Nurmi and Fedrizzi [46]).

It is important to note here that Montero [59, 60] analized the impact of fuzziness in Group Decision Making and also proposed an escape from Arrow's imposibility result to reach a consensus-based solution respecting democratic compelling properties. Even more, when such imposibility result is unavoidable in some ratio-

[5] This is, for example, the key for the Kemeny [48] rule, inspired by the treatment of statistical data. In fact, as demonstrated by Young [71], this method is just the "maximum likelihood rule" tacitly formulated by Condorcet [14]. See Klamler [50, 51] and Nurmi [62] for Kemeny-type and other distance functions on the set of choice functions, García-Lapresta and Pérez-Román [28] and the references cited there concerning other voting methods using distances for obtaining a collective outcome, and Meskanen and Nurmi [58] who present a comprehensive approach to methods which can be viewed as distance minimizing ones, focusing on the measurement of disagreement.

[6] The Borda [8] and Condorcet [14] approaches have been called "*constructive* consensus methods" by Hornik and Meyer [43], but they assert that both rules agree in absence of cycles in the collective outcome, which is not true. Emerson [22, 23] advocates for the Borda rule, which promotes consensus and avoids the "tyranny of the majority". This is the reason why he has devised a modified Borda count system of consensus voting (see Emerson [24]). It is interesting to point out that Emerson [23] also defines a "level of consensus", but specifically intended for his system.

nal amalgamation operations, Cutello and Montero [17] have searched for (and then characterized) "good enough" solutions.

4 SOME EXAMPLES OF DEALING WITH CONSENSUS

This section glimpses three practical cases where the idea of consensus is taken into account for different purposes in distinct contexts.

4.1 Penalizing disagreement through Majority Judgement

According to Balinski and Laraki [4], Galton's only comment related to Social Choice Theory "was discarded as a small contribution by Black [7]". These are Galton's words (cited in Black [7]), which support the idea of penalizing the disagreement[7]:

> Each voter [...] has equal authority with each of his colleagues. How can the right conclusion be reached, considering that there may be as many different estimates as there are members? That conclusion is clearly *not* the *average* of all the estimates, which would give a voting power to cranks in proportion to their crankiness. One absurdly large or small estimate would leave a greater impress on the result than one of reasonable amount, and the more an estimate diverges from the bulk of the rest, the more influence would it exert. I wish to point out that the estimate to which least objection can be raised is the *middlemost* estimate, the number of votes that it is too high being exactly balanced by the number of votes that it is too low. Every other estimate is condemned by a majority of voters as being either too high or too low, the middlemost alone escaping this condemnation.

This argument has also been advocated by Balinski and Laraki, the idea of considering the median (middlemost) of the voters' assessments being the key of a new voting system proposal: the *Majority Judgement*, which has been empirically tested at Orsay during the 2007 presidential elections (see Balinski and Laraki [4, 5] for details both of the theoretical background and concerning the aforementioned experiment).

[7] As mentioned, a notion of penalization is also taken into account by García-Lapresta [27] and García-Lapresta and Pérez-Román [28].

4.2 Consensus for publishing in Wikipedia

Consensus plays a main role in the publishing policy of Wikipedia (see Figure 1, where the process is shown through a flowchart), and might entail a rejected entry if agreement for acceptance is not present after a reasonable time period. According to Wikipedia rules "consensus need not be fully opposed; if consensus is neutral or unclear on the issue and unlikely to improve, the proposal is likewise rejected" (see [72] for more details). Thus, Wikipedia seems to respect the ancient criterion of truth[8] known as *consensus gentium* (L.: Agreement of the peoples), stating that "which is universal among men carries the weight of truth" (cited in Runes [67]).

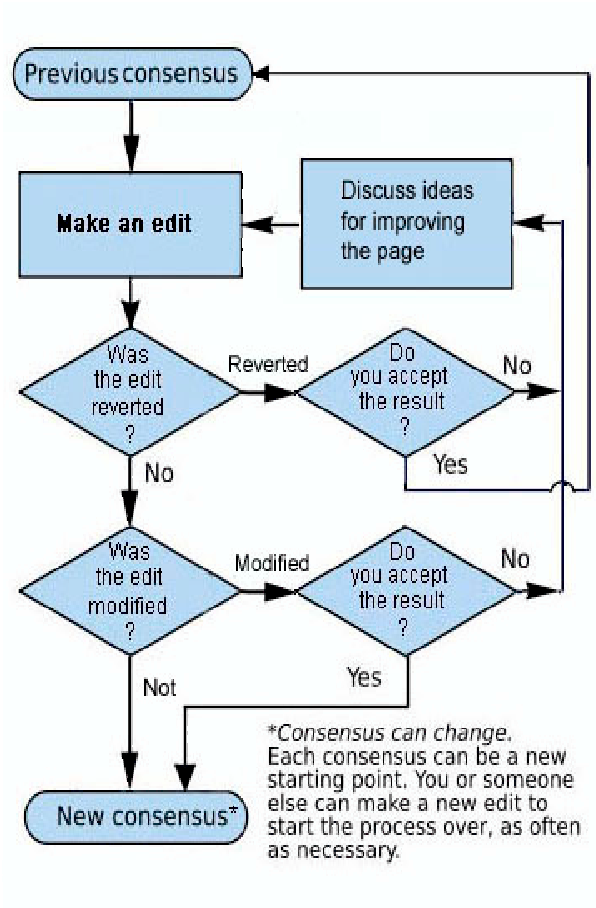


Figure 1: Wikipedia publishing consensus process. Source: Wikipedia

4.3 Consensus in Biomathematics

According to Day and McMorris [18], who are the authors of a handbook relating Consensus Theory to

[8] In fact, there are theories where the concept of truth crucially relies on consensus as an essential component.

Biomathematics[9], the archetypal problem of aggregation concerning group choice pioneered by Arrow [2] and then refined and extended by many other authors, can be applied in areas far from their origins in Group Choice and Welfare Economics. As pointed out by Janowitz in the foreword to the mentioned book:

> Since consensus theory has its roots in the theory of elections, many contributions have been (and are being) made by political scientists, sociologists and economists. In the context of human behavior, consensus research is carried out by psychologists. Facing with conflicting evidence on evolutionary history, systematic biologists appeal to concepts of consensus, and molecular biologists attempt to apply consensus theory in areas of DNA research. Market researchers find the discipline relevant since it can viewed as a theory of how large organizations make decisions based on possibly conflicting lines of evidence. Contemporary applications of consensus theory involve airplane and missile navigational systems, methods to combat bioterrorism, drug development based on DNA research, marketing and manufacturing decisions of large companies, and stock market predictions".

Particularly, Biomathematics and Molecular Biology are concerned in consensus ("bioconsensus") in areas where the objects of interest represent partitions of a set, hierarchical structures, phylogenetic trees and alignment of molecular sequences (for deeper explanations see Day and McMorris [18], especially chapter 1, and Janowitz, Lapointe, McMorris, Mirkin and Roberts [41]).

5 CONCLUSION

Along this paper, in a non-exhaustive way, we have outlined the idea of consensus from classic authors to nowadays researchers, from theoretical philosophical points of view to practical applications. Progress have been achieved, but much remains to be done, provided that, as mentioned, part of the work is currently in progress.

Acknowledgements

This paper has been partially supported by Spanish Ministerio de Educación y Ciencia (Project SEJ2006-04267/ECON) and ERDF. The author wish to thank José Luis García-Lapresta for his encouragement and suggestions. Helpful comments have also been provided by Christian Klamler, Stefano Vannucci and two anonymous referees.

[9] This book has been interestingly reviewed by Saari [68] mainly attending its Social Choice background.

References

[1] J. Alcalde and M. Vorsatz. The measurement of consensus in Social Choice, *ESTYLF 2008*.

[2] K. Arrow. *Social Choice and Individual Values*, Cowles Foundation Monographs, No. 12, Wiley, 1951. Revised second edition, 1963. Reprinted by Yale University Press, 1978.

[3] S. Alonso. *Group Decision Making with Incomplete Fuzzy Preference Relations*, Ph. D. Thesis, Universidad de Granada, 2006.

[4] M. Balinski and R. Laraki. A theory of measuring, electing and ranking, *Ecole Polytechnique – Centre National de la Rechereche Scientifique*. Cahier 2006-11, 2006.

[5] M. Balinski and R. Laraki. Election by Majority Judgement: Experimental evidence, *Ecole Polytechnique – Centre National de la Rechereche Scientifique*. Cahier 2007-28, 2007.

[6] M.A. Ballester and J.L. García-Lapresta. Sequential consensus for selecting qualified individuals of a group, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, forthcoming.

[7] D. Black. *The Theory of Committees and Elections*, Cambridge University Press, 1958.

[8] J.C de Borda. Mémoire sus les élections au scrutin, *Historie de l'Academie Royale des Sciences*, 1784.

[9] R. Bosch. *Characterizations of Voting Rules and Consensus Measures*, Ph. D. Dissertation, Tilburg University, 2005.

[10] P. Britos. Legitimidad democrática: ¿consenso o votación? http://www.bu.edu/wcp/Papers/Poli/PoliBrit.htm

[11] J. Buchahan and G. Tullock. *The Calculus of Consent. Logical Foundations of Constitutional Democracy*, University of Michigan Press, 1962.

[12] F.J. Cabrerizo, S. Alonso, L.J. Pérez, E. Herrera-Viedma. A consensus model for Group Decision Making in unbalanced fuzzy linguistic contexts, *ESTYLF 2008*.

[13] C. Carlsson, D. Ehrenberg, P. Eklund, M. Fedrizzi, P. Gustafsson, P. Lindholm, G. Merkuryeva, T. Riissanen and A.G.S. Ventre. Consensus in distributed soft environments, *European Journal of Operational Research* 61, pp. 165-85, 1992.

[14] Condorcet, Marquis de. Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la pluralité des voix, *L'Imrimierie Royale*, 1785.

[15] W.D. Cook and L.M. Seiford. Priority ranking and consensus formation, *Management Science* 24, pp. 1721-32, 1978.

[16] W.D. Cook and L.M. Seiford. The Borda-Kendall consensus method for priority ranking problems, *Management Science* 28, pp. 621-37, 1982.

[17] V. Cutello and J. Montero. A characterization of rational amalgamation operations, International Journal of Approximate Reasoning 8, pp. 325-44, 1993.

[18] W.H.E. Day and F.R. McMorris. *Axiomatic Consensus Theory in Group Choice and Biomathematics*, SIAM, 2003.

[19] F. Dietrich. A generalised model of judgment aggregation, *Social Choice and Welfare* 28, pp. 529-65, 2007.

[20] F. Dietrich and C. List. Arrow's theorem in Judgement Aggregation, *Social Choice and Welfare* 29, pp. 19-33, 2007.

[21] P. Eklund, A. Rusinowska and H. De Swart. Consensus reaching in committees, *European Journal of Operational Research* 178, pp. 185-93, 2007.

[22] P. Emerson. *Consensus Voting Systems*, Samizdat, 1991.

[23] P. Emerson. *Beyond the Tyranny of the Majority. Voting Methodologies in Decision-making and Electoral Systems*, The *de Borda* Institute, 1998.

[24] P. Emerson. Consensus voting and conflict resolution, http://www.opendemocracy.net/deborda/articles/consensus_voting_and_conflict_resolution

[25] M. Fedrizzi, J. Kacprzyk and S. Zadrozny. An interactive multi-user decision support system for consensus reaching processes using Fuzzy Logic with Linguistic Quantifiers. *Decision Support Systems* 4, pp. 313-27, 1988.

[26] M. Freimer and P.L. Yu. Some new results on compromise solutions for group decision problems. Management Science, Vol. 22, No. 6, pp. 688-93, 1976.

[27] J.L. García-Lapresta. Favoring consensus and penalizing disagreement in Group Decision Making, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, forthcoming.

[28] J.L. García-Lapresta and D. Pérez-Román. Some measures of consensus generated by distances on weak orders, *ESTYLF 2008*.

[29] J. González-Pachón and C. Romero. Distance-based consensus methods: A Goal Programming approach, *Omega, The International Journal of Management Science* 27, pp. 341-47, 1999.

[30] J. González-Pachón and C. Romero. Inferring consensus weights from pairwise comparison matrices without suitable properties, *Annals of Operations Research* 154, pp. 123-32, 2007.

[31] G.J. Graham Jr. Rousseau's concept of consensus, *Political Science Quarterly*, Vol. 85, No. 1, pp. 80-98, 1970.

[32] B. Grofman and S.L. Feld. Rousseau's general will: A Condorcetian perspective, *American Political Science Review* Vol. 82, No. 2, pp. 567-76, 1988.

[33] J. Habermas. *The Theory of the Communicative Action*, Beacon Press, Vol. One, 1984; Vol. Two, 1987.

[34] S. Hartmann, G. Pigozzi and J. Sprenger. Reliable methods of Judgement Aggregation, http://philsci-archive.pitt.edu/archive/00003593/01/RMJA.pdf

[35] S. Hartmann and J. Sprenger. Consensus, compromise and Judgement Aggregation, http://www.accessecon.com/pubs/SCW2008/SCW2008-08-00296S.pdf

[36] F. Herrera, E. Herrera-Viedma, J.L. Verdegay. A model of consensus in group decision making under linguistic assessments, *Fuzzy Sets and Systems* 78 , pp. 73-87, 1996.

[37] F. Herrera, E. Herrera-Viedma, J.L. Verdegay. A rational consensus model in Group Decision Making using linguistic assessments, *Fuzzy Sets and Systems* 88, pp. 31-49, 1997.

[38] E. Herrera-Viedma, S. Alonso, F. Chiclana, F. Herrera. A consensus model for Group Decision Making with incomplete fuzzy preference relations, *IEEE Transactions on Systems Fuzzy Systems* Vol 15, No. 5, pp. 863-77, 2005.

[39] E. Herrera-Viedma, F. Herrera, F. Chiclana. A consensus model for Multiperson Decision Making with different preference structures, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* Vol. 32, No.3, pp. 394-402, 2002.

[40] E. Herrera-Viedma, L. Martínez, F. Mata, F. Chiclana. A consensus support system model for Group Decision-making problems with multigranular linguistic preference relations, *IEEE Trans. on Fuzzy Systems* Vol. 13, No. 5, pp. 644-58, 2005.

[41] M. Janowitz, F.J. Lapointe, F.R. McMorris, B. Mirkin and F.S. Roberts (eds.). *Bioconsensus*, DIMACS Series in Mathematics and Theoretical Computer Science, American Mathematical Society, 2003.

[42] M. Jiménez de Parga. Palabras en el acto de presentación de *Ciudadanía y Valores*, http://www.funciva.org/descargas/acto_presentacion/discurso.doc

[43] K. Hornik and D. Meyer. Deriving consensus rankings from benchmarking experiments, *Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series* 33, 2006.

[44] J. Kacprzyk and M. Fedrizzi. A "soft" measure of consensus in the setting of partial (fuzzy) preferences, *European Journal of Operational Research* 34, pp. 316-25, 1988.

[45] J. Kacprzyk, M. Fedrizzi and H. Nurmi. Group decision making and consensus under fuzzy preferences and fuzzy majority, *European Journal of Operational Research* 49, pp. 21-31, 1992.

[46] J. Kacprzyk, H. Nurmi and M. Fedrizzi (eds.), *Consensus under Fuzziness*, Kluwer, 1996.

[47] J. Kacprzyk, S. Zadrozny, M. Fedrizzi and H. Nurmi. On group decision making, consensus reaching, voting and voting paradoxes under fuzzy preferences and a fuzzy majority: A survey and some perspectives, in H. Bustince et al. (eds.): *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, pp. 263-95, Springer, 2007.

[48] J. Kemeny. Mathematics without numbers, *Daedalus* 88, pp. 571-91, 1959.

[49] M. Kendall. *Rank Correlation Methods*, Hafner, 1962.

[50] C. Klamler. On some distance aspects in Social Choice, in B. Simeone and F. Pukelsheim (eds.): Mathematics and Democracy, pp. 97-104, Springer, 2006.

[51] C. Kalmler. A distance measure for choice functions. *Social Choice and Welfare* 30, pp. 419-25, 2008.

[52] J. Knight and J. Johnson. Aggregation and deliberation: On the possibility of Democratic legitimacy, *Political Theory*, Vol 22, No. 2, pp. 277-96, 1994.

[53] L. Lipsitz. Consensus, in D. Sills (dir.): *Enciclopedia Internacional de las Ciencias Sociales*, Aguilar, 1974.

[54] C. List. Judgment aggregation: A bibliography on the discursive dilemma, doctrinal paradox and decisions on multiple propositions, http://personal.lse.ac.uk/list/doctrinalparadox.htm

[55] F. Mata. *Modelos para Sistemas de Apoyo al Consenso en Problemas de Toma de Decisión en Grupo definidos en Contextos Lingüísticos Multigranulares*, Ph. D. Thesis, Universidad de Jaén, 2006.

[56] F. Mata, L. Martínez and J.C. Martínez. Penalizing manipulation strategies in Consensus processes, *ESTYLF 2008*.

[57] D. van Mill. The possibility of rational outcomes from democratic discourse and procedures, *The Journal of Politics*, Vol 58, No. 3, pp. 734-52, 1996.

[58] T. Meskanen and H. Nurmi. Distance from consensus: A theme and variations, in B. Simeone and F. Pukelsheim (eds.): Mathematics and Democracy, pp. 117-32, Springer, 2006.

[59] J. Montero. Arrow's theorem under fuzzy rationality, *Behavioral Science* 32, pp. 267-73, 1987.

[60] J. Montero. The impact of fuzziness in social choice paradoxes, *Soft Computing* 12, pp. 177-82, 2008.

[61] D. Negro Pavón. Rousseau y los orígenes de la política de consenso, http://www.cepc.es/rap/Publicaciones/Revistas/3/REPNE_008_063.pdf

[62] H. Nurmi. A comparison of some distance-based choice rules in ranking environments, *Theory and Decision* 57, pp. 5-24, 2004.

[63] P.H. Partridge. *Consent & Consensus*, Pall Mall, 1971.

[64] S. Règnier. Sur quelques aspects mathématiques des problèmes de classification automatique, *ICC Bulletin*, pp. 175-91, 1965.

[65] W.H. Riker. *Liberalism Against Populism. A Confrontation between the Theory of Democracy and the Theory of Social Choice*, Waveland Press, 1982.

[66] J.J. Rousseau. *Du Contrat Social ou Principes de Droit Politique*, Marc Michel, 1762.

[67] D.D. Runes (ed.), *Dictionary of Philosophy*, Littlefield, Adams and Company, p. 64, 1962.

[68] D.G. Saari. Review to [18], http://www.maa.org/reviews/axiomaticonsensus.html

[69] N. Saccamano. Rhetoric, consensus, and the law in Rousseau's *Contrat Social*, *MLN*, Vol. 107, No. 4, pp. 730-51, 1992.

[70] D. Samet and D. Schmeidler. Between liberalism and democracy. *Journal of Economic Theory* 110, pp. 213-33, 2003.

[71] H.P. Young. Optimal ranking and choice from pairwise comparisons, in B. Grofman and G. Owen (eds.): *Information Pooling and Group Decision Making*, JAI Press, pp. 113-22, 1986.

[72] Wikipedia publishing policy: Consensus, http://en.wikipedia.org/wiki/Wikipedia:Consensus

[73] P.L. Yu. A class of solutions for group decision making. *Management Science*, Vol. 19, No. 8, pp. 936-46, 1973.

SOME MEASURES OF CONSENSUS GENERATED BY DISTANCES ON WEAK ORDERS

José Luis García-Lapresta[1] **David Pérez-Román**[2]

[1] PRESAD Research Group, Universidad de Valladolid, lapresta@eco.uva.es
[2] PRESAD Research Group, Universidad de Valladolid, david@emp.uva.es

Abstract

In this paper we introduce some classes of consensus measures based on metrics on weak orders, and we analyze some of their properties. Taking into account these consensus measures, we have also introduced indices of contribution to consensus for each decision maker for prioritizing them in order of their contributions to consensus. We have also proposed a decision procedure where the individual opinions are weighted by the corresponding indices of contribution to consensus.

Keywords: Consensus, Metrics, Weak orders.

1 Introduction

When the members of a committee show their preferences over a set of alternatives, it is important to find not only a collective preference but to measure the agreement or disagreement between the members of the committee and the result (see Nurmi [21]). In this paper we focus our attention in how to measure consensus in groups of voters when they show their preferences over a fixed set of alternatives or candidates by means of weak orders (complete preorders).

We note that consensus has different meanings[1]. One of them is related to iterative procedures where voters must change their preferences to improve agreement. Usually, a moderator advise voters to modify some opinions (see, for instance, Eklund, Rusinowska and de Swart [10]). However, in this paper consensus is related to the degree of agreement in a committee, and voters do not need to change their preferences.

From a technical point of view, it is interesting to note that the problem of measuring the concordance or discordance between two linear orders has been widely explored in the literature. In this way, different *rank correlation indices* have been considered for assigning grades of agreement between two rankings (see Kendall and Gibbons [17]). Some of the most important indices in this context are Spearman's rho [24], Kendall's tau [16], and Gini's cograduation index [13]. On the other hand, some natural extensions of the above mentioned indices have been considered for measuring the concordance or discordance among more than two linear orders (see Hays [14]). For details and references, see for instance Borroni and Zenga [6].

In the field of Social Choice, Bosch [7] introduced the notion of *consensus measure* as a mapping that assigns a number between 0 and 1 to every profile of linear orders, satisfying three properties: (strong) *unanimity* (in every subgroup of voters, the highest degree of consensus is only reached whenever all individuals have the same ranking), *anonymity* (the degree of consensus is not affected by any permutation of voters) and *neutrality* (the degree of consensus is not affected by any permutation of alternatives).

Recently, Alcalde and Vorsatz [1] have introduced some consensus measures in the context of linear orders –related to some of the above mentioned rank correlation indices– and they provide some axiomatic characterizations.

In this paper we extend the Bosch's notion of consensus measure to the context of weak orders (indifference among different alternatives is allowed)[2], and we consider some additional properties that such measures could fulfill. After that, we introduce a class of consen-

[1] For a survey and references, see Martínez-Panero [18].

[2] Recently, García-Lapresta [11] has introduced a class of agreement measures in the context of weak orders when voters classify alternatives within a finite scale defined by linguistic categories with associated scores. These measures are based on distances among individual and collective scores generated by an aggregation operator.

sus measures based on the distances among individual weak orders[3]. We pay special attention to some specific metrics based on discrete, Minkowski, Chebishev and cosine distances.

We also introduce two consensus measures related to the Borda rule [5]. Although the Borda rule [4] was initially introduced under the assumption of individual preferences are linear orders, there exist some extensions to weak orders (see Black [4], Gärdenfors [12] and Young [25], among others). We follow one extended Borda count where relative positions of tied alternatives are the average of corresponding positions after a linearization of the initial weak order (see Smith [23], Black [4] and Cook and Seiford [9]).

The mentioned consensus measures related to the Borda rule are defined through distances among individual preferences and the outcome provided by the Borda rule. These distances among weak orders have been defined as Euclidean distances among vectors that represent positions of alternatives in corresponding weak orders.

We analyze properties of the introduced consensus measures, and we propose a decision procedure that weighs opinions of voters by their relative contributions to consensus. This process penalizes voters who are far from majority opinion and it favor voters that more contribute to consensus.

The paper is organized as follows. Section 2 is devoted to introduce basic terminology and distances used along the paper. In Section 3 we introduce consensus measures and we analyze their properties. Section 4 includes a proposal for measuring the marginal contribution to consensus of each voter based on a previously fixed consensus measure. Moreover, we also propose a decision procedure that weights the opinions of voters by their relative contributions to consensus.

2 Preliminaries

Consider a set of voters $V = \{v_1, \ldots, v_m\}$ $(m \geq 3)$ who show their preferences over a set of alternatives $X = \{x_1, \ldots, x_n\}$ $(n \geq 3)$. With $L(X)$ we denote the set of linear orders on X, and with $W(X)$ the set of weak orders on X. Given $R \in W(X)$, the *inverse* of R is the weak order R^{-1} defined by $x_i \, R^{-1} \, x_j \;\Leftrightarrow\; x_j \, R \, x_i$, for all $x_i, x_j \in X$.

A *profile* is a vector $\boldsymbol{R} = (R_1, \ldots, R_m)$ of weak or linear orders, where R_i contains the preferences of the voter v_i. Given a profile $R = (R_1, \ldots, R_m)$, we denote $R^{-1} = (R_1^{-1}, \ldots, R_m^{-1})$.

If π is a permutation on $\{1, \ldots, m\}$ and $\emptyset \neq I \subseteq V$, we denote $\boldsymbol{R}_\pi = (R_{\pi(1)}, \ldots, R_{\pi(m)})$ and $I_\pi = \{v_{\pi^{-1}(i)} \mid v_i \in I\}$, i.e., $v_j \in I_\pi \;\Leftrightarrow\; v_{\pi(j)} \in I$.

Given a permutation σ on $\{1, \ldots, n\}$, we denote with $\boldsymbol{R}^\sigma = (R_1^\sigma, \ldots, R_m^\sigma)$ the profile that results of recalling in $\boldsymbol{R}$ the alternatives according to σ, i.e., $x_i \, R_k \, x_j \;\Leftrightarrow\; x_{\sigma(i)} \, R_k^\sigma \, x_{\sigma(j)}$ for all $i, j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, m\}$.

The cardinal of any subset I is denoted by $|I|$. With $\mathcal{P}(V)$ we denote the power set of V, i.e., $I \in \mathcal{P}(V) \Leftrightarrow I \subseteq V$, and $\mathcal{P}_2(V) = \{I \in \mathcal{P}(V) \mid |I| \geq 2\}$. Notice that $|\mathcal{P}_2(V)| = |\mathcal{P}(V)| - |V| - 1 = 2^m - m - 1$.

2.1 Codification of weak orders

We now introduce a system for codifying linear and weak orders by means of vectors which represent the relative position of each alternative in the corresponding order. Similar procedures have been considered in the generalization of scoring rules from linear orders to weak orders (see Smith [23], Black [4] and Cook and Seiford [9], among others).

Given a profile $(R_1, \ldots, R_m) \in L(X)^m$ of linear orders, consider the mapping $o_i : X \longrightarrow \{1, \ldots, n\}$ which assigns the position of each alternative in R_i. Thus, the vector $(o_i(x_1), \ldots, o_i(x_n)) \in \{1, \ldots, n\}^n$ determines the corresponding linear order, and we denote $R_i \equiv (o_i(x_1), \ldots, o_i(x_n))$.

There does not exist a unique system for codifying weak orders. We propose one based on linearizing the weak order and to assign each alternative the average of the positions of the alternatives within the same equivalence class. Thus, given a profile of weak orders $(R_1, \ldots, R_m) \in W(X)^m$, the mapping

$$o_i : X \longrightarrow \{1,\ 1.5,\ 2,\ 2.5, \ldots,\ n - 0.5,\ n\}$$

assigns the relative position of each alternative in R_i.

So, if 7 alternatives are arranged in the weak order $x_2 \sim x_3 \sim x_5 \succ x_1 \succ x_4 \sim x_7 \succ x_6$, then this weak order is codified by the vector $(4,\ 2,\ 2,\ 5.5,\ 2,\ 7,\ 5.5)$.

Remark 1 Because of ties, weak orders on $W(X)$ not necessarily assign all positions $1, \ldots, n$ to alternatives. However, it is clear that the sum of all positions is $1 + 2 + \cdots + n$, as in linear orders. Consequently, for

[3] The use of distances for obtaining the outcome in a voting system has been widely considered in the literature. On this, see Kemeny [15], Slater [22], Baigent [2, 3], Nurmi [21] and Meskanen and Nurmi [19, 20], among others.

[4] The Borda rule plays the same role that arithmetic mean does for obtaining a representative value from a set of real numbers. In fact, the position of an alternative in the aggregated preference provided by the Borda rule is just the average of individual positions of the considered alternative.

every $R_i \in W(X)$, it holds

$$\sum_{j=1}^{n} o_i(x_j) = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Remark 2 Notice that if $R_i \in W(X)$ is codified by the vector $(o_i(x_1), \ldots, o_i(x_n))$, then R_i^{-1} is codified by $(n+1-o_i(x_1), \ldots, n+1-o_i(x_n))$.

2.2 Metrics

We now introduce a simple procedure for constructing a metric on $W(X)$ from a metric on $\mathbb{R}^n$.

Definition 1 *Given a metric $d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow [0,\infty)$, the* metric on $W(X)$ induced by d *is the mapping $\bar{d} : W(X) \times W(X) \longrightarrow [0,\infty)$ defined by $\bar{d}(R_1, R_2) = d\big((o_1(x_1), \ldots, o_1(x_n)), (o_2(x_1), \ldots, o_2(x_n))\big)$, for all $R_1, R_2 \in W(X)$.*

Example 1 Typical examples of metrics in $\mathbb{R}^n$ are the following:

1. The discrete metric d',

$$d'\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) = \\ = \begin{cases} 1, & \text{if } (a_1, \ldots, a_n) \neq (b_1, \ldots, b_n), \\ 0, & \text{if } (a_1, \ldots, a_n) = (b_1, \ldots, b_n). \end{cases}$$

2. For every $p \geq 1$, the Minkowski metric d_p,

$$d_p\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) = \left(\sum_{i=1}^{n} |a_i - b_i|^p \right)^{\frac{1}{p}}.$$

   For $p = 1$ and $p = 2$ we have the Manhattan and Euclidean metrics, respectively.

3. The Chebyshev metric d_∞,

$$d_\infty\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) = \\ = \max\big\{|a_1 - b_1|, \ldots, |a_n - b_n|\big\}.$$

4. The cosine metric d_c,

$$d_c\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) = \\ = 1 - \frac{\sum_{i=1}^{n} a_i\, b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\, \sqrt{\sum_{i=1}^{n} b_i^2}}.$$

Definition 2 *A metric $d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow [0,\infty)$ is* neutral *if for every permutation σ on $\{1, \ldots, n\}$, it holds*

$$d\big((a_{\sigma(1)}, \ldots, a_{\sigma(n)}), (b_{\sigma(1)}, \ldots, b_{\sigma(n)})\big) = \\ = d\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big),$$

for all $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{R}^n$.

Remark 3 All the metrics introduced in Example 1 are neutral.

3 Consensus measures

Consensus measures have been analyzed by Bosch [7] in the context of linear orders. We now extend this concept to the framework of weak orders. Notice that we use a unanimity condition weaker than that of Bosch.

Definition 3 *A* consensus measure *on $W(X)^m$ is a mapping $\mathcal{M} : W(X)^m \times \mathcal{P}_2(V) \longrightarrow [0,1]$ that satisfies the following conditions:*

1. Weak unanimity. *For every $\boldsymbol{R} \in W(X)^m$, it holds*

$$\mathcal{M}(\boldsymbol{R}, V) = 1 \;\Leftrightarrow\; R_1 = \cdots = R_m.$$

2. Anonymity. *For all permutation π on $\{1, \ldots, m\}$, $\boldsymbol{R} \in W(X)^m$ and $I \in \mathcal{P}_2(V)$, it holds*

$$\mathcal{M}(\boldsymbol{R}_\pi, I_\pi) = \mathcal{M}(\boldsymbol{R}, I).$$

3. Neutrality. *For all permutation σ on $\{1, \ldots, n\}$, $\boldsymbol{R} \in W(X)^m$ and $I \in \mathcal{P}_2(V)$, it holds*

$$\mathcal{M}(\boldsymbol{R}^\sigma, I) = \mathcal{M}(\boldsymbol{R}, I).$$

Weak unanimity means that the maximum consensus in the set of all decision makers is only achieved when all opinions are the same. Anonymity requires symmetry with respect to decision makers, and neutrality means symmetry with respect to alternatives.

We now introduce other properties that a consensus measure could satisfy.

Definition 4 *Let $\mathcal{M} : W(X)^m \times \mathcal{P}_2(V) \longrightarrow [0,1]$ be a consensus measure.*

1. *$\mathcal{M}$ satisfies* strong unanimity *if for all $\boldsymbol{R} \in W(X)^m$ and $I \in \mathcal{P}_2(V)$, it holds*

$$\mathcal{M}(\boldsymbol{R}, I) = 1 \;\Leftrightarrow\; R_i = R_j \;\text{ for all }\; v_i, v_j \in I.$$

2. *$\mathcal{M}$ satisfies* maximum dissension *if for all $\boldsymbol{R} \in W(X)^m$ and $v_i, v_j \in V$ such that $i \neq j$, it holds $\mathcal{M}(\boldsymbol{R}, \{v_i, v_j\}) = 0 \;\Leftrightarrow\; (R_i, R_j \in L(X)$ and $R_j = R_i^{-1})$.*

3. *$\mathcal{M}$ is* reciprocal *if for all $\boldsymbol{R} \in W(X)^m$ and $I \in \mathcal{P}_2(V)$, it holds*

$$\mathcal{M}(\boldsymbol{R}^{-1}, I) = \mathcal{M}(\boldsymbol{R}, I).$$

4. $\mathcal{M}$ *is* homogeneous *if for all* $\boldsymbol{R} \in W(X)^m$, $I \in \mathcal{P}_2(V)$ *and* $t \in \mathbb{N}$, *it holds*

$$\mathcal{M}^t(t\boldsymbol{R}, t\,I) = \mathcal{M}(\boldsymbol{R}, I),$$

where $\mathcal{M}^t : W(X)^{tm} \times \mathcal{P}_2(t\,V) \longrightarrow [0,1]$, $t\boldsymbol{R} = (\boldsymbol{R}, \overset{t}{\ldots}, \boldsymbol{R}) \in W(X)^{tm}$ *is the profile defined by* t *copies of* $\boldsymbol{R}$ *and* $t\,I = I \uplus \overset{t}{\cdots} \uplus I$ *is the set of voters defined by* t *copies of* I.

Strong unanimity means that the maximum consensus in every subset of decision makers is only achieved when all opinions are the same. Obviously, strong unanimity implies weak unanimity.

Maximum dissension means that in each subset of two voters[5], the minimum consensus is only reached whenever preferences of voters are linear orders and each one is the inverse of the other.

Reciprocity means that if all individual weak orders are reversed, then the consensus does not change.

Homogeneity means that if we replicate a subset of voters, then the consensus in that group does not change.

3.1 Consensus measures based on metrics

Definition 5 *Given a metric* $\bar{d}$ *on* $W(X)$, *the mapping*

$$\mathcal{M}_{\bar{d}} : W(X)^m \times \mathcal{P}_2(V) \longrightarrow [0,1]$$

is defined by

$$\mathcal{M}_{\bar{d}}(\boldsymbol{R}, I) = 1 - \frac{\sum_{v_i, v_j \in I\,,\, i<j} \bar{d}(R_i, R_j)}{\binom{|I|}{2} \cdot \Delta_n}\,,$$

where

$$\Delta_n = \max\{\bar{d}(R_i, R_j) \mid R_i, R_j \in W(X)\}.$$

Proposition 1 *For every metric* $\bar{d}$ *on* $W(X)$, $\mathcal{M}_{\bar{d}}$ *satisfies strong unanimity and anonymity.*

If $\mathcal{M}_{\bar{d}}$ is neutral, then we say that $\mathcal{M}_{\bar{d}}$ is the *consensus measure associated with* $\bar{d}$.

Proposition 2 *If* $d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow [0, \infty)$ *is a neutral metric, then* $\mathcal{M}_{\bar{d}}$ *is a consensus measure.*

Proposition 3 *If* $\bar{d}$ *is the metric on* $W(X)$ *induced by* d', d_p, *with* $p \geq 1$, d_∞ *or* d_c, *then* $\mathcal{M}_{\bar{d}}$ *is a reciprocal consensus measure.*

[5]It is clear that a society reach maximum consensus when all the opinions are the same. However, in a society with more than two members it is not an obvious issue to determine when there is minimum consensus (maximum disagreement).

Proposition 4 *If* $\bar{d}$ *is the metric induced by* d_p, *with* $p > 1$, *or* d_c, *then* $\mathcal{M}_{\bar{d}}$ *satisfies the maximum dissension property.*

Proposition 5 *If* $\bar{d}$ *is the metric induced by* d', d_1 *or* d_∞, *then* $\mathcal{M}_{\bar{d}}$ *does not satisfy the maximum dissension property.*

Proposition 6 *The consensus measure* $\mathcal{M}_{\bar{d}}$ *is not homogeneous for any distance* $\bar{d}$ *on* $W(X)$.

Proposition 7 *If* $\boldsymbol{R} \in L(X)^2$ *and* $R_2 = R_1^{-1}$, *then for any distance* $\bar{d}$ *on* $W(X)$ *it holds:*

$$\lim_{t \to \infty} \mathcal{M}^t_{\bar{d}}(t\,\boldsymbol{R}, t\,I) = \frac{1}{2}.$$

3.2 Borda consensus measures

Given a profile of weak orders $(R_1, \ldots, R_m) \in W(X)^m$ and $I \in \mathcal{P}_2(V)$, let $o^*_I : X \longrightarrow \mathbb{R}$ be the mapping defined by

$$o^*_I(x_j) = \frac{1}{|I|} \sum_{i \in I} o_i(x_j)\,,$$

which assigns the average position to each alternative for the set of voters I.

Notice that the weak order R^*_I defined by $x_j\, R^*_I\, x_k \Leftrightarrow o^*_I(x_j) \leq o^*_I(x_k)$ corresponds to the *Borda rule* for the set of voters I.

Definition 6 *The* absolute Borda consensus measure *is the mapping*

$$\mathcal{M}_B : W(X)^m \times \mathcal{P}_2(V) \longrightarrow [0,1]$$

defined by

$$\mathcal{M}_B(\boldsymbol{R}, I) = 1 - \frac{\sum_{v_i \in I} \bar{d}(R_i, R^*_V)}{|I| \cdot \Delta_n},$$

where R^*_V *is the weak order associated with the outcome provided by the Borda rule for the whole set of voters* V, $\bar{d}$ *the metric on* $W(X)$ *induced by the Euclidean metric and*

$$\Delta_n = \max\{\bar{d}(R_i, R_j) \mid R_i, R_j \in W(X)\}.$$

Proposition 8 $\mathcal{M}_B$ *is a consensus measure that satisfies reciprocity and homogeneity.*

Proposition 9 $\mathcal{M}_B$ *does not satisfy strong unanimity.*

Definition 7 *The* relative Borda consensus measure *is the mapping*

$$\mathcal{M}'_B : W(X)^m \times \mathcal{P}_2(V) \longrightarrow [0,1]$$

defined by

$$\mathcal{M}'_B(\boldsymbol{R}, I) = 1 - \frac{\sum_{v_i \in I} \bar{d}(R_i, R_I^*)}{|I| \cdot \Delta_n},$$

where R_I^* is the weak order associated with the outcome provided by the Borda rule for the set of voters I and $\bar{d}$ the metric on $W(X)$ induced by the Euclidean metric.

Proposition 10 *$\mathcal{M}'_{\mathcal{B}}$ is a consensus measure that satisfies strong unanimity, reciprocity and homogeneity.*

Proposition 11 *$\mathcal{M}_B$ and $\mathcal{M}'_B$ do not satisfy the maximum dissension property.*

We summarize the properties of the analyzed consensus measures in Table 1.

Table 1: Summary

	S.U.	M.D.	R.	H.
$\mathcal{M}_{\bar{d}}$ for d'	Yes	No	Yes	No
$\mathcal{M}_{\bar{d}}$ for d_1	Yes	No	Yes	No
$\mathcal{M}_{\bar{d}}$ for d_p, $p > 1$	Yes	Yes	Yes	No
$\mathcal{M}_{\bar{d}}$ for d_∞	Yes	No	Yes	No
$\mathcal{M}_{\bar{d}}$ for d_c	Yes	Yes	Yes	No
$\mathcal{M}_B$	No	No	Yes	Yes
$\mathcal{M}'_B$	Yes	No	Yes	Yes

S.U.: Strong unanimity; M.D.: Maximum dissension; R.: Reciprocity; H.: Homogeneity.

4 Favoring consensus in group decision making

We now introduce an index for each voter that shows what is his/her contribution to consensus. This index has some similarities with the Shapley value in cooperative game theory.

Definition 8 *Given a consensus measure $\mathcal{M}$, the mapping $\boldsymbol{c} : W(X)^m \longrightarrow [-1,1]^m$ is defined by $\boldsymbol{c}(\boldsymbol{R}) = \big(c_1(\boldsymbol{R}), \ldots, c_m(\boldsymbol{R})\big)$ with*

$$c_i(\boldsymbol{R}) = \frac{\sum_{I \in S(i)} \Big(\mathcal{M}\big(\boldsymbol{R}, I \cup \{v_i\}\big) - \mathcal{M}(\boldsymbol{R}, I)\Big)}{|S(i)|},$$

$i = 1, \ldots, m$, where $S(i) = \{I \in \mathcal{P}_2(V) \mid v_i \notin I\}$.

We say that $c_i(\boldsymbol{R})$ is the marginal contribution to consensus *of voter v_i with respect to $\boldsymbol{R} \in W(X)^m$.*

Proposition 12 *Let $\bar{d} : W(X) \times W(X) \longrightarrow [0, \infty)$ be a metric. For the consensus measure $\mathcal{M}_{\bar{d}}$ and $\mathcal{M}_{\mathcal{B}}$, it holds $c_1(\boldsymbol{R}) + \cdots + c_m(\boldsymbol{R}) = 0$, for every profile $\boldsymbol{R} \in W(X)^m$.*

We can use the vector $\boldsymbol{c}(\boldsymbol{R})$ for prioritizing the voters in order of their contribution to consensus, as suggested by Cook, Kress and Seiford [8]. In this way, we introduce a new index

$$c'_i(\boldsymbol{R}) = c_i(\boldsymbol{R}) - \min\{c_1(\boldsymbol{R}), \ldots, c_m(\boldsymbol{R})\},$$

for every $i \in \{1, \ldots, m\}$. We now define a weight for each voter:

$$w_i(\boldsymbol{R}) = \begin{cases} \dfrac{c'_i(\boldsymbol{R})}{c'_1(R) + \cdots + c'_m(R)}, \\ \text{if } \; c'_1(\boldsymbol{R}) + \cdots + c'_m(\boldsymbol{R}) \neq 0 \\ \dfrac{1}{m}, \; \text{if } \; c'_1(\boldsymbol{R}) + \cdots + c'_m(\boldsymbol{R}) = 0, \end{cases}$$

for $i = 1, \ldots, m$.

Notice that $w_i(\boldsymbol{R}) \in [0,1]$ for every $i \in \{1, \ldots, m\}$, and $w_1(\boldsymbol{R}) + \cdots + w_m(\boldsymbol{R}) = 1$.

Now, for each decision maker $v_i \in V$ we multiply the position of each alternative $o_i(x_j)$ by his/her weight $w_i(\boldsymbol{R})$ for assigning collective positions to the alternatives:

$$O(x_j) = \sum_{i=1}^{m} w_i(R) \cdot o_i(x_j), \quad j = 1, \ldots, n.$$

Thus, we can order the alternatives through the weak order $\succeq$ on X defined by

$$x_j \succeq x_k \;\Leftrightarrow\; O(x_j) \leq O(x_k).$$

4.1 An illustrative example

To illustrate the above decision making procedure, we consider the set of voters $V = \{v_1, v_2, v_3, v_4, v_5\}$ that rank order the alternativas of $X = \{x_1, \ldots, x_7\}$ by means of the weak orders included in Table 2.

Table 2: Individual orders

R_1	R_2	R_3	R_4	R_5
$x_2\ x_3\ x_5$	x_5	x_3	$x_1\ x_2$	$x_1\ x_4$
x_1	$x_1\ x_3$	x_4	x_5	x_7
$x_4\ x_7$	x_2	x_1	x_3	x_2
x_6	$x_4\ x_6\ x_7$	$x_2\ x_5$	x_4	$x_3\ x_5\ x_6$
		$x_6\ x_7$	$x_6\ x_7$	

If we consider the consensus measure $\mathcal{M}_{\bar{d}}$ for d_c, we obtain the marginal contributions to consensus included in Table 3.

Table 3: Marginal contributions to consensus

$$c_1(\boldsymbol{R}) = 0.05114639$$
$$c_2(\boldsymbol{R}) = 0.03468305$$
$$c_3(\boldsymbol{R}) = 0.01296238$$
$$c_4(\boldsymbol{R}) = 0.07872037$$
$$c_5(\boldsymbol{R}) = -0.17751219$$

Table 4: Individual weights

$$w_1(\boldsymbol{R}) = 0.25762578$$
$$w_2(\boldsymbol{R}) = 0.23907680$$
$$w_3(\boldsymbol{R}) = 0.21460449$$
$$w_4(\boldsymbol{R}) = 0.28869292$$
$$w_5(\boldsymbol{R}) = 0$$

These marginal contributions to consensus induce the weights included in Table 4.

We now weigh the position of each alternative by these individual weights, and we obtain the collective positions included in Table 5.

Table 5: Collective positions

$$O(x_1) = 2.705047994$$
$$O(x_2) = 2.870318382$$
$$O(x_3) = 2.482319753$$
$$O(x_4) = 4.724076212$$
$$O(x_5) = 2.586127353$$
$$O(x_6) = 6.509274488$$
$$O(x_7) = 6.122835819$$

Taking into account these aggregated positions we have the following linear order on X:

$$\mathcal{M}_{\bar{d}} \text{ for } d_c \quad x_3 \succ x_5 \succ x_1 \succ x_2 \succ x_4 \succ x_7 \succ x_6$$

If we apply the same procedure taking into account consensus measures $\mathcal{M}_{\bar{d}}$ for metrics d_p, with $p = 2$, and d_∞, and the Borda consensus measures, then we obtain the following orders:

$$\begin{array}{ll} \mathcal{M}_{\bar{d}} \text{ for } d_2 & x_5 \succ x_3 \succ x_1 \succ x_2 \succ x_4 \succ x_7 \succ x_6 \\ \mathcal{M}_{\bar{d}} \text{ for } d_\infty & x_5 \succ x_2 \succ x_1 \sim x_3 \succ x_4 \succ x_7 \succ x_6 \\ \mathcal{M}_B & x_3 \succ x_5 \succ x_1 \succ x_2 \succ x_4 \succ x_7 \succ x_6 \\ \mathcal{M}'_B & x_5 \succ x_3 \succ x_1 \succ x_2 \succ x_4 \succ x_7 \succ x_6 \end{array}$$

Remark 4 The Borda count is equivalent to use

$$O_B(x_j) = \frac{1}{m}\sum_{i=1}^{m} o_i(x_j)\,, \quad j = 1, \ldots, n,$$

and the weak order $\succeq_B$ on X defined by

$$x_j \succeq_B x_k \Leftrightarrow O_B(x_j) \leq O_B(x_k).$$

If we apply the Borda count to the weak orders of Table 2, then we would obtain the following collective order:

$$\succeq_B \quad x_1 \succ x_3 \succ x_2 \succ x_5 \succ x_4 \succ x_7 \succ x_6$$

quite different to those obtained by applying the proposed weighted decision making procedure.

4.2 Some remarks

It is interesting to emphasize the following features of our decision making procedure:

- Due to the marginal contribution to consensus indices and corresponding weights are usually irrational numbers, so when the number of decision makers is high, it is unlikely that the decision procedure provides ties among alternatives.
- Since the proposed decision procedure penalizes individuals who are far from majority positions, this provides incentives for decision makers to moderate their opinions which, otherwise, may be excluded or underestimated.
- It is easy to implement computer programs for generating the final ranking on the set of alternatives.

Acknowledgements

This paper has been partially supported by Spanish Ministry of Education and Science grant SEJ2006-04267/ECON, and ERDF. The authors are grateful to Jorge Alcalde-Unzu, Miguel Ángel Ballester and Christian Klamler for their suggestions and comments.

References

[1] Alcalde, J., Vorsatz, M. (2008): The measurement of consensus in social choice. *ESTYLF 2008.*

[2] Baigent, N. (1987): Preference proximity and anonymous social choice. *The Quarterly Journal of Economics* 102, pp. 161-169.

[3] Baigent, N. (1987): Metric rationalization of social choice functions according to principles of social choice. *Mathematical Social Sciences* 1, pp. 59-65.

[4] Black, D. (1976): Partial justification of the Borda count. *Public Choice* 28, pp. 1-15.

[5] Borda, J.C. de (1784): *Mémoire sur les élections au scrutin*. Historie de l'Academie Royale des Sciences, Paris.

[6] Borroni, C.G., Zenga, M. (2007): A test of concordance based on Gini's mean difference. *Statistical Methods and Applications* 16, pp. 289-308.

[7] Bosch, R. (2005): *Characterizations of Voging Rules and Consensus Measures*. Ph. D. Dissertation, Tilburg University.

[8] Cook, W.D., Kress, M., Seiford, L.M. (1996): A general framework for distance-based consensus in ordinal ranking models. *European Journal of Operational Research* 96, pp. 392-397.

[9] Cook, W.D., Seiford, L.M. (1982): On the Borda-Kendall consensus method for priority ranking problems. *Management Science* 28, pp. 621-637.

[10] Eklund, P., Rusinowska, A., de Swart, H. (2007): Consensus reaching in committees. *European Journal of Operational Research* 178, pp. 185-193.

[11] García-Lapresta, J.L. (2008): Favoring consensus and penalizing disagreement in group decision making. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, in press.

[12] Gärdenfors, P. (1973): Positionalist voting functions. *Theory and Decision* 4, pp. 1-24.

[13] Gini, C. (1954): *Corso di Statistica*. Veschi, Rome.

[14] Hays, W.L. (1960): A note on average tau as a measure of concordance. *Journal of the American Statistical Association* 55, pp. 331-341.

[15] Kemeny, J. (1959): Mathematics without numbers. *Daedalus* 88, pp. 571-591.

[16] Kendall, M.G. (1962): *Rank Correlation Methods*. Griffin, London.

[17] Kendall, M., Gibbons, J.D. (1990): *Rank Correlation Methods*. Oxford University Press, New York.

[18] Martínez-Panero, M. (2008): Approaches to consensus: The state of the art. *ESTYLF 2008*.

[19] Meskanen, T., Nurmi, H. (2006): Analyzing political disagreement. *JASM Working Papers* 2, University of Turku, Finland.

[20] Meskanen, T., Nurmi, H. (2007): Distance from consensus: A theme and variations. In *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice* (eds.: Simeone, B., Pukelsheim, F.). Springer, pp. 117-132.

[21] Nurmi, H. (2004): A comparison of some distance-based choice rules in ranking environments. *Theory and Decision* 57, pp. 5-24.

[22] Slater, P. (1961): Inconsistencies in a schedule of paired comparisons. *Biometrika* 48, pp. 303-312.

[23] Smith, J. (1973): Aggregation of preferences with variable electorate. *Econometrica* 41, pp. 1027-1041.

[24] Spearman, C. (1904): The proof and measurement of association between two things. *American Journal of Psychology* 15, pp. 72-101.

[25] Young, H.P. (1974): An axiomatization of Borda's rule. *Journal of Economic Theory* 9, pp. 43-52.

Penalizing manipulation strategies in Consensus processes

F. Mata L. Martínez J.C. Martínez

Dept. of Computer Science. University of Jaén, 23071 - Jaén, Spain.
{fmata, martin}@ujaen.es, jcmc0006@estudiante.ujaen.es

Abstract

In this work, how to minimize the effects of manipulation strategies in group decision making (GDM) problems is addressed. In this kind of problems a consensus phase may be carried out in order to achieve a minimum level of agreement before making a decision on a set of alternatives. In this setting, we discuss the possibility that an expert tries to reach his personal goal rather than the group one. One way to do this is manipulating the opinions that the expert provides to group opinion and another one is not changing his opinions like a strategy to impose his opinion to the group. The goal of this work is to propose a consensus model that provides a penalization mechanism in order to prevent the effects of both kinds of malicious behaviors. To achieve it, the mechanism assigns to each expert an importance which may change according to expert' behavior during the consensus process.

Keywords: Consensus, group decision making, penalization, manipulation.

1 INTRODUCTION

GDM problems are decision situations where given a set of possible alternatives, a group of experts try to achieve a common solution taking into account their individual opinions. Usually, GDM problems have been solved carrying out Selection Processes where experts obtain the best set of alternatives to solve the problem from their opinions [6]. However, some experts might consider that their opinions have not been taken into account in order to obtain the solution, and therefore they might not agree with the proposed solution. To avoid this situation, it is suitable to carry out a consensus process consisting of several rounds (see Figure 1) where experts discuss and change their opinions in order to reach a minimum agreement before making a decision [2, 7, 8, 9].

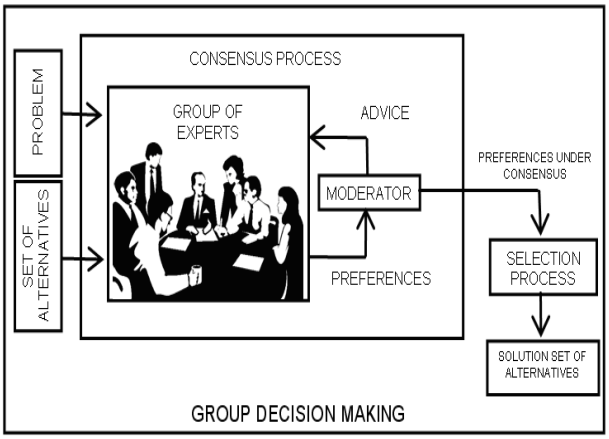


Figure 1: Resolution process of a GDM problem

A consensus process is usually coordinated by a moderator who evaluates the level of agreement of each round and helps experts to make their opinions closer to each other. However, the moderator may not be objective and to manage the consensus process toward his own goals. One way to avoid this situation is to automate the consensus reaching processes [7, 8, 10].

Most consensus processes are carried out into impartiality environments where all the experts' opinions are treated at the same level of importance [9]. However, the true goal of an expert could be the selection of his most preferred alternative and so he could be tempted to strategically manipulate his assessments in order to attain this goal [11, 13, 16]. In addition, an expert can decide not to follow the advice suggested by the moderator like a strategy to enhance his opinions, forcing the other group members to change their opinions

towards the opinion of the manipulative expert.

To prevent the first type of manipulation (called *strategic manipulation*), some mechanisms have been proposed in the literature [5, 16, 17]. Also we can find an approach to combat the second type of malicious behavior in [12], but there is no any proposal to deal with both kinds of no desirable behaviors in consensus processes.

The aim of this contribution is to propose a consensus model that includes an initial approach of a penalization mechanism to minimize the effects of both types of manipulation attempts. To do so, in each consensus round, the model calculates the importance degree associated with each expert in basis to his predisposition for agreement. Then the collective opinion is obtained by means of an uninorm aggregation operator that takes into account the importance degrees. In this way, the model achieves to penalize manipulative experts. Finally, the model generates a set of changes for those experts whose assessments are more distant from the collective opinion.

This work is organized as follows. In Section 2, we introduce the GDM problems addressed in this work. In Section 3, the consensus model is briefly explained. In Section 4 we describe the penalization mechanism. In Section 5 the aggregation process based on uninorms is presented. Finally, some conclusions are drawn in Section 6.

2 GDM PROBLEMS AND PREFERENCES EXPRESSION

GDM problems are classically defined as decision situations in which a set of individuals (also called experts) $E = \{e_1, e_2, \ldots, e_m\}$ $(m \geq 2)$, try to choose the best alternative/s from a set of alternatives $X = \{x_1, x_2, \ldots, x_n\}$ $(n \geq 2)$. Utility vectors are usual preference structures used by the experts to give their opinions [3, 14],

$$U_{e_i} = \{u_i^1, u_i^2, \ldots, u_i^n\}$$

where each $u_i^k \in [0, 1]$ represent the utility or value given by expert e_i for the alternative x_k. It is assumed that a greater value of u_i^k, implies a greater satisfaction of expert e_i on the alternative x_k. In this contribution, we deal with GDM problems where experts use utility vectors to provide their opinions.

3 MODEL DESCRIPTION

In this work we propose a penalization mechanism for a consensus reaching model in GDM problems. This model consists of three phases (see Figure 2):

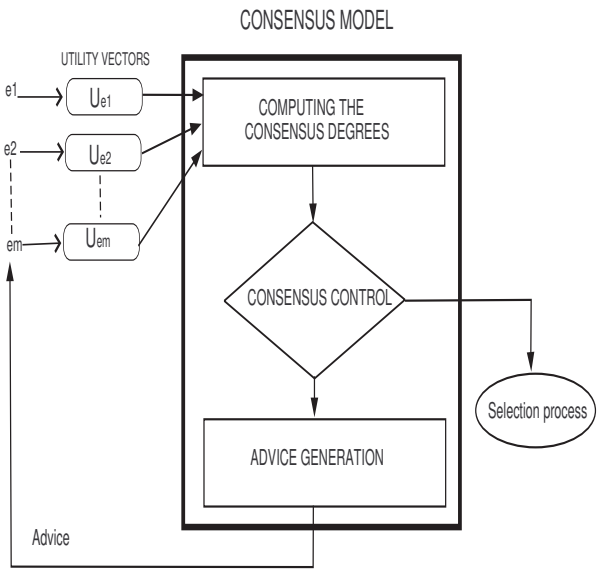


Figure 2: Consensus model

1. *Computing the consensus degree, cr.* The consensus degree cr measures the level agreement among the experts. Firstly, the model obtains a similarity vector for each pair of experts e_i, e_j with $(i < j)$

$$sv_{i,j} = \{s(u_i^1, u_j^1), s(u_i^2, u_j^2), \ldots, s(u_i^n, u_j^n)\}.$$

Each element of the vector is the similarity at level of alternatives between two experts' opinions. It is obtained from similarity function $s(\cdot)$:

$$s(u_i^k, u_j^k) = 1 - \mid u_i^k - u_j^k \mid, \qquad (1)$$

where u_i^k and u_j^k are the utility values given to alternative x_k by experts e_i and e_j respectively. Next, the model computes a consensus vector CV,

$$CV = \{cv^1, \ldots, cv^n\},$$

where each element of vector is obtained as the arithmetic mean of similarity vectors

$$cv^k = \theta(sv_{i,j}^k),\ i, j = 1, \ldots, m,\ i < j,\ k = 1, \ldots, n.$$

Finally, the consensus degree cr, which represents the overall agreement among all experts, is calculated as the average of values contained in CV,

$$cr = \frac{\sum_{k=1}^{n} cv^k}{n}.$$

2. *Consensus control.* In this phase the level of agreement cr is checked. If cr is greater or equal than a given consensus threshold γ then the consensus reaching process would stop. Otherwise,

the consensus process keeps going. The consensus threshold γ is the desired minimum consensus before starting the selection process and it should be fixed in advance.

3. *Advice generation.* If cr is not high enough, the model suggests a set of changes to the experts in order to make their preferences closer to each other. To do so, the model computes the collective utility vector by aggregating all experts' opinions

$$COLV = \{u_c^1, u_c^2, \ldots, u_c^n\},$$

where each

$$u_c^k = \delta(u_1^k, u_2^k, \ldots, u_m^k).$$

In this work, we propose to use an uninorm based operator δ as aggregation operator such as we explain in Section 5.

Moreover, a proximity vector

$$pv_i = \{s(u_i^1, u_c^1), s(u_i^2, u_c^2), \ldots, s(u_i^n, u_c^n)\}$$

and a proximity value p_i,

$$p_i = \frac{\sum_{k=1}^{n} pv_i^k}{n}$$

are calculated for each expert e_i. These values represent the similarity between expert's opinion and the collective one.

Then, the model proceeds as follow:

i. It selects the set of experts with less proximity, called $EXPCH$. These experts are the furthest from collective opinion and therefore they should change their opinions in order to increase the level of agreement in the next consensus round.

ii. It identifies the set of alternatives where the level of agreement is lower,

$$ALT = \{x_l \in X \mid cv^l < cr\}.$$

iii. For each expert $e_i \in EXPCH$ and for each alternative $x_l \in ALT$, the model suggest the direction of the changes of opinion, i.e to increase or decrease the current utility value. To do this, the model takes into account the individual opinion as well as the collective one and it applies one of the following three "direction rules":

DR.1. If $(u_i^l - u_c^l) < 0$, the expert e_i should increase the utility given to alternative x_l.

DR.2. If $(u_i^l - u_c^l) > 0$, the expert e_i should decrease the utility given to alternative x_l.

DR.1. If $(u_i^l - u_c^l) = 0$, the expert e_i has not to modify the utility given to alternative x_l.

4 PENALIZATION MECHANISM

The purpose of a consensus reaching process is to achieve a high level of agreement before making a decision. We may assume that all experts want to reach the agreement and therefore they are willing to renounce their initial assessments in order to reach it. However, some experts may have other intentions and try to manipulate the consensus reaching process in order to increase their personal interests or benefits. The manipulation may be seem at least from two points of view:

i) *Strategic manipulation of the preferences.* This is the case of an expert who tries to impose his preferences by means of enhancing his most preferred alternative. For example, e_i can maximize the utility value of the alternative x_1 and minimize the utility values of the rest of alternatives, $U_{e_i} = \{1, 0, 0, \ldots, 0\}$, like a strategy to impose his opinion. In this case, e_i may be manipulating his opinions in order to make the collective opinion closer to his particular opinion.

ii) *Disobedient behavior.* On the other hand, let us suppose that an expert decides not to follow the advice suggested by the model forcing the other group members to change their opinions towards his individual opinion. This behavior may also be seem as a strategy to manipulate the consensus process and perhaps to obtain personal goals instead of group goals.

In this section a mechanism to penalize both types of manipulation strategies are proposed. This mechanism is based on computing an importance degree or weight for each expert. Experts' weights are taken into account by the aggregation operator when the collective preference is obtained. So, an expert who completely discounts all the alternatives except his favorite, or he manifests a disobedient behavior by ignoring the model's advice, he would have little influence on the construction of the group opinion. Both types of malicious behaviors are considered to compute the experts' importance such as we present in the following subsections.

4.1 STRATEGIC PREFERENCES MANIPULATION

In [17, 18] Yager proposed that the importance of an expert would be directly related to the total degree of satisfaction allocated, i.e., according to the assessments given on all the alternatives. The more satisfied expert is the most important, thus, "happy" experts

get more importance. So, if an expert rejects all alternatives except his preferred one, the expert's importance should be very low or null in the group opinion. This could be detected by checking the "extreme" utility values and computed its difference.

Let $t_i = 1 - (Max_j[u_i^j] - Min_j[u_i^j])$. The importance of an expert e_i that deliberately manipulates his preferences, wsm_i, is obtained as:

$$wsm_i = \frac{t_i}{Max_{(1 \leq k \leq m)}[t_k]} \quad (2)$$

That is, the expert is more important when his most extreme utility values are not very different.

4.2 DISOBEDIENT BEHAVIOR

This may happen when an expert decides not to apply the advice suggested by the model as strategy to impose his opinion. The importance of a "disobedient" expert, wdb_i, is computed as:

$$wdb_i = Max(0, 1 - \frac{n_rejections}{max_rejections}) \quad (3)$$

where $n_rejections$ is the number of times that expert e_i has refused to make the changes proposed by the model, and $max_rejections$ is the maximum number of times that an expert might refuse them. The $max_rejections$ value should be fixed before starting the consensus process.

It is obvious that in the first consensus round the system has not suggested any advice yet, and therefore $wdb_i = 1$ for all $e_i \in E$.

About the disobedient behavior there are several open questions that we are studying how to solve. For example, what happens if an expert follows the advice by just changing his assessments in a minimal fraction? Let us suppose expert e_i gives $U_{e_i} = \{0.4, 0.5, 0.9\}$ and the model suggests decreasing the first and third value and increasing the second one. Then the expert provides $U_{e_i} = \{0.3999, 0.5001, 0.8999\}$. He has applied the advice, but clearly he is not being honest. On the other hand, what happens if the expert does not apply all the advice? Let us suppose e_i decides to change only the first two values, ignoring the third one. The discussion about when an expert shows a disobedient behavior is not trivial. We can try to identify disobedient experts by defining two parameters: a minimum change tax and a minimum percentage of changes. For example, we can establish that the minimum variation of the utility values should be ± 0.25 and the minimum percentage of changes carried out by the expert bigger than 85%. So an expert which does not observe these conditions, he might be considered "disobedient".

4.3 COMPUTING THE IMPORTANCE DEGREE OF AN EXPERT

When both weights have been calculated, then the importance degree for an expert w_i can be obtained. This value is computed by aggregating both weights using the Min operator,

$$w_i = Min(wsm_i, wdb_i). \quad (4)$$

In this way, the model always chooses the lowest weigh and therefore the penalization is higher.

Afterwards the model recalculates the utility values given by each expert taking into account the w_i value. The new values obtained are aggregated in order to build the collective utility vector $COLV$. This aggregation process is described in the next section.

5 USING UNINORMS AS AGGREGATION OPERATOR

Several aggregation operators may be used to obtain the collective opinion. In [8], we used the mean operator to aggregate the experts' preferences. However this operator has a disadvantage, an expert cannot know whether his assessments will increase or decrease the "support" for an alternative because it depends on the other experts' assessments.

In this work we propose to use an aggregation operator based on uninorms. Uninorms are an unification and generalization of the t-norm and t-conorm operators [1, 4, 15]. A uninorm is a mapping $R : I \times I \rightarrow I$, having the following properties:

1. $R(x, y) = R(y, x)$: Commutativity.
2. $R(x, y) \geq R(u, v)$ for $x \geq u, y \geq v$: Monotonicity.
3. $R(x, R(y, z)) = R(R(x, y), z)$: Associativity.
4. There exists some $g \in [0, 1]$ such that for all $x, R(x, g) = x$: Identity element.

Using uninorms, we can see that additional arguments have an effect on the aggregation depending on their relationship to a boundary point, the identity g, but this boundary point doesn't vary with the prior value aggregated, i.e., g is fixed. The identity provides an "a priori" knowledge of the effect of the values provided by the expert, so by providing a greater value than the identity, an expert knows "a priori" that he is helping to increase support for an alternative. On the other hand, providing a smaller value than the identity g, an expert knows he is helping to decrease support for an alternative, and providing a value equal to g the expert

expresses his indifference with respect an alternative [18]. The behavior of a uninorm is very important in our case since according to property 4,

$$R(a_1, a_2, \ldots, a_n, g) = R(a_1, a_2, \ldots, a_n) \qquad (5)$$

then we can see that an utility value $u_i^k = g$ can be seen as an indifferent assessment with respect to the alternative x_k.

In our model, the result of the aggregating process is a collective utility vector

$$COLV = \{u_c^1, u_c^2, \ldots, u_c^n, \}$$

where each element is

$$u_c^k = R_g(u_1^k, u_2^k \ldots, u_m^k)$$

being R a uninorm operator with identity g.

In order to incorporate the experts' importance degrees in the aggregation process, we associate each u_i^k with its importance degree $w_i \in [0, 1]$ calculated in (4),

$$u_c^k = R_g((w_1, u_1^k), (w_2, u_2^k) \ldots, (w_m, u_m^k)).$$

Then, $COLV$ can be obtained as weighted uninorm aggregation. Let h_g be the importance transformation function that penalizes the experts' opinion and let $\hat{u}_i^k = h_g(w_i, u_i^k)$ be the penalized value, then we have that

$$u_c^k = R_g(\hat{u}_1^k, \hat{u}_2^k \ldots, \hat{u}_m^k).$$

The question now becomes: what is the form of function h_g. Several properties can be assigned to that kind of functions [17]:

i) It is required that $h_g(1, a) = a$, and $h_g(0, a) = g$, the identity of the uninorm. That is, the opinion from an expert that has 0 importance will not contribute to the aggregation.

ii) h_g should be monotonic with respect the value. Specifically, if a and $\hat{a}$ are two arguments such that $a \geq \hat{a}$ then any modification by importance will not interchange their order, and $\forall w$ $h_g(w, a) \geq h_g(w, \hat{a})$.

iii) $h_g(w, a)$ is bounded by its values for $w = 0$ and $w = 1$.

iv) As fourth condition we require that $h_g(w, a)$ transitions from $w = 0$ to $w = 1$ in a consistent monotonic manner. Let w and $\tilde{w}$ such that $\tilde{w} > w$. Then

- if $h_g(1, a) > h_g(0, a)$ $(a > g)$, then for all $\tilde{w} > w$ we have $h_g(\tilde{w}, a) \geq h_g(w, a)$;
- if $h_g(1, a) < h_g(0, a)$ $(a < g)$, then for all $\tilde{w} > w$ we have $h_g(\tilde{w}, a) \leq h_g(w, a)$;
- if $h_g(1, a) = h_g(0, a)$ $(a = g)$, then for all $\tilde{w} > w$ we have $h_g(\tilde{w}, a) \leq h_g(w, a)$.

In [17] we can find some interesting examples of such functions. Here there are two of them that are valid for any identity element g:
$h_g(w, a) = wa + \overline{w}g$,
$h_g(w, a) = (w \wedge a) \vee (\overline{w} \wedge g) \vee (a \wedge g)$,
with $\overline{w} = 1 - w$.

It can be seen that using a function h_g to transform the pairs (w_i, u_k^i) into a single numeric value $\hat{u}_i^k$ together with an uninorm to aggregate these values, we obtain an aggregation operator able to deal with experts with different importance degree and to penalize manipulative behaviors.

Example

Let us consider three experts and three alternatives. We use the function $h_g(w, a) = wa + \overline{w}g$ to obtain the penalized utility for any expert. As aggregation operator we use the called three Π operator proposed by Yager in [15],

$$R(x, y) = \frac{xy}{\overline{xy} + xy} \qquad (6)$$

where $\overline{x} = 1 - x, \overline{y} = 1 - y$. In this uninorm, the identity $g = 0.5$. Let us suppose the three experts' utility vectors are:

$U_{e_1} = \{1, 0, 0\}, U_{e_2} = \{0.4, 0.5, 0.9\}$ and $U_{e_3} = \{0.8, 0.7, 0\}$.

From those opinions, we might think e_1 is trying to strategically manipulate the consensus process. If we use (2) to compute the experts' importance associated with strategic manipulation behavior, we obtain the following results:

$$wsm_1 = 0, \quad wsm_2 = 1 \; and \; wsm_3 = 0.4.$$

Let us suppose we are in the first consensus round, and therefore $wdb_i = 1$ for all e_i (there is not disobedient behavior because the system has not generated any advice yet). Therefore, according to (4), the importance associated with each expert is

$$w_1 = 0, \quad w_2 = 1, \quad and \; w_3 = 0.4$$

We can see that the most important expert is e_2, and the less important expert is e_1. Taking into account these weights and using the function h_g described above, we obtain the penalized utility vector for each expert:

$\hat{U}_{e_1} = \{0.5, 0.5, 0.5\}$, $\hat{U}_{e_2} = \{0.4, 0.5, 0.9\}$ and $\hat{U}_{e_3} = \{0.62, 0.58, 0.3\}$.

Note that penalized preferences for expert e_2 (whose importance is $w_2 = 1$) have not changed, i.e, he has not been really penalized. However, all the penalized utilities of e_1 (whose importance degree is $w_1 = 0$) are 0.5, i.e. the value of g, the identity of the uninorm. This implies that according to (5), the opinions of expert e_1 will not affect to the collective vector, $COLV = \{0.52, 0.58, 0.79\}$, and therefore the other experts will not have to modify their opinions in order to be close to expert e_1.

6 CONCLUSIONS

In this contribution we have considered the possibility when experts try to manipulate a consensus process in order to impose their preferences on the group opinion. This manipulation may come from preferences that have been deliberately manipulated by the experts or from disobedient experts which do not change their preferences in order to attain their personal goals. A mechanism for a consensus model is suggested in order to minimize the negative effects of these attempts of manipulation. This mechanism is based on associating an importance degree with each expert and to apply a uninorm aggregation operator to obtain the collective preference.

Acknowledges

This work has been supported by the Research Project TIN2006-02121.

Referencias

[1] B. De Baets. *Fuzzy Logic and Intelligent Technologies for Nuclear Science and Industry*. World Scientific, Singapore, 1998.

[2] C. Carlsson, D. Ehrenberg, P. Eklund, M. Fedrizzi, P.Gustafsson, P. Lindholm, G. Merkuryeva, T. Riissanen, and A.G.S.Ventre. Consensus in distributed soft environments. *European Journal of Operational Research*, 61:165–185, 1992.

[3] L. Dombi. *A General Framework for the Utility-Based and Outranking Methods*, pages 202–208. in: Fuzzy Logic and Soft Computing. World Scientific, 1995.

[4] J.C. Fodor, R.R. Yager, and A. Rybalov. Structure of uni-norms. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5:411–427, 1997.

[5] J.L. Garcia-Lapresta. Favoring consensus and penalizing disagreement in group decision making. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Forthcoming.

[6] F. Herrera, E. Herrera-Viedma, and J.L. Verdegay. A sequential selection process in group decision making with linguistic assessment. *Information Sciences*, 85:223–239, 1995.

[7] E. Herrera-Viedma, F. Herrera, and F. Chiclana. A consensus model for multiperson decision making with different preference structures. *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans*, 32(3):394–402, 2002.

[8] E. Herrera-Viedma, L. Martínez, F. Mata, and F. Chiclana. A consensus support system model for group decision-making problems with multigranular linguistic preference relations. *IEEE Transactions on Fuzzy Systems*, 13(5):644–658, 2005.

[9] J. Kacprzyk, M. Fedrizzi, and H.Nurmi. *Consensus Under Fuzziness*. Kluwer Academic Publishers, 1997.

[10] N.I. Karacapilidis and C.P. Pappis. A framework for group decision making support systems: Combining AI tools and OR techniques. *European Journal of Operational Research*, 103:373–388, 1997.

[11] J.S. Kelly. Almost all social choice functions are highly manipulable, but a few aren't. *Social Choice and Welfare*, 10:161–175, 1993.

[12] F. Mata, L. Martínez, E. Herrera-Viedma, and F. Herrera. A penalized consensus model for group decision making problems with multigranular linguistic information. In *Proceedings of EUROFUSE Workshop on New Trends in Preference Modeling (EUROFUSE'07)*, pages 23–29, Jaen (Spain), April 2007.

[13] D.A. Smith. Manipulability of commond social choice functions. *Social Choice and Welfare*, 16:639–661, 1999.

[14] T. Tanino. *On Group Decision Making Under Fuzzy Preferences*, pages 172–185. in: Multiperson Decision Making Using Fuzzy Sets and Possibility Theory. Kluwer Academic Publishers, 1990.

[15] R.R. Yager. Uninorm aggregation operators. *Fuzzy Sets and System*, 80:111–120, 1996.

[16] R.R. Yager. Penalizing strategic preference manipulation in multi-agent decision making. *IEEE Transactions of Fuzzy Systems*, 9(3):393–403, 2001.

[17] R.R. Yager. Defending against strategic manipulation in uninorm-based multi-agent decision making. *European Journal of Operational Research*, 141:217–232, 2002.

[18] R.R. Yager. *Using Importances in Group Preference Aggregation to Block Strategic Manipulation*, pages 177–191. in: Studies in Fuzziness and Soft Computing, Aggregation operators. Physica-Verlag Heidelberg, 2002.

Sistemas Basados en Reglas Difusas en Clasificación: Nuevos Retos

Alberto Fernández Hilario
Departamento de C.C.I.A.
E.T.S. de Ingeniería en Informática y Telecomunicaciones
Universidad de Granada
E-mail: alberto@decsai.ugr.es

María José del Jesus
Departamento de Informática
Escuela Politécnica Superior
Universidad de Jaén
E-mail: mjjesus@ujaen.es

Francisco Herrera
Departamento de C.C.I.A.
E.T.S. de Ingeniería en Informática y Telecomunicaciones
Universidad de Granada
E-mail:herrera@decsai.ugr.es

Resumen

Los Sistemas de Clasificación Basados en Reglas Difusas (SCBRDs) son modelos de clasificación que utilizan reglas difusas para representar el conocimiento. Los SCBRDs se encuentran muy extendidos en la actualidad, con numerosas aplicaciones y estudios de su comportamiento y efectividad.

En este trabajo presentamos algunas líneas de investigación que entendemos son interesantes y que pueden ser consideradas como nuevos retos en el desarrollo de estudios en el ámbito de los SCBRDs.

Palabras Clave: Sistemas de Clasificación Basados en Reglas Difusas, Métodos de Inferencia, Clasificación No Balanceada, Alta Dimensionalidad, Medidas de Complejidad.

1. INTRODUCCIÓN

Los SCBRDs [21] son una herramienta ampliamente utilizada en el ámbito del aprendizaje automático puesto que permiten la incorporación de toda la información disponible en el modelado de sistemas, tanto la que proviene de expertos como la que tiene su origen en medidas empíricas y modelos matemáticos, haciendo posible el tratamiento de información con incertidumbre y permitiendo la representación del conocimiento de una forma comprensible para los usuarios del sistema.

En la actualidad podemos encontrar múltiples estudios en el ámbito de los SCBRDs desde diferentes perspectivas: estudios sobre sus componentes, métodos de aprendizaje de SCBRDs y aplicaciones. En este trabajo haremos una breve introducción a estos tres aspectos, y analizaremos algunas líneas de investigación que pueden ser consideradas como nuevos retos en el ámbito de los SCBRDs.

Nos centraremos en los siguientes temas, algunos de ellos con estudios recientes, pero que todavía tienen un amplio recorrido de cara a disponer de modelos de calidad en esos ámbitos: clasificación con clases no balanceadas, problemas de clasificación con alta dimensionalidad y caracterización de los problemas de clasificación mediante medidas de complejidad y comportamiento de los SCBRDs con estas medidas. Siendo importante el problema de la interpretabilidad y el equilibrio entre interpretabilidad y precisión, no lo abordaremos porque otro trabajo de la presente sesión especial está dedicado a este tema.

El presente trabajo se organiza de la siguiente forma: En la Sección 2 haremos una introducción a los SCBRDs, y los estudios recientes sobre sus componentes, métodos de aprendizaje y aplicaciones. En la Sección 3 abordaremos los tres problemas mencionados como retos en el ámbito de los SCBRDs. Finalmente, en la Sección 4 comentaremos las conclusiones de nuestro trabajo.

2. ESTUDIOS RECIENTES SOBRE SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS DIFUSAS

En esta sección definiremos en primer lugar qué son los SCBRDs y cuáles son sus componentes. A continuación trataremos de realizar un rápido repaso a la evolución en el ámbito de la investigación respecto a este tipo de modelos y citaremos las principales áreas de aplicación de los SCBRDs. Por último expondremos algunos estudios recientes sobre las componentes de los SCBRDs y nombraremos varios métodos de aprendizaje propuestos en la literatura especializada.

2.1. Sistemas de Clasificación Basados en Reglas Difusas

Un SCBRD está compuesto por una Base de Conocimiento (BC) y un Método de Razonamiento Difuso (MRD) que, utilizando la información de la BC, determina una clase para cualquier patrón de datos admisible que llegue al sistema. La potencia del razonamiento aproximado reside en la posibilidad de obtención de un resultado (una clasificación) incluso cuando no tengamos compatibilidad exacta (con grado 1) entre el ejemplo y el antecedente de las reglas.

La BC está formada por dos componentes:

- La *Base de Datos*, que contiene la definición de los conjuntos difusos asociados a los términos lingüísticos utilizados en la Base de Reglas.
- La *Base de Reglas*, formada por un conjunto de reglas de clasificación

$$R = \{R_1, ..., R_L\} \tag{1}$$

Es muy común el uso de reglas difusas con una clase en el consecuente y un peso asociado a dicha regla:

$$R_k: \quad \text{Si } X_1 \text{ es } A_1^k \text{ y } \ldots \text{ y } X_N \text{ es } A_N^k \text{ entonces } Y \text{ es } C_j \text{ con peso } P_k \tag{2}$$

El MRD es un procedimiento de inferencia que utiliza la información de la BC para predecir una clase ante un ejemplo no clasificado. Tradicionalmente en la literatura especializada se ha utilizado el MRD del máximo, también denominado MRD clásico o de la regla ganadora, que considera la clase indicada por una sola regla teniendo en cuenta el grado de asociación del consecuente de la regla sobre el ejemplo. Otros MRDs que combinan la información aportada por todas las reglas que representan el conocimiento de la zona a la que pertenece el ejemplo se estudian en [7].

A continuación presentamos el modelo general de razonamiento difuso que combina la información proporcionada por las reglas difusas compatibles con el ejemplo.

En el proceso de clasificación del ejemplo $e = (e_1, \ldots, e_N)$, los pasos del modelo general de un MRD son los siguientes:

1. Calcular el grado de emparejamiento del ejemplo con el antecedente de las reglas.
2. Calcular el grado de asociación del ejemplo a la clase consecuente de cada regla mediante una función de ponderación entre el grado de emparejamiento y el grado de certeza de la regla con la clase asociada.
3. Determinar el grado de asociación del ejemplo con las distintas clases.
4. Clasificación. Aplicaremos una función de decisión F sobre el grado de asociación del ejemplo con las clases que determinará, en base al criterio del máximo, la etiqueta de clase c a la que corresponda el mayor valor.

2.2. Evolución de las Publicaciones en Revistas y Áreas de Aplicación

Nuestra intención es mostrar una reprospectiva del impacto de los SCBRD. Utilizamos como herramienta el "ISI Web of Science[1]" que proporciona acceso a información multidisciplinar para cerca de 8.700 revistas de investigación más prestigiosas y de mayor impacto. Esta herramienta permite realizar búsquedas avanzadas sobre diversos campos, en concreto consideramos la siguiente consulta (con fecha Abril de 2008):

TS=(("fuzzy rule*" OR "fuzzy system*") AND ("classification*" OR "pattern recognition")) Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, IC, CCR-EXPANDED [back to 1840].

En la Figura 1 se representa el resultado de la consulta anterior, en la que se observa la evolución durante cada año de las 643 publicaciones encontradas.

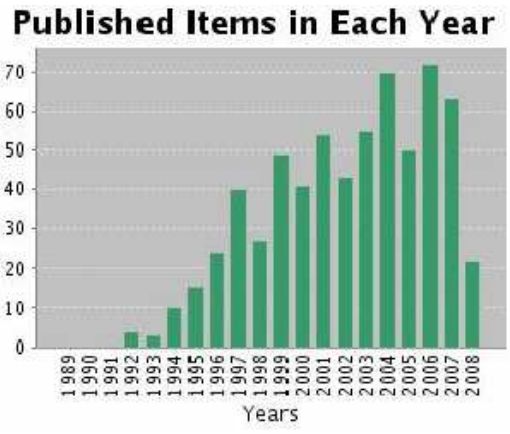


Figura 1: Número de artículos sobre SCBRDs indexados en el "ISI Web of Science"

Observamos un número creciente de publicaciones con más de 50 artículos por año en los últimos 5 años. Estos datos nos permiten señalar que el campo de los SCBRDs ha alcanzado un estado de madurez en la última

[1] http://scientific.thomson.com/products/wos/

década, con una amplia comunidad de investigadores trabajando en SCBRDs.

Debemos subrayar que los SCBRDs proporcionan un modelo altamente interpretable para el usuario final. Por ello se han empleado en muchas aplicaciones reales, como detección de intrusiones [46], estimación de la nubosidad a través de imágenes de satélite [16], procesamiento de imágenes [45], medicina [1, 47], etc.

Continuando con la consulta anterior en el ISI, la Figura 2 muestra la clasificación de las revistas en las áreas de fundamentos en las que hay mayor número de trabajos, donde las áreas de Informática, Ingeniería y Matemáticas muestran una clara supremacía.

Field: Subject Area	Record Count	% of 643	Bar Chart
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	279	43.3904 %	
ENGINEERING, ELECTRICAL & ELECTRONIC	164	25.5054 %	
COMPUTER SCIENCE, THEORY & METHODS	133	20.6843 %	
MATHEMATICS, APPLIED	63	9.7978 %	
STATISTICS & PROBABILITY	62	9.6423 %	
COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS	55	8.5537 %	
AUTOMATION & CONTROL SYSTEMS	51	7.9316 %	
COMPUTER SCIENCE, CYBERNETICS	42	6.5319 %	
COMPUTER SCIENCE, INFORMATION SYSTEMS	33	5.1322 %	
ENGINEERING, BIOMEDICAL	20	3.1104 %	
OPERATIONS RESEARCH & MANAGEMENT SCIENCE	20	3.1104 %	
MEDICAL INFORMATICS	19	2.9549 %	
COMPUTER SCIENCE, HARDWARE & ARCHITECTURE	15	2.3328 %	
ENGINEERING, MULTIDISCIPLINARY	15	2.3328 %	

Figura 2: Clasificación de publicaciones por área de investigación ("ISI Web of Science")

En la Figura 3 mostramos cómo se distribuyen los trabajos en diferentes areas de aplicación.

Por último, en la Tabla 1 realizamos un resumen de la distribución de los trabajos en algunos temas de aplicación, donde hemos ordenado los resultados por el número de publicaciones encontradas.

Cuadro 1: Distribución de las publicaciones sobre SCBRDs según el área de aplicación asociada

Área de Aplicación	**#Publicaciones**
Diagnosis clínica	68 trabajos
Procesamiento de Imágenes	64 trabajos
Medicina	26 trabajos
Visión Artificial	20 trabajos
Química	10 trabajos
Energía	10 trabajos

Como ya hemos mostrado, son muchos los campos en los que los SCBRDs pueden ser aplicables, debido en gran parte, como ya apuntabamos anteriormente, a la facilidad para recoger la información del sistema y la alta interpretabilidad hacia el usuario final.

ENVIRONMENTAL SCIENCES	13	2.0218 %
INSTRUMENTS & INSTRUMENTATION	12	1.8663 %
COMPUTER SCIENCE, SOFTWARE ENGINEERING	10	1.5552 %
ENGINEERING, CIVIL	9	1.3997 %
GEOSCIENCES, MULTIDISCIPLINARY	9	1.3997 %
REMOTE SENSING	9	1.3997 %
CHEMISTRY, ANALYTICAL	8	1.2442 %
ENGINEERING, MECHANICAL	8	1.2442 %
IMAGING SCIENCE & PHOTOGRAPHIC TECHNOLOGY	8	1.2442 %
ROBOTICS	7	1.0886 %
TELECOMMUNICATIONS	7	1.0886 %
WATER RESOURCES	7	1.0886 %
ECOLOGY	6	0.9331 %
GEOCHEMISTRY & GEOPHYSICS	5	0.7776 %
MANAGEMENT	5	0.7776 %
OPTICS	5	0.7776 %
ENERGY & FUELS	4	0.6221 %
ENGINEERING, CHEMICAL	4	0.6221 %
HEALTH CARE SCIENCES & SERVICES	4	0.6221 %
MATHEMATICAL & COMPUTATIONAL BIOLOGY	4	0.6221 %
METEOROLOGY & ATMOSPHERIC SCIENCES	4	0.6221 %
PHYSICS, APPLIED	4	0.6221 %

Figura 3: Clasificación de publicaciones por área de aplicación ("ISI Web of Science")

2.3. Algunos Estudios sobre Componentes y Aprendizaje

Desde los primeros trabajos sobre SCBRDs, los investigadores siempre han buscado mejorar el rendimiento de estos modelos mediante el estudio de las componentes o actualizando los métodos de aprendizaje.

Sobre las componentes del modelo, destaca la importancia del peso de las reglas [20], por lo que muchos estudios han buscado nuevas heurísticas para definir dicho peso. Ishibuchi en [23] propone diversas metodologías para el cálculo del peso de las reglas, y actualmente Zolghadri ha realizado distintas propuestas también en este campo [31, 52, 53]. Respecto a otras componentes, Ghosh en [15] define un modelo basado en funciones de pertenencia tipo π y un operador de agregación producto.

En el caso de los métodos de aprendizaje, durante los últimos años se han realizado distintas metodologías incluyendo propuestas basadas en redes neuronales [33, 27], enfriamiento simulado [42, 34], algoritmos genéticos [24, 51], sistemas inmunes [26] o basados en reglas de asociación [37].

De este modo, comprobamos que los SCBRDs están en constante evolución. Es importante destacar que es posible con las nuevas herramientas disponibles hoy en día, la creación de nuevos modelos de aprendizaje más robustos y una mejor optimización de todas las componentes del modelo difuso.

3. NUEVOS RETOS PARA SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS DIFUSAS

En esta sección presentaremos algunas líneas de investigación que estimamos interesantes y por tanto podrían ser consideradas como nuevos retos en el ámbito de los SCBRDs: clasificación con clases no balanceadas, problemas con alta dimensionalidad y el comportamiento de los SCBRDs frente a las medidas de complejidad en conjuntos de datos.

3.1. Clasificación No Balanceada

El aprendizaje sobre conjuntos de datos no balanceados es un tema importante que ha aparecido recientemente en la comunidad del aprendizaje automático [6]. Cuando se trabaja con conjuntos de datos no balanceados, una o más clases están representadas por un gran número de ejemplos, mientras que el resto se representan por unos pocos.

El problema de los conjuntos de datos no balanceados es extremadamente importante puesto que está implícito en la mayoría de aplicaciones reales, como detección de fraudes [11], gestión de riesgos [19], y especialmente en diagnóstico médico [32].

En clasificación, este problema provoca una tendencia hacia una menor sensibilidad para detectar los ejemplos de la clase minoritaria durante el entrenamiento de los clasificadores. La mayor desventaja en los conjuntos de datos no balanceados es el solapamiento entre los ejemplos de la clase minoritaria (llamada comunmente clase positiva) y mayoritaria (clase negativa), dada la dificultad de la mayoría de algoritmos de aprendizaje para detectar los llamados "pequeños datos disjuntos" [25]. Este hecho se muestra gráficamente en la Figura 4.

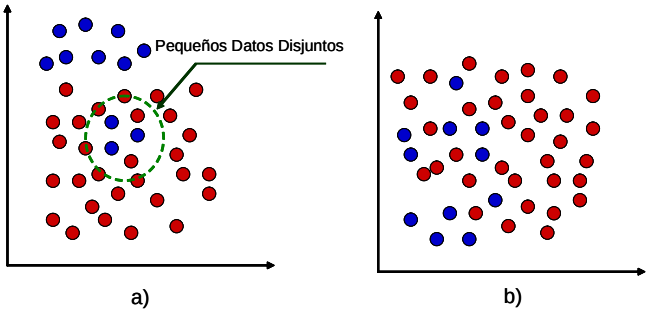


Figura 4: Ejemplo del no balanceo: a) pequeños datos disjuntos b) solapamiento entre las clases.

Con respecto a los SCBRDs, tan solo hay unos pocos trabajos en la literatura especializada que estudian su uso sobre conjuntos de datos no balanceados. Algunos de estos estudios incluyen sistemas difusos aproximativos [48, 49], mientras que otros presentan tres propuestas de aprendizaje diferentes: uno usando árboles de decisión difusos [8], otro se basa en la extración de reglas difusas utilizando grafos difusos y algoritmos genéticos [44], y el último se basa en un procedimiento específico de generación de reglas lingüísticas para datos no balanceados, incorporando pesos de coste en el cómputo del grado de confianza y soporte de la reglas, y llamado E-Algorithm [50]. Nuestro trabajo en el ámbito de los conjuntos de datos no balanceados se ha basado en el análisis de la cooperación entre algunos métodos de preprocesamiento y los SCBRDs, además del estudio de los distintos componentes del modelo difuso [12, 13, 14].

Debemos enfatizar que existen líneas de investigación abiertas en este campo, por ejemplo la caracterización de un modelo de aprendizaje específico que tenga en cuenta las características de este problema, el estudio en los distintos niveles de no balanceo, puesto que los SCBRDs tienen un comportamiento distinto dependiendo del porcentaje de no balanceo [14], o el análisis de los pequeños datos disjuntos.

3.2. Problemas con Alta Dimensionalidad

Uno de los retos de la minería de datos, es diseñar algoritmos de aprendizaje para obtener información a partir de grandes conjuntos o bases de datos [39]. Esta necesidad surge de la relación entre el rendimiento del modelo y el tamaño del conjunto de entrenamiento. Además, hay que evitar el problema del sobreaprendizaje cuando existe un gran número de variables en problemas con pocos ejemplos.

Los modelos de aprendizaje de SCBRDs tienen problemas de escalabilidad cuando trabajan con problemas de alta dimensionalidad. Esto es debido al crecimiento exponencial del número de reglas cuando aumenta el número de variables de entrada del problema o cuando el número de patrones es también elevado. De este modo se producen dos problemas principales:

- Por un lado el proceso de aprendizaje se vuelve mucho más complejo.
- Por otro lado se genera una pérdida importante de la interpretabilidad de los SCBRDs.

Para resolver el problema de la alta dimensionalidad en el número de atributos se puede emplear un proceso de selección de características, tanto como una etapa implícita del modelo [4, 5, 28], como con la inclusión de etiquetas tipo "Don't Care" en las variables difusas y/o limitando el número de etiquetas difusas en el an-

tecedente de las reglas [22]. Asimismo, es muy usual encontrar este tipo de procedimientos en aplicaciones reales de data mining o reconocimiento de patrones con multiples caracteristicas [38, 46], puesto que no solo se reduce el tiempo de entrenamiento e inferencia, si no que también se suelen obtener un mejor rendimiento del clasificador.

En el caso de conjuntos de datos con un gran número de ejemplos, se podría proceder mediante un preprocesamiento basado en selección de instancias [3] o el uso de particiones distribuidas de datos [36].

En este sentido, se debe enfocar el estudio hacia SCBRDs más compactos y con mayor robustez frente al problema de la alta dimensionalidad, buscando técnicas que permitan, con un menor número de reglas, adaptarse con un comportamiento adecuado a este tipo de conjuntos, tomando en consideración las propuestas mencionadas en esta sección.

3.3. Sistemas de Clasificación Basados en Reglas Difusas y Medidas de Complejidad

Las medidas de complejidad permiten caracterizar la dificultad de un conjunto de datos en problemas de clasificación. Pueden ser útiles para analizar la dificultad de los mismos para diferentes algoritmos de aprendizaje [43]. En concreto, los problemas de clasificación pueden ser difíciles por tres razones diferentes:

- Ciertos problemas son conocidos por tener un error de Bayes no nulo [17]. Algunas clases pueden ser ambiguas intrínsecamente o debido a medidas incorrectas de los atributos.
- Algunos problemas pueden presentar límites de decisión complejos, por lo que no se puede ofrecer una descripción compacta de los mismos [41].
- Muestras de tamaño reducido y la dispersión inducida por la alta dimensionalidad afectan a las reglas [30, 40].

Un breve repaso al trabajo realizado es el siguiente: en [18] Ho y Basu definen medidas de complejidad para conjuntos de datos de dos clases. Singh en [43] ofrece una revisión de medidas de complejidad y propone dos nuevas medidas. Dong y Kothaire en [9] proponen un algoritmo de selección de características basado en una medida de complejidad definida por Ho y Basu. Bernadó y Ho en [10] investigan el dominio de competencia de XCS por medio de una metodología que caracteriza la complejidad de un problema de clasificación mediante un conjunto de descriptores geométricos. En [29], Li y otros analizan algunos árboles de decisión omnivariados usando la medida de complejidad basada en la densidad de datos propuesta por Ho y Basu. Baumgartner y Somorjai en [2] definine medidas específicas para clasificadores lineales regularizados, usando las medidas de Ho y Basu como referencia. Mollineda y otros en [35] extienden algunas de las definiciones de medidas de Ho y Basu para problemas con dos o más clases, analizando estas medidas generalizadas en dos problemas clásicos de Selección de Prototipos. Finalmente, Sánchez y otros en [41] analizan el efecto de la complejidad de datos en el clasificador de vecino más cercano.

Una medida de complejidad comunmente estudiada es la *razón discriminante de Fisher*. Se trata de una medida geométrica de solapamiento, propuesta por Ho y Basu y calcula la separación entre dos clases de acuerdo a una característica en concreto. Compara la diferencia entre la media de las clases con la suma de la varianza de las clases. La razón discriminante de Fisher se define como:

$$f = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \tag{3}$$

donde μ_1, μ_2, σ_1 y σ_2 son las medidas y varianzas de las dos clases respectivamente. Valores pequeños indican que las clases tienen un alto grado de solapamiento. Las Figuras 5 a 8 muestran un ejemplo ilustrativo generado artificialmente con 2 variables en el rango [0,0; 1,0] y dos clases como ejemplo.

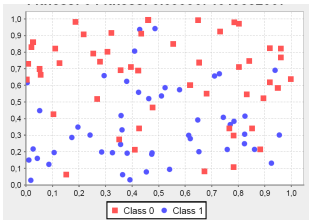


Figura 5: $F1 = 0{,}6994$

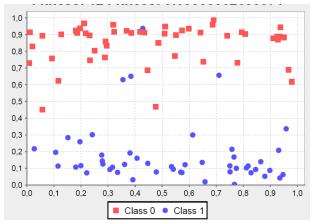


Figura 6: $F1 = 9{,}69$

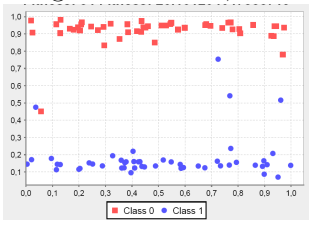


Figura 7: $F1 = 26{,}16$

Figura 8: $F1 = 48{,}65$

Este tipo de análisis puede ser aplicado a los SCBRDs para analizar su comportamiento con las diferentes medidas y diseñar modelos de aprendizaje de acuerdo al tipo de problemas.

4. CONCLUSIONES

En este trabajo hemos presentado algunas líneas de investigación que pueden ser consideradas como nuevos retos en el desarrollo de estudios en el campo de los SCBRDSs, como son la clasificación en el ámbito de los datos no balanceados, el problema de la alta dimensionalidad y el estudio de las medidas de complejidad de los datos.

Agradecimientos

Este trabajo de investigación ha sido posible gracias a la subvención del proyecto TIN2005- 08386-C05-01 y TIN-2005-08386-C05-03.

Referencias

[1] M. R. Akbarzadeh-Totonchi and M. Moshtagh-Khorasani. A hierarchical fuzzy rule-based approach to aphasia diagnosis. *Journal of Biomedical Informatics*, 40(5):465–475, 2007.

[2] R. Baumgartner and R. Somorjai. Data complexity assessment in undersampled classification. *Pattern Recognition Letters*, 27:1383–1389, 2006.

[3] J. R. Cano, F. Herrera, and M. Lozano. Evolutionary stratified training set selection for extracting classification rules with trade-off precision-interpretability. *Data and Knowledge Engineering*, 60:90–108, 2007.

[4] J. Casillas, O. Cordón, M. J. del Jesus, and F. Herrera. Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional problems. *Information Sciences*, 136(1):135–157, 2001.

[5] D. Chakraborty and N. R. Pal. A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification. *IEEE Transactions on Neural Networks*, 15(1):110–123, 2004.

[6] N. V. Chawla, N. Japkowicz, and A. Kolcz. Editorial: special issue on learning from imbalanced data sets. *SIGKDD Explorations*, 6(1):1–6, 2004.

[7] O. Cordón, M. J. del Jesus, and F. Herrera. A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, 20(1):21–45, 1999.

[8] K. Crockett, Z. Bandar, and J. O'Shea. On producing balanced fuzzy decision tree classifiers. In *IEEE International Conference on Fuzzy Systems*, pages 1756–1762, 2006.

[9] M. Dong and R. Kothari. Feature subset selection using a new definition of classifiability. *Pattern Recognition Letters*, 24(9-10):1215–1225, 2003.

[10] E.Bernadó-Mansilla and T.K.Ho. Domain of competence of XCS classifier system in complexity measurement space. *IEEE Transactions on Evolutionary Computation*, 9(1):82–104, 2005.

[11] T. Fawcett and F. J. Provost. Adaptive fraud detection. *Data Mining and Knowledge Discovery*, 1(3):291–316, 1997.

[12] A. Fernández, S. García, M. J. del Jesus, and F. Herrera. An analysis of the rule weights and fuzzy reasoning methods for linguistic rule based classification systems applied to problems with highly imbalanced data sets. In *International Workshop on Fuzzy Logic and Applications (WILF07)*, volume 4578 of *Lecture Notes on Computer Science*, pages 170–179. Springer-Verlag, 2007.

[13] A. Fernández, S. García, M. J. del Jesus, and F. Herrera. A study on the use of the fuzzy reasoning method based on the winning rule vs. voting procedure for classification with imbalanced data sets. In *9th International Work-Conference on Artificial Neural Networks (IWANN07)*, volume 4507 of *Lecture Notes on Computer Science*, pages 375–382. Springer-Verlag, 2007.

[14] A. Fernández, S. García, M. J. del Jesus, and F. Herrera. A study of the behaviour of linguistic fuzzy rule based classification systems in the framework of imbalanced data-sets. *Fuzzy Sets and Systems*, In Press (2008), doi:10.1016/j.fss.2007.12.023, 2008.

[15] A. Ghosh, S. K. Meher, and B. U. Shankar. A novel fuzzy classifier based on product aggregation operator. *Pattern Recognition*, 41(3):961–971, 2008.

[16] A. Ghosh, N. Pal, and J. Das. A fuzzy rule based approach to cloud cover estimation. *Remote Sensing of Environment*, 100(4):531–549, 2006.

[17] T. Ho and H. Baird. Large-scale simulation studies in image pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(10):1067–1079, 1997.

[18] T. Ho and M. Basu. Complexity measures of supervised classification problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):289–300, 2002.

[19] Y. M. Huang, C. M. Hung, and H. C. Jiau. Evaluation of neural networks and data mining methods on a credit assessment task for class imbalance problem. *Nonlinear Analysis: Real World Applications*, 7(4):720–747, 2006.

[20] H. Ishibuchi and T. Nakashima. Effect of rule weights in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 9(4):506–515, 2001.

[21] H. Ishibuchi, T. Nakashima, and M. Nii. *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining*. Springer-Verlag, 2004.

[22] H. Ishibuchi and T. Yamamoto. Comparison of heuristic criteria for fuzzy rule selection in classification problems. *Fuzzy Optimization and Decision Making*, 3(2):119–139, 2004.

[23] H. Ishibuchi and T. Yamamoto. Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 13:428–435, 2005.

[24] H. Ishibuchi, T. Yamamoto, and T. Nakashima. Hybridization of fuzzy GBML approaches for pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics- Part B: Cybernetics*, 35(2):359–365, 2005.

[25] T. Jo and N. Japkowicz. Class imbalances versus small disjuncts. *SIGKDD Explorations*, 6(1):40–49, 2004.

[26] Z. Lei and L. Ren-hou. Designing of classifiers based on immune principles and fuzzy rules. *Information Sciences*, 178(7):1836–1847, 2008.

[27] G. Leng, T. McGinnity, and G. Prasad. An approach for on-line extraction of fuzzy rules using a self-organising fuzzy neural network. *Fuzzy Sets and Systems*, 150(2):211–243, 2005.

[28] Y. Li and Z.-F. Wu. Fuzzy feature selection based on min-max learning rule and extension matrix. *Pattern Recognition*, 41:217–226, 2008.

[29] Y.-H. Li, M. Dong, and R. Kothari. Classifiability-based omnivariate decision trees. *IEEE Transactions On Neural Networks*, 16(6):1547–1560, 2005.

[30] M. Liwicki and H. Bunke. Handwriting recognition of whiteboard notes - studying the influence of training set size and type. *International Journal of Pattern Recognition and Artificial Intelligence*, 21(1):83–98, 2007.

[31] E. G. Mansoori, M. J. Zolghadri, and S. D. Katebi. A weigthing function for improving fuzzy classification systems performance. *Fuzzy Sets and Systems*, 158(5):583–591, 2007.

[32] M. Mazurowski, P. Habas, J. Zurada, J. Lo, J. Baker, and G. Tourassi. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural Networks*, 21(2-3):427–436, 2008.

[33] S. Mitra and Y. Hayashi. Neuro-fuzzy rule generation: survey in soft computing framework. *IEEE Transactions on Neural Networks*, 11(3):748–768, 2000.

[34] H. Mohamadia, J. Habibia, M. S. Abadeh, and H. Saadi. Data mining with a simulated annealing based fuzzy classification system. *Pattern Recognition*, 41(5):1824–1833, 2008.

[35] R. Mollineda, J. Sánchez, and J. Sotoca. Data characterization for effective prototype selection. In *Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA 2005)*, volume 3523 of *Lecture Notes in Computer Science*, pages 27–34, 2005.

[36] Y. Nojima, I. Kuwajima, and H. Ishibuchi. Data set subdivision for parallel distributed implementation of genetic fuzzy rule selection. In *IEEE International Conference on Fuzzy Systems*, pages 2006–2011, 2007.

[37] F. P. Pach, A. Gyenesei, and J. Abonyi. Compact fuzzy association rule-based classifier. *Expert Systems with Applications*, 34(4):2406–2416, 2008.

[38] J. Paetz and G. Schneider. A neuro-fuzzy approach to virtual screening in molecular bioinformatics. *Fuzzy Sets and Systems*, 152(1):67–82, 2005.

[39] F. Provost and V. Kolluri. A survey of methods for scaling up inductive algorithms. *Data Mining and Knowledge Discovery*, 3(2):131–169, 1999.

[40] X. Qiu and L. Wu. Nearest neighbour discriminant analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(8):1245–1259, 2006.

[41] J. Sánchez, R. Mollineda, and J. Totoca. An analysis of how training data complexity affects the nearest neighbours classifiers. *Pattern Analysis & Applications*, 10:189–201, 2007.

[42] L. Sánchez, I. Couso, and J. Corrales. Combining GP operators with SA search to evolve fuzzy rule based classifiers. *Information Sciences*, 136(1–4):175–191, 2001.

[43] S. Singh. Multiresolution estimates of classification complexity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1534–1539, 2003.

[44] V. Soler, J. Cerquides, J. Sabria, J. Roig, and M. Prim. Imbalanced datasets classification by fuzzy rule extraction and genetic algorithms. In *IEEE International Conference on Data Mining - Workshops*, pages 330–336, 2006.

[45] T.Nakashima, G. Schaefer, Y. Yokota, and H. Ishibuchi. A weighted fuzzy classifier and its application to image processing tasks. *Fuzzy Sets and Systems*, 158:284–294, 2007.

[46] C. Tsang, S. Kwong, and H. Wang. Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection. *Pattern Recognition*, 40(9):2373–2391, 2007.

[47] S. A. Vinterbo, E.-Y. Kim, and L. Ohno-Machado. Small, fuzzy and interpretable gene expression based classifiers. *Bioinformatics*, 21(9):1964–1970, 2005.

[48] S. Visa and A. Ralescu. Fuzzy classifiers for imbalanced, complex classes of varying size. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 393–400, 2004.

[49] S. Visa and A. Ralescu. The effect of imbalanced data class distribution on fuzzy classifiers - experimental study. In *IEEE International Conference on Fuzzy Systems*, pages 749–754, 2005.

[50] L. Xu, M. Y. Chow, and L. S. Taylor. Power distribution fault cause identification with imbalanced data using the data mining-based fuzzy classification E-Algorithm. *IEEE Transactions on Power Systems*, 22(1):164–171, 2007.

[51] E. Zhou and A. Khotanzad. Fuzzy classifier design using genetic algorithms. *Pattern Recognition*, 40(12):3401–3414, 2007.

[52] M. J. Zolghadri and E. G. Mansoori. Weighting fuzzy classification rules using receiver operating characteristics (ROC) analysis. *Information Sciences*, 177(11):2296–2307, 2007.

[53] M. J. Zolghadri and M. Taheri. A proposed method for learning rule weights in fuzzy rule-based classification systems. *Fuzzy Sets And Systems*, 159:449–459, 2008.

Equilibrio entre Interpretabilidad y Precisión en Sistemas Basados en Reglas Difusas: Nuevos retos

José M. Alonso Luis Magdalena

European Centre for Soft Computing
Edificio Científico-Tecnológico
Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Asturias, España
{jose.alonso,luis.magdalena}@softcomputing.es

Resumen

Este artículo revisa la evolución histórica de los procesos de modelado de sistemas basados en reglas difusas. Dichos sistemas se caracterizan por dos propiedades fundamentales: Precisión e Interpretabilidad. Aunque inicialmente se dio prioridad a la Interpretabilidad, esta pronto quedó en segundo plano y la atención se centró en la Precisión. Actualmente el principal desafío consiste en encontrar un equilibrio óptimo entre ambas propiedades. Sin embargo, para ello hay que afrontar una tarea previa: Caracterizar la Interpretabilidad y encontrar un índice apropiado para evaluarla.

Palabras Clave: Interpretabilidad, precisión, equilibrio, índices de medida.

1 INTRODUCCIÓN

La lógica difusa introducida por Zadeh [27] en 1965 es reconocida a nivel mundial por su capacidad para modelar conceptos lingüísticos, así como por su utilización para extracción y representación de conocimiento en el modelado de sistemas. La expresividad semántica de la lógica difusa, mediante el uso de reglas [19] y variables lingüísticas [29], es muy próxima al lenguaje natural empleado por los humanos. Esto favorece la Interpretabilidad del modelo final construido, lo que se considera la principal virtud de los Sistemas Basados en Reglas Difusas (SBRDs) frente a otras técnicas de Soft Computing como por ejemplo las redes neuronales.

El modelado difuso, es decir, el modelado de SBRDs constituye una línea de investigación muy activa desde los inicios de la lógica difusa hasta nuestros días. Entre 1965 y 1990, la atención se centró en construir sistemas formados por pequeños grupos de reglas lingüísticas (fáciles de interpretar) a partir de conocimiento experto. Sin embargo, pronto se vieron las limitaciones de esta clase de sistemas ya que para abordar problemas complejos no es suficiente considerar sólo el conocimiento experto, es necesario completar dicho conocimiento con el extraído a partir de datos, haciendo uso de técnicas de aprendizaje automático. Como consecuencia, entre 1990 y 2000 se pasó a generar SBRDs formados por muchas reglas complejas capaces de alcanzar una gran Precisión. Desgraciadamente, la Interpretabilidad quedó en segundo plano hasta que a partir del año 2000 se planteó un nuevo enfoque, construir sistemas precisos e interpretables que combinen conocimiento experto y conocimiento extraído de forma automática a partir de datos experimentales.

En esa línea y dependiendo de la propiedad que se considere prioritaria existen dos grandes tendencias dentro del modelado difuso. Por un lado, aquellos trabajos que maximizan la Interpretabilidad del modelo y después aplican alguna técnica de optimización para aumentar la Precisión [8]. Por otro lado, aquellos trabajos que defienden la estrategia contraria, primero maximizan la Precisión y después intentan refinar el modelo para mejorar su Interpretabilidad [10]. Según la clasificación hecha en [1], el primer enfoque se denomina modelado difuso lingüístico con mejora de Precisión, y el segundo modelado difuso preciso con mejora de Interpretabilidad.

El modelado difuso admite dos formas de refinamiento: Extensión del proceso de modelado incorporando nuevos algoritmos para aprendizaje de particiones y reglas; extensión de la estructura del modelo con modificadores lingüísticos, reglas con pesos, reglas con excepciones, reglas por defecto, etc.

Cuanto más flexible sea el proceso de modelado mayor es la Precisión que se puede alcanzar, pero a cambio de un mayor coste computacional (demasiados grados de

libertad pueden hacer imposible encontrar el óptimo). Por el contrario, imponer restricciones al proceso de modelado favorece la Interpretabilidad pero reduce significativamente la Precisión máxima que se puede alcanzar. Como resultado, buscar un buen equilibrio entre ambas propiedades es una de las tareas más difíciles del modelado difuso [9].

La búsqueda del equilibrio óptimo entre Precisión e Interpretabilidad se puede plantear como un problema multiobjetivo porque ambas propiedades representan objetivos contradictorios. Tal y como establece el principio de incompatibilidad enunciado por Zadeh [28]: *A medida que la complejidad de un problema crece, disminuye la posibilidad de analizarlo en términos precisos y relevantes de forma simultánea.* Como consecuencia, las técnicas de optimización multiobjetivo han irrumpido con fuerza en el modelado difuso [16]. No obstante, para que dichas técnicas tengan éxito primero es necesario hacer un esfuerzo por caracterizar la Interpretabilidad. La eterna asignatura pendiente del modelado de SBRDs es definir un índice para evaluar dicha propiedad de forma equivalente a como se evalúa la precisión. Por esta razón, en este trabajo estudiaremos el equilibrio entre Interpretabilidad y Precisión en SBRDs profundizando en el estudio de la Interpretabilidad.

El resto del trabajo se ha estructurado de la siguiente forma. En la sección 2 se presentan los principales logros alcanzados hasta ahora. A continuación, la sección 3 presenta los problemas que quedan abiertos, enumerando nuevos retos y apuntando posibles soluciones. Por último, la sección 4 cierra el artículo con unas breves conclusiones.

2 LOGROS CONSEGUIDOS HASTA AHORA

Para estudiar el equilibrio entre Precisión e Interpretabilidad primero hay que caracterizar ambas propiedades por separado.

El estudio de la Precisión de los SBRDs no entraña gran dificultad. En problemas de regresión lo más habitual es usar el error cuadrático medio. En problemas de clasificación se cuenta el número de casos mal clasificados. En ambos casos la medida es objetiva y los índices son universalmente aceptados.

La tarea es mucho más compleja cuando se trata de evaluar la Interpretabilidad de los SBRDs. El principal problema radica en que su evaluación es muy subjetiva, depende mucho de la persona, de sus conocimientos y experiencia previa. Tanto los criterios de evaluación como la heurística seguida varían bastante de unas personas a otras. Ni siquiera existe una definición de Interpretabilidad reconocida por todo el mundo.

Las siguientes subsecciones intentan aclarar el asunto. Primero revisaremos las definiciones existentes en la literatura. A continuación enumeraremos los factores que influyen a la hora de evaluar la Interpretabilidad, y presentaremos las propuestas de índices existentes.

2.1 CARACTERIZACIÓN DE LA INTERPRETABILIDAD

La primera definición formal de Interpretabilidad que se encuentra en la literatura es la formulada por Tarski [26] en 1953. Se trata de una definición matemática en el contexto de la lógica clásica con la que Tarski pretendía sentar las bases para identificar teorías interpretables. En resumen, asumiendo que T y S son dos teorías formales, se dice que T es interpretable en S, si y sólo si es posible pasar de T a S de forma que cada teorema de T se puede traducir y probar en S.

En la literatura especializada en sistemas difusos encontramos una definición muy similar aunque mucho más reciente e informal [20]. En vez de hablar de teorías se hace referencia a dos lenguajes L y L'. El primero es la lógica difusa y sirve para describir el modelo de manera formal. El segundo se utiliza para dar una explicación del modelo a un usuario del mismo, normalmente es el lenguaje natural. Si el sistema es interpretable, el paso de L a L' debe ser inmediato. Evaluar la Interpretabilidad del sistema es equivalente a medir la complejidad de la traducción entre L y L'.

Bodenhofer y Bauer realizaron una definición axiomática más formal [4, 5]: *La Interpretabilidad de un sistema es la posibilidad de estimar su comportamiento a partir de la lectura y comprensión de la descripción de su base de reglas.* Puesto que la legibilidad de las reglas depende a su vez de la legibilidad de las expresiones lingüísticas que estas incluyen, Bodenhofer considera que el análisis se debe realizar a nivel de particiones garantizando que estas respeten las relaciones de orden e inclusión entre conjuntos difusos. En conclusión, la Interpretabilidad de las particiones es un prerrequisito para que un SBRD sea interpretable.

2.1.1 Interpretabilidad de variables

Aunque el uso de variables lingüísticas favorece la Interpretabilidad, no la garantiza. Por ello es necesario imponer ciertas restricciones en la fase de diseño. Este es un tema que ha sido estudiado en profundidad por muchos autores, y por suerte existe un tipo especial de particiones difusas, las denominadas particiones difusas fuertes [23], que satisfacen todas las restricciones exigidas para que el sistema resultante sea interpretable [11, 14]:

- Definir un número moderado de etiquetas, términos lingüísticos, por variable. Los psicólogos establecen 7 ± 2 como un límite en el número de conceptos simultáneos que admite la capacidad de procesamiento humano [21].
- Asociar de forma clara un término lingüístico con cada función de pertenencia, de manera que los conjuntos difusos se distingan sin ambigüedades.
- Utilizar funciones de pertenencia que cubran todo el universo de discurso en que se define la variable, y que cumplan la propiedad de normalidad.
- Trabajar con conjuntos difusos convexos [27].

El uso de particiones difusas fuertes garantiza la Interpretabilidad a nivel de variables. Sin embargo, supone una gran restricción desde el punto de vista de la Precisión. Fazendeiro y Oliveira [12] demostraron como es posible incrementar la Precisión de un SBRD prescindiendo del uso de particiones fuertes. Cuanto más se degrada la partición mayor es la Precisión que se puede alcanzar, pero la Interpretabilidad disminuye drásticamente.

2.1.2 Interpretabilidad de reglas

Hay que destacar que la legibilidad de la base de reglas es mayor si se define una semántica global previa a la generación de reglas, es decir, si todas las reglas comparten los mismos términos lingüísticos definidos por los mismos conjuntos difusos, lo que permite comparar las reglas directamente a nivel lingüístico. Una vez estudiada la Interpretabilidad de las variables queda discutir la Interpretabilidad de las reglas. Existen dos tipos de reglas difusas:

1. Mamdani [19]: La conclusión es un conjunto difuso.
2. Takagi-Sugeno [25]: La conclusión es un valor numérico obtenido como una combinación de los valores de las entradas.

El uso de reglas tipo Mamdani favorece la Interpretabilidad porque son reglas lingüísticas de la forma:

$$\textbf{Si} \quad \underbrace{\underbrace{X_a \text{ es } A_a^i}_{\textit{Proposición } P_a} \quad \textbf{Y} \ \ldots \ \textbf{Y} \quad \underbrace{X_z \text{ es } A_z^j}_{\textit{Proposición } P_z}}_{\textit{Premisa}}$$

$$\textbf{Entonces} \quad \underbrace{Y \text{ es } C^n}_{\textit{Conclusión}}$$

Una base de reglas construida con este tipo de reglas es una forma normal disyuntiva, es decir, una disyunción de conjunciones. Las premisas están compuestas por proposiciones lingüísticas que relacionan duplas (variable, término lingüístico) de la forma X_a *es* A_a^i donde X_a es el nombre de la variable a, y A_a^i representa la etiqueta i de esa variable. Además, el uso de términos compuestos formados por combinaciones de etiquetas primarias a través de disyunciones y/o negaciones ayuda a construir reglas más compactas y parecidas a las definidas por expertos. El uso de modificadores lingüísticos (*más o menos*, *muy*, *ligeramente*, etc.), operadores que modifican las funciones de pertenencia de los conjuntos difusos asociados a las etiquetas, permite aumentar la Precisión del modelo pero puede empeorar su legibilidad [1]. Existen otras formas de aumentar la Precisión del modelo. Una alternativa es el uso de técnicas de optimización que ajustan las particiones borrosas para un conjunto fijo de reglas [7], o que hacen una selección de reglas [2]. Otra posibilidad es añadir pesos a las reglas [17].

2.2 EVALUACIÓN DE LA INTERPRETABILIDAD

El análisis previo nos permite concluir que los sistemas difusos más interpretables son aquellos que consideran reglas lingüísticas puras (sin pesos ni excepciones), con semántica global y particiones difusas fuertes. No obstante, no todos los sistemas alcanzan el mismo grado de Interpretabilidad. Esta se puede evaluar según dos puntos de vista:

1. **Descripción global**. Para entender un SBRD en su conjunto, éste debe tener unas dimensiones reducidas. El análisis se debe hacer a dos niveles. A nivel de variables, si se trabaja con particiones fuertes, el único parámetro que se debe considerar es el número de etiquetas. A nivel de reglas, hay que tener en cuenta el número de reglas, pero también su complejidad, es decir, el número de variables utilizadas por regla y la complejidad de las proposiciones lingüísticas que forman las premisas.
2. **Explicación puntual**. Para entender el comportamiento de un SBRD ante una situación concreta hay que analizar los operadores difusos utilizados, así como el número de reglas que pueden activarse simultáneamente dado un vector de entrada. Entender la salida de un SBRD implica comprender los mecanismos de inferencia, agregación y defusificación.

Ambos puntos de vista son contradictorios. El primero fomenta la utilización de reglas generales y compactas que sólo utilizan subconjuntos de las variables de entrada. El segundo lleva a preferir reglas completas,

mucho más específicas, que utilizan todas las entradas porque cuanto más generales sean las reglas, mayor número de ellas se pueden activar a la vez.

Hasta ahora, todos los intentos por definir un índice para medir la Interpretabilidad se han centrado en el primer punto de vista, a priori el más asequible.

El primero en proponer un índice para evaluar la Interpretabilidad fue Nauck [22]. Se trata de una fórmula matemática que combina la complejidad del clasificador (número de clases dividido por el número total de premisas), la complejidad de las particiones (número de etiquetas definidas por variable), y el grado de cobertura de la partición (igual a 1 para particiones fuertes). Desgraciadamente, este trabajo no ha tenido continuidad y no hemos podido encontrar otros trabajos que demuestren la utilidad del índice propuesto. Es un índice muy sencillo que no tiene en cuenta explícitamente el número de reglas, ni la complejidad de las mismas.

Una propuesta posterior, inspirada en el índice de Nauck pero mucho más ambiciosa plantea la posibilidad de medir la Interpretabilidad mediante un sistema experto difuso [3] que hace una evaluación a nivel de particiones y reglas. Asumiendo que el SBRD es interpretable, es decir, incluye particiones fuertes, semántica global y reglas lingüísticas, el sistema difuso propuesto analiza el grado de Interpretabilidad alcanzado. La principal ventaja de este índice es que permite analizar los factores que influyen en el resultado de la medida, y se puede adaptar de forma relativamente sencilla a las preferencias del usuario. Puesto que la medida de Interpretabilidad es muy subjetiva, es inútil utilizar una fórmula matemática.

Alternativamente, Fazendeiro y Oliveira [12] propusieron un índice que podría combinarse con el anterior. Se trata de un índice numérico para evaluar la Interpretabilidad sólo a nivel de particiones difusas. El valor máximo se alcanza cuando la partición es fuerte, pero a medida que esta propiedad se degrada la Interpretabilidad disminuye. Existen otras propuestas similares [6, 18] que definen índices para medir la distinguibilidad entre conjuntos difusos, la cobertura de la partición, etc.

3 PROBLEMAS ABIERTOS Y NUEVOS RETOS

El concepto de Interpretabilidad aparece en la literatura bajo distintos nombres (legibilidad, transparencia, comprensibilidad, etc.) pero siempre implica una interacción con personas. Puesto que los seres humanos pensamos y nos comunicamos a través de palabras, parece lógico pensar que la única forma de obtener un índice de Interpretabilidad que sea verdaderamente útil pasa por dar el salto a la computación con palabras. Deberíamos seguir el consejo de Zadeh y dar el paso de la computación con números a la computación con palabras, del análisis de medidas numéricas al análisis de percepciones [30].

La propuesta inicial de Zadeh es del año 1975 cuando puso de manifiesto la necesidad de utilizar variables lingüísticas para superar la incapacidad de los sistemas tradicionales para resolver problemas donde el comportamiento del sistema modelado está fuertemente influenciado por la opinión de personas, percepciones y emociones [29]: *Para hacer algunas evaluaciones significativas, habría que abandonar los estándares de rigor y precisión característicos de los análisis matemáticos clásicos, y ser más tolerantes con nuevos enfoques que son aproximados por naturaleza.*

A pesar de los intentos hechos recientemente, todavía no existe un índice de Interpretabilidad universalmente aceptado. La caracterización y evaluación de la Interpretabilidad de los SBRDs sigue siendo un problema abierto. Por lo tanto, podemos establecer como un primer gran reto de futuro la definición de un índice para medir la Interpretabilidad, haciendo uso de la nueva teoría de computación con palabras propuesta por Zadeh.

También es interesante explorar otros campos de la ciencia que pueden estar relacionados con esta tarea, como por ejemplo las investigaciones realizadas en torno a las técnicas de procesamiento de lenguaje natural. Como sabemos, la expresividad de las reglas lingüísticas es muy próxima al lenguaje natural, y por ello su uso favorece la Interpretabilidad de los modelos difusos dado que el proceso de comprensión humano se realiza en términos de lenguaje natural. El filósofo Paul Grice propuso cuatro máximas conversacionales que se deducen de la pragmática del lenguaje natural y el sentido común [13]:

1. *Máxima de Calidad*: No decir aquello que se cree falso. No decir nada para lo que no se tiene evidencia suficiente de ser verdadero.

2. *Máxima de Cantidad*: Adecuar el mensaje al objetivo de la comunicación. No incluir en la conversación más información de la estrictamente necesaria.

3. *Máxima de Relevancia*: Sólo incluir información relevante.

4. *Máxima de Forma*: Evitar el uso de expresiones que generan ambigüedad. Hacer un discurso breve, conciso y ordenado.

Tener en mente las máximas de Grice a lo largo de todo el proceso de modelado difuso puede ayudar a obtener SBRDs más fáciles de interpretar: La base de reglas debe ser coherente (*Máxima de Calidad*), no debe incluir redundancias (*Máxima de Cantidad*), ni tampoco ambigüedades (*Máxima de Forma*). Además, la selección de las reglas más importantes (*Máxima de Relevancia*) producirá sistemas más robustos y compactos, favoreciendo el equilibrio entre Interpretabilidad y Precisión.

Otro reto importante consiste en construir sistemas capaces de explicar el resultado de la inferencia difusa. Hasta ahora los esfuerzos por medir la Interpretabilidad se han centrado en evaluar la complejidad de la descripción lingüística del SBRD (*Descripción global*). Sin embargo, como resultado de la regla composicional de inferencia lo habitual es que varias reglas se disparen a la vez para un mismo vector de entrada. Esto implica que el análisis del comportamiento del sistema (*Explicación puntual*) a partir de su descripción no es inmediato, puede llegar a ser muy complicado dependiendo del mecanismo de agregación y la estrategia de defusificación. Evidentemente, el análisis será más sencillo si el modelo se construye pensando en favorecer su Interpretabilidad. Cuando la salida del sistema se obtiene como resultado de la combinación de un conjunto de reglas, lo más habitual, entender la descripción del modelo es un requisito necesario pero no suficiente para comprender su comportamiento. Esto hace que algunos investigadores pongan en duda la reconocida Interpretabilidad de los SBRDs: *Mientras que los sistema clásicos se ven como cajas negras donde una fórmula matemática establece la relación entrada-salida, los sistemas difusos se pueden ver como cajas blancas (o grises según el proceso de modelado seguido) en el sentido de que cada elemento del conjunto puede ser analizado y entendido de forma sencilla por una persona.*

Por ejemplo, de forma similar a como se generan resúmenes textuales para explicar el tiempo que va a hacer a partir de los datos de previsión meteorológicos [24], se podría intentar generar resúmenes textuales para explicar el resultado de la inferencia de un SBRD.

Como alternativa se puede construir SBRDs que sólo admitan como máximo dos entradas por regla [15], de forma que todas las reglas admitan una representación gráfica sencilla en dos dimensiones.

Volviendo a la búsqueda del equilibrio entre Interpretabilidad y Precisión, el último reto que nos gustaría plantear es la utilización de los índices de Interpretabilidad propuestos en problemas multiobjetivo reales a fin de evaluar su utilidad. Hasta ahora, en estos problemas, la Interpretabilidad se evalúa con índices muy elementales [16]: Número de reglas, Número total de premisas, Número medio de premisas por regla, etc.

En el contexto del análisis multiobjetivo con algoritmos genéticos [2], lo ideal sería poder parar la ejecución del algoritmo y tener un experto que evaluase los SBRDSs, construyese un frente de Pareto según su experiencia, y seleccionase los miembros de la población para la generación siguiente. Como la ejecución manual es inviable y todo el proceso debe ser automático, el resultado final depende mucho del índice que se use para medir la Interpretabilidad. Disponer de un índice difuso capaz de adaptarse de forma sencilla a las preferencias del usuario y al contexto del problema bajo estudio sería muy útil y contribuiría decisivamente a la obtención de SBRDs con un mejor equilibrio entre Precisión e Interpretabilidad.

4 CONCLUSIONES

El modelado de SBRDs representa un área de gran actividad investigadora y la búsqueda del equilibrio entre Precisión e Interpretabilidad es uno de sus grandes desafíos. Aunque hace años que se trabaja en ello, el problema sigue abierto porque la tarea no es sencilla y plantea varios retos entre los que hay que destacar la caracterización y evaluación de la Interpretabilidad.

Se ha hecho un gran esfuerzo por identificar los factores que influyen en la Interpretabilidad de SBRDs pero todavía no existe un índice universalmente reconocido y utilizado en aplicaciones reales.

Para superar este reto hay que dar el salto de la computación numérica a la computación con palabras. Aprovechar la experiencia de otras disciplinas en el análisis de lenguaje natural y la evaluación de percepciones puede ser de gran utilidad. La medida de Interpretabilidad es muy subjetiva, los criterios de evaluación varían de unas personas a otras.

Referencias

[1] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto, and F. Herrera. Hybrid learning methods to get the interpretability-accuracy trade-off in fuzzy modeling. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 10(9):717–734, July 2006.

[2] R. Alcalá, M. J. Gacto, F. Herrera, and J. Alcalá-Fdez. A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(5):539–557, 2007.

[3] José M. Alonso, S. Guillaume, and L. Magdalena. A hierarchical fuzzy system for assessing interpretability of linguistic knowledge bases in classification problems. In *IPMU 2006. Information Processing and*

Management of Uncertainty in Knowledge-Based Systems, pages 348–355, Paris, France, July, 2-7, 2006.

[4] U. Bodenhofer and P. Bauer. Towards an axiomatic treatment of "interpretability". In *Proc. IIZUKA2000*, pages 334–339, Iizuka, October 2000.

[5] U. Bodenhofer and P. Bauer. A formal model of interpretability of linguistic variables. *In [10]*, pages 524–545, 2003.

[6] A. Botta, B. Lazzerini, F. Marcelloni, and D. Stefanescu. Exploiting fuzzy ordering relations to preserve interpretability in context adaptation of fuzzy systems. In *FUZZ-IEEE07*, pages 1137–1142, London, 2007.

[7] E. Van Broekhoven, V. Adriaenssens, and B. De Baets. Interpretability-preserving genetic optimization of linguistic terms in fuzzy models for fuzzy ordered classification: An ecological case study. *International Journal of Approximate Reasoning*, 44:65–90, 2007.

[8] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. *Accuracy improvements in linguistic fuzzy modeling*, volume 129. Studies in Fuzziness and Soft Computing, Springer-Verlag, Heidelberg, 2003.

[9] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: An overview. *In [10]*, pages 3–22, 2003.

[10] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. *Interpretability issues in fuzzy modeling*, volume 128. Studies in Fuzziness and Soft Computing, Springer-Verlag, Heidelberg, 2003.

[11] J. Valente de Oliveira. Semantic constraints for membership function optimization. *IEEE Transactions on Systems, Man and Cybernetics. Part A, Systems and Humans*, 29(1):128–138, January 1999.

[12] P. Fazendeiro and J. Valente de Oliveira. A working hypothesis on the semantics/accuracy synergy. In *Joint EUSFLAT-LFA 2005*, pages 266–271, Barcelona, Spain, September, 7-9, 2005.

[13] H. P. Grice. Logic and conversation. *Cole, P. and Morgan, J. (eds.) Syntax and semantics, New York: Academic Press*, 3:43–58, 1975.

[14] S. Guillaume. Designing fuzzy inference systems from data: An interpretability-oriented review. *IEEE Transactions on Fuzzy Systems*, 9(3):426–443, 2001.

[15] H. Ishibuchi, Y. Kaisho, and Y. Nojima. Designing fuzzy rule-based classifiers that can visually explain their classification results to human users. In *3rd International Workshop on Genetic and Evolving Fuzzy Systems*, pages 5–10, 2008.

[16] H. Ishibuchi and Y. Nojima. Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. *International Journal of Approximate Reasoning*, 44:4–31, 2007.

[17] H. Ishibuchi and T. Yamamoto. Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 13(4):428–435, August 2005.

[18] Y. Jin, W. von Seelen, and B. Sendhoff. On generating FC^3 fuzzy rule systems from data using evolution strategies. *IEEE Transactions on Systems, Man, and Cybernetics*, 29(6):829–845, 1999.

[19] E. H. Mamdani. Application of fuzzy logic to approximate reasoning using linguistic systems. *IEEE Transactions on Computers*, 26(12):1182–1191, 1977.

[20] C. Mencar, G. Castellano, and Anna M. Fanelli. Some fundamental interpretability issues in fuzzy modeling. In *Joint EUSFLAT-LFA 2005*, pages 100–105, Barcelona, Spain, September, 7-9, 2005.

[21] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2):81–97, 1956.

[22] D. D. Nauck. Measuring interpretability in rule-based classification systems. In *FUZZ-IEEE03*, volume 1, pages 196–201, St. Louis, Missouri, USA, May, 25-28, 2003.

[23] E. H. Ruspini. A new approach to clustering. *Information and Control*, 15(1):22–32, July 1969.

[24] S. G. Sripada, E. Reiter, J. Hunter, and J. Yu. Generating english summaries of time series data using the gricean maxims. In *Pragmatics and Content Selection SIGKDD'03, ACM Special Interest Group on Knowledge Discovery and Data Mining*, pages 187–196, Washington, USA, 2003.

[25] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modelling and control. *IEEE Transactions on Systems, Man, and Cybernetics. Part C: applications and reviews*, 15:116–132, 1985.

[26] A. Tarski, A. Mostowski, and R. Robinson. *Undecidable Theories*. North-Holland, 1953.

[27] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[28] Lotfi A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, 3:28–44, 1973.

[29] Lotfi A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Parts I, II, and III. Information Sciences*, 8, 8, 9:199–249, 301–357, 43–80, 1975.

[30] Lotfi A. Zadeh. From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. *IEEE Transactions on Circuits and Systems - I: Fundamental theory and applications*, 45(1):105–119, 1999.

FUZZY TEMPORAL RULE-BASED SYSTEMS: NEW CHALLENGES

Senén Barro Alberto J. Bugarín Purificación Cariñena Félix Díaz-Hermida Manuel Mucientes

Departamento de Electrónica y Computación. Universidad de Santiago de Compostela
{senen, alberto, puri}@dec.usc.es, felixdh@usc.es, manuel@dec.usc.es

Abstract

Applications in control and/or monitoring usually demand an explicit representation and management of time. Fuzzy Temporal Rules (FTRs) introduce an explicit fuzzy representation of time, allowing relative occurrences of events, quantification, and other types of operators. In this model fuzziness can appear not only in the temporal references, but also in the sets of values and operators involved. This paper presents the most relevant features and applications of knowledge representation and reasoning within the FTRs framework. New challenges, both regarding automated learning of knowledge bases involving these rules and potential new fields of application in a number of areas, are also described.

Keywords: Fuzzy Temporal Rules, Fuzzy Temporal Reasoning, Evolutionary Fuzzy Systems

1 INTRODUCTION

Modelling of applications in control and/or monitoring by means of fuzzy rule-based systems usually demand an explicit representation and management of time. In propositions such as "*The temperature has been very low for a few seconds*", "*Vomiting started more than 15 minutes after the beginning of the radiation exposure*" or "*The distance to the wall has kept low in part of the last measurements*" time appears as a temporal reference for a number of events ("for a few seconds", "in the last measurements") or as a relationship between the occurrence of different events ("15 minutes after"). This explicit representation of time in propositions and the subsequent models for performing reasoning using these pieces of knowledge cannot be directly accomplished within the non-temporal propositions scheme "*X is A*" usual in fuzzy rule-based systems (or at least not in a direct and intuitive way).

Other expressions involving different types of operators, either describing relationships between the variables ("*the temperature in most of the tanks is high*", "*the minimum of the high pressure values*", "*the mean of the velocities in the vicinity of an obstacle*"), or stating the occurrence of events in a fuzzy temporal window ("*the mean of high temperatures in the last ten minutes*", "*the minimum of the recent values of high pressure*"), or even combinations of both ("*for a minute the mean value of the temperature in heater 1 was bigger than the mean value of the temperature in heater 2*") are examples of knowledge representation needs when modelling complex dynamic systems. Furthermore, in other fields like flexible querying or data mining it may become necessary to evaluate the fulfilment of complex relationships such as "*most of the people who entered the ER recently are young*", "*the maximum salary between middle-aged employees with low income*". To sum up, modelling such expressions demands an adequate semantic model that take into account an explicit fuzzy representation of time and consider that fuzziness can appear not only in the temporal references, but also in the sets of values and operators involved.

With the purpose of increasing the flexibility of the usual model for fuzzy propositions, some models that include time more or less explicitly as another decision variable have been proposed [15, 18, 9, 6, 8]. In [3] a model of fuzzy temporal rules (FTRs) that extended these proposals aiming to allow the fuzzy representation and reasoning on temporal and non-temporal references was described. A language was introduced which allows experts to describe their knowledge (including the temporal component) in a legible and flexible way, allowing relative occurrences of events, quantification and other types of operators.

Since this FTRs model is oriented to real-time applications, the signal values are assumed to be obtained with a given periodicity, or at least will be associated with successive time points. In a dynamic system the variables are ob-

tained from measurements of external signals, whose values are changing in time, or can be inferred at particular time points. This makes it necessary to consider not only the current values of the signals, but also those previously observed, which allows to reason about the history of signal values.

The general expression of a FTR in this model is:

IF P_1 and P_2 and ... and P_M THEN C_1 and C_2 and ... and C_N

where P_m, $m = 1,...,M$ and C_n, $n = 1,...,N$ are Fuzzy Temporal Propositions (FTPs). For the sake of clarity, we will describe in what follows some of the most relevant features of the FTPs.

Propositions in the antecedent part of the rule represent facts where both temporal and non-temporal information can be fuzzy and can be given in an absolute manner, or relative to the occurrence of other facts. As an example, we present in Table 1 some of the rewriting rules of the grammar to construct these FTPs (the reader is referred to [3, 4] for a more complete description of the rules).

A **value constraint** is a non-temporal fuzzy value, defined on the universe of discourse of a given variable. Three types of value constraints can be applied to a signal: absolute ("Pressure is *high*"), relative to another reference signal through a point to point *value relationship*, as in "Pressure in heater 1 is *much greater than pressure in heater 2*", and quantified value constraints ("Pressure is *high in most of the heaters*").

A *filtered signal* represents a signal defined in the $[0, 1]$ interval, obtained after applying a value constraint to a variable, by means of a **linguistic filtering** process [14]: evaluation of the possibility distribution defining the constraint on each one of the signal values.

Sometimes it is necessary to establish **temporal constraints** on the set of time points where a signal is being evaluated. For example, it may be required that the value constraint is applied not over all the history of signal values, but only for a subset of time points. This subset will be defined by a temporal constraint, which can be an instant or an interval, fuzzy or not. Temporal constraints can also be absolute or relative, so they require the consideration of different relationships between temporal entities. We consider the following basic temporal entities: fuzzy temporal instant, fuzzy temporal duration, fuzzy temporal interval and fuzzy temporal relationships, assuming that all the basic temporal relationships (both qualitative and quantitative) between instants and/or intervals, and between temporal durations can be reduced to relationships between time points and durations. Some examples of this kind of relationships are [1]: *before, at the end of, when,...*.

A temporal constraint can also act as a *temporal context* [17] for signal evaluation, establishing a temporal window within which the degree of fulfilment of the proposition has to be obtained. Basically, a temporal reference can be described in an absolute manner ("*at 20:00*"), relative to the current time point ("*ten minutes ago*") or relative to the occurrence of an event ("*a little after an increase in pressure*", "*between 30 minutes and 2 hours after the beginning of the irradiation*" [8]).

Three types of operators are considered: quantifiers, specification operators and reduction operators. **Quantifiers** in the temporal part of FTPs allow to model expressions of the form $X\ is\ A\ in\ Q\ of\ T$ (for example, "*the temperature has been high in* most of the *previous half an hour*"), which introduce the central concept of *persistence* [2], related to the use of fuzzy temporal intervals as temporal references in the propositions: when a condition on a variable must be evaluated over a fuzzy temporal interval, it may be required that the value or condition is fulfilled, at least to some degree, for all the time points in the interval, or just for any of them. Between both extremes there is a continuous of situations ("*in most of the previous hour*") that have not yet been fully developed within the scope of temporal reasoning.

A signal **specification operator** selects a candidate among several, according to specific non-temporal or temporal, and maybe fuzzy, criteria. Examples of this type of operators are *first, last, maximum, minimum*... Thus, a proposition can refer to "*the maximum of pressure values in the last 30 minutes*", "*the last value of high temperatures*", "*the minimum of the velocity values in the proximity of an obstacle*"...

A **reduction operator** resumes a series of signal values in a single one, defined on the same universe of discourse. When describing dynamic systems we usually find expressions involving operators that act on the values observed or inferred for a variable on a temporal reference, in order to obtain a new value that does not necessarily have to match any of the original ones. Examples of these operators are: *mean value* or *accumulated value*.

Considering this type of FTP, expressions containing value relationships between variables can be described ("*Temperature in heater 1 is much greater than temperature in heater 2*"), with different degrees of structural complexity ("*Temperature in heater 1 in the previous seconds was much greater than temperature in heater 2 two hours ago*"). More generic structures can be described, where the temporal and value constraints in the proposition are rewritten in a more elaborate way: "*Pressure 1 was much higher than pressure 2, a little after temperature 1 was lower than temperature 2*", "*Temperature is low while pressure is high*". Related to the use of different operators, other expressions are: "*Most of the temperature values through the previous*

Table 1: FTP: General structure of a propositional clause.

(R1) <Propositional clause>::= (<Proposition> | <Proposition><Value relationship><Proposition>)[<Temporal interval>]
(R2) <Proposition>::= (<Filtered signal> | <Signal>) <Temporal instant> |
| (<Operator> | <Quantifier>) (<Filtered signal> | <Signal>) [<Value constraint>] <Temporal interval>
(R3) <Filtered signal>::= <Signal> <Value constraint>
(R4) <Signal>::= TEMPERATURE | PRESSURE | ...
(R5) <Value constraint>::= (<Relative value constraint> | <Absolute value constraint>) [<Quantified value constraint>]
(R6) <Relative value constraint>::= <Value relationship> <Signal>
(R7) <Absolute value constraint>::= HIGH | LOW | MORE THAN 30 | ...
(R8) <Quantified value constraint>::= <Quantifier> [<Signal>] (<Absolute value constraint> | <Relative value constraint>)
(R9) <Value relationship>::= GREATER THAN | LESSER THAN | SIMILAR TO |...
(R10) <Temporal instant> [...]
(R11) <Temporal interval> [...]
(R24) <Operator>::= <Reduction operator> | <Specification operator>
(R27) <Quantifier>::= <Absolute quantifier> | <Relative quantifier>
(R28) <Absolute quantifier>::= [APPROXIMATELY] (MORE_THAN_2 | LESS_THAN_7 | BETWEEN_3_AND_5 | ...)
(R29) <Relative quantifier>::= [APPROXIMATELY] (ALL | MOST_OF | HALF | PART OF | ...)

minutes have been high" (quantification), "*The maximum of temperature values between 6 and 8:30 was very low*" (specification) or "*The mean value of temperature in the last 48 hours was moderate*" (reduction).

1.1 APPLICATIONS OF FUZZY TEMPORAL RULES

The first real application of the FTR model was in the field of mobile robotics, namely in the implementation of two behaviours on a Nomad 200 robot: moving objects avoidance and wall-following [13, 12]. When the robot has to avoid a collision with a moving object, it is specially worthy to estimate the trend of the object. In that way, instead of taking a control action based on the current values of some variables, a more appropriate behavior can be obtained through the analysis of the movement of the object in the recent past (last few seconds). FTRs play a central role in this analysis, as they can cope with noisy measurements coming from the robot sensors and, at the same time, evaluate the trend of each variable. The control system was implemented with 114 rules divided in three modules: object course evaluation, behaviour selection, and behaviour implementation. A typical FTP for this behaviour is: "*the speed of the object is not increasing in part of the last seconds*".

The wall-following behaviour was implemented by several authors using fuzzy controllers. However, on the real robot this behaviour is quite sensible to spurious sensor measurements, which is reflected in oscillations of the robot along the path. That situation can be mostly solved taking control actions based on several consecutive measurements, discarding those that are noisy. A 321 FTRs system was impemented to that purpose. The system has been divided in three rule bases: straight wall, convex and concave corner. FTPs are like: "*right-hand distance diminishes a little approximately in the last measurements*". With this model, the persistence in the occurrence of certain events in time, and also temporal and value relationships between signals can be taken into account.

Another practical application of the FTR model was in the field of landmark detection for mobile robots in indoor environments (door-detection in a corridor [5]). This task constitutes a vital part within a complex navigation system in closed environments, since it provides information for robot localization in the environment, fundamental for planning trajectories and movements. It was developed again using real data from the ultrasound sensors of the robot, and good results have been obtained given the requirements of the system.

Door detection requires the consideration of two aspects: temporal relationship between consecutive detections in some sensors, and value relationships between the signals. Translating the knowledge associated to door detection into our FTRs model, two rules are obtained for the detection of doors on the robot's left and right side, respectively. To describe these rules, the temporal relationship "a few instants before" between the events "sharp rise" and "sharp fall" of the ultrasound signals was considered since the temporal distance between the current distance value provided by one of the sensors, and previous values provided by the other is the key for the detection of a door.

This temporal distance between the events marking the beginning and end of each door depends on the velocity of the robot. This velocity is not constant, but changes according to the environment (it decreases when the robot turns, or when an irregularity in the wall is detected, and it is bigger in straight wall situations [12]), so we take as a reference the velocity when the robot is getting close to a door. Since the velocity presents important oscillations, it is more adequate to consider the mean velocity in the neighbourhood of the door we try to detect. Thus, a FTP describing this situation, for doors on the left side, is: "*The mean value of the velocity in the previous instants to a sharp fall in* $UltrasoundSensor6$". The particular semantic expres-

sions used in this application of FTRs to door detection are detailed in [5].

Although the structure of the propositions involved in these three tasks was quite simple (this meaning they did not made use of all the expressiveness of the FTRs model) the quality of the results obtained made us feel optimistic about the applicability of the model. Nevertheless further work should be done in order to successfully extend the application of the FTRs model to different real complex applications. Given the huge number of rules that have to be defined for complex behaviours such as the ones previously mentioned, a natural continuity of the work done was to introduce some automatic learning mechanism for the generation of the FTRs, at least for a part of the model (if not for the complete set of rules in the grammar). In this way, the process of construction of the fuzzy temporal knowledge base (FTKB) would be automatic, saving development time and allowing the model to be applied in many different fields, by users not specially skilled in the underlaying formalism. The next section presents the application of an evolutionary algorithm to learn a FTKB for the classification of moving objects in mobile robotics.

2 AUTOMATED EXTRACTION OF FTRS FROM DATA USING EVOLUTIONARY COMPUTATION

Learning of a FTKBs in the field of mobile robotics for the detection of people and other moving objects was described in [11]. A correct classification of these objects is fundamental for the development of a number of tasks for human-robot interaction. This is a usual scenario in real environments, where people walk together, and carry objects (like suitcases, trolleys, etc.). In these situations, the density of moving objects is high, and there are occlusions between moving objects, so the detection and tracking difficulties are higher [10] than for individual objects. Detection of people moving in groups is therefore a truly complex pattern classification task that demands evaluating spatial-temporal information rather than classifying a pattern as a moving object by solely using the current values of some variables. In this realm, quantifying the fulfillment of a linguistic label by a set of data (spatial pattern), and analyzing the persistence of this fulfillment in a temporal reference (temporal pattern) is accomplished by means of FTRs made up of FTPs involving both spatial and temporal quantification. In this section we describe a pattern classifier system for the detection of moving objects using laser range finders data that has been learned through an evolutionary algorithm [11]. The laser sensors emit several beams, each one in a given direction. When a beam hits an obstacle, it is reflected and registered by the scanner's receiver. The time between the transmission and the reception of the pulse is known as the time of flight. With this information, the distance measured in the direction of each beam can be calculated. The existence of moving objects in the surrounding of the robot can be determined by two basic features:

1. A moving object appears in the distance histograms of the lasers as a local minimum.

2. As other static objects (like the legs of a table, a wastepaper basket, etc.) can share this characteristic, the object that has generated the local minimum must also be new at that position. This condition is verified using an occupancy grid map, and calculating the probability $P_{new}^{i,j}$ that the detected object is new in that cell.

The rules for the detection of moving objects must take into account information about the size of the local minimum (the gap between the object that produces the local minimum and the obstacles behind), the number of beams of the local minimum, and the probability of a new object in each of the cells of the local minimum. This last characteristic distinguishes between static and moving objects. The analysis of $P_{new}^{i,j}$ can also generate false positives. These errors can be reduced if, instead of taking into account $P_{new}^{i,j}$, the system analyzes this probability in different time instants. Moreover, some cells of the local minimum can have a high probability of containing a new object, but others not. The system should also quantify how many cells must have a P_{new} above a threshold.

A FTP that is able to perform the spatial filtering of successive values of $P_{new}^{i,j}$ is the following:

$$P_{new} \textit{is high in most of the cells in part of the last instants} \quad (1)$$

This type of FTP (*X is A <in* Q_s *of>* S_s *<in* Q_t *of>* S_t), where X ($P_{new}^{i,j}$) is a linguistic variable, A ($high$) represents a linguistic value of X, Q_s ("*most of*") and Q_t ("*part of*") are fuzzy quantifiers, S_s is the (fuzzy) set of reference ("*cells*"), and S_t (" *the last instants*") is a temporal reference, can be formally described in the framework of the FTR model.

The syntactic tree representing FTP in (1) according to the grammar proposed in [3] is shown in Figure 1. In this proposition quantifiers are applied both to the non-temporal part of the proposition (*in most of the cells*) and to the temporal part (*in part of the last instants*).

The design of this FTP involves the definition and tuning of ten parameters (linguistic labels and quantifiers) per rule for this application. Moreover, drastic changes in the characteristics of the environment or the moving objects could affect the accuracy of the pattern classifier system, making useless the tunned parameters. Therefore, an evolutionary algorithm based on the cooperative-competitive approach

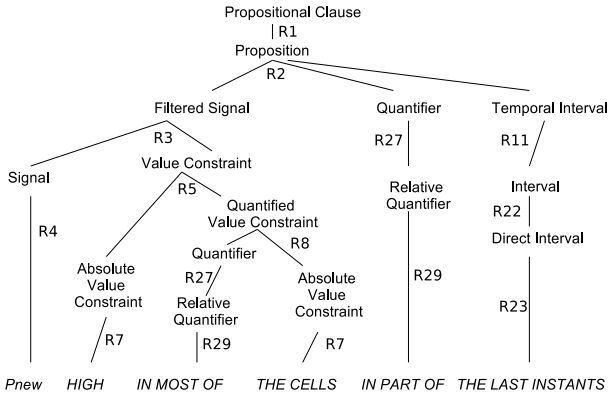


Figure 1: Syntactic decomposition of FTP: $P_{new}^{i,j}$ *is high in most of the cells in part of the last instants*.

has been used to learn a pattern classifier that was tested with data obtained with a *Pioneer II* robot equipped with two laser range scanners.

Experiments have been performed with a five-fold cross-validation. Results showing the average and standard deviation values of the five-fold cross-validation for the number of rules of the knowledge base, false positives (*fp*), false negatives (*fn*), and percentage of examples correctly classified (% correct) over the training and test examples sets are shown in Table 2. The percentage of examples correctly classified over the test set is 96,10%, despite the number of moving objects, their high concentration in small areas of the hall, and the movement of the robot (see [11] for more details).

Table 2: Results of the five-fold cross-validation.

	Training ($\overline{x} \pm \sigma$)	Test ($\overline{x} \pm \sigma$)
#rules	25.60 $\pm$ 5.32	—
fp	11.40 $\pm$ 5.73	2.20 $\pm$ 2.95
fp	19.60 $\pm$ 7.57	7.40 $\pm$ 3.21
% correct	96.85 $\pm$ 0.61	96.10 $\pm$ 1.74

The fact that people move in groups increases the difficulties in the detection, as compared with a single person, because cells that were originally free are occupied in successive time instants by different moving objects (legs). Thus the values of $P_{new}^{i,j}$ are lower. During the experiment up to ten people moved around the robot, which means 20 moving objects at the same time. Such a high number of moving objects concentrated in a few and small areas of the environment generates partial (and total) occlusions of the objects, modifying the values of the gaps and the number of beams of the detected moving objects. Also, more than half of the time the robot was moving, making harder the discrimination of new objects as the scan matching errors increase.

3 OPEN ISSUES AND NEW CHALLENGES

Taking into account the FTR model features there are a number of open issues, both in the area of the most appropriate strategies and techniques for automatic discovery of rules and in the area of real applications of the model.

3.1 AUTOMATIC LEARNING OF FTRS

One of the challenges that must be tackled to extend the use of FTRs is the adaptation of the existing genetic fuzzy systems to the FTR paradigm. FTRs have a complex and variable structure. This structure can change not only at the rule level (as a conventional fuzzy rule), but also at the proposition level. Therefore, it is necessary to define a context-free grammar to generate only those structures that are valid. Genetic programming seems to be the most adequate frame to learn the rules, due to its ability to work with chromosomes with very different shapes. Moreover, the size of the search space is much larger than that of a conventional knowledge base, due to the complex syntax of the rules, the dependence on the temporal reference framework, the existence of reduction, specification and quantification operators, temporal membership functions, etc. In this context parallel and/or distributed approaches for the execution of the algorithm should be taken in consideration as options.

Another challenge is the construction of the training dataset. Generally, examples only provide information of the spatial part of the propositions, i.e., they do not contain temporal information. Thus, the temporal part of the proposition has to be learned from scratch, trying to improve the performance of the corresponding rule or knowledge base. Usually, the initial population is composed of conventional fuzzy rules that evolve into FTRs, because the consideration of the temporal part of the proposition contributes to an increase in the performance. Of course, the inclusion of temporal information in the initial population could accelerate the convergence of the learning algorithm, and improve the results. However, how this temporal information could be included remains an open issue.

3.2 NEW FIELDS OF APPLICATION OF FTRS

FTRs have been successfully applied in the field of mobile robotics, but there are many other fields which could benefit from the extended capabilities of this model as opposed to traditional non-temporal rules.

3.2.1 Medicine

In the process of diagnosing a disease a lot of information is obtained from the patient, producing a variety of data stored in different places: administrative data, physio-

logical signals (ECG,...), results from clinical tests (blood analysis,...), etc. Specially in the case of data from laboratory tests, diagnosis or therapeutical procedures, most of the data are temporal. In some cases, it is relevant to know the relative occurrence of an event with regard to other events; in other cases, what is interesting is the duration of an event. Temporal patterns can be used to distinguish between different pathologies, or to support the choice of a given therapy.

When the amount of temporal data stored in a database relative to a particular patient is too large to be globally examined by a human expert, it would be useful to have an automatic tool that can extract relevant information from these data, which can assist the expert in diagnosing a given disease or the degree of severity of an illness.

Development of an automatic tool for standard polysomnographic analysis is currently undergoing, aiming to provide the physician with new information from the intelligent processing of the set of polysomnography signals. The use of techniques for temporal abstraction of the information from two perspectives (solutions based on domain knowledge, and solutions based on supervised learning) needs to be complemented with the design, development and implementation of new computational tools for information fusion and analysis of association relationships in the information obtained from patients follow-up.

In particular, it could be of interest to extract fuzzy temporal association rules from a database of temporal events, which have been obtained from a preprocessing of the polysomnographic data. In this sense, existing algorithms for association rules extraction must be adapted, since none of the proposals in the literature deals with actual fuzzy temporal information, but only time stamps associated to the data. Furthermore, it would be interesting to include the concept of more complex relationships/operators such as quantification in the associations extracted from the rules, in order to evaluate for instance the severity of an illness.

3.2.2 Trends analysis in economy temporal series

In fields like economy, the problem of qualitatively describing information provided by indicators (e.g. the trend of a temporal series: increasing, decreasing or approximately constant) is far more complex than in more analytical-friendly fields. Although statistical techniques can be used to analyse temporal series, we want to stress that a linguistic qualitative description of these signals (as usual for example in finances reports or economic argumentation) is very useful to provide accurate and brief information that compiles all the numeric data. As an example, an "approximate trend" of a signal S could be defined involving temporal information and quantifiers as:

- S *is approximately increasing/decreasing around* t: when the percentage variation is *significantly superior/inferior* to 0 in *most of the* points close to t.
- S *is approximately constant around* t: when the percentage variation is *significantly close* to 0 in *most of the* points close to t.

Here, *significantly superior/inferior* to 0 is a fuzzy number, and *most of the* a fuzzy quantifier.

There are plenty of common-use examples in these areas that need to be modelled using similar fuzzy semantic approaches, such as: "*the oil production is becoming stagnant in the last years*" (reinterpreted as "*in at least about the 80% of the approximate period of the last 5 years, the increasing in oil production was negative or only slightly superior to 0*"); "*the variation of the yearly variations of home prices has been close to 0 in the last months*"; "*the correlation between the significant falls in the yearly variation of the number of initiated homes, and the significant increases in the unemployment rate in the real state sector six months later, has been very high in the last year*" (in this case using correlation quantifiers such as those described in [7]).

Several promising application fields of this kind of expressions can be identified:

- **Fuzzy filtering.** Applying quantifiers to percentage variations extends the possibilities of fuzzy filtering. For example, in Figure 2 (top) the yearly percentage variation of the number of initiated homes is shown for the period Jan. 1992-Oct. 2007. Figure 2 (bottom) shows the result of evaluating the quantified expression "*In most of the last six months the yearly variation in the number of initiated houses has significantly fallen*". Periods of negative trends are easily detected by means of a fuzzy quantified expression.
- **Argument assessment.** Reasoning in economy is plenty of examples of expressions that can be modelled and evaluated by means of a combination of percentage series and fuzzy quantifiers. For example, the degree of fulfillment of expression "*the oil production is becoming stagnant in the last years*" could be evaluated in order to check the validity of the argument.
- **Linguistic summaries of series.** This would allow automatic detection of positive, negative and stagnant trends in a temporal series. For example, the proposition "*the unemployment trend is increasing in the last year*" (reinterpreted as "in most of the last twelve months, the yearly unemployment variation was increasing") could be an automatically generated summary as a result of a fuzzy temporal rule-based analysis of the unemployment rate.

- **Temporal causal analysis.** Let us consider the expression "*the correlation between the significant falls of the yearly percentage variation of initiated homes, and the significant increases of the yearly percentage variation in the unemployment rate in the real state sector six months after that, has been very high in the last year*". Evaluating the expression on experimental data, if for some range of the temporal difference (six months in the example) the fulfillment of the expression is very high a possible causal relation could be detected.

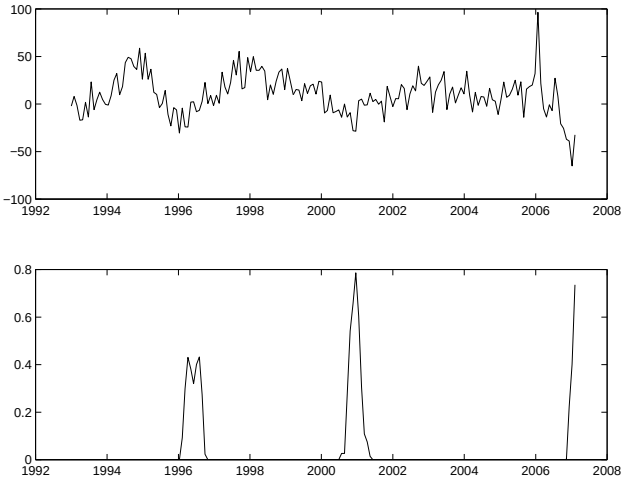


Figure 2: Fuzzy percentage quantified filtering of the number of initiated houses.

The most interesting features of this approach are:

- The approximation is linguistic, fuzzy and easily interpretable.
- Quantifiers, temporal granularity, and linguistic variables are all fuzzy thus being able to cope with uncertainty.
- Fuzzy quantifiers let us avoid outliers (usually due to noise); in this way, the result does not depend on a particular increase in a decreasing series.
- Working with percentage variations allows us to operate with independence of the base signal. As absolute values of series are not involved, certain signal independence is achieved.
- And perhaps the most remarkable, *common sense* expressions (e.g., economic arguments) can be evaluated.

3.2.3 Business intelligence

Organizations and companies have carried out in the recent years a huge effort to improve their productivity by means of automating their business processes. Such automation was achieved through defining workflows that describe the coordination of the execution of the activities that have to be performed in order to achieve the goals stated in the business process [16].

Execution of workflows is typically performed in the company information system, where all the data that are generated during the execution are stored. Examples of such data are: who was the person/process that performed a given activity, when it occurred, which was the information used,... Starting from such data a process mining stage can be done, aiming [16] (i) to reconstruct or discover processes for an automatic definition of the workflow associated to the activities performed by the members of the organization during the execution of a business process, (ii) to evolve processes, aiming to monitor the workflow execution in order to detect exceptions or execution conditions of activities that may produce changes in the initial workflow.

The techniques proposed for process mining are typically based on the identification of causal temporal relationships among the events recorded in the information system. Temporal rules in this context could allow to detect the execution order of the activities or to avoid repetition of activities or deadlocks. Within this context, frameworks for temporal representation and reasoning involving representation of changing relationships over time (instants, intervals, and topological relationships among them) are now starting to be used for development of semantic-based approaches to business intelligence applications.

3.2.4 Industrial and other applications

Monitoring, supervision, diagnosis and control of processes are areas where a number of elements coexist, making them a potential field of application for temporal information processing techniques. In complex environments such as chemical or environmental processes, power and water-treatment plants a huge number of signals have to be processed, temporal evolution of indicators is relevant, and also temporal relationships between events needs to be considered. Furthermore knowledge-based approaches in these fields are becoming a common-use tool for the design of advanced intelligent systems and researchers are paying increasing attention to them. Developing a support software suite for defining and executing Fuzzy Temporal Knowledge Bases (those involving FTRs), with facilities for rule editing, debugging, code generation, and automated learning of rules from data is a must for solving complex applications in industrial systems using FTRs.

4 REMARKS

Although Fuzzy Temporal Rules provide a general framework for the representation of knowledge and reasoning

in complex systems, the applications developed to date (mostly in the field of intelligent control and classification in robotics) do not fully exploit all the syntactic and semantic richness of the model. The natural niche for application of FTRs is monitoring, control, diagnosis and other complex tasks in dynamic systems, and these fields demand that tools and methodologies are available for making easier the design and/or automatic learning of fuzzy temporal systems. This is in our opinion the main challenge to be faced in the near future for achieving a successful use of temporal reasoning based systems in these fields.

Acknowledgements

Authors wish to acknowledge Spanish Ministry for Education and Science and Xunta de Galicia for their financial support under grants TIN2005-08521-C02-01, TIN2006-15460-C04-02 and 07SIN005206PR.

References

[1] S. Barro, R. Marín, J. Mira and A. Patón. A model and a language for the fuzzy representation and handling of time. *Fuzzy Sets and Systems*, vol. 61, pp. 153–175, 1994.

[2] P. Cariñena, A. Bugarín, and S. Barro. Un modelo para la cuantificación de la persistencia en proposiciones temporales borrosas. In *Actas del VIII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF'98)*, 1998, pp. 109–116.

[3] P. Cariñena, A. Bugarín, M. Mucientes, F. Díaz-Hermida, and S. Barro. *Technologies for Constructing Intelligent Systems 2. Tools*, ser. Studies in Fuzziness and Soft Computing. Springer-Verlag, 2002, vol. 90, ch. Fuzzy Temporal Rules: A Rule-based Approach for Fuzzy Temporal Knowledge Representation and Reasoning, pp. 237–250.

[4] P. Cariñena. Un modelo de Reglas Temporales Borrosas para el razonamiento sobre sistemas dinámicos. Ph.D. Dissertation. University of Santiago de Compostela, 2003.

[5] P. Cariñena, C.V. Regueiro, A. Otero, A. Bugarín and S. Barro. Landmark detection in mobile robotics using fuzzy temporal rules. *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 423–435, 2004.

[6] B. Carse, T. Fogarty and A. Munro. Artificial evolution of fuzzy rule bases which represent time: a temporal fuzzy classifier system. *International Journal of Intelligent Systems*, vol. 13, pp. 905–927, 1998.

[7] F. Díaz-Hermida, P. Cariñena, A. Bugarín. Efficient evaluation of similarity quantified expressions in the temporal domain. *Soft Methods for Integrated Uncertainty Modelling -SMPS2006*, pp. 191–198. ISBN: 3-540-34776-3.

[8] H. Kindler, D. Densow, and T. Fliedner. A pragmatic implementation of medical temporal reasoning for clinical medicine. *Computers in Biology and Medicine*, 28, pp. 105–120, 1998.

[9] H. Maeda, S. Asaoka and S. Murakami. Dynamical fuzzy reasoning and its application to system modeling. *Fuzzy Sets and Systems*, 80, pp. 101–109, 1996.

[10] M. Mucientes and W. Burgard. Multiple hypothesis tracking of clusters of people. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 692–697, 2006.

[11] M. Mucientes and A. Bugarín. People Detection with Learned Quantified Fuzzy Temporal Rules. In *Proc. 16th IEEE Conference on Fuzzy Systems (FUZZ-IEEE'07)*, pp. 1149–1154, 2007.

[12] M. Mucientes, R. Iglesias, C.V. Regueiro, A. Bugarín, and S. Barro. A fuzzy temporal rule-based velocity controller for mobile robotics. *Fuzzy Sets and Systems*, 134, pp. 83–99, 2003.

[13] M. Mucientes, R. Iglesias, C.V. Regueiro, A. Bugarín, P. Cariñena, and S. Barro. Fuzzy temporal rules for mobile robot guidance in dynamic environments. *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*, 31(3), pp. 391–398, 2001.

[14] J. Presedo, J.A. Vila, S. Barro, F. Palacios, R. Ruiz, A. Taddei, and M. Emdim. Fuzzy Modelling of the expert's knowledge in ECG-based ischaemia detection. *Fuzzy Sets and Systems*, 77, pp. 63–75, 1996.

[15] D. Qian. Representation and use of imprecise temporal knowledge in dynamic systems. *Fuzzy Sets and Systems*, 50, pp. 59–77, 1992.

[16] W.M.P. van der Aaslt, B.F. van Dongen, J. Herbst, L. Maruster, G. Schimm and A.J.M. Weijters. Workflow mining: A survery of issues and approaches. *Data Knowledge Engineering*, 47(2), pp. 237–267, 2003.

[17] A. van der Maas, A. ter Hofstede and P. de Vries Robbé. Formal description of temporal knowledge in case reports. *Artificial Intelligence in Medicine*, 16(3), pp. 251–282, 1999.

[18] J. Virant and N. Zimic, Attention to time in fuzzy logic. *Fuzzy Sets and Systems*, 82, pp. 39–49, 1996.

SISTEMAS BASADOS EN REGLAS DIFUSAS CON DATOS IMPRECISOS: NUEVOS RETOS

Luciano Sánchez[1] **Inés Couso**[2]

[1] Departamento de Informática, Universidad de Oviedo, luciano@uniovi.es
[2] Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, couso@uniovi.es

Resumen

En las bases de conocimiento difusas, los conjuntos borrosos representan conceptos imprecisos, contenidos en los antecedentes y consecuentes de las reglas. Un sistema basado en reglas difusas (SBRD) contiene, además de la base de conocimiento, un motor de inferencia, con interfaces que le permiten admitir y producir datos numéricos. En este trabajo discutiremos cómo alterar este motor de inferencia de forma que el SBRD pueda ser aplicado al procesamiento de datos borrosos.

Además, analizaremos los retos que plantea el uso de este nuevo tipo de información: cómo representar datos vagos mediante conjuntos borrosos, cómo preprocesar esta información, qué esquemas de razonamiento son más adecuados, qué tipos de algoritmos permiten dotar a los SBRD de capacidad de aprendizaje a partir de datos de baja calidad y cómo validar los resultados de estos sistemas.

Palabras Clave: Sistemas Basados en Reglas Difusas, Datos Imprecisos.

1. INTRODUCCIÓN

Las bases de conocimiento borrosas contienen reglas "SI-ENTONCES", cuyos antecedentes y consecuentes contienen sentencias basadas en la lógica borrosa. Un Sistema Basado en Reglas Difusas (SBRD) combina una base de conocimiento y un motor de inferencia. Este último contiene un sistema de inferencia, e interfaces de fuzzificación y defuzzificación [6]. Tales interfaces convierten datos numéricos en información borrosa, compatible con las reglas que componen la base de conocimiento, y transforman los resultados borrosos de la base de conocimiento en valores nítidos. Como es bien conocido, esta estructura permite que los SBRDs puedan aproximar una amplia familia de funciones con precisión arbitraria [5]. Comparados con otros aproximadores universales, los SBRDs tienen un valor adicional, porque balancean la precisión con la interpretabilidad linguistica [4], explotando el potencial de la lógica borrosa como un esquema de representación y cálculo de nociones inciertas o vagas.

Paradójicamente, los SBRDs, pese a estar basados en la lógica borrosa, no se aplican con frecuencia a datos imprecisos. La práctica común en SBRDs consiste en asociar conjuntos borrosas a palabras y usarlos para modelar afirmaciones vagas, pero hacer que los modelos que dependen de esas afirmaciones usen datos nítidos, y produzcan resultados nítidos. En este trabajo estudiaremos cómo superar esta limitación, y expondremos los nuevos retos que plantea el uso de datos de baja calidad en el contexto de los SBRDs.

La estructura de este documento es la siguiente: En la sección 2 se discute la representación de un dato vago mediante conjuntos borrosos, y se establecen los beneficios esperados del uso de datos imprecisos en problemas de clasificación y regresión. En la sección 3 se consideran algunas herramientas de análisis gráfico y preprocesamiento que permiten mejorar la calidad de la información imprecisa. En la sección 4 se estudia el diseño del motor de inferencia de un SBRDs que procesa datos vagos. En la sección 5 se recogen algunas medidas de precisión de tales sistemas. En la sección 6, el problema de validar los resultados de esos sistemas de forma comparada. Por último, en la sección 7 (conclusiones) se resumen todos los problemas abiertos.

2. REPRESENTACION DE DATOS DE BAJA CALIDAD

Los modelos basados en datos imprecisos no son exclusivos de la literatura relacionada con los sistemas borrosos, y constituyen un campo actual de investigación en varias disciplinas. Por ejemplo, en [14] se proponen ocho fuentes de información que pueden ca-

racterizarse mediante intervalos: informes de pasa-no llega, dígitos significativos, mediciones intermitentes, fallos en detección, datos censurados, datos agrupados, datos perdidos y total ignorancia.

Estos modelos tienen una fuerte relación con algunas tendencias recientes en estadística borrosa, donde el tratamiento formal de los datos vagos está basado en el concepto de Variable Aleatoria Borrosa (VAB). Una VAB es una función, que asocia cada resultado de un experimento aleatorio con un subconjunto borroso de la recta real. En general, puede considerarse que estos subconjuntos borrosos modelan la imprecisión en la observación del evento, mientras que la aleatoriedad en el experimento modela la variabilidad de los resultados [3, 10].

2.1. Definiciones de VAB

Hay diferentes definiciones de VAB, y la elección de una de ellas condiciona el uso conjunto de la lógica y la estadística borrosas necesario para aplicar SBRDs a datos vagos, por lo que a continuación se enumeran los modelos existentes en la literatura y se propone la adopción de uno de ellos.

2.1.1. Interpretación clásica

En [23], Puri y Ralescu consideran que las observaciones de algunos experimentos aleatorios no producen resultados numéricos, sino términos lingüísticos imprecisos. Si la VAB tiene un número finito de imágenes, pueden asignárseles valores de probabilidad. Por ejemplo, el resultado es "alto" con probabilidad 0.5, "medio" con probabilidad 0.25 y "bajo" con probabilidad 0.25, donde "alto", "medio" y "bajo" forman una partición borrosa.

Observe que, aunque esta definición considera que los resultados del experimento tienen un significado lingüístico, no necesariamente es compatible con el uso de estos términos en lógica borrosa. Por ejemplo, en el modelo anterior, la probabilidad de que el resultado sea el conjunto borroso "alto o medio" (entendiendo la disyunción como combinación de las pertenencias de ambas etiquetas con una t-conorma) es cero, porque el conjunto borroso asociado a esa expresión lógica no coincide exactamente con ninguna de las imágenes de la variable.

2.1.2. Modelo de Segundo Orden

Según [10], que a su vez extiende la definición de [20], una variable aleatoria borrosa representa un conocimiento impreciso acerca de una variable aleatoria clásica, que es llamada "variable aleatoria original". Es decir: supongamos que el resultado del experimento aleatorio es el valor ω, y que la observación del experimento aleatorio es un conjunto borroso con pertenencia $X(\omega)$. Entonces, la pertenencia de un punto x al conjunto borroso $X(\omega)$ representa el grado de posibilidad de la afirmación "La imagen de ω es x". O lo que es lo mismo, una VAB es un conjunto borroso que asocia una pertenencia a cada una de las posibles variables aleatorias clásicas. Dicha pertenencia indica el grado de certeza que se tiene acerca de que esta variable sea la original.

2.1.3. Modelo de Primer Orden

De acuerdo con [2], se supone que, si el resultado del experimento aleatorio es ω, la probabilidad de percibir el valor x es $p(x|\omega)$, y que esta segunda probabilidad no es completamente conocida: sólo se sabe que está dominada por la posibilidad $\Pi(x|\omega)$. En este modelo, la función de pertenencia $X(\omega)$ de la VAB es la posibilidad $\Pi(x|\omega)$. En otras palabras, una variable aleatoria borrosa se corresponde con una familia de distribuciones de probabilidad. Esto, como se indica en la figura 1, es una generalización directa de los modelos estocásticos usados habitualmente en ingeniería para caracterizar el ruido en las observaciones.

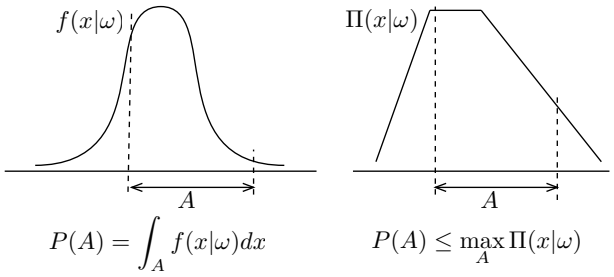


Figura 1: Modelos estocástico y posibilístico de primer orden del error de observación. **Izquierda:** Modelo estocástico de una observación con ruido. Se conoce de forma precisa la probabilidad de percibir un valor contenido en el conjunto A, cuando el resultado del experimento es ω. **Derecha:** modelo posibilístico de una observación vaga: se conoce un intervalo de valores que contiene a $P(A)$.

2.1.4. Discusión

En general, el modelo clásico no siempre es válido para representar datos vagos en el contexto de los SBRDs, ya que es problemático conciliar el uso de términos lingüísticos desde el punto de vista de la lógica borrosa con esta definición, como se ha visto en el ejemplo anterior. Por el contrario, tanto el modelo de primer orden como el modelo de segundo orden son adecuados. El tratamiento del modelo de primer orden tiene una implementación considerablemente más sencilla, y por

esta razón es preferible. Adicionalmente, de acuerdo con este modelo, los α-cortes de los conjuntos borrosos son intervalos de confianza del valor percibido de la variable, lo que permite representar mediante conjuntos borrosos algunos tipos de datos de baja calidad que tradicionalmente no se consideraban borrosos. Por ejemplo, en [24] se utilizaron conjuntos borrosos para reconciliar diferentes opiniones acerca de un valor, y en [22] se usó esta propiedad para asimilar las coordenadas proporcionadas por un GPS a conjuntos borrosos.

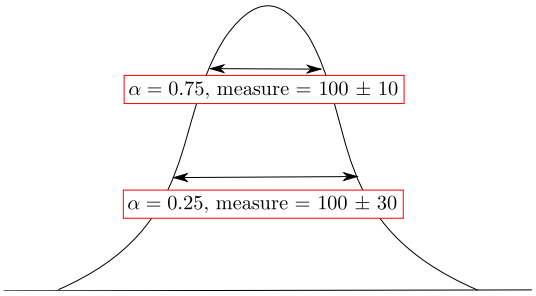


Figura 2: Visión alternativa del modelo de primer orden: conjunto borroso como familia de intervalos. Cada α-corte contiene al resultado del experimento con probabilidad $\geq 1 - \alpha$.

2.2. Beneficios esperados del uso de datos imprecisos

Una utilidad obvia de la extensión de los SBRDs a datos de baja calidad es la posibilidad de solucionar problemas en que no se dispone de mejor información. No obstante, esta técnica también proporciona ventajas en algunos problemas con datos nítidos. Por ejemplo, observe la figura 3:

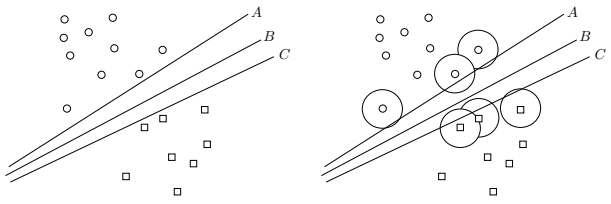


Figura 3: Los clasificadores A, B, C tienen error cero con respecto a los centros de los datos, pero sólo el clasificador 'B' tiene error cero si se tiene en cuenta la imprecisión de las mediciones.

Tres superficies de decisión A, B y C separan las dos categorías de la muestra, sin errores. Ahora bien, si cada dato se envuelve en un círculo, sólo la superficie B separa las dos clases, porque es la que más dista de los puntos más cercanos de la muestra. Esto parece indicar que la adición artificial de imprecisión a datos nítidos, y la optimización de un clasificador sobre los datos extendidos, podría producir una mejora en la capacidad de generalización. También podría existir una relación entre el diseño de un clasificador de margen mínimo y el diseño de un clasificador con datos imprecisos.

Por otra parte, en la figura 4 se muestra un problema de regresión, donde también podrían encontrarse ventajas. Tanto en la parte izquierda de la figura, como en la derecha, los modelos de regresión tienen error cuadrático medio cero. Sin embargo, si cada uno de los datos se envuelve en un intervalo, el error cuadrático es más específico en el modelo de la izquierda. De nuevo, la extensión de los datos sería equivalente a añadir al error cuadrático medio un término de regularización, lo que tal vez, como en el caso anterior, pueda explicarse estudiando la relación con las máquinas de vectores de soporte.

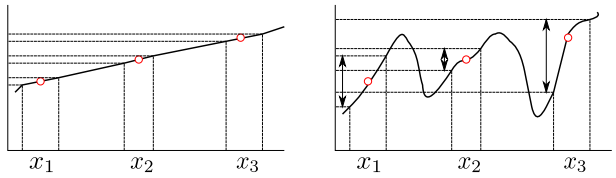


Figura 4: Ambos modelos de regresión tienen error cuadrático cero sobre los datos nítidos, pero el modelo de la izquierda tiene un error más específico sobre los datos extendidos, porque su pendiente en los puntos de la muestra es menor.

2.3. Problemática

El tratamiento de datos imprecisos es interesante tanto en problemas nítidos como en problemas con datos imprecisos, u otros que puedan codificarse mediante una pertenencia borrosa. Este tratamiento depende del concepto de dato vago (entendido como instancia de una VAB), y afecta al diseño del motor de inferencia del SBRD. Pero además, todas las fases del desarrollo del sistema necesitan adaptarse al nuevo tipo de datos: para diseñar un SBRD se necesitan nuevas técnicas de preprocesamiento y mejora de calidad de los datos, aprendizaje de bases de conocimiento y validación de resultados. Estos cambios se estudian en las siguientes secciones.

3. PREPROCESAMIENTO DE DATOS VAGOS

El preprocesado de datos imprecisos ha sido estudiado de forma desigual. Por una parte, no tenemos constancia de que se hayan desarrollado algoritmos de selección de instancias vagas. Por otra, hay varias extensiones borrosas de la selección de características y transformaciones de datos en problemas de clasificación y regresión. Ambas partes se discuten a continuación.

3.1. Selección de instancias

En problemas con datos precisos, la selección de instancias se basa en detectar las zonas de mayor densidad entre los ejemplos: una instancia puede ser redundante si existen en el fichero de datos otros ejemplos similares. Por el contrario, en problemas con datos vagos, es posible que una instancia intrínsicamente aporte poca información (como caso extremo, un ejemplo en que todos los datos de entrada son desconocidos siempre es redundante). Sería posible investigar la relación entre los diferentes costes de obtención de una instancia con distintas tolerancias, y el aporte de información que en cada caso esa instancia le haría a la base de conocimiento final.

3.2. Selección de características y transformaciones

En problemas de clasificación, puede determinarse el subconjunto más relevante de características mediante algoritmos basados en medidas de dependencia entre variables (como por ejemplo [27, 31]) o con técnicas de búsqueda (como [33]). En problemas de regresión es habitual emplear transformaciones, como el análisis de componentes principales y el análisis factorial (desarrollados en [15, 21] para datos borrosos). Las versiones borrosas de otros enfoques modernos, como el análisis de componentes independientes, están dirigidas a mejorar la robustez cuando se trabaja con datos nítidos [19] y no admiten datos vagos; su extensión es una tarea pendiente. Por el contrario, el escalado multidimensional sí ha sido generalizado al caso difuso [12, 16, 34], de forma que la proyección de los datos no sólo incluye información acerca de las similaridades entre las instancias sino también acerca de la vaguedad de cada una (ver figura 5).

4. DISEÑO DEL MOTOR DE INFERENCIA

La regla de inferencia composicional (extensión cilíndrica del dato de entrada, intersección con el grafo borroso codificado en la base de conocimiento, proyección en la salida) no necesariamente es compatible con la interpretación posibilística de un conjunto borroso vista en la sección 2. Para clarificar nuestra opinión, observe la siguiente base de conocimiento:

```
SI X ES PEQUEÑO ENTONCES Y ES GRANDE
SI X ES GRANDE ENTONCES Y ES PEQUEÑO
```

donde las pertenencias de los valores lingüísticos son $\texttt{PEQUEÑO}(t) = 1 - t$, y $\texttt{GRANDE}(t) = t$ para $t \in [0, 1]$. El grafo borroso de este sistema es

$$G(x, y) = \text{máx}\{\text{mín}\{x, 1 - y\}, \text{mín}\{1 - x, y\}\}. \quad (1)$$

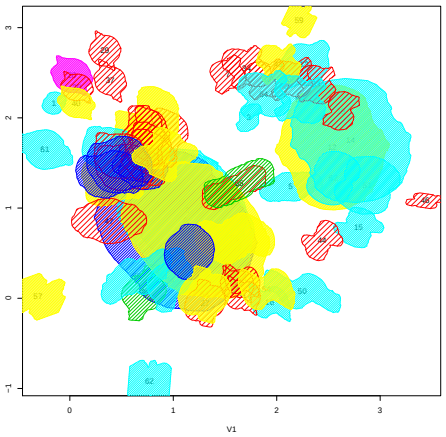


Figura 5: Extensión de MDS al caso borroso: se muestra gráficamente la similaridad entre las instancias y la vaguedad de cada una.

Por tanto, dada una entrada imprecisa X, la salida es el conjunto borroso

$$H_X(y) = \sup_x \text{mín}\{G(x, y), X(x)\}, \quad (2)$$

que no es normal (ver figura 6) y que por tanto no se corresponde con una distribución de posibilidad.

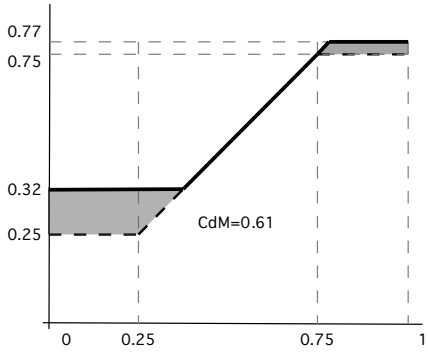


Figura 6: Inferencia con datos vagos (I): la curva con guiones es la imagen de la entrada crisp 0.25 mediante la regla de inferencia composicional. El valor defuzzificado es 0.62. La curva continua es la imagen del conjunto borroso triangular (0.15;0.25;0.35). El valor defuzzificado es 0.61

Para obtener una salida compatible con el modelo de imprecisión, puede extenderse la salida defuzzificada mediante el principio de extensión. Esta salida defuzzificada es

$$g(x_0) = \text{CdM}(H_{x_0}(y)) \quad (3)$$

donde CdM significa "centro de masas". La extensión tiene la forma que sigue:

$$F(X, y) = \sup\{\alpha \mid g^{-1}(y) \in X_\alpha\}. \quad (4)$$

La imagen de la misma entrada usada en la figura 6, ahora mediante la ecuación 4, puede verse en la figura 7. En esta segunda figura se aprecia que, con este tipo de inferencia, el que la entrada sea el valor vago $0{,}25 \pm 0{,}1$ implica que la salida tenga una tolerancia $\pm 0{,}05$.

El inconveniente de este método está, sin embargo, en la carga computacional que supone calcular la respuesta del sistema. Sería interesante desarrollar una versión de la regla de inferencia composicional, y tal vez un nuevo conjunto de operadores, que permitiesen producir una aproximación razonable al resultado mostrado en esta última figura. En este sentido, creemos que es posible que haya puntos en común entre los mecanismos de razonamiento de los SBRDs y los algoritmos utilizados para realizar inferencia en redes credales [11].

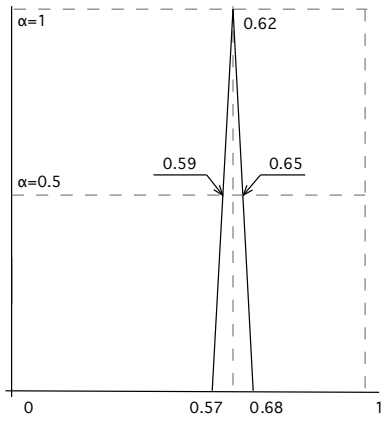


Figura 7: Inferencia con datos vagos (II): Imagen del conjunto borroso triangular (0.15;0.25;0.35) mediante la ecuación 4. El punto modal es la imagen del punto modal de la entrada.

5. PRECISION DE UN SBRD. OPTIMIZACIÓN GENÉTICA

La medición de la calidad de un SBRD se basa en la interpretabilidad lingüística y en la precisión del mismo. Sólo las medidas de precisión se ven afectadas cuando los datos son imprecisos. En ese caso, el modelo de VAB es determinante en la estimación de la precisión.

En general, se considerará que la precisión de un SBRD sobre datos vagos es, a su vez, un valor conocido de forma imprecisa. En trabajos anteriores [24, 28, 29] hemos adaptado la definición de la varianza de una VAB, que para el modelo de primer orden propuesto en la sección 2 se reduce a un intervalo [7, 10]. En [26] se propuso otro criterio específico para problemas de clasificación, que produce un número borroso discreto. En ambos casos, los algoritmos de aprendizaje de reglas borrosas que optimicen este criterio deben optimizar una función intervalo-valorada o borrosa [29, 30].

5.1. Sistemas borroso-genéticos

Los algoritmos genéticos han probado ser muy eficaces para solucionar el problema del aprendizaje de un SBRD a partir de datos nítidos [6, 17, 18]. En trabajos anteriores, hemos conseguido obtener algunos resultados que permiten formular sistemas borroso-genéticos a partir de información vaga, siguiendo los enfoques Pittsburgh [24], Iterativo [25] y GCCL [32] en problemas de modelado.

Por el contrario, hasta el momento no se ha publicado ninguna implementación del cuarto enfoque, los algoritmos tipo Michigan. Tampoco tenemos conocimiento de que los sistemas borroso-genéticos hayan sido aplicados a problemas de clasificación con datos imprecisos.

6. VALIDACION

El uso de medidas de calidad intervalares o borrosas plantea algunos problemas con la comparación estadística de los resultados de diferentes algoritmos. Aunque existen numerosos trabajos teóricos de estadística borrosa acerca de test estadísticos según el modelo clásico de probabilidad, los estudios relacionados con los modelos de primer o segundo orden son menos numerosos [13, 20, 35]. Es interesante observar que estos dos últimos modelos, cuando la observaciones son intervalo-valoradas, guardan relación con los trabajos en probabilidades imprecisas [3] y con algunos enfoques recientes en modelos econométricos [1]. Recientemente, se está trabajado en la base teórica de los tests de remuestreo que se podrían aplicar en un futuro a la comparación de las calidades de diferentes algoritmos [8, 9].

Por otro lado, es interesante destacar que no existe un catálogo de problemas sobre los que se puedan evaluar las propiedades de diferentes algoritmos. Hasta la fecha, cada autor propone sus propios problemas de validación. Sería deseable que se elaborase un catálogo de problemas con diferentes complejidades y con resultados conocidos.

7. CONCLUSIONES Y PROBLEMAS ABIERTOS

En este estudio hemos mostrado las líneas generales del uso de datos vagos con sistemas basados en reglas borrosas. Se ha propuesto combinar las herramientas de la lógica borrosa con las de la estadística borrosa, aprovechando resultados recientes acerca de la representación de la imprecisión en la observación de los

resultados de un experimento aleatorio.

La principal conclusión de este estudio es que muchos de los problemas del uso de datos vagos con SBRDs están sin resolver de forma satisfactoria. A modo de resumen, los retos mencionados en este trabajo han sido:

1. El estudio teórico de las relaciones entre los problemas con imprecisión añadida y las máquinas de vectores de soporte.
2. El desarrollo de las técnicas de análisis de componentes principales e independientes para datos imprecisos.
3. Nuevos algoritmos de selección de características e instancias para datos vagos.
4. Nuevos mecanismos de razonamiento borroso que permitan aproximar la solución basada en el principio de extensión.
5. Otras medidas de calidad para clasificadores y modelos.
6. Estudio teórico y práctico de algoritmos de optimización de funciones intervalo-valoradas y borrosas.
7. Tests estadísticos que permitan comparar la calidad de dos algoritmos sobre datos vagos.
8. Bancos de prueba para comparar los nuevos algoritmos.

Agradecimientos

Este trabajo ha sido soportado por el Ministerio de Ciencia e Innovación, a través del proyecto de investigacón TIN2005-08036-C05-05.

Referencias

[1] Beresteanu, A., and Molinari, F.: Asymptotic Properties for a Class of Partially Identified Models, Econometrica, 2008 (En prensa)

[2] Baudrit, C., Couso, I., Dubois, D. Joint propagation of probability and possibility in risk analysis: Towards a formal framework, International Journal of Approximate Reasoning, 45, pp 82–105. 2007.

[3] Baudrit, C., Dubois, D., Perror, N. Representing parametric probabilistic models tainted with imprecision. Fuzzy Sets and Systems, 2008 (En prensa)

[4] Casillas, J, Cordon, O., Herrera, F., Magdalena, L. Interpretability issues in fuzzy modeling. Springer. Heidelberg, Germany. 2003.

[5] Castro, J., Delgado, M. Fuzzy Systems with Defuzzification are Universal Approximators. IEEE Trans. on Syst., Man, and Cybern 26(1), pp 149–152. 1996.

[6] Cordón, O., Herrera, F., Hoffmann, F., Magdalena, L. Genetic fuzzy systems: Evolutionary tuning and learning of fuzzy knowledge bases. World Scientific Publishing Company, Singapore. 2001.

[7] Couso, I., Dubois, D. Montes, S., Sánchez, L.,On various definitions of the variance of a fuzzy random variable. ISIPTA 07 - Fifth International Symposium on Imprecise Probability: Theories and Applications, 135-144. 2007

[8] Couso, I., Sánchez, L., Mark-recapture techniques in statistical tests for imprecise data. International Journal of Approximate Reasoning (sometido)

[9] Couso, I., Sánchez, L., Defuzzification of fuzzy p-values. Fourth International Workshop on Soft methods in probability and statistics SPMS 08 (sometido)

[10] Couso, I., Sánchez, L. Higher Order Models for Fuzzy Random Variables. Fuzzy Sets and Systems 159 (3) pp. 237-258. 2008

[11] Ide, J. S., Cozman, F. G., Approximate algorithms for credal networks with binary variables. International Journal of Approximate Reasoning 48(1), pp 275-296. 2008.

[12] Denoeux, T., Masson, M. H., Multidimensional scaling of interval-valued dissimilarity data. Pattern Recognition Letters 21, pp 83-92. 2000.

[13] Denoeux, T, Masson, M. H. , Hebert, P.A. Nonparametric Rank-Based Statistics and Significance Test for Fuzzy Data, Fuzzy Sets and Systems, 153. pp 1-28. 2005.

[14] Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W., Ginzburg, L., Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty. Tech. Rep. SAND2007-0939, Sandia National Laboratories, 2007.

[15] Giordani, P., Kiers, H., Principal Component Analysis of symmetric fuzzy data. Computational Statistics and Data Analysis 45(3). pp 519-548. 2004.

[16] Hebert, P. A., Masson, M. H., Denoeux, T. Fuzzy multidimensional scaling. Computational Statistics and Data Analysis 51. pp 335-359. 2006.

[17] Herrera, F. Genetic Fuzzy Systems: Status, Critical Considerations and Future Directions. International Journal of Computational Intelligence Research. 1 (1). pp 59-67. 2005.

[18] Herrera, F. Genetic Fuzzy Systems: Taxonomy, Current Research Trends and Prospects. Evolutionary Intelligence 1. pp 27-46. 2008.

[19] Honda, K., Ichihashi, H. Fuzzy local independent component analysis with external criteria and its application to knowledge discovery in databases. International Journal of Approximate Reasoning 42(3), pp. 159-173. 2006.

[20] Kruse, R., Meyer, K. D., Statistics with vague data. D. Reidel Publishing Company. 1987.

[21] Nakamori, Y., Watada, J. Factor analysis for fuzzy data. Sixth Intl. Conf. Fuzzy Systems, pp. 1115-1120. 1997.

[22] Otero, A., Otero, J., Sanchez, L., Villar, J. R. Longest path estimation from inherently fuzzy data acquired with GPS using genetic Algorithms. Proc. of the 2006 IEEE International Conference on Evolutionary Fuzzy Systems, Ambleside, UK. pp 30-305. 2006.

[23] Puri, M. L., Ralescu, D. Fuzzy random variables. J. Math. Anal. Appl., 114 pp 409-422, 1986.

[24] Sánchez, L. Couso, I. Casillas, J. A multiobjective genetic fuzzy system with imprecise probability fitness for vague data. Proc. of the 2nd International Symposium on Evolving Fuzzy Systems (EFS 2006), 131-136, 2006

[25] Sánchez, L., Otero, J., Villar, J. R. Boosting of fuzzy models for high-dimensional imprecise datasets. Proc. of the 2006 International Conference in Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris, France. pp 1965-1973. 2006.

[26] Sánchez, L., Couso, I., Advocating the use of imprecisely observed data in Genetic Fuzzy Systems. IEEE Trans Fuzzy Sys 15(4). pp 551-562. 2007.

[27] Sánchez, L.; Suarez, M. R.; Villar, J. R.; Couso, I. Some results about mutual information-based feature selection and fuzzy discretization of vague data. FUZZ-IEEE 2007. pp 23-26. 2007

[28] Sánchez, L., Otero, J., Learning Fuzzy Linguistic Models from low quality data by genetic algorithms. FUZZ-IEEE 2007. pp 23-26. 2007

[29] Sánchez, L. Couso, I. Casillas, J. Modelling Vague Data with Genetic Fuzzy Systems under a Combination of Crisp and Imprecise Criteria, in Proc. of the 2007 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making (MCDM'2007), pp. 30–37, Honolulu, Hawaii, USA. 2007

[30] Sánchez, L., Villar, J. R. Obtaining transparent models of chaotic systems with multiobjective simulated annealing algorithms. Information Sciences 178(4). pp 952-970. 2008

[31] Sánchez, L., Suárez, M. R., Villar, J. R., Couso, I. Mutual Information-based Feature Selection and Fuzzy Discretization of Vague Data. International Journal of Approximate Reasoning (Admitido)

[32] Sánchez, L., Otero, J., Couso, I. Obtaining Linguistic Fuzzy Rule-based Regression Models from Imprecise Data with Multiobjective Genetic Algorithms. Soft Computing (Sometido)

[33] Sánchez, L., Villar, J. R., Couso, I. Genetic Feature Selection for Low Quality Data. IPMU08 (Admitido). Torremolinos, Málaga. 2008

[34] Sánchez, L., Palacios, A., Suárez, A., Couso, I. Selecting the most relevant tests from vague data in the early diagnosis of dyslexia IPMU08 (Admitido). Torremolinos, Málaga. 2008

[35] Zaffalon, M. Conservative rules for predictive inference with incomplete data, in: Fourth International Symposium on Imprecise Probabilities and their Applications, pp. 406-415. 2005.

REGLAS DE ASOCIACIÓN DIFUSAS: NUEVOS RETOS*

Miguel Delgado[1], **M. Dolores Ruiz**[1], **Daniel Sánchez**[1]

[1] Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada
ETSIIT, C/Periodista Daniel Saucedo Aranda s/n, 18071, Granada
{mdelgado}@ugr.es,{mdruiz,daniel}@decsai.ugr.es

Resumen

Las reglas de asociación difusas desde su aparición a finales de los años noventa han sufrido un gran desarrollo y se han aplicado en numerosas situaciones. Este trabajo pretende hacer una revisión de las aportaciones hechas hasta el momento y comentar posibles líneas de investigación en este área.

Palabras Clave: Reglas de asociación difusas, nuevos retos, medidas de interés, transacciones difusas.

1. INTRODUCCIÓN

Las reglas de asociación han recibido mucha atención y estudio desde su primera aparición en los trabajos de Agrawal et al. [1]. La idea de las reglas de asociación se originó en el contexto del análisis de la bolsa de la compra donde observaron la compra conjunta de varios artículos de la tienda. En lo sucesivo para referirnos a dichos artículos emplearemos el término *ítem* muy usado en la literatura. Las reglas de asociación materializan la observación de que habitualmente los clientes compran conjuntamente algunos artículos determinados.

De manera informal podríamos decir que las reglas de asociación son expresiones de la forma $X \rightarrow Y$, donde X e Y son conjuntos de items. Un ejemplo de regla de asociación sería $\{sobres, papel\} \rightarrow \{sellos\}$. Se han empleado numerosas medidas para evaluar la validez o la representatividad de las reglas de asociación. Las más usadas son el soporte y la confianza. El *soporte*, es el porcentaje de transacciones que contienen a X y a Y; y la *confianza*, mide la frecuencia de aparición de Y cuando la transacción contiene a X.

Las técnicas de minería de datos y en particular las reglas de asociación, pretenden identificar conocimiento que sea válido, novedoso, útil y *comprensible*. Esto justifica la importancia de trabajar con representaciones del conocimiento que sean *semánticamente significativas* para el usuario. El usuario es la pieza clave en la búsqueda de conocimiento. Este debe guiar al sistema y juzgar la calidad de las relaciones obtenidas para determinar si el conocimiento obtenido es *interesante* o no. Para representar este tipo de conocimiento y los posibles tipos de imprecisión en la información, parece adecuado el uso de la Teoría de Subconjuntos Difusos propuesta por L.A. Zadeh [30].

Este trabajo presenta una breve revisión sobre las aportaciones más importantes hechas sobre reglas de asociación difusas así como dar a conocer los hitos más importantes que se encuentran en la literatura más reciente.

2. DE REGLAS CUANTITATIVAS A REGLAS DIFUSAS

Los primeros trabajos que empezaron a hablar de reglas de asociación difusas se centraban en el uso de etiquetas difusas para conseguir reglas de

*Este trabajo ha sido parcialmente sufragado por los proyectos TIN2006-15041-C04-01 y TIN2006-07262.

asociación cuantitativas. Las reglas de asociación cuantitativas se basan en dividir el dominio del atributo en intervalos y después extraer las reglas cuyos items son los pares $\langle atributo, intervalo\rangle$ en lugar de $\langle atributo, valor\rangle$. Pero al dividir el dominio en un número fijo de intervalos surge el llamado *problema de la frontera* en el que podemos rechazar algunos intervalos interesantes al excluir algunos elementos potenciales cerca de su frontera. (Ver [18]).

Para solucionar este problema algunos autores han propuesto dividir el dominio del atributo en intervalos solapados. Pero con la Teoría de subconjuntos difusos, podemos evitar el problema de los bordes del intervalo utilizando conjuntos difusos en lugar de intervalos. De esta forma un elemento pertenecerá a un conjunto difuso con un grado de pertenencia en el intervalo $[0, 1]$. Además podemos dotar a dichos conjuntos de una semántica significativa para el entendimiento del usuario utilizando para ello términos lingüísticos.

En esta línea se encuentran los primeros trabajos [19], [3], [18] y [15] que datan de finales de los años 90. A continuación desarrollaremos la propuesta hecha en [15] y [18] basada en la extracción de reglas de asociación difusas en bases de datos relacionales crisp.

Sea $D = \{t_1, \ldots, t_n\}$ una base de datos transaccional, y sea $ER = \{A_1, \ldots, A_m\}$ el conjunto de todos los atributos que aparecen en D. Cada atributo tendrá asociado varios conjuntos difusos que darán una descripción del rango de valores del atributo. Sea $F_{A_k} = \{a_k^1, \ldots, a_k^l\}$ el conjunto de todos los conjuntos difusos asociados al atributo A_k. Con esta notación un ítem será de la forma $\langle A_k, a_k^j\rangle$ con $A_k \in ER$ y $a_k^j \in F_{A_k}$. Llamaremos itemset a un conjunto de items de la forma $\langle A_i, a_i\rangle \cup \cdots \cup \langle A_j, a_j\rangle$ que notaremos por $\langle X, A\rangle$ donde X contiene todos los atributos y A los conjuntos difusos seleccionados para cada atributo en X.

Por ejemplo, si el atributo *Sueldo* tiene tres conjuntos difusos asociados: *alto*, *medio* y *bajo*, tendremos que $F_{Sueldo} = \{alto, medio, bajo\}$. Dichos conjuntos difusos y sus correspondientes funciones de pertenencia vendrán dados por expertos en dicho dominio.

Definición 1. [15] Una *regla de asociación difusa* es una expresión de la forma:

$$\text{Si } X \text{ es } A \text{ entonces } Y \text{ es } B,$$

donde $X, Y \subseteq ER$, es decir, $X = \{x_1, \ldots, x_p\}$ e $Y = \{y_1, \ldots, y_q\}$ son conjuntos de atributos disjuntos y $A = \{a_{x_1}, \ldots, a_{x_p}\}$ y $B = \{b_{y_1}, \ldots, b_{y_q}\}$ contienen los conjuntos difusos asociados a los correspondientes atributos de X e Y.

A la primera parte de la regla 'X es A' se le llama antecedente y 'Y es B' es el consecuente de la regla. El significado de la regla es que cuando se *satisface* que X es A entonces Y es B también se cumple. Aquí satisfacer significa que hay un número suficiente de transacciones que contribuyen a superar un umbral mínimo especificado por el usuario. De ahora en adelante notaremos dicha regla por $\langle X, A\rangle \rightarrow \langle Y, B\rangle$.

Para calcular el soporte del itemset $\langle X, A\rangle$ usan la siguiente fórmula:

$$FSop(X, A) = \frac{\sum_{t_i \in D} \prod_{x_j \in X} \mu^i_{a_j}(x_j)}{|D|} \qquad (1)$$

donde $|D|$ es el número de transacciones de la base de datos y $\mu^i_{a_j}(x_j)$ es el grado de pertenencia del atributo x_j en la transacción i-ésima al conjunto difuso a_{x_j}.

Observemos que en la fórmula (1) podríamos utilizar en lugar del producto cualquier t-norma. De forma análoga al caso crisp, para calcular el soporte y la confianza de la regla de asociación difusa $\langle X, A\rangle \rightarrow \langle Y, B\rangle$, usaremos las fórmulas:

$$\begin{aligned} FSop(\langle X, A\rangle \rightarrow \langle Y, B\rangle) &= FSop(Z, C) \\ FConf(\langle X, A\rangle \rightarrow \langle Y, B\rangle) &= \frac{FSop(\langle Z, C\rangle)}{FSop(\langle X, A\rangle)} \end{aligned} \qquad (2)$$

donde $Z = X \cup Y$ y $C = A \cup B$. Además del soporte y la confianza, otras medidas del cumplimiento para reglas de asociación han sido tenidas en cuenta para su extensión al caso difuso.

El resto de esta sección tratará de mostrar un enfoque distinto al anterior para tratar con la incertidumbre en la base de datos a analizar. En los trabajos [6] y [7] se generaliza el concepto de

transacción y regla de asociación para el caso en el que éstas sean difusas. Veremos los conceptos más importantes a continuación.

Definición 2. [6] Sea $I = \{i_1, \ldots, i_m\}$ un conjunto finito de items. Una transacción difusa es un subconjunto difuso no vacío $\tilde{\tau} \subseteq I$.

Para cada item $i \in I$ y cada transacción $\tilde{\tau}$ tendremos que i pertenece a $\tilde{\tau}$ con grado[1] $\tilde{\tau}(i)$, donde $\tilde{\tau}(i)$ es un número real en el intervalo $[0, 1]$.

Sea A un itemset de I, es decir, un subconjunto de items, en una transacción difusa $\tilde{\tau}$, se define el grado de pertenencia de $A \subseteq I$ a la transacción difusa $\tilde{\tau}$ como $\tilde{\tau}(A) = \min_{i \in A} \tilde{\tau}(i)$.

Usando la definición 2 una transacción será un caso especial de transacción difusa donde todos los items tienen valor 0 ó 1 según se encuentren o no en la transacción.

Ejemplo 1. Consideremos el conjunto de items $I = \{i_1, i_2, i_3, i_4\}$ y la tabla

	i_1	i_2	i_3	i_4
$\tilde{\tau}_1$	0	0.6	0.7	0.9
$\tilde{\tau}_2$	0	1	0	1
$\tilde{\tau}_3$	1	0.5	0.75	1
$\tilde{\tau}_4$	1	0	0.1	1
$\tilde{\tau}_5$	0.5	1	0	1
$\tilde{\tau}_6$	1	0	0.75	1

En particular, podemos observar que $\tilde{\tau}_2$ es una transacción crisp. Algunos grados de inclusión podrían ser: $\tilde{\tau}_1(\{i_3, i_4\})$ =0.7, $\tilde{\tau}_1(\{i_2, i_3, i_4\})$ =0.6 y $\tilde{\tau}_4(\{i_1, i_4\})$ =1.

Definición 3. [6] Sea I un conjunto de items, D un conjunto de transacciones difusas y $A, B \in I$ dos itemsets disjuntos, es decir, $A \cap B = \emptyset$. Una regla de asociación difusa $A \rightarrow B$ se cumple en D si y sólo si,

$$\tilde{\tau}(A) \leq \tilde{\tau}(B) \tag{3}$$

para todo $\tilde{\tau} \in D$, esto es, el grado de inclusión de B es mayor que el grado de inclusión de A para todas las transacciones difusas $\tilde{\tau}$ en D.

Esta definición preserva el significado de las reglas de asociación crisp ya que si necesitamos que $A \subseteq \tilde{\tau}$ en algún sentido, necesitaremos que también $B \subseteq \tilde{\tau}$ se cumpla, en nuestro caso sería $\tilde{\tau}(A) \leq \tilde{\tau}(B)$. De esta forma, como una transacción crisp es un caso especial de transacción difusa, una regla de asociación será un caso especial de regla de asociación difusa.

Usando este modelo se presenta en [6] y [7] una generalización de las medidas de soporte y confianza usando una aproximación semántica basada en la evaluación de sentencias cuantificadas. La evaluación de una sentencia cuantificada será un valor en el intervalo $[0, 1]$ que nos dirá el grado de cumplimiento de una expresión del tipo "Q de los F son G", donde F y G son dos subconjuntos de un conjunto finito X, y Q es un cuantificador relativo difuso. Ejemplos de cuantificadores son "la mayoría", "muchos", o "casi todos".

Definición 4. [6] Sea $A \subseteq I$. El *soporte* de A en D es la evaluación de la sentencia cuantificada

$$Q \text{ de los } D \text{ son } \tilde{\Gamma}_A \tag{4}$$

donde $\tilde{\Gamma}_A$ es un conjunto difuso en D definido como $\tilde{\Gamma}_A(\tilde{\tau}) = \tilde{\tau}(A)$.

Definición 5. [6] El *soporte* de la regla de asociación difusa $A \rightarrow B$ en el conjunto de transacciones difusas D es $FSop(A \cup B)$, es decir, la evaluación de la sentencia cuantificada

$$Q \text{ de los } D \text{ son } \tilde{\Gamma}_{A \cup B}. \tag{5}$$

Definición 6. [6] La *confianza* de la regla de asociación difusa $A \rightarrow B$ en el conjunto de transacciones difusas D es la evaluación de la sentencia cuantificada

$$Q \text{ de los } \tilde{\Gamma}_A \text{ son } \tilde{\Gamma}_B. \tag{6}$$

Conviene señalar que estas definiciones establecen familias de medidas de soporte y confianza, dependiendo del método de evaluación y del cuantificador Q elegido. En los trabajos [6] y [7] se justifica el uso del método GD para evaluar las sentencias y el cuantificador $Q = Q_M$ donde Q_M es la mayoría, es decir, $Q_M(x) = x$.

2.1. OTROS ENFOQUES

La evaluación de reglas de asociación difusas ha sido estudiada en la sección anterior de dos formas

[1] Observemos que $\tilde{\tau}(i)$ no es más que $\mu_{\tilde{\tau}}(i)$ donde $\mu_{\tilde{\tau}} : I \longrightarrow [0, 1]$ es la función de pertenencia del conjunto difuso $\tilde{\tau}$ sobre el referencial $I =$ { conjunto de items }.

distintas. Algunos autores basan el problema de la evaluación de reglas de asociación difusas en la generalización del caso crisp al difuso únicamente mediante la extensión de las llamadas medidas de interés sin tener en cuenta la propia naturaleza imprecisa del lenguaje humano y por tanto, de la representación del conocimiento.

También merecen mención algunos ámbitos de aplicación de la extracción de reglas difusas. En problemas de clasificación han sido utilizadas con muy buenos resultados, también se han extendido al caso de reglas temporales y han sido numerosos los campos en los que se han aplicado como por ejemplo en Aprendizaje Automático [17].

3. ÚLTIMOS TRABAJOS Y NUEVOS RETOS

Desde el surgimiento de las reglas de asociación difusas han sido numerosas las contribuciones en este campo y de muy diversa índole, así como el amplio abanico de posibilidades de aplicación. En esta sección comentaremos algunas de las más significativas y más recientes.

La búsqueda de nuevas medidas de interés[2] ha sido ampliamente estudiado y explotado desde la aparición de las reglas difusas. Las primeras contribuciones extendían algunas de las medidas ya propuestas para el caso crisp. Otras se centraban en el aspecto semántico y también en la forma de representación del conocimiento [16].

[13], [5] y [28] proveen una representación del conocimiento extraíble de la base de datos mediante una tabla de contingencia en la que se clasifican los datos (de acuerdo a dos itemsets pre-establecidos) en ejemplos positivos, negativos e irrelevantes[3] en [13] y una división parecida con intersecciones no vacías en ejemplos positivos, no-positivos, negativos y no-negativos en [5]. Esta forma de representar el conocimiento no es la primera vez que se usa en el ámbito de las reglas de asociación, pero sí es novedosa en cuanto, a que dichos trabajos proveen una forma sencilla y semántica de extender las medidas de interés del caso crisp al caso difuso. En esta línea se encuentran también los trabajos [26], [10], y [12]. En el último de ellos se presenta, usando una idea parecida a la anterior, una extensión del modelo estudiado para reglas de asociación crisp en [26] para el caso difuso. Dicha extensión es ampliamente estudiada en [12] y ofrece una nueva perspectiva para estudiar la evaluación de reglas de asociación difusas manteniendo las propiedades del caso crisp. Esta propuesta ofrece además la posibilidad de reutilizarse para otras áreas en las que el uso de las reglas de asociación difusas están siendo utilizadas para medir tendencias, gradualidad, temporalidad y otros aspectos debido a su sencillez y al tipo de algoritmo de extracción de reglas que es fácilmente extrapolable con tan sólo usar dicha representación.

Conviene destacar algunos trabajos que utilizan de alguna forma reglas difusas pero no las usan directamente para extraer el conocimiento. Estos trabajos se centran en el uso de medidas de interés subjetivas[4] para la evaluación de reglas de asociación. Más concretamente, el usuario establece sus creencias en forma de reglas difusas [22], [21] o bien en forma de taxonomía difusa [27], y después se establece un criterio para medir la similitud entre las reglas de asociación encontradas y las creencias establecidas por el usuario, de manera que aquellas reglas que más difieran de las creencias del usuario serán más interesantes.

En [14] se utiliza una ontología de dominio para extraer reglas de asociación difusas más interesantes. Dicha ontología viene predefinida por el usuario pero no contiene información imprecisa como en el caso anterior de la taxonomía difusa.

Entre los trabajos más recientes podemos mencionar el uso de reglas de asociación difusas para la extracción de dependencias graduales [24], [17]; para el refinamiento de consultas en minería web [8]; y para extraer conocimiento en el área de Biología [23]. Gracias a las buenas propiedades para extraer el conocimiento de una forma comprensible y que además se ajusta a la percepción de la realidad, las reglas de asociación difusas se

[2] En la literatura podemos encontrarnos con medidas de interés, medidas de calidad, y medidas para la evaluación de reglas de asociación difusas.

[3] En [28] se utiliza la misma partición pero con distinto nombre: ejemplos, contraejemplos y tuplas irrelevantes.

[4] Las medidas subjetivas son aquellas que tienen en cuenta la participación activa del usuario.

han modificado y adaptado para la extracción de conocimiento de tipo más específico. En [29] se han usado para expresar restricciones temporales; en [20] se propone un nuevo tipo de reglas de asociación difusas periódicas; y se han ideado mecanismos de derivación de reglas difusas en [4]. Algunas investigaciones han ido un paso más por delante y han propuesto las llamadas *asociaciones lingüísticas* [25].

Además se han hecho aportaciones paralelas para automatizar la extracción de reglas de asociación difusas [2] mediante mecanismos que escojan los umbrales para el soporte y la confianza, así como un proceso para automatizar la definición de las funciones de pertenencia de los conjuntos difusos.

La habilidad de representar conceptos graduales mediante conceptos difusos es una de las características clave de los conjuntos difusos. En [11] se pone de manifiesto el gran potencial de los conjuntos difusos para expresar y manejar cantidades de una forma coherente y semántica. En este trabajo se da un paso más en la utilización de reglas difusas y dependencias graduales en bolsas[5] y bolsas difusas, conceptos que generalizan al de transacción y al de transacción difusa.

En los párrafos anteriores se ha manifestado el amplio rango de actuación de las reglas de asociación difusas. Esto es debido a la gran potencia de la teoría de conjuntos difusos para representar incertidumbre, gradualidad y granularidad de conceptos, así como su gran interpretabilidad del mundo que nos rodea. Esta última característica ha sido tomada en cuenta a la hora de incorporar el conocimiento del usuario para dotar a los sistemas de capacidad para decidir cuál es el conocimiento interesante para el usuario. Además hay que destacar la facilidad que ofrece en la agregación, combinación y fusión de la información. Pero aún no se le ha sacado el jugo suficiente a todas estas características.

Hemos comentado algunas propuestas que incluyen subjetividad en el campo de la minería de datos, más concretamente en el preprocesamiento de los datos, bien usando taxonomías difusas o bien expresando el conocimiento del usuario mediante reglas difusas. Pero no es suficiente, habría que dar un paso más para integrar en el proceso completo el conocimiento del usuario y obtener realmente aquellas reglas, por supuesto difusas, que reflejen realmente el conocimiento novedoso, útil y de forma comprensible para el usuario.

Además la existencia de diferentes tipos de reglas difusas sugiere que en cada momento, según las necesidades del usuario, se debería guiar el proceso para obtener el conocimiento que interese extraer. Por tanto, parece comprensible que la evaluación del cumplimiento de la asociación no debería ser independiente de su interpretación. De esta forma queda reflejado que debería ser decisión del usuario (dependiendo del conocimiento que se quiera extraer) usar un tipo de evaluación u otro que se ajuste a sus exigencias.

Más aún, puesto que las medidas de evaluación pueden ser generalizadas de diversas formas, sería necesario buscar un soporte lógico o una base axiomática que refleje la elección de la medida específica. En esta línea se encuentran los trabajos [26], [10] para reglas de asociación crisp, y [5], [28], [13] y [12] para reglas de asociación difusas, que intentan establecer una ampliación del modelo crisp al difuso.

Una de las líneas de trabajo en el ámbito no difuso de las reglas de asociación se centra en el descubrimiento de reglas, patrones o tendencias que son atípicas [9]. Puesto que el usuario le interesa el conocimiento novedoso, esta es una buena línea en la que trabajar puesto que extrae un conocimiento no existente que puede ser de utilidad. Estas reglas atípicas o peculiares podrían extraerse usando reglas de asociación difusas para una mejor comprensión del conocimiento extraído.

En cuanto a las medidas de interés para las reglas de asociación difusas un nuevo reto que tenemos es una evaluación acorde a la propia semántica que encierran dichas reglas, que debería ser mediante la participación del usuario. Incluso sería bueno crear distintos filtros de extracción de reglas para ir extrayendo en cada paso el conocimiento de utilidad en cada situación.

Aunque han sido numerosas las áreas de conocimiento en las que han sido usadas, las buenas propiedades de la teoría de conjuntos difusos

[5] Una bolsa podría definirse como un conjunto de items donde éstos pueden repetirse.

para expresar el conocimiento de forma comprensible para el usuario se hace patente que aún hay que buscar nuevas áreas de aplicabilidad y adaptar el tipo de regla o el tipo de evaluación que necesitemos en cada momento.

Referencias

[1] R. Agrawal, T. Imielinski, and A. Swami. Mining associations between sets of items in massive databases. In *ACM-SIGMOD International Conference on Data*, pages 207–216, 1993.

[2] R. Alcala, J. Alcala-Fdez, M.J. Gacto, and F. Herrera. Genetic Learning of Membership Functions for Mining Fuzzy Association Rules. *Fuzzy Systems Conference. FUZZ-IEEE.*, pages 1–6, 2007.

[3] W.H. Au and K.C.C. Chan. Mining fuzzy association rules. In *Proc. 6th International Conference Information Knowledge Management*, pages 209–215, 1997.

[4] S. Ayouni and S.B. Yahia. Extracting compact and information lossless set of fuzzy association rules. *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pages 1–6, 2007.

[5] M. De Cock, C. Cornelis, and E.E. Kerre. Elicitation of fuzzy association rules from positive and negative examples. *Fuzzy Sets and Systems*, 149(1):73–85, 2005.

[6] M. Delgado, N. Marín, D. Sánchez, and M.A. Vila. Fuzzy association rules: General model and applications. *IEEE Transactions on Fuzzy Systems*, 11(2):214–225, 2003.

[7] M. Delgado, N. Marín, D. Sánchez, and M.A. Vila. Mining fuzzy association rules: An overview. *BISC International Workshop on Soft computing for Internet and Bioinformatics*, pages 351–374, 2003.

[8] M. Delgado, M.J. Martín-Bautista, D. Sánchez, J.M. Serrano, and M.A. Vila. *Fuzzy Association Rules for Query Refinement in Web Retrieval.* Springer.

[9] M. Delgado, M.D. Ruiz, and D. Sánchez. A logic approach for exceptions and anomalies in association rules. *Mathware & Soft Computing*, Submitted, 2007.

[10] M. Delgado, M.D. Ruiz, and D. Sánchez. Studying interest measures for association rules through a logical model. *IJUFKS*, Submitted, 2007.

[11] M. Delgado, M.D. Ruiz, and D. Sánchez. Pattern extraction from bag databases. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, To appear, 2008.

[12] M. Delgado, M.D. Ruiz, and M.A. Vila. Un modelo lógico para la representación de reglas de asociación difusas mediante niveles de restricción. In *ESTYLF'08*, Mieres, Spain, 2008, submitted.

[13] D. Dubois, E. Hüllermeier, and H. Prade. A systematic approach to the assessment of fuzzy association rules. *Data Min. Knowl. Discov.*, 13(2):167–192, 2006.

[14] Z. Farzanyar, M. Kangavari, and S. Hashemi. A New Algorithm for Mining Fuzzy Association Rules in the Large Databases Based on Ontology. *Proceedings of the Sixth IEEE Int. Conference on Data Mining-Workshops*, pages 65–69, 2006.

[15] A. Gyenesei. A fuzzy approach for mining quantitative association rules. *Acta Cybern.*, 15(2), 2001.

[16] E. Hüllermeier. Fuzzy association rules: Semantic issues and quality measures. *LNCS*, 2206:380–391, 2001.

[17] E. Hüllermeier. Fuzzy sets in machine learning and data mining. *Applied Soft Computing, to appear*, 2008.

[18] C.M. Kuok, A.W. Fu, and M.H. Wong. Mining fuzzy association rules in databases. *SIGMOD Record*, 27(1):41–46, 1998.

[19] J.H. Lee and H.L. Kwang. An extension of association rules using fuzzy sets. In *IFSA'97*, Prague, Czech Republic, 1997.

[20] W.J. Lee, J.Y. Jiang, and S.J. Lee. Mining fuzzy periodic association rules. *Data & Knowledge Engineering*, 2008.

[21] B. Liu, W. Hsu, S. Chen, and Y. Ma. Analyzing the subjective interestingness of association rules. *IEEE Intelligent Systems*, 15(5):47–55, 2000.

[22] B. Liu, W. Hsu, L. Mun, and H. Lee. Finding interesting patterns using user expectations. *IEEE Transactions on Knowledge and Data Engeneering*, 11(6):817–832, 1999.

[23] F.J. Lopez, A. Blanco, F. Garcia, and A. Marin. Extracting Biological Knowledge by Fuzzy Association Rule Mining. *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pages 1–6, 2007.

[24] C. Molina, J.M. Serrano, D. Sánchez, and M.A. Vila. Measuring variation strenght in gradual dependecies. In *EUSFLAT 2007*, To appear.

[25] V.Ñovak, I. Perfilieva, A. Dvorak, C. Chen, Q. Wei, and P. Yan. Mining pure linguistic associations from numerical data. *International Journal of Approximate Reasoning (to appear).*

[26] J. Rauch. Logic of association rules. *Applied Intelligence*, 22:9–28, 2005.

[27] B. Shekar and R.Ñatarajan. A framework for evaluating knowledge-based interestingness of association rules. *Fuzzy Optimization and decision Making*, 3:157–185, 2004.

[28] T. Sudkamp. Examples, counterexamples, and measuring fuzzy associations. *Fuzzy Sets and Systems*, 149(1):57–71, 2005.

[29] T. Sudkamp. Refinement of temporal constraints in fuzzy associations. *International Journal of Approximate Reasoning*, 48:23–35, 2008.

[30] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

Modelado Causal en Marketing mediante Aprendizaje no Supervisado de Reglas de Asociación Difusas

Albert Orriols-Puig[1] **Jorge Casillas**[2] **Francisco J. Martínez-López**[3]

[1] Grup de Recerca en Sistemes Intel·ligents, Enginyeria i Arquitectura La Salle,
Universitat Ramon Llull, 08022 Barcelona, aorriols@salle.url.edu
[2] Dpto. Ciencias de la Computación e Inteligencia Artificial,
Universidad de Granada, 18071 Granada, casillas@decsai.ugr.es
[3] Dpto. Comercialización e Investigación de Mercados,
Universidad de Granada, 18071 Granada, fjmlopez@ugr.es

Resumen

En marketing es habitual que se haga uso de modelos complejos, con múltiples relaciones de dependencia, para aproximarse al entendimiento diversos fenómenos de su interés. Este tipo de modelos son especialmente utilizados en la disciplina del comportamiento del consumidor. Están formados por lo que se conoce como variables latentes o constructos, esto es, variables que tienen un nivel de abstracción tal que no permiten medirse de manera unívoca por una única evaluación del consumidor, sino que precisan ser medidas por varias evaluaciones a través de escalas multi-ítem. Este hecho añade incertidumbre a las variables. Recientemente, algunos autores han planteado soluciones metodológicas basadas e inteligencia artificial que permiten trabajar con este tipo de dato, permitiendo, por tanto, aplicarse en el modelado de marketing. El trabajo presente sigue esta línea y propone afrontar el problema con Fuzzy-CSar, un sistema clasificador que extrae reglas de asociación difusas para el modelado del comportamiento del consumidor. La novedad de nuestra aproximación con respecto a técnicas existentes es que no se presupone ningún modelo causal previo, sino que se buscan relaciones fuertes entre las variables del sistema. Fuzzy-CSar se aplica a un problema de marketing real y los resultados se comparan con los obtenidos con un sistema GCCL diseñado expresamente para el mismo problema. Los resultados muestran las ventajas de evolucionar reglas de asociación y la competitividad de Fuzzy-CSar en general.

Palabras Clave: Lógica difusa; algoritmos genéticos; reglas de asociación; Marketing; comportamiento del consumidor

1 INTRODUCCIÓN

En las disciplinas de empresa y, en particular, en la de marketing, el uso de modelos ha sido tradicional para dirigir los procesos de análisis de datos. Como sabemos, este hecho supone que una estructura de relación entre los elementos del modelo (es decir, variables), conocida a priori, se utilice para, por medio de métodos analíticos de datos, describir o predecir la relación entre los mismos. Este enfoque de análisis de los datos respeta el procedimiento clásicamente establecido por el método científico, por el que un investigador previamente plantea unas hipótesis de relación entre variables para, posteriormente, utilizar métodos de análisis mediante los que contrastarlas en los datos y, finalmente, extraer conclusiones de los resultados obtenidos. Básicamente, en eso consiste el proceso de modelado en marketing, siendo habitual su utilización en el ámbito práctico/profesional por parte de los analistas de mercados. Esto es, se utiliza un modelo para dirigir el proceso de búsqueda de información en las bases de datos con el objeto de apoyar decisiones de marketing.

Pero, dicho lo anterior, ¿podría prescindirse de los modelos? Si bien la utilidad de los modelos es indudable, paradójicamente, su uso puede limitar el proceso de agregación de valor a los datos en ciertos tipos de problemas de decisión. Esto tiene implicaciones sobre todo en el ámbito práctico, en el de la toma de decisiones en marketing, donde los procesos de búsqueda de información están mucho menos condicionados por los procedimientos metodológicos, si lo comparamos con el académico, y mucho más por la calidad de la información generada. En otras palabras, en determinadas circunstancias, el conocimiento que generan métodos analíticos basados en estructuras relacionales de variables puede no tener toda la efectividad que requeriría un gestor de marketing. Esto se debería a que la estructura de base utilizada para dirigir el proceso de extracción de conocimiento de los datos y, conse-

cuentemente, la estructura de la información generada, no es adecuada para resolver el problema de decisión.

En particular, existen determinados escenarios de decisión, caracterizados por estar regular o, en su caso, deficientemente estructurados o delimitados, donde no existe información a priori de base para orientar el proceso de búsqueda de información, etc., en los que se precisa un cambio de mentalidad a la hora de desarrollar y aplicar los métodos analíticos de bases de datos. Para apoyar la toma de decisiones en marketing se utilizan multitud de métodos, importados generalmente de otras disciplinas, de entre las que destacan la estadística y la inteligencia artificial. En cierta medida, la evolución de los sistemas de apoyo a la gestión de marketing (SAGMk) se comprende por el tipo de métodos en los que se basaran principalmente. No obstante, desde finales de los 80, estos sistemas han mostrado una tendencia clara hacia lo que se denominan SAGMk basados en el conocimiento, gran categoría de agrupación de sistemas que se caracterizan por utilizar métodos analíticos basados en inteligencia artificial [1]. Sin duda, estamos convencidos de que el presente y futuro de las soluciones para este tipo de problemas de decisión que comentamos se encuentra en esta disciplina, concretamente en el campo de la extracción de conocimiento de bases de datos (KDD). En este sentido, el proceso de KDD y, en particular, en su etapa de minería de datos, se pueden distinguir dos grandes estilos de ejecución, así como un tercero fruto de la combinación de los mismos [7]: el enfoque "top-down", donde el experto parte de una estructura relacional a priori entre las variables que conforman la base de datos a analizar; y el enfoque "bottom-up", donde, en contraposición al anterior, no se hace uso de ninguna información a priori sobre la estructura, utilizándose métodos que exploran automáticamente los datos y que sugieren los patrones más significativos subyacentes a los mismos. Es lógica, por tanto, la asociación del primer y el segundo enfoque con los estilos de aplicación en minería de datos supervisada y no supervisada respectivamente.

Puede que los métodos de KDD basados en un enfoque "bottom-up" parezcan poco ortodoxos, e incluso poco útiles a primera vista. En cierta medida, se puede pensar que las soluciones de este tipo de métodos implicarían que habría que encontrar sentido a muchos patrones de información obtenidos, pudiendo resultar algo caótico. No obstante, no perdamos de vista el tipo de problemas de decisión donde sería más procedente su uso: los problemas poco o mal estructurados, donde es más difícil orientar el sentido de las decisiones, puesto que el conocimiento previo del que se dispone no permite abordar su resolución de manera inmediata. Bajo este escenario, los patrones extraídos de la base pueden ayudar precisamente a encontrar sentido al problema de información que se aborda. En cierta medida, los resultados de métodos de KDD diseñados con este enfoque actuarían estimulando la creatividad del usuario/tomador de decisiones de marketing que se apoye en ellos. Por tanto, utilizando la clasificación ORAC propuesta por Wierenga y Van Bruggen [5, 6] para la resolución de problemas en marketing, este tipo de soluciones podrían encuadrarse bajo el modo creativo. Este modo de resolución de problemas se caracteriza por propiciar un razonamiento divergente, esto es, abierto a todas las opciones posibles, contemplando incluso otras soluciones aparentemente no válidas o descabelladas si se aplicara un razonamiento, más ortodoxo, convergente.

En este trabajo proponemos un sistema de aprendizaje evolutivo, con una aproximación descriptiva, para bases de datos de marketing; en particular, la aplicación la realizamos sobre una base de datos que recoge variables de comportamiento del consumidor. En concreto, el diseño del algoritmo se ha realizado para un escenario de decisión caracterizado por los siguientes puntos: (1) se parte de un conjunto de variables cuya relación estructural se desconoce o se percibe deficiente, por lo que la utilización de un modelo para la extracción de conocimiento no es posible o, en su caso, conveniente; (2) el gestor de marketing está interesado especialmente en una variable objetivo, que quiere explicar en función de las anteriores, aunque la extracción de reglas que expliquen el comportamiento de otras variables también se considera valioso. Para cumplir este objetivo, proponemos Fuzzy-CSar un sistema clasificador basado en algoritmos genéticos que evoluciona un conjunto de reglas de asociación. El sistema usa una representación difusa que le permite tratar con datos imprecisos, los cuales caracterizan el problema de marketing abordado. De esta forma, cada variable latente se compone de un conjunto de variables observadas (ítemes) que aportan información parcial sobre la misma. Por tanto, la información de cada ítem debe ser agregada para definir la variable en cuestión. Para este fin, se usa el emparejamiento multi-ítem definido en [3]. Los resultados experimentales y la comparación del sistema con una solución GCCL [3] expresamente diseñada para encontrar un conjunto de reglas que definan una sola variable objetivo demuestran la robustez de nuestra propuesta.

La organización del resto del artículo es la siguiente. La sección 2 presenta brevemente los conceptos básicos de reglas de asociación. La sección 3 describe Fuzzy-CSar en detalle y la sección 4 aplica el sistema para extraer reglas de asociaciones interesantes del problema de marketing. Finalmente, la sección 5 concluye

el trabajo y presenta líneas de trabajo futuro.

2 REGLAS DE ASOCIACIÓN DIFUSAS

Dado un conjunto de ejemplos $I = \{i_1, i_2, \ldots, i_n\}$, un conjunto de atributos $A = \{a_1, a_2, \ldots, a_\ell\}$ y una colección de conjuntos difusos asociados a cada variable $F_{a_i} = \{f_{i1}, f_{i2}, \ldots, f_{im}\}$, una regla de asociación difusa presenta la siguiente estructura [2]:

$$\begin{aligned} &\textbf{si } X = \{x_1, x_2, \ldots, x_p\} \text{ es } \{F_{x_1}, F_{x_2}, \ldots, F_{x_p}\} \\ &\textbf{entonces } Y = \{y_1, y_2, \ldots, y_p\} \text{ es } \{G_{x_1}, G_{x_2}, \ldots, G_{x_p}\}, \end{aligned} \quad (1)$$

donde X e Y son dos conjuntos de atributos disjuntos tal que $X \in A$ e $Y \in A$.

El objetivo de los sistemas de aprendizaje de reglas de asociación es generar reglas que denoten asociaciones interesantes entre las variables del problema. Aunque se pueden encontrar muchas medidas de interés en la literatura, las dos más usadas son el soporte (*sop*), un estimador de la frecuencia de aparición del concepto descrito en la regla, y la confianza (*conf*), un estimador de la veracidad de la asociación descrita. En este trabajo, las dos métricas se calculan como

$$\text{sop(R)} = \frac{1}{N}\sum_{e=1}^{N} \mu_A(x^{(e)}) \cdot \mu_B(y^{(e)}) \quad (2)$$

$$\text{conf(R)} = \frac{\sum_{e=1}^{N} \left(\mu_A(x^{(e)}) \cdot \max\left\{1-\mu_A(x^{(e)}), \mu_B(y^{(e)})\right\}\right)}{\sum_{e=1}^{N} \mu_A(x^{(e)})} \quad (3)$$

donde μ_A y μ_B es el grado de emparejamiento de la regla con el ejemplo correspondiente (ver la siguiente sección para detalles de los cálculos).

3 DESCRIPCIÓN DE FUZZY-CSAR

Fuzzy-CSar es una arquitectura incremental de evaluación de reglas de asociación difusas que usa un algoritmo genético (AG) para el descubrimiento de reglas interesantes, es decir, reglas con alto soporte y confianza. A continuación, se describen en detalle los diferentes componentes del sistema.

3.1 REPRESENTACIÓN DEL CONOCIMIENTO

Fuzzy-CSar evoluciona un conjunto o *población* de individuos o *clasificadores*, los cuales están formados por una regla de asociación difusa ($X \rightarrow Y$) y una serie de parámetros. La regla difusa se representa del siguiente modo:

si x_i es $\widetilde{A_i^k}$ y $\cdots$ y x_j es $\widetilde{A_j^k}$ **entonces** y_c es $\widetilde{B_c^k}$,

donde el antecedente X es representado por un conjunto de ℓ_a variables $x_i \ldots x_j$ ($0 < \ell_a < \ell$, siendo ℓ el número de variables del problema) y el consecuente Y contiene una única variable y_c que no existe en el antecedente. Por consiguiente, se permite que las reglas tengan un número arbitrario de variables en el antecedente. Cada variable se representa con una disyunción de etiquetas lingüísticas $\widetilde{A_i^k} = \{A_{i1} \wedge \ldots \wedge A_{in_i}\}$. Para evitar tener reglas con variables muy generales que aporten poca información al experto humano, el sistema permite fijar el número máximo de etiquetas lingüísticas permitidas en las variables del antecedente ($maxEtiqAnt$) y del consecuente ($maxEtiqCons$).

Cada clasificador tiene asociado seis parámetros: (1) el soporte *sop*, un indicador de la frecuencia de aparición de la regla y de su grado de emparejamiento con los ejemplos de entrada; (2) la confianza *conf*, que indica la fuerza de la implicación; (3) el *fitness* F, que denota la cualidad de dicha regla; (4) la experiencia *exp*, el número de veces que la regla se ha emparejado con grado mayor que cero con un ejemplo de entrada; (5) la numerosidad *num*, el número de copias del clasificador en la población; y (6) el tamaño del conjunto de asociación *as*, la media de los tamaños de los conjuntos de asociación en los que la regla ha participado.

3.2 EMPAREJAMIENTO MULTI-ÍTEM

Fuzzy-CSar usa el concepto de *emparejamiento multi-ítem* propuesto en [3] para considerar apropiadamente el conjunto de ítemes que representan cada variable. La idea es que se une la información aportada para cada ítem de una misma variable. El método computa el grado de emparejamiento de una variable i con el atributo de entrada correspondiente e_i como

$$\mu_{\widetilde{A_i}}(e_i) = max_{h_i=1}^{p_i} \mu_{\widetilde{A_i}}(x_i^{(e_i)}), \quad (4)$$

donde p_i es el número de ítemes que representan la variable i. Así pues, el grado de emparejamiento de la variable es el grado máximo de emparejamiento de sus ítemes.

Además, el emparejamiento multi-ítem trata especialmente las variables/constructos de segundo orden, esto es, variables que se reflejan en otros de primer orden, en lugar de en ítemes. Debido a que las variables de segundo orden representan la intersección de distintas variables de primer orden, se calcula el grado de este tipo de variables como la T-norma (mínimo) de cada conjunción difusa, es decir,

$$\mu_{\widetilde{A_i}}(e_i) = min_{k_i=1}^{s_i} \mu_{\widetilde{A_i}}(x_i^{(e_i)}). \quad (5)$$

3.3 INTERACCIÓN DE APRENDIZAJE

En cada iteración de aprendizaje, Fuzzy-CSar recibe un ejemplo de entrada $e = (e_1, e_2, \ldots, e_\ell)$. Seguida-

mente, el sistema crea el *conjunto de emparejamiento* [M] con todos los clasificadores de la población que tienen un grado de emparejamiento con e superior a 0. Si [M] tiene menos de θ_{mna} clasificadores, se dispara el *operador de cubrimiento*, el cual genera nuevas reglas hasta que [M] tenga θ_{mna} clasificadores. A continuación, se crean distintos *conjuntos de asociación* [A] con los clasificadores de [M] siguiendo la metodología que se detalla en el apartado 3.3.2. Se da una probabilidad de selección a cada [A] proporcional a la media de la confianza de los clasificadores de dicho [A]. Luego, se aplica el operador de subsumisión al [A] seleccionado (ver apartado 3.3.3) y se actualizan los parámetros de los clasificadores en [P]. Al final de la iteración se aplica el AG sobre [A] si el tiempo medio desde su última aplicación sobre dicho [A] es superior a θ_{GA}. A continuación se explican más detalladamente (1) el operador de cubrimiento, (2) la metodología de creación de conjuntos de asociación, (3) la subsumisión en el conjunto de asociación, y (4) la actualización de parámetros.

3.3.1 Operador de Cubrimiento

El operador de cubrimiento usa el ejemplo de entrada e para generar un clasificador con el máximo grado de emparejamiento posible con e. Para cada variable, el operador decide aleatoriamente (con probabilidad $1 - P_{\#}$) si la variable debe estar en el antecedente de la regla, con la restricción de que se deben seleccionar un mínimo de 1 variable y un máximo de $\ell - 1$ variables. Seguidamente, se escoge una de las variables no seleccionadas para formar el consecuente de la regla. Cada una de las variables seleccionadas se inicializa con la etiqueta lingüística que maximiza el grado de emparejamiento con el valor correspondiente del atributo (e_i). Finalmente, las variables se generalizan introduciendo cada etiqueta lingüística no seleccionada con probabilidad $P_{\#}$.

3.3.2 Creación de los Conjuntos de Asociación

Este procedimiento persigue crear conjuntos de reglas con características similares para establecer una competición entre ellas con el objetivo que solamente se mantengan las reglas de más calidad en la población. Persiguiendo este fin, Fuzzy-CSar aplica el siguiente procedimiento heurístico. Primeramente, se hace una ordenación ascendente de las reglas de la población según la variable que tienen en el consecuente. Dadas dos reglas r_1 y r_2 que tienen la misma variable en el consecuente, se considera que $r_1 < r_2$ si $\ell_1 < \ell_2$ o ($\ell_1 = \ell_2$ y $u_1 > u_2$), donde ℓ_1, u_1, ℓ_1 y u_2 son, respectivamente, la posición del centro del primer y último conjunto difuso de las variables del consecuente de las reglas r_1 y r_2.

A continuación se construyen los conjuntos de asociación. Se crea el primer [A] y se le inserta el primer clasificador de la población. Entonces, el siguiente clasificador k también se inserta si tiene la misma variable en el consecuente y ℓ_k es inferior que el u_i mínimo entre todos los clasificadores de [A]. El proceso se repite hasta que se encuentra el primer clasificador que no satisface la condición. En este caso, se crea un [A] nuevo y se aplica el mismo proceso para insertar clasificadores en él. La idea que reside en esta estrategia es que las reglas que explican la misma región de un mismo consecuente seguramente denotan las mismas asociaciones entre las variables de entrada.

3.3.3 Subsumisión en el Conjunto de Asociación

Fuzzy-CSar aplica un proceso de subsumisión de reglas con el objetivo de eliminar reglas muy parecidas de la población. Para cada regla del conjunto de asociación seleccionado, se comprueba que la regla no pueda ser subsumida por ninguna de las otras reglas que están en el mismo conjunto de asociación. Se considera que una regla r_1 subsume a otra regla r_2 si se cumplen las cuatro condiciones siguientes: (1) r_1 tiene suficiente experiencia y alta confianza ($conf_1 > conf_0$ y $exp_1 > \theta_{exp}$, donde $conf_0$ y θ_{exp} son dos parámetros de configuración); (2) todas las variables presentes en el antecedente de r_1 también existen en el antecedente de r_2 (además, r_2 puede tener otras variables en el antecedente); (3) la variable en el consecuente de r_1 y r_2 es la misma; y (4) r_1 es más general que r_2, es decir, para cada variable en el antecedente o consecuente de r_1, esta variable tiene definidas, como mínimo, las mismas etiquetas lingüísticas que la misma variable en r_2.

3.3.4 Actualización de Parámetros

Al final de cada iteración de aprendizaje se actualizan los parámetros de todos los clasificadores de la población. Primeramente, se actualiza el soporte de la regla según

$$sop_{t+1} = \frac{sop_t \cdot (lt - 1) + \mu_{\widetilde{A}}(x^{(e)}) \cdot \mu_{\widetilde{B}}(y^{(e)})}{lt}, \quad (6)$$

donde lt es el tiempo de vida del clasificador, es decir, el número de iteraciones que ha estado en la población. A continuación, se actualiza la confianza de la regla:

$$conf_{t+1} = \frac{sum_imp_{t+1}}{sum_mat_{t+1}}. \quad (7)$$

sum_imp y sum_mat se definen como

$$\begin{aligned} sum_imp_{t+1} &= sum_imp_t + \mu_{\widetilde{A}}(x^{(e)}) \cdot \\ &\quad max\{1 - \mu_{\widetilde{A}}(x^{(e)}), \widetilde{B}(y^{(e)})\} \quad (8) \end{aligned}$$

$$sum_mat_{t+1} = sum_mat_t + \mu_{\widetilde{A}}(x^{(e)}). \quad (9)$$

Seguidamente, se computa el *fitness* de cada clasificador:

$$F = conf^{\nu}. \quad (10)$$

Finalmente, se actualiza as el cual mantiene la media aritmética del número de clasificadores de los conjuntos de asociación en los que el clasificador ha participado.

3.4 DESCUBRIMIENTO DE REGLAS

Fuzzy-CSar usa un AG *steady-state* para crear nuevas reglas prometedoras. El AG selecciona dos clasificadores con probabilidad proporcional a su *fitness*. Estos dos padres se cruzan para generar dos hijos cl_1 y cl_2 con probabilidad P_{chi}; de lo contrario, los hijos son una copia exacta de los padres. Fuzzy-CSar utiliza un cruce uniforme que contempla la restricción que los hijos deben tener, como mínimo, una variable en antecedente. Los hijos resultantes pueden sufrir tres tipos distintos de mutación: (1) introducción/borrado de variables en antecedente (con probabilidad $P_{I/B}$); (2) mutación de las etiquetas lingüísticas de variables existentes (con probabilidad P_{μ}); (3) mutación de la variable de consecuente (con probabilidad P_C). El primer tipo de mutación selecciona aleatoriamente si se debe añadir una de las variables no usadas (sólo aplicable si hay variables no usadas en la regla) o quitar una variable del antecedente (sólo aplicable si el antecedente tiene más de una variable). El segundo tipo de mutación modifica los términos lingüísticos de una de las variables de la regla acorde con el método definido en [3]. El tercer tipo de mutación cambia la variable del consecuente, escogiendo aleatoriamente una de las posibles variables de salida. En cualquiera de los tres tipos de mutación, cuando se añade una nueva variable, esta se inicializa con la etiqueta lingüística que maximiza el grado de emparejamiento con el atributo correspondiente del ejemplo de entrada.

Después del cruce y de la mutación, los descendientes se introducen en la población, borrando clasificadores si se sobrepasa su tamaño máximo. En este caso, los clasificadores reciben una probabilidad de borrado proporcional a su parámetro as. Además, esta probabilidad de borrado se aumenta si el *fitness* del clasificador es inferior a $\delta\overline{F}$, donde $\overline{F}$ es el *fitness* medio de la población y δ es un parámetro de configuración.

3.5 REDUCCIÓN DE LA POBLACIÓN

Al final del proceso de aprendizaje se usa la siguiente heurística para eliminar reglas poco interesantes de la población final. En primer lugar, se eliminan todas las reglas cuya experiencia es inferior a θ_{exp}. Entonces, para cada regla en la población, se mira si existe otra regla que la pueda subsumir. Se usa el mismo criterio que el mencionado en el apartado 3.3.3 para decidir si una regla r_1 puede subsumir otra regla r_2 pero con una excepción: se requiere que la confianza de r_1 sea mayor que la confianza de r_2 en vez de requerir que la confianza de r_1 sea mayor que $conf_0$.

4 EXPERIMENTACIÓN

4.1 DESCRIPCIÓN DEL PROBLEMA

La experimentación parte del modelo causal presentado en [4], el cual se basa en el análisis del estado de *flow* del usuario al navegar por la web. Se han utilizado los datos proporcionados por los mismos autores; esta práctica es habitual en escenarios de investigación en los que se evalúa empíricamente el rendimiento de una herramienta analítica basada en inteligencia artificial para modelado en marketing. El concepto de *flow* en Internet se alcanza cuando el usuario esta totalmente inmerso en la navegación, hasta tal punto que "nada más parece importarle" [4].

El modelo que consideramos en la experimentación consta de 9 variables: *inicio de uso de la web*, *implicación en la web*, *habilidad/control*, *velocidad de interacción*, *desafío/estímulo*, *nivel de atención*, *telepresencia/distorsión de tiempo* y *flow*. Todas las variables toman valores enteros de 1 a 9 excepto *inicio de uso de la web*, que puede tomar intervalos de años de experiencia como usuario de la Web, cuyas marcas de clase con las siguientes: {0,25, 0,75, 1.50, 2,50, 3,50 y 4,50}. Se ha usado una semántica difusa con tres etiquetas lingüísticas triangulares distribuidas uniformemente para todas las variables del sistema excepto *inicio de uso de la web*. Para esta última variable se usa una semántica no uniforme con seis etiquetas lingüísticas, cada una centrada en uno de los seis posibles valores de la variable.

Se han hecho tres tipos distintos de experimentos con este problema fijando el grupo de variables que podían estar en el antecedente y el grupo de variables que podían estar en el consecuente.

- En el primer experimento (*4in1out*), se han considerado únicamente cuatro de las nueve posibles variables en el antecedente (*habilidad/control*, *velocidad de interacción*, *desafío/estímulo* y *telepresencia/distorsión de tiempo*) pues son las variables que tienen una relación directa con *flow* según el modelo causal proporcionado en [4].
- En el segundo experimento (*9in1out*), se ha con-

siderado que las reglas de asociación podían contener todas las variables en el antecedente excepto *flow* y el consecuente se restringe a la variable *flow*.

- En el tercer experimento (*9in9out*), se permite que cualquiera de las nueve variables (incluyendo *flow*) puedan estar tanto en el antecedente como en el consecuente de la regla.

Los resultados obtenidos con Fuzzy-CSar se han comparado con los proporcionados por el sistema GCCL presentado en [3], un sistema multi-objetivo basado en NSGA-II que obtiene un frente de Pareto con las mejores reglas encontradas (esto es, las reglas no dominadas en términos de soporte y confianza). Es importante destacar la diferencia entre ambas aproximaciones: mientras la solución GCCL tiene como objetivo optimizar el frente Pareto, el objetivo de Fuzzy-CSar es evolucionar un conjunto de reglas de asociación diversas que tengan máxima confianza. Los dos sistemas se han configurado como se especifica a continuación. Para Fuzzy-CSar se han usado los siguientes parámetros: tamaño_pop = 6.400, $\theta_{GA} = 50$, $P_{\#} = 0{,}5$, $\theta_{exp} = 1.000$, $conf_0 = 0{,}95$, ν=1, $P_{\chi} = 0{,}8$, $P_{I/B} = 0{,}1$, $P_{\mu} = 0{,}1$, $P_C = 0{,}1$, $\delta = 0{,}1$. Para el sistema GCCL, se ha usado (ver [3] para detalles sobre la notación): número de generaciones 100, tamaño de la población 100, probabilidad de cruce 0,7 y probabilidad de mutación por cromosoma 0,1. Todos los resultados que se presentan a continuación son medias de 10 semillas aleatorias distintas.

4.2 RESULTADOS

La figura 1 compara los frentes Pareto (media de 10 ejecuciones) obtenidos con los sistemas Fuzzy-CSar y GCCL cuando se usan ocho variables en el antecedente y la variable *flow* en el consecuente (problema *8in1out*). Los resultados son equivalentes a usar 4 variables en el antecedente (problema *4in1out*); por falta de espacio y por similitud con la figura 1, se omite la figura de los frentes Pareto para el problema *4in1out*. Esta similitud entre los frentes Pareto generados con ambos problemas soporta la hipótesis que las cuatro variables escogidas en el primer experimento son las más relevantes para definir el concepto de *flow*.

La tabla 1 complementa los resultados mostrando el tamaño de la población evolucionada por Fuzzy-CSar, el número de reglas no dominadas dentro de esta población y el tamaño del conjunto Pareto creado por GCCL; asimismo, la tabla 2 indica la distancia de *crowding* media entre las soluciones de la población para cada caso. Los resultados muestran claramente que Fuzzy-CSar evoluciona un frente Pareto muy similar al obtenido por GCCL; en ambos casos, las soluciones están uniformemente distribuidas a lo largo del Pareto. Además, Fuzzy-CSar también genera un conjunto variado de reglas (479 y 1260 en media para los problemas *4in1out* y *9in1out* respectivamente) que, aun no estando en el Pareto, pueden poseer conocimiento interesante para el experto humano. Cabe destacar las diferencias entre ambos sistemas para ver el significado de estos resultados. Fuzzy-CSar es un sistema incremental de aprendizaje que utiliza un sistema implícito de *niching* genotípico y optimiza en función de la confianza de cada regla; así pues, Fuzzy-CSar no tiene el objetivo implícito de optimizar el frente Pareto, sino de mantener una solución distribuida en el espacio de búsqueda en base a nichos. En cambio, el sistema GCCL es un multi-objetivo que persigue la optimización del frente Pareto en base dos objetivos: confianza y soporte. Para conseguir este fin, GCCL usa ordenación por *ranking*, *niching* explícito mediante una función de *crowding* y un esquema de se-

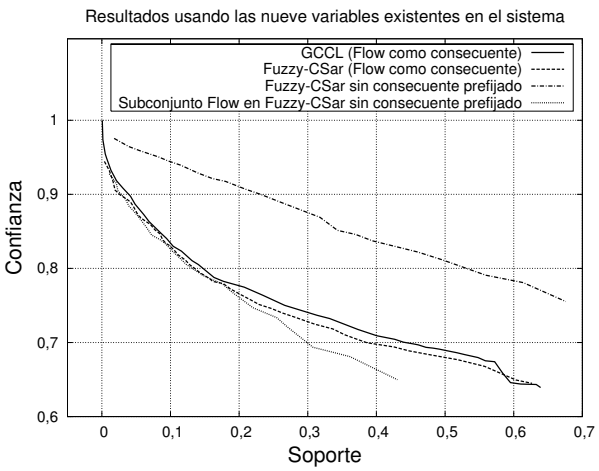


Figura 1: Comparación del frente Pareto medio de Fuzzy-CSar y GCCL

Tabla 1: Número de reglas evolucionadas por Fuzzy-CSar, número de estas reglas que están en el Pareto y tamaño del Pareto obtenido por GCCL

Problema	*FCSar*	*FCSar Par.*	*GCCL Par.*
4in1out	479,2	76,3	82,6
8in1out	1259,7	105,9	84,4
9in9out	1752,5	468,3	—

Tabla 2: Media de la distancia de *crowding* de los frentes Pareto evolucionados con Fuzzy-CSar y GCCL

Problema	*FCSar All*	*FCSar Par.*	*GCCL Par.*
4in1out	$2.36 \cdot 10^{-3}$	$1.53 \cdot 10^{-2}$	$1.49 \cdot 10^{-2}$
8in1out	$9.54 \cdot 10^{-4}$	$1.07 \cdot 10^{-2}$	$1.49 \cdot 10^{-2}$
9in9out	$6.88 \cdot 10^{-4}$	$2.58 \cdot 10^{-3}$	—

lección de torneo basado en la ordenación de las soluciones en distintos frentes. Por tanto, el hecho que Fuzzy-CSar presente un frente Pareto similar a un sistema especializado en optimizar el frente Pareto como es GCCL confirma la robustez Fuzzy-CSar.

En segundo lugar, la figura 1 también recoge los resultados del tercer experimento, esto es, cuando se permite que cualquier variable esté en el antecedente o consecuente de las reglas de Fuzzy-CSar; en este caso, Fuzzy-CSar se usa para buscar asociaciones entre cualquier variable del problema. Por consiguiente, en esta aproximación dejamos que Fuzzy-CSar descubra las asociaciones más importantes, esto es, el modelo estructural del sistema, sin considerar el modelo causal proporcionado inicialmente. Los resultados muestran el potencial diferenciador de nuestra aproximación. En una misma ejecución, Fuzzy-CSar es capaz de evolucionar un conjunto de reglas con alto soporte y confianza, las cuales tienen distintos consecuentes, resultando en un frente Pareto que supera ampliamente al obtenido con el método GCCL. Por último, la figura muestra el frente Pareto de esta ejecución cuando sólo se consideran las reglas que tienen la variable *flow* en el consecuente. En este caso, el Pareto se asemeja al original por valores de confianza altos y se ve ligeramente degradado a medida que la confianza de las soluciones decrece. Este comportamiento se puede explicar con la siguiente hipótesis. En este experimento, el número de reglas que Fuzzy-CSar debe evolucionar crece con el número de consecuentes considerados, lo que promueve la competición entre reglas para mantenerse en la población. Como Fuzzy-CSar optimiza en función de la confianza de las reglas, el sistema prioriza las reglas con máxima confianza, lo que va en detrimento de reglas con menos confianza pero con más soporte.

Se podrían obtener resultados similares con la aproximación GCCL mediante la ejecución de nueve experimentos, considerando cada vez una de las nueve variables como salida. Esto resultaría en nueve Paretos distintos que se deberían juntar para obtener el Pareto final. Aún así, es importante destacar que Fuzzy-CSar proporciona un soporte natural para la extracción automática de reglas de asociación interesantes con distintos consecuentes, evolucionando paralelamente las distintas soluciones distribuidas en el espacio de características y manteniendo aquellas que tienen confianza más alta.

4.3 EJEMPLO DE INTERPRETACIÓN DEL CONOCIMIENTO EXTRAÍDO

Finalmente, de las múltiples reglas generadas en el proceso de minería de datos, seleccionamos una muestra ilustrativa del atractivo de la información ofrecida por el método de aprendizaje no supervisado que presentamos. En concreto, sin perjuicio del interés de otras reglas generadas en las que encontramos una estructura de antecedentes y consecuentes más acorde con el modelo estructural finalmente propuesto en [4], a continuación mostramos algunos casos con una relación de antecedentes y consecuente no contemplada en dicho modelo. Precisamente, éste es el atractivo de las salidas ofrecidas por nuestro método, al no condicionarse el proceso de aprendizaje automático por ninguna estructura de conocimiento *a priori*.

R_1: **SI** *implicaciónWeb* **es** Media **y** *habilidad/control* **es** {Pequeña o Media} **y** *nivelAtención* **es** {Pequaño o Medio} **y** *flow* **es** {Pequeño o Medio} **ENTONCES** *comportamExploratorio* **es** Medio [Sop: 0,22; Conf: 0,87]

En primer lugar, la regla R_1 muestra una relación de antecedentes para el consecuente "comportamiento exploratorio". En el modelo propuesto en [4] sólo existía un determinante directo de este tipo de comportamientos de navegación del consumidor en Internet, es decir, el *nivel de telepresencia* y *pérdida de la noción del tiempo* cuando se navega. En cambio, el patrón de información en R_1 deja al margen este elemento. En particular, considera las siguientes variables del consumidor: *implicación con la Web*, *habilidad/control* percibido y *nivel de atención* en el proceso de navegación y, finalmente, el estado de *flow*. *Grosso modo*, nos dice que niveles intermedios de estos antecedentes explican, con un nivel de confianza de 0,87, estados moderados de comportamientos exploratorios, sin problema inmediato de compra que resolver, en la Web. Brevemente, esta regla que puede inspirar al experto de marketing a considerar nuevos caminos o relaciones no contempladas *a priori*, directamente, entre los antecedentes y consecuente citados. En general, por los antecedentes contenidos en la regla y, en especial, por incluir a *flow* como determinante de comportamientos exploratorios en la Web. Esta relación, que es bastante razonable, siendo contemplada *a priori* en el modelo de base de [4], fue descartada finalmente tras el refinado del modelo. No obstante, reglas como R_1 nos advierten de la conveniencia de su consideración.

R_2: **SI** *implicaciónWeb* **es** {Pequeña o Media} **y** *desafío/estímulo* **es** {Pequeño o Medio} **y** *telepres_distTiempo* **es** Medio **y** *comportamExploratorio* **es** {Medio o Grande} **ENTONCES** *nivelAtención* **es** Medio [Sop: 0,21; Conf: 0,84]

A continuación hemos seleccionado una regla en la que aparece "nivel de atención" como consecuente. En el modelo validado en [4], nivel de atención es una consecuencia de *implicación con la Web* y desafío/estímulo. De nuevo, en la regla R_2 aparecen relaciones no consideradas en [4]. Aparte de *implicación con la Web*

y *desafío/estímulo*, sí contempladas en [4], R_2 sugiere que variables cómo *telepresencia/pérdida* de la noción del tiempo y *flow* pueden determinar niveles de atención moderados en el proceso de navegación. Esta información, más allá de los niveles tomados por los antecedentes y el consecuente, es interesante por la relación de prelación que establece entre las variables, contraria a la concluida en [4]. Esto es razonable si consideramos el sistema de generación de las reglas de nuestro método. Al contrario que en [4], donde el método estadístico utilizado analiza la procedencia de un modelo establecido por el experto, en nuestro caso el patrón de información que se genera no responde a condicionantes estructurales de ningún tipo. En algunos casos, los patrones obtenidos no tendrán sentido. Sin embargo, no es el caso del patrón en R_2, que podría justificarse en un proceso dinámico en el que la experiencia del proceso de navegación on-line se retroalimentara cambiando, por tanto, la secuencia de relaciones entre ciertas variables como las analizadas en este caso. En definitiva, R_2 plantea un escenario no contemplado *a priori* por los expertos en [4] que, de haber aplicado un método como el que presentamos en este trabajo, se podría haber estudiado de cara al proceso de diseño y depuración de su modelo.

Más allá de estos ejemplos, pensemos en el soporte que métodos de aprendizaje no supervisado como éste pueden tener para un gestor de marketing, cuando tenga que afrontar problemas de decisión estructurados deficientemente, sin información *a priori*, al proporcionar patrones de información no predefinidos, imprevistos, y en ocasiones inspiradores de soluciones.

5 CONCLUSIONES Y TRABAJO FUTURO

En este artículo se ha aplicado Fuzzy-CSar, un sistema clasificador difuso que evoluciona reglas de asociación, para la extracción de modelos casuales en un entorno de marketing. Fuzzy-CSar usa una representación difusa para tratar la incertidumbre de los datos del problema. Se ha comparado el comportamiento del sistema con el presentado por un método multi-objetivo GCCL. Los resultados han mostrado la competitividad de Fuzzy-CSar. Más concretamente, se ha observado empíricamente que Fuzzy-CSar puede conseguir frentes Pareto muy similares a GCCL cuando se fija una variable en el consecuente de las reglas. Además, Fuzzy-CSar tiene la ventaja que puede aprender reglas de asociación sin que previamente se tengan que fijar que variables son de entrada y cuales son de salida. Los experimentos realizados en este entorno han mostrado que Fuzzy-CSar es capaz de extraer reglas de asociación con alto soporte y confianza, dibujando un frente Pareto muy superior al obtenido con GCCL fijando la variable de salida. Asimismo, Fuzzy-CSar continua generando reglas interesantes para el consecuente *flow*.

Como trabajo futuro, se va a analizar más detalladamente la diversidad de las poblaciones evolucionadas por Fuzzy-CSar. Los resultados presentados indican que el sistema evoluciona una gran diversidad de reglas, de las cuales solo un pequeño porcentaje forman el frente Pareto. Aún así, reglas que estén dominadas por soporte y confianza por otras reglas pueden ser igualmente interesantes si descubren nuevas relaciones entre variables. Por tanto, se investigará más en detalle la semántica de las reglas de la población y se diseñarán metodologías para su comparación.

Referencias

[1] G. H. V. Bruggen and B. Wierenga. Broadening the perspective on marketing decision models. *International Journal of Research in Marketing*, 17:159–168, 2000.

[2] T. P. Hong, C. S. Kuo., and S. C. Chi. Trade-off between computation time and number of rules for fuzzy mining from quantitative data. *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 9(5):587–604, 2001.

[3] F. Martínez-López and J. Casillas. Marketing intelligent systems for consumer behaviour modelling by a descriptive induction approach based on genetic fuzzy systems. *Industrial Marketing Management*, doi:10.1016/j.indmarman.2008.02.003, 2008.

[4] T. Novak, D. Hoffman, and Y. Yung. Measuring the customer experience in online environments: A structural modelling approach. *Marketing Science*, 19(1):22–42, 2000.

[5] B. Wierenga and G. H. V. Bruggen. The integration of marketing problem-solving modes and marketing management support systems. *Marketing Management Support Systems*, 61:21–37, 1997.

[6] B. Wierenga and G. H. V. Bruggen. *Marketing management support systems: principles, tools and implementation*. Kluwer Academic Publishers, 2000.

[7] M. Zhang. Introduction to data mining and KDD process. Technical report, Victoria University of Wellington. Disponible en: http://www.mcs.vuw.ac.nz/courses/COMP422/2004T2/Lecturenotes/week01.pdf, 2004.

CLASIFICACIÓN DIFUSA DE LA MORTALIDAD POR INFECCIÓN EN PACIENTES CON QUEMADURAS GRAVES MEDIANTE ALGORITMOS EVOLUTIVOS MULTIOBJETIVO

Alcaraz, J.M.[1] **Jiménez, F.**[2] **Juárez, J.M.**[2] **Palma, J.**[2] **Sánchez, J.F**[3] **Sánchez, G.**[2]

[1] Universidad de Murcia, josem.alcaraz@carm.es
[3] Universidad de Murcia, jsanchez@ditec.um.es
[2] Universidad de Murcia, {fernan, jmjuarez, jtpalma, gracia}@um.es

Resumen

La clasificación de la supervivencia en pacientes infectados con quemaduras graves es un problema actual de gran relevancia en el campo de la medicina. En este trabajo proponemos, por un lado, un modelo de optimización multiobjetivo con restricciones para obtener modelos difusos de clasificación basándose en criterios de precisión e interpretabilidad. Por otro lado proponemos un enfoque multiobjetivo evolutivo para la clasificación difusa a partir de datos con atributos reales y discretos. El enfoque evolutivo multiobjetivo se ha evaluado mediante tres esquemas evolutivos diferentes: Preselección con nichos, NSGA-II y ENORA. Los resultados obtenidos han sido comparados en términos de eficacia mediante el uso de técnicas estadísticas.

Palabras Clave: Algoritmos evolutivos multiobjetivo. Clasificación Difusa. Aplicaciones en Medicina.

1 INTRODUCCIÓN

La Unidad de Cuidados Intensivos (UCI) es un servicio médico hospitalario que proporciona atención médica de pacientes en estado crítico. La muerte por infección generalizada en España se eleva a 15000 pacientes de UCI al año, siendo una causa común de mortalidad en pacientes con quemaduras graves. Uno de los problemas fundamentales es categorizar la gravedad real de estos pacientes debido a la dificultad que supone su diagnóstico. Recientes estudios de la SEMYCYUC[1] establecen que hay evidencias científicas suficientes para afirmar que una clasificación precoz y una terapia temprana puede mejorar notoriamente su supervivencia [13].

Una forma eficaz de abordar el problema es mediante modelado difuso. En modelado difuso uno de los problemas más habituales es la identificación de un modelo difuso usando datos del tipo entrada-salida. Dado un conjunto de datos para el que presumimos alguna dependencia funcional, el problema consiste en derivar reglas (difusas) a partir de los datos que caracterizan la función desconocida, de forma tan precisa e interpretable como sea posible. Se han propuesto numerosas aproximaciones para generar automáticamente reglas *if-then* a partir de datos numéricos. Por la complejidad del problema, la *Computación evolutiva* ha sido una de las técnicas más aceptadas.

La computación evolutiva [5] ha sido aplicada con éxito para aprender modelos difusos [6, 9]. Esto ha desembocado en muchos algoritmos complejos y a menudo [12] no se le ha dado importancia a la interpretabilidad del conjunto de reglas resultante. En estos casos, el modelo difuso se convierte en una caja negra, y puede cuestionarse la decisión de aplicar modelado difuso en lugar de otras técnicas.

Por otra parte, los algoritmos evolutivos han sido reconocidos como técnicas apropiadas para la optimización multiobjetivo porque llevan a cabo una búsqueda de múltiples soluciones en paralelo [2, 3]. Los actuales enfoques evolutivos para la optimización multiobjetivo consisten en algoritmos evolutivos multiobjetivo basados en la idea de optimalidad de Pareto, en el que todos los objetivos se optimizan simultáneamente para encontrar múltiples soluciones no-dominadas en una única ejecución del algoritmo. El decisor puede, entonces, elegir la solución más apropiada de acuerdo con el entorno de decisión actual al finalizar la ejecución del algoritmo. Además, si el

[1] SEMYCYUC: Sociedad Española de medicina Intensiva, Crítica y Unidades Coronarias

entorno de decisión cambia, no siempre es necesario ejecutar el algoritmo de nuevo; se puede escoger otra solución del conjunto de soluciones no-dominadas que ya se ha obtenido.

El modelado difuso puede considerarse desde un enfoque de optimización evolutiva multiobjetivo [8]. En su mayor parte, las actuales líneas de desarrollo en modelización difusa abordan mejorar la precisión en los modelos descriptivos, y mejorar la interpretabilidad en modelos aproximativos [1]. Este artículo enfoca el problema por medio de optimización multiobjetivo en la que los criterios de precisión e interpretabilidad son considerados simultáneamente.

En el artículo, proponemos un enfoque de optimización multiobjetivo evolutiva para generar modelos difusos de clasificación teniendo en cuenta criterios de precisión e interpretabilidad. La sección 2 describe el modelo difuso de clasificación y los criterios tenidos en cuenta en el proceso de optimización. Se propone un modelo de optimización multiobjetivo con restricciones. La sección 3 muestra los componentes principales de los tres algoritmos evolutivos multiobjetivo utilizados en este artículo. La sección 4 muestra los experimentos efectuados y los resultados obtenidos para el problema de clasificación la mortalidad por infección en pacientes con quemaduras graves. La sección 5 concluye el artículo.

2 MODELO DIFUSO DE CLASIFICACIÓN

2.1 IDENTIFICACIÓN DEL MODELO

Consideramos un conjunto de N datos de aprendizaje constituidos de la siguiente forma:

- Un vector de atributos reales de entrada $\boldsymbol{x} = (x_1, \ldots, x_p)$, $x_i \in [l_i, u_i] \subset \Re$, $p \geq 0$,
- Un vector de atributos discretos de entrada $\boldsymbol{w} = (w_1, \ldots, w_q)$, $w_i \in \{1, \ldots, v_i\}$, $q \geq 0$, siendo v_i el número de clases para el atributo discreto de entrada,
- Un valor para el atributo discreto de salida $y \in \{1, \ldots, z\}$, siendo z el número de clases para el atributo discreto de salida.

Consideramos un modelo difuso de clasificación formado por M reglas $R_1, \ldots, R_M$. Cada regla R_j tiene la siguiente estructura:

$$R_j: \quad \textbf{Si} \quad \begin{array}{l} x_1 \text{ es } \widetilde{A}_{1j} \quad \textbf{y} \ \ldots \ \textbf{y} \quad x_p \text{ es } \widetilde{A}_{pj} \quad \textbf{y} \\ w_1 \text{ es } B_{1j} \quad \textbf{y} \ \ldots \ \textbf{y} \quad w_q \text{ es } B_{qj} \end{array} \quad \textbf{entonces} \quad y \text{ es } C_j$$

donde $\widetilde{A}_{ij}$ son los conjuntos difusos asociados a los atributos reales de entrada, B_{ij} los valores discretos asociados a los q atributos discretos de entrada, y C_j el valor discreto asociado al atributo discreto de salida. Cada conjunto difuso $\widetilde{A}_{ij}$ se describe mediante su función de pertenencia $\mu_{\widetilde{A}_{ij}}$. En nuestro modelo se han utilizado funciones de pertenencia gaussianas:

$$\mu_{\widetilde{A}_{ij}}(x_i) = \exp\left[-\frac{1}{2}\left(\frac{x_i - a_{ij}}{\sigma_{ij}}\right)^2\right]$$

donde $a_{ij} \in [l_i, u_i]$ es el centro, y $\sigma_{ij} > 0$ la varianza. La salida del modelo ψ corresponde a la clase de salida C cuyo valor de activación λ_C sea máximo:

$$\psi(\boldsymbol{x}, \boldsymbol{w}) = \arg_C \max_{C=1}^{z} \lambda_C(\boldsymbol{x}, \boldsymbol{w})$$

El valor de activación λ_C se calcula sumando los grados de disparo de cada regla R_j cuyo valor para el atributo discreto de salida C_j sea igual a C:

$$\lambda_C(\boldsymbol{x}, \boldsymbol{w}) = \sum_{\substack{j=1,\ldots,M \\ C_j = C}} \left((\phi_j(\boldsymbol{w}) + 1) \prod_{i=1}^{p} \mu_{\widetilde{A}_{ij}}(x_i) \right)$$

donde $\phi_j(\boldsymbol{w})$ es el número de atributos discretos de entrada tales que $w_j = B_{ij}$.

2.2 CRITERIOS DE MODELADO DIFUSO

Consideramos tres criterios principales: precisión, transparencia y compacidad.

Precisión. *Tasa de clasificación* (*Classification Rate*): $CR = \Phi/N$, donde Φ es el número de datos para los cuales $\psi(\boldsymbol{x}, \boldsymbol{w}) = y$.

Transparencia. *Similitud* (S) [7]

$$S = \max_{\substack{i,k=1,\ldots,M \\ j=1,\ldots,p \\ \widetilde{A}_{ij} \neq \widetilde{A}_{kj}}} S(\widetilde{A}_{ij}, \widetilde{A}_{kj})$$

donde $S(\widetilde{A}, \widetilde{B}) = \max\left\{ \frac{|\widetilde{A} \cap \widetilde{B}|}{|\widetilde{A}|}, \frac{|\widetilde{A} \cap \widetilde{B}|}{|\widetilde{B}|} \right\}$

Compacidad Número de reglas (M), y el número de conjuntos difusos distintos (L).

2.3 MODELO DE OPTIMIZACIÓN

De acuerdo con las anteriores observaciones, proponemos el siguiente modelo de optimización con dos objetivos y una restricción:

$$\begin{array}{ll} Maximizar & CR \\ Minimizar & M \\ Sujeto\ a & S \leq g_s \end{array} \tag{1}$$

Se supone que el número de conjuntos difusos diferentes (L) es menor en aquellos modelos con un número de reglas (M) menor. Por tanto en el modelo de optimización se ha utilizado sólo el segundo de los valores. Por otra parte, nos interesa minimizar la

similitud (S) sólo hasta un cierto umbral (g_s), ya que modelos con conjuntos difusos poco similares no serán suficientemente precisos. De esta forma se restringe el espacio de búsqueda y se aumenta la eficacia de los algoritmos.

3 ALGORITMOS EVOLUTIVOS MULTIOBJETIVO

Proponemos un sistema de aprendizaje híbrido para encontrar múltiples soluciones Pareto-óptimas simultáneamente, teniendo en cuenta criterios de precisión, transparencia y compacidad. Estudiamos distintos algoritmos evolutivos multiobjetivo para desarrollar la estructura y los parámetros del conjunto de reglas, junto con un operador de simplificación del conjunto de reglas para conseguir transparencia y compacidad.

3.1 REPRESENTACIÓN DE SOLUCIONES

Se utiliza una representación con codificación de números reales y discretos utilizando un enfoque Pittsburgh. Cada individuo de la población contiene un número de reglas M que puede variar entre min y max, donde los valores min y max los elige un decisor teniendo en cuenta que debe cumplirse que debe haber al menos una regla por cada clase de salida, es decir, $z \leq min \leq M \leq max$. Para cada regla R_j se representan:

- Los conjuntos difusos asociados a los atributos reales de entrada x_i, $i = 1, \ldots, p$, mediante números reales $a_{ij} \in [l_i, u_i]$ y $\rho_{ij} = 2\sigma_{ij}^2 > 0$, que definen los centros y varianzas respectivamente.
- Los valores discretos asociados a los atributos discretos de entrada w_i, $i = 1, \ldots, q$, mediante números enteros $b_{ij} \in \{1, \ldots, v_i\}$.
- El valor discreto asociado al atributo discreto de salida mediante un número entero $c_j \in \{1, \ldots, z\}$.
- Para llevar a cabo cruce y mutación adaptativos, cada individuo tiene dos parámetros discretos $d \in (0, \delta)$ y $e \in (0, \epsilon)$ asociados al cruce y mutación, donde δ es el número de operadores de cruce y ε es el número de operadores de mutación.

3.2 POBLACIÓN INICIAL

Los individuos se inicializan generando aleatoriamente una distribución uniforme dentro de los límites del espacio de búsqueda. Debe cumplirse también:

$$\alpha_i = \frac{u_i - l_i}{\gamma} \leq \rho_{ij} \leq u_i - l_i = \beta_i, \ i = 1, \ldots, p$$

3.3 MANEJO DE RESTRICCIONES

Se utiliza la regla propuesta en [10] para el manejo de restricciones.

3.4 OPERADORES DE VARIACIÓN

Se definen diferentes operadores de variación, los cuales se aplican de forma adaptativa mediante los parámetros d y e y siendo p_v su probabilidad de variación.

Cruce de Conjuntos Difusos. Intercambia dos conjuntos difusos seleccionados aleatoriamente.

Cruce de Reglas. Intercambia dos reglas seleccionadas aleatoriamente.

Cruce de Reglas por Incremento. Añade a cada individuo una regla seleccionada aleatoriamente del otro individuo.

Mutación del Centro de un Conjunto Gaussiano. Mutación del centro de uno o más conjuntos difusos seleccionados aleatoriamente.

Mutación de la Varianza de un Conjunto Gaussiano. Mutación de la varianza de uno o más conjuntos difusos seleccionados aleatoriamente.

Mutación de un Conjunto Gaussiano. Cambia un conjunto difuso seleccionado aleatoriamente por otro conjunto difuso seleccionado aleatoriamente.

Mutación por Incremento de Reglas. Añade una nueva regla con valores aleatorios.

Mutación de un Valor Discreto. Mutación de un valor discreto seleccionado aleatoriamente.

3.5 SIMPLIFICACIÓN DEL CONJUNTO DE REGLAS

La simplificación del conjunto de reglas se realiza mediante la técnica propuesta en [7] y [12], realizando la unión de aquellos conjuntos con $\eta_1 < S(\widetilde{A}, \widetilde{B})$ y separando aquellos con $\eta_1 < S(\widetilde{A}, \widetilde{B}) < \eta_2$.

Los valores η_1 y η_2 son el umbral para efectuar la fusión o la separación y debe cumplirse $0 < \eta_2 < \eta_1 < 1$.

3.6 PROCESO DE DECISIÓN

Mientras el conjunto de soluciones no dominadas S sea no vacío:

> Extraer la solución $s \in S$ con menor número de reglas y de conjuntos difusos.
>
> Si s es suficientemente interpretable y precisa, devolver s.

Si S es vacío, no hay solución satisfactoria.

4 EXPERIMENTOS Y RESULTADOS

Este trabajo se basa en los datos recogidos por el Sistema de Información Clínico de una UCI durante los

años 1999 y 2002, seleccionando, bajo asesoramiento médico, un total de 99 pacientes con diferentes complicaciones infecciosas. A partir de la información reflejada en el historial de cada paciente, los especialistas consideran que los parámetros mostrados en la tabla 1 recogen las evidencias clínicamente más relevantes para el establecimiento de la supervivencia de pacientes con infecciones en la UCI.

Para este estudio se han seleccionado, de forma aleatoria, 89 pacientes para formar parte del conjunto de entrenamiento, y 10 para la evaluación. La importancia de cada parámetro respecto a la supervivencia de los pacientes se ha establecido realizando una clasificación de parámetros mediante la medida de Información Mútua, basado en la Entropía de Shanon [4]. Los pesos mostrados en la tabla 1 muestran los resultados obtenidos y, como resultado, se han descartado para este estudio los parámetros con peso igual a 0. Tenemos, por tanto, un problema con $p = 2$ entradas reales pertenecientes al intervalo $[0, 85]$ y $q = 10$ entradas discretas de tipo booleano.

Una de las técnicas más comunes de la Inteligencia Artificial utilizadas en el campo médico para la toma de decisiones son los *árboles de decisión*. Con éstas técnicas se han obtenido las siguientes tasas de clasificación en evaluación: 0.55 (ID3), 0.70 (J48) y 0.75 (ADT). En nuestro trabajo utilizamos tres algoritmos evolutivos multiobjetivo de tipo Pareto: Preselección con nichos, ENORA y NSGA-II. Preselección con nichos y ENORA son algoritmos propuestos con los autores en [7] y [11] respectivamente, mientras NSGA-II es el conocido algoritmo evolutivo multiobjetivo propuesto por Deb en [3]. Los algoritmos se han ejecutado 100 veces con los parámetros mostrados en la tabla 2. La tabla 4 muestra las mejores soluciones no-dominadas obtenidas y se marca en negrita la solución elegida tras el proceso de decisión para cada algoritmo. Finalmente, se elige la solución obtenida con ENORA con 10 reglas que se muestra en la tabla 4.

Tabla 2: Parámetros de ejecución de los algoritmos.

$N = 100$	$p_v = 0.1$
$N^o\ Evaluaciones = 10^5$	$g_s = 0.1$
$min = 10$	$\gamma = 100$
$max = 20$	$\eta_1 = 0.6$, $\eta_2 = 0.1$
$N^o\ Hijos = 10$ (Preselección con nichos)	
$N^o\ min\ Nichos = 5$ (Preselección con nichos)	
$N^o\ max\ Nichos = 35$ (Preselección con nichos)	

Para comparar los algoritmos, utilizamos el indicador del hipervolumen (ν) que calcula la fracción del espacio objetivo que está no-dominada por ninguna de las soluciones obtenidas por el algoritmo [3]. Las estadísticas mostradas en la tabla 4 indican que ENORA obtiene

Tabla 3: Soluciones no-dominadas.

M	L	CR-Train	CR-Eval	S
Preselección con nichos				
10	**9**	**0.820225**	**0.8**	**0.065554**
10	10	0.820225	0.8	0.045618
14	11	0.853933	0.9	0.056930
11	11	0.842697	0.7	0.072503
20	14	0.910112	0.8	0.099939
NSGA-II				
10	8	0.910112	0.8	0.077349
11	**8**	**0.932584**	**0.8**	**0.077349**
ENORA				
10	**9**	**0.943820**	**0.8**	**0.083848**
15	10	0.955056	0.8	0.093499

menores valores de localización y dispersión que Preselección con nichos y NSGA-II. Por último, los intervalos de confianza del 95% para la media obtenidos con t-test muestran que ENORA obtiene menores valores que Preselección con nichos y NSGA-II. Así que, los conjuntos de aproximación obtenidos con ENORA son preferibles a los de Preselección y los de NSGA-II bajo el indicador del hipervolumen ν. El t-test es robusto con muestras mayores de 30 individuos, por tanto los resultados son significativos y podemos concluir que existe diferencia estadística entre los valores de hipervolumen obtenidos por los algoritmos. Los diagramas de caja mostrados en la figura 1 confirman las conclusiones afirmadas anteriormente. El análisis estadístico muestra, por tanto, que para el tipo de problemas multiobjetivo que estamos considerando, la búsqueda Pareto basada en la partición del espacio de búsqueda en tramos radiales es más eficiente que las estrategias de búsqueda generales basadas exclusivamente en funciones de diversidad, como NSGA-II, o en esquemas de diversidad con formación explícita de nichos, como el algoritmo de Preselección con nichos.

Tabla 5: Estadísticas para el hipervolumen.

	Preselec.	**NSGA-II**	**ENORA**
Mínimo	0.2929	0.2593	0.2425
Máximo	0.3336	0.3821	0.3296
Media	0.3201	0.3137	0.2827
D.E.	0.0086	0.0235	0.0137
I.C. Inferior	0.3184	0.3090	0.2799
I.C. Superior	0.3218	0.3184	0.2854

D.E. = Desviación estándar de la media
I.C. = Intervalo de confianza de la media (95%)

5 CONCLUSIONES

Este artículo muestra importantes resultados en la combinación de algoritmos evolutivos multiobjetivo y modelado difuso para clasificación. Se ha considerado un problema real de gran relevancia en el dominio

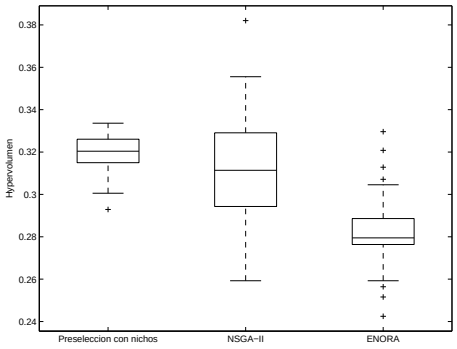


Figura 1: Diagramas de caja para el hipervolumen.

clínico consistente en la clasificación de la mortalidad por infección en pacientes con quemaduras graves. Se ha propuesto un modelo de optimización multiobjetivo con restricciones para clasificación difusa el cual tiene en cuenta criterios como precisión, transparencia y compacidad. Se han implementado tres algoritmos evolutivos multiobjetivo (Preselección con nichos, ENORA y NSGA-II) en combinación con técnicas de simplificación de reglas.

Los resultados obtenidos mejoran los de otras técnicas usadas habitualmente en medicina, con la ventaja de que la técnica propuesta identifica un conjunto de soluciones alternativas. Los test estadísticos se han realizado sobre los valores del indicador de calidad del hipervolumen para comparar los algoritmos y se ha mostrado que, para el problema considerado, ENORA consigue mejores resultados que los algoritmos de preselección con nichos y NSGA-II.

Agradecimientos

Desarrollo financiado en parte por el MEC y FEDER bajo los proyectos PET2006_0406 y TIN2006-15460-C04. Nuestro agradecimiento al Dr. Francisco Palacios Ortega del Hospital Universitario de Getafe por su asesoramiento y colaboración.

Referencias

[1] Casillas, J., Cordón, O., Herrera, F., Magdalena, L. (2003) Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview, in J. Casillas, O. Cordón, F. Herrera, L. Magdalena (Eds.), Interpretability Issues in Fuzzy Modeling, Studies in Fuzziness and Soft Computing, Springer, pp. 3-22.

[2] Coello, C.A., Veldhuizen, D.V., Lamont, G.B. (2002). Evolutionary Algorithms for Solving Multi-Objective Problems. Kluwer Academic/Plenum publishers, New York.

[3] Deb, K., Pratap, A., Agarwal S., Meyarivan T. A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transaction on Evolutionary Computation 6:2 (2002) 181-197.

[4] Klir, G.J., Folger, T.A. (1992) Fuzzy Sets, Uncertainty, and Information. Prentice-Hall.

[5] Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley.

[6] Gómez-Skarmeta, A.F., Jiménez, F. (1999). Fuzzy modeling with hibrid systems. Fuzzy Sets and Systems, 104:199-208.

[7] Gómez-Skarmeta, A.F., Jiménez, F., Sánchez, G. (2007). Improving Interpretability in Approximative Fuzzy Models via Multiobjective Evolutionary Algorithms. International Journal of Intelligent Systems, 22:943-969, 2.007

[8] Ishibuchi, H., Murata, T., Trksen, I. (1997). Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. Fuzzy Sets and Systems, 89:135-150.

[9] Ishibuchi, H., Nakashima, T., Murata, T. (1999). Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. IEEE Transactions on Systems, Man, and Cubernetics - Part B: Cybernetics, 29(5):601-618.

[10] Jiménez, F., Gómez-Skarmeta, A.F., Sánchez, G., Deb, K. (2002). An evolutionary algorithm for constrained multi-objective optimization, in: Proceedings IEEE World Congress on Evolutionary Computation.

[11] Sánchez, G., Jiménez, J., Vasant, P. (2007). Fuzzy Optimization with Multi-Objective Evolutionary Algorithms: a Case Study. IEEE Symposium of Computational Intelligence in Multicriteria Decision Making (MCDM). Honolulu, Hawaii

[12] Setnes, M., Babuska, R., Verbruggen, H. B. (1998). Rule-based modeling: Precision and transparency. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications & Reviews, 28:165-169.

[13] Spanish Society of Intensive-Critical Medicine and Coronary Units and Spanish Society of Emergency. (2007). Generalized infection mortality could be 20% off (in Spanish).

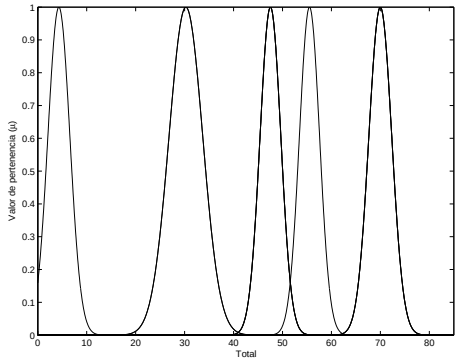
[14] Takagi, T., Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control. IEEE Transactions on Systems, Man and Cybernetics, 15:116-132.

Tabla 1: Parámetros considerados de los pacientes. El tipo puede ser real (R) o booleano (B).

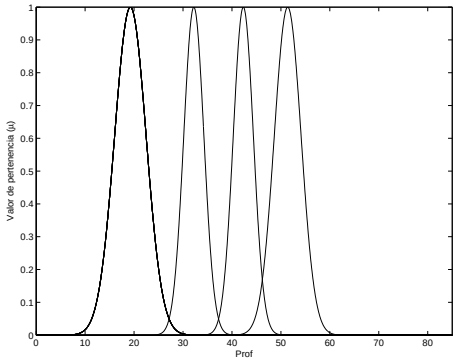
Nombre	**Descripción**	**Tipo**	**Peso**	**Atributo**	**Límites**
Total	Superficie total quemada en %	R	0.164613	x_1	0 − 85
Prof	Superficie quemadura profunda %	R	0.115335	x_2	0 − 85
SAPS II	Indicador de gravedad general	R	0	−	
Peso	Peso del paciente	R	0	−	
Edad	Edad del paciente	R	0	−	
Pneumonia	Presencia de infección pulmonar	B	0.207799	w_1	1 , 2
Sexo	Sexo del paciente	B	0.069403	w_2	1 , 2
Inh	Utilización de inhibidores	B	0.067112	w_3	1 , 2
Wound-Infect	Infección por herida quirúrjica	B	0.034936	w_4	1 , 2
Sida-drogas	Consumo de drogas con VIH	B	0.025781	w_4	1 , 2
Co-Hepática	Problemas hepáticos previos	B	0.024223	w_5	1 , 2
Bacteremia	Presencia de bacterias en sangre	B	0.023324	w_6	1 , 2
Co-Cardiaca	Cardiopatías previas	B	0.008141	w_7	1 , 2
Co-Respir	Problemas respiratorios previos	B	0.007035	w_8	1 , 2
HTA	Hipertensión arterial	B	0.001650	w_9	1 , 2
Diabetes	Paciente diabético	B	0.000541	w_{10}	1 , 2
Co-Renal	Problemas renales previos	B	0	−	
Muere	Previsión de fallecimiento	B	−	y	1 , 2

Tabla 4: Modelo difuso obtenido con ENORA.

x_1	x_2	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	y
(30.26; 23.72)	(19.25; 20.79)	1	1	1	1	2	1	2	2	1	2	1	1
(47.54; 9.02)	(19.25; 20.79)	2	1	1	2	1	1	2	2	1	1	2	1
(4.31; 10.11)	(19.25; 20.79)	1	2	2	2	2	1	2	2	1	1	2	1
(55.51; 8.50)	(42.32; 8.50)	2	1	1	2	1	1	1	2	2	2	1	1
(69.98; 10.95)	(19.25; 20.79)	1	1	1	2	1	1	2	1	2	2	1	1
(47.54; 9.02)	(51.39; 15.07)	2	2	2	1	1	2	2	2	1	2	1	2
(69.98; 10.95)	(19.25; 20.79)	1	2	1	2	1	2	1	1	1	1	1	2
(30.26; 23.72)	(19.25; 20.79)	2	2	2	2	1	2	2	1	2	2	1	2
(47.54; 9.02)	(19.25; 20.79)	2	2	1	2	1	2	1	1	1	1	1	2
(69.98; 10.95)	(32.21; 8.50)	1	1	1	2	2	2	1	2	2	2	1	2



$\boldsymbol{x_1}$, $S = 0.073930$.



$\boldsymbol{x_2}$, $S = 0.083848$.

APLICACIONES DE LA LÓGICA BORROSA EN EL EQUIPO DE FÚTBOL ROBÓTICO TeamChaos

Francisco Martín[1] **Carlos E. Agüero**[1] **José M. Cañas**[1] **Pablo Barrera**[1] **Vicente Matellán**[2]

[1] Depto. de Sistemas Telemáticos y Computación. Universidad Rey Juan Carlos, {fmartin, caguero, barrera, jmplaza}@gsyc.es
[2] Depto. de Ingenierías Mecánica, Informática y Aeroespacial. Universidad de León, vicente.matellan@unileon.es

Resumen

El equipo TeamChaos ha participado en las últimas ediciones de la RoboCup en la categoría de robots con 4 patas. Varios de los subsistemas del software que controla los cuatro robots del equipo utilizan la lógica borrosa. En concreto en este artículo describimos el sistema de auto-localización de los robots que se basa en una mezcla de un sistema basado en una rejilla uniforme borrosa y una población de tamaño variable de filtros extendidos de Kalman.

Palabras Clave: Localización, robot, robocup, lógica borrosa.

1. INTRODUCCIÓN

Una de las habilidades más básicas que ha de ser capaz de exhibir un robot móvil es la de auto-localización [4]. Puede definirse como la capacidad de un robot de determinar su posición usando sus propios sensores. Esa posición puede calcularse sobre un mapa ya compilado del entorno, o determinarse mientras se construye el mapa (SLAM).

Las técnicas clásicas que se ha propuesto para resolver este problema varían enormemente dependiendo de los sensores disponibles en cada tipo de robot, del entorno y de la aplicación. Las más exitosas, consideradas ahora mismo estándar en la literatura, son las aproximaciones probabilísticas. Así por ejemplo, en [9, 8, 6, 5, 2] se han propuesto soluciones a la localización de robots móviles con ruedas equipados con sensores de ultrasonidos y/o láser en entornos de interiores no muy dinámicos.

La propuesta presentada en este artículo también utiliza esta misma aproximación. La diferencia principal de estos trabajos con nuestra propuesta es el tipo de robot al que está dirigida y el entorno. En nuestro caso robots se trata de robots con patas, donde la información odométrica es muy poco fiable y que no disponen de información instantánea de todo el entorno del robot (de los 360 grados alrededor), mientras que las aproximaciones anteriores están realizadas para robots con odometría muy precisa y una abundante información del entorno de 360º.

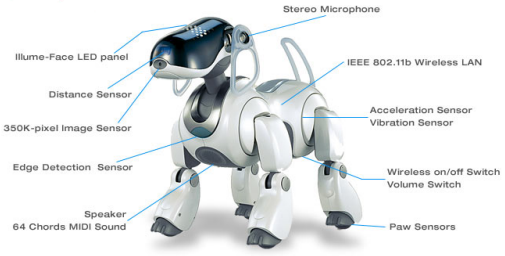


Figura 1: Sony aiBo ERS7

En concreto, en este trabajo utilizaremos el robot AIBO (figura 1). Este robot tiene como sensor principal una cámara situada en su cabeza. Las imágenes que se obtienen de la cámara han de ser procesadas para obtener información de ella. Otra característica de este robot es la de tener como principales actuadores de locomoción cuatro patas, lo que hace difícil obtener una información odométrica precisa.

Este modelo de robot se usa en la competición de la RoboCup[1] en la categoría de cuatro patas[2], que es como indicábamos la otra característica diferenciadora, el dinamismo y los requisitos de vivacidad de la competición. El campo de juego de esta categoría es el que refleja la figura 2, donde se puede apreciar que hay elementos fácilmente distinguibles por color, como las porterías, las balizas y las líneas del campo.

[1] http://www.robocup.org
[2] http://www.tzi.de/4legged/bin/view/Website/WebHome

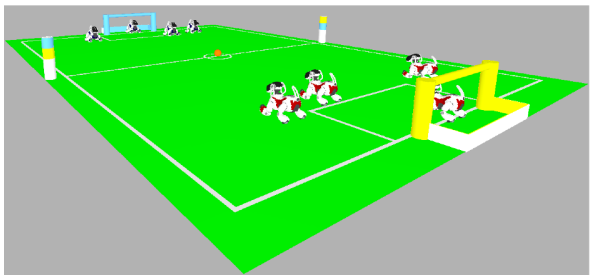


Figura 2: Campo de la RoboCup de la liga de 4 patas.

En este entorno es importante que los robots estén localizados en todo momento para que su comportamiento sea coherente: los robots no deben de salirse de los límites del campo; en determinadas situaciones deben situarse en una posición inicial conocida, como por ejemplo, al iniciar el partido o después de cada gol; también deben saber hacia dónde han de dirigir la pelota; en que parte del campo se encuentran para poder generar una estrategia común entre los miembros de un equipo, etc.

Los métodos de localización usados por los demás equipos son muy variados. La mayor parte de los equipos usan métodos basados en Filtros de Partículas. Nuestro equipo[3] ha usado tradicionalmente un método basado en lógica difusa que proporcionaba una localización global satisfactoria en el campo orignl[1], pero que resultaba demasiado costoso en el campo ampliado que se usa en las últimas competiciones. En este artículo describiremos los trabajos para mejorar esta localización.

El resto de este artículo se organiza de la siguiente forma. En la segunda sección presentamos los fundamentos de la localización probabilística borrosa que se ha venido utilizando en el equipo TeamChaos. En la tercera presentamos una posible mejora basada en la utilización de filtros de Kalman para mejorar el funcionamiento local. En la cuarta presentamos algunos resultados y en la quinta resumimos nuestra opinión.

2. Localización probabilística borrosa

El método que se ha venido utilizando hasta la fecha se ha denominado Fuzzy-Markov, o abreviadamente FMK. Una descripción detallada del mismo se puede encontrar en [3]. Básicamente, consiste en dividir el campo de juego en una cuadrícula G_t tal que $G_t(x, y)$ representa la probabilidad ($[0, 1]$) de que el robot se encuentre en la posición (x, y). Cada una de las celdas, que definiremos en adelante como *fcell*, contiene información sobre la probabilidad de que el robot esté en esa celda, e información sobre cual es el rango de orientaciones más probables para el robot, es decir, se trata realmente de una cuadrícula de $2\frac{1}{2}$D

Cada *fcell* se representa por medio de un trapezoide difuso. En la figura 3 podemos ver este trapezoide, definido por la tupla:

$$< \theta, \Delta, \alpha, h, b >$$

Intuitivamente, si h es bajo, la probabilidad de estar en esta celda es baja. Si h es alto, es muy probable que el robot esté en esta posición. Si el trapezoide es ancho, existe gran incertidumbre sobre la orientación del robot. Si el trapezoide es estrecho, o tiene incluso forma de triángulo (porque Δ es prácticamente nulo), la incertidumbre de orientación es tan baja que podemos afirmar que es θ.

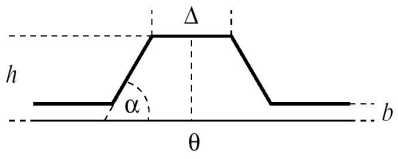


Figura 3: *fcell* representando el ángulo θ

El proceso de localización usado en este método es iterativo, teniendo cada ciclo un paso de predicción y otro de actualización. La fase de predicción se realiza cada vez que se realiza un movimiento, lo que hace que la probabilidad se distribuya (difumine) en la dirección del movimiento. En la fase de actualización se incorpora la información visual. Cada observación de una marca visual se compone de una distancia y un ángulo a cada una de las marcas visibles. Para codificar la información de una marca visual conocida en el instante t, construimos la distribución de probabilidad $S_t(\cdot|r)$, tal que $S_t(x, y|r)$ es la posibilidad de que el robot se encuentre en la posición (x, y), siendo la distancia a una marca visual determinada r. Para cada observación, actualizamos G_t como se muestra en la ecuación 1.

$$G_t(x, y) = G_t(x, y) \times S_t(x, y|r) \qquad (1)$$

donde $\times$ es un operador de intersección difuso.

Un ejemplo de la aplicación secuencial de estas dos operaciones para la localización de un robot se puede observar en la figura 4. El robot parte de un estado de total incertidumbre, y mediante la información odométrica y la información de la posición relativa de la portería y de la baliza superior derecha, termina localizándose correctamente.

[3] http://www.teamchaos.es

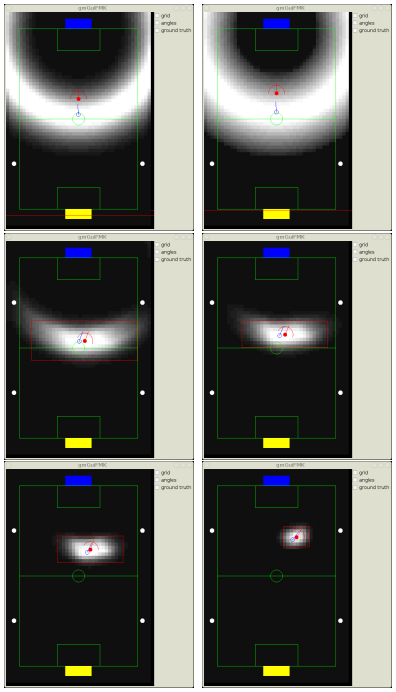


Figura 4: Proceso de localización del robot mediante grid borroso..

El proceso anterior calcula la probabilidad de estar en una posición de la cuadrícula, pero no aporta información angular. Podría ser natural considerar un cubo 3D $G_t(x, y, \theta)$, pero el tiempo de computación de todas las posibles posiciones del robot haría que el algoritmo no fuera abordable computacionalmente. En lugar de eso se mantiene una cuadrícula 2D de *fcell*, que es una forma compacta de representar información angular. Con el procedimiento descrito anteriormente se obtiene la componente h de esta *fcell*, pero aún es necesario describir como obtener el resto de las componentes cuando se produce una observación. Esta operación es muy ligera y calcula las nuevas probabilidades de orientación a partir de una estimación anterior y una nueva observación.

Este algoritmo es un método global, es decir, se almacena la probabilidad de que el robot esté en cada cuadrícula y orientación. El tamaño de cada celda es de 100 mm de lado, lo que significa que para el campo de juego original, cuyas medida eran 2700×4200 mm, se necesitaban 1134 celdas. Para una división típica en rejilla que almacenara 16 posibles orientaciones por cada celda, que es lo mínimo para tener una precisión en cuanto a la orientación aceptable, supondría tener 18144 estados.

El campo ha aumentado sus dimensiones. Las nuevas dimensiones, 2700×4200 aumentan el número de celdas a más del doble, 2400. El aumento de coste compu-

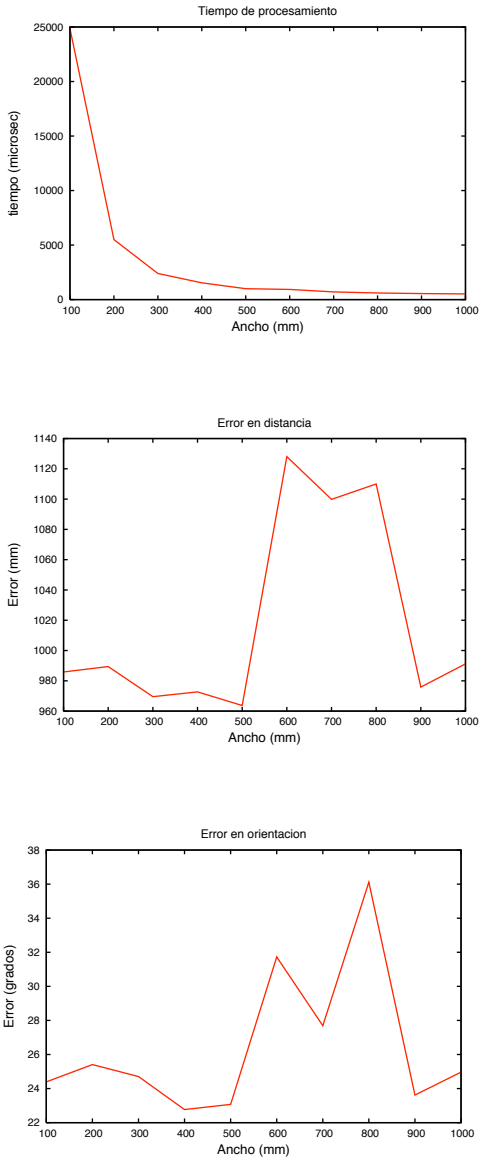


Figura 5: Variación del tiempo de computación (izquierda), error en la estimación de x, y (centro) y error en estimación θ dependiendo del tamaño de la celda.

tacional, al ser FMK un sistema $O(n^2)$, ha hecho que el tiempo de cálculo del sistema de localización impida que el sistema completo de control del robot pueda ejecutarse en un solo ciclo de control, que corresponde al tiempo entre que se procesa una imagen y la siguiente. La solución inicial consistió en aumentar el tamaño de la celda, pero esto supone que en el mejor de los casos el sistema tiene una precisión menor, lo cual resultaba inaceptable.

El tiempo de proceso de los algoritmos de localización es crítica. Hay que tener en cuenta que el robot está ejecutando múltiples módulos (coordinación, asignación de roles, visión, cálculo de movimientos de las patas, etc.). La suma de los tiempos de cada uno no debe exceder del tiempo de ciclo. Este tiempo de ciclo está marcado por la frecuencia en la que se recibe una imagen desde la cámara, que desencadena todo el procesamiento subsiguiente. El robot percibe una nueva imagen, en media, cada $33810\mu s$.

Uno de los principales requisitos como ya hemos mencionado es la reducción del tiempo de computación del módulo de localización. El tiempo de este módulo depende sobre todo del número de celdas que conforman la rejilla que divide al campo.

En la figura 5 mostramos un análisis del tiempo de cómputo y la precisión obtenida en distancia y en orientación que persigue averiguar el tamaño de celda óptimo. Los tamaños de celda que se han analizado varían desde 100 mm hasta 1000 mm para el ancho de cada celda. En la primera figura se observa como la diferencia entre 100 mm, que es el ancho de celda que se usa en el algoritmo *FMK* y tamaños de 300 mm a 400 mm es considerable, no obteniendo mejoras sensibles a partir de 500 mm. La gráfica central muestra el error que el algoritmo *FMK* presenta. El mínimo se sitúa en 500 mm y tamaños mayores no son capaces de localizar al robot. El mismo resultado se observa en el caso de la gráfica situada a la derecha, que corresponde al error en la orientación. Este comportamiento del algoritmo se debe a que el tamaño de celda es tan elevado que es verosímil obtener medidas sensoriales válidas en ellas.

3. Combinación de FMK con EKF

La solución que hemos adoptado ha consistido en aumentar el tamaño de la celda para mantener un sistema de localización global, y conseguir la precisión local requerida mediante un Filtro Extendido de Kalman (no indicaremos los detalles del filtro, únicamente indicar que estima la posición y la orientación en el plano).

La alternativa de utilizar únicamente un método de localización local como es el EKF no es válida. Básicamente, los problemas están relacionadas con la incapacidad del filtro para localizar al robot partiendo de una posición desconocida o cuando el robot sea desplazado manualmente de un lugar a otro del entorno. También es complicado evaluar si la estimación de la posición del robot se ha degradado debido a errores sensoriales.

En esta sección se presentará la forma en la que la combinación del método global *FMK* y uno o varios EKFs pueden solucionar estos problemas. En concreto, plantearemos dos estrategias de combinación de los métodos: combinar un filtro extendido de Kalman con *FMK* (*KFMK*) o varios EKFs con *FMK* (*NKFMK*) . En las siguientes secciones se mostrará en detalle como se combinan ambos métodos.

3.1. *KFMK*

La primera estrategia de combinación de varios métodos combina un único filtro extendido de Kalman con *FMK*. Como se puede observar en la figura 6, ambos métodos se ejecutan de manera independiente, teniendo las mismas entradas, que son la información sensorial en forma de odometría y la posición relativa a los objetos.

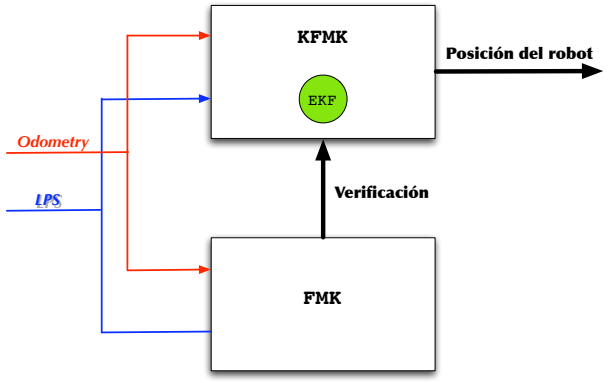


Figura 6: Esquema de combinación de un filtro extendido de Kalman y *FMK*.

3.2. Combinación de varios EKF con FMK

El método descrito en la sección anterior combina un *EKF* y *FMK*. Como la experimentación mostró que era poco costoso desde el punto de vista computacional mantener varios EKFs, nos planteamos usar mas de uno. Esto podría aportar una mejor respuesta a situaciones de secuestro y manejar varias hipótesis.

En resumen, esta segunda estrategia de combinación usa varios filtros extendidos de Kalman que se ejecutan en paralelo con *FMK*. El esquema mostrado en la figura 7 es parecido a la estrategia de combinación des-

crita anteriormente, pero con varios filtros extendidos de Kalman.

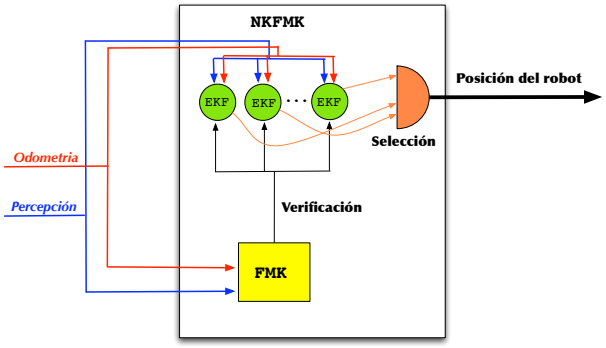


Figura 7: Esquema de combinación de varios filtros extendidos de Kalman y *FMK*.

El número de *EKF* varía durante la operación del robot según las necesidades del sistema. Al principio se crea un único *EKF*. Cada vez que *FMK* diverge de la estimación del EKF se inicia un nuevo filtro en la posición que estima FMK hasta alcanzar el número máximo de filtros permitidos. Si ya se ha alcanzado el número máximo de filtros permitidos, se comprueba la calidad del filtro que depende de la incertidumbre.

Si dos o más filtros convergen a una posición muy cercana, se elimina el filtro que tiene peor calidad. Esto hace que se tenga el menor número de filtros posible que cubran todas las hipótesis posibles. La posición del mejor filtro será la que se transmita hacia los otros módulos.

4. Validación experimental

El objetivo de la experimentación es comprobar el grado de precisión de las diferentes opciones, en concreto vamos a analizar el uso de un EKF, del sistema FMK original frente a las dos opciones propuestas en la sección anterior, FMK combinado con un EKF y FMK combinado con n EKFs.

Se han realizado diferentes trayectorias, en concreto analizaremos el recorrido descrito en la figura 8 que incluye una trayectoria por el campo y un "secuestro", es decir, un traslado de posición del robot por un agente extraño al robot (no por los propios movimientos del robot). En color rojo se muestra el recorrido que realiza el robot sobre el terreno de juego. Después del punto de control número 2, se produce un secuestro del robot desde el punto "so" (secuestro origen) al punto "sd" (secuestro destino). Este experimento también se ha repetido 28 veces, siendo la duración del mismo de unos 35 a unos 40 segundos.

Todos los resultados presentados en esta sección han

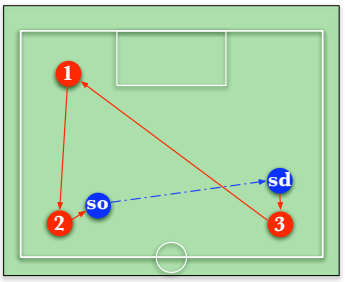


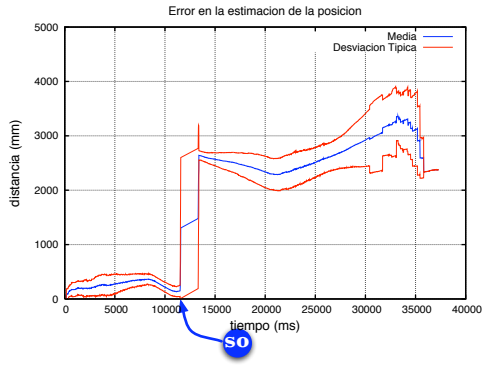
Figura 8: Recorrido del experimento con secuestro.

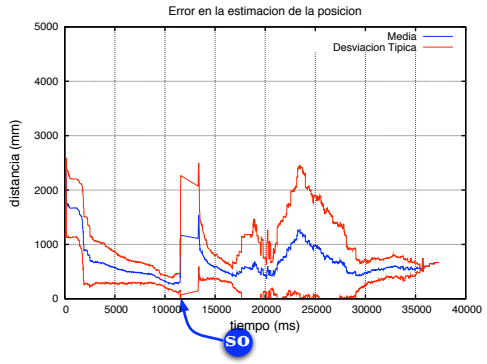
sido generados a partir de datos procedentes de experimentación con un robot real. Durante los experimentos se recogió en ficheros la información del sistema de "verdad absoluta" y la información perceptiva que el robot usa como entrada al módulo de auto-localización. Posteriormente se han analizado las ejecuciones en un ordenador. Durante estas ejecuciones es posible evaluar distintos métodos de auto-localización con distintas configuraciones, y obtener los resultados experimentales presentados.

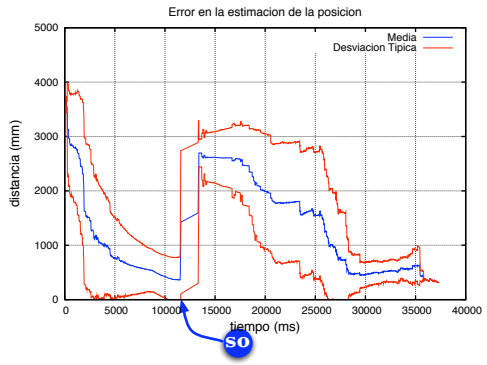
Las gráficas de la figura 9 presentan el error medio después de 28 ejecuciones de la trayectoria descrita y el secuestro para cada uno de los métodos (EKF, FMK, EKF+FMK y FMK+nEKF respectivamente de arriba hacia abajo en la figura). La gráfica muestra el error medio y la desviación típica de dicho error en distancia (también se ha estudiado el error en el ángulo pero no se incluye por motivos de espacio) en el recorrido para cada uno de los métodos.

La comparación de las medias de las gráficas anteriores se muestran en la figura 10, donde se observa que la mayor parte del tiempo, salvo en el instante inmediatamente posterior al secuestro, el método *FMK+nEKF* tiene el error más bajo de todos los métodos comparados.

Sin embargo, no es el error medio la métrica más interesante para la aplicación del fútbol robótico. El dato del porcentaje de tiempo que el robot mantiene el error en la estimación por debajo de un umbral es muy descriptivo. Estos datos se encuentran recogidos en las tablas 1. Casi la mitad del tiempo, el error del método *FMK+nEKF* se encuentra por debajo de los treinta centímetros, lo que significa que la estimación de la posición del robot es muy buena. Por debajo del umbral de un metro, que puede entenderse como que el robot no está "perdido", los métodos combinados se encuentran el noventa por ciento del tiempo, mientras que el método *FMK* un sesenta por ciento. El método *FMK+EKF* queda en evidencia con estos datos, demostrando que el escaso error en media proviene del veinticuatro por ciento del tiempo que corresponde a







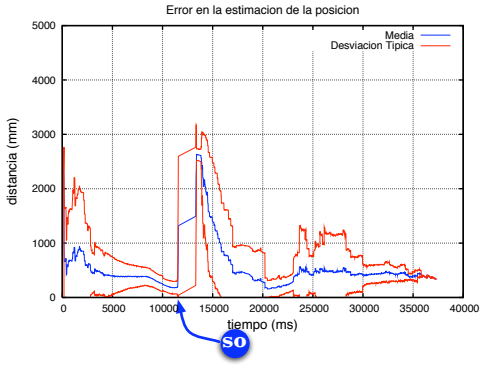


Figura 9: Media de error de en distancia entre la estimación en la posición y la posición real del robot.

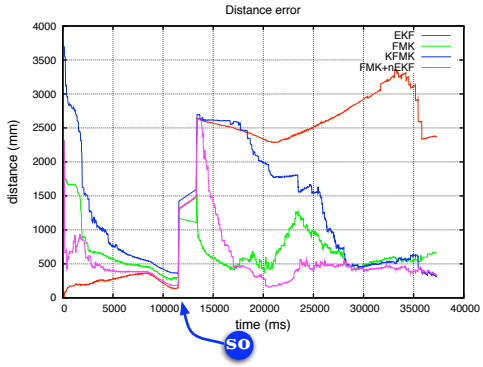


Figura 10: Media de error en distancia entre la estimación en la posición y la posición real del robot para EKF, FMK, EKF+FMK y FMK+nEKF

Rango	*EKF*	*FMK*	*FM* +EKF	*FMK* +nEKF
<= 300	24,60	24,09	21,25	45,32
<= 500	33,93	50,53	37,34	75,39
<= 1000	35,93	83,76	53,47	88,76
<= 1500	36,34	90,42	59,238	90,56

Cuadro 1: Análisis del porcentaje de tiempo que el error en distancia de la estimación de la posición del robot es inferior a un rango determinado (en mm).

su etapa inicial antes del secuestro en que su error era cercano a 0.

5. Conclusiones

En este artículo hemos presentado dos mecanismos para mejorar la precisión del mecanismo de autolocalización borrosa para robots con patas en el entorno de la RoboCup que hemos desarrollado. El experimento descrito demuestra que el algoritmo más interesante desde el punto de vista de la fiabilidad y precisión en distancia es el de fusión del FMK con n EKFs. Experimentos adicionales han demostrado que es igualmente el más fiable en ángulo y que su coste computacional es asumible.

Agradecimientos

Los autores quieren agradecer al resto de miembros del equipo TeamChaos, en particular a los Dres. Humberto Martínez y David Hernández su inmeso trabajo en el equipo. Este trabajo ha sido parcialmente financiado con fondos del Ministerio de Educación y Ciencia (DPI2007-66556-C03-01) y de la Comunidad Autónoma de Madrid (RoboCity 2030: S-0505/DPI/0176)

Referencias

[1] P. Buschka, A. Saffiotti, and Z.Wasik. Fuzzy landmark-based localization for a legged robot. In *Proceedings of the International Conference on Intelligent Robots and Systems 2000*, Takamatsu, Japan, October 2000.

[2] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research*, 11:391–427, 1999.

[3] D. Herrero-Pérez, H. Martínez-Barberá, and A. Saffiotti. Fuzzy self-localization using natural features in the four-legged league. *Lecture Notes in Computer Science. Robocup 2004*, 3276, 2005.

[4] J.Borenstein, B.Everett, and L.Feng. *Navigating mobile robots: Systems and techniques*. Ltd. Wesley, MA, 1996.

[5] Jana Koseck¡E1¿and Fayin li. Vision based topological markov localization. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, Barcelona (Spain), April 2004.

[6] María E. López, Luis Miguel Bergasa, and M.S.Escudero. Visually augmented POMDP for indoor robot navigation. *Applied Informatics*, pages 183–187, 2003.

[7] Humberto Martínez, Vicente Matellán, and Miguel Cazorla. Teamchaos technical report. Technical report, TeamChaos, 2006.

[8] Dandapani Radhakrishnan and Illah Nourbakhsh. Topological localization by training a vision-based transition detector. In *Proceedings of IROS 1999*, volume 1, pages 468 – 473, October 1999.

[9] Reid Simmons and Sven Koening. Probabilistic navigation in partially observable environments. In *Proceedings of the 1995 International Joint Conference on Artificial Intelligence*, pages 1080–1087, Montreal (Canada), July 1995.

CONTROLADOR PARA NAVEGACIÓN REACTIVA EVITANDO OBSTÁCULOS MEDIANTE TÉCNICAS NEURO-FUZZY

Andrés A. Gersnoviez [1] **Iluminada Baturone** [2]

[1] Dpto. de Arquitectura de Computadores, Electrónica y Tecnología Electrónica, Universidad de Córdoba (España); el1gemia@uco.es
[2] Instituto de Microelectrónica de Sevilla, (IMSE-CNM-CSIC) y Dpto. Electrónica y Electromagnetismo, Universidad de Sevilla (España); lumi@imse.cnm.es

Resumen

Este artículo muestra cómo la combinación de técnicas neuro-fuzzy con análisis geométrico ofrece mejores resultados que las resoluciones mediante heurística pura y enfoque puramente físico a la hora de resolver el problema de navegación de robots tipo coche. El controlador descrito, que sigue una técnica reactiva, genera trayectorias de longitud mínima cuando no se detectan obstáculos, y en presencia de éstos genera desviaciones mínimas con respecto a ellas. Todas estas trayectorias de referencia respetan las restricciones cinemáticas de los robots tipo-coche. Además de su eficiencia, el controlador propuesto es bastante simple y lingüísticamente interpretable. Ha sido diseñado y verificado al completo usando las herramientas de CAD del entorno de trabajo Xfuzzy 3.

Palabras Clave: Control neuro-fuzzy, navegación reactiva, robots tipo coche, interpretabilidad lingüística, sistemas empotrados.

1 INTRODUCCIÓN

Una tarea básica que debe realizar un robot móvil autónomo es navegar de forma óptima y segura evitando posibles obstáculos que se interpongan en su camino. Puesto que se trata de una tarea para la cual se dispone de conocimiento heurístico, son varios los controladores difusos que se han desarrollado empleando sólo este tipo de conocimiento. Estos controladores son simples pero no proporcionan trayectorias óptimas [1], [5]. Durante nuestro estudio hemos comprobado este hecho con un controlador que diseñamos a partir sólo de heurística. Otros autores han ajustado el controlador mediante datos numéricos obtenidos a partir de las trayectorias proporcionadas por un experto [4], [7]. La desventaja de estos controladores es que pierden su significado lingüístico.

La aproximación que se describe en este artículo es emplear conocimiento heurístico para definir la estructura del controlador junto con información numérica para definir los distintos módulos constitutivos. La información numérica se ha obtenido a partir de un análisis geométrico aproximado que tiene en cuenta las características cinemáticas y dinámicas del robot. Se ha considerado un robot tipo coche que, por tanto, posee restricciones no holónomas, consistentes en que la dirección de la velocidad debe ser siempre tangente a la trayectoria y, además, el radio de curvatura del robot está mecánicamente limitado a un valor mínimo. Las ecuaciones resultantes de aplicar el análisis geométrico se han aproximado mediante interpoladores que, además, tienen interpretación lingüística. El controlador al completo ha sido diseñado gracias a las distintas herramientas de CAD del entorno de trabajo Xfuzzy 3, desarrollado en el IMSE [6]. En el apartado 3 se muestran los resultados obtenidos por simulación. Por último, en el apartado 4 se exponen las conclusiones.

2 EL ALGORITMO DE NAVEGACIÓN

El modelo cinemático que vamos a utilizar para describir el robot usa como variables básicas la posición cartesiana y la velocidad del punto situado en el centro del eje trasero del vehículo, la orientación ϕ del eje del vehículo y su velocidad angular. Estas variables se muestran en la figura 1. En este modelo se supone que no se produce deslizamiento en las ruedas durante el movimiento debido al suficiente rozamiento y a que las velocidades con las que vamos a trabajar son suficientemente bajas. Este modelo ha sido ya validado experimentalmente por los autores con otros controladores distintos al que aquí se describe [8].

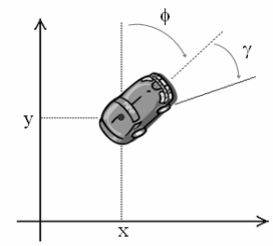


Figura 1: Modelo cinemático del vehículo.

En nuestro caso, no se usa ningún modelo que no sea el robot en cuestión para la navegación, y la información que se recibe es la de sus sensores. En particular, consideramos el láser 2D situado en la parte delantera del vehículo que realiza un barrido de 180º. El láser identifica los puntos de posibles obstáculos a través de su distancia, *h*, y ángulo crítico, φ. El problema de navegación consiste en generar una trayectoria libre de colisiones desde una configuración inicial (x, y, ϕ, v, γ) a otra objetivo. En el sistema de coordenadas global que hemos seleccionado, la configuración objetivo es (0, 0, 180º, 0, 0).

Como perseguimos el diseño de un controlador en tiempo real de bajo coste, se ha elegido una estrategia de navegación reactiva que sólo explora un pequeño área para detectar obstáculos; dicha área debe ser lo suficientemente grande para permitir realizar la maniobra de esquivación. El análisis de un área mayor implicaría un coste computacional mucho mayor e incrementaría innecesariamente el número de obstáculos a considerar. Para el diseño del controlador, hemos seguido la siguiente heurística:

- Si hay un obstáculo *muy cerca* del robot, el robot debe detenerse para evitar la colisión.
- Si hay un obstáculo *cerca* del robot, el robot debe realizar un maniobra para esquivar dicho obstáculo.
- Si el obstáculo está *lejos* del robot, en ese caso no es necesario esquivarlo (por el momento), siguiendo la trayectoria para alcanzar la posición objetivo.

2.1. OBSTÁCULOS MUY CERCANOS

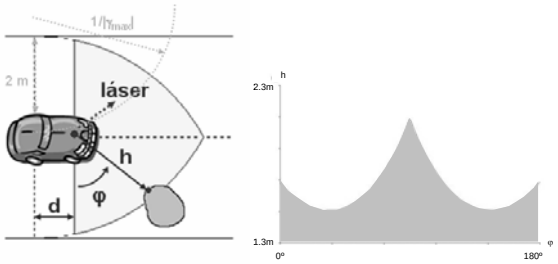
En primer lugar vamos a centrar nuestro estudio en el tipo de obstáculos que son imposibles de evitar. Teniendo en cuenta que la curvatura del robot está limitada a γ_{max} y que existe una zona de seguridad de 2m a ambos lados del robot, en los cuales no deben entrar ningún obstáculo (el robot considerado tiene 1m de anchura y la velocidad máxima considerada es de 1m/s), haciendo un análisis geométrico del problema obtenemos que, aproximadamente, los obstáculos que entran en el área sombreada de la figura 2a son obstáculos *muy cercanos* y, por tanto, inevitables.

Tomando las coordenadas del obstáculo respecto al sistema de referencia cuyo origen de coordenadas se sitúa en el punto medio del eje trasero del robot, el área de obstáculos muy cercanos vendría descrita por la ecuación (1):

$$h^2 + 2h\left[d|\sin\varphi| + R_{\min}|\cos\varphi|\right] < 4(R_{\min}+1) - d^2 \qquad (1)$$

donde R_{min} es el radio del arco de máxima curvatura.

El área sombreada en la figura 2b contiene los puntos que satisfacen la ecuación (1).



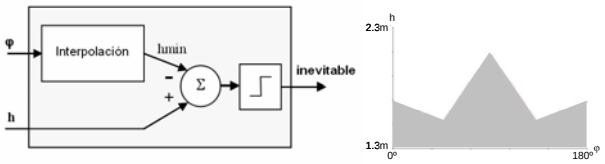
(a) (b)

Figura 2: (a) Zona de obstáculos *muy cercanos*; La misma zona según (1)

Como podemos ver, la ecuación (1) no explica lingüísticamente lo que debe hacer el robot. Para diseñar un módulo difuso que interpole esta información, tenemos dos opciones: un clasificador monolítico de dos entradas, *h* y φ, y una única salida, que nos informa si el obstáculo es inevitable o no; y un sistema jerárquico.

Aunque los resultados obtenidos por el sistema monolítico son bastante buenos, el caso en el que usamos jerarquía simplifica más nuestro problema y por eso es el que hemos seleccionado. La estructura jerárquica considerada se muestra en la figura 3a. En este caso estudiamos la frontera de obstáculos muy cercanos en función de *h* y φ, esto es, considerando en la ecuación (1) "=" en lugar de "<", creando un sistema difuso ("*interpolación*") cuya única entrada es φ y cuya salida nos da el valor de *h* para dicha frontera, es decir, el valor mínimo que puede tomar *h* para que el obstáculo no sea considerado como inevitable. A continuación se le resta al valor de entrada *h* el valor de la salida del sistema difuso "*interpolación*", de modo que si se obtiene un resultado negativo nos encontramos ante un obstáculo *muy cercano*, no siéndolo en caso contrario. Usando las herramientas de Xfuzzy 3, el sistema difuso obtenido está compuesto por las siguientes reglas:

1 – Si φ está *muy a la izquierda* o *muy a la derecha*, entonces h_{min} es 1.7m.
2 – Si φ está *a la izquierda* o *a la derecha*, entonces h_{min} es 1.5m.



(a) (b)

Figura 3: (a) Sistema jerárquico para detectar obstáculos *muy cercanos*; (b) Salida del sistema jerárquico simplificado

3 – Si φ está *enfrente* del robot, entonces h_{min} es 2.1m.

El área sombreada en la figura 3b contiene los puntos que este sistema clasifica como inevitables.

2.2. OBSTÁCULOS CERCANOS

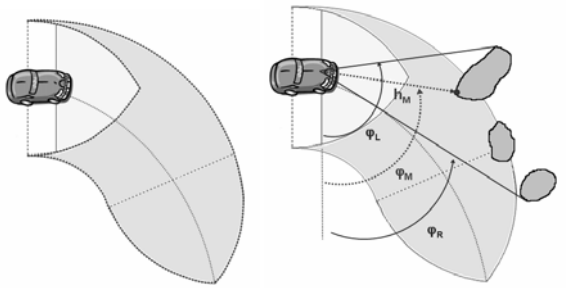
Consideremos ahora los obstáculos que están *cerca* del robot y pueden ser esquivados. Para ello vamos a prever en dónde se encontrará el robot en los próximos segundos, teniendo en cuenta su velocidad y su curvatura actual, γ.

El pasillo de detección estará formado por los dos arcos de circunferencia con radios (R+2) y (R-2), siendo R el radio de la curvatura actual, ya que por las razones de seguridad que comentamos anteriormente, el pasillo debe tener 2m a cada lado, y ahora añadimos 2m en la dirección de movimiento (debido a que la velocidad del vehículo es de 1m/s). Los obstáculos que se adentren en dicho pasillo serán considerados *cercanos* y, por tanto, deben ser esquivados. El pasillo en cuestión se puede ver representado en la figura 4a, donde la zona sombreada de oscuro es la de obstáculos *cercanos*. La forma del pasillo se corresponde al caso en que el robot está girando a la derecha, y su forma será más o menos abrupta según si el valor de γ es mayor o menor. Considerando un sistema de coordenadas relativo al robot, las ecuaciones que definen dicha área son:

$$(2-\frac{1}{|\gamma|})^2 < x_r^2 + (y_r + \frac{1}{\gamma})^2 < (2+\frac{1}{|\gamma|})^2 \text{ and } x_R < f(y_R) \quad (2)$$

donde (x_r, y_r) son las coordenadas del obstáculo referidas al sistema de referencia relativo y (x_R, y_R) son las coordenadas del obstáculo referidas al sistema de referencia desplazado 2m en la dirección de navegación:

$$\begin{aligned} x_r &= d + h\cdot\sin\varphi \qquad y_r = -h\cdot\cos\varphi \\ x_R &= (x_r - \frac{\sin 2\gamma}{\gamma})\cos 2\gamma - (y_r + \frac{1-\cos 2\gamma}{\gamma})\sin 2\gamma \\ y_R &= (x_r - \frac{\sin 2\gamma}{\gamma})\sin 2\gamma + (y_r + \frac{1-\cos 2\gamma}{\gamma})\cos 2\gamma \end{aligned} \quad (3)$$



(a) (b)

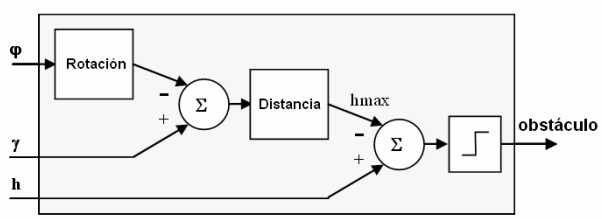
Figura 4: (a) Área de obstáculos *cercanos*; (b) Parámetros para calcular la nueva curvatura

La función f() en la ecuación (2) representa la relación existente entre x_R e y_R cuando se sustituye "<" por "=" en la ecuación (1).

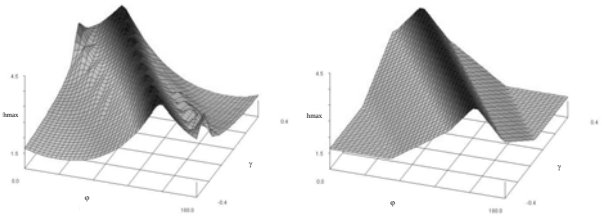
Al igual que en el caso de obstáculos *muy cercanos*, tenemos dos opciones para definir un clasificador difuso que decida si un obstáculo es *cercano* o no empleando los datos numéricos (φ, γ, h_{max}) obtenidos a partir de la ecuación (2): un sistema monolítico y otro jerárquico. El sistema monolítico tiene ahora tres entradas, *h*, φ, y γ, y una única salida. Como consecuencia de esto, se ha seleccionado directamente la opción del sistema jerárquico (figura 5a), que sólo contiene bloques con una entrada y una salida. El bloque llamado "*rotación*" estima una relación entre γ y φ que permite al bloque llamado "*distancia*" estimar el valor máximo de *h*, h_{max}, con el que un obstáculo puede ser considerado como *cercano*. El siguiente paso es restarle a *h* el valor obtenido de h_{max} y, si el resultado es negativo se trata de un obstáculo "cercano", no siéndolo en caso contrario. Usando las herramientas de Xfuzzy, se han extraído 2 reglas para la base de reglas *"rotación"* y 3 para la base de reglas *"distancia"*. Las últimas pueden expresarse lingüísticamente de la siguiente forma:

- Si el obstáculo se encuentra *enfrente* del robot en futuras configuraciones, entonces h_{max} es 4.0m.
- Si el obstáculo se encuentra *hacia la derecha* o *hacia la izquierda* del robot en futuras configuraciones, entonces h_{max} es 2.0m.
- Si el obstáculo se encuentra *bastante a la derecha* o *bastante a la izquierda* del robot en futuras configuraciones, entonces h_{max} es 1.6m.

De acuerdo con (2) y (3), la superficie de la figura 5b muestra los valores máximos de *h*, h_{max}, que tienen los obstáculos considerados como "cercanos" dependiendo de su posición angular φ, y la curvatura del robot, γ. La superficie de la figura 5c muestra la salida de la base de reglas "distancia" en el sistema jerárquico de la figura 5a.



(a)



(b) (c)

Figura 5: (a) Sistema jerárquico para clasificar si un obstáculo es *cercano* o no; (b) Distancia máxima a la que un obstáculo es considerado *cercano* según (2); (c) Salida del sistema neuro-fuzzy jerárquico

2.3. EVITACIÓN DE OBSTÁCULOS

Una vez se ha detectado un obstáculo *cercano*, el controlador debe modificar la trayectoria del vehículo para evitar la colisión. La mínima magnitud de la curvatura viene determinada por el punto más cercano del obstáculo al robot (h_M y φ_M en la figura 4b). Teniendo en cuenta que debemos mantener una distancia de 2m a cada lado del corredor y que la curvatura de referencia aportada por el controlador no se adopta instantáneamente por el robot sino que sufre un retraso, mediante un análisis geométrico, obtenemos que el valor de curvatura mínima para evitar dicho obstáculo es:

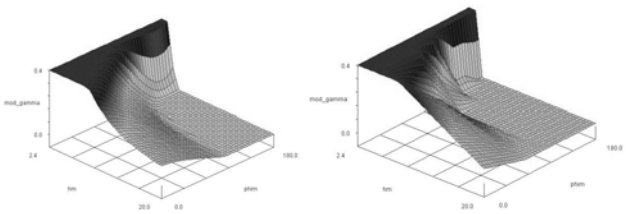
$$|\gamma| = \frac{8 + 2dh_M \cos\varphi_M}{d^2 + h_M^2 + 2dh_M \sin\varphi_M - 16} \qquad (4)$$

Sacando provecho de la simetría del problema ($\gamma = f(\varphi_M, h_M)$, cuando se gira a la derecha y $\gamma = f(180-\varphi_M, h_M)$ cuando se gira a la izquierda), solamente ha sido analizado el problema del giro hacia la derecha. En la figura 6a se muestra los valores de γ (en el caso de giro a la derecha) frente a φ_M y h_M.

Al igual que en los casos anteriores, se tiene la opción de diseñar tanto un sistema monolítico como uno jerárquico. En este caso, hemos seleccionado el sistema monolítico puesto que su complejidad era menor (5 reglas frente a las 9 obtenidas de forma jerárquica), consiguiendo una mejor aproximación. El sistema obtenido, que decide cuánto debe girarse a la derecha (una vez que se ha decidido girar a la derecha), fuerza al robot a girar más mientras más peligroso sea el obstáculo a esquivar. Esto se puede deducir observando las reglas que aplica dicho módulo:

- Si h_M es *muy pequeña*, entonces gira al máximo a la derecha.
- Si h_M es *pequeña* y φ_M no está *demasiado a la izquierda*, entonces gira al máximo a la derecha.
- Si h_M es *mediana* y φ_M está *enfrente* o *a la derecha* del robot, entonces gira al máximo a la derecha.
- Si h_M es *grande* y φ_M está *a la derecha*, entonces gira al máximo a la derecha.
- En otro caso, continua la trayectoria.

Como estas reglas emplean conceptos difusos, los valores que el controlador proporciona no cambian de forma abrupta, sino que lo hacen de forma suave, como se puede apreciar en la figura 6b.



(a) (b)

Figura 6: (a) Curvatura para esquivar obstáculos según (4); (b) Salida del sistema neuro-fuzzy monolítico

Para concretar el signo de la curvatura, el controlador evalúa la apertura angular (φ_L, φ_R) vista por el láser, tal y como podemos apreciar en la figura 4b, aplicando las siguientes reglas heurísticas:

1 - Si el obstáculo se encuentra a la derecha, entonces gira a la izquierda (y viceversa).
2 - Si el obstáculo se encuentra enfrente del robot, entonces:
a) Si el obstáculo ya haya sido detectado en ciclos anteriores, gira en la misma dirección que en el ciclo de control previo.
b) Si es la primera vez en que se detecta el obstáculo, gira en el mismo sentido en el que se giraría si no hubiera obstáculo.

2.4. NAVEGACIÓN HACIA LA CONFIGURACIÓN OBJETIVO

En el caso de no detectarse obstáculos, el robot debe navegar hacia la configuración final (0, 0, 180º, 0, 0) siguiendo el camino más corto. Los caminos más cortos para un robot tipo-coche consisten en una secuencia finita de dos componentes elementales: arcos de circunferencia y segmentos de línea recta, como demuestra Dubins en [3]. Analizando geométricamente los caminos más cortos, se puede encontrar un ángulo, α, que define la orientación

del segmento que contiene el punto (x, y) de la configuración actual del robot y su tangente al arco de circunferencia que define el final del camino (figura 7).

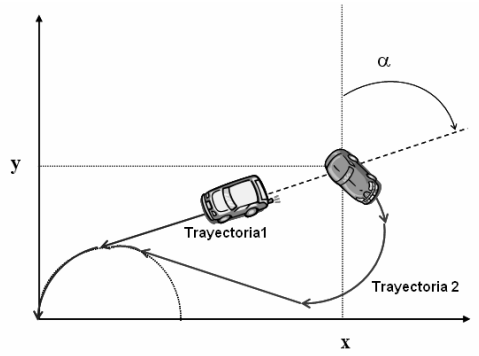


Figura 7: Caminos de longitud mínima para robots tipo-coche

Dependiendo de la diferencia entre este ángulo y la orientación actual del robot, éste girará hacia la derecha o hacia la izquierda o directamente no girará (si la diferencia es cero). La dependencia del ángulo α con *x* e *y* se describe en la ecuación (5):

$$\alpha = \begin{cases} sign(x)\cdot ar\cos(\psi) & \text{si} \quad \left(|x| - R_{\min}\right)^2 + y^2 > {R_{\min}}^2 \\ sign(x)\cdot ar\cos\left(1 - \dfrac{|x|}{R_{\min}}\right) & \text{si} \quad \left(|x| - R_{\min}\right)^2 + y^2 \le {R_{\min}}^2 \end{cases} \quad (5)$$

donde R_{min} es el radio de mínima curvatura y el valor de ψ viene dado por la ecuación (6):

$$\psi = \frac{R_{\min}\left(R_{\min} - |x|\right) + y\sqrt{x^2 + y^2 - 2R_{\min}|x|}}{x^2 + y^2 - 2R_{\min}|x| + R^2} \quad (6)$$

Una vez más, podemos optar tanto por un sistema difuso monolítico como por otro jerárquico. En este caso, debido a la mayor complejidad presentada por el sistema monolítico al tener tres entradas (x,y, ϕ), se ha optado por elegir directamente la opción jerárquica. El esquema jerárquico seleccionado se muestra en la figura 8a. Como en casos anteriores, dicho sistema está formado por dos bases de reglas conectadas en serie. La primera de ellas se ha generado y posteriormente simplificado hasta tener 9 reglas. Esta base genera un valor aproximado del ángulo α, dependiendo de las variables de entrada *x* e *y*. En las figuras 8b y 8c se ilustra cómo este módulo aproxima el análisis expuesto en la ecuación (5). La segunda base de reglas (con 2 reglas, sacando partido a la simetría del problema), genera un valor para la curvatura dependiendo de la diferencia entre ϕ y α. Dichas reglas son las siguientes:

- Si (ϕ – α) es *un tanto más pequeño que 180º*, entonces gira un poco a la derecha.

- Si (ϕ – α) es *mucho más pequeño que 180º*, entonces gira a la derecha al máximo.

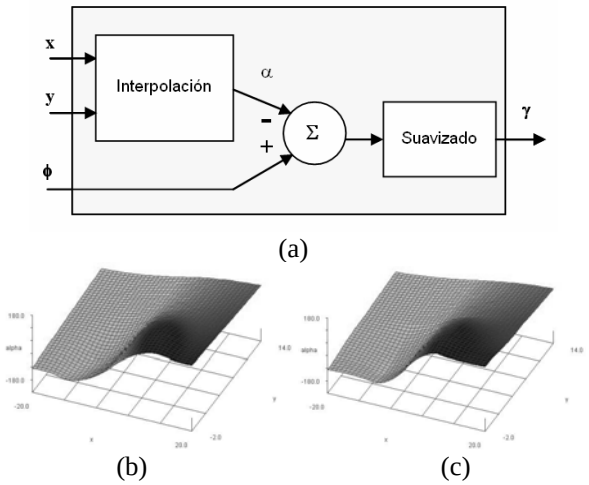


Figura 8: (a) Sistema jerárquico para la navegación sin obstáculos; (b) Valores de alfa según (6); (c) Salida del sistema neuro-fuzzy simplificado

3 RESULTADOS

Aparte del uso extensivo de las herramientas de identificación, simplificación y ajuste (*xfdm*, *xfsp* y *xfsl*) de Xfuzzy 3 para diseñar cada uno de los módulos, la totalidad del controlador neuro-difuso ha sido descrito mediante la herramienta *xfedit* de Xfuzzy 3, que permite conectar tanto bases de reglas difusas o no difusas como módulos aritméticos entre sí. En la figura 9 se muestra una captura de la ventana principal de dicha herramienta en la que se puede observar parte de dicho controlador.

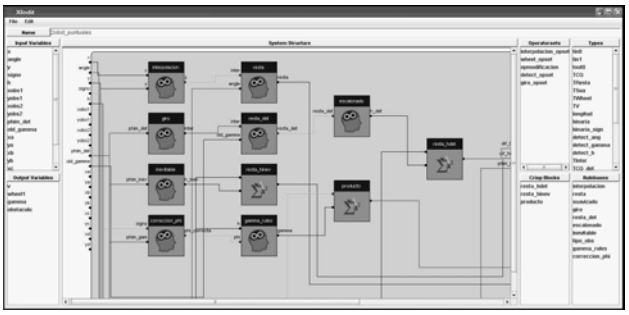


Figura 9: Descripción del sistema neuro-fuzzy mediante Xfuzzy 3

Se han realizado distintas simulaciones con la herramienta *xfsim*. Ésta simula el controlador trabajando en lazo cerrado con un modelo del robot. El modelo del robot (que incluye los modelos de sus sensores) genera una nueva configuración del robot y la nueva información dada por el láser. Este modelo se introduce en *xfsim* como una clase Java.

En la figura 10 podemos ver distintos ejemplos de cómo el robot es controlado para llegar a la configuración (x, y, ϕ, γ, v) = (0, 0, 180º, 0, 0) sin colisionar con los obstáculos. Como se puede apreciar, si no se detecta ningún obstáculo, el robot sigue hacia la configuración objetivo mediante una trayectoria de longitud cercana a la mínima, formada por arcos de circunferencia de radio mínimo junto a segmentos de línea recta. Y, en el caso de tener obstáculos cerca, éstos son esquivados con desviaciones mínimas de la trayectoria casi-óptima.

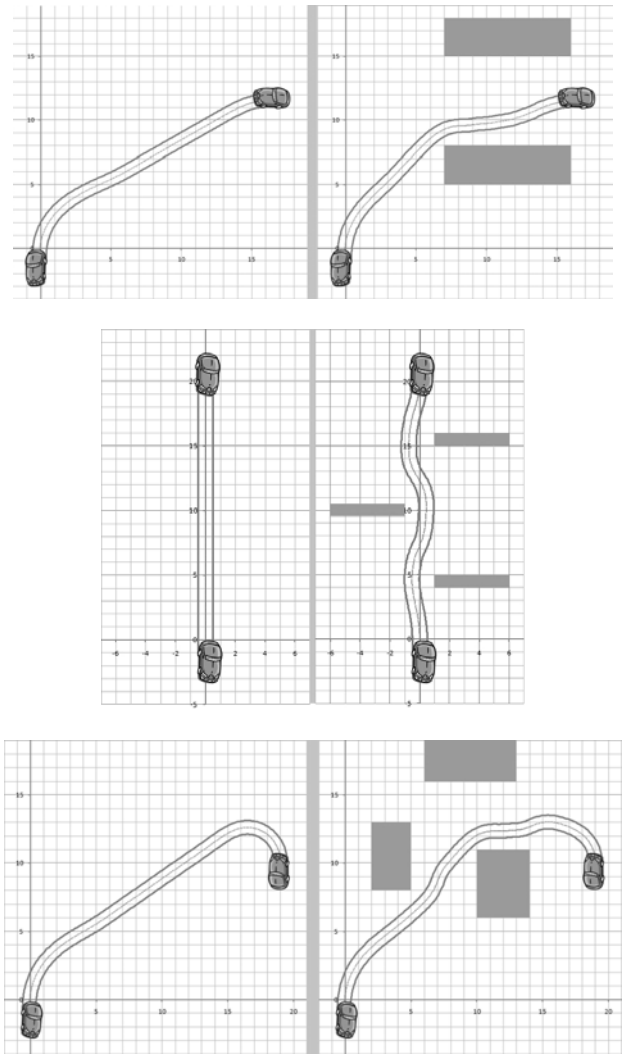


Figura 10: Simulación del sistema con Xfuzzy 3, realizando trayectorias con y sin obstáculos

4 CONCLUSIONES

A la hora de diseñar un sistema que controle la navegación de un robot móvil autónomo se consiguen muy buenos resultados cuando se combina el conocimiento heurístico con un análisis geométrico aproximado del problema. Esta solución representa un buen compromiso entre hacer un costoso análisis exhaustivo del problema o solo emplear conocimiento heurístico que luego requiera un tedioso ajuste por ensayo y error. Con ello, hemos conseguido un controlador que proporciona trayectorias de referencia casi óptimas que respetan las restricciones tanto cinemáticas como dinámicas de los robots tipo-coche. La automatización de todo el diseño (esto es, extracción de base de reglas a partir de datos numéricos y conocimiento heurístico, simplificación y ajuste de las mismas, además de la verificación del sistema al completo) ha sido posible gracias a las distintas herramientas de CAD del entorno Xfuzzy 3. Otra ventaja adicional del controlador obtenido es que sus reglas son lingüísticamente interpretables, lo cual facilita enormemente la comunicación entre un humano y el robot.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia y Tecnología mediante los proyectos TEC2005-04359 y DPI2005-002293, y por la Junta de Andalucía mediante los proyectos TIC2006-635 y TEP2006-375.

Referencias

[1] B. Kimiaghalam, A. Homaifar, B. Suttikulvet, B. Sayyarrodsari, "A multi-layered multi fuzzy inference systems for autonomous robot navigation and obstacle avoidance", *Proc. IEEE Int. Conf. on Fuzzy Systems*, pp. 340-343, Dec. 2001.

[2] I. Baturone, F. J. Moreno-Velo, A. A. Gersnoviez, "A CAD approach to simplify fuzzy system description", *Proc. IEEE Int. Conf. on Fuzzy Systems*, Vancouver, 2006.

[3] L.E. Dubins, "On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents", *American Journal Math.*, vol. 79, pp. 497-516, 1957.

[4] M. Sugeno, M. Nishida, "Fuzzy control of a model car", *Fuzzy Sets and Systems*, vol. 16, pp. 103-113, 1985.

[5] N. Zhang, D. Beetner, D. C. Wunsch II, B. Hemmelman, A. Hasan, "An embedded real-time neuro-fuzzy controller for mobile robot navigation", *Proc. IEEE Int. Conf. on Fuzzy Systems*, pp. 319-324, 2005.

[6] Página principal de Xfuzzy: http://www.imse.cnm.es/Xfuzzy.

[7] X. Wang, S. X. Yang, "A neuro-fuzzy approach to obstacle avoidance for a nonholonomic mobile robot", *Proc. IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, pp. 29-34, 2003.

[8] I. Baturone, F. J. Moreno-Velo, S. Sánchez-Solano, and A. Ollero, "Automatic design of fuzzy controllers for car-like autonomous robots". *IEEE Trans. on Fuzzy Systems*, Vol. 12, No. 4, Aug. 2004, pp. 447-465.

APRENDIZAJE DE REGLAS DIFUSAS PARA LA CLASIFICACIÓN DE COMPORTAMIENTOS EN UN SISTEMA DE VIDEOVIGILANCIA COGNITIVA

Javier Albusac[1] **J.J Castro-Schez**[1] **David Vallejo**[1] **Luis Jiménez-Linares**[1]

[1]Paseo de la Universidad n^o 4. Escuela Superior de Informática, Ciudad Real.
Universidad de Castilla-La Mancha.
{JavierAlonso.Albusac@uclm.es, JoseJesus.Castro, David.Vallejo, Luis.Jimenez}@uclm.es

Resumen

En los últimos años, el aumento generalizado de las medidas de seguridad, junto con la evolución tecnológica y el abaratamiento del hardware, han dado lugar a un incremento de interés por parte de la comunidad científica en el campo de la vigilancia inteligente. En este trabajo se describe cómo ha sido diseñado un componente de análisis de normalidad, basado en lógica difusa, para el estudio de relaciones espaciales entre los objetos móviles y zonas de una escena captadas por una cámara de vigilancia. Este componente será integrado dentro de un sistema de vigilancia cognitivo basado en técnicas de softcomputing, con una arquitectura dividida en varios niveles.

Palabras Clave: Vigilancia Inteligente, Lógica Difusa, Aprendizaje Inductivo.

1 INTRODUCCIÓN

Gran parte de los sistemas de seguridad implantados actualmente son una evolución de los sistemas de primera generación [3]. Dichos sistemas están formados por un conjunto de cámaras que envían la señal a una sede central donde un vigilante puede observar varios monitores. Según Tan Kok Kheng, vicepresidente de la división OEM de WPG Systems, - una de las principales distribuidoras de sistemas de vigilancia avanzados -, "Tras 20 minutos de vigilancia, la atención humana a los detalles del vídeo disminuye hasta niveles inaceptables y la videovigilancia deja de tener sentido. La videovigilancia tradicional ya no puede cumplir las, cada vez mayores, demandas del sector". La solución a estas deficiencias pasa por la utilización de sistemas de vigilancia inteligentes [3] capaces de interpretar lo que está ocurriendo. "La vídeovigilancia ya no puede ser simple y reactiva, necesita ser inteligente y proactiva", enfatiza el Sr. Kheng.

Entre las grandes dificultades a las que deben enfrentarse los Sistemas de Vigilancia Inteligentes (SVI) cabe destacar dos de ellas. En primer lugar, estos tipos de sistemas suelen observar subdominios del mundo real no deterministas, conviviendo en todo momento con la incertidumbre y la vaguedad, sin poder afirmar con total certeza qué clase de objetos están participando y qué eventos se están produciendo. Dependiendo de si se lleva a cabo una vigilancia forense o predictiva [5], existirá un mayor grado de incertidumbre o vaguedad. Cuando se realiza una vigilancia predictiva, el objetivo principal es anticiparse a posibles situaciones de riesgo antes de que ocurran, por lo que existe un alto grado de incertidumbre. En cambio, si la vigilancia es forense, es decir, se analizan los hechos una vez que han ocurrido, puede que la incertidumbre sea mínima y la vaguedad sea mayor.

Por otra parte, para que un SVI pueda clasificar un objeto a partir de sus propiedades o pueda determinar si las conductas y los eventos son normales, es necesario definir y formalizar el conocimiento del dominio. El segundo problema nace en la dificultad para reutilizar el conocimiento, debido en gran parte a la fuerte dependencia con el entorno observado. Por ejemplo, cada lugar observado por una cámara tiene sus zonas con sus características y normas propias. Un objeto observado se comportará de forma normal siempre y cuando cumpla dichas normas. Esta dificultad repercute directamente en el proceso de implantación o ampliación del sistema, ya que la ubicación de nuevos sensores implicaría la ampliación de la base de conocimiento por parte del experto. En este caso, los algoritmos de aprendizaje semi-automáticos y automáticos juegan un papel fundamental para agilizar este proceso y eliminar la fuerte dependencia con el experto [6, 2].

En el presente artículo se presentará un componente, basado en lógica difusa, integrado en un SVI en el que se está trabajando actualmente. Este componente estudia principalmente las relaciones espaciales que existen entre los objetos móviles y las zonas observadas, y clasifica los eventos simples a partir de un conjunto de reglas difusas generadas por un algoritmo de aprendizaje inductivo.

2 ARQUITECTURA DEL SVI

En la Figura 1 se puede observar la arquitectura multicapa del SVI con tres niveles bien diferenciados. En el primer nivel o nivel cero, se encuentran los sensores encargados de capturar el entorno y los algoritmos de visión que realizan la segmentación [7] y el tracking [8] de los objetos. En el segundo nivel se analizan los eventos que se producen en una escena y se determina si estos son normales o, por el contrario, implican algún peligro para el entorno.

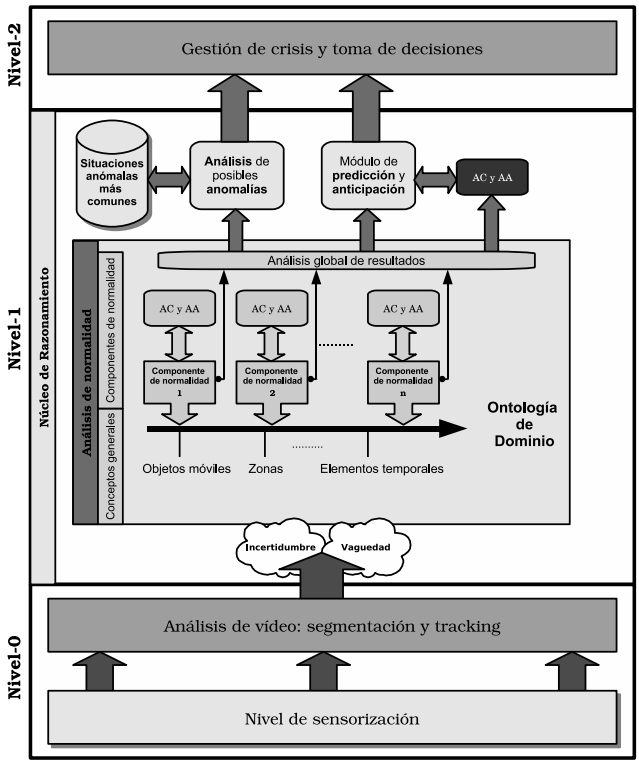


Figura 1: Arquitectura de un sistema de videovigilancia inteligente.

La mayoría de autores, para detectar situaciones anómalas, optan por definir las anomalías más comunes y tratan de encontrar similitudes entre éstas y las situaciones que se producen. El gran inconveniente de esta aproximación es que la mayoría de anomalías son imprevisibles incluso para un experto en el dominio. Por tanto, cuando se produce alguna de las situaciones anómalas no definidas, el sistema es incapaz de responder ante ellas. En cambio, la normalidad de un dominio es perfectamente conocida y, por este motivo, se ha optado por definir cuáles son las situaciones normales en un escenario, de tal forma que todo aquello que no sea normal debería ser tratado con especial atención. Además, para afrontar los problemas con mayor garantía de éxito, se definen las situaciones anormales más comunes para que sean identificadas. Por tanto, la peor situación se daría cuando el sistema fuera incapaz de recocer qué tipo de anomalía se está produciendo pero, al menos, sabría que está pasando algo que no es normal.

La formalización del conocimiento de dominio en el nivel intermedio se lleva a cabo mediante ontologías. En nuestro caso, la ontología de dominio está constituida, por una parte, por conceptos generales válidos para cualquier escena, como por ejemplo las clases de objetos móviles (personas, vehículos, etc) y sus propiedades, definición de zonas, elementos temporales, etc. Y por otra parte, los componentes de normalidad con módulos de adquisición y aprendizaje (AC y AA), que analizan la normalidad desde diferentes puntos de vista. Cada uno de estos componentes es independiente del resto y se pueden conectar o desconectar del sistema en cualquier momento. Así, el sistema puede por ejemplo disponer de un módulo encargado de analizar las trayectorias de los objetos móviles, otro componente para determinar si es apropiado que un tipo de objeto invada una determinada zona, otro componente para analizar las velocidades, etc. Cada componente ofrece respuestas en instantes de tiempo diferentes, dependiendo del peso del tipo de razonamiento. Uno de los requisitos fundamentales de cualquier sistema de seguridad es que ofrezca respuestas cercanas al tiempo real. Por tanto, son necesarios los componentes de normalidad ligeros, ya que a pesar de que no realizan un análisis exhaustivo de la situación actual, ofrecen una respuesta rápida.

Finalmente, el análisis de normalidad, la similitud con posibles situaciones anómalas y la predicción de acciones futuras, son los elementos que maneja el tercer nivel para tomar decisiones y gestionar una crisis en el caso de que ésta se produzca.

3 COMPONENTE DE NORMALIDAD DIFUSO

El objetivo principal del componente es la generación de reglas difusas que permitan clasificar los eventos de una escena, a partir de las relaciones espaciales que existen entre los objetos y las zonas. Se trata de un componente de normalidad ligero, ya que ofrece una

respuesta en un espacio breve de tiempo.

El motor de inferencia, a partir de las reglas obtenidas, tiene capacidad para determinar qué clases de objetos pueden invadir ciertas zonas y en qué grado (proporción de la superficie del objeto que invade la zona). Por ejemplo, una posible regla sería que un coche no puede invadir una zona ajardinada pero, si éste la invade parcialmente, el sistema podría decidir no activar la alarma ya que no sería un motivo suficiente. Por tanto, tratar las relaciones espaciales entre objetos y zonas de forma difusa, garantiza mayor flexibilidad al sistema y evita un número elevado de activaciones innecesarias de alarmas.

3.1 APRENDIZAJE INDUCTIVO DE REGLAS DIFUSAS

Sea ε un conjunto de ejemplos (eventos simples de una escena que ocurren en instantes de tiempo concretos) ε=$\{e_1, e_2, ..., e_n\}$, donde cada uno de ellos tiene la estructura $e_i = ((x_{i1}, x_{i2}, ..., x_{in}), o_i)$; siendo $x_{i1}, x_{i2}, ..., x_{in}$ los valores de entrada de las variables pertenecientes a V y o_i es la clase a la que pertenece el ejemplo.

El algoritmo de aprendizaje automático empleado por el componente utiliza como entrada el conjunto de ejemplos ε, las variables que describen cada ejemplo V, los dominios de definición de dichas variables DDV, y el conjunto de posibles salida [4]. Dicho algoritmo genera un conjunto de reglas difusas, que pueden ser fácilmente interpretadas, y cuya forma es la siguiente:

SI ν_1 es ZD_1 y ν_2 es ZD_2 y ... y ν_n es ZD_n **ENTONCES** clase = O_x,

donde cada ZD_i es un conjunto de valores disyuntivamente asociados con la variable ν_i, que toma valores de su dominio de definición DDV_i, verificando que $ZD_i \subseteq DDV_i$. Otra forma de escribir la regla es $((ZD_{1i}, ZD_{2i}, ..., ZD_{ni}),O_x)$.

El algoritmo se compone de dos pasos principales. El primero consiste en *"convertir los ejemplos de entrada en reglas particulares"*. Es decir, cada uno de los ejemplos del conjunto de entrenamiento se transforma en una regla, formada por la clase a la que pertenece el ejemplo y una etiqueta lingüística para cada variable (proceso de fusificación). El segundo paso consiste en *"Construir el conjunto de reglas definitivas maximales"* (utilizadas por el sistema para clasificar las situaciones) mediante un proceso de amplificación.

El principal problema en el proceso de clasificación es el de ambigüedad, que ocurre cuando un ejemplo de entrada cumple varias reglas pertenecientes a diferentes clases. En este caso, para resolver la ambigüedad, se elige aquella regla con mayor *grado de conveniencia*.

$$grado_de_conveniencia = min\{\varphi_i(x_j)\} \qquad (1)$$

Donde φ_i son las funciones de pertenencia asociadas a cada variable de entrada X_j en la regla R_i. La funciones φ_i se construyen en función de las etiquetas presentes en la regla para cada X_j, por lo que éstas varían para cada una de las reglas.

3.2 VARIABLES DE DOMINIO

El componente de normalidad maneja un conjunto formado por cinco variables $V = \{v_1, v_2, v_3, v_4, v_5\}$, cada una de ellas con un dominio de definición $DDV = \{DDV_1, DDV_2, DDV_3, DDV_4, DDV_5\}$. Cada DDV_i de DDV se define mediante un conjunto $DDV_i = \{L_{1i}, L_{2i}, ..., L_{mi}\}$ que contiene todos los posibles valores que la variable v_i puede tomar (siendo L_{ji} la etiqueta lingüística del valor j en la variable v_i, definida mediante una función trapezoidal con los parámetros $(a_{ji}, b_{ji}, c_{ji}, d_{ji})$).

La variable v_1 representa la *clase* a la que pertenece un objeto observado (ver Fig. 2). El proceso de clasificación de un objeto se lleva a cabo en las capas inferiores del SVI durante el proceso de segmentación, liberando al componente de esta tarea. Cuando un objeto es clasificado, se le asigna un identificador que lo diferencia del resto y un valor numérico que representa la clase a la que pertenece. Finalmente, el componente de normalidad difuso sustituye el valor numérico por una etiqueta lingüística, necesaria para el algoritmo de aprendizaje.

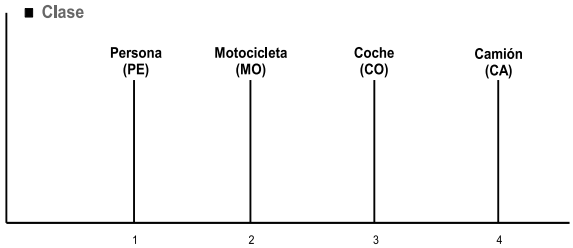


Figura 2: Variable v_1, clase a la que pertenece un objeto móvil. DDV_1={PE, MO, CO, CA}

La variable v_2 referencia el *tamaño* del objeto. El componente recibe de los niveles inferiores información acerca de la elipse que envuelve al objeto (el punto central y los dos radios cuya longitud se mide en píxeles). El valor lingüístico que representa el tamaño de un objeto se calcula a partir del área de la elipse que lo envuelve

$$\max_i\{\mu_i(\pi \times r_1 \times r_2)\} \mid i \in DDV_2$$

.

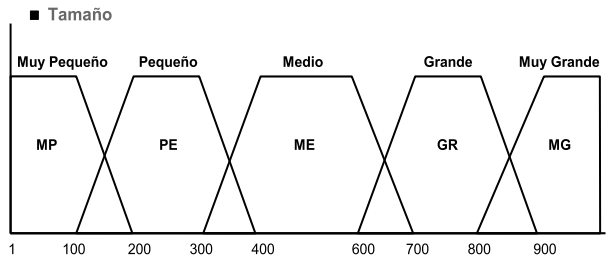


Figura 3: Variable v_2, tamaño de un objeto. DDV_2={MP, PE, ME, GR, MG}

Por otra parte, la variable v_3 representa la *intersección* entre un objeto y una zona (porcentaje de la superficie de un objeto que invade una zona). Para calcular el valor de dicha variable (Ec. 3), se recorre píxel a píxel el rectángulo mínimo que envuelve a la elipse de un objeto en la imagen capturada por la cámara (ver Fig. 3.2). En dicho recorrido se determina el número de puntos que pertenecen a la elipse (Ec. 2) y a la zona [1] en el rectángulo mínimo.

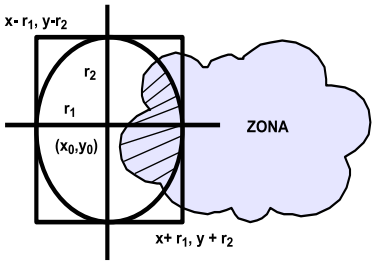


Figura 4: Rectángulo mínimo que envuelve a la elipse de un objeto e intersección con una zona

$$\frac{(x-x_0)^2}{r_1^2}+\frac{(y-y_0)^2}{r_2^2} \qquad (2)$$

Siendo (x_0,y_0) el centro de la elipse del objeto analizado y r_1, r_2, los radios de la elipse que representan el ancho y la altura respectivamente. La pertenencia del punto (x, y) a la elipse dependerá del valor obtenido en 2:

- < 1, el punto (x, y) pertenece a la elipse.
- $= 1$, el punto (x, y) pertenece al perímetro de la elipse.
- > 1, el punto (x, y) no pertenece a la elipse.

$$intersec(obj_x, z_y) = area(x \mid x \in obj_x \wedge x \in z_y) \qquad (3)$$

El último paso consiste en dar una interpretación lingüística de la intersección entre la elipse y la zona mediante el uso de DDV_3 (ver Fig. 5)

$$\max_i\{\mu_i(\frac{intersec(obj_x, zona_y)}{area(obj_x)})\} \mid i \in DDV_3$$

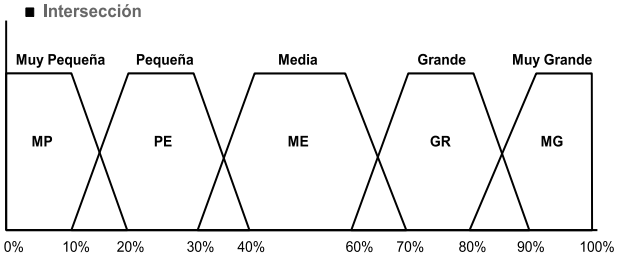


Figura 5: Variable v_3, tamaño de la intersección entre la elipse de un objeto y una zona. DDV_3={MP, PE, ME, GR, MG}

La variable v_4 representa la *velocidad* de un objeto móvil en una escena y depende del desplazamiento en píxeles desde un frame al siguiente. Por tanto, para determinar la velocidad de un objeto entre dos frames f_1 y f_2, se calcula la distancia (Ec. 4) entre los puntos (x_1, y_1) y (x_2, y_2) que corresponden con el centro de la elipse que envuelve al objeto en f_1 y f_2 respectivamente.

$$distancia(p_1, p_2) = \sqrt{(x_1-x_2)^2+(y_1-y_2)^2} \qquad (4)$$

Finalmente, se lleva a cabo una representación lingüística de los valores obtenidos a partir de la Ec. 4 con la utilización de DDV_4 (ver Fig. 6).

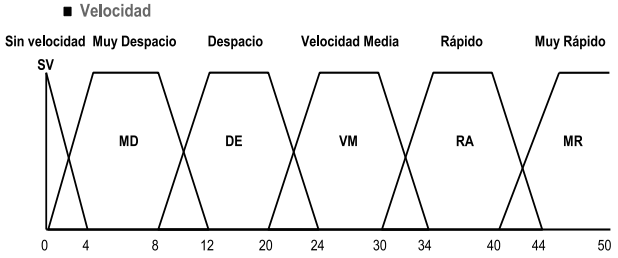


Figura 6: Variable v_4, velocidad de un objeto. DDV_4={SM, MD, DE, VM, RA, MR}

La última variable, v_5, representa la *zona* con la que intersecta el objeto. Un objeto puede invadir varias

zonas al mismo tiempo, por lo que la intersección con distintas zonas puede ser no nula simultáneamente. Cuando se da esta situación, el sistema trata cada una de las intersecciones de forma individual. De esta forma, si un objeto invade dos zonas al mismo tiempo se podría dar el caso en el que la relación espacial entre el objeto y una de las zonas fuera normal, mientras que con la otra podrían existir ciertas anomalías. Por otra parte, las zonas definidas en DDV_5 son totalmente dependientes de la escena observada, y en el caso de la Fig. 7 las zonas se corresponden con las mostradas en la Fig. 8 (ver Sección 4).

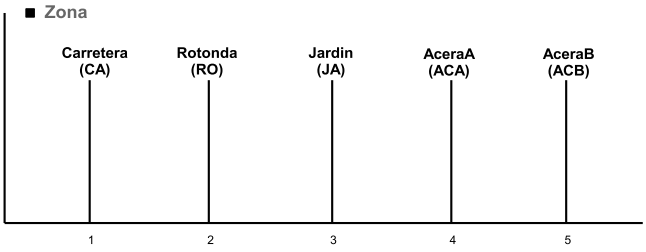


Figura 7: Variable v_5, zonas observadas en la escena. DDV_5={CA, RO, JA, ACA, ACB}

4 RESULTADOS EXPERIMENTALES

Para probar el componente de normalidad difuso se ha utilizado el vídeo capturado por una cámara IP situada en la ventana del Grupo de Investigación ORETO de la Universidad de Castilla-La Macha. La definición de los DDV_i para este ejemplo son los que se han mostrado en la seccción anterior. En la Figura 8, se muestran las zonas definidas para la escena observada. Dos aceras (Acera A y Acera B) por la que sólo pueden circular personas; la carretera por la que pueden circular tanto personas como vehículos; una zona ajardinada que separa dos carriles con sentidos opuestos (ningún objeto debería invadir esta zona); y finalmente, una rotonda que tampoco debería ser invadida por ningún objeto.

Observando la Figura 8, se puede intuir que la *perspectiva* puede ser una de las mayores fuentes de ruido, y por tanto, uno de los factores que pueden influir en mayor medida en los resultados. En la Figura 9 se pude apreciar como la elipse de un objeto, situado completamente en la carretera, invade parte del jardín.

La lógica difusa nos permite tratar con éxito este problema. El componente difuso y en concreto el algoritmo de aprendizaje, podría generar un conjunto de reglas que indiquen que los objetos con tamaño medio o grande, invadiendo el jardín muy poco o poco, es

Figura 8: Definición de zonas en la escena vigilada.

Figura 9: Problema de la perspectiva.

una situación normal. De esta forma, se elimina el ruido generado debido a los efectos ópticos producidos por la perspectiva. Por tanto, las reglas no tienen en cuenta la posición real y absoluta de un objeto en una escena, sino que analiza la posición de un objeto desde el punto de vista de la cámara.

A modo de ejemplo, en la Tabla 1 se muestran algunas de las reglas generadas a partir de un conjunto de ejemplos formado por 205 situaciones, de los cuales el 80% han sido empleados en el proceso de aprendizaje y el 20% para el proceso de clasificación. Ante la dificultad de obtener escenas de vídeo donde suceden situaciones anómalas, se ha optado por realizar simulaciones en las que se han ubicado objetos móviles virtuales en la escena.

En la Tabla 2 se muestran los resultados del proceso de aprendizaje y clasificación en diez tests diferentes. La mayoría de errores en el proceso de clasificación provienen de la ambigüedad que existe entre las reglas. En el contexto de los sistemas de seguridad, la peor situación ocurre cuando un evento anómalo es clasificado como normal. Por otra parte, si el algoritmo genera un número elevado de reglas, aumentan

Tabla 1: Reglas generadas de forma automática

Clase O_i	**Regla** R_i
O_1: Normal	R_0: Si v_3 no es {GR,MG} y v_5 no es {JA,ACA,ACB} R_1: Si v_3 no es {GR,MG} y v_4 no es {MR} y v_5 no es {ACB} R_2: Si v_2 no es {ME} y v_3 no es {ME,MG} y v_5 es {ACB,CA} R_3: Si v_2 no es {GR} y v_3 no es {ME,GR} y v_5 es {ACB}
O_2: Vehículo en acera	R_4: Si v_1 no es {PE} y v_2 no es {MP,PE} y v_4 es {GR,MG} y v_5 es {ACA,ACB}
O_3: Invasión del jardín	R_5: Si $v3$ no es {MP,PE} y v_5 es {JA}
O_4: Invasión de la rotonda	R_6: Si v_3 no es {MP,PE} y v_4 no es {MD} y v_5 es {RO} R_7: Si v_3 es {GR,MG} y v_5 es {RO}
O_5:Persona detenida en la carretera	Si v_1 es {PE} y v_2 es {MP,PE} y v_3 no es {MP,PE} y v_4 es {SV} y v_5 es {CA}

los conflictos entre ellas y, en consecuencia, la probabilidad de que el problema de la ambigüedad se acentúe. Por tanto, para reducir la ambigüedad en la mayor medida de lo posible, el experto debe configurar un conjunto de entrenamiento lo más representativo posible para así reducir el número de reglas y conflictos.

Tabla 2: Resultados de los tests

Test	**Reglas**	**Aciertos**	**Fallos**	**%**
1	17	39	1	97.5%
2	13	40	0	100%
3	15	37	3	92.5%
4	15	37	3	92.5%
5	13	38	2	95%
6	14	37	3	92.5%
7	15	38	2	95%
8	15	39	1	97.5%
9	16	36	4	90%
10	14	37	3	92.5%

5 CONCLUSIONES

En el presente artículo se ha presentado un componente difuso, integrado en un Sistema de Vigilancia Video Cognitivo, cuya función principal es clasificar los eventos que se producen en una escena observada por una cámara, mediante el análisis de las relaciones espaciales entre los objetos y las zonas del entorno.

Como línea de trabajo futuro, se pretende desarrollar otros componentes integrables que permitan analizar la normalidad desde otros puntos de vista. Además, se establecerá un mecanismo para relacionar la información procedente de diversos sensores heterogéneos ubicados en la misma escena, con el objetivo de reforzar la creencia y el porcentaje de éxito en las interpretaciones realizadas.

Agradecimientos

Este trabajo ha sido financiado gracias a los Proyectos PAC06-0141 y PBC06-0064 de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1] Shimrat, M. Algorithm 112, Position of Point Relative to Polygon. Comm. ACM 5(8), Pág.424, 1962.

[2] Makris, D. and Ellis, T. Automatic Learning of an activity-based semantic scene model. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03), Pág.183-188, 2003.

[3] Valera, M. and Velastin, SA. Intelligent distributed surveillance systems: a review. IEEE Proceedings Vision, Image and Signal Processing, 152(2), Pág.192-204, 2005.

[4] Castro, JL., Castro-Schez, JJ and Zurita, J. Learning maximal structure rules in fuzzy logic for knowledge acquisition in expert systems. Fuzzy Sets and Systems, Elsevier, 101(3), Pág.331-342, 1999.

[5] Martínez-Tomás R., Ricón, M.,Bachiller, M. and Mira, J. On the correspondence between objects and events for the diagnosis of situations in visual surveillance task. Elsevier, Pattern Recognition Letters. 2007.

[6] Jiangungm, L., Qifeng, T., Tieniu, T. and Weiming, H. Semantic interpretation of object activities in a surveillance system. Proceedings of the 16 th International Conference on Pattern Recognition, ICPR'02. Pág.777-780, 2002.

[7] Rodríguez-Benítez, L., Moreno-García, J., Albusac, J., Castro-Schez, J.J. and Jiménez-Linares, L. An approximate reasoning technique for segmentation on compressed mpeg video. International Conference on Computer Vision Theory and Applications, VISAPP'07. Pág.184-191, 2007.

[8] Rodríguez-Benítez, L., Moreno-García, J., Albusac, J., Castro-Schez, J.J. and Jiménez-Linares, L. Seguimiento de objetos representados lingsticamente utilizando tcnicas de razonamiento aproximado. II Simposio sobre Lgica Fuzzy y Soft Computing (LFSC'07), II Congreso Espaol de Informtica, (CEDI'07). Pág.237-244, 2007.

Segmentación de imágenes de ultrasonidos de próstata

Humberto Bustince[1] **Miguel Pagola**[1] **Edurne Barrenechea**[1] **Hamid Tizhoosh**[2]

[1] Dpto. Automática y Computación, Universidad Pública de Navarra, Pamplona, España
{bustince, miguel.pagola, edurne.barrenechea}@unavarra.es
[2] Pattern Analysis and Machine Intelligence Lab, University of Waterloo, Waterloo, Canadá
tizhoosh@uwaterloo.ca

Resumen

En este trabajo desarrollamos un algoritmo basado en conjuntos intervalo valorados difusos para la segmentación de imágenes de ultrasonidos. Presentamos una generalización del algoritmo difuso de umbralización clásico, en donde utilizamos dos funciones de pertenencia para representar a la imagen. Definimos unas funciones que llamaremos funciones de desconocimento, que nos permiten representar la incertidumbre que tiene el experto en la elección de las funciones de pertenencia. Construimos los conjuntos intervalo valorados difusos mediante estas funciones de desconocimiento.
Palabras Clave: Imágenes de Ultrasonidos; Conjuntos intervalo valorados difusos; Umbralización

1 Introducción

El cáncer de próstata es uno de los más frecuentes entre los hombres mayores de 50 años. Afecta a un hombre de cada 8, con un ratio de mortalidad de 1/28 [3]. Las imágenes de ultrasonidos se utilizan en el diagnóstico, tratamiento y seguimiento de la enfermedad. Esta técnica se utiliza debido a que no es invasiva, es barata y los resultados se obtienen instantaneamente.

En muchos casos, la imagen de la próstata es segmentada manualmente por el experto. En los últimos años se han publicado un considerable número de artículos presentando métodos semiautomáticos de segmentación de próstata en imágenes de ultrasonidos [6]. Debido a las características de la imagen de ultrasonidos, los métodos convencionales para detectar bordes o segmentar imágenes no han dado buenos resultados. Todos los trabajos están referidos a un tipo de imágenes llamadas *imágenes de ultrasonidos transrectales* (IUT).

En este trabajo proponemos un nuevo método de segmentación de imágenes de ultrasonidos utilizando teécnicas difusas. En un primer paso realizamos un realce de la imagen. En el segundo paso utilizamos un algoritmo de umbralización para separar la zona que es próstata del resto de la imagen.

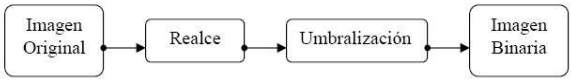


Figura 1: Fases de la segmentación de IUT.

El algoritmo que proponemos, utiliza conjuntos intervalo valorados difusos (IVFSs) para representar la incertidumbre en el proceso de umbralización de la imagen. De forma natural, el nuevo método nos ha llevado también a presentar una *generalización del algoritmo difuso*.

En este algoritmo difuso generalizado, las imágenes las vamos a representar mediante dos conjuntos difusos (uno para la zona de próstata y otro para el fondo). Por este motivo, en el algoritmo que proponemos introducimos el concepto de *funciones de desconocimiento* G_u. Dichas funciones, son una forma de representar el desconocimiento que tiene el experto en la elección de las dos funciones de pertenencia que van a representar a la imagen.

Los algoritmos de segmentación podemos dividirlos en tres categorías diferentes desde el punto de vista de las posibilidades del usuario en la optimización del algoritmo a su problema específico:

1. Algoritmos generalistas, de forma que el usuario ejecuta el algoritmo y obtiene la solución. Si el algoritmo realiza una segmentación adecuada para el tipo de imágenes con las que está trabajando,

usará ese algoritmo, sin embargo si la solución aportada por el algoritmo no es satisfactoria, el usuario buscará otro algoritmo (No supervisado).

2. Algoritmos paramétricos, el usuario puede ir modificando una serie de parámetros de funcionamiento del algoritmo hasta que obtiene una solución adecuada para sus imágenes. Normalmente el ajuste adecuado de estos parámetros se realiza cuando el usuario ya tiene experiencia (Semisupervisado).

3. Algoritmos que incorporan un proceso de optimización. Partiendo de un conjunto de imágenes ideales, el algoritmo, tras un proceso de aprendizaje segmenta de forma "óptima" ese tipo de imágenes. Sin embargo en otro tipo de imágenes el algoritmo seguramente obtendrá muy malos resultados (Supervisado).

La calidad de las imágenes de ultrasonidos es muy variable y la cantidad de ruido, iluminación dependen de muchos parámetros, temperatura, caracterísitcas fisiologicas del paciente, etc. Por tanto, los algoritmos supervisados y no supervisados tienen problemas en su uso ya que las imágenes obtenidas pueden tener diferentes carcaterísticas.

En este sentido en nuestra propuesta és el médico el que debe escoger las funciones de pertenencia cada vez que quiere realizar una segmentación. Por tanto, es un algoritmo de tipo semisupervisado. En nuestra propuesta, el usuario debe escoger unas funciones de pertenencia que el cree mejor representan a los objetos de la imagen. Este es el principal problema de nuestro algoritmo, es evidente que la eficacia de nuestro algoritmo depende de las funciones escogidas.

Evidentemente el proceso de elección de las funciones de pertenencia, (una vez que el medico ya tiene experiencia), es más rápido que si el médico segmenta a mano la imagen. La segmentación manual suele costar entre 10 y 15 minutos, mientras que escoger las funciones de pertenencia y lanzar el algoritmo, puede realizarse en menos de 1 minuto.

El trabajo está organizado de la siguiente forma: primero mostramos la técnica utilizada para realzar las imágenes de ultrasonidos. En la siguiente sección mostramos el algoritmo generalizado difuso de umbralización, las funciones de desconocimiento y el algoritmo IVFS que proponemos. En la sección de resultados experimentales presentamos las condiciones del experimento que hemos realizado y comparamos los resultados con otros métodos.

2 Realce de la imagen

En el primer paso de nuestro método para segmentar IUTs, utilizamos el algoritmo propuesto por Sabha et. al. [8] para realzar las imágenes de ultrasonidos. La idea principal es eliminar el ruido de las imágenes y realzar el nivel de gris de una zona seleccionada. Primero eliminamos el ruido utilizando el filtro de la mediana (7×7 ó 9×9). Debido a que en las imágenes suelen aparecer varias zonas para analizar, el usuario debe seleccionar el punto central de la zona que quiere segmentar. Después cada pixel es fusificado según su nivel de gris utilizando una función de pertenencia que tiene en cuenta el valor medio de gris alrededor del punto central, (para mas detalles [10]).

3 Umbralización de la imagen realzada

3.1 Generalización del algoritmo difuso de umbralización

En el algoritmo clásico difuso de umbralización (ver [5, 4]) el usuario escoge una única función de pertenencia para representar la imagen:

(a) Asignar L conjuntos difusos $\tilde{Q}_t$ a la imagen Q. Cada uno está asociado con un nivel de intensidad t, $(t = 0, 1, \cdots, L-1)$, siendo L la escala de grises utilizada.

(b) Calcular la entropía de cada uno de los L conjuntos difusos $\tilde{Q}_t$ asociados con Q.

(c) Tomar como *mejor umbral* t el nivel de gris asociado con el conjunto difuso de menor entropía.

Algoritmo difuso clásico

Como estudiamos en [1], una forma general de construir funciones de pertenencia en el algoritmo clásico difuso de umbralización es la siguiente:

$$\mu_{\tilde{Q}_t}(q) = \begin{cases} F(REF(q, m_B(t))) \; si \; q \leq t \\ F(REF(q, m_O(t))) \; si \; q > t \end{cases} \quad (1)$$

Donde q es el nivel de intensidad del pixel, $h(q)$ el número de píxeles en la imagen con intensidad q y $m_B(t)$, $m_O(t)$ son:

$$m_B(t) = \frac{\sum_{q=0}^{t} qh(q)}{\sum_{q=0}^{t} h(q)} \quad (2)$$

$$m_O(t) = \frac{\sum_{q=t+1}^{L-1} qh(q)}{\sum_{q=t+1}^{L-1} h(q)}. \quad (3)$$

Siendo REF una función de equivalencia restringida [1]:

Definición 1 *(see [2]) Una función* $REF : [0,1]^2 \rightarrow [0,1]$ *es una función de equivalencia restringida, si satisface las siguientes condiciones:*
1) $REF(x,y) = REF(y,x)$ *para todo* $x, y \in [0,1]$*;*
2) $REF(x,y) = 1$ *si y sólo si* $x = y$*;*
3) $REF(x,y) = 0$ *si y sólo si if* $x = 1$ *e* $y = 0$ *o* $x = 0$ *e* $y = 1$*;*
4) $REF(x,y) = REF(c(x),c(y))$ *para todo* $x, y \in [0,1]$*,* c *siendo* c *una negación fuerte;*
5) *Para todo* $x, y, z \in [0,1]$*, si* $x \leq y \leq z$*, entonces* $REF(x,y) \quad REF(x,z)$ *y* $REF(y,z) \quad REF(x,z)$.

A continuación presentamos unos ejemplos de funciones de equivalencia restringida.

Example 1

$$REF_1(x,y) = 1 - |x - y| \tag{4}$$

$$REF_2(x,y) = (1 - |x - y|)^2 \tag{5}$$

$$REF_3(x,y) = 1 - |x - y|^2 \tag{6}$$

Y $F : [0,1] \rightarrow [0.5,1]$ una función tal que

$$\begin{cases} F(x) = 1 \text{ si y sólo si } x = 1 \\ F(x) = 0.5 \text{ si y sólo si } x = 0 \\ \text{creciente} \end{cases} \tag{7}$$

En nuestra propuesta, el usuario debe escoger dos funciones: una para representar la zona de la próstata y otra para representar el fondo. De esta forma, el experto es capaz de utilizar una mejor representación para los píxeles que no esta seguro de su pertenencia al objeto (próstata) o al fondo:

$$\mu_{\tilde{Q}_{Bt}}(q) = F_1(REF_1(q, m_B(t))) \tag{8}$$

$$\mu_{\tilde{Q}_{Ot}}(q) = F_2(REF_2(q, m_O(t))) \tag{9}$$

En la figura 2 mostramos una imagen IUT original, su imagen realzada, el histograma de la imagen realzada y un par de funciones de pertenencia que representan $\mu_{\tilde{Q}_{Bt}}$, $\mu_{\tilde{Q}_{Ot}}$ de la imagen.

En futuros trabajos queremos desarrollar una metodología para escoger estas funciones, teniendo en cuenta parámetros como la intensidad media, porcentaje de ruido, etc. Además queremos asociar las funciones de pertenencia $\mu_{\tilde{Q}_{Bt}}$ y $\mu_{\tilde{Q}_{Ot}}$ a etiquetas lingüisticas como"condensado" o "disperso", de forma que la elección de la función de pertenencia sea un proceso sencillo y que no esté basado únicamente en la experiencia del usuario.

3.2 Funciones de desconocimiento y píxeles dudosos

En esta sección introducimos el concepto de función de desconocimiento. Estas funciones representan el

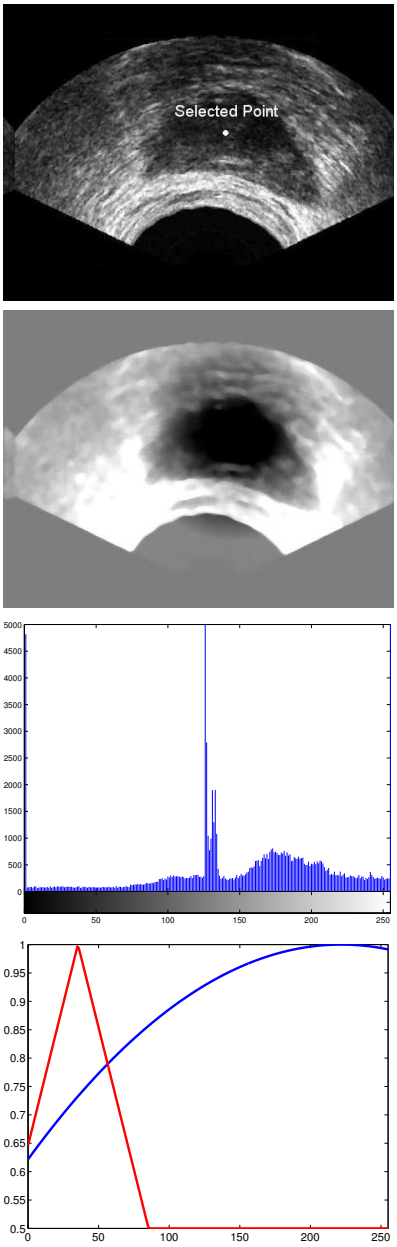


Figura 2: De arriba abajo: Imagen IUT original, imagen realzada, histograma de la imagen realzada, funciones de pertenencia que representan a la imagen.

desconocimiento que tiene el usuario al escoger las funciones de pertenencia usadas para representar la imagen. Por tanto, en este nuevo algoritmo vamos a asociar a cada pixel tres valores numéricos:

- Un valor para representar la pertenencia a la zona de próstata. Este valor lo interpretamos como el conocimiento que tiene el experto de la pertenencia de ese pixel a la zona de la próstata.
- Un valor para representar la pertenencia al fondo. Este valor lo interpretamos como el conocimiento

que tiene el experto de la pertenencia de ese pixel a la zona del fondo.

- Un valor para representar el desconocimiento del experto sobre la pertenencia de ese pixel al fondo o a la próstata. Este desconocimiento impide al experto hacer una construcción exacta de las funciones de pertenencia de la zona de la próstata y la zona del fondo. Cuanto menor sea el valor del desconocimiento, mejor habrán sido escogidas las funciones de pertenencia. Evidentemente, existen píxeles en una imagen para los cuales el experto no sabe exactamente el grado de pertenencia al fondo o a la zona de la próstata, es precisamente este desconocimiento el que intentamos modelizar.

A una función de desconocimiento le vamos a exigir que cumpla las siguientes propiedades:

0.- La función de desconocimiento debe depender de $\mu_{\tilde{Q}_{Bt}}(q)$ y $\mu_{\tilde{Q}_{Ot}}(q)$.

1.- La función de desconocimiento no depende del orden en que tomemos las pertenencias al fondo y al objeto.

2.- Representación del conocimiento total: el desconocimiento del experto debe ser cero si y solo si está completamente seguro que el pixel pertenece al fondo o al objeto.

3.- Representación de la duda total: si $\mu_{\tilde{Q}_{Bt}}(q) = 0.5$ y $\mu_{\tilde{Q}_{0t}}(q) = 0.5$; esto es, si el experto no es capaz de distinguir si el pixel pertenece al fondo o al objeto. Diremos que el desconocimiento que tiene el experto sobre la pertenencia de ese pixel es uno.

4.- Si la pertenencia del pixel al fondo y la pertenencia al objeto son mayores que 0.5, entonces el desconocimiento debe decrecer respecto a ambas pertenencias.

5.- El desconocimiento no debe presentar una reacción caótica frente a pequeños cambios en la opinión del experto respecto a la pertenencia de un pixel. En otras palabras la expresión del desconocimiento debe ser continua.

Las funciones de pertenencia del algoritmo clásico se definen en el intervalo $[0.5, 1]$. Esto es debido a que el algoritmo minimiza la entropía total de la imagen fusificada. Al estar definida la función de pertenencia entre 0.5 y 1, cuando tomamos el conjunto con menor entropía (paso (c)), entonces ese conjunto es el conjunto con mayor número de píxeles con pertenencia cercana a 1, por tanto tomamos el conjunto que mejor clasifica los píxeles (ver [5] y [1]).

Teniendo en cuenta todas las consideraciones anteriores y que las funciones de pertenencia en el algoritmo difuso las definimos en $[0.5, 1]$, nos han llevado a dar la siguiente definición:

Definición 2 *Llamaremos a una función* $G_u : [0.5, 1]^2 \to [0, 1]$ *función de desconocimiento si satisface las siguientes condiciones:*

(G_u1) $G_u(x, y) = G_u(y, x)$ *para todo* $x, y \in [0.5, 1]$;

(G_u2) $G_u(x, y) = 0$ *si y sólo si* $x = 1$ *or* $y = 1$;

(G_u3) *Si* $x = 0.5$ *y* $y = 0.5$ *entonces* $G_u(x, y) = 1$;

(G_u4) G_u *es decreciente;*

(G_u5) G_u *es continua.*

Aunque hemos definido las funciones de desconocimiento a partir de dos funciones de pertenencia cualesquiera, en nuestro algoritmo $G_u(1, 1)$ no puede darse debido a la forma de construcción de las funciones de pertenencia. Además al interpretar el valor de las funciones de pertenencia como el conocimiento que tiene el experto de la pertenencia del pixel a una zona, el valor 0.5 significa desconocimiento total del experto. Al imponer que las funciones de pertenencia esten en el intervalo $[0.5, 1]$ estamos obligando al experto a no dar conocimiento en negativo. Por ejemplo $\mu_{\tilde{Q}_{Bt}}(q) = 0$ significaría que el pixel no pertenence al fondo. Esto es debido a que no queremos que el experto de su opinión relacionando la pertenencia a una clase con la negación de la pertenencia a otra clase.

En el siguiente teorema presentamos un método de construcción de funciones G_u a partir de t-normas.

Teorema 1 *Sea* T *una t-norm continua tal que*

$$T(x, y) = 0 \text{ si y solo si } x \cdot y = 0.$$

Bajo estas condiciones la función:

$$G_u(x, y) = \frac{T(1 - x, 1 - y)}{T(0.5, 0.5)} \tag{10}$$

es una función de desconocimiento.

3.3 Algoritmo utilizando conjuntos intervalo valorados difusos

En esta sección presentamos el nuevo algoritmo que utiliza conjuntos intervalo valorados difusos para umbralizar imágenes. Los conjuntos intervalo valorados difusos están construidos a partir de $\tilde{Q}_B$ y $\tilde{Q}_O$, y son tales que la amplitud del intervalo que representa la pertenencia de un pixel a la imagen describe el desconocimiento que tiene el experto en la construcción de las pertenencias de dicho pixel a los conjuntos difusos $\tilde{Q}_B$ y $\tilde{Q}_O$. Evidentemente vamos a utilizar funciones de desconocimiento para construir dichos intervalos.

3.3.1 Conjuntos intervalo valorados difusos

Un conjunto intervalo valorado difuso es tal que el grado de pertenencia de cada elemento al conjunto esta dado por un subintervalo cerrado del intervalo [0,1]. El concepto de conjunto difuso tipo 2 fue introducido por Zadeh ([11]) como una generalización de un conjunto difuso ordinario. En los conjuntos difusos de tipo 2, el grado de pertenencia de cada elemento al conjunto es así mismo un conjunto difuso en [0,1].

Un conjunto difuso de tipo 2 intervalar $\overline{\overline{A}}$ en U esta definido por

$$\overline{\overline{A}} = \{(u, A(u), \mu_u(x))|u \in U, A(u) \in L([0,1])\},$$

donde $A(u) = [\underline{A}(u), \overline{A}(u)]$ es una función de pertenencia; esto es, es un subintervalo cerrado de $[0,1]$, y la función $\mu_u(x)$ representa el conjunto difuso asociado con el elemento $u \in U$ obtenido cuando x engloba al intervalo $[0,1]$; $\mu_u(x)$ es dada de la siguiente forma:

$$\mu_u(x) = \begin{cases} a \text{ si } \underline{A}(u) \leq x \leq \overline{A}(u) \\ 0 \text{ en otro caso} \end{cases},$$

donde $0 \leq a \leq 1$.

En este trabajo utilizamos como entropía intervalo valorada difusa la expresión dada por Sambuc [9].

$$\mathcal{E}(Q) = \sum_{i=0}^{N} (\overline{Q}(u_i) - \underline{Q}(u_i)) \tag{11}$$

3.3.2 Algoritmo IVFS

Los pasos del nuevo algoritmo que proponemos son:

(a) Construir L conjuntos difusos $\tilde{Q}_{Bt}$ asociados con la imagen Q.

(b) Construir L conjuntos difusos $\tilde{Q}_{Ot}$ asociados con la imagen Q.

(c) Para cada t, $(t = 0, 1, \cdots, L-1)$, construir un conjunto intervalo-valorado difuso tal que la longitud del intervalo de pertenencia de cada elemento cuantifique el desconocimiento existente al construir las funciones de pertenencia de los conjuntos difusos $\tilde{Q}_{Bt}$ y $\tilde{Q}_{Ot}$.

(d) Tomar como *mejor umbral* t el nivel de gris asociado con el conjunto intervalo valorado difuso de menor entropía intervalo valorada.

Algoritmo IVFS

En la siguiente proposición proponemos un método de construcción de IVFS a partir de conjuntos difusos, utilizando funciones de desconocimiento, que desde nuestro punto de vista, es uno de los mas sencillos y eficaces.

Proposición 1 *Sean $\tilde{Q}_B$ y $\tilde{Q}_O$ los conjuntos difusos asociados al fondo y al objeto de la imagen Q construidos por un experto. Sea G_u la función desconocimiento asociada a la construcción de los conjuntos difusos anteriores. Bajo estas condiciones, si definimos:*

$$\begin{aligned} \Phi: \; & \mathcal{FS}s(U) \times \mathcal{FS}s(U) \longrightarrow \mathcal{IVFS}s(U) \text{ dado por} \\ & \Phi(\tilde{Q}_B, \tilde{Q}_O) = \{(u_i, \mathbf{x}(u_i))|u_i \in U\} \text{ tal que} \\ & \mathbf{x}(u_i) = [G_u(0.5, 0.5) - G_u(\mu_{\tilde{Q}_B}(u_i), \\ & \qquad \mu_{\tilde{Q}_O}(u_i)), G_u(0.5, 0.5)] \end{aligned}$$

entonces para cada pertenencia $\mathbf{x}(u_i)$ definida por Φ se verifica que su amplitud es: $G_u(\mu_{\tilde{Q}_B}(u_i), \mu_{\tilde{Q}_O}(u_i))$.

4 Resultados experimentales

Para evaluar nuestro método utilizamos un conjunto de 10 pares de imágenes consistentes en una imagen original y la segmentación ideal generada por un experto radiólogo. Cada imagen tiene asociado un punto en el centro de la región a segmentar elegido por el experto. A partir de ese punto realizamos el realce de cada imagen según lo visto en la sección 2 y umbralizamos la imagen utilizando el algoritmo intervalo valorado. Además, el experto debe escoger dos funciones de pertenencia diferentes, una para representar los píxeles de la zona de la próstata y otra para representar los píxeles del fondo. En este caso hemos escogido las siguientes funciones:

$$\mu_{PROST_t}(q) = e^{-log(2)\frac{|q - m_{PROST}(t)|}{L-1}} \tag{12}$$

$$\mu_{FONDO_t}(q) = 0.5\left(2 - \left|\frac{q - m_{FONDO(t)}}{L-1}\right|^2\right) \tag{13}$$

Para construir los conjuntos intervalo valorados utilizamos la t-norma mínimo como función de desconocimiento. Además utilizamos la eq. (11) para calcular la entropía intervalo valorada.

En las figuras 3 y 4 mostramos las imágenes segmentadas por nuestro algoritmo y las comparamos con las imágenes ideales.

Para evaluar la eficacia del método de segmentación utilizamos la medida del solapamiento de áreas (S_A) entre el área de la imagen segmentada ideal y la imagen segmentada por el sistema IT2FLS.

$$S_A = \frac{|Ideal \cap IBF|}{|Ideal \cup IBF|} \tag{14}$$

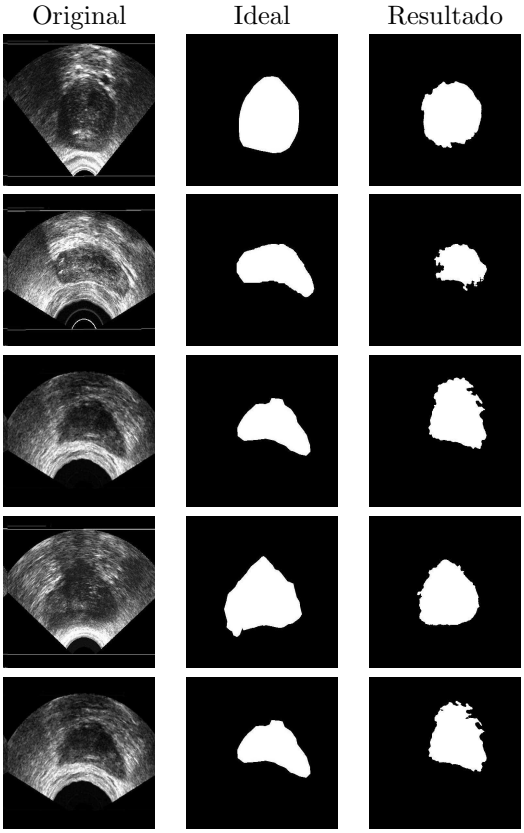


Figura 3: Comparación de las imágenes ideales con las obtenidas nuestro método.

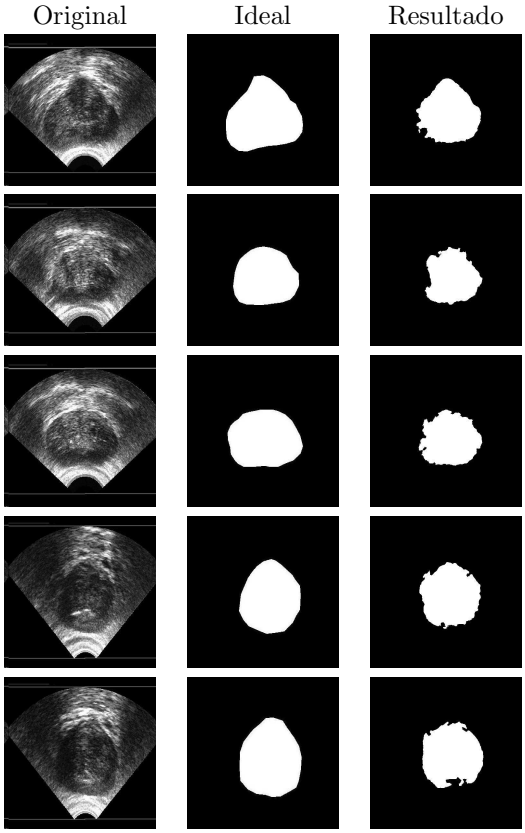


Figura 4: Comparación de las imágenes ideales con las obtenidas nuestro método.

Donde $Ideal$ es la imagen binaria segmentada por el experto, IBF es la imagen binaria obtenida utilizando nuestro método, $\cap$ y $\cup$ la intersección y la unión entre conjuntos crisp respectivamente.

En la tabla 1 mostramos los valores obtenidos por nuestro algoritmo y los obtenidos utilizando el método clásico de Otsu [7].

Al estudiar las imágenes binarias obtenidas (fig. 3 y 4) vemos que los bordes de las zonas segmentadas son muy irregulares. Esto es debido a que el algoritmo de realce no preserva la nitidez en los bordes del quiste. El porcentaje de solapamiento está alrededor del 60%. Es un porcentaje bajo con respecto a otras técnicas supervisadas que alcanzan el 90% (ver [6]). Sin embargo, es un porcentaje prometedor debido a que el algoritmo de realce de la imagen puede mejorarse y además no se ha realizado ningún postproceso de la imagen binaria obtenida.

5 Conclusiones

Hemos propuesto un nuevo algoritmo de segmentación de imágenes de ultrasonidos. El usuario puede escoger dos expresiones diferentes para representar el objeto y el fondo. De esta forma se pueden obtener imágenes segmentadas con un número de píxeles muy pequeño en los que existe duda sobre su pertenencia. Considerando los resultados en terminos de apariencia visual y los porcentajes del solapamiento de las áreas, este método puede utilizarse para una primera estimación de la zona que ocupa la próstata y servir para un posterior refinamiento utilizando otras técnicas. Comprobamos que nuestro método obtiene mejores soluciones que el algoritmo clásico de Otsu. Debemos estudiar como el usuario puede escoger siempre las funciones de pertenencia que mejor representan al objeto y al fondo. Además debemos estudiar otros métodos de realce de las imágenes de ultrasonidos que preserven la forma de la próstata y no se pierda definición en sus bordes.

Agradecimientos

Este trabajo ha sido financiado en parte por el proyecto TIN2007-65981.

Referencias

[1] H. Bustince, E. Barrenechea, M. Pagola, Image thresholding using restricted equivalence func-

Imagen	1	2	3	4	5	6	7	8	9	10
IVFS $S_A(\%)$	76	54	58	71	60	67	56	57	61	63
OTSU $S_A(\%)$	16	65	14	80	14	82	62	67	19	20

Tabla 1: Solapamiento del área.

tions and maximizing the measures of similarity, *Fuzzy Sets and Systems*, **158**, 496–516 (2007).

[2] Bustince H., Barrenechea E., Pagola M., Restricted Equivalence Functions, *Fuzzy Sets and Systems*, 157, 2333-2346 (2006).

[3] Canadian Cancer Society. Prostate cancer stats 2003.

[4] M.G. Forero. Fuzzy thresholding and histogram analysis, in *Fuzzy Filters for Image Processing*, Edited by M. Nachtegael, D. Van der Weken, D. Van de Ville and E.E. Kerre, Springer, pages 129-152, 2003.

[5] L.K. Huang and M.J. Wang. Image thresholding by minimizing the measure of fuzziness, *Pattern recognition*, Vol. 28(1), pages 41-51, 1995.

[6] J.A. Noble and D. Boukerroui, Ultrasound image segmentation: A survey, IEEE Transactions on Medical Imaging 25 (2006), pp. 987-1010.

[7] N. Otsu. A threshold selection method from gray level histograms, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 9, pages 62-66, 1979.

[8] F.Sahba, H.R. Tizhoosh, M.M. Salama, A coarse-to-fine approach to prostate boundary segmentation in ultrasound images, BioMedical Engineering OnLine 2005, 4:58.

[9] R. Sambuc. Function Φ-Flous, Application a l'aide au Diagnostic en Pathologie Thyroidienne, *These de Doctorat en Medicine*, University of Marseille, 1975.

[10] Tizhoosh, H.R., Observer-Dependent Image Enhancement, Proceeding of FUZZ-IEEE 2001.

[11] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning – I, *Information Science*, **8**, 199–249, 1975.

Clasificador Adaptativo para el Control Difuso de Vehículos Marinos

Germán Villalba Madrid[1] **David Herrero Pérez**[2] **Humberto Martínez Barberá**[1]
germanvm@um.es **dherrero@ing.uc3m.es** **humberto@um.es**

[1] Universidad de Murcia [2] Universidad Carlos III

Resumen

El presente trabajo trata sobre el estudio del control de una embarcación autónoma. Para ello se ha diseñado un sistema organizado en dos niveles, uno inferior en el que se encuentran una serie de controladores específicos para distintas condiciones de mar, y uno superior en el que se encuentra un clasificador que, en función de las condiciones estimadas de mar, selecciona uno u otro controlador específico, combinación que proporciona un control adaptativo efectivo.

Al ser el ajuste de los controladores un problema complejo, se ha desarrollado un esquema de aprendizaje mediante el cual se entrenan tanto los controladores específicos como el clasificador. Se han realizado diversos experimentos para contrastar la idoneidad de la solución propuesta en un simulador realista.

Palabras Clave: Adaptativo; Algoritmo Evolutivo; Cinemática; Embarcación; PD Difuso

1 Introducción

La utilización de ASC (Autonomous Surface Craft) es altamente potencial en tareas de exploración, toma de muestras e incluso en apoyo a operaciones especiales. El desarrollo de este tipo de embarcaciones presenta diversos problemas nuevos e interesantes, uno de los cuales es el control autónomo en general, y en particular el ajuste de los controladores.

La eficacia y eficiencia de controladores que guíen convenientemente una embarcación siguiendo una secuencia de waypoints depende en gran medida de las condiciones del entorno: mar de fondo, batimetría, viento, etc. Frente a otros robots autónomos donde la dinámica se puede considerar despreciable o poco importante frente a un modelo cinemático, en un ASC la interacción de la embarcación con el oleaje y la fuerza del viento no se puede obviar, y por tanto el modelo dinámico es un punto importante [4] [6] [7]. Por tanto, surgen problemas relativos al modelado de la embarcación. La simulación de controladores para el gobierno de embarcaciones es amplicamente utilizado para reducir los costes, pero la complejidad debido a los 6 grados de libertad (DOF) posibles en un barco, motiva a realizar simplificaciones considerando despreciables algunos de ellos, como en [4] donde se reducen en algunos tipos de barcos a 4 DOF.

En [7] se distingue entre control para maniobra y para mantenimiento de las condiciones de mar y velocidad. Así se establecen condiciones de mar diferentes (mar en calma para las maniobras y efecto del oleaje para el segundo caso) y distintos sistemas de referencia, como son un sistema inercial (n - frame) ligado a la tierra y otro sistema de referencia sujeto a la embarción (b-frame).

Otros estudios [1], utilizan controladores difusos aplicados a veleros, donde el problema introduce otras variables como son la fuerza y dirección del viento, además de la respuesta de la embarción respecto al vector viento en cuanto a trimado de las velas (polares).

Una variante al problema se estudia en [8], donde se trata el posicionamiento de un ROV (Remotely Operated Vehicle) en las proximidades de estructuras sumergidas con el fin de realizar acciones de exploración y supervisión, utilizando como controladores unos PD adaptativos. Sin embargo, desprecia las fuerzas y momentos inducidos por la mar, debido a que se encuentra fuera de su influencia (la profundidad de operación es mayor de 20 metros).

Los costes asociados a las pruebas de optimización de

controladores para embarcaciones marinas obliga a la utilización de plataformas software donde se puedan simular las condiciones de trabajo, y así facilitar la optimización de la solución buscada. Se ha desarrollado un entorno de simulación de oleaje basado en un modelo lineal de oleaje [3] [5]. Así, dados unos valores de altura de ola y periodo en aguas profundas (límite de contorno), estas perturbaciones se propagan hacia la costa, siendo su altura y velocidad dependientes de las características de la batimetría. Se pretende pues obtener un controlador que permita llevar nuestra embarcación a través de una serie de waypoints en diferentes condiciones de mar.

Para resolver el problema, se ha desarrollado un sistema en dos niveles, donde en el nivel más bajo se encuentran una serie de controladores difusos específicos optimizados para unas condiciones de mar determinadas, mientras que en el nivel superior un clasificador adaptativo decide cual será el controlador específico a utilizar según las condiciones de mar estimadas. Así, se obtiene un control con un mejor comportamiento medio en todas las condiciones. Los criterios que conforman la función objetivo a minimizar están vinculados con la minimización de la energía consumida.

2 Entorno de Simulación

El entorno de trabajo software está desarrollado en JAVA, incluye un simulador de oleaje lineal, simulación de la cinemática de una embarcación e implementa los recursos necesarios para trabajar con controladores difusos y algoritmos de optimización.

El sistema permite iniciar un nuevo aprendizaje, continuar un aprendizaje detenido, modificar las condiciones de entorno durante un aprendizaje e iniciar una simulación. Todo esto, escogiendo el escenario, con la zona de navegación y el tipo de perturbación. Además, carga los datos de la orografía a partir de ficheros **DEM** *(Digital Elevation Model)*, mientras que los resultados obtenidos se almacenan en ficheros XML.

2.1 Simulación del Entorno Marino

Para simular los efectos del oleaje y viento, se ha desarrollado un simualdor de oleaje y viento basado en el modelo de propagación de oleaje lineal [3]:

$$\omega - kU = \sqrt{gktanh(kh)} \tag{1}$$

donde ω es la velocidad angular, k es el número de onda, U es la velocidad de la corriente, g es la constante gravitacional y h es la profundidad en el punto dado.

La amplitud máxima de la perturbación queda limitada, según [9] a:

$$H_{max} = 0.142Ltanh(kd) \tag{2}$$

siendo H la altura de la ola, L la longitud de onda, k el número de onda y d la profundidad en la posición (x, y) de cálculo.

Quedando el sistema dependiente de las condiciones de altura de ola y periodo en aguas profundas, así como de la batimetría en cada punto.

$$H(x, y) = f(H_0, T_0, d) \tag{3}$$

siendo $H(x, y)$ la amplitud de la ola en metros para una posición (x, y), H_0 la amplitud de la ola en los límites de aguas profundas, T_0 el periodo de la ola (invariable durante la propagación) y d la profundidad del fondo marino en un punto dado.

El diferente comportamiento del oleaje se clasifica en zonas dependiendo de la profundidad y la longitud de onda, según la tabla 1:

Tabla 1: Profundidad Relativa.

Shallow	**Transitional**	**Deep**
$\frac{d}{L} < \frac{1}{25}$	$\frac{1}{25} < \frac{d}{L} < \frac{1}{2}$	$\frac{d}{L} > \frac{1}{2}$

El entorno de simulación permite realizar experimentos con las condiciones límites de contorno invariables a lo largo del tiempo, manteniendo los valores de altura de ola a la media de la escala Douglas, o también, presentando variaciones según una función aleatoria gaussiana dentro de la escala correspondiente para la altura de ola.

2.2 Cinemática de la Embarcación

Se ha desarrollado un simulador básico de una embarcación [10], utilizando una simplificación de la dinámica y cinemática correspondiente a una embarcación que cuenta con un único propulsor y un timón. De esta forma [2] se considera la siguiente ecuación de movimiento:

$$m\dot{u} = X_\xi(\xi) + R(u) + T \tag{4}$$

donce m es la masa de la embarcación, X_ξ es la fuerza longitudinal de la ola, que depende del parámetro adimensional ξ, R es la fuerza de resistencia que depende de la velocidad u y T es el empuje.

Así la cinemática implementada consiste en la variación del rumbo a partir de la acción sobre el timón, mediante la estimación del radio de giro, teniendo en cuanta las restricciones impuestas por la dinámica, tanto de la embarcación como del actuador. Respecto a la navegación se ha fijado la velocidad máxima a $3.5m/s$, con una aceleración máxima de $0.4m/s^2$. Respecto a las fuerzas de resistencia se ha tenido en cuenta la presentada por el timón cuando modifica su ángulo respecto a crujía.

Además, se han incluido unas limitaciones en la dinámica del actuador como son la velocidad de cambio y los valores máximos [4]. Comercialmente, para el tipo de embarcación propuesta, se pueden encontrar servomotores para control del timón con velocidades angulares comprendidas entre $0.1sec/60°$ hasta $0.7sec/60°$. Aún así, tomamos como referencia un actuador ligeramente más lento, de $1.0sec/60°$. El ángulo máximo corresponderá a $\pm 45°$.

2.3 Interacción Mar - Embarcación

Para la navegación, se ha relacionado el desplazamiento de la embarcación provocada por la propulsión del mismo, más la perturbación generada por la ola en las coordenadas instantáneas de su posición.

La posición de la embarcación se modifica por la transferencia de energía cinética y potencial instantánea de la ola a la embarcación, siendo función de la altura de la ola, teniendo en cuenta que ésta depende de la posición en el mapa [5]. Dicha energía se transfiere en la dirección de propagación de la perturbación, solo en el frente de avance.

3 Controladores Específicos

El control de la embarcación, para unas condiciones de viento y oleaje determinados, se realiza mediante un controlador PD-difuso, entrenado para tal fin mediante un algoritmo evolutivo. Dicho controlador tiene como entradas el error de posicionamiento de la embarcación con respecto a la trayectoria deseada (fijada por un waypoint de inicio y otro de llegada) y la variación del mismo. La salida del controlador es el ángulo que se ha de girar el timón para controlar la nave. En todo momento se asume una velocidad de crucero, salvo en las proximidades de los waypoints, necesario para negociarlos correctamentes. Una vez fijadas dichas velocidades (normalmente fijados por requisitos de consumo), el problema del control del timón es altamente complejo, debido principalmente a la imposibilidad del modelado preciso del conjunto embarcación-entorno, a la no linealidad del sistema, y las imprecisiones e incertidumbres del conjunto del sistema de control.

Tras diversos ensayos de simulación se ha establecido como modelo adecuado controladores PD-difusos MISO tipo Sugeno, compuestos por una matriz de 15 reglas (tabla 2), con 5 conjuntos difusos para el término de error y 3 para la derivada del error. La estructura de las reglas y de los conjuntos de entrada presenta ciertas simetrías, que se han tenido en cuenta a la hora de implementar el algoritmo evolutivo (tabla 2).

Tabla 2: Base de Reglas de Controlador Difuso.

de / e	**NL**	**NS**	**Z**	**PS**	**PL**
N	AC0	AC5	AC6	-AC3	-AC2
Z	AC1	AC4	Z	-AC4	-AC1
P	AC2	AC3	-AC6	-AC5	-AC0

3.1 Codificación

La codificación de los controladores difusos se ha realizado mediante tres cromosomas, que representan: los conjuntos difusos para los antecedentes (uno para el término de error y otro para el término de la derivada) y los consecuentes crisp de la base de reglas difusas, contando con 3, 2 y 7 genes respectivamente.

Los cruces solo se pueden realizar entre cromosomas iguales, intercambiando genes entre ellos. Se permite tanto el cruce o mutación de varios cromosomas como el bloqueo a la modificación de un solo cromosoma por generación. Una vez cruzados o mutados, se someten a un test de factibilidad, y después se evalúan para poder ser ordenados con el resto de la población.

3.2 Proceso de Aprendizaje

Se parte de una población de tamaño máximo fijo N, generada de forma aleatoria. Todos los individuos se evalúan en los entornos de aprendizaje, asignándoles un fitness $g(x)$ según los criterios de distancia recorrida, tiempo necesario y acciones de control realizadas. Utilizando el fitness como parámetro de ordenación, se escogen como élite un número dado de controladores que minimizan la función objetivo.

Posteriormente se crea la nueva generación de individuos, teniendo en cuenta las probabilidades establecidas de mutación y cruce, p y q, respectivamente. Estos nuevos controladores se evalúan en las condiciones de aprendizaje, para poder ordenarlos según fitness con los anteriores individuos. Para evitar perder heterogeneidad, aquellos individuos que tengan una edad en generaciones superior al límite *age* establecido son eliminados, excepto los seleccionados para formar parte de la élite. El valor de *age* se relaciona inversamente con la probabilidad de mutación p para obtener

el efecto deseado. De los individuos que permanecen, se toman los N con menor fitness, descartando el resto.

Además de la exploración en busca de nuevos mínimos locales, cuando el valor de fitness de un determinado individuo es menor de un valor predeterminado, ese controlador en concreto entra en un proceso de explotación del óptimo encontrado.

3.3 Función Objetivo

La función objetivo a minimizar corresponde a la suma ponderada de la distancia, tiempo y acciones de control. La distancia corresponde a un criterio relacionado con las trayectorias seguidas, favoreciendo aquellas que requieran un coste energético menor para alcanzar el waypoint de destino (esto, en función del oleaje, no tiene por qué corresponder a una línea recta). El tiempo corresponde a un criterio que intenta obtener la trayectoria más rápida. Por último, las acciones de control estarán vinculadas con la estabilidad de la embarcación, el confort en la misma y el consumo del actuador. Para poder comparar los resultados obtenidos en distintos entornos de simulación (variando las distancias, la posición de los waypoints, atc) es necesario realizar una normalización en cada uno de los criterios.

Distancia normalizada: corresponde a la distancia en línea recta siguiendo el recorrido establecido, y descontando el offset en cada waypoint.

$$distNorm = \sum_{i=0}^{n-1} (dist|wp_{i+1} - wp_i|) \tag{5}$$

siendo wp los waypoints seleccionados, y n el número de waypoints por los que debe pasar.

Tiempo normalizado: es el tiempo que tardaría en recorrer la distancia normalizada a 2/3 de la velocidad máxima.

$$timeNorm = \frac{distNorm}{\frac{2}{3}V_{max}} \tag{6}$$

Acciones de control normalizada: se determina empíricamente, siendo muy crítica en la optimización. Este valor, junto al peso asignado a esta función, son más difíciles de ajustar que las correspondientes a la distancia y el tiempo.

$$actNorm = f(t_N, K) \tag{7}$$

siendo t_N el tiempo de normalización y K un parámetro que depende de la geometría de las balizas. Se distingue entre acciones de control durante la maniobra y cuando debe mantener rumbo a siguiente waypoint, basada en [7].

El método de la suma ponderada $g(x)$ de $f_i(x) \in F$, reduce el problema multiobjetivo a simple objetivo [11].

$$g(x) = \sum_{i=1}^{n} w_i f_i(x) \tag{8}$$

siendo $f_i(x)$ las funciones a minimizar y w_i los pesos, verificando que $0.0 < w_i < 1.0$.

4 Clasificador Adaptativo

Una vez obtenidos controladores difusos para unas condiciones específicas de mar, se pretende que un clasificador sea capaz de seleccionar cuando tiene que intervenir cada uno de los controladores específicos, adaptando las acciones de control a las condiciones en de mar. Puesto que no existe ningún sensor práctico que sea capaz de establecer en qué condiciones de mar se encuentra navegando una embarcación en un instante determinado (se puede conocer la meteorología general, pero no la condición de oleaje instantánea), el clasificador ha de estimar dichas condiciones a partir de la observación de la variación del error y del efecto de las acciones de control (estas últimas dependen en gran medidad de las condiciones de mar). Debido a que dicha estimación presenta una gran incertidumbre y no haber una relación lineal, se ha optado por diseñar un clasificador adaptativo difuso.

Tras diversos ensayos de simulación se ha establecido como modelo adecuado un clasificador difuso MIMO tipo Sugeno, compuestos por una matriz de 9 reglas, con 3 conjuntos difusos para el término de error y 3 para la acción de control. La estructura de las reglas y de los conjuntos de entrada presenta ciertas simetrías, que se han tenido en cuenta de forma similar a los controladores específicos.

Se ha utilizado el mismo algoritmo evolutivo que para los controladores específicos, con la única diferencia que la codificación de las reglas incluyen genes adicionales para las nuevas salidas. Así, la salida del clasificador es, para cada controlador específico, la idoneidad del mismo para las condiciones instantáneas. En cada momento se selecciona el controlador más idóneo, rompiendo empates mediante la selección de aquel que sea menos malo en las condiciones para las que no ha sido entrenado.

5 Resultados Experimentales

Debido al tiempo necesario para realizar las simulaciones y producir un aprendizaje efectivo, se consideran únicamente dos controladores específicos para condiciones de mar *Rizada* y *Gruesa*, aunque se ha utilizado una condición específica adicional *Marejada* para validar el controlador adaptativo. Las características de amplitud y periodo de las condiciones anteriores en aguas profundas están definidas en la tabla 3. Hay que tener en cuenta que al no incluir oleaje de viento, la perturbación es de baja frecuencia para todos los casos analizados.

Tabla 3: Condiciones de Mar.

	Rizada	**Marejada**	**Gruesa**
H (m)	0.05	0.875	3.25
T (s)	1.0	5.0	10.0

Se han dispuesto recorridos delimitados por 4 balizas formando un rombo, siguiendo unas trayectorias a lo largo del perímetro y su diagonal menor. Dichos recorridos se ha ubicado en tres zonas de navegación diferentes, como se indican a continuación;

- Aguas Profundas: situada en un área donde la batimetría no influye en la perturbación del oleaje.
- Aguas Costeras: la zona de navegación está influenciada por la profundidad, además de tener una baliza a sotavento del origen de la perturbación.
- Aguas de Estrecho: la zona de navegación está influenciada por la profundidad y sometida a la perturbación en todo el recorrido.

El algoritmo evolutivo, para cada caso, se ha simulado en las tres zonas de navegación con la condición de mar correspondiente. Para la función de fitness, se han tenido en cuenta los siguientes pesos (tabla 4):

Tabla 4: Asignación de pesos.

Criterio de Evaluación	**Peso**
Distancia	0.34
Tiempo	0.36
Acciones	0.30

Cada ensayo se ha limitado a 2.5 veces el tiempo normalizado, con el fin de establecer un máximo en el tiempo de evaluación. Después de la evaluación, se generan los nuevos individuos que proceden de los cruces y mutaciones, y se evalúan estos nuevos controladores. Los individuos se ordenan según fitness, tanto la nueva como la antigua población, seleccionando los 30 mejores después de eliminar aquellos que sobrepasan en edad el valor de $age = 8$. Se ha comprobado que este valor junto con el utilizado para las mutaciones y cruces, permite que la información se transfiera a las generaciones sucesivas antes de desaparecer.

La probabilidad de mutación y cruce con la que se ha trabajado es de $p = 0.12$ y $q = 0.30$, respectivamente. Y se conserva el elitismo únicamente para el mejor individuo. Los demás pueden desaparecer de una generación a la siguiente.

5.1 Controladores Específicos

Se han optimizado dos controladores específicos, uno para mar Rizada y otro para mar Gruesa, que se comentan a continuacin.

5.1.1 Controlador para mar Rizada

El controlador para mar rizada, cuya superficie de control se muestra en la figura 1, se ha obtenido tras 2500 generaciones. Se muestran en la figura 2 los conjuntos difusos correspondientes al término de error y al de la derivada del error.

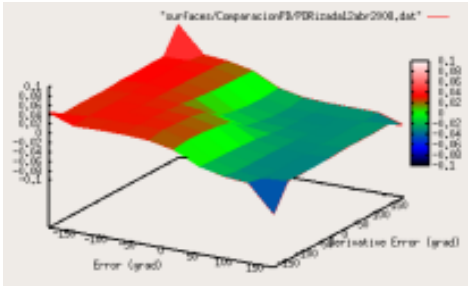


Figura 1: Superficie de control para mar Rizada

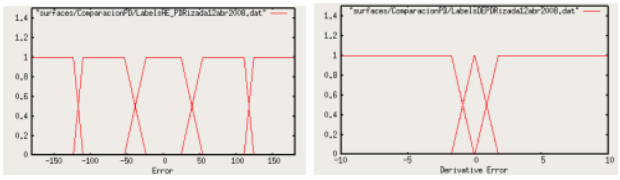


Figura 2: Conjuntos difusos para mar Rizada

5.1.2 Controlador para mar Gruesa

El controlador para mar rizada, cuya superficie de control se muestra en la figura 3, se ha obtenido tras 2500 generaciones. Se muestran en la figura 4 los conjuntos

difusos correspondientes al término de error y al de la derivada del error.

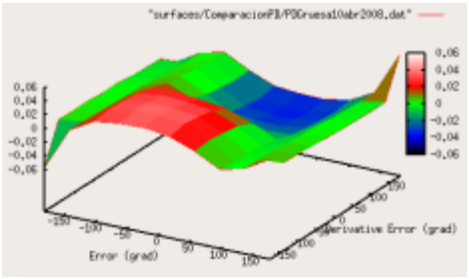


Figura 3: Superficie de control para mar Gruesa

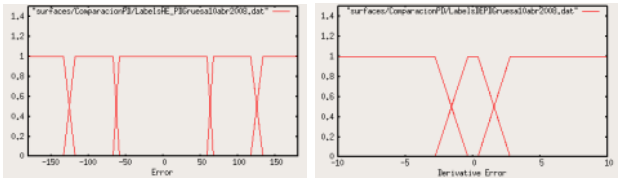


Figura 4: Conjuntos difusos para mar Gruesa

5.1.3 Anlisis de los controladores

En las figuras 2 y 4 se muestra cómo el controlador de mar Gruesa amplía la banda muerta, frente al de mar Rizada. Esto permite al PD-Gruesa no sobrecontrolar cuando el mar está turbulento, mientras que cuando la perturbación es mínima precisa de que el efecto integral de la perturbación sobre la embarcación sea suficiente para que produzca un cambio en la acción de control.

Para las derivadas de error, teniendo en cuenta que su entrada se mide en grados, la diferencia entre ambos controladores aparenta ser mínima. Sin embargo, se comprobó que si los límites de la etiqueta difusa *ZERO* se excedían de ± 3.5, la embarcación no era capaz de finalizar el recorrido, por lo que la diferencia en los conjuntos difusos correspondientes son considerables.

En la figura 5 se muestran las perturbaciones (línea azul) y las acciones de control (línea roja) de ambos controladore sen sus respectivos entronos de aprendizaje. Se observa que las perturbaciones de mar Gruesa son considerablemente mayores que las de Rizada, y que , correspondientemente, se producen más accioens de control.

5.2 Clasificador Adaptativo

En la figura 6 se muestran las acciones de control del clasificador adaptativo obtenido tras 2500 generaciones.

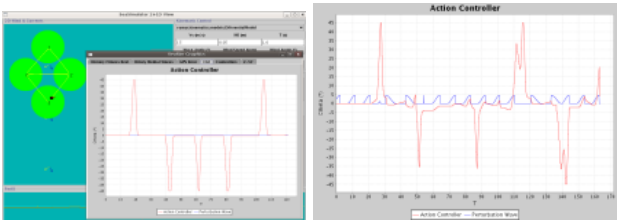


Figura 5: Acciones de control para mar Rizada y Gruesa

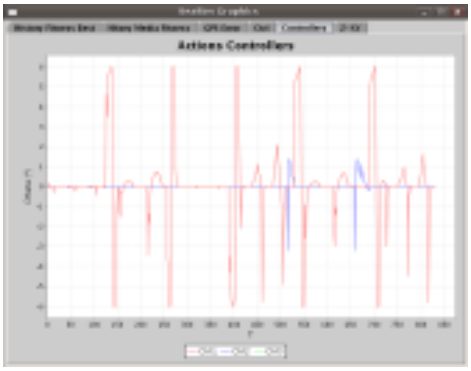


Figura 6: Acciones de control del clasificador adaptativo

El clasificador adaptativo proporciona, según la tabla 5, una salida intermedia entre los dos controladores utilizados. Dicha tabla muestra los controladores especficos de mar Rizada y Gruesa y el controlador adaptativo evaluados en las tres condiciones de operacin Rizada, Marejada y Gruesa. Como se observa, los especficos son mejores, con diferencia, para sus condiciones de entrenamiento, mientras que son peores en las demás. El controlador adaptativo, para todas las condiciones de mar, está comprendido entre los dos controladores específicos, aunque se percibe una mayor sintonía o aproximación hacia el específico de mar Rizada, e.d. si la perturbacin apreciada es pequeña selecciona este, pero si es grande selecciona el de mar Gruesa, como cabía esperar.

Tabla 5: Tabla de Resultados: Fitness.

Controlador	Rizada	Marejada	Gruesa
Esp. Rizada	0.78	6.19	24.02
Esp. Gruesa	4.06	3.69	14.43
Adaptativo	1.27	5.91	18.36

6 Conclusiones

El presente trabajo muestra una aproximación inicial al diseño de un controlador adaptativo en dos niveles que permita controlar de una manera efectiva una embarcación bajo distintas condiciones de mar. Por un lado se tienen controladores específicos para distintas condiciones de mar, y por otro un clasificador que selecciona el controlador más adecuado según las condiciones de mar observadas. En todos los casos, los controladores han sido entrenados mediante un algoritmo evolutivo integrado en un simulador realista de las condiciones de mar y la dinámica de la embarcación.

Se han realizado una serie de ensayos experimentales en simulador que confirman que el sistema funciona adecuadamente. Los controladores específicos dan un buen resultado en sus condiciones de entrenamiento, mientras que el clasificador adaptativo ofrece un resultado adecuado en todas las condicioens de mar. La ventaja que ofrece el modelo de controlador multinivel es que permite realizar entrenamientos parciales, que son mucho menos costosos computacionalmente, ya que el espacio de búsqueda es menor.

Tras una serie de experimentos encaminados a comprobar la validez de todo el sistema, ahora surgen nuevas posibilidades que se van a ir encarando en el futuro. La primera y más obvia es la comprobacin de si el uso de controladores específicos adicionales mejoraría sustancialmente el resultado, en particular uno específico para Marejada. Tanbién se está planteando la posibilidad de reformular el entrenamiento como un problema multi-objetivo, ya que, en algunas cirscuntancias los términos de la función de fitness son contradictorios. Adicionalmente, se van a realizar pruebas en una embarcacin prototipo para validar los resultados del simulador.

Agradecimientos

El presente trabajo ha sido financiado por el proyecto CICYT DPI2007-66556-C03-02 y por el programa de ayuda a los grupos de excelencia de la fundación Séneca 04552/GERM/06

Referencias

[1] Dovalo, J.A., Salom, J. and Calvo, O. (1995), "Lógica Difusa a Bordo de un Velero", ESTYLF95

[2] Harris, D., Thomas, G. and Renilson, M. (2000), "Towards Predicting the Behavior of Yachts in Following Seas", 7^{th} International Conference on Stability of Ships and Ocean Vehicles. Launceston, Tasmania, 7 - 11 Febreary

[3] Jonsson, I.G. (1990) "Wave-current interactions". The Sea. Le Méhauté and Hanes, ed., John Wiley and Sons, New York, Chapter 7, 65-120

[4] López, E., and others, (2004) "Mathematical Models of Ships for Manoeuvring Simulation and Control", Journal of Maritime Research, Vol. I. No. 2, pp. 5-20. SEECMAR.

[5] McKee Smith, J., Sherlok, A. and Resio, D. (2001), "STWave Steady State Spectral Wave Model", report ERDC / CHL SR-01-1, USACE

[6] Pérez, T. and Blanke, M., (2002) "Mathematical Ship Modeling for Control Applications". Technical Report. University of Newcastle. Australia

[7] Pérez, T. (2005) "A Review of Geometrical Aspects of Ship Motion in Manoeuvring and Seakeeping and the Use of a Consistent Notation" MSS Technical Report (MSS-TR-001-2005) Marine System Simulator (MSS) Group

[8] Quang Hoang, N. and Kreuzer, E., (2006) "Adaptative PD-Controller for positioning of a remotely operated vehicle close to an underwater structure: Theory and experiments". http://www.sciencedirect.com

[9] Smith, J.M. (1999), "Wave Breaking on an Opposing Current", Coastal Engineering Technical Note IV-17, U.S. Army Engineer Research and Development Center, Vicksburg

[10] SNAME (1950) Nomenclature for treating the motion of a submerged body through a fluid. Technical Report Bulletin 1-5. Society of Naval Architecs and Marine Engineers, New York, USA.

[11] Weise, T. "Global Optimization Algorithms - Theory and Application" 2007 - http://www.it-weise.de/

A Case-based Reasoning Strategy for Classifying Leukemia Patients

Sara Rodríguez, Juan F. De Paz, Javier Bajo and Juan M. Corchado

Departamento de Informática y Automática, Universidad de Salamanca
Plaza de la Merced s/n, 37008, Salamanca, España
{srg, fcofds, jbajope, corchado}@usal.es

Abstract

The use of computational methods is fundamental in cancer research. One of the possibilities is the use of Artificial Intelligence techniques. Several of these techniques have been used to analyze expression arrays. However, the new Exon arrays which work with a large amount of data require novel solutions. This paper presents a Case-based reasoning (CBR) system for automatic classification of leukemia patients from Exon array data. The proposed CBR system incorporates novel algorithms for data filtering and classification. The system has been tested and the results obtained are presented in this paper.

Keywords: Case-based Reasoning, ESOINN neural network, leukemia classification.

1 Introduction

During recent years there have been great advances in the field of Biomedicine [21]. The incorporation of computational and artificial intelligence techniques to the field of medicine has yielded remarkable progress in predicting and detecting diseases [21]. Genomics deals with the study of genes, their documentation, their structure and how they interact [22]. We distinguish different fields of study within the genome. One is transcriptome, which deals with the study of ribonucleic acid (RNA), and can be studied through expression analysis [13]. This technique studies RNA chains thereby identifying the level of expression for each gene studied. It consists of hybridizing a sample for a patient and colouring the cellular material with a special dye. This offers different levels of luminescence that can be analyzed and represented as a data array. Traditionally, methods and tools have been developed to work with expression arrays containing about 50.000 data. The emergence of the Exon arrays [7], holds important potential for biomedicine. However, the Exon arrays require novel tools and methods to work with very large 5.500.000 amounts of data.

For some time now, we have been working on the identification of techniques to automate the reasoning cycle of several CBR systems applied to complex domains [16]. The objective of this work is to develop a CBR system that allows the identification of patients with various types of cancer. The model aims to improve the cancer classification based on microarray data. The system proposed in this paper presents a new synthesis that brings several artificial intelligence subfields together (filter techniques, clustering, and artificial neural networks). Specifically, the system presented in this paper uses a model which takes advantage of two novel methods for analyzing Exon array data: a technique for filtering data, and a technique ESOINN [6] for clustering. The first method combines various filtering techniques to dramatically reduce the dimensionality of the data.

The paper is structured as follows: The next section presents the problem that motivates this research, i.e., the classification of leukemia patients from samples obtained through Exon arrays. Section 3 and Section 4 describe the proposed CBR model and how it is adapted to the problem . Finally, Section 5 presents the results and conclusions obtained after testing the model.

2 Computational Methods in Cancer Research

Hematological malignancies like leukemia, have been the ground of chromosomal and genetic analysis for many years. Recently, conventional array-based expression profiling has demonstrated that chromosomal alterations are associated with distinctive patterns of expression. Leukemia is a blood cancer form, originating in a malfunctioning bone marrow that tends to produce abnormal red and white cells at an increased rate [17]. The four most important types of Leukemia are acute and chronic myelogenous leukemia (AML;CML) and acute and chronic lymphocytic Leukemia(ALL; CLL). About 25.000 new cases of both acute and chronic Leukemia appear every year. Most cases appear in adults and persons over 60 years, but acute lymphocytic Leukemia has an increased rate in children. About 12.000 adult cases are diagnosed annually as acute myelogenic

Leukemia, 8.000 as chronic lymphocytic Leukemia, 500 as chronic myelogenous forms, and about 3.500 as acute forms of lymphocytic Leukemia. The rest of the cases are unclassified blood cancer types. A recent study [19], shows that an estimated 19,900 new cases of myeloma were diagnosed in the United States in 2007.
Microarray technology is based on a database of gene fragments called expressed sequence tags (ESTs), which are used to measure target abundance using the scanned intensities of fluorescence from tagged molecules hybridized to ESTs [14]. The high dimensionality of the data provided by each Exon array presents problems of handling and processing. An expression analysis consists basically of two stages: normalization and filtering; clustering and classification. These stages can be automated and included in a CBR system.

3 CBR System for Classifying Exon Array Data

The CBR developed tool receives data from the analysis of chips and is responsible for classifying of individuals based on evidence and existing data. CBR is a type of reasoning based on the use of past experiences [11]. The primary concept when working with CBRs is the concept of case. A case can be defined as a past experience, and is composed of three elements: A problem description, which delineates the initial problem; a solution, which provides the sequence of actions carried out in order to solve the problem; and the final stage, which describes the state achieved once the solution was applied. The way cases are managed is known as the CBR cycle, and consists of four sequential phases: retrieve, reuse, revise and retain.

3.1. RETRIEVE

Contrary to what usually happens in the CBR, our case study is unique in that the number of variables is much greater than the number of cases. This leads to a change in the way the CBR functions so that instead of recovering cases at this stage, important variables are retrieved. Traditionally, only the similar cases to the current problem are recovered, often because of performance, and then adapted. In the case study, the number of cases is not the problem, rather the number of variables. For this reason variables are retrieved at this stage and then, the other stages of the CBR are carried out. This phase will be broken down into 5 stages:

3.1.1. RMA

The RMA (Robust Multi-array Average) [8] algorithm is frequently used for pre-processing Affymetrix microarray data. RMA consists of three steps: (i) Background Correction; (ii) Quantile Normalization; (iii) Expression Calculation.

3.1.2. Control and Errors

During this phase, all probes used for testing hybridization are eliminated. On occasion, some of the measures made during hybridization may be erroneous; not so with the control variables. In this case, the erroneous probes that were marked during the implementation of the RMA must be eliminated.

3.1.3. Variability

Once both the control and the erroneous probes have been eliminated, the filtering begins. The first stage is to remove the probes that have low variability. This work is carried out according to the following steps:

1. Calculate the standard deviation for each of the probes j

$$\sigma_{\cdot j} = +\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\overline{\mu}_{\cdot j} - x_{ij}\right)^2} \qquad (1)$$

   Where N is the number of items total, $\overline{\mu}_{\cdot j}$ is the average population for the variable j, x_{ij} is the value of the probe j for the individual i.
2. Standardize the above values z≡N(0,1)

$$z_i = \frac{\sigma_{\cdot j} - \mu}{\sigma} \qquad (2)$$

3. Discard of probes whose value of $z < -1.0$.

3.1.4. Uniform distributions

Finally, all remaining variables that follow a normal distribution are eliminated. The variables that follow a uniform distribution will not allow the separation of individuals. Therefore, the variables that do not follow this distribution will be really useful variables in the classification of the cases. The contrast followed is the Kolmogorov-Smirnov [3] test (KS-Test). $\alpha = 0.05$.

3.1.5. Correlations

At the last stage of the filtering process, correlated variables are eliminated so that only the independent variables remain. To this end, the linear correlation index of Pearson is calculated and the probes meeting the following condition are eliminated.

$$r_{x_i y_{\cdot j}} > \alpha \qquad (3)$$

being: $\alpha = 0.95$ $r_{x_i y_{\cdot j}} = \dfrac{\sigma_{x_i x_{\cdot j}}}{\sigma_{x_i}\sigma_{x_j}}$, Where $\sigma_{x_i x_{\cdot j}}$ is the covariance between the probes i and j.

3.2. REUSE

Once filtered and standardized, the probes produce a set of values x_{ij} with i = 1 ... N, j = 1 ... s where N is the total number of cases, s the number of end probes. The next step is to perform the clustering of individuals. Since the problem on which this study is based contained no prior classification with which training could take place, a technique of unsupervised classification was used. There is a wide range of possibilities. Some of these techniques are artificial neural networks such as SOM [10] (Self-Organizing Map), GNG [5] (Growing Neural Gas) resulting from the union of techniques CHL [15] (competitive Hebbian Learning) and NG [5] (Neural Gas), GCS [5] (Growing Cell Structure), Growing Grid or the SOINN [20] (self-organizing incremental neuronal network). Some of the methods, such as self-organized maps, set the number of clusters in the initial phase of training when using the algorithm of the k-means. The self-organized maps have other variants of learning methods that base their behaviour on methods similar to the NG. They create a mesh that is adjusted automatically to a specific area. The greatest disadvantage, however, is that the number of neurons that are distributed over the surface and the degree of proximity are set beforehand, resulting in the number remaining constant throughout the entire training process, thus complicating, the adaptation of the mesh. Unlike the self-organizing maps based on meshes Growing Grid or GCS do not set the number of neurons, or the degree of connectivity, but they do establish the dimensionality of each mesh. This complicates the separation phase between groups.

After analyzing different techniques and checking the problems they might present so that they might be applied to the problem at hand, we have decided to use a variation of neural network SOINN, called ESOINN [6] (Enhanced self-organizing incremental neuronal network). Unlike the SOINN, ESOINN consists of a single layer, so it is not necessary to determine the manner in which the training of the first layer changes to the second. With a single layer, ESOINN is able to incorporate both the distribution process along the surface and the separation between low density groups. The operation and training of the network presents many similarities with those used in GCS. Nevertheless, it more closely resembles a merger between a CHL and a NG. The training phase and the various algorithms applied at every stage are detailed below:

1. Create an empty set of nodes A
2. Create an empty set of interconnections between nodes $C \subset AxA$
3. Insert two nodes in the A and assign random values to weights W.
4. Select a pattern p $\xi_p = (x_{p1},...,x_{ps})$ of the data set I with dimension R^s where x_{pj} represents the intensity of luminescence in probe j of individual i.
5. Search nodes: a_1 node closest to the pattern of entry and a_2 the second node closest to the pattern of entry, with $a_i \in A$, using Euclidean distance as the measure distance.

$$a_1 = \arg\min_{a \in A} \left\| \xi_p - W_a \right\|, \quad a_2 = \arg\min_{a \in A \setminus \{a_1\}} \left\| \xi_p - W_a \right\| \tag{4}$$

Up to this point the same steps are taken without changing any of the CHL algorithm techniques. Next the part concerning the ESOINN network begins, and modifications are made in order to automatically adjust various parameters.

6. If the distance $a_1 > T_i$ or $a_2 > T_i$ add a new node in that position and continue with the step

$$T_i = \begin{cases} \max_{j \in N_i} \left\| W_i - W_j \right\| & N_i \neq \phi \\ \max_{j \in N_i} \left\| W_i - W_j \right\| & N_i = \phi \end{cases} \tag{5}$$

Being N_i the set of neighbouring nodes of i

7. Increase the age of nodes connected with a_1

$$e_i(t+1) = e_i(t) + 1 \tag{6}$$

8. If necessary, establish a new connection or delete it between a_1 and a_2 according to the constraints:
   - If either of the nodes is not associated with any subclass or both are in the same, then create a new connection between them and assign an age value of zero.
   - If both nodes are in different subclass A, B, then to calculate the greater density for every subclass A_{max} B_{max}. If some of the next conditions is true, both subclass are joint and a new connections is created, otherwise if there was a connection it is deleted.

$$\min(a_1) > \alpha_A A_{\max} \text{ or } \min(a_1) > \alpha_B B_{\max} \tag{7}$$

Being

$$\alpha_a = \begin{cases} 0.0 & 2.0 * mean_A \geq A_{\max} \\ 0.5 & 3.0 mean_A \geq A_{\max} > 2.0 * mean_A \\ 1.0 & A_{\max} > 3.0 * mean_A \end{cases}$$

$$mean_A = \frac{1}{N_A} \sum_{i \in A} h_i$$

This ensures that connections with nodes with a higher age be subsequently deleted.

9. Update the density of a_1 depending on the distance to the neighbour nodes

$$h_i = \frac{1}{N} \sum_{j=1}^{n} \sum_{k=1}^{\lambda} p_{ik} \tag{8}$$

Where N is the number of $\sum_{k=1}^{\lambda} p_{ik} \neq 0$, n the number of periods, λ the number of pattern for period $\lambda = 50 + \sqrt{\#D}$

$$p_{ik} = \begin{cases} \dfrac{1}{\left(1 + \dfrac{1}{m}\sum_{j=1}^{m} \|W_i - W_j\|\right)^2} & i \text{ winner} \\ 0 & eoc \end{cases} \quad (9)$$

Where m is the number of neighbours to node *i*.

10. Increase the number of times winner of the neuron

$$M_{a_1}(t+1) = M_{a_1}(t) + 1 \quad (10)$$

11. Update the weights of neurons by following a process similar to the SOINN, but introducing a new definition for the learning rate in order to provide greater stability for the model. This learning rate has produced good results in other networks such as SOM [4].

$$\Delta W_{a_1} = n_1(M_{a_1})(\xi - W_{a_1})$$
$$\Delta W_i = n_2(M_{a_i})(\xi - W_{a_i}) \text{ with } i \in N_i \quad (11)$$

Being $n_1(x) = \frac{1}{\sqrt{x}}$, $n_2(x) = \frac{1}{\sqrt{2 + x^2}}$

12. Delete the connections with higher age. The ages are typified z≡N(0,1) and are removed those whose values are in the region of rejection with k>0. α =0.05.
13. If the number of iterations is a multiple of λ
    1. Update subclass to which each neuron belongs bearing in mind the highest local density of each neuron with its neighbours.
    2. Delete nodes that meet any of the following conditions
        1. $\forall a \in A / \#N = 2 \Rightarrow h_a < c_1 \sum_{j=1}^{N_A} h_j / \#N$

           N is the set of neighbour nodes of a, c1=0.001
        2. $\forall a \in A / \#N = 1 \Rightarrow h_a < c_2 \sum_{j=1}^{N_A} h_j / \#N$

           N is the set of neighbour nodes of a, c_2=1.0.
14. The clustering of elements is carried out bearing in mind the connections among the neurons.
15. If all input patterns have been passed then a KS-Test [3] is carried out in order to determine if the density distribution for the neurons in each group follows a normal distribution. If so then the learning procedure is finished. α =0.05.

Once, the clusters have been made, the new sample is classified. Its association is carried out bearing in mind the similarity of the new case with the recovered variables in the first phase. The similarity measure used is::

$$d(n,m) = \sum_{i=1}^{s} f(x_{ni}, x_{mi}) * w_i \quad (12)$$

Where s is the total number variables, n and m the cases, w_i the value obtained in the uniform test and f the Minkowski [7] Distance that is given for:

$$f(x,y) = \sqrt[p]{\sum_i |x_i - y_j|^p} \text{ con } x_i, y_j \in R^p \quad (13)$$

This dissimilarity measure weighs those probes that have a less uniform distribution, since these variables don't allow a separation. In order to validate the selected distance, the Kruskal-Wallis [1] test was carried out. It was verified whether the proportion of errors in the classification of each one of the individuals were the same for each group, bearing in mind the different measure. The results are shown in table 1 α =0.05.

Table 1. Comparing errors. * Different median and = equal, (-) median of column less than median of row.

	Minkowski	Eclidean	Max Absolute
Minkowski			
Eclidean	=		
Max Absolute	*(-)	*(-)	

3.3. REVISE/RETAIN

The revision is carried out by an expert who determines the correction with the group assigned by the system. If the assignation is considered correct, then the retrieve and reuse phases are carried out again so that the system is ready for the next classification.

4 Case Study

In the case study presented in the framework of this research 248 samples are available from analyses performed on patients either through punctures in marrow or blood samples, which have been hybridized and analyzed through Exon arrays manufactured by Affymetrix. The aim of the tests performed is to determine whether the system is able to classify new patients based on the previous cases analyzed and stored.

Figure 1 shows a scheme of the bio-inspired model intended to resolve the problem described in Section 3. The proposed model follows the procedures that are performed in medical centres. As can be seen in Figure 1, a previous phase, external to the model, consists of a set of tests which allow us to obtain data from the chips and are carried out by the laboratory personnel. The chips are hybridized and explored by means of a scanner, obtaining information on the marking of several genes based on the luminescence. At that point, the CBR-based model starts to process the data obtained from the Exon arrays.

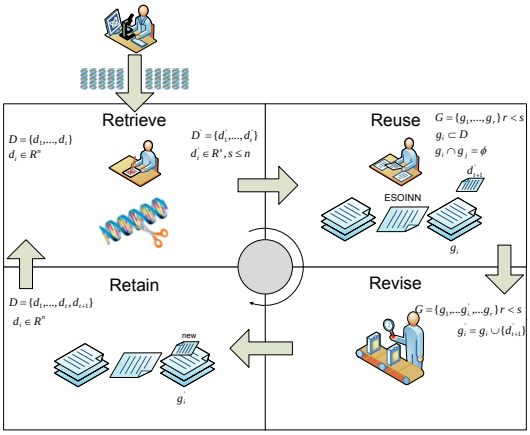


Figure 1. Proposed CBR model

The retrieve phase receives an array with a patient's data as input information. The retrieve step filters genes but never patients. The set of patients is represented as $D = \{d_1,...,d_t\}$, where $d_i \in R^n$ represents the patient *i* and *n* represents the number of probes taken into consideration. As explained in Section 3.1. during the retrieve phase the data are normalized by the RMA algorithm [8] and the dimensionality is reduced bearing in mind, above all, the variability, distribution and correlation of probes. The new set of patients is defined through *s* variables $D' = \{d_1',...,d_t'\}\ d_i' \in R^s, s \le n$.

The reuse phase uses the information obtained in the previous step to classify the patient into a leukemia group. The patients are first grouped into clusters. The data coming from the retriever phase consists of a group of patients $D' = \{d_1',...,d_t'\}$ con $d_i' \in R^s, s \le n$, each one characterized by a set of meaningful attributes $d_i = (x_{i1},...,x_{is})$, where x_{ij} is the luminescence value of the probe *i* for the patient *j*. In order to create clusters and consequently obtain patterns to classify the new patient, the reuse phase implements a novel neural network based on the ESOINN [6] one. The structure of this neural network has been described in detail in Section 3.2. The network classifies the patients by taking into account their proximity and their density, in such a way that the result provided is a set G where $G = \{g_1,...,g_r\} r < s$. $g_i \subset D$, $g_i \cap g_j = \phi$ with $i \neq j$ and $i, j < r$. The set G is composed of a group of clusters, each of them containing patients with a similar disease. Once the clusters have been obtained, the system can classify the new patient by assigning him to one of the clusters. The new patient is defined as d_{t+1}' and his membership to a group is determined by a similarity function defined in (12). The result of the reuse phase is a group of clusters $G = \{g_1,...g_i',...g_r\} r < s$ where $g_i' = g_i \cup \{d_{t+1}'\}$.

An expert from the Cancer Institute determines if $g_i' = g_i \cup \{d_{t+1}'\}$ can be considered as correct. In the retain phase the system learns from the new experience. If the classification is considered successful, then the patient is added to the memory case $D = \{d_1,...,d_t,d_{t+1}\}$.

5 Results and Conclusions

This paper has presented a CBR system which allows automatic cancer diagnosis for patients using data from Exon arrays. The model combines techniques for the reduction of the dimensionality of the original data set and a novel method of clustering for classifying patients.

In the study of leukemia on the basis of data from Exon arrays, the process of filtering data acquires special importance. In the experiments reported in this paper, we worked with a database of bone marrow cases from 248 adult patients with five types of leukaemia, plus a group of 16 samples belonging to healthy persons (no leukemias). Each case (microarray experiment) stores 289.961 ESTs corresponding to the expression level of thousands of genes. The data consisted of 5.500.000 scanned intensities. The retrieve stage of the proposed CBR system presents a novel technique to reduce the dimensionality of the data. The total number of variables selected in our experiments was reduced to 883. In addition, the selected variables resulted in a classification similar to that already achieved by experts from the laboratory of the Institute of Cancer of Salamanca. The error rates have remained fairly low especially for cases where the number of patients was high. Figure 2 shows an Intensity Plot with the classification performed for patients from groups CLL and ALL. As can be seen in Figure 2a, represented in black, most of the people of the CLL group are together, coinciding with the previous classification given by the experts at the Institute of Cancer. Only a small portion of the individuals departed from the initial classification, it is represented by the black lines that appear separated. Figure 2b shows the classification obtained for the LAL patients. We can see that, although the ranking is not bad, the proportion of individuals misclassified is higher. Groups that have fewer individuals are those with a higher classification error.

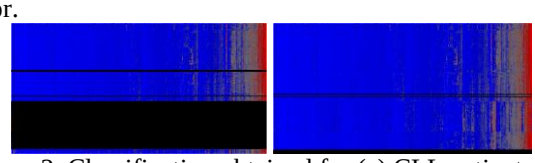


Figure 2. Classification obtained for (a) CLL patients and (b) ALL patients.

In a similar way we proceeded to evaluate the classification for the rest of the groups. Figure shows the total number of patients from each group and the number of misclassifications. As can be seen in Figure 3, groups with fewer patients are those with a greater error rate.

Figure 3b, shows the percentage of error in each group. Once the validity of the method of filtration for selecting the most important variables for classification is verified, the next step in the evaluation was to assess the functioning of the classification process. The system was tested with 15 new patients. The patients were assigned to the expected groups. Only one of the patients was misclassified.

Figure 3. Classification errors (a) numerical (b) percentage

The final classification was compared with the data obtained using a dendogram [18] and PAM [9] (Partitioning Around Medoids). The proportion of errors in every group was calculated and the Kurskal-Wallis [12] test was applied to determinate if the median of these proportions was equal. The results are shown in table 2.

Table 2. Comparation of methods. * different median and = equal, (-) median of column less than median of row

	CBR	Dendogram	PAM
CBR			
Dendogram	*(-)		
PAM	*(-)	*(-)	

As demonstrated, the proposed system allows the reduction of the dimensionality based on the filtering of genes with little variability and those that do not allow a separation of individuals due to the distribution of data. It also presents a technique for clustering based on the use of neural networks ESOINN. The results obtained allow the detection of pathology, and the facilitation of a classification and reliable diagnosis, as shown by the results presented in this paper

Acknowledge

This development has been partially supported by the Spanish Ministry of Science project TIN2006-14630-C03-03. Special thanks to the Salamanca Institute of Cancer for the information and technology provided.

References

[1] Aaronson, J., Juergen, H, Overton, G.: Knowledge Discovery in GENBANK. In Proceedings of the First International Conference on Intelligent Systems for Molecular Biology, (1993) 3–11.

[2] Affymetrix, GeneChip Human Exon 1.0 ST Array, www.affymetrix.com/products/arrays/specific/Exon.affx

[3] Brunelli, R.: Histogram Analysis for Image Retrieval. Pattern Recognition, Vol. 34, (2001) 1625-1637.

[4] Corchado, J. M., Bajo, J., De Paz Y., De Paz J. F. Integrating Case Planning and RPTW Neuronal Networks to Construct an Intelligent Environment for Health Care. International Journal of Computational Intelligence in Bioinformatics and System Biology, In Press.

[5] Fritzke, B. A growing neural gas network learns topologies. Cambridge MA: G. Tesauro, D. S. Touretzky, and T. K. Leen, Advances in Neural Information Processing Systems 7, (1995) 625-632.

[6] Furao, S, Ogura, T, Hasegawa, O.: An enhanced self-organizing incremental neural network for online unsupervised learning. Neural Networks, Vol. 20, (2007). 893-903.

[7] Gariepy, R., Pepe, W. D.: On the Level sets of a Distance Function in a Minkowski Space. Proceedings of the American Mathematical Society, Vol. 31, No. 1 (1972), 255-259

[8] Irizarry, R.A., Hobbs,B., Collin,F., Beazer-Barclay,Y.D., Antonellis,K.J., Scherf,U. Speed,T.P.: Exploration, Normalization, and Summaries of High density Oligonucleotide Array Probe Level Data. Biostatistics, (2003) 4, 249-264.

[9] Kaufman, L., Rousseeuw, P.J.: Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York. (1990).

[10] Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological Cybernetics, (1982) 59-69.

[11] Kolodner J.: Case-Based Reasoning. Morgan Kaufmann (1993).

[12] Kruskal W, Wallis W.: Use of ranks in one-criterion variance analysis. Journal of American Statistics Association (1952)

[13] Lander E.S. et al.: Initial sequencing and analysis of the human genome. Nature (2001). 409, 860 - 921

[14] Lipshutz, R. J., Fodor, S. P. A., Gingeras, T. R., and Lockhart, D. H.: High density synthetic oligonucleotide arrays. Nature Genetics, (1999) 21:20–24.

[15] Martinetz, T.: Competitive Hebbian learning rule forms perfectly topology preserving maps. Amsterdam : Springer,. ICANN'93: International Conference on Artificial Neural Networks. (1993). 427-434.

[16] Riverola F., Díaz F., Corchado J. M.: Gene-CBR: a case-based reasoning tool for cancer diagnosis using microarray datasets. Computational Intelligence. ISSN 0824-7935. Vol. 22 , num. 3-4, 254-268.(2006)

[17] Rubnitz, J.E., Hijiya, N., Zhou, Y., Hancock, M.L., Rivera ,G.K., Pui C.: Lack of benefit of early detection of relapse after completion of therapy for acute lymphoblastic leukemia. Pediatric Blood & Cancer, Vol. 44 (2), (2005). 138-141

[18] Saitou, N., Nie, M.: The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. (1987) Evol. 4: 406-425

[19] SEER (Surveillance Epidemiology and End Results), U.S. National Cancer Institute http://seer.cancer.gov/. (2007)

[20] Shen, F.: An algorithm for incremental unsupervised learning and topology representation. Tokyo: Ph.D. thesis. Tokyo Institute of Technology, (2006).

[21] Shortliffe, E.H., Cimino, J.J.: Biomedical Informatics: Computer Applications in Health Care and Biomedicine. Springer. (2006).

[22] Tsoka S., Ouzounis C.: Recent developments and future directions in computational genomics. FEBS Letters, Vol. 480 (1), (2000) 42-48

Making biological image sequence analysis simple: singular defuzzication of bivariate fuzzy temporal random sets

Teresa Leon[1] **Guillermo Ayala**[1] **Maria Elena Diaz**[2]

[1] Dept. de Estadistica e I.O., Universidad de Valencia {teresa.leon,guillermo.ayala}@uv.es
[2] Dept. de Informatica , Universidad de Valencia, elena.diaz@uv.es

Abstract

Endocytosis over the plasma membrane is the process by which cells carry traffic from the plasma membrane to endosomes. Total Internal Reflection Fluorescence Microscopy (TIRFM) allows to image this process and proteins are observed as overlapped areas forming random clumps of different sizes, shapes and durations. In this work we deal with pairs of gray-level image sequences, associated with two pairs of fluorescent-tagged proteins simultaneously imaged by TIRFM, considering them as realizations of a bivariate fuzzy random closed set. Spatial-temporal interactions are quantified by means of the fuzzy counterparts of the pair-correlation function and the Ripley's K-function. Differences in signal intensities between components requires that we match α-level sets, therefore we propose a procedure based on the Vorob'ev average which also allows the selection of thresholds in a unsupervised manner.

Keywords: Temporal random sets, fuzzy sets, clathrin mediated endocytosis, Total Internal Reflection Fluorescence Microscopy (TIRFM)

1 INTRODUCTION

TIRFM [12] permits to researchers to image two fluorophores simultaneously in space and time, thus allowing for the analysis of complex cellular mechanisms. We are interested in the study of the interdependencies between clathrin-GFP protein and other proteins which are involved in the endocytosis process, a cellular mechanism by which cells carry traffic from the plasma membrane to intracellular compartments. The study of the spatial relationship between proteins in living cells is essential.The reason is that although biochemical assays can show that given proteins can interact, in the cellular context what matters is if these proteins do targeted to the same regions. This spatial correlation is not a proof of an interaction, although can be very suggestive when, combined with additional manipulations, can prove that such interactions are meaningful. For clathrin mediated endocytosis (CME), many of the key proteins players are known ([6]), but understanding their spatial and spatial temporal inter-dependencies in normal and perturbed conditions will allow new insight into their function. Currently in this field most of the analysis is confined to visual inspection or limited statistical analysis of only isolated endocytic spots which is done on a one-by-one basis ([3]). Interactions between two proteins is determined by randomly selecting spots followed by manually scoring for colocalization. Colocalization is observed when two or more of the emission signals overlap in the final image and usually occurs when fluorescently labeled molecules bind to targets that lie in very close or identical spatial positions.

To address these and related questions, a probabilistic modeling is required. Random sets have been used since they were introduced by Matheron in the late sixties. Applications include material science, chemistry, geostatistics, microscopy, medicine or cellular communication networks. Definitions and statistical analysis of these models can be found in([5, 8, 11]. Recently, temporal Boolean models, a particular case of random sets, have been proposed ([2, 9]) to model randomly placed particles with random durations and successfully applied to thresholded image sequences of fluorescent-tagged clathrin proteins in order to count endocytic events and estimate their duration distribution. In this work we consider the gray-level image sequences of two fluorescent-tagged proteins simultane-

ously imaged by TIRFM as a realization of a bivariate fuzzy temporal random closed set. The need of making use of fuzzy set theory arises from the ambiguity within the pixels inherent to gray-level images due to the possible multivalued levels of brightness ([7]).

2 NOTATION AND DEFINITIONS

Our data are two gray-level image sequences, each one corresponding to one type of event (protein). Let us introduce some notations and definitions. A thresholded frame is a (stochastic) binary image which is considered as a realization of a random closed set, denoted as RACS ([5]). A thresholded image sequence is modeled as a temporal random closed set, denoted as TRACS. If two sequences of binary images are considered, we have a bivariate-TRACS. The whole sequence of gray-level images is considered as a realization of a fuzzy temporal random set (denoted as FTRACS). Higher gray-levels are interpreted as higher membership degrees. Basic references about random closed sets are ([11, 5]). An overview over fuzzy random variables can be found in ([4]).

Let us denote by K the compact subsets in R^2, and by $\mathcal{F}$ the closed subsets in R^2, for every $K \in K$, the set $\mathcal{F}_{\mathcal{K}} = \{\mathcal{F} \in \mathcal{F} : \mathcal{F} \cap \mathcal{K} \neq \emptyset\}$ can be defined, σ_f denotes the σ-algebra generated by $\mathcal{F}_{\mathcal{K}}$ for all $K \in K$. A *random closed set*, Φ, is a random variable taking values in $(\mathcal{F}, \sigma_{\{}$). So Φ is an application $\Phi : (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow (\mathcal{F}, \sigma_{\{})$ where $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space. A *bivariate random closed set* is a pair of random closed sets : $\Phi = (\Phi_1, \Phi_2)$. We define a *temporal random set* in $R^2 \times R_+$ as the union of a family of RACS indexed by the observation time $t \geq 0$: $\Phi = \bigcup_{t \geq 0} \Phi(t)$ (see Fig. 1).

A *fuzzy random closed set* is a fuzzy set Φ such that for every $\alpha \in (0, 1]$, the corresponding α-level set $L_\alpha(\Phi)$ is a random closed set. The notation $\Phi(x, t) = (\Phi_1(x, t), \Phi_2(x, t))$ shows the dependency on the time t and the location x of the bivariate FRACS Φ. If we consider different $\alpha's$ for the level-sets of each component of Φ, we can define $L_\alpha(\Phi(x, t)) = (L_{\alpha_1}(\Phi_1(x, t)), L_{\alpha_2}(\Phi_2(x, t)))$, where $\alpha = (\alpha_1, \alpha_2)$. For $i = 1, 2$ and $\alpha \in (0, 1]$, the level-sets $L_\alpha(\Phi_i)$ is defined as $L_\alpha(\Phi_i) = \{(x, t) \in R^2 \times R_+ : \Phi_i(x, t) \geq \alpha\}$. For a given α, this set is composed of the points (x, t) with

gray-level intensity higher than α. Notice it holds that $L_\alpha(\Phi_i) = \bigcup_{t>0}\{x \in R^2 : \Phi_i(x, t) \geq \alpha\} = \bigcup_{t \geq 0} L_\alpha(\Phi_i(\cdot, t)) = \bigcup_{x \in R^2} L_\alpha(\Phi_i(x, \cdot))$. If we are interested in any location at a particular time t, we should study $L_\alpha(\Phi_i(\cdot, t))$. However, if our interest is in the evolution along time of a given pixel x we should consider $L_\alpha(\Phi_i(x, \cdot))$.

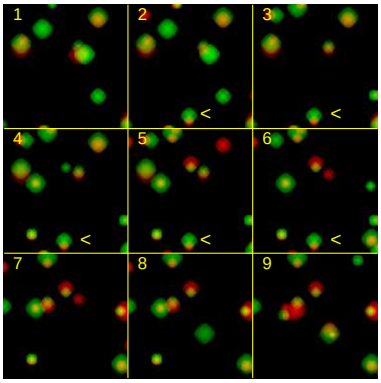


Figure 1: Nine consecutive frames of a simulated FTRACS in which spatial temporal dependencies between events were generated.

Throughout this paper we assume that $\Phi = (\Phi_1, \Phi_2)$ is jointly stationary, i.e. invariant against spatial temporal translations. This implies that Φ_1 and Φ_2 are (marginally) stationary. This assumption is a natural and pragmatic simplification. It is tenable in many real applications and justifies the use of relatively simple nonparametric summaries. Under stationarity the area fractions for a given $\alpha \in (0, 1]$ can be defined as

$$p_i^{(\alpha_i)} = P(\mathbf{0} \in L_{\alpha_i}(\Phi_i(\cdot, 0))), \quad (1)$$

for $i = 1, 2$, where the spatial origin $\mathbf{0}$ denotes an arbitrary location and the temporal origin 0 denotes an arbitrary time. For each $\alpha \in (0, 1]$, $p_i^{(\alpha)}$ gives the probability that an arbitrary point at an arbitrary time belongs to the α-level-set of the i-th component of the random set and can be interpreted as the mean volume covered by the i-th temporal random set per unit area and time at level α.

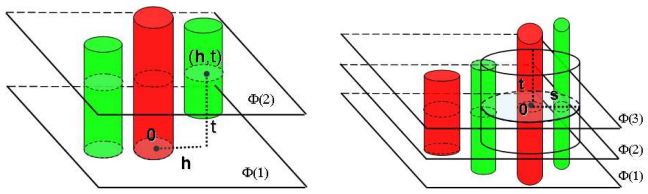
For $\mathbf{h} \in R^2$ and $t \in R_+$, we define the spatial temporal cross-covariance function of $(L_{\alpha_1}(\Phi_1), L_{\alpha_2}(\Phi_2))$ as

$$C_{12}^{(\alpha)}(\mathbf{h}, t) = P(\mathbf{0} \in L_{\alpha_1}(\Phi_1(\cdot, 0)), \mathbf{h} \in L_{\alpha_2}(\Phi_2(\cdot, t))). \quad (2)$$

This is a generalization of the covariance for random sets, details of theoretical aspects can be found in ([8, 10, 11]). This function gives the probability that an arbitrary point observed at an arbitrary time belongs to the first component of the temporal random set and its translation by a vector $\mathbf{h}$ and a time t belongs to the second component (see Fig. 2 (a)).

From the cross-covariance we define the $K_{12}^{(\alpha)}$-function as

The product $p_j^{(\alpha_j)} K_{ij}^{(\alpha)}(s, t)$ is the mean volume covered by $L_{\alpha_j}(\Phi_j)$ within a cylinder with radius s and height $2t$ centered at an arbitrary point of $L_{\alpha_i}(\Phi_i)$, for $i, j = 1, 2$ (see Fig. 2 (b)).



(a) (b)

Figure 2: (a) The spatial temporal cross-covariance and (b) spatial temporal K-function. Red corresponds to type-1 events and green to type-2.

Finally, we define the pair-correlation function as

$$g_{12}^{(\alpha)}(s,t) = \frac{C_{12}^{(\alpha)}(s,t)}{p_1^{(\alpha_1)} p_2^{(\alpha_2)}}. \tag{3}$$

Apart from the functions $K_{12}^{(\alpha)}$ and $g_{12}^{(\alpha)}$, the function $L_{12}^{(\alpha)}$ will be used. This function stabilizes the variance and is defined as

$$L_{12}^{(\alpha)}(s,t) = \sqrt{\frac{K_{12}^{(\alpha)}(s,t)t}{2\pi}}. \tag{4}$$

Under spatial temporal independence between the two types of events, the cross-covariance function at a level $\alpha = (\alpha_1, \alpha_2)$ is $C_{12}^{(\alpha)}(\mathbf{h},t) = P(\mathbf{0} \in L_{\alpha_1}(\Phi_1(\cdot,0)))P(\mathbf{h} \in L_{\alpha_2}(\Phi_2(\cdot,t))) = p_1^{(\alpha_1)} p_2^{(\alpha_2)}$. Therefore, from (3) it holds that $g_{12}^{(\alpha)}(s,t) = 1$, from (**??**) $K_{12}^{(\alpha)}(s,t) = 2\pi s^2 t$ and from (4) $L_{12}^{(\alpha)}(s,t) - st = 0$.

In order to analyze spatial associations, we consider

$$K_{12}^{(\alpha)}(s) = \int_{B(\mathbf{0},s)} \frac{C_{12}(\mathbf{u},0)}{p_1^{(\alpha_1)} p_2^{(\alpha_2)}} d\mathbf{u}, \tag{5}$$

and

$$L_{12}^{(\alpha)}(s) = \sqrt{\frac{K_{12}^{(\alpha)}(s)}{\pi}}. \tag{6}$$

Under spatial independence, $K_{12}^{(\alpha)}(s) = \pi s^2$ and $L_{12}^{(\alpha)}(s) - s = 0$.

3 MATCHING α-LEVEL SETS BETWEEN COMPONENTS

Differences in signal intensities between components, which often appear in real applications due to fluorochrome concentration fluctuations, photobleaching, non equivalent electronic channel settings, etc., suggested us the matching of α-level sets between components, in such a way, that thresholding with matched α-level sets $(\widehat{\alpha}_1, \widehat{\alpha}_2)$ will produce similar binary images. The proposed procedure also allows the selection of thresholds in a unsupervised manner. This procedure is based on the definition of Vorobév Fuzzy Set Average, which was introduced in [1]. Let u be a fuzzy set and let Φ be a random compact defined as: $\Phi = L_\alpha(u)$ where α is a random value in $(0,1]$ with probability density f. The Vorobév average of the fuzzy set u, denoted $A_v u$, is defined as

$$A_v u = L_{\alpha^*}(u), \tag{7}$$

where

$$\nu(L_{\alpha^*}(u)) = \int_{[0,1]} \nu(L_\alpha(u)) f(\alpha) d\alpha. \tag{8}$$

$\nu(\cdot)$ stands for the area. If different α^*-levels verifies the former equation, the minimum is chosen. Intuitively, the Vorobév fuzzy set average is a particular α-cut, i.e. the one whose area equals the mean area of Φ.

We propose a procedure based on this idea to obtain, in an efficient and unsupervised way, the thresholding levels. Let us choose randomly a value α_i within the unit interval $(0,1]$ with probability density function ξ, and consider the area fraction $p_i^{(\alpha_i)} = P(\mathbf{0} \in L_{\alpha_i}(\Phi_i))$, as a random variable. Let μ_i and σ_i^2 denote its mean and standard deviation, i.e. $\mu_i = \int_{(0,1]} p_i^{(\alpha_i)} \xi(\alpha_i) d\alpha_i$ and $\sigma_i^2 = \int_{(0,1]} (p_i^{(\alpha_i)} - \mu_i)^2 \xi(\alpha_i) d\alpha_i$. Then, we define α_i^* as the α-value which verifies:

$$P(0 \in L_{\alpha_i^*}(\Phi_i(\cdot,0))) = \mu_i \tag{9}$$

for $i = 1, 2$. In this way, we have a pair (α_1^*, α_2^*) of thresholds. Note that they represent the same property for each random set. Let us go further and generate a whole set of pairs. We define $\alpha_i^*(\beta)$ as the α-value which verifies:

$$P(0 \in L_{\alpha_i^*(\beta)}(\Phi_i(\cdot,0))) = \mu_i + \beta\sigma_i \tag{10}$$

for $i = 1, 2$. If different $\alpha_i^*(\beta)$-levels verify the former equation, the minimum is taken. Again, for a given β, both values $\alpha_1^*(\beta)$ and $\alpha_2^*(\beta)$ represent the same concept. Finally, the matching function m is defined as

$$m(\alpha_1^*(\beta)) = \alpha_2^*(\beta). \tag{11}$$

We have used a beta distribution with parameters (3,3) for ξ.

4 ESTIMATION AND TESTING FOR INDEPENDENCE

We propose estimators of $p_i^{(\alpha_i)}$, $C_{12}^{(\alpha)}$, and $K_{12}^{(\alpha)}$. Essentially, we replace the (unknown) continuous versions

of these functions with the corresponding (known) discrete versions. We assume that the sampling times are $t = 0, \delta, 2\delta, \ldots, n\delta$, i.e., $n+1$ equally spaced frames, where δ is the temporal delay between two consecutive frames (the acquisition ratio). Our sampling information consists of two sequences of gray-level images $\{(\phi_1(\cdot, k\delta) \cap W, \phi_2(\cdot, k\delta) \cap W)\}_{k=0,\ldots,n}$, where W is the sampling window. The volume fraction is computed as the mean of the area fractions observed at each thresholded frame. Determining critical values for a test of independence for bivariate processes requires maintaining the univariate spatial pattern of each process but breaking any dependence between them. The method of toroidal shifts provides a nonparametric way to test independence. We have implemented a software tool in Matlab for estimating, simulation and testing procedures.

Acknowledgements

This paper has been supported by projects HFSP RGY40/2003 and TIN2006-10134.

References

[1] G.Ayala, T. Leon and V. Zapater. Different Averages of a Fuzzy Set with an Application to Vessel Segmentation. *IEEE Transactions on Fuzzy Systems*, 13, 384-393, 2005.

[2] G. Ayala, R. Sebastian, M.E. Díaz, E. Díaz, R. Zoncu, and D. Toomre. Analysis of spatially and temporally overlapping events with application to image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 1707-1712, 2006.

[3] K. Bellve, Leonard, Standley, Lifshitz, Tuft, Hayakawa, Corvera, and Fogarty. Plasma membrane domains specialized for clathrin-mediated endocytosis in primary cells. *The Journal of Biological Chemistry*, 281(23), 16139-16146, 2006.

[4] M. Gil, M. Lopez-Diaz, and D. Ralescu. Overview on the development of fuzzy random variables. *Fuzzy Sets and Systems*, 157, 2546-2557, 2006.

[5] G. Matheron. *Random sets and Integral Geometry.*Wiley, London, 1975.

[6] M. Metzler, Legendre-Guillemin, Gan, Chopra, Kwok, McPherson, and Hayden. Hip1 functions in clathrin-mediated endocytosis through binding to clathrin and adaptor protein 2. *The Journal of Biological Chemistry*, 276(42), 39271-39276, 2001.

[7] S. Mitra and S. K. Pal. Fuzzy sets in pattern recognition and machine intelligence. *Fuzzy sets and systems*, 156, 381-386, 2005.

[8] I.Molchanov. *Theory of Random Sets (Probability and its Applications)*. Springer, 2005.

[9] R. Sebastian, M. E. Díaz, G. Ayala, E. Díaz, R. Zoncu, and D. Toomre. Studying endocytosis in space and time by means of temporal boolean models. *Pattern Recognition*, 39, 2175-2185, 2006.

[10] J.Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, 1982.

[11] D. Stoyan, Kendall, and Mecke. *Stochastic Geometry and its Applications*. Wiley, Berlin, second edition, 1995.

[12] . D.Toomre, and D. Manstein. Lighting up the cell surface with evanescent wave microscopy. *Trends Cell Biol.*, 11, 298-303, 1991.

BOOLEAN NETWORKS: A STUDY ON MICROARRAY DATA DISCRETIZATION

Cyntia Velarde[1] **Cristina Rubio-Escudero**[2] **Rocío Romero-Zaliz**[3]

[1] Dpto. Computación, FCEyN, Universidad de Buenos Aires, Argentina, cvelarde@dc.uba.ar
[2] Dpto. Lenguajes y Sistemas Informaticos, Universidad de Sevilla, España, crubioescudero@us.es
[3] DECSAI, Universidad de Granada, España, rocio@decsai.ugr.es

Abstract

Biomedical research has been revolutionized by high-throughput techniques and the enormous amount of biological data they are able to generate. Genetic networks arise as an essential task to mine these data since they explain the function of genes in terms of how they influence other genes. Genetic networks based on discrete states, such as boolean networks, have been widely used and have shown abilities to model some of the complex dynamics of gene expression networks. In this work we propose a new method for the discretization of gene expression data based on the fuzzification of already proposed techniques. The proposal is applied to the microarray data obtained from a problem on the inflammation and host response to injury in human beings.

Keywords: Microarray, Boolean network, Discretization process.

1 INTRODUCTION

Microarray technology has revolutionized modern biomedical research by its capacity to monitor the behavior of thousands of genes simultaneously [2]. In particular, time-series expression experiments are an increasingly popular method for studying a wide range of biological phenomena. The reconstruction of genetic networks is becoming an essential task to understand the enormous amount of information generated by this high-throughput technique data [4].

Systems biology research arises at this point as the field to explore the life regulation processes in a cohesive way making use of the new technologies. In regulatory networks, proteins have a main role in the regulation of genes [13], but unfortunately, for the vast majority or biological datasets available, there is no information about the level of protein activity. Therefore, we use the expression level of the genes as an indicator of the activity of proteins they generate. Gene networks represent these gene interactions. A gene network can be described as a set of nodes which usually represent genes, proteins or other biochemical entities and the transcription factors that regulate the system.

Given the complexity of this problem, most approaches work on discretized versions of the gene expression data. In fact, discrete state, time and models have been extensively used to model biological networks, and have shown to be able to model, at least in part, the complex dynamics of these networks.

In this paper, we describe a number of methods for the discretization of gene expression data and we propose a new method based on the fuzzification of the already described methods capable to reduce the rate of misdiscretized data. We apply these methods to the microarray data obtained from a problem on the inflammation and host response to injury in human beings.

2 BACKGROUND

In this section we will introduce the discretization methods used in the experimental section and the inference process developed to extract a boolean network from a set of microarray time-series data.

2.1 DISCRETIZATION PROCESS

Several discretization techniques have been used in expression data analysis. These techniques can be grouped in two high level categories: (1) discretization using expression absolute values, and (2) discretization using expression variations between time points [9].

2.1.1 Notation

Let A' be an n row by m column gene expression matrix, where A'_{ij} represents the expression level of gene i under condition j. The matrix A is defined by its set of rows, I, and its set of columns, J. Moreover, let A'_{IJ} denote the average value in the expression matrix A' and A'_{iJ} and A'_{Ij} denote the mean of row i and condition j, respectively. Let let H_{IJ} denote the maximum (high) value in the expression matrix A and H_{iJ} and H_{Ij} denote the maximum value of row i and column j, respectively. In the same way, let M_{IJ} denote the median value in the expression matrix A and M_{iJ} and M_{Ij} denote the median value of row i and column j, respectively.

In this work, we are concerned with applications that use a discretized version of the matrix, where each element in A' is mapped to one element of an alphabet, Σ. Each different symbol represents a distinct activation level. In the simpler case, Σ may contain only two symbols, one used for *regulation* and other for *non-regulation*. In this case, the expression matrix is usually transformed into a binary matrix, where 1 means regulation and 0 means no regulation. After the discretization process, matrix A' is transformed in matrix A and A_{ij} represents the discretized value of the expression level of gene i under condition j.

We define the *gene expression pattern* for gene i as the discretized expression level of gene i under all conditions $j = 1...m$ in matrix A. Note that for gene a and gene b, the expression levels for conditions $j = 1...m$ in matrix A' might result in the same discretized values for conditions $j = 1...m$ in matrix A (Figure 1). Then we say that gene a and gene b *share the same gene expression pattern*.

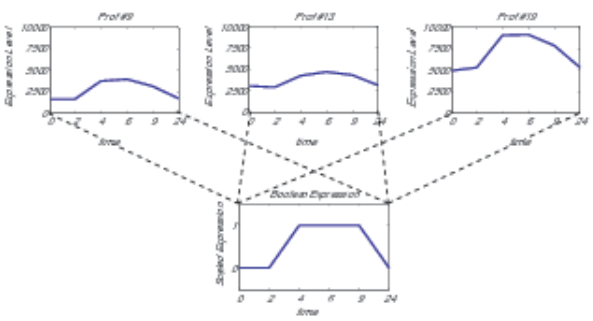


Figure 1: Genes with different expression values at each condition might end up sharing a common expression pattern after the discretization process.

2.1.2 Discretization methods using expression absolute values

Several discretization methods that use expression absolute values [1, 12] will be used in our experiments:

- *A_{IJ} method*: Using the *average* expression value alone, either computed using all the values in the matrix (Equation 1), by row, or by column.

$$A_{ij} = \begin{cases} 1 & if\ A'_{ij} \geq A'_{IJ} \\ 0 & otherwise \end{cases} \quad (1)$$

- *M_{IJ} method*: Using the highest and lowest expression values for each gene i and computing its *median* using all the values in the matrix (Equation 2), by row, or by column.

$$A_{ij} = \begin{cases} 1 & if\ A'_{ij} \geq M_{IJ} \\ 0 & otherwise \end{cases} \quad (2)$$

- *H_{IJ} method*: Using the maximal expression value as the cut-off, observed for the whole matrix, for each gene i or for each condition j. If X is a fixed percentage of this maximal expression value, it can either be computed using all the values in the matrix (Equation 3), by row, or by column.

$$A_{ij} = \begin{cases} 1 & if\ A'_{ij} \geq H_{IJ}(1 - X\%) \\ 0 & otherwise \end{cases} \quad (3)$$

- *$TOP\ X\%$ method*: Using a percentage $X\%$ of the highest values, the expresssion values inside this percentage are discretized to 1 and the remaining to 0. Again, this procedure can be computed using all the values in the matrix (Equation 4), by row , or by column.

$$A_{ij} = \begin{cases} 1 & if\ rank(A'_{ij}, sort(A')) \geq \frac{X \times n \times m}{100} \\ 0 & otherwise \end{cases} \quad (4)$$

where $sort(x)$ returns all elements of x in ascending order, and $rank(x, y)$ returns the position of element x in the list y.

2.1.3 Discretization methods using expression variations between time points

- *Transitional state discrimination (TSD) method* [10]. For this discretization process all values need to be standarized to z-scores (Equation 5):

$$z = \frac{x - \mu}{\sigma} \quad (5)$$

NOT A = A NAND A
A AND B = (A NAND B) NAND (A NAND B)
A OR B = (A NAND A) NAND (B NAND B)

Table 1: Boolean functions obtained only using the $NAND$ function.

where x is a raw score to be standardized, σ is the standard deviation of the population and μ is the mean of the population. The quantity z represents the distance between the raw score and the population mean in units of the standard deviation. If z is negative, then the raw score is below the mean and positive when above.

After this standardization of A' data, each gene expression profile is discretized using two state transitions (Equation 6, $\alpha = 0$).

$$A_{ij} = \begin{cases} 1 & if\ |A'_{ij} - A'_{i(j-1)}| \geq t \\ 0 & otherwise \end{cases} \quad (6)$$

- Another alternative is to first compute the standard deviation for time point 0, $std(0)$ and providing a parameter α to compute a discretization threshold $t = std(0) \times \alpha$ [6] (Equation 6, $\alpha > 0$).

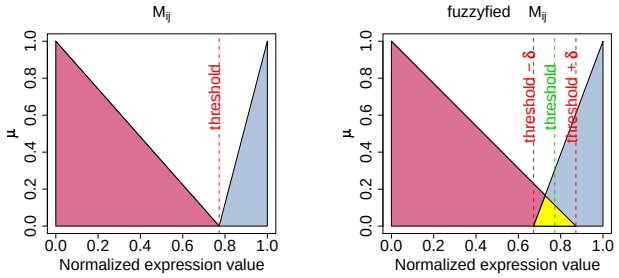
2.2 BOOLEAN NETWORK CONSTRUCTION

A boolean network is composed by a set of nodes n which represent genes, proteins or other biochemical entities. These nodes can take on/off values. The net is determined by a set of at maximum n boolean functions, each of them having the state of k specific nodes as input, where k depends on each node. Therefore, each node has its own boolean function which determines the next state (state at time t_{+1}) based on the actual state (state at time t) of the input nodes. The changes in the net are assumed to occur at discrete time intervals.

The algorithm applied to build the boolean network with our data is the GeneYapay [5]. It performs an exhaustive search of boolean functions over the data, where a number of nodes, less or equal than k, univocally determine the output of some other gene. All possible subsets of $1, 2, ..., k$ elements are visited calculating the number of inconsistencies of the boolean functions in relation to the output value of each gene. The algorithm stops the search for each node when a subset of nodes is found which defines the expression profile. The implementation applied [15] only uses the $NAND$ function since all other boolean function -AND, OR, NOT- can be expressed using $NAND$ (Table 1).

3 PROPOSAL

Crisp discretization methods are widely used in microarray experiments [1, 12, 8, 10]. However, there is a drawback in these methodologies when the expression level is near the threshold value specified by the discretization method. Therefore, several points in the microarray may be misdiscretizied. This is mainly due to the crisp characteristics of the discretization methods (Figure 2 (a)). In this work, we propose an alternative methodology to avoid potential misdiscretization errors by using fuzzy sets (Figure 2 (b)).



(a) (b)

Figure 2: Example of a fuzzyfied discretization method M_{IJ}. (a) Original method. (b) Fuzzified version, where δ is a given small value.

The basic idea is to extend the crisp set (Figure 2 (a)) to a fuzzy set by overlapping both crisp sets using a small value δ as a deviation from the calculated threshold (e.g., $threshold \pm \delta$) (Figure 2 (b)). The data in the intersection of both fuzzy sets (i.e., zero set $ZERO$ and one set ONE) can be discretize to 0 or to 1 as they have almost the same membership μ to both fuzzy sets (Figure 2 (b)). We suggest that expression values with $\mu(ZERO) > 0$ and $\mu(ONE) > 0$ to both fuzzy sets are to be discretized to 0 or 1 depending on their correlation with the elements belonging to the 0 or 1 set. That is, using the Pearson correlation coefficient [11] we calculate if a given expression value is more likely to be similar to the elements in the 0 set (i.e., $\mu(ZERO) > 0$ and $\mu(ONE) = 0$) or to the 1 set (i.e., $\mu(ONE) > 0$ and $\mu(ZERO) = 0$), an discretize it accordingly.

Then, we can redefine the discretization methods introduced in Section 2.1 using this new methodology. We will notate them with an extra tilde (e.g., M'_{IJ}, TSD') to differentiate them from the original versions.

4 EXPERIMENTS

Experiments using all the discretization functions described in Section 2.1 and the new proposed versions

were applied to an immuno-inflammatory response problem. This study, in part carried out at the Cellular Injury and Adaptation Laboratory, Washington University School of Medicine, is a piece of a large-scale research project devoted to profile leukocyte gene expression and plasma proteins of burn and trauma patients [3]. The host response to trauma and burns is a collection of biological and pathological processes that depends critically upon the regulation of the human immuno-inflammatory response. The objective of this study is to identify significant relationships that regulate the integration of the human complex biological system. For this purpose, 48 GeneChips®HG-U133A v2.0 from Affymetrix Inc. were analyzed, derived from samples taken from human blood of eight patients: four treated with intravenous endotoxin and four with a placebo and expression retrieved over time at hours 0, 2, 4, 6, 9 and 24.

The microarray data has been normalized by scaling (Invariant Set Approach) and normalizing (Model-Based Expression Index) [7].

4.1 Results for the original versions

Several interesting results were obtained after discretizing the microarray information with the different methods (Tables 2 and 3). We analyze patient 1 as an example of a treatment patient and patient 5 as an example of a control patient. Also, we analyze the results obtained by the all patient expression patterns.

We can extract from Table 3 that all A_{IJ}, M_{IJ} and H_{IJ} obtain the same amount of patterns. M_{IJ} and H_{IJ} discretization methods have a very similar behavior since they discretize almost all genes in the same fashion since the threshold for each method is almost the same (Table 3). On the other hand, the comparison between the other methods (Table 2) show a high dissimilarity in the discretized patterns. The union of all patients shows similar results than the patients taken independently (Table 2).

Thresholds calculated as a result of the different discretization process are very close to each other (Table 3). Nevertheless, a small difference in the threshold value can affect the discretization of many genes.

The results obtained by the $TOP\ X\%$ discretization methods (Table 3) produce different patterns, nevertheless most of the genes are clustered in two major groups for all $X = 10, 20, 30$, therefore providing a very poor discretization.

Analyzing patterns I and II from Table 4, we can see that $TSD\ \alpha = 0, 1$ discretize them equally, while $TSD\ \alpha = 2$ discretize them differently. We can observe in Figure 3 that their expression levels are not so different. Something similar happens with patterns III and IV. An interesting result is obtained by the discretization of patterns V and VI. We can observe in Figure 3, that they are almost opposite patterns, but $TSD\ \alpha = 1, 2$ discretize them equally, while $TSD\ \alpha = 0$ discretize them correctly.

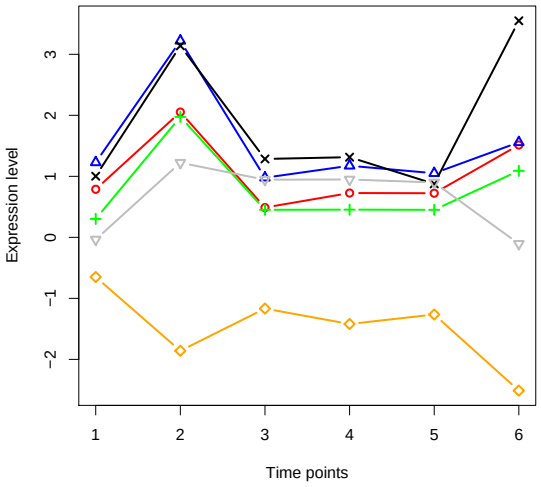


Figure 3: Expression levels for the patterns in Table 4. Color-codes: expression pattern I (red, circle), expression pattern II (blue, triangle), expression pattern III (green, plus sign), expression pattern IV (black, cross), expression pattern V (orange, diamond) and expression pattern VI (gray, upside-down triangle).

Our experiments reveal that the studied discretization methods are sensible to the threshold used for each of them, therefore, data located near this threshold would have high probability of misdiscretization.

4.2 Results for the new versions

We applied the modified versions of the discretization methods introduced in Section 2.1. Due to space restrictions, we will only show the results obtained for method M'_{IJ} for patient 1 as an example. We decided to use a small threshold ($\delta = 0.01$) since greater values will force to unecessarily re-analyze a great amount of discretizations.

The correlation between values inside the threshold and a cluster prototype for each group (i.e., zero set $ZERO$ and one set ONE) are calculated. The discretization of these values is determined by the best correlated group. Each cluster prototype is selected between patters with expression level values outside of the threshold.

Results show that correlation values of the new discretization method is much better (Table 5). We can also see in Figure 4 that the expression levels are more

Table 2: Comparison between discretization methods A_{IJ} (Equation 1), M_{IJ} (Equation 2) and H_{IJ} (Equation 3) described in Section 2.1.2. Each cell shows the number of genes, and its percentage of the total dataset, discretized differently by each pair of methods. All methods applied to the eight patients, here we show only patient 1 (treatment) and patient 5 (control).

METHODS	TREATMENT PATIENT	CONTROL PATIENT	ALL PATIENTS
A_{IJ} vs. M_{IJ}	13752 (62%)	9650 (43%)	78725 (50%)
A_{IJ} vs. H_{IJ}	13769 (62%)	9677 (43.5%)	78878 (50.7%)
M_{IJ} vs. H_{IJ}	106 (0.47%)	126 (0.56%)	826 (0.53%)

Table 3: Discretization results. Discretization methods A_{IJ} (Equation 1), M_{IJ} (Equation 2) and H_{IJ} (Equation 3) described in Section 2.1.2. Discretization method $TOP\ X\%$ (Equation 4) described in Section 2.1.2 with $X = 10, 20, 30$. Discretization method TSD (Equation 6) described in Section 2.1.3 with $\alpha = 0, 1, 2$. All methods applied to the eight patients, here we show only patient 1 (treatment) and patient 5 (control).

	#PATTERNS		
METHOD	TREATMENT	CONTROL	THRESHOLD
A_{IJ}	64	64	0.22242430
M_{IJ}	64	64	0.21305742
H_{IJ}	64	64	0.21303658
TOP 10%	63	64	0.22853785
TOP 20%	55	54	0.23952316
TOP 30%	47	43	0.24580431
$TSD\ \alpha = 0$	64	64	-
$TSD\ \alpha = 1$	34	31	-
$TSD\ \alpha = 2$	29	13	-

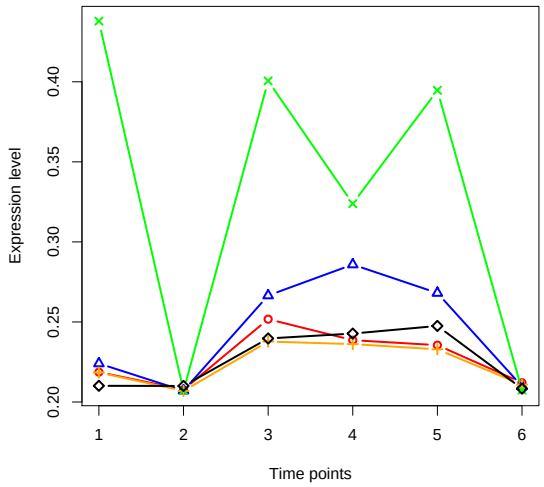


Figure 4: Expression levels for the patterns in Table 5 against M_{IJ} and M'_{IJ} prototypes. Color-codes: expression pattern I (red, circle), expression pattern II (blue, triangle), expression pattern III (orange, plus sign), M_{IJ} prototype (green, cross) and M'_{IJ} prototype (black, diamond).

similar to the new method's prototype than the original method's prototype. Thus, the discretization values will be more appropriate.

As an example of the results obtained by the new methodology proposed, Table 6 shows some genes whose M_{IJ} discretization is corrected by the new M'_{IJ}.

We then calculate the boolean networks for the original M_{IJ} (Figure 5) and the modified M'_{IJ} method (Figure 6). The networks obtained have some portions in common, but also some other protions are quite different. Due to the discretization changes between M_{IJ} and M'_{IJ}, several genes belonging to a pattern in the original method, may now belong to a different pattern in the new method. We studied the biological annotation of every gene inside each of these patterns for both discretization methods using the Onto-CC software [14]. We could see that the new discretization method M'_{IJ} produces more homogeneous sets, that is, the genes in each set have a very simmilar annotation. For instance, pattern G33 (`100001`) have 360 genes in the M_{IJ} method, while the same pattern share the same information plus five extra genes. These five genes are annotated with several terms, like "ubiquitin cycle" and "protein targeting" wich are also associated with the other 360 genes. Therefore, the boolean networks calculated using the new methodology are more reliable since they are constructed using information with

Table 4: Example of genes discretized differently by $TSD\ \alpha = 0$, $TSD\ \alpha = 1$ and $TSD\ \alpha = 2$.

ID	TSD $\alpha = 0$	TSD $\alpha = 1$	TSD $\alpha = 2$	**EXPRESSION VALUES**					
I	010101	011000	000000	0.788009	2.053057	0.488498	0.726193	0.723284	1.513306
II	010101	011000	001000	1.228279	3.221180	0.977528	1.173404	1.050649	1.557317
III	010101	011000	001000	0.302274	1.974590	0.451306	0.454151	0.450831	1.090290
IV	010101	011001	010001	1.002203	3.13927	1.284922	1.314039	0.876443	3.548847
V	101010	010001	000000	−0.649956	−1.859078	−1.167513	−1.42009	−1.26441	−2.511316
VI	110100	010001	000000	−0.035346	1.223311	0.944796	0.948552	0.899278	−0.105929

Table 5: Correlation between genes to M_{IJ} and M'_{IJ}.

ID	**EXPRESSION VALUES**						M_{IJ}	M'_{IJ}
I	0.2186793	0.2076971	0.2517480	0.2386211	0.2355112	0.2120486	0.63664744	0.88130940
II	0.2240480	0.2070605	0.2665070	0.285858p	0.2680730	0.2097774	0.55217325	0.95665902
III	0.2184243	0.2069158	0.2377598	0.2361701	0.2328981	0.2108157	0.67215942	0.93229356

less missclassification errors.

5 CONCLUSIONS

Different discretization methods produce significant differences in the analysis of microarray experiments. Thus, selecting an appropriate method will directly affect the quality of the results. Classical discretization methods show major differences in their results when applied to the same dataset. Moreover, expression levels near the threshold value in each method are questionably discretized and produce potential misdiscretizations. We have proposed new fuzzy based methods that uses correlation to calculate the appropriate discretization values for these dubious expression levels. Thus, the boolean networks are calculated using discretized expression values with a smaller missclassification errors. Further studies with different datasets will provide more detailed analysis.

References

[1] C. Becquet, S. Blachon, B. Jeudy, J-F. Boulicaut, and O. Gandrillon. Strong-association-rule mining for large-scale gene-expression data analysis: a case study on human sage data. *Genome Biology*, 3(12):Electronic publication, 2002.

[2] P. Brown and D. Botstein. Exploring the new world of the genome with dna microarrays. *Nature Genetics*, 21:33–37, 1999.

[3] S.E. Calvano et. al. A network-based analysis of systemic inflammation in humans. *Nature*, 437(7061):1032–7, 2005.

[4] G.W. Carter. Inferring network interactions within a cell. *Bioinformatics*, 6(4):380–389, 2005.

[5] D. D'Onia, L. Tam, J.P. Cobb, and I. Zwir. A hierarchical reverse-forward methodology for learning complex genetic networks. In *Proceedings of the 3rd International Conference on Systems Biology (ICSB), Stockholm, Sweden*, 2003.

[6] S. Erdal, O. Ozturk, D. Armbruster, H. Ferhatosmanoglu, and W. C. Ray. A time series analysis of microarray data. In *Proceeding of the 4rd IEEE Symposium on Bioinformatics and Bioengineering*, pages 366–374, 2004.

[7] C. Li and W. H. Wong. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA*, 98(1):31–6, 2001.

[8] S. Lonardi, W. Szpankowski, and Q. Yang. Finding biclusters by random projections. In *Proceedings of the 15th Annual Symposium on Combinatorial Pattern Matching*, pages 102–116. Springer, 2004.

[9] S.C. Madeira and A.L. Oliveira. An evaluation of discretization methods for non-supervised analysis of time-series gene expression data. Technical Report 42, INESC-ID, 2005.

[10] C. Mollr-Levet, S. Cho, and O. Wolkenhauer. Microarray data clustering based on temporal variation: Fcv and tsd preclustering. *Applied Bioinformatics*, 2(1):35–45, 2003.

[11] K. Pearson. Mathematical contributions to the theory of evolution. iii. regression, heredity and

Table 6: Example of genes discretized differently by M_{IJ} and M'_{IJ}.

M_{IJ}	M'_{IJ}	**EXPRESSION VALUES**					
111111	101111	0.273675	0.225677	0.246043	0.250797	0.254966	0.248701
001111	001110	0.212099	0.211830	0.245788	0.245442	0.254923	0.214511
001000	000000	0.212132	0.209505	0.214234	0.212397	0.211857	0.209414

panmixia. *Philos. Trans. Royal Soc. London Ser. A*, 187:253–318, 1896.

[12] R.G. Pensa, C. Leschi, J. Besson, and J. Boulicaut. Assessment of discretization techniques for relevant pattern discovery from gene expression data. In *4th Workshop on Data Mining in Bioinformatics*, 2004.

[13] J. Rice and G. Stolovitzky. Making the most of it: Pathway reconstruction and integrative simulation using the data at hand. *Biosilico*, 2(2):70–77, 2004, 2(2), 70.

[14] R. Romero-Zaliz, C. del Val, J.P. Cobb, and I. Zwir. Onto-cc: a web server for identifying gene ontology conceptual clusters. *Nucleic Acids Res*, 2008.

[15] C. Velarde. Una metodologia para la construccion de redes booleanas basadas en microarrays. Master's thesis, University of Buenos Aires, 2008. In preparation.

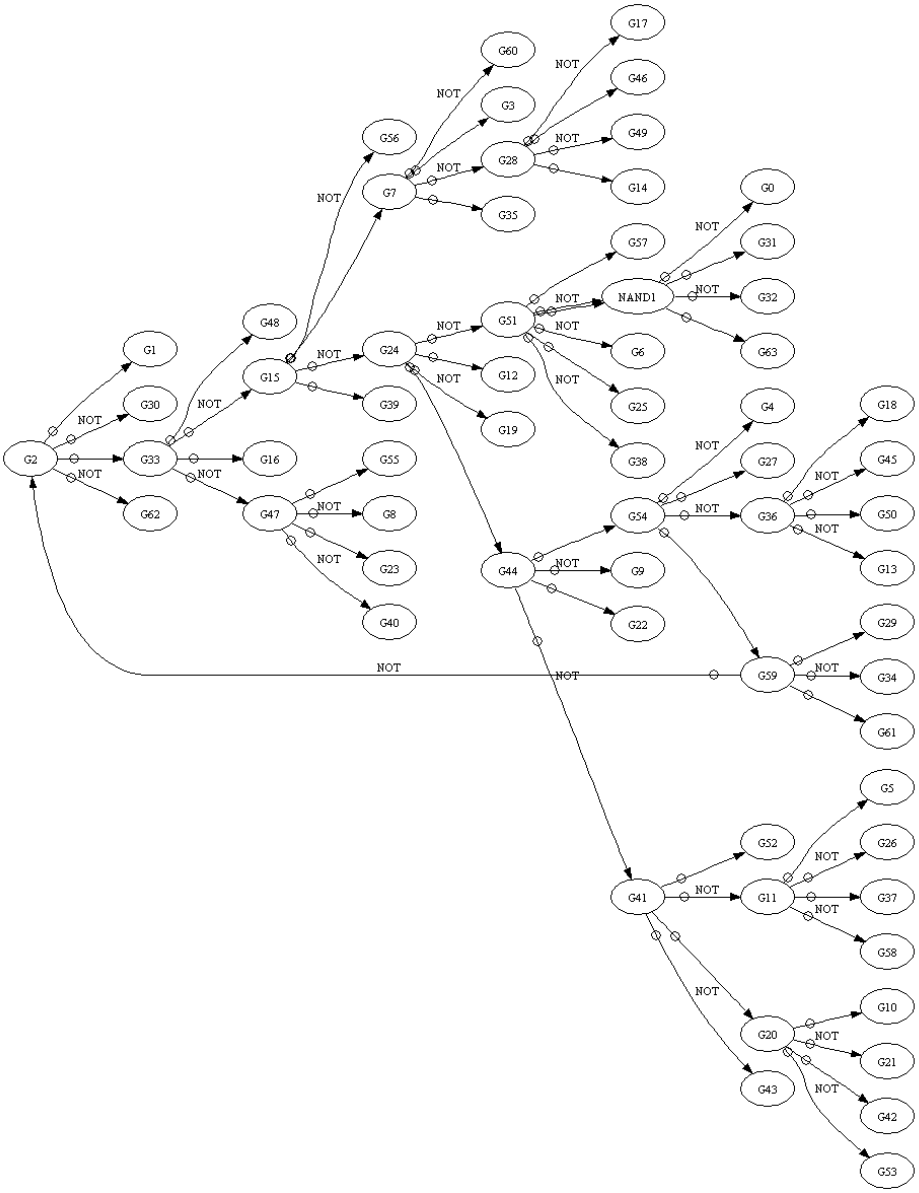


Figure 5: Boolean network obtained for patient 1 using M_{IJ} discretization method. Each node correspond to a an expression pattern.

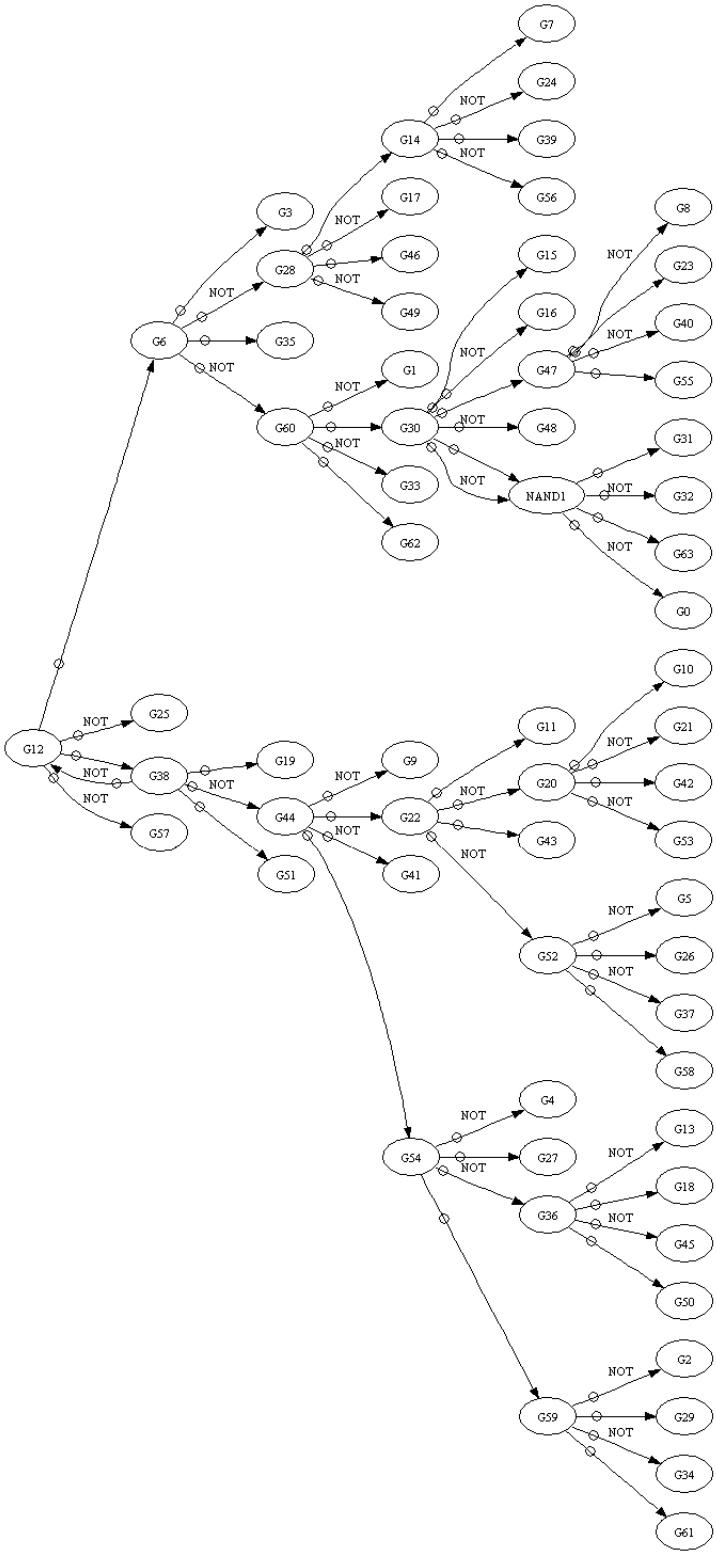


Figure 6: Boolean network obtained for patient 1 using M'_{IJ} discretization method. Each node correspond to a an expression pattern.

APLICACIÓN DE REGLAS DE ASOCIACIÓN DIFUSAS EN PROTEÓMICA

F. Javier López Domingo[1] **Armando Blanco Morón**[1] **Fernando García Alcalde**[1]
Sergio Blanco Manrique[2]

[1] Dpto. de Ciencias de la Computación e IA, Universidad de Granada, {fjavier,armando,fernan}@decsai.ugr.es
[2] Dpto. IESIA, Universidad de Huelva, blanco.sergio@gmail.com

Resumen

La identificación del conjunto de interacciones proteína-proteína es necesaria para entender los comportamiento celulares. Un gran número de trabajos en la literatura estudian este tipo de interacciones, habiendo relacionado algunos de ellos los perfiles de expresión génicos y las interacciones proteína-proteína. Los datos biológicos tienden a ser imprecisos y ruidosos. Sin embargo, la mayoría de las técnicas empleadas hasta el momento para su tratamiento son técnicas crisp clásicas. Las técnicas difusas han demostrado un buen comportamiento ante este tipo de información imprecisa. Así, en este trabajo presentamos una primera aproximación al modelado de las interacciones prot-prot mediante reglas de asociación difusas, niveles de expresión génica e información acerca de la localización celular.

Palabras Clave: Interacciones proteína-proteína, reglas de asociación difusas, microarrays, Gene Ontology.

1 INTRODUCCIÓN

La secuenciación completa de los genomas de diversas especies (entre ellos el genoma humano), así como el desarrollo de nuevas tecnologías como los microarrays, han cambiando nuestra forma de plantear problemas y extraer conclusiones, ofreciendo una imagen compleja y completa de los genomas. Sin embargo, se han producido gran cantidad de datos cuyo análisis e interpretación no resulta trivial, siendo así imprescindible uso de técnicas computacionales para la extracción de información de los resultados [15].

Actualmente, el estudio de las redes de interacciones proteicas es considerado de vital importancia [3]. Las proteínas generalmente interactúan unas con otras y forman largos ensamblajes en un tiempo y espacio dado. Al mismo tiempo los complejos proteicos proporcionan una diversidad funcional que es codificada por la combinación contextual de productos génicos. Por ejemplo, es conocido que cada proceso central de la célula está catalizado por un conjunto de diez o más proteínas unidas entre sí y altamente coordinadas. De esta forma, las interacciones proteína-proteína constituyen la base de los complejos que llevan a cabo los diferentes procesos celulares. Por tanto, la identificación del conjunto de interacciones proteína-proteína en un organismo es necesaria para descifrar los mecanismos moleculares subyacentes a los diferentes procesos biológicos y para descubrir nuevas funciones de proteínas basándose en las proteínas con las que interaccionan. Sin embargo, incluso para organismos modelo como por ejemplo la levadura, la mayoría de las interacciones proteína-proteína no se han descubierto aún [17].

El uso de diversas técnicas experimentales a gran escala (*high throughput techniques*) ha dado lugar a la identificación de grandes conjuntos de interacciones proteína-proteína para diversas especies. Sin embargo, la utilidad de estos conjuntos se ve mermada por la gran cantidad de falsos positivos que presentan [3]. Un gran número de trabajos en la literatura han estudiado este tipo de interacciones, habiendo relacionado algunos de ellos los perfiles de expresión génicos y las interacciones proteína-proteína [9, 11, 3]. Las técnicas computacionales, y en particular las reglas de asociación, se han aplicado también tanto al estudio como a la predicción de interacciones proteína-proteína [16, 2, 8].

Es bien conocido que la información biológica tiende a ser imprecisa y a contener gran cantidad de ruido. Aún así, como norma general se suele hacer uso de técnicas crisp clásicas, como las mencionadas ante-

riormente para el análisis de datos biológicos, mientras que otras metodologías que han demostrado ser más apropiadas para el tratamiento de este tipo de información, como por ejemplo las metodologías difusas, se usan raramente.

En este trabajo presentamos un nuevo enfoque para el modelado de las interacciones proteína-proteína mediante reglas de asociación difusas, niveles de expresión génica y localizaciones celulares de la *Gene Ontology* (GO) [19]. De este modo, proponemos la obtención de reglas de asociación difusas de la forma $GenA[+] \rightarrow GenB[+]$, $GenA[-] \rightarrow GenB[-]$ (+ y - indican sobre y sub-expresion respectivamente), para capturar las asociaciones existentes entre los niveles de expresión de los genes. Una vez obtenida la base de reglas, ésta se comparará con las interacciones proteína-proteína más fiables que hay descritas en la literatura, teniendo en cuenta además la información de la localización celular de las proteínas proporcionada por GO.

La teoría de conjuntos difusos es especialmente apropiada para el modelado de datos imprecisos y ruidosos (como por ejemplo los datos de microarrays), mientras que las reglas de asociación son de gran utilidad para estudiar relaciones entre un gran número de variables (en este caso los perfiles de expresión de cada gen), por tanto las reglas de asociación difusas son una metodología apropiada para nuestro propósito, que de forma general consiste en el estudio de las interacciones proteína-proteína a partir de los niveles de expresión.

En 1993, Agrawal propuso un algoritmo para la extracción de reglas de asociación de grandes bases de datos [1]. Desde entonces se les ha prestado una especial atención a este tipo de técnicas, y hoy en día siguen siendo en el campo del Knowledge Discovery in Databases (KDD), una de las principales técnicas para la búsqueda de patrones en grandes conjuntos de datos.

Dada una base de datos de transacciones D, donde cada transacción es un conjunto de pares atributo-valor o *items*, el objetivo de estas técnicas es encontrar un conjunto de expresiones de la forma $X \rightarrow Y$, donde X e Y son pares de conjuntos atributo-valor o *itemsets*. A esta expresión se la llama *regla de asociación* e indica que cuando se da X es también probable que se dé Y. A esta probabilidad se la denomina *confianza* de la regla, y se define como el porcentaje de transacciones que contienen X e Y respecto del total de transacciones que contienen X. Al porcentaje de transacciones de la base de datos que contienen tanto X como Y se le denomina *soporte* de la regla.

La teoría de conjuntos difusos fue propuesta por Zadeh en 1965 para modelar matemáticamente la imprecisión inherente a ciertos conceptos [18]. En este caso en particular el algoritmo de extracción de reglas de asociación necesita tratar conceptos como la "SUB-EXPRESIÓN" o la "SOBRE-EXPRESIÓN" de los genes. Así pues, hay que "*fuzzificar*" el dominio de los niveles de expresión definiendo sobre él los conjuntos difusos que representarán estas etiquetas lingüísticas. Por lo tanto, las reglas de asociación difusas son también expresiones de la forma $X \rightarrow Y$, pero en este caso X e Y son conjuntos de pares atributo-valor difusos. Valores de confianza y soporte difusos medirán la fiabilidad de las reglas y además, para evitar ciertos inconvenientes que presenta el uso de estas dos medidas de fiabilidad, se utilizarán también los *Factores de Certeza* [4].

El conjunto de datos utilizado en este trabajo se compone de información del genoma de la levadura (*Saccharomyces cerevisae*), ya que la gran cantidad de experimentación llevada a cabo sobre este organismo modelo ha proporcionado conjuntos de datos de gran calidad así como una extensa bibliografía acerca del mismo. La mayoría de la información acerca de la levadura se encuentra almacenada en bases de datos como la *Saccharomyces Genome Database* (SGD [7, 21]), la *Comprehensive Yeast Genome Database* (CYGD [10, 20]) y otras.

El resto del trabajo se organiza como sigue: la sección 2.1 describe el conjunto de datos utilizado en los experimentos. En la sección 2.1 se comenta el algoritmo seguido para la extracción de las reglas de asociación difusas. A continuación, en el apartado 2.3 se describen los pasos seguidos para modelar las interacciones proteína-proteína a partir de las reglas de asociación difusas y la localización celular GO. Finalmente, en la sección 3 se muestran los resultados obtenidos y en el apartado 4 se proporcionan ciertas conclusiones y trabajos futuros.

2 MÉTODOS

2.1 CONJUNTO DE DATOS

2.1.1 Microarrays

Los microarrays son dispositivos miniaturizados que permiten conocer el nivel de expresión de miles de genes bajo una condición experimental determinada. Esta tecnología se ha convertido en la fuente de una enorme cantidad de datos de expresión génica organizados en las llamadas *matrices de expresión*.

Para este trabajo se han recopilado 27 conjuntos de datos de microarrays cada uno de ellos obtenido de trabajos ya publicados acerca del genoma de la levadura y representando un total de 746 condiciones experimentales diferentes, relativas a diferentes estados celulares

tras daar el ADN, diferentes instantes durante el ciclo celular etc[1]. Todos los conjuntos de datos descargados se encontraban ya preprocesados y contienen valores de expresión relativos transformados mediante un log_2. Todos estos datos son accesibles públicamente a través del servidor FTP de la SGD [2]. Los 27 conjuntos se han integrado en una única matriz de expresión formada por 746 filas (una por condición experimental) y 7075 columnas (una por gen). Es importante también destacar que los llamados *valores perdidos* se mantienen como tales sin necesidad de inferir para ellos ningún valor, representando esto una ventaja de las reglas de asociación respecto a otras técnicas de minería de datos como el clústering o el biclústering en las que es imprescindible que todos los genes tengan un valor de expresión para todas las condiciones [12, 14]. Incluyendo esta gran variedad de condiciones experimentales se pretende que en la matriz de expresión utilizada queden reflejados la mayor cantidad posible de estados de la célula, de forma que las asociaciones obtenidas mediante las reglas sean lo más fiables posible.

La definición de los conjuntos difusos SOBRE-EXPRESADO, SUB-EXPRESADO y NO-MODIFICADO se ha llevado a cabo siguiendo el consejo experto tal y como muestra la figura 1

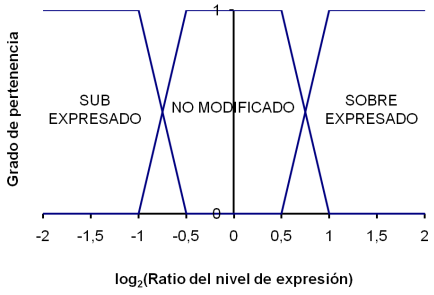


Figura 1: Etiquetas lingüísticas definidas para los niveles de expresión

2.1.2 Gene Ontology

La Gene Ontology (GO) surgió para satisfacer la necesidad de tener descripciones consistentes de los genes. Está formada por tres vocabularios estructurados (ontologías) que permiten describir los genes en función de su localización en la célula (Cellular component), su función molecular (Molecular Function) o el proceso biológico (Biological process) en el que participan.

En este trabajo se ha hecho uso de una versión reducida del GO llamada *GO Slim* (versión del 25 Marzo de 2008). Se trata de un subconjunto del GO que pretende dar una visión a más alto nivel del mismo, conteniendo sólo los principales términos de las tres ontologías. Además, sólo se incluyen en el GO Slim las anotaciones más fiables del GO.

2.1.3 Interacciones proteína-proteína

En este punto se han descargado dos conjuntos diferenciados de interacciones proteína-proteína. El primero de ellos está formado por las 9074 interacciones de alta fiabilidad descritas por Sean R. Collins et al. [3].

Ronald Jansen et al. llevan a cabo un estudio de un conjunto de interacciones de proteinas en el que distinguen entre interacciones *permanentes* y *transitorias* [11]. Definen las interacciones *permanentes* como aquellas que se mantienen bajo la mayoría de las condiciones celulares, y llegan a la conclusión de que existe mayor correlación entre los niveles de expresión de genes cuyas proteínas pertenecen a complejos *permanentes*, que entre perfiles de expresión de genes cuyas proteínas pertenecen a complejos *transitorios*. Así, el segundo conjunto de interacciones proteína-proteína se obtiene a partir de los complejos proteicos definidos como *permanentes* en [11]. Haciendo uso del catálogo de complejos proteicos disponible en la CYGD, se han descargado las proteínas que componen los complejos que se muestran en la Tabla 1. Para cada complejo se combinan entre sí las proteínas que lo forman y así se obtiene el segundo conjunto con 13570 interacciones proteína-proteína (el número de interacciones definidas para cada complejo se muestra también en la Tabla 1).

2.2 EXTRACCIÓN DE REGLAS DE ASOCIACIÓN

El algoritmo utilizado para la obtención de las reglas de asociación ha sido una modificación difusa del algoritmo Apriori [1] tal y como se describe en [4].

Para obtener la base de reglas el umbral de soporte se fijó en 0.01 mientras que el umbral para la confianza y el factor de certeza se estableció en 0.1. Estos valores se establecieron basándose en estudios previos del grupo en la aplicación de reglas de asociación difusas sobre datos biológicos [13].

[1] Una descripción más detallada de las condiciones experimentales se puede obtener en ftp://genome-ftp.stanford.edu/pub/yeast/systematic_results/expression_data/expression_connection_data/README.gz

[2] ftp://genome-ftp.stanford.edu/pub/yeast/systematic_results/expression_data/expression_connection_data

Tabla 1: Complejos permamentes utilizados para definir interacciones proteína-proteína

COMPLEJO	NUM. PROT.	NUM. INTER.
cytoplasmic ribosomes	138	9453
cytoplasmic ribosomal small subunit	57	1596
cytoplasmic ribosomal large subunit	81	3240
mitochondrial ribosomes	76	2850
mitochondrial ribosomal small subunit	31	465
mitochondrial ribosomal large subunit	44	946
20S proteasome	15	105
cytochrome bc1 complex	10	45
F0/F1 ATP synthase	18	153
succinate dehydrogenase complex	4	6
cytochrome-c oxidase	11	55
RNA polymerase I	23	169
RNA polymerase II	13	78
RNA polymerase III	13	78

2.3 REGLAS, NIVELES DE EXPRESIÓN Y LOCALIZACIÓN CELULAR

Tal y como se ha comentado anteriormente, diversos trabajos han establecido previamente la relación existente entre que dos genes tengan perfiles de expresión parecidos y que dichos genes interaccionen para formar un complejo proteico [3, 9, 11]. Dos proteínas no pueden interaccionar si no se encuentran en el mismo espacio celular, por lo tanto, en este trabajo proponemos un nuevo enfoque basado en reglas de asociación difusas y en la localización celular GO para cuantificar la relación existente entre que dos genes presenten perfiles de expresión parecidos y que sus proteínas interaccionen entre sí.

De este modo, el primer paso consiste en aplicar el algoritmo de extracción de reglas de asociación sobre la matriz de expresión descrita en el apartado 2.1.1. Una vez obtenida la base de reglas difusas diremos que podría haber una interacción potencial entre dos proteínas A y B si ambas comparten alguna localización celular GO y aparecen en la base de reglas las asociaciones que se muestran en la Tabla 2

Obsérvese en este punto que la coexpresión se ha definido de 4 formas diferentes en función de las asociaciones que aparezcan en la base de reglas relacionando una pareja dada de genes. Unas de estas definiciones son más restrictivas y otras menos para apreciar también la influencia de la aparición de más o menos asociaciones entre los niveles de expresión de una pareja de genes.

Tabla 2: Distintas formas de definir la coexpresión

TIPO	ASOCIACIONES
0	$GenA[+] \rightarrow GenB[+]$ $GenA[-] \rightarrow GenB[-]$
1	$GenA[+] \rightarrow GenB[+]$ $GenA[-] \rightarrow GenB[-]$ $GenA[noModif] \rightarrow GenB[noModif]$
2	$GenA[+] \rightarrow GenB[+]$ $GenA[-] \rightarrow GenB[-]$ $GenB[+] \rightarrow GenA[+]$ $GenB[-] \rightarrow GenA[-]$
3	$GenA[+] \rightarrow GenB[+]$ $GenA[-] \rightarrow GenB[-]$ $GenB[+] \rightarrow GenA[+]$ $GenB[-] \rightarrow GenA[-]$ $GenA[noModif] \rightarrow GenB[noModif]$ $GenB[noModif] \rightarrow GenA[noModif]$

Finalmente, una vez seleccionados los casos de la base de reglas que satisfacen las dos condiciones anteriores, se compararán con los diferentes conjuntos de interacciones proteína-proteína descritos en 2.1.3 y se analizarán las co-ocurrencias.

3 RESULTADOS

Las coincidencias de los conjuntos de interacciones proteína-proteína con las ocurrencias en la base de reglas que cumplen las condiciones del apartado 2.3, se muestran de forma detallada en la Tabla 3. El número total de pares de proteínas co-expresados y co-localizados se muestra en la Tabla 4. Tal y como se puede observar, las co-ocurrencias de co-expresión y co-localización celular capturan casi el 80% de las interacciones proteína-proteína definidas a partir de los complejos proteicos permanentes. Esto era esperado, ya que tal y como comentan Jansen et al. [3] los genes cuyas proteínas que pertenecen a complejos proteicos permanentes presentan, en general, una alta correlación entre sus perfiles de expresión. El número de interacciones capturadas del conjunto de Collins et al. es bastante menor, seguramente porque en este caso no hay distinción entre interacciones permanentes y transitorias. El hecho de que haya ciertas interacciones de los complejos permanentes que no quedan capturadas puede atribuirse en nuestra opinión a la gran cantidad de ruido e imprecisión que afecta a los datos de micorarrays, así como a imprecisiones en las anota-

ciones GO. Asimismo, no se debe olvidar que a pesar de la gran cantidad de condiciones experimentales introducidas, no se pueden captar a la perfección todos lo posibles estados de la célula. En cualquier caso, y dado que las interacciones de algunos complejos se capturan mejor que otras, un estudio más profundo de estos casos es necesario para extraer más conclusiones.

Por otra parte, aparecen bastantes pares proteína-proteína co-expresados y co-localizados que no tienen correspondencia en los conjuntos de interacciones. Esto era también esperado, dado que el hecho de que en un espacio de la célula aparezca cantidad de diferentes proteínas no necesariamente implica que todas tengan que interaccionar con todas. Para que dos proteínas interaccionen sus estructuras tienen que poder unirse. Así, el sistema más común de interacción entre dos proteínas consiste en un acoplamiento preciso entre dos superficies rígidas, una de cada proteína. Estas interacciones pueden ser muy específicas, permitiendo que una proteína seleccione otra proteína determinada de entre los muchos miles de proteínas que presenta una célula eucariota [17]. En este sentido, la introducción en la metodología de información de las estructuras de las proteínas y de sus dominios es imprescindible. En cualquier caso, es necesario que haya cantidad de proteína suficiente para que puedan interaccionar, y dada la ingente cantidad de combinaciones posibles proteína-proteína (el genoma de la levadura contiene más de 7000 genes), y la gran especificidad de las interacciones, en nuestra opinión esta primera aproximación usando la metodología difusa proporciona un primer filtrado de todas esas combinaciones bastante significativo. Además, hay que tener en cuenta que tal y como se comentó en la introducción, la mayoría de las interacciones entre proteínas incluso en organismos modelo como la levadura no se han descubierto aún.

Además de todo lo anterior, no se aprecian grandes diferencias en las coincidencias según el tipo de definición que se haga de la co-expresión. Sin embargo, sí que se aprecian grandes diferencias en el número de pares de proteínas co-expresadas y co-localizadas según el tipo de definición que se haga. Así, parece más adecuado utilizar definiciones en las que aparezcan asociaciones bilaterales, ya que capturan prácticamente las mismas interacciones que las unilaterales y producen un menor número de pares de proteínas co-expresados y co-localizados. Es decir, parece adecuado considerar no sólo que aparezcan las asociaciones de la forma $GenA[nivelExpresionA] \rightarrow GenB[nivelExpresionB]$, sino que se observen también las relaciones $GenB[nivelExpresionB] \rightarrow GenA[nivelExpresionA]$.

4 CONCLUSIONES

Este trabajo es un primer paso en el desarrollo de una metodología basada en reglas de asociación difusas para el modelado de las interacciones proteína-proteína. Utilizando únicamente datos de expresión génica y de localización celular de las proteínas se consiguen capturar prácticamente todas las interacciones entre proteínas de los complejos permanentes estudiados por Jansen et al. [11]. Aunque se piensa que la no distinción en el conjunto de Collins et al. entre interacciones permanentes y transitorias es la principal causa de que se capturen menos interacciones de este conjunto, un estudio más profundo de estos casos es necesario, ya que la introducción de otro tipo de información en el sistema puede permitir capturar estas ocurrencias.

La metodología proporciona un filtrado efectivo de la gran cantidad de combinaciones proteína-proteína posibles en una célula eucariota como la de la levadura. La introducción en la misma de información estructural de las proteínas así como de sus dominios es imprescindible, ya que dos proteínas no pueden interaccionar a menos que sus estructuras puedan acoplarse, lo cual a su vez hace que las interacciones sean altamente específicas.

Por tanto, la integración de más información en la metodología así como la variación de ciertos parámetros como la granularidad en la definición de los conjuntos difusos serán el objeto de trabajos futuros.

Agradecimientos

Este trabajo se ha llevado a cabo como parte de los proyectos TIC-640 de la Junta de Andalucía, Sevilla, y TIN2006-13177 del MEC, Madrid.

Referencias

[1] R. Agrawal, I. Imielinski, A. Swami. Mining association rules between sets of items in large databases. *In Proc. Of the ACM SIGMOD INTL Conf. on Management of Data. ACM SIGMOD 93*, Pág.207-216, 1993.

[2] G. Bebek, J. Yang. PathFinder: mining signal transduction pathway segments from protein-protein interaction networks. *BMC Bioinformatics*, 8, Pág.335-347, 2007.

[3] S.R. Collins et al. Toward a Comprehensive Atlas of the Physical Interactome of Saccharomyces cerevisiae. *Mol. Cell. Proteomics*, 6, Pág.439-450, 2007.

[4] M. Delgado et al. Fuzzy association rules: General model and applications. *IEEE Trans. Fuzzy Systems*, 11, Pág.214-225, 2003

[5] M. Delgado et al. Mining fuzzy association rules: an overview. *In Proceedings Of the BISC International Workshop on Soft Computing for Internet and Bioinformatics*, 2003

[6] W. Dubitzky et al. Introduction to Microarray Data Analysis. *A Practical Approach to Microarray Data Analysis*, Pág.91-109, 2004.

[7] S.S. Dwight et al. Saccharomyces Genome Database (SGD) provides secondary gene annotation using the Gene Ontology (GO). *Nucleic Acids Res.*, 30, Pág.69-72, 2002.

[8] H. Fei-Hung, C. Hung-Wen. Protein-Protein interaction prediction based on association rules of protein functional regions. *Second International Conference on Innovative Computing, Information and Control, 2007. ICICIC '07*, Pág.359-362, 2007.

[9] H. Ge et al. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. *Nature Genet.*, 29, Pág.482-486, 2001.

[10] U. Güldener et al. CYGD: the Comprehensive Yeast Genome Database. *Nucleic Acids Research*, 33, Pág.D364-D368, 2005.

[11] R. Jansen, D. Greenbaum, M. Gerstein. Relating wholegenome expression data with protein-protein interactions. *Genome Res.*, 12, Pág.37-46, 2002.

[12] D. Jiang et al. Cluster analysis for gene expression data: A survey. *IEEE Transaction on Knowledge and Data Engineering*, 16(11), Pág.1370-1386, 2004.

[13] F.J. Lopez et al. Fuzzy association rules for biological data analysis: a case study on yeast. *BMC Bioinformatics*, 9, Pág.107, 2008.

[14] S. Madeira, A. Olivera. Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1(1), Pág.24-45, 2004.

[15] A. Narayanan, E.C. Keedwell, B. Olsson. Artificial intelligence techniques for bioinformatics. *Appl. Bioinf.*, 1, Pág.191-222, 2002.

[16] A. Valencia, F. Pazos. Computational Methods for the prediction of protein interactions. *Curr. Opin. Struc. Biol.*, 12, Pág.368-373, 2002.

[17] J. Wilson, T. Hunt. Molecular Biology of The Cell, Fourth Edition: A Problems Approach. *Ed. Garland Science*, 2002.

[18] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3), Pág.338-353, 1965.

[19] The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. *Nature Genet.*, 25, Pág.25-29, 2000.

[20] The Comprehensive Yeast Genome Database [http://mips.gsf.de/genre/proj/yeast].

[21] The Saccharomyces Genome Database [http://www.yeastgenome.org]

Tabla 3: Coincidencias entre los conjuntos de interacciones y los pares co-expresados y co-localizados.

CONJ. INTER.	**TIPO 0**	**TIPO 1**	**TIPO 2**	**TIPO 3**
cytopl. ribosomes	8940 (95%)	8938 (95%)	8874 (94%)	8814 (93%)
cytop. ribosomal small sub.	1580 (99%)	1580 (99%)	1580 (99%)	1580 (99%)
cytop. ribosomal large sub.	2956 (91%)	2956 (91%)	2916 (90%)	2884 (89%)
mitoch. ribosomes	1488 (52%)	1413 (50%)	1315 (46%)	1216 (43%)
mitoch. ribosomal small sub.	195 (42%)	179 (38%)	159 (34%)	146 (31%)
mitoch. ribosomal large sub.	589 (62%)	567 (60%)	532 (56%)	490 (52%)
20S proteasome	42 (40%)	42 (40%)	42 (40%)	42 (40%)
cytochrome bc1 comp.	36 (80%)	36 (80%)	36 (80%)	36 (80%)
F0/F1 ATP synthase	124 (81%)	95 (62%)	96 (63%)	76 (50%)
succinate dehydr. complex	6 (100%)	6 (100%)	6 (100%)	6 (100%)
cytochr. c oxidase	31 (56%)	31 (56%)	31 (56%)	31 (56%)
RNA poly.I	50 (30%)	50 (30%)	50 (30%)	50 (30%)
RNA poly. II	14 (18%)	13 (17%)	13 (17%)	13 (17%)
RNA poly. III	29 (37%)	29 (37%)	29 (37%)	29 (37%)
Conjunto completo complejos	10770 (79%)	10663 (79%)	10052 (74%)	10323 (76%)
Collins et al.	4010 (44%)	3987 (44%)	3899 (43%)	3869 (43%)

Tabla 4: Pares co-expresados y co-localizados.

TIPO 0	**TIPO 1**	**TIPO 2**	**TIPO 3**
313986	229994	164534	152550

Fuzzy Contact Maps Overlap for Protein Comparison: the Roles of Fitness and Normalization

Lluvia Morales[1] **Juan R. González**[2] **David Pelta**[2]

Dept. of Computer Science and AI, University of Granada, 18071 Granada,Spain
[1] lluviamorales@ugr.es
[2] {jrgonzalez,dpelta}@decsai.ugr.es

Abstract

The comparison of protein structures is an important problem in Bioinformatics, and Soft Computing techniques were recently introduced for achieving a better representation and potentially, for getting better solving strategies. In this paper we propose new alternatives for measuring the cost in the Generalized Maximum Fuzzy Contact Map Overlap problem, and we analyze the role of normalization, when protein classification is performed.

Keywords: Fuzzy Contact Maps, Protein Comparison.

1 INTRODUCTION

A protein is a complex molecule composed by a linear arrangement of amino acids. Each amino acid is a multi-atom compound. Usually, only the "residue" part of these amino acids are considered when studying protein structures for comparison purposes. Thus a protein's *primary sequence* is usually thought-of as composed of "residues". Under specific physiological conditions, the linear arrangement of residues will *fold* and adopt a complex three dimensional shape. The shape thus adopted is called the *native state* (or tertiary structure) of the protein. In its native state, residues that are far away along the linear arrangement may come into proximity in three dimensional space in a fashion similar to what occurs with the extremes of a sheet of paper when used to produce complex origami shapes. The proximity relation between residues in a protein can be captured by a mathematical construct called a "contact map".

A contact map [9, 8] is a concise representation of a protein's 3D structure. Formally, a map is specified by a 0-1 matrix S, with entries indexed by pairs of protein residues, such that:

$$S_{i,j} = \begin{cases} 1 & \text{if residue } i \text{ and } j \text{ are in contact} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Residues i and j are said to be in "contact" if their Euclidean distance is at most $\Re$ (a threshold measured in Angstroms) in the protein's native fold.

The comparison of proteins through their contact maps is equivalent to solving the maximum contact map overlap problem MAX-CMO [3, 1] (when the maps are crisp) or the generalized fuzzy contact map overlap problem GMAX-FCMO [6] (when the maps are fuzzy). Both belong to the NP-Hard class of complexity.

The GMAX-FCMO has not been as extensively tested as its "crisp" counterpart. In previous work we addressed the comparison of fuzzy contact maps against crisp contact maps[7] and we showed that if we first solved the problem trough GMAX-FCMO model, and then such solutions were measured as in MAX-CMO, the results obtained could be better than those obtained when MAX-CMO is solved directly.

In this paper we wish to extend our research in the GMAX-FCMO in two aspects: a) proposing new alternatives to measure the cost of a solution, and b) to discuss the role of normalization, when the aim is to perform protein structure classification.

The paper is organized as follows: In Section 2, fuzzy contact maps as well as their comparison on the MAX-CMO and GMAX-FCMO models is presented. Section 3 describes the experiments and results obtained. Finally, Section 4 is devoted to the conclusions and future work.

2 FUZZY CONTACT MAPS MODEL DESCRIPTION

Fuzzy contact maps were introduced in [6] with two aims: a) to take into account potential measurements errors in atom coordinates, and b) to allow highlighting features that occurs at different thresholds.

We define a *fuzzy contact* as that made by two residues that are *approximately*, rather than exactly, at a distance $\Re$. Formally, a fuzzy contact is defined by:

$$F_{i,j} = \mu(\overline{[i,j]}, \Re) \tag{2}$$

where $\mu()$ is a particular definition of (fuzzy) contact, $\overline{[i,j]}$ stands for the Euclidean distance between residues i, j, and $\Re$ is the threshold as for the crisp contacts. The standard, i.e. crisp, contact map is just a special case of the fuzzy contact map when a user-defined α-cut is specified.

Figure 1 (a), (b) and (c) shows three alternative definitions for "contact". Each panel in the Figure is a fuzzy contact map where a dot appears for each pair of residues having $F_{i,j} > 0$ (i.e. the support of the corresponding fuzzy set).

Fuzzy contact maps are further generalized by removing the constraint (in the original model) of having only one threshold $\Re$ as a reference distance. In this way, besides having a membership value, a contact will have a "type". The formal definition of a General Fuzzy Contact is given by:

$$F_{i,j} = max\{\mu_1(\overline{[i,j]}, \Re_1), \ldots, \mu_n(\overline{[i,j]}, \Re_n)\} \tag{3}$$

with the contact map C defined as:

$$C^{r \times r} = (F_{i,j}) \; with \; 0 \leq i,j \leq r \tag{4}$$

i.e. up to n different thresholds and up to n different semantic interpretations of "contact" are used to define the $r \times r$ contact map being r the number of residues in the protein.

2.1 PROTEIN COMPARISION TROUGH CONTACT MAPS OVERLAP

A solution for the comparison of two contact maps under the crisp Maximum Contact Map Overlap model consists of an alignment or pairing between the nodes of the two contact maps. The value of an alignment is the number of cycles of length four that appear between the corresponding graphs after the pairings.

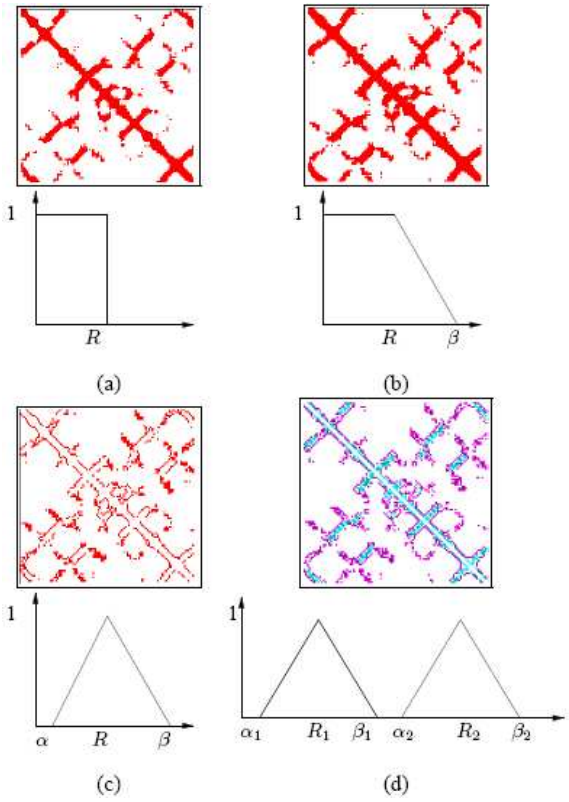


Figure 1: Four examples of contact maps. In (a) the standard model; (b) the simplest fuzzy generalization; (c) another generalization; (d) a two threshold, two membership functions fuzzy contact maps.

This value is called the overlap (fitness) of the two contact maps and the goal is to maximize it.

For example, Figure 2 shows a sample solution for the comparison of two contact maps of 5 and 7 residues respectively. In the crisp model, we can omit the colors of the arcs. Three residues are paired, as shown with a dotted line (first index corresponds to the bottom graph): $1 \leftrightarrow 1$, $2 \leftrightarrow 4$ and $3 \leftrightarrow 5$, and the overlap value is two because there are two cycles of length four.

The first cycle is formed by the pairing $1 \leftrightarrow 1$, the arc $1 \leftrightarrow 5 \in P_2$, the pairing $3 \leftrightarrow 5$ and the arc $3 \leftrightarrow 1 \in P_1$. The second one begins with the pairing $2 \leftrightarrow 4$, follows the arc $4 \leftrightarrow 5 \in P_2$, then the pairing $3 \leftrightarrow 5$ and finally, the arc $3 \leftrightarrow 2 \in P_1$.

Solutions for the GMAX-FCMO model have exactly the same structure as in the crisp MAX-CMO, but the overlap value is computed differently. Now, the arcs have weights and types (colors), so the contribution of each cycle to the global fitness is calculated as a function of the membership values of the con-

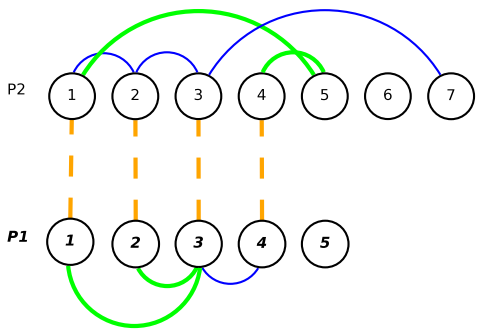


Figure 2: Two levels of contacts in a fuzzy contacts graph

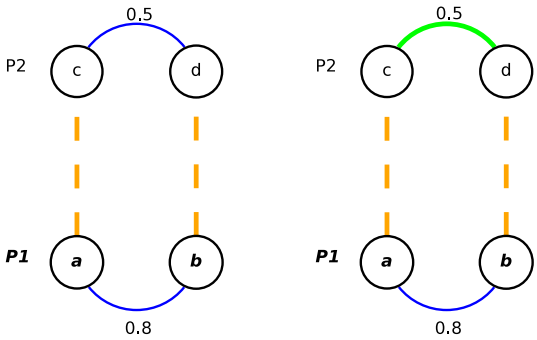


Figure 3: A cycle of length four making contact of the same (left) and different (right) types

tacts involved and their types. In the original GMAX-FCMO model, both membership values are multiplied and then, if both contacts have the same type, then the contribution is added to the fitness. Otherwise, when contacts of different type are involved in a cycle, the contribution is subtracted from the fitness. In this way, alignments between contacts of different types are penalized.

So, as the contact maps in Figure 2 are fuzzy, the contribution of one cycle to the fitness is the product of the membership value of the contact $1 \leftrightarrow 3$ of the bottom protein and that of the contact $1 \leftrightarrow 5$ of the upper protein, with positive sign because the two contacts are of the same type; and, this value is added to the product of the membership of contact $2 \leftrightarrow 3$ and the membership of contact $4 \leftrightarrow 5$.

3 EXPERIMENTS AND RESULTS

The aims of the experiments are: 1) to propose new alternatives for measuring the cost in the Generalized Maximum Fuzzy Contact Map Overlap problem, and 2) to analyze the role of normalization, when protein classification is performed.

3.1 ALTERNATIVES FOR COST CALCULATIONS

The value of an overlap is the sum of the contribution of every cycle of length four. In the crisp model, every cycle contributes a unity to the overlap.

Cycles in the fuzzy model has the appearance shown in Fig. 3. The contribution of a cycle is calculated as $C = \mu(a,b) \times \mu(c,d) \times F(t(a,b), t(c,d))$, where $t(a,b), t(c,d)$ stand for the type (color) of the contact between a, b and c, d respectively. The function F simply returns 1, if both contacts are of the same type, and -1 in other case. So, the cost of the cycles in the example are $0.8 \times 0.5 \times 1$ in Fig. 3 (left) and $0.8 \times 0.5 \times -1$ in the right cycle.

In this experiment, we will use fuzzy contact maps generated from one membership function, thus avoiding the use of function F. Having this in mind, we propose the following set of alternatives to measure the contribution of an individual cycle:

1. Product: $C = \mu(a,b) \times \mu(c,d))$
2. Min: $C = min(\mu(a,b), \mu(c,d))$
3. Max: $C = max(\mu(a,b), \mu(c,d))$
4. Avg: $C = (\mu(a,b) + \mu(c,d))/2$

where the first alternative corresponds to the original GMAX-FCMO cycle contribution.

3.2 NORMALIZATION ALTERNATIVES

Overlap values per se, are not useful (at least in the crisp model) for classification purposes, as such values depend of the size of the proteins being compared. Once the GMAX-FCMO is solved, a normalization scheme should be applied and it is claimed that this scheme may play a crucial role in protein classification.

Following the ideas posed in [7, 5], we will use four alternatives in our experiments:

1. $Norm1(P_i, P_j) =$
   $overlap(P_i, P_j)$ / $min(contactsP_i, contactsP_j)$
2. $Norm2(P_i, P_j) =$
   $2 * overlap(P_i, P_j)/(contactsP_i + contactsP_j)$
3. $Norm3(P_i, P_j) =$
   $\begin{cases} 0 & \text{if the contacts difference is greater than 75\%} \\ Norm1(P_i, P_j) & \text{otherwise} \end{cases}$

4. $NormFuzzy(P_i, P_j) = NormF(P_i, P_j) = overlap(P_i, P_j)\ /\ max(selfSim(P_i),\ selfSim(P_j))$

where the self-similarity ($selfSim$) of a protein is the value of the overlap between a protein with itself.

3.3 COMPUTATIONAL EXPERIMENTS

The global idea of the experiment is as follows: we will conduct queries on a protein database, solving a GMAX-FCMO problem for comparing the query with each protein in the database. Then, we will have a list of overlap values that should be normalized and after that, we will analyze what are the most similar proteins in the database for every query performed (classes are known).

Our protein database consist on 150 selected protein structures from the Nh3D v3.0 test dataset[2]. This dataset has been compiled by selecting well resolved representatives from the Topology level of CATH database and contains 806 topology representatives belonging to 40 architectures, that can be further classified in terms of "class". We selected as the query set the structures that have the nereast to average size for each of the 15 architectures with at least 10 topology representatives. Then, to build the test database we took all the proteins of this 15 architectures removing the query proteins and we picked randomly 10 proteins of each architecture. Each query was then compared against every structure in the test dataset.

For every protein, a fuzzy contact map is constructed using the membership function described in Figure 4 which have a threshold of 8Å and where a decaying slope reduces the level of membership of a contact as the distance approaches from 6.5 to 8.

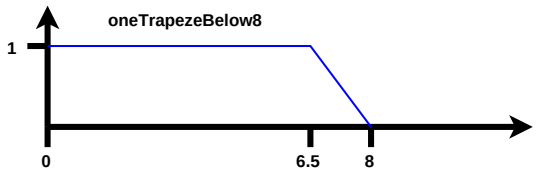


Figure 4: Experimental fuzzy function

Then for solving GMAX-FCMO between the query and every protein, we adapted a Multi-Start Variable Neighborhood Search (VNS) metaheuristic developed for MAX-CMO that was presented in [5]. It is also publicly available online as one of the methods used in the ProCKSi server [4]. This algorithm follows a standard VNS algorithm structure with just a few changes: there is an extra Multi-Start loop to better explore the solution space and reduce the big influence of a single initial solution; and a "simplify" function that is used after every local search to remove pairings that do not contribute to the solution fitness (it helps to avoid the saturation of the solution with useless pairings). The algorithm also uses reduced solution evaluation to recompute the cost of a neighbor solution considering only the changes from the current one, thus significantly reducing the computational time needed.

For every alternative for measuring the cost, we run the VNS metaheuristic and we normalize the overlap values with every normalization alternative proposed. The results are analyzed using ROC curve analysis and the area under the curve (AUC) values, both in terms of classification at the level of architecture and at the level of class.

Table 1: AUCs for the classification at the level of arquitecture

	Fitness	Norm1	Norm2	Norm3	NormF
PRODUCT	0,565	0,468	0,625	0,542	0,622
MIN	0,571	0,479	0,637	0,552	0,631
MAX	0,565	0,470	0,628	0,546	0,623
AVG	0,569	0,476	0,636	0,551	0,629

Table 2: AUCs for the classification at the level of class

	Fitness	Norm1	Norm2	Norm3	NormF
PRODUCT	0,569	0,419	0,553	0,491	0,580
MIN	0,573	0,426	0,557	0,498	0,584
MAX	0,574	0,420	0,561	0,494	0,585
AVG	0,575	0,426	0,561	0,497	0,586

As we can see by looking at any column on Tables 1 and 2, the AUC values do not change significantly for any of the proposed cycle contributions. More precisely, the AUC values within each column have differences of at most 0.011, so the classification performance is mostly unaffected whatever the cycle constribution is. Considering this results it is clear that the actual value for the contribution of a cycle is not important as long as it adds to the final fitness value. This probably comes from the fact that the VNS algorithm will try to pair any residues that lead to more contributing cycles. Therefore, all the relevant pairings will get added as all of them improve the solution. In this manner, the solution obtained with the VNS will remain similar with all the different cycle contributions, and that is the reason why the classification performance (AUC values) is almost identical.

As could be infered from the areas described in Tables 1 and 2 for different normalization kinds, the further

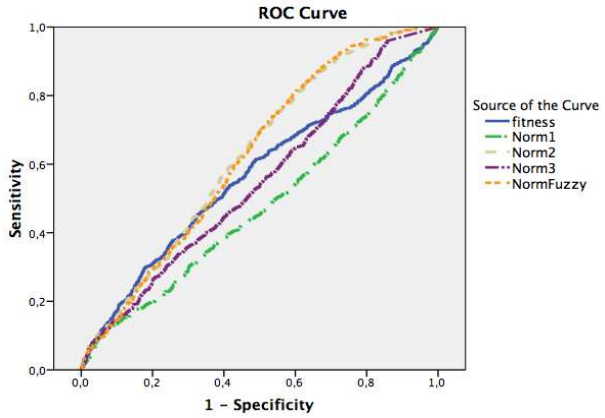


Figure 5: ROC curves for the same architecture of proteins and all kinds of contributions

analysis of the results using ROC curves shows that *Norm2* and *NormFuzzy* normalizations are the best options for post processing the overlap in order to classificate the proteins set studied. This statement can be seen visually in figure 5 for proteins with the same architecture and in figure 6 for proteins with the same class. These results are similar to the results obtained for the crisp model [5] in the sense that the normalization done is again an important factor on the fuzzy model. The differences are very significative among the normalizations with the aforementioned *Norm2* and *NormFuzzy* having reasonably good classification performance at the same time that *Norm1* and *Norm3* lead to worse than random performance (AUCs below 0.5).

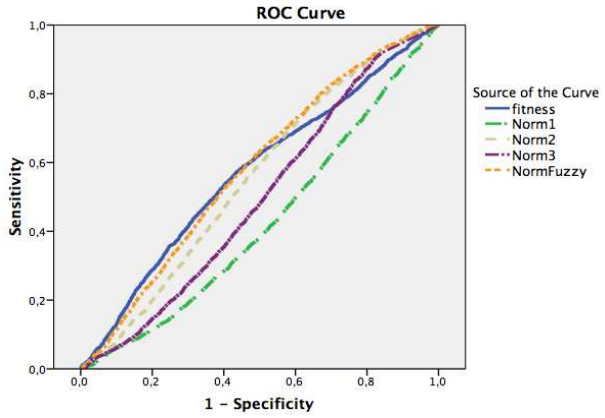


Figure 6: ROC curves for the same class of proteins and all kinds of contributions

4 CONCLUSIONS

This paper has analyzed the influence of the computation of the contribution of each cycle and the normalization scheme when doing protein classification using the GMAX-FCMO model for protein structure comparison.

The results indicate that the strategy used to compute the contribution of each cycle to the solution is not relevant while the normalization is playing a key role. This emphasizes the importance of normalization as it has been probed to be very important both for the crisp and the fuzzy model.

As a future work we plan to conduct more analysis on the fuzzy model in two main lines: how to compute the value of a cycle when there is more than one type of contact; and which values for the distances (in Angstroms) should be used as the basis for the fuzzy functions that define each type of contact.

Acknowledgements

This work is supported in part by Project HeuriCosc (TIN2005-08404-C04-01) from the Spanish Ministry of Education and Science; and Projects MINAS (TIC-00129-JA) and P07-TIC02970 from the Regional Government of Andalusia.

L. Morales is supported by the Mexican National Council on Science and Technology (CONACYT).

References

[1] A. Caprara, R. Carr, S. Istrail, G. Lancia, and B. Walenz, "1001 optimal pdb structure alignments: integer programming methods for finding the maximum contact map overlap." *J Comput Biol*, vol. 11, no. 1, pp. 27–52, 2004.

[2] B. Thiruv, G. Quon, S. A. Saldanha, and B. Steipe, "Nh3D: A Reference Dataset of Non-homologous protein structures", in *BMC Structural Biology 2005 5(12).*

[3] B. Carr, W. Hart, N. Krasnogor, E. Burke, J. Hirst, and J. Smith, "Alignment of protein structures with a memetic evolutionary algorithm," in *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference.* Morgan Kaufman, 2002.

[4] D. Barthel D, J. Hirst, J. Blazewicz, E. Burke, N. Krasnogor. "ProCKSI: a decision support system for Protein (Structure) Comparison, Knowledge, Similarity and Information". *BMC Bioinformatics, 8, 416*, 2007.

[5] D. A. Pelta, J. R. González, and J. M. Moreno-Vega, "A simple and fast heuristic for protein structure comparison", *BMC Bioinformatics, 9, 161*, 2008.

[6] D. Pelta, N. Krasnogor, C. Bousono-Calzon, J. L. Verdegay, J. Hirst, and E. Burke, "A fuzzy sets based generalization of contact maps for the overlap of protein structures," *Journal of Fuzzy Sets and Systems*, vol. 152, no. 1, pp. 103–123, 2005.

[7] J. R. González and D. A. Pelta, "On Using Fuzzy contact Maps for Protein structure Comparison", in *IEEE International Fuzzy Systems Conference. FUZZ-IEEE 2007.*

[8] L. Mirny and E. Domany, "Protein fold recognition and dynamics in the space of contact maps," *Proteins*, vol. 26, pp. 391–410, 1996.

[9] S. Lifson and C. Sander, "Antiparallel and parallel beta-strands differ in amino acid residue preferences," *Nature*, vol. 282, pp. 109–111, 1979.

RESÚMENES DE DIALOGOS, SNIPPETS Y SIMILITUD.

Alejandro Sobrino Cerdeiriña

Dpto. de Lógica y Filosofía Moral. Facultad de Filosofía
Universidad de Santiago de Compostela
España
alejandro.sobrino@usc.es

Resumen

En este trabajo se presentan las analogías y diferencias entre los resúmenes de diálogos entre humanos, en los que se pide información preguntando, y los diálogos persona-máquina focalizada en los buscadores de Internet, donde la introducción de una palabra en la caja de búsqueda puede ser interpretada como hacer una pregunta a una máquina. Se examina también el papel de la similitud en los resúmenes de los diálogos humanos y en los resúmenes de páginas web, conocidos como *snippets*.

Palabras Clave: diálogos humanos, diálogo persona-máquina, resúmenes de diálogos, *snippets*, similitud.

1 INTRODUCCIÓN

Sparck-Jones [7] definió un resumen como '*a reductive transformation of source text to summary text through content condensation by selection and/or generalisation on what is important in the source*". Zadeh [9], por su parte, ha puesto como ejemplos familiares de computar con palabras "*parking a car, driving in heavy traffic, playing golf, riding a bicycle, undestanding speech and summarizing a story*".

Este trabajo presenta algunos aspectos de los resúmenes automáticos basados en preguntas, así como algunas características propias de los resúmenes de documentos web, una de cuyas expresiones más familiares son los *snippets*. Para ello procederemos a analizar las similitudes y diferencias entre resumir un diálogo humano-humano (diálogo basado en preguntas) y un diálogo humano-máquina (diálogo con un buscador), así como el papel del *computing with words* y la similitud en ambos. Los asuntos que se tratan en este artículo son conceptuales o teóricos (qué concepción del significado representa un hipervínculo como respuesta a una consulta web) y no aplicados (cuál es la medida de similitud adecuada para resolver tal o cual problema).

Un diálogo es un proceso mediante el cual dos personas intercambian datos o razonamientos con el objetivo de que el otro comprenda o acepte lo que el primero le dice. Hay varios tipos de diálogo, pero muchos de ellos se inician para conocer cosas que no sabemos. Recabar información sobre lo desconocido ha sido y es una constante en la vida humana.

Buscar información es un acto relevante tanto en el mundo científico como en el ámbito del ocio. En el mundo científico porque el saber se asienta en el conocimiento; en el área del ocio, porque acceder a datos acerca de las cosas más variadas es, sin duda, un entretenimiento. Un ámbito reciente en el que la búsqueda de información tiene un papel relevante es el de la localización de contenidos en la gran biblioteca digital, de conocimiento y de ocio, que es Internet.

Dado su enorme tamaño, para informarse en Internet es imprescindible ayudarse de software especializado genéricamente conocido como 'buscadores'. Admiten preguntas en forma de palabras clave y proporcionan respuestas en términos de un hipervínculo a una página y del *snippet* o breve resumen de la misma. La interacción de un humano con un buscador constituye, así pues, un diálogo simplificado [1].

En la procura de información, el diálogo humano-humano es más complejo que el diálogo humano-buscador. Al dialogar, los humanos no sólo hacemos preguntas y respuestas en turnos convenidos, sino que la cadena de preguntas o respuestas constituyen, a menudo, razonamientos. Una lógica erotética debe proporcionar patrones para evaluar este tipo de argumentos, pero no sólo atendiendo a sus aspectos inferenciales, sino asegurándose de compatibilizar creencias y 'homologar' contextos previos, a fin de que esas inferencias provoquen retroalimentación propiciando una nueva pregunta y, en consecuencia, haciendo que el diálogo progrese.

El diálogo humano-máquina es, en cambio, mucho más restringido. En el caso de los buscadores, se basa fundamentalmente en el emparejamiento de cadenas de caracteres que devuelven páginas como respuesta a

preguntas y en explotar la tecnología del hipertexto asociada a las páginas web.

Este trabajo hará un recorrido por estos temas. En el punto 2, se exponen las características principales del diálogo humano. En el 3 se detallan algunas notas del diálogo humano-máquina (en este caso, humano-buscador). En el 4 se analizan los resúmenes de diálogos humanos y algunos procedimientos basados en *soft-computing* y similitud que facilitan su obtención. En el apartado 5 se estudian los *snippets* y algunas maneras, basadas en similitud, de generarlos. Finalmente, en 6 se resumen algunas conclusiones.

2. DIÁLOGO HUMANO-HUMANO.

En un contexto humano, como apuntó Delancy [2], dialogar es intercambiar información para convencer al otro*: "In its communicative functions, language is a set of tools with which we attempt to guide another mind to creative within itself a mental representation that approximates one we have"*. Hay distintas formas de convencer asociadas a otros tantos tipos de diálogo. Walton propone en [8] la siguiente tipología:

Tipos de diálogo	Situación inicial	Objetivo del participante	Objetivo del diálogo
Persuasión	Conflicto de opiniones	Persuadir a la otra parte	Resolver o clarificar el problema
Negociación	Conflicto de intereses	Alcanzar lo que se desea	Acuerdo razonable que no perjudique mucho a ninguno
Búsqueda de Información	Necesidad de información	Adquirir o dar información	Intercambiar información
Deliberación	Dilema o elección	Coordinar objetivos y acciones	Decidir el mejor curso de acción disponible
Erístico	Conflicto personal	Descalificar verbalmente al oponente	Revelar las bases profundas del conflicto
Investigación, pesquisa	Necesidad de una prueba	Encontrar y verificar la evidencia	Probar o desaprobar la hipótesis

En este trabajo nos interesa el tipo de diálogo usado para buscar información. En este, como en cualquier otro, es posible diferenciar varios componentes:
- lingüístico, que muestra su forma en términos de turnos de pregunta-respuesta
- intencional, que atiende al propósito del discurso: convencer, conminar, probar,…
- atencional, que recoge los objetos, propiedades y relaciones más destacadas en cada paso del diálogo.

Diálogos y resúmenes se vinculan de la siguiente manera: informarse de documentos muy extensos o de naturaleza muy heterogénea es difícil. En ese caso, el peticionario de la información suele hacer una pregunta que actúa de hilo conductor para las posibles respuestas, concretando el foco al que debe restringirse el que responde en el diálogo. Este debe ponerse en el lugar del que pregunta para entender la cuestión y dar una respuesta cabal a la misma. Ello supone compatibilizar creencias, aproximar contextos y cumplir una serie de preceptos que aseguren la comunicación.

Grice [4] indicó que algunos preceptos o máximas que todo diálogo debe cumplir y a los que denominó implicaturas conversacionales:
- Máxima de cantidad (informatividad): proporciona siempre a tu interlocutor la información necesaria, ni más ni menos.
- Máxima de calidad (verosimilitud): trata de que lo que digas sea verdadero.
- Máxima de la relevancia: no digas cosas superfluas, sino importantes para el curso de la conversación.
- Máxima de la cortesía: se breve y ordenado; evita la oscuridad y la ambigüedad.

En un diálogo, los argumentos tienen el siguiente formato: una parte hace suya lo que dice otra como premisas y, usando reglas de inferencia, logra probar una conclusión a la otra parte, que a su vez puede aceptar o no, construyendo un contra-argumento que se base en nueva información o que recuse la argumentación anterior por algún defecto de forma. En los diálogos, los argumentos no se construyen con significados binarios (significa(A, B)), sino ternarios (significa(A, B, C)) -A significa B para C-. Si C no capta el significado B que A le intenta transmitir, A debe revisar sus creencias o verificar el grado de afinidad de sus contextos, para hacer más asequible B a C. El razonamiento dialogado es fundamentalmente un razonamiento revisable y analógico: la compatibilización de creencias o contextos suele ser parcial, no absoluta aunque, en cualquier caso, suficiente para el progreso del mismo.

En un formalismo con sabor lógico algunos de esos requisitos se podrían explicitar así:
- Predicados:
quiere(A, X): El agente A quiere hacer X
cree(A, X): El agente A cree X
sabe(A, X): El agente A conoce X
pretende(A, cree(B, X)): A pretende que B crea X
- Axiomas:
$(sabe(A, X), (X \rightarrow Y)) \rightarrow sabe(A, Y)$
$(cree(A, X), (X \rightarrow Y)) \rightarrow cree(A, Y)$
$(cree(A, X), X) \rightarrow sabe(A, X)$
Ya que las creencias son personales, estos axiomas no deberían usarse para inferir verdades universales o definitivas. Así, si A cree X con intensidad α e Y se sigue de X con grado β, el grado de intensidad con el que B crea Y debería ser una función de α y β.

Las creencias juegan un papel importante en el diálogo y los seres humanos son capaces de sincronizar creencias. P. ej., considérese la siguiente escena: acudo a un centro comercial para comprar una consola de videojuegos. Leo

el folleto de ofertas y compruebo que están todas las grandes marcas del mercado. Pero a la hora de pedir desconozco que una de ellas se ha agotado, extremo que sí sabe el dependiente del comercio

D: ¿Puedo ayudarle en algo?

C: Quería información sobre las mejores consolas del mercado.

D: Existen tres modelos que destacan sobre los demás: ps3, x_box y wii. ps3 y x_box tienen mejores servidores para juegos on-line y wii posee mejores interfaces.

C: Creo que me decanto por la wii. Justamente valoro mucho la interactividad con la máquina.

D: Lamentablemente, la wii está agotada.

C: En ese caso, esperaré a que la tengan.

D: En breve dispondremos de ella. Buenas tardes, señor.

C: Buenas tardes. Hasta luego.

Este diálogo es una pequeña muestra de sincronización de creencias y a este proceso se le conoce como *grounding*. Las creencias compartidas y la sincronización de intenciones pueden ser expresadas [5] de esta manera:

cree(A, P) ∧ cree (B, P) → creencia compartida (A, B, P)

En el caso del ejemplo anterior, esta creencia compartida se hace explícita en el tercer turno del diálogo y fuerza el fin del mismo. De modo explícito, la sincronización de creencias e intenciones se podría esquematizar así: (supongamos que el dependiente se llama Juan y el consumidor, Pedro):

Creencias del dependiente	Creencias del consumidor
cree(dependiente(juan), disponible(p_station)). cree(dependiente(juan), disponible(x_box)).	cree(consumidor (pedro), disponible(ps3)). cree(consumidor (pedro), disponible(x_box)). cree(consumidor (pedro), disponible(wii)).

Intenciones del dependiente	Intenciones del consumidor
aconsejar (dependiente(juan), consola(ps3)). aconsejar(dependiente(juan), consola (x_box)). aconsejar(dependiente(juan), consola (wii)). vender(dependiente(juan), disponible(ps3)). vender(dependiente(juan), disponible(x_box)).	informarse(consumidor(pedro), consola(ps3)) informarse(consumidor(pedro), consola(x_box)) informarse(consumidor(pedro), consola(wii)) comprar(consumidor (pedro), consola(wii)).

Los diálogos humanos comportan la siguiente tipología:

1. Dos agentes conversacionales: A (proponente), B (el que responde). A debe hacer una propuesta clara, para que B pueda sacar conclusiones. Para el primero es clave expresar su 'intención'; para el segundo lo relevante es captar la intención de A y, haciendo uso de su *background*, realizar una inferencia que le de pie a una nueva pregunta.
2. Un conjunto de actos de habla. Los agentes (a) hacen preguntas; (b) hacen aserciones; (c) se retractan de las aserciones y (d) proponen argumentos para el debate. Se pueden expresar actos de habla, sencillos o complejos. P. ej.:

sabe(A, P) ∧ quiere(A, informar(A, B, P)),
→ sabe(B, P)

recoge el acto de haba de que alguien que sabe algo informa de ello a un tercero que, a partir de ese momento, también lo sabe.

desea(A, (cree(B, p) ∧
sabe(A, (q_1, q_2,...,q_k)) ∧
demuestra(A, se_sigue(p, q_1, q_2,...,q_k))
→ cree(B, p).

formaliza el acto de habla complejo de desear que alguien crea algo y verlo consumado después de habérselo demostrado.

3. Una vocación clara de que el diálogo progrese.

Como podremos ver a continuación, los aspectos que tienen que ver con creencias o intenciones todavía no juegan papel alguno ni han sido implementados en los resúmenes de diálogos humanos.

3. RESÚMENES DE DIÁLOGOS HUMANOS

Hay diferentes tipos de resúmenes. Se pueden clasificar atendiendo al

- género: indicativos, informativos, críticos
- formato: texto, tablas, esquemas
- escala: palabras clave, textos breves, documentos largos
- especialización: genéricos o focalizados en la pregunta.

En este trabajo nos ocupamos de la última categoría y, más en concreto, de los resúmenes motivados por una pregunta (*query-driven*). El resumen guiado por preguntas va 'de arriba abajo': el interés que suscita la pregunta ha de aparecer explícito en la estrategia de búsqueda de la respuesta y el sistema ha de usarlo para filtrar o analizar las porciones de texto que sean una buena contestación de la pregunta planteada. Una ventaja de los resúmenes basados en preguntas es que permiten una evaluación clara, en términos de la ayuda que suponen para la tarea a la que se aplican, para propósitos relacionados con la pregunta.

La idea básica para resumir un diálogo se basa en cuantificar el grado de similitud semántica entre determinadas palabras (generalmente nombres), o el concepto que las representa, y el diálogo como un todo. Se han propuesto varias métricas de similitud semántica. Entre ellas:

1. *Normalized Path Length* (de Leacock y Chodorov)

La similitud semántica entre dos palabras p_1 y p_2 es

$$sim(c_1, c_2) = -\log \frac{len(c_1, c_2)}{2 \times D}$$

donde c_1 y c_2 son los conceptos –proporcionados p. ej., por WordNet- correspondientes a p_1 y p_2; $len(c_1,c_2)$ es la

longitud del camino más corto entre ellos y D es la profundidad máxima de la taxonomía.

2. *Extended Gloss Overlaps Measure* (de Banerjee y Pedersen). Se basa en el número de palabras compartidas (*overlaped*) en las definiciones (*glosses*) de los conceptos de WordNet. Se puede extender la definición para incluir aquellos conceptos relacionados con el concepto fuente en la jerarquía de WordNet. La similitud entre dos conceptos es:

$$sim(c_1, c_2) = \sum score(R_1(c_1), R_2(c_2))$$

donde R es el conjunto de relaciones semánticas, *score(_)* es una función que acepta dos definiciones *(glosses)* como entrada, encuentra los solapes y devuelve una puntuación de salida.

Estas medidas permiten hallar el grado de similitud entre una expresión y el diálogo como un todo. Son usadas por varios sistemas para resumir diálogos. El de Gurevych y Strube [3], procede así:

- segmenta el diálogo, D, en expresiones, Exp_i; {Exp_1, Exp_2, …, Exp_n}.
- halla la representación conceptual de Exp_n (CR_{Exp^n}) y de D (CR_D) usando WordNet.
- Para cada expresión Exp_n crea una matriz bidimensional $\#CR_D \times CR_{Exp}$ y computa, para cada par, su similitud semántica usando una métrica de similitud conceptual, obteniéndose una *puntuación(i, j)*.
- La puntuación final es la similitud media promediada para cada par entre todos los conceptos CR_{Exp^n} y CR_D:

$$puntuación_{final} = \frac{\sum_{i=1}^{\#Exp_n} \sum_{j=1}^{CR_D} puntuación(i, j)}{\#CR_{Exp_n} \times CR_D}$$

- Finalmente se ordenan las expresiones por sus puntuaciones. A continuación, se aplica una ratio de compresión COMP con rango [0, 100] que permita seleccionar un número determinado de expresiones relevantes PN:

$$PN_r = (\frac{COMP}{100}) \times Numero_{total}$$

-Esas expresiones configurarán, en un orden adecuado, el resumen.

P. ej., para el siguiente diálogo:

A	Expr1	Okay
	Expr2	Tell me about your home
B	Expr3	Well, it's an older home
	Expr4	It was made back in the early sixties
	Expr5	It's a pier beam house
A	Expr6	Huh-uh
B	Expr7	Got three bedrooms, one bath
	Expr8	And that just makes me scream
A	Expr9	That´s pretty tough
	Expr10	What area do you live in?
		I live in Houston

El algoritmo facilita el siguiente resumen, atendiendo a ratios de compresión del 20 y 35%

COMP	Expresiones seleccionadas
20%	I live in Houston Got three bedrooms, one bath
35%	I live in Houston Got three bedrooms, one bath Tell me about your home Well, it´s an older home

4. DIÁLOGO HUMANO-MÁQUINA: LA BÚSQUEDA EN LA WEB COMO UNA ESPECIE DE DIÁLOGO.

Hasta ahora, el proceso de búsqueda de información en la Web supone, casi siempre, la utilización de un software que hace de intermediario. Se le conoce como motor de búsqueda y proporciona, para una pregunta dada, una lista ordenada de enlaces a aquellas páginas que deberían ser o contener una respuesta relevante. Este proceso presenta, entre otras, estas notas distintivas:

1. *Respecto a la interacción con el usuario:*

• Es cerrado. El usuario hace la pregunta y el buscador ofrece un orden de posibles respuestas. No puede terciar nadie más en el proceso.

• Es directo. Las respuestas se actualizan ejecutando el hipertexto y no se permite la intervención del que hace la búsqueda para especificar algún tipo de matiz o contexto.

• Es estático. El hipertexto lleva de una información a otra de manera bidireccional (va hacia delante y hacia atrás), pero los vínculos no cambian cuando se vuelve hacia atrás. Una ligera variación en la búsqueda supone iniciar un camino totalmente diferente.

2. *Respecto al diálogo humano-humano.*

• No hay un sistema de turnos. La persona pregunta una vez y la máquina da respuesta (orden de páginas) y facilita el acceso a otras páginas (hipervínculos)

• Hay un uso empobrecido del lenguaje, que se reduce a poner palabras clave.

• El clic en el hiperenlace es un símil del significado como denotación por ostensión, en tanto que proporciona la respuesta a una interrogación apuntando simplemente a la página a la que enlaza.

• En la web, al contrario de lo que suele suceder en un diálogo humano, la atención que se le presta a la respuesta para continuar el diálogo es superficial. La mayoría de las veces simplemente se ojea el contenido de la página para decidir si seguir o no elicitando al buscador o pinchando en uno u otro hiperenlace.

• En un proceso de búsqueda mecánica, cuando se va hacia atrás no es tanto para revisar lo anterior, sino para desestimar lo actual.

• En la web, las preguntas se hacen generalmente en contextos potencialmente muy diferentes; por tanto, suelen formularse de una manera muy general o vaga. Al introducir un término *t* en la caja de búsqueda se está, la

mayor parte de las veces, solicitando información que presente algún tipo de similitud con *t*.
3. *Respecto al uso de la tecnología de hipertexto:*
• Una página puede ser una respuesta adecuada a una pregunta porque contiene la información deseada o porque dirige a otra que contiene esa información.
• Las posibilidades de preguntar se sacrifican al diseño. Las herramientas de búsqueda tienen que ser intuitivas y atractivas para invitar a su uso. Por eso, la mayoría de los buscadores sólo tiene en su página principal a la caja de búsqueda y, como opción, hiperenlaces a otras utilidades. Entre ellas suele encontrarse la búsqueda avanzada, opción desestimada en la mayoría de las búsquedas.
• La velocidad de respuesta o latencia es crítica en los buscadores. Por eso, el proceso de búsqueda tiene que ser muy sencillo (usualmente se basa en el igualamiento léxico) para ofrecer respuestas inmediatas, primando siempre la cobertura sobre la precisión.

5. RESÚMENES DE DOCUMENTOS WEB

Ser capaz de describir de modo resumido y coherente el contenido de las páginas web ofrece ventajas. Se pueden usar para presentarlos como los resultados de una búsqueda o para etiquetar de modo breve *bookmarks* personalizadas. Hay dos tipos fundamentales de resúmenes de contenidos web:
-los generados automáticamente (*snippets*)
-los hechos por humanos (resúmenes para Directorios como el Open Directory Project)
Estos segundos se usan, habitualmente, como pruebas para evaluar a los primeros.
Actualmente hay varias formas de generar *snippets*:
• basados en técnicas lingüístico/estadísticas
• basados en metaetiquetas insertadas por los autores de las páginas web
En general, resumir documentos web presenta características distintivas:
- No hay dominios restringidos. Todo entra en la web, desde artículos de periódico a listas de hipervínculos.
- Las páginas web carecen de marcadores sintácticos. Sólo tienen marcas físicas; tales como parágrafo, título o encabezamientos, y se asume que el autor del documento los inserta de manera adecuada.
- En muchos casos, los documentos contienen solamente unas pocas palabras o fracciones de oraciones sin cohesión alguna entre sí.
- Las técnicas de resumen usadas en aplicaciones *on-line* no deben requerir pre-procesamiento o entrenamiento. Deberían ser completadas tan rápido como fuera posible.
- La información de las páginas web puede incluir imágenes y otras entidades no textuales (e. g. audio, video, etc.) que podrían tener importancia en un resumen.

Hay fundamentalmente tres formas de obtener resúmenes de documentos web:
- *Basada en parágrafos.* Se trata de identificar un parágrafo sencillo o un segmento de texto que resuma el contenido principal del mismo. Este método se emplea, sobre todo, cuando el texto es grande y refiere a más de un tópico. Pero presenta dificultades.
(a) La estructuración de la las páginas web no siempre permite recuperar un parágrafo sencillo.
(b) Si hay hipervínculos, el texto se suele segmentar en diferentes nodos y un documento sencillo puede resultar demasiado corto para segmentarlo en parágrafos.
(c) Este procedimiento no permite siquiera dar una breve indicación del contenido de una imagen.
- *Basada en oraciones.* Está técnica se basa en la repetición de términos o frases en el texto para extraer las oraciones más relevantes o las frases clave y juntarlas de modo conveniente. Asume que un resumen no tiene porque tener cohesión, se sacrifica la coherencia por ofrecer tanta información como sea posible del texto fuente. Puede usar heurísticas, como preferir los términos que aparecen en los títulos, en negrilla, etc. Esta técnica, aunque tampoco puede tratar el contenido gráfico, puede ser más útil que la anterior, ya que hace uso de pequeñas porciones de información.
- *Basado en el modelo de discurso*. Usa técnicas léxicas como el orden de las palabras, la identificación de los nombres propios, las reiteraciones, la sinonimia, la anáfora, las conectivas, etc. Esta técnica une oraciones del texto rescatadas según esta heurística y el resumen resultante es un 'collage' de lo que se puede encontrar en el texto. Ya que algunas características son específicas de dominio, se ajusta a un tipo de texto y a géneros específicos de escritura. Esto hace que la implementación de tales sistemas para uso extensivo, como reclamaría la Web, sea difícil.

Existen varios métodos basados en similitud para generar resúmenes automáticos de páginas web. Expondremos dos [6]:
1. *Identificación de contenido por análisis de marcas*.
Se intenta comprender la intención del autor de la página identificando categorías y funciones de los objetos. Los objetos se clasifican en básicos (ob) y compuestos (oc). Un *ob* es, p. ej., un fichero jpeg. En HTML, un *ob* es un elemento que no admite separación dentro de dos etiquetas o un objeto anidado. Un *oc* es un conjunto de *ob* que hacen una función. Pueden ser detectados por un análisis del diseño de las páginas web. La idea básica es que los objetos que pertenecen a la misma categoría tienen estilos visuales consistentes, de tal manera que puedan ser separados de los objetos de otras categorías. Después de detectados todos los *ob* y *oc* de una página web, se identifica la categoría de cada objeto usando heurísticas. Estas son algunas de las categorías:
• Objeto de información
• Objeto de navegación
• Objeto de interacción
• Objeto de decoración

• Objeto de función especial: logo, copyright, etc.

A continuación se define el *Content Body* (CB) de una página web como los principales objetos relacionados con el contenido de esa página. Este es el algoritmo que se emplea para hallarlo:

- Considerar cada objeto seleccionado como un documento sencillo y hallar el índice *tf×idf* para él.
- Calcular la similitud entre dos objetos cualesquiera con la función coseno y añadir un enlace entre ellos si supera un umbral. El umbral se escoge empíricamente. Después de haber procesado todos los pares de objetos se obtiene un grafo conectado de diferentes objetos.
- En el grafo se define el *Core Objetc* (CO) como el elemento que tiene el mayor número de enlaces
- CB estará formado por todos los objetos que tienen un enlace con el CO.
- Se le asigna puntuación 1 a aquellos objetos que está en CB (0 en caso contrario). Las verbalizaciones de los objetos con puntuación 1 se incorporan al resumen de la página web en cuestión.

2. *Resúmenes supervisados*

Basado en aproximaciones de aprendizaje automático, los resúmenes supervisados parten de extraer, en primer lugar, un conjunto de características de un página web. Sobre estas características actúa un algoritmo de aprendizaje supervisado que entrena al resumidor automático, con el fin de que identifique si una oración debe ser seleccionada o no para figurar en el resumen. El algoritmo utiliza un total de 8 características, 5 propias de documentos textuales y 3 específicas de los enlaces web. Dado un conjunto de oraciones S_i (i=1,…,n), se definen las 8 características así:

f_{i1} Mide la posición de la oración S en un cierto parágrafo.

f_{i2} Mide la longitud de una oración S, esto es, el número de palabras que contiene.

f_{i3} Se halla $\sum TF_w \times SF_w$. Tiene en cuenta no sólo el número de palabras, sino también su distribución entre oraciones. Se usa para penalizar palabras localmente frecuentes.

f_{i4} Similitud entre S_i y el título, calculada como el coseno de los vectores TF correspondientes.

f_{i5} Similitud como coseno entre S_i y el texto de la página.

f_{i6} Similitud como coseno entre S_i y los metadatos de la página.

f_{i7} Número de ocurrencias de la palabras de S_i en un clase especial de palabras (aquellas que están escritas en itálica, negrilla o subrayado).

f_{i8} Tamaño promedio de la fuente de las palabras de S_i. En general, a mayor tamaño, mayor importancia.

Una vez extraídas esas características, se aplica un clasificador bayesiano ingenuo para entrenar al resumidor automático:

$$p(s\in S \mid f_1, f_2, \ldots, f_8) = \frac{\prod_{j=1}^{8} p(f_j \mid s\in S)\, p(s\in S)}{\prod_{j=1}^{8} p(f_j)}$$

donde $p(s\in S)$ está por la ratio de compresión del sistema, que puede ser adaptada a diferentes utilidades, $p(f_i)$ es la probabilidad de cada característica i y $p(f_i|s\in S)$ es la probabilidad condicional de cada característica i. Los últimos dos vectores pueden ser estimados a partir del corpus de entrenamiento. Cada oración tendrá una puntuación por la ecuación de arriba, que será denotada S_{sup}. En el resumen figurarán las oraciones que tengan puntuación mayor.

6. CONCLUSIONES

Se ha hecho un acercamiento a algunos problemas de los resúmenes de diálogos y de los resúmenes de contenidos web, así como a las diferencias y analogías entre ambos. También se ha presentado el papel fundamental de la computación con palabras, desde una perspectiva léxica o conceptual, en la generación de ambos.

Agradecimientos

Financiación a cargo de los proyectos HUM 2007-66607-C04-02 del MEC y SCAIWEB PAC06-0059 de la Junta de Castilla La Mancha

Bibliografía.

[1] Dale, R. et al., 'The Web as a Dialogue: the role of natural language generation in hypertext', *Proc. of the AAAI'97 Spring Symp.: Natural Language for the World Wide Web*, Stanford Univ., March 24-26

[2] Delancy, S., "What an innatist argument schould look like", in T. Haukioja, et. al. (eds), *SKY 1997* pp. 7-24

[3] Gurevych, I. & Strube, M., 'Semantic Similarity Applied to Spoken Dialogue Summarization', In: *Proc. of the 20th Intern. Conf. on Computational Linguistics*, Geneva, 23-27 August 2004, pp.764-770.

[4] Grice, P., 'Logic and Conversation' In P. Cole and J. L. Morgan (eds.), *Syntax and Semantics, III*. Speech Acts. Academic Press, New York, 41-58, 1975

[5] Nugues, P. M., *An introduction to language processing with Perl and Prolog*. Springer, 2006.

[6] Shen, D. et al., 'Noise reduction through summarization for Web-page classification', *Information Processing and Management*, 43, 6, 2007, 1735-1747.

[7] Sparck-Jones, K., 'Automatic Summarising: the state of the art', *Information Processing and Management*, 43, 6, 2007, 1443-1481.

[8] Walton, D. 'The place of dialogue theory in logic, computer science and communication studies', *Synthese*, 123, 3, 327-346, 2000

[9] Zadeh, L. A., 'From Computing with numbers to computing with words. From manipulation of measurements to manipulation of perceptions-', *Int. J. Appl. Math. Comput. Sci*., 2002, vol. 12, nº 3, 307-324.

RECUPERAR INFORMACIÓN CON CONOCIMIENTO

Sara Carrera Carrera[1] **Milagros Fernández Gavilanes**[2] **Manuel Vilares Ferro**[1]

[1] Departamento de Informática, Universidad de Vigo
Campus As Lagoas s/n, 32004 Ourense, {sccarrera,vilares}@uvigo.es
[2] Departamento de Computación, Universidad de A Coruña
Campus de Elviña s/n, 15071 A Coruña, {mfgavilanes@udc.es}

Resumen

Introducimos una propuesta en recuperación de información basada en la consideración de recursos sintácticos y semánticos generados automáticamente a partir de la propia colección documental. Efectuamos un estudio comparativo entre dos estrategias que pueden actúar de forma complementaria. En la primera se clasifican los conceptos relevantes en el dominio a partir de una extracción de términos. En la segunda analizamos las dependencias sintácticas existentes en la colección documental, de manera a obtener una asociación de los elementos que la conforman.

Palabras Clave: Adquisición del conocimiento, procesamiento del lenguaje natural, recuperación de información.

1. INTRODUCCIÓN

El contínuo incremento de textos disponibles en formato digital hace necesaria la creación de herramientas que capturen el conocimiento, de forma a poder manipularlo y explotarlo. Nos estamos refiriendo a la necesidad de automatizar la adquisición del conocimiento, para mejorar la calidad de los sistemas de *recuperación de información* (RI), suministrándoles estructuras conceptuales que reflejen el contenido de un documento.

La eficiencia de las herramientas en este campo se vincula con la observación de los datos semánticos relevantes descritos en los términos y conceptos del dominio considerado. Este tipo de recursos se suele importar de módulos externos y genéricos [1]. Consecuencia de esto es la inoperancia en el tratamiento de dominios aún sin explotar, la dificultad de mantenimiento y la imposibilidad de abordar el procesamiento de colecciones dinámicas como son las ligadas a Internet. Para solucionar esta conyuntura y producir resultados prácticos y comprensibles, debemos permitir la integración del conocimiento subyacente en los documentos; intentando aprovechar las estructuras lingüísticas que le sirven de soporte.

El estado del arte actual distingue dos corrientes para plantear el problema descrito. Se pueden agrupar en acercamientos basados en *similaridad* y aproximaciones basadas en *conjuntos teóricos*. El primer tipo utiliza una medida de distancia para calcular la similaridad entre pares de vectores de palabras [3]. En la segunda aproximación, los conjuntos teóricos ordenan de manera parcial los objetos de acuerdo a la posible existencia de relaciones de inclusión entre sus atributos [6]. Ambas técnicas adoptan el modelo de espacio vectorial y representan un término como un vector de atributos, identificados mediante características sintácticas.

Nuestra propuesta se centra en facilitar la tarea de adquisición de conocimiento a través de una aproximación híbrida. Combinamos técnicas propias del *procesamiento del lenguaje natural* (PLN) tales como el análisis sintáctico superficial y los marcadores semánticos por una parte, con técnicas de aprendizaje basadas en métodos estadísticos. Aunque es posible imaginar que el ámbito ideal de aplicación de una herramienta de este tipo sería el representado por Internet, haciendo que la tarea de RI fuese más precisa y completa, nuestra propuesta se expondrá sobre un sencillo corpus paralelo de economía en francés y español con el objeto de facilitar su descripción.

2. EXTRAER EL CONOCIMIENTO

De manera intuitiva, nuestro interés se centra en recopilar las relaciones semánticas que surjan del texto. Esto conlleva la organización de los términos en clases de acuerdo a su similaridad. Hay que destacar el propósito experimental de nuestro trabajo, que no es otro que el de establecer un marco de evaluación de dis-

tintas técnicas de adquisición de conocimiento, una de ellas basada en información léxica y la otra en información sintáctica. Se considera, además, la posibilidad de complementar ambas capacidades. Generando a partir de la extracción de términos un conjunto de clases semánticas que se utilizarían para inicializar el proceso de clasificación por dependencias. En cualquier caso, nuestro marco de trabajo se apoya en el modelo de distribución semántica de Harris [4] que establece que aquellas palabras que aparecen en los mismos contextos son semánticamente próximas.

2.1. ORDENACIÓN DE TÉRMINOS

En esta sección se describe la puesta en marcha de un prototipo rápido de clasificación de terminología. Partimos de un extractor de términos de propósito general. El corpus de entrada para esta tarea será analizado con un etiquetador léxico que asigna a cada palabra una categoría y sus posibles lemas. El extractor de términos nos proporciona el conjunto de términos base, junto con sus variantes sintácticas o morfo/sintácticas, combinando técnicas lingüísticas y estadísticas.

En un primer momento recogemos todos aquellos términos cuya frecuencia de aparición sea mayor a un umbral dado, estos datos se guardan en un fichero XML[1]. En este punto es necesario definir una restricción sobre los términos, así consideramos que cada término está formado por una *cabeza* y una *expansión*. La cabeza es el núcleo alrededor del cual asumimos que se estructura el significado del término mientras que la expansión es su modificador, tal como se ilustra en el Cuadro 1. Consideraremos para el proceso sólo aquellas *cabezas relevantes*, esto es, cabezas que bien sean sustantivos, bien nombres propios o bien acrónimos. Estimamos tales categorías léxicas como las que aportan mayor información semántica, sobre un enfoque en régimen nominal del problema.

La idea parte de utilizar los contextos en los cuales se usan los términos, y aproximar aquellos cuyo contexto es similar. Así, dado un elemento y un término cualquiera donde aparece ese elemento, su contexto será bien la cabeza bien la expansión que lo acompaña. El proceso de clasificación se basa en un algoritmo iterativo que mide la similaridad entre los contextos de aparicición de las cabezas relevantes.

Cuadro 1: Ejemplo de la extracción de términos

Término	**cabeza**	**expansión**
sociedad de gestión	sociedad	de gestión
fondo luxemburgués	fondo	luxemburgués
sesión de subida	sesión	de subida

Generamos una tabla de colisión cuyas entradas son las cabezas relevantes que llamaremos *elementos principales* y el valor de cada entrada es una lista de los contextos de aparición del elemento principal correspondiente. Partimos de una estructura global que captura el significado esencial del texto. El siguiente paso consiste en agrupar los términos en clases semánticas. Para ello, recorremos la tabla de colisión comparando los distintos contextos aplicando como medida de similaridad[2], el coeficiente DICE, definido como sigue:

$$\text{DICE}(C_1, C_2) = \frac{|C_1 \cap C_2|}{(|C_1| + |C_2|)/2} \qquad (1)$$

donde C_1 y C_2 son los contextos de dos términos y $|C_i|$ representa el cardinal de $C_i,\ i = 1, 2$. Intuitivamente, calculamos los contextos comunes entre los términos y aplicamos normalización. En cada paso del procedimento, se unen pares de elementos principales cuyo valor para el coeficiente DICE es el mayor en esa iteración. De este modo, la tabla de colisión se reduce en un elemento en cada iteración. El proceso finaliza cuando ningún contexto es compartido y el coeficiente es cero en todos los casos.

Una vez acaba el algoritmo, las entradas en la tabla de colisión son palabras relacionadas semánticamente junto con los contextos que comparten. Su lectura final produce una salida que será almacenada en un fichero OWL[3].

2.2. ORDENACIÓN DE DEPENDENCIAS

Partimos de un análisis sintáctico robusto basado en una cascada de autómatas finitos [8]. Identificamos términos relevantes en frases nominales y verbales, es decir, aquellos sustantivos y verbos que aportan la información semántica esencial junto con las relaciones entre ellos. Como resultado obtenemos un grafo de dependencias del tipo *gobernante/gobernado*, como se muestra en la Figura 1.

2.2.1. Depurando las dependencias

Los analizadores sintácticos trabajan con datos léxicos que aportan la información sobre las categorías gramaticales y los lemas de las palabras. En nuestro caso, el analizador sintáctico nos devuelve un conjunto de dependencias entre pares de palabras. Sin embargo, tanto a nivel léxico como a nivel sintáctico pueden surgir una serie de ambigüedades que es necesario solventar para extraer la semántica latente en los documentos. La idea consiste en recopilar información del corpus para detectar y eliminar este tipo de estructuras.

[1] ver http://www.w3.org/XML/

[2] definimos la similaridad entre entidades como el número de propiedades comunes compartidas por ellas.

[3] ver http://www.w3.org/TR/owl-features/

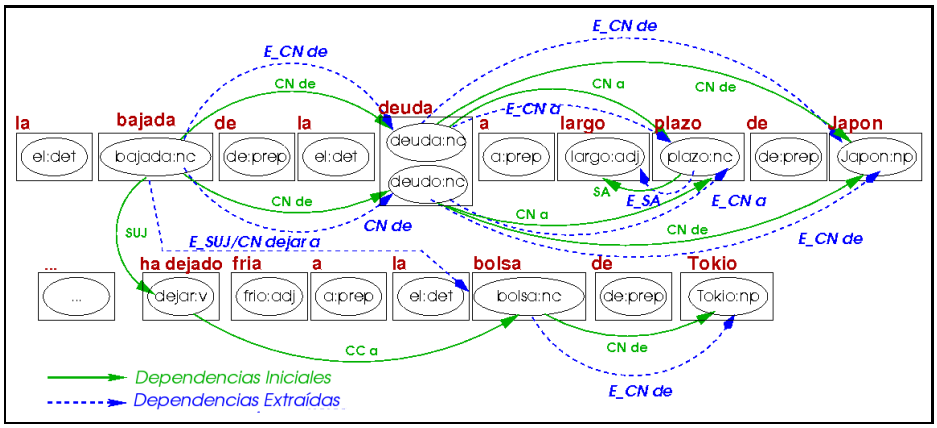


Figura 1: Grafo de dependencias del análisis sintáctico

Introducimos la notación y nomenclatura, a través de un ejemplo que utilizaremos en el resto de este apartado. Tomamos la frase *"La bajada de la deuda a largo plazo de Japón ... ha dejado fría a la bolsa de Tokio"*, de la Figura 1. Las formas rectangulares son llamadas *grupos* que indican las posiciones que ocupan los distintos elementos en la frase. Los lemas junto con su correspondiente categoría gramatical se representan mediante elipses llamadas *nodos*. Con líneas continuas se representan dependencias binarias entre palabras a través de alguna construcción sintáctica.

Tratamos de llegar a la fase de análisis semántico sin descartar ninguna posibilidad a dicha interpretación. Ello permite generar una estructura semánttica sobre la que aplicar, sin restricciones previas, procesos de adquisición del conocimiento que supondrían un filtrado de interpretacions inviables o inadecuadas. Así, la ambigüedad léxica ilustrada en la Figura 1 en el grupo correspondiente a la palabra *"deuda"*, debería decidirse en favor de la primera alternativa porque tenemos la certeza intuitiva de que *"deuda"* se refiere a la obligación que una persona tiene de pagar un dinero y tal situación no se refiere a un familiar. Al trabajar con un corpus especializado, ésto podrá confirmarse explorando el mismo en profundidad. Ello conlleva la consideración de un proceso iterativo convergente en un punto fijo en el que cada paso de la cadena secuencial re-inyecta en la siguiente iteración la información previa obtenida en forma de peso probabilístico [2].

Nuestro interés se centra por una parte en las dependencias entre sustantivos y adjetivos. Esto justifica el hecho de guardar las dependencias con líneas discontiuas mostradas en la Figura 1. Así, la palabra *"plazo"* (sustantivo) se relaciona con *"largo"* (adjetivo). Por otra parte, centramos nuestro interés en la extracción de dependencias entre sustantivos unidos por una preposición como *"bolsa de Tokio"* o unidos por un verbo, como en la dependencia *"bajada dejar_a bolsa"*, extraída de dos relaciones: *"bajada"* con *"dejar"* y *"dejar"* con *"bolsa"*.

El algoritmo intenta eliminar este tipo de ambigüedades añadiendo una fase de simplificación que aplica una restricción sintáctica simple, a saber, en una frase una palabra dada sólo puede tener a otra como gobernante. Así vemos en el ejemplo de la Figura 1, que *"Japón"* es gobernado por *"deuda"* (obligación de pago de dinero) y por *"deuda"* (familiar o pariente). Es necesario eliminar una de estas dependencias. No se considera ninguna otra restricción topológica y por consiguiente una palabra gobernante puede tener más de un gobernado, como el caso de "bajada" como gobernante de *"deuda"* y de *"bolsa"*. Además, una palabra puede ser a la vez gobernante y gobernado, como en *"plazo"*; gobernante de *"largo"*, pero también gobernada por *"deuda"*.

2.2.2. Agrupando los términos

La generación de clases semánticas se inspira en una propuesta de exploración de errores[4] originalmente diseñada para identificar la información errónea en entornos de análisis sintáctico [7]. Esta técnica combina dos procesos iterativos complementarios. Dada una iteración, el primero calcula, para cada par gobernante/gobernado en una frase, la probabilidad de la correspondiente dependencia. El segundo proceso calcula, a partir del primero, la clase semántica más probable para ser asignada a los términos de la dependencia. Así en cada iteración se procede a una desambiguación tanto a nivel sintáctico como a nivel semántico.

Ilustramos la agrupación de términos en nuestro ejemplo de ejecución de la Figura 1, centrándonos en la dependencia etiquetada como [_CNde] que se refiere a *"deuda"* y a *"Japón"*. Introducimos ambos procesos

[4] en terminología anglosajona, *error mining*.

Cuadro 2: Extracción de clases para *"deuda de Japón"*

$$1.\quad P(\text{deuda:uc:moneda, [_CNde], Japón:up:países})_{\text{local(0)}} = \frac{P(\text{deuda:uc, [_CNde], Japón:up})_{\text{local(0)}}}{\Sigma_{X,Y}\, P(\text{deuda:uc:X})_{\text{local(0)}}\; P(\text{Japón:up:Y})_{\text{local(0)}}}$$

$$2{,}1\quad P(\text{deuda:uc:monedas, [_CNde]}, X)_{\text{global(n+1)}} = \frac{\Sigma_X\, P(\text{deuda:uc:monedas,[_CNde]},X)_{\text{local(n)}}}{\#\text{dep}_{\text{local(n)}}(\text{deuda})}$$

$$2.\quad 2{,}2\quad P(Y, \text{[_CNde], Japón:up:países})_{\text{global(n+1)}} = \frac{\Sigma_Y\, P(Y,\text{[_CNde],Japón:up:países})_{\text{local(n)}}}{\#\text{dep}_{\text{local(n)}}(\text{Japón})}$$

$$2{,}3\quad P(\text{deuda:uc:monedas, [_CNde], Japón:up:países})_{\text{global(n+1)}} = P(\text{deuda:uc:monedas, [_CNde]}, X)_{\text{global(n+1)}}\; P(Y, \text{[_CNde], Japón:up:países})_{\text{global(n+1)}}$$

$$3.\quad P(\text{deuda:uc:monedas, [_CNde], Japón:up:países})_{\text{local(n+1)}} = \frac{P(\text{deuda:uc:monedas, [_CNde], Japón:up:países})_{\text{local(n)}}\; P(\text{deuda:uc:monedas, [_CNde], Japón:up:países})_{\text{global(n+1)}}}{\Sigma_{X,Y}\quad P(\text{deuda:uc:X, [_CNde], Japón:up:Y})_{\text{local(n)}}\; P(\text{deuda:uc:X, [_CNde], Japón:up:Y})_{\text{global(n+1)}}}$$

iterativos para este caso particular y hablaremos sin distinción de peso, probabilidad o preferencia para referirnos al mismo concepto.

1. Empezamos calculando la probabilidad local de la dependencia en cada frase, que depende del peso de cada palabra que a su vez depende de que su categoría léxica sea la correcta. El peso inicial asignado a una categoría es calculado por el algoritmo de exploración de errores. También tenemos en cuenta la probabilidad inicial de la dependencia observada que es una simple proporción de todas las posibles dependencias implicadas en las categorías consideradas. La normalización se calcula en base a la preferencia por una posible categoría léxica de los términos implicados.

2. Re-inyectamos localmente en la frase las probabilidades calculadas sobre el corpus en conjunto, para así re-calcular el peso de todas las posibles dependencias, después de lo cual podremos estimar globalmente las más probables. La normalización viene dada por el número de dependencias que conectan los términos considerados.

3. El valor local en la nueva iteración debería tener en cuenta tanto las preferencias globales como la re-incorporación local de esas preferencias en la frase. La normalización se calcula a partir del peso local previo y del peso global de la dependencia actual, teniendo en cuenta todas las posibles categorías léxicas asociadas con cada uno de los términos considerados.

Con respecto a la asignación de clases semánticas, la secuencia de pasos se puede ver en el Cuadro 2, que muestra la probabilidad de que la palabra *"deuda"* se refiera al grupo de `Monedas` y *"Japón"* al grupo de `Países`. Las clases `Monedas` y `Países` se definen *a priori* en una lista de clases semánticas:

1. En cada frase se calcula la probabilidad local de la dependencia para saber si sus elementos pertenecen a una u otra clase. Iniciamos el proceso a partir del peso local calculado y teniendo en cuenta la preferencia inicial de que los términos envueltos se correspondan a las clases consideradas[5]. La normalizacion viene dada por las probabilidades de todas las posibles clases incluyendo cada uno de los términos considerados que aquí se representa por las variables X e Y.

2. Entonces calculamos este peso a nivel global, reinyectando el resultado obtenido a nivel de corpus, para así poder re-calcular los pesos de todas las posibles clases en la frase. Esto se consigue calculando primero la probabilidad a nivel de corpus (2.1 y 2.2) para cada término y clase semántica, sin tener en cuenta ni el contexto izquierdo ni el derecho, representado por las variables X e Y respectivamente. La probabilidad final (2.3) es una combinación de las dos anteriores.

3. Después de cada iteración re-introducimos el valor global previo para obtener un nuevo valor local. La normalización se hace sumando las preferencias correspondientes a los términos y clases implicadas en la dependencia, para todas las posibles clases semánticas consideradas.

Una vez aplicadas éstas dos aproximaciones es posible construir una jerarquía teniendo en cuenta los elementos obtenidos en cada clase. Esta jerarquía, como en

[5]Las probabilidades se obtienen como una proporción de clases consideradas. Si el término se encuentra en una lista asociada, éste parte de una probabilidad mayor.

el caso de los términos, es almacenada en un fichero OWL.

3. CONOCIMIENTO Y RI

Las corrientes de investigación actuales en el campo de la RI centran sus esfuerzos en la mejora de los procesos de indexación y en la gestión inteligente de las consultas. El uso de diferentes tipos de estructuras de conocimiento tales como jerarquías u ontologías brindan claras ventajas para este cometido [5]. La cobertura y la precisión mejoran y las peticiones del usuario pueden concretarse de forma automática.

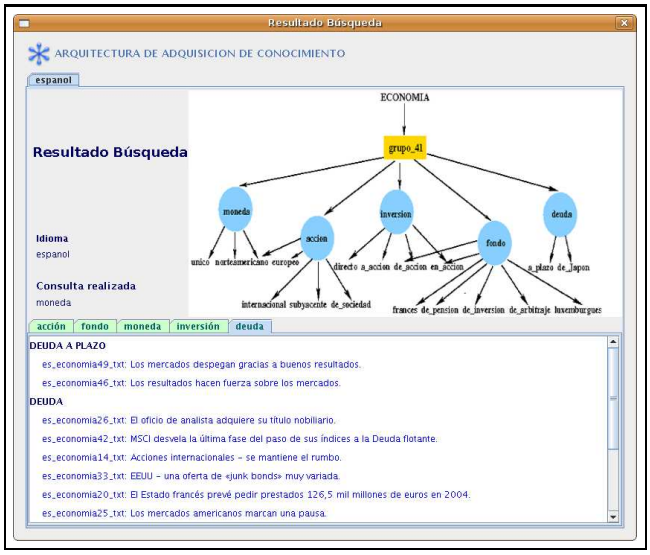


Figura 2: Sub-jerarquía para la consulta *"acción"* usando una estrategia basada en términos

3.1. PROPUESTA DE RI

Es necesario plantear nuevos métodos a la hora de crear y manipular los índices, haciéndolos más flexibles y ricos semánticamente. Esto debería extenderse al tratamiento de consultas, pues a menudo el usuario no tiene una idea exacta de la información contenida en la colección documental, lo que condiciona fuertemente sus interrogaciones. Una estructura jerárquica permite guiar las dos tareas aquí tratadas, el proceso de recuperación y la manipulación de consultas.

Nuestro acercamiento se apoya en una jerarquía de conceptos construida a partir de las relaciones semánticas que emergen del texto. Una herramienta de visualización permite la exploración del espacio de información; acadando así un entorno adecuado para la presentación de los documentos relevantes a una consulta. El punto de partida de esta propuesta es el de aprovechar, en el momento de la recuperación, la vinculación existente entre la jerarquía de conceptos y los textos.

El corpus tratado genera una estructura conceptual organizada alrededor de clases. A través de la proyección de las consultas se obtiene una conexión automática entre la estructura y la colección documental. El todo se apoya en una interfaz de visualización que vemos en las Figuras 2 y 3, donde se muestran la sección de la jerarquía relacionada con una determinada petición y en la parte inferior los documentos seleccionados.

3.2. PLANTEAMIENTO PRÁCTICO

En la práctica nos interesa analizar cómo es posible, partiendo de la adquisición de conocimiento, explotar las estructuras creadas y mejorar así los resultados obtenidos en la recuperación de documentos. En los apartados previos se han descrito dos técnicas de adquisición de conocimiento, una basada en términos y otra basada en dependencias ambas se integran para delimitar un marco de evaluación.

El motor de búsqueda utilizado es LUCENE[6], que asume la creación del índice una vez finalizado el análisis sintáctico. La jerarquía de conceptos entra en juego en la fase de consulta, en lo que hemos denominado *fase de expansión*. Ésta consiste en comparar la consulta del usuario con las estructuras del grafo conceptual y descubrir así la existencia de elementos relacionados con la petición. En caso afirmativo, estos elementos se toman de la jerarquía y se buscan en el índice, recuperando los documentos relevantes.

En caso de introducir una petición con más de una palabra, el proceso se lanza sin intervención de la jerarquía. Una consulta de este tipo es filtrada por una lista de parada y a continuación pasada al motor de búsqueda. Las peticiones que incluyan operadores lógicos, como AND y OR, son tratadas directamente por LUCENE. En cualquier otro caso, en un primer momento se lleva a cabo la expansión de la consulta haciendo uso de la estructura conceptual para, seguidamente, pasar los resultados al índice.

Una vez introducida una petición, el sistema busca la clase correspondiente en la estructura de conocimiento y recupera los conceptos asociados con esa clase. Esta función se ejecuta indistintamente de si estamos operando sobre la jerarquía de términos o sobre la jerarquía de dependencias. Sin embargo, los resultados varían en función de la estrategia elegida, pues el número de relaciones semánticas implicadas cambiarán si estamos considerando una u otra aproximación. Para ilustrarlo, describiremos la respuesta del sistema a la consulta *"acción"* planteada para ambas estrategias.

Nos centramos en la Figura 2 resultado de la consulta

[6] ver http://lucene.apache.org/

usando la ontología por términos. Vemos que *"acción"* pertenece a la misma clase que *"deuda"*, *"moneda"*, *"fondo"* e *"inversión"*. Estos elementos son cabezas de términos y se representan en un círculo. Sus expansiones vienen dadas por las flechas, como en *"acción de sociedad"*. El cuadro da nombre a la clase, un nombre genérico y asignado automáticamente.

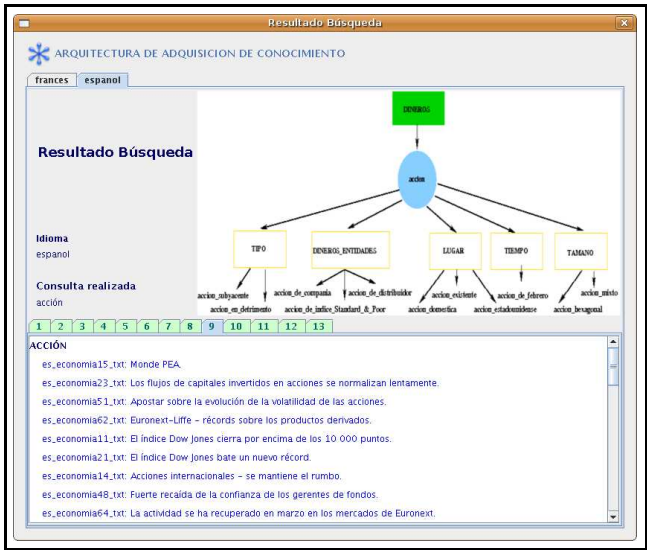


Figura 3: Sub-jerarquía para la consulta *"acción"*, usando una estrategia basada en dependencias

Nos centramos, ahora, en la Figura 3 resultado de la consulta usando la ontología por dependencias. Se definen dos tipos de clases: `conceptos` y `propiedades`. Las propiedades se consideran características de los conceptos. La interpretación es la siguiente: el gobernante es la consulta actual y está representado por un círculo. El cuadro superior indica la clase a la que pertenece el gobernante, en este caso es un concepto, `Dineros`. Los rectángulos representan el tipo de relación que se establece entre el gobernante y sus gobernados. Así, ambos pueden ser conceptos como *"acción"* y *"Standard and Poor's"* donde la relación tiene por nombre `Dineros_Entidades`. El gobernante puede ser un concepto y el gobernado una propiedad como en *"acción subyacente"*. Esta interpretación se puede extender al caso en el cual el gobernante es una propiedad y el gobernado un concepto.

4. CONCLUSIÓN

Hemos descrito una estrategia en el ámbito de la RI, basándonos en una indexación inteligente que se beneficia de las relaciones semánticas entre los conceptos presentes en la colección de textos. En contraste con acercamientos previos, elegimos una técnica que implica la ejecución de un número importante de tareas en modo no supervisado. Presentamos un conjunto de ejemplos en el campo de la economía.

Generamos de forma dinámica una estructura de conceptos que sirve de base para el módulo de RI, aproximación interesante para explorar nuevos dominios de conocimiento. Se trata de eliminar, en lo posible la traza de subjetividad introducida por el programador, reduciendo el factor humano en la toma de decisiones. Se procede, además, a un refinamiento de las consultas en base a la estructura conceptual.

Agradecimientos

Trabajo parcialmente subvencionado por el MEC en el proyecto HUM2007-66607-C04-02 y por la Xunta de Galicia en los proyectos PGIDIT05PXIC30501PN, 07SIN005206PR y la Red Gallega para el Procesamiento del Lenguaje y la RI.

Referencias

[1] Aussenac-Gilles, N. and Mothe J. 2004. Ontologies as background knowledge to explore document collections. In *RIAO*, pp. 129–142.

[2] Fernández M., E. de la Clergerie and Vilares M. Knowledge acquisition through error mining. 2007 In *RANLP, Proceedings*, pp. 220–224.

[3] Grefenstette, G. 1994. Explorations in Automatic Thesaurus Discovery. isbn: 0792394682

[4] Harris, Z.S. 1968. *Mathematical Structures of Languages*. J. Wiley & Sons.

[5] Masolo, C. 2001. Ontology driven information retrieval. Report of the IKF, 3180:371–380

[6] Petersen, W. 2001. A set-theoretical approach for the induction of inheritance hierarchies. *Electr. Notes Theor. Comput. Sci.*, 53.

[7] Sagot, B. and É. Villemonte de La Clergerie. 2006. Error mining in parsing results. In *Proc. of the 21st Int. Conf. on Computational Linguistics*, pages 329–336.

[8] Vilares, J., M.A. Alonso, and M. Vilares. 2004. Morphological and syntactic processing for text retrieval. *LNCS ISSN:0302-9743*, 3180:371–380.

Interpretando Información en Lenguaje Natural como Restricciones Generalizadas

Andrés Soto [1, 2] **José A. Olivas** [2] **Manuel E. Prieto** [2]

[1] Dpto. Ciencias de la Computación, Universidad Autónoma del Carmen,
Ave. 31, Ciudad del Carmen, CP 24160 Campeche, México.
email: soto_andres@yahoo.com
[2] Dpto. de Tecnologías y Sistemas de Información, Universidad de Castilla La Mancha,
Paseo de la Universidad 4, CP 13071 Ciudad Real, España.
email: JoseAngel.Olivas@uclm.es, Manuel.Prieto@uclm.es

Resumen

El lenguaje natural (LN) es básicamente un sistema para describir percepciones, las cuales son intrínsecamente imprecisas. Para hacer frente a esta situación Zadeh plantea un nuevo enfoque denominado NL-Computation (Cómputo en LN), el cual emplea herramientas tales como las restricciones generalizadas (RG) y las formas prototípicas (FP). Dada una proposición en LN, si ésta se puede expresar mediante RG, entonces se puede considerar precisa, al menos en cierto grado.
En este trabajo, se analizan diferentes estructuras del LN tales como frases nominales y oraciones copulativas, las cuales expresan restricciones. Se propone el uso de las RG para expresarlas. De esta manera, mucha de la información existente en documentos en LN podría ser convertida en RG y, luego, usada para deducir nueva información.

Palabras Clave: restricciones generalizadas, lenguaje natural, formas prototípicas.

1 INTRODUCCION

La Web es hoy en día el mayor repositorio de información que ha existido jamás en la historia de la Humanidad con millones y millones de megabytes de información. Desde este punto de vista, la Web pudiera considerarse como una gigantesca base de datos con un potencial increíble de información disponible para todo aquel que tenga acceso a la misma. Descontando el aspecto de la credibilidad que puede tener dicha información, el mayor impedimento en la actualidad es la dificultad para acceder a la información relevante de acuerdo con las necesidades de cada usuario.

Los buscadores existentes son capaces de recuperar millones de páginas en unos pocos segundos, pero muchas de dichas páginas no son relevantes para el usuario que realizó la consulta. La mayoría de los buscadores como Google utilizan algún tipo de técnica de comparación de términos para hallar las páginas relevantes. Pero dichos buscadores solamente tienen en cuenta los aspectos sintácticos de la información contenida en dichas páginas, tales como el número de veces que un término aparece en una página o aspectos estadísticos como el número de enlaces que la página recibe. Sin embargo, la relevancia es una relación entre un documento y una persona que busca información: dicha relación se cumple si el contenido al que hace referencia el documento (i.e. su significado, el sentido de la información que contiene) se corresponde con la información que la persona está buscando. De ahí que la relevancia esté más vinculada con la semántica, lo que trae consigo que muchas de las páginas que se recuperan por los buscadores son de poca utilidad al usuario.
Aún más limitada es la capacidad de los buscadores actuales para sintetizar o deducir información a partir de la información ya existente en la Web. Existen diferentes obstáculos para que los buscadores no puedan realizar deducciones, pero el problema fundamental es la dificultad para que los ordenadores puedan entender el lenguaje natural (LN), el problema básico de la imprecisión del significado. Para hacer frente a dichos obstáculos, Zadeh [19] propone algunas nuevas herramientas tales como: las Restricciones Generalizadas (RG) y la Teoría de Formas Prototípicas (FP). Dichas herramientas forman parte de lo que Zadeh [18] ha denominado Cómputo con Información descrita en Lenguaje Natural (Computation with Information Described in Natural Language) o también Cómputo en Lenguaje Natural o Cómputo LN. Dicha rama de la computación está llamada a tener una importancia enorme por cuanto la mayor parte del conocimiento humano está expresado en LN.

Para que los ordenadores puedan entender la información expresada en LN se requiere hacer frente a la imprecisión inherente a las proposiciones extraídas de dicho lenguaje. La solución que Zadeh [19] plantea para este problema es que la proposición debe estar expresada en un lenguaje matemáticamente bien definido, utilizando las restricciones generalizadas (RG), lo cual equivaldría a estar expresada de una forma precisa. En este sentido, el concepto de restricción generalizada puede considerarse como un enlace entre los LN y las matemáticas.

Zadeh [20] define como Precisiable Natural Language (LNP en español) al subconjunto del LN compuesto por todas aquellas proposiciones del mismo que pueden ser precisadas. En LNP, una percepción es equivalente a su descripción en LN.

Una forma prototípica o protoform [16] puede considerarse como un resumen abstracto o expresión simbólica, la cual define la estructura semántica esencial de las proposiciones, preguntas, órdenes, etc. El conjunto de todas las proposiciones del LN que admiten ser expresadas de forma precisa, junto con las reglas que gobiernan la propagación de las restricciones generalizadas constituye el llamado Lenguaje de Formas Prototípicas. El Lenguaje de Formas Prototípicas surge a partir de abstraer los aspectos fundamentales de los elementos del Lenguaje de Restricciones Generalizadas (LRG)

La idea básica propuesta por Zadeh [16] es la siguiente: dada una descripción de una percepción (i.e. una proposición p) en LN, se transforma la misma en una restricción generalizada RG(p). Luego, dicha restricción generalizada se transforma en una protoform Pf(p) en forma de una abstracción de RG(p). Posteriormente, utilizando las reglas deductivas asociadas a las formas prototípicas, sería posible deducir nueva información relevante.

Para poder aplicar las ideas de Zadeh es necesario identificar las proposiciones en LN y transformarlas en RG. Pero, lo que no está claro es cómo transformar las proposiciones en LN a RG, por lo cual se requiere desarrollar métodos y algoritmos que permitan realizar dicha transformación.

Desde hace años se ha planteado la posibilidad de usar técnicas de procesamiento de lenguaje natural para reconocer patrones lingüísticos con información sintáctica y semántica para mejorar la precisión en sistemas de Recuperación de Información [4, 6, 7, 14, 15]. El grupo SMILe de la Universidad de Castilla La Mancha ha desarrollado distintos trabajos [2, 10, 13] con el objetivo de utilizar aspectos semánticos para mejorar la precisión en sistemas de Recuperación de Información.

A partir de las experiencias antes mencionadas, en este trabajo se concibe la posibilidad de considerar que los adjetivos, al relacionarse con un nombre en una frase nominal, restringen el significado de dicho nombre, en la misma medida en que lo describen [11]. Otro tanto puede plantearse con respecto a la función del adverbio como modificador de un adjetivo [12]. Por tanto, en este trabajo se plantea que las frases nominales actúan como restricciones, especificando características del nombre que constituye la cabeza de dichas frases. Combinando este enfoque con las ideas de Zadeh antes mencionadas, se plantea también que, reconociendo las frases nominales existentes en un documento en LN y convirtiéndolas en RG, es posible extraer gran cantidad de información de dichos documentos que, al expresarla en forma de RG, sería posible utilizarla para deducir nueva información no existente previamente en dichos documentos. Avanzando un paso más en este enfoque, las oraciones copulativas y comparativas son consideradas como relaciones que restringen el significado del sujeto de la oración y que, por tanto, pueden ser transformadas a su vez en RG y operar con ellas. Los resultados obtenidos con este enfoque se presentan a continuación.

2 INTERPRETANDO INFORMACION EN LENGUAJE NATURAL

Zadeh plantea en [16, 17] que un LN es básicamente un sistema para describir percepciones, las cuales son intrínsecamente imprecisas, reflejando la limitada habilidad humana para resolver y almacenar información detallada. Por ejemplo, la proposición *p*: *Maria es joven* (*Mary is young*) expresa que la edad de *Maria* tiene un valor que la caracteriza como *joven*. Por tanto, independientemente de que su edad no es conocida con precisión, sabemos que la misma está comprendida entre un conjunto de valores que caracteriza a las personas jóvenes. De esta manera, el significado de *joven* es claro, aunque la edad no se conoce con precisión.

Las Gramáticas Libres de Contexto (*Context-Free Grammar*) constituyen el sistema matemático más utilizado para modelar la estructura de las oraciones en lenguaje natural. Con ayuda de analizadores lexicográficos y sintácticos [1, 9] es posible procesar documentos en LN y expresar la estructura sintáctica de una oración mediante el árbol sintáctico correspondiente.

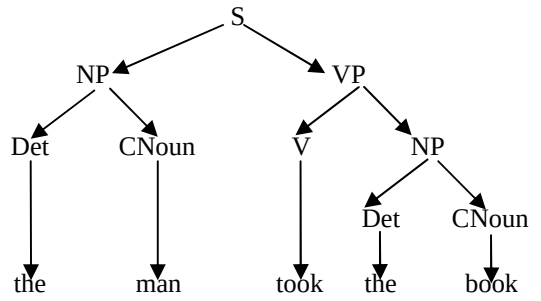


Figura 1 Árbol sintáctico de una oración

Otra forma de representar los árboles sintácticos es a través de la denominada notación lineal entre corchetes como la que se muestra a continuación:

[S [NP [Det the] [CNoun man]]
[VP [V took] [NP [Det the] [CNoun book]]]]

En el grupo SMILE se han desarrollado programas para la detección de frases nominales y oraciones en documentos en LN con resultados satisfactorios: más de 70% de éxito en el reconocimiento de FN. En el resto del trabajo se considerará que con ayuda de uno de estos programas es posible disponer de la notación lineal entre corchetes de cada oración.
Las palabras, frases, cláusulas y oraciones constituyen una jerarquía gramatical[1] (ver Tabla 1).

Tabla 1 Jerarquía gramatical

Oraciones: constan de una o más…
Cláusulas: constan de una o más…
Frases: constan de una o más…
Palabras

Las frases son grupos de palabras unidas gramaticalmente sin incluir sujeto, ni predicado, a diferencia de las cláusulas que sí incluyen ambos. Las frases están formadas por, al menos, una palabra principal o elemento cabeza (*Head*). Dicha cabeza puede ir precedida por un grupo de palabras modificadoras que forman la cadena previa *pre-Head*. Y puede ir seguida a su vez de otro grupo modificador posterior, *post-Head*. Dichas cadenas pueden constituir nuevas frases o incluso cláusulas (ver Tabla 2). Según el tipo de palabra que constituye el elemento cabeza de la frase, las mismas se pueden clasificar en diferentes tipos (ver Tabla 3).

Tabla 2 Estructura de una frase nominal

pre-Head	**Head**	**post-Head**
[NP *the small*	*children*	*at the window*]

Tabla 3 Tipos de frases

Tipo de frase	**Abreviatura**
Nominal	NP
Verbal	VP
Adjetival	AP
Adverbial	AdvP
Preposicional	PP

Las frases nominales son unidades sintácticas que aparecen en la mayoría de las oraciones y expresan restricciones con respecto al nombre al cual van asociadas y dan información acerca del mismo. Son las más frecuentes, pudiendo aparecer desempeñando diversas funciones en la oración, tanto en el sujeto como en el predicado. Como parte de los modificadores previos del nombre pueden aparecer:

- Determinante (*determiner*[2]): Artículos (*the*, *a*), demostrativos (*this*, *that*), numerales (*two*, *five*, etc.), posesivos (*my*, *their*, etc.), cuantificadores (*some*, *many*, etc.)
- Adverbios (*a* ***very*** *young man*)
- Adjetivos (*the* ***red*** *ball*),
- Frases preposicionales (*the student* ***of physics***),
- Cláusulas that (*the claim* ***that the earth is round***);

Por otra parte los modificadores que suceden a la cabeza (i.e. al nombre) pueden ser frases preposicionales (*the man* ***with long hair***) o cláusulas relativas (*the house* ***where I live***). De esta manera, el ejemplo de la Tabla 2 quedaría:

[NP [Det *the*] [Adj *small*] [CNoun *children*]
[PP [Prep *at*] [NP [Det *the*] [CNoun *window*]]]
]

Utilizando la Programación Orientada a Objetos se desarrolló un programa que procesa las oraciones expresadas en forma de árbol sintáctico y las transforma en una estructura correspondiente representada mediante objetos y clases. En este sentido, se definieron clases de tipo oración, frase, cláusula. En el caso de las frases, éstas tienen componentes tales como *head*, *prehead* y *posthead* correspondiéndose así con la estructura general de una frase. A partir de esta definición general se definieron subclases para los diferentes tipos de frase: *noun phrase*, *verb phrase*, etc. De esta manera, procesando la notación lineal del árbol sintáctico de una oración es posible transformarlo en un objeto oración compuesto a su vez por diferentes objetos de tipo frase o cláusula, incluyendo el caso de oraciones compuestas y complejas. Por tanto, en este trabajo se plantea el enfoque de transformar de esta manera el documento original en lenguaje original en una cadena o sucesión de objetos de tipo oración, en correspondencia con los existentes en el documento original.
Posteriormente se desarrolló un intérprete que procesa los objetos oración (las oraciones representadas mediante objetos) uno a uno, sucesivamente. Producto de la interpretación se van creando otros objetos que se corresponden con las diferentes entidades a las cuales se hace referencia en las frases nominales (FN) del documento. Estos objetos-entidades tienen atributos que se definen mediante el proceso de interpretación de las oraciones, en particular de las FN.

[1] KS3 Grammar, Phonetics & Linguistics Dept., University College London, 2008, http://www.phon.ucl.ac.uk/home/dick/tta/KS3.htm

[2] Los lingüistas hoy en día diferencian los adjetivos de los determinantes, considerándolos como dos categorías lexicológicas diferentes. Un determinante es una palabra que expresa la referencia a un nombre en el contexto indicando su definición, cantidad u otra propiedad similar. Mientras que un adjetivo es una palabra que modifica a un nombre o pronombre, ampliando la información que se tiene sobre el mismo.

3 INTERPRETANDO LAS FRASES NOMINALES COMO RESTRICCIONES

Cada una de las cosas existentes en el mundo (e incluso las imaginarias) posee una serie de características que la definen como tal. Dichas características o atributos tales como tamaño, talla, color, forma, contorno, sabor, consistencia, etc. tienen valores lingüísticos asociados que permiten especificar y diferenciar las diferentes entidades existentes. Cada una de estas características define una dimensión en la cual se expresa dicha característica. Algunas características se tienen o no se tienen, mientras que otras se pueden tener en un grado mayor o menor.

Los nombres[3] son palabras que se usan para designar cualquier realidad, concreta o abstracta, y que sirven para referirse a ella, reconocerla y distinguirla de otra. Los nombres comunes se utilizan para designar agrupaciones de personas, animales o cosas que pertenecen a una misma clase, especie o familia, significando su naturaleza o sus cualidades; p. ej., *naranja* es un nombre común, que se aplica a todos los objetos que poseen las propiedades de forma, color, olor, sabor, etc., que distinguen a una naranja de cualquier otra cosa[4]. Los nombres comunes representan una fuerte tendencia para que ocurran de manera simultánea una serie de valores particulares correspondientes a diferentes dimensiones, los cuales permiten caracterizar entidades reales o abstractas. El nombre *naranja* en sí mismo no designa ninguna naranja en particular sino toda una categoría de objetos que comparten una serie de atributos comunes. En otras palabras, la categoría *naranja* es una agrupación de características que ocurren de manera simultánea. Por otra parte, los nombres propios se utilizan para designar y diferenciar a unos seres de otros de su misma especie.

En una frase nominal como *naranja jugosa* (*juicy orange*), la palabra *jugosa* es un atributo que caracteriza algunos miembros de la categoría *naranja*, indicando un valor posible en una de las dimensiones conceptuales que definen dicha categoría. La frase le atribuye la característica de tener jugo a la naranja en cuestión. La palabra *jugosa* actúa como adjetivo para limitar o completar el significado del sustantivo al que acompaña y define una subcategoría, *naranja jugosa*, de la categoría *naranja* [3]. Los adjetivos que aparecen antes del nombre como parte de la FN se denominan atributivos.

Por otra parte, los adverbios modifican a los adjetivos sobre los que actúan en las FN. Los adverbios *very* y *extremely* se denominan adverbios de grado o nivel, por cuanto especifican el grado de aplicación del adjetivo asociado. El valor expresado por el adjetivo (i.e. *young*) se transforma en un nuevo nivel (i.e. *very young*) mediante el uso del adverbio.

De acuerdo con lo anterior, cada vez que una FN es interpretada, se define un objeto de tipo entidad sobre el cual recae el efecto de la interpretación. Según se van creando, los objetos se van colocando secuencial y ordenadamente en un área de trabajo o módulo. Si la FN hace referencia a un objeto que ya había sido creado (i.e. la cabeza de la FN está constituida por un pronombre como *she* o *he*), se seleccionará el primer objeto del módulo en orden inverso que corresponda en género y número con el pronombre. En el caso de un nombre común, se considerará que hace referencia a una clase de objetos (i.e. *girl*, *bean*, *car*) y que los adjetivos que lo acompañan identifican atributos de subclases de la misma (i.e. *young girl* es una subclase de *girl* que tiene el atributo *young*). En el caso de un nombre propio, se considerará por omisión que hace referencia a un objeto específico de la clase persona.

Por tanto, suponiendo que se tiene la FN de la Tabla 2, la presencia del determinante *the* presupone que la entidad ya había sido mencionada previamente y que, por tanto, debe corresponder con algún objeto existente en el módulo o área de trabajo. Si no se detecta (puede que el algoritmo no sea capaz de detectar a qué entidad se hace referencia), se crea una nueva entidad de tipo *children*, a la cual se le asocia el atributo *small* y se crea una relación de tipo *at* con la entidad *window* a la que se hace referencia.

4 INTERPRETANDO LAS ORACIONES COMO RESTRICCIONES

En inglés, los adjetivos pueden aparecer no sólo antes del nombre al que describen, sino después, en cuyo caso aparecerán después del verbo, como parte del predicado, haciendo la función de complemento en las oraciones copulativas. Estos adjetivos se denominan predicativos.

En las oraciones copulativas, el verbo no expresa en sí una acción o condición, si no que sirve para enlazar el sujeto con el predicado, en la forma de un complemento o una frase adverbial. Estos verbos se denominan "verbo de enlace" (*linking verb* o *copula verb*), de los cuales el más importante es el verbo "*to be*".

Desde cierto punto de vista, se puede interpretar que el verbo de enlace relaciona dos cosas como subconjuntos:

- *John is a doctor*: incluye al objeto *John* de la clase persona en la clase *doctor*.
- *Doctors are educated*: incluye la clase de los doctores en la clase de las personas educadas.
- *Mary is running*: el estado actual de *Mary*, *running*, pertenece al conjunto de los estados de movimiento.
- *Running is fun*: el elemento *running* perteneciente al conjunto de los estados de movimiento también pertenece al conjunto de los estados de diversión. (consecuentes con diversión).

Las oraciones copulativas son de las más frecuentes en inglés. De las 10 estructuras básicas para las oraciones simples que plantean Kolln & Funk [8], 5 corresponden a oraciones copulativas y 3 de ellas son específicas para el verbo "*to be*":

[3]Diccionario de la lengua española © 2005 Espasa-Calpe S.A., Madrid

[4]Diccionario de la RAE Esencial de VV.AA. Espasa-Calpe, S.A. Madrid 2006

- El complemento es un adjetivo: *The team is good*. Incluye a un objeto específico de la clase *team* en la subclase *team* con el atributo *good*.
- El complemento es un adverbio: *The car is outside*. Incluye a un objeto específico de la clase *car* en la clase de los objetos cuya posición es *outside*.
- El complemento es una frase nominal: *The Raiders is the winning team*. Incluye a un objeto que representa un conjunto específico de personas, identificado como *Raiders* (nombre propio) como elemento de la subclase *team* con el atributo *winning*.

Por tanto, si tenemos la oración *Mary is young*, la misma provoca que se verifique la existencia de un objeto con nombre *Mary* y devuelve un apuntador al mismo; si no existe, se crea. En ambos casos, se le asocia el atributo *young*.

verify_name(pn('*Mary*'), P1),
send(P1, set_attrib, adjp([adj(*young*)]))

Lo mismo sucederá si la frase adjetival es *very young*

send(P1, set_attrib,
adjp([advp(adv(*very*)), adj(*young*)]))

teniendo en cuenta que *young* (y *very young*) representan conjuntos borrosos, a los cuales pertenece el valor del atributo correspondiente (i.e. la edad) de *Mary* con un grado de pertenencia mayor que cero (0).

Otro tipo de oraciones muy importante comprende las oraciones comparativas formadas por dos cláusulas unidas por una fórmula comparativa definida por dos palabras (*Mary is more intelligent than Tom*). La estructura de una comparación consta de: la forma positiva (grado absoluto) del adjetivo o adverbio, más el sufijo *–er* (o precedida del modificador *more* o *less*). La comparación se completa con la preposición "*than*" seguida de la otra cláusula[5]

- FN *is more*/*less* Adj *than* FN. *The village is less picturesque than the town nearby*.
- FN *is as* Adj *as* FN. *It's as big as a house*
- FN *is not so/as* Adj *as* FN. *It's not so/as good as this one*
- FN *is the same* Adj *as* FN. *It's the same difficulty as the other one.*

Suponiendo que se tiene la oración *Tom is few years older than Mary*, la misma corresponde al tipo de estructura comparativa que se muestra en el esquema de la Tabla 4.

Tabla 4 Estructura de una oración comparativa

FN1	is	FN2	more	Adj	than	FN3
Tom	is	few years	more	old	than	Mary

donde por cada FN se obtendría un puntero a un objeto diferente: P1 y P3 a los objetos de tipo persona *Tom* y *Mary*, mientras que P2 apuntaría a un objeto de tipo *year* con atributo *few*. Como es una oración comparativa, para definir el atributo de *Tom* debo tener en cuenta el atributo correspondiente de *Mary*. Considerando que:

a) *old* es antónimo de *young*,
b) *old* y *young* son valores borrosos en la dimensión *time* tales que *young* < *old*
c) *years* es una unidad de medición de la dimensión *time*
d) *more* equivale en cierto sentido a más

entonces el atributo de *Tom* será *young* + *few years* y su grado de pertenencia al conjunto borroso *old* será mayor que el de *Mary*.

Otro ejemplo, más complejo, con la misma estructura (ver Tabla 5) es el siguiente: *This new fine car is several hundred dollars more expensive than the old one over there*.

Tabla 5 Ejemplo de oración comparativa

FN1	is	FN2	more	Adj	than	FN3
This new fine car	*is*	*several hundred dollars*	*more*	*expensive*	*than*	*the old green car over there*

donde por cada FN se obtendría un puntero a un objeto diferente: P1 y P3 a objetos de tipo *car*, P2 a un objeto de tipo *dollars* con diferentes atributos. Como es una oración comparativa, para definir el atributo *expensive* de P1 debo tener en cuenta el atributo correspondiente de P3. Considerando que:

a) *expensive* es un valor borroso en la dimensión *price*,
b) *dollars* es una unidad de medición de la dimensión *price*
c) *more* equivale en cierto sentido a más

entonces el atributo de P1 será *expensive* + *several hundred dollars* y su grado de pertenencia al conjunto borroso *expensive* será mayor que el de P3.

6 CONCLUSIONES

El enfoque denominado Cómputo en Lenguaje Natural y el uso de las Restricciones Generalizadas (RG) permite ampliar el papel del lenguaje natural en las teorías científicas [18]. En este trabajo se considera que las frases nominales (FN) y las oraciones copulativas y comparativas expresan restricciones entre los componentes de una oración y que pueden ser interpretadas a manera de RG, partiendo de su representación mediante árboles sintácticos.

En el trabajo se plantean algunas ideas básicas que se aplicaron para desarrollar un algoritmo que interpreta las FN y algunos tipos de oraciones, convirtiéndolas en objetos computacionales y relaciones entre dichos objetos. Los adjetivos y adverbios se interpretan como referencias a las características de las entidades a las cuales aluden los nombres que aparecen en las oraciones y se convierten en valores borrosos que describen los atributos de dichos objetos-entidades.

[5] Según los gramáticos, la comparación correcta sería: He is taller than his father is

También se plantea la posibilidad de utilizar las oraciones comparativas como relaciones de comparación entre atributos con valores borrosos y operar con los objetos a los cuales están asociados dichos atributos.

Los programas desarrollados para convertir los árboles sintácticos en estructuras basadas en objetos y para la posterior interpretación de dichas estructuras se han probado con ejemplos sencillos como los que aparecen en el trabajo.

Los resultados planteados, limitados hasta el momento, permiten adentrarse un paso más allá de esa nueva frontera que Zadeh [18] denominara Cómputo con Información descrita en Lenguaje Natural (Computation with Information Described in Natural Language)

Agradecimientos

Financiado parcialmente a través de los proyectos PAC06-0059 SCAIWEB, JCCM, y TIN2007-67494 F-META, MEC-FEDER, España.

Referencias

1. S. Bird, E. Loper. NLTK: The Natural Language Toolkit**.** *Proceedings of the* Association for Computational Linguistics *ACL demonstration session.* Pág. 214-217, Barcelona, 2004.
2. P. J. Garcés, J. A. Olivas, F. P. Romero. FIS-CRM: A Representation Model Based on Fuzzy Interrelations for Internet Search. *ICAI-02*, Pág. 219-224, 2002
3. M. Gasser, *How Language Works*, Edición 3.0, Indiana University, 2006.
4. S. Harabagiu, M. Surdeanu. Infrastructure for Open-Domain Information Extraction. In *Proceedings of Human Language Technology 2002*, *HLT'02* San Diego, California USA, 2002.
5. D. Jurafsky, J. H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, *Computational Linguistics, and Speech Recognition. Prentice Hall PTR*, Pág. 934, 2000.
6. B. Katz, J. Lin. REXTOR: A System for Generating Relations from Natural Language. *Proceedings of the ACL-2000 workshop on Recent advances in natural language processing and information retrieval: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics*, Vol. 11 Pág. 67-77, 2000.
7. B. Katz. Annotating the World Wide Web using Natural Language. *International Conference On Computational Linguistics***,** *Proceedings of the 2nd workshop on NLP and XML*, Vol. 17**,** Pág. 1-8, 2002
8. M. Kolln, R. Funk. Understanding English Grammar. *Prentice Hall, 5th edition*, 496 págs., 1997
9. N. Madnani. Getting Started on Natural Language Processing with Python**,** *ACM Crossroads*, Vol. 13, Issue 4, 2007.
10. J. A. Olivas, P. J. Garces, F. P. Romero. An application of the FIS-CRM model to the FISS metasearcher: Using fuzzy synonymy and fuzzy generality for representing concepts in documents. *Int. J. Approx. Reasoning 34, (2-3)*, Pág. 201-219, 2003
11. A. Soto, J.A. Olivas, M.E. Prieto. Using Generalized Constraints and Protoforms to Deal with Adjectives. *Proceedings IEEE International Conference on Fuzzy Systems, Imperial College, London, UK.* 2007.
12. A. Soto, J.A. Olivas, M.E. Prieto. Using Generalized Constraints and Protoforms to deal with adverbs. *In New Dimensions in Fuzzy Logic and Related Technologies. Proceedings of the 5th Conference European Society for Fuzzy Logic and Technology EUSFLAT 07, University of Ostrava, Ostrava, Czech Rep.* Pág. 119-127, 2007.
13. A. Soto, J. A. Olivas, M. E. Prieto. Fuzzy Approach of Synonymy and Polysemy for Information Retrieval. *Granular Computing: At the Junction of Rough Sets and Fuzzy Sets. Studies in Fuzziness and Soft Computing. Springer Berlin / Heidelberg*, Vol. 224, Pág. 179-198, 2008
14. J. Vilares, M. A. Alonso, F. J. Ribadas, M. Vilares. COLE experiments at CLEF 2002 Spanish monolingual track. In Carol Peters (ed.), *Results of the CLEF 2002 Cross-Language System Evaluation Campaign, Working Notes for the CLEF 2002 Workshop*, Pág. 153-160, Roma, Italia, 2002.
15. A. Voutilainen. NPtool a detector of English noun phrases *Proceedings of Workshop on Very Large Corpora* WVLC, Ohio State University, Pág. 48-57, 1993.
16. L. A. Zadeh, "A note on web intelligence, world knowledge and fuzzy logic", *Elsevier, Data & Knowledge Engineering* 50, Pág. 291–304, 2004.
17. L. A. Zadeh, "From Search Engines to Question-Answering Systems— The Problems of World Knowledge, Relevance, Deduction and Precisiation", Keynote Speech, *IRI 2005 IEEE International Conference on Information Reuse and Integration*, Las Vegas, Nevada, USA, 2005.
18. L. A. Zadeh, "A New Frontier in Computation – Computation with Information Described in Natural Language", *3rd IEEE Conference On Intelligent Systems*, University of Westminster, United Kingdom, Pág. 4-6, 2006.
19. L. A. Zadeh, From Search Engines to Question Answering Systems: The Problems of World Knowledge, Relevance, Deduction and Precisiation. En E. Sanchez (ed.): Fuzzy Logic and the Semantic Web, Elsevier 2006
20. L. A. Zadeh, and M. Nikravesh, "Perception-Based Intelligent Decision Systems", *ONR Summer 2002 Program Review*, University of California, Los Angeles UCLA, USA. 2002.

UN SISTEMA DE RECOMENDACIONES PARA TRATAMIENTOS DE LUMBALGIAS CRÓNICAS VIA WEB

Carlos Porcel[1] **Bernabé Esteban**[2] **Enrique Herrera-Viedma**[2]

[1] Universidad de Jaén, cporcel@ujaen.es
[2] Universidad de Granada, {bernabe@ugr.es,viedma@decsai.ugr.es}

Resumen

En este trabajo presentamos el diseño de un sistema de recomendaciones lingüístico difuso de asistencia al tratamiento y prevención de las lumbalgias crónicas. El sistema prestará asesoramiento y ayuda a los pacientes en los tratamientos de sus patologías, con el objetivo de mejorar su calidad de vida y fomentar entre los usuarios hábitos de vida saludables.

Palabras Clave: sistemas de recomendaciones, modelado lingüístico difuso, lumbalgias.

1 INTRODUCCIÓN

En las sociedades modernas, el dolor lumbar afecta a más de dos tercios de los adultos en algún momento de su vida [24]. Se trata, casi siempre, de episodios de dolor autolimitado, con tendencia a la mejoría completa de forma espontánea, progresiva y rápida (en días o en pocas semanas), pero se produce hasta en un 60% de recurrencias y el dolor se hace crónico en un 10%-30% de los casos [2].

En la práctica se han planteado una amplia y creciente variedad de tipos de cuidados sanitarios conservadores, usándose a menudo varias intervenciones simultáneas. Desde hace tiempo, se admite de forma consensuada, que el ejercicio es una terapia activa que desempeña un papel clave en el tratamiento del dolor lumbar mecánico inespecífico, encontrándose entre los tratamientos más prescritos en lumbalgias crónicas [9]. Esto obliga a individualizar las actuaciones sanitarias. Las modalidades terapéuticas pasivas tienen efecto sólo a corto plazo y, por tanto, su papel es limitado en el dolor lumbar crónico. La prevención del dolor lumbar es uno de los pilares fundamentales de su tratamiento [8]. Como consecuencia de la elevada incidencia de la patología lumbar, se han introducido programas de prevención en el ámbito laboral y social como una tentativa de disminuir su incidencia y a la vez, disminuir el absentismo y los costes económicos subsecuentes [27]. Estos programas de prevención incluyen la educación del individuo para ayudarle a conocer su problema, manejarlo y evitarlo. Dentro de los programas de prevención del dolor lumbar, el más extensamente utilizado con todas sus variantes, es la Escuela de Columna [21]. Aunque su uso es relativamente reciente, la gran preocupación ha sido siempre su efectividad real para cambiar los hábitos posturales y de vida de los individuos.

En este trabajo proponemos el diseño de una herramienta web que posibilite a los pacientes prevenir y/o seguir un tratamiento de la lumbalgia crónica desde su domicilio, puesto de trabajo o cualquier otro lugar. La herramienta se va a desarrollar integrando un sistema de recomendaciones híbrido [5, 10, 25] con el Modelado Lingüístico Difuso (MLD) multi-granular, para representar y gestionar la información de forma flexible mediante el uso de etiquetas lingüísticas [3, 12, 28]. En el sistema, los perfiles de usuario representan las patologías de los pacientes, las cuales son caracterizadas mediante unos parámetros obtenidos a raíz de la realización de una serie de tests [7]. Es decir, a la población de estudio se le realiza una serie de tests [7] y en función de los resultados obtenidos, se establecen en uno u otro grupo de diagnóstico. A cada uno de estos grupos de diagnóstico le corresponde un programa de trabajo específico, pero que se va personalizando en función de la patología del paciente. Entonces, el sistema recomienda a cada paciente un conjunto de ejercicios en función de la patología presentada. El paciente realizará durante un tiempo el tratamiento recomendado, y tanto él como los expertos, podrán ir realizando el seguimiento a través de la propia herramienta web. En la sección 2 presentamos los conceptos fundamentales sobre sistemas de recomendaciones y el

MLD. En la sección 3 se presenta el diseño del sistema. Por último, en la sección 4 apuntamos las conclusiones obtenidas.

2 PRELIMINARES

2.1 SISTEMAS DE RECOMENDACIONES

Para mejorar el acceso a la información a través de la Web, los usuarios necesitan herramientas para filtrar la gran cantidad de información disponible. Los sistemas de recomendaciones son herramientas útiles que ayudan a difundir la información exclusivamente a quiénes la necesitan, distinguiendo entre información relevante e irrelevante [5, 25].

Hay diversos tipos de sistemas de recomendaciones [5], en función del método de generación de recomendaciones que se use. Nosotros vamos a adoptar un enfoque híbrido, en el que para generar las recomendaciones se tienen en cuenta las características usadas para la representación de los tratamientos, las usadas para representar las patologías de los pacientes y las evaluaciones que los pacientes y expertos dan sobre los tratamientos seguidos.

Para mantener este esquema, la actividad de generación de recomendaciones es seguida por una fase de realimentación en la que los usuarios valoran la relevancia de las recomendaciones suministradas por el sistema, y el sistema usa estas evaluaciones para actualizar automáticamente los perfiles de los usuarios [10, 25].

Dada su utilidad, los sistemas de recomendaciones son una tecnología extensamente aplicada en muy diversos campos, como venta electrónica [6, 16], recomendaciones de películas [19, 22], música [18, 23], turismo [20], etc.

2.2 MODELADO LINGÜÍSTICO DIFUSO

Hay situaciones en las que la información no puede ser valorada precisamente de forma cuantitativa, pero podría serlo de forma cualitativa. El uso de la Teoría de Conjuntos Difusos ha dado muy buenos resultados para el modelado de información cualitativa [28] y se ha demostrado su utilidad en multitud de problemas [11, 14, 15]. Se trata de una herramienta basada en el concepto de *variable lingüística* [28].

2.2.1 Modelado lingüístico difuso 2-tuplas

El MLD 2-tuplas [12] es un modelo continuo de representación de la información. Para definirlo, tenemos que establecer el modelo de representación y el modelo computacional de las 2-tuplas para representar y agregar respectivamente la información lingüística.

Consideremos que $\mathcal{S} = \{s_0, ..., s_g\}$ es un conjunto de términos lingüísticos con cardinalidad impar, donde la semántica asociada con cada una de las etiquetas viene dada por medio de funciones de pertenencia triangulares y consideramos todos los términos distribuidos sobre una escala sobre la que hay establecida una relación de orden total. Si mediante un método simbólico de agregación de información lingüística, obtenemos un valor $\beta \in [0, g]$, y $\beta \notin \{0, ..., g\}$, podemos representar β como una 2-tupla (s_i, α_i), donde s_i representa la etiqueta lingüística, y α_i es un valor numérico que expresa la traslación de β con respecto al índice de la etiqueta más cercana, i, en el conjunto de términos lingüísticos ($s_i \in \mathcal{S}$). Este modelo define un conjunto de funciones de transformación entre valores numéricos y 2-tuplas: $\Delta(\beta) = (s_i, \alpha)$ y $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ [12].

Para establecer el modelo computacional, definimos operadores de negación, de comparación y de agregación. Usando las funciones Δ y Δ^{-1}, podemos extender cualquiera de los operadores de agregación ya definidos para trabajar con 2-tuplas [12].

2.2.2 Modelado lingüístico difuso multi-granular

En función del grado de incertidumbre que un experto encargado de cualificar un fenómeno tenga sobre el mismo, el conjunto de términos lingüísticos elegido para proporcionar ese conocimiento tendrá más o menos términos. Cuando distintos expertos tienen diferentes grados de incertidumbre sobre el fenómeno o cuando un experto tiene que valorar distintos conceptos, se necesitan varios conjuntos de términos lingüísticos de diferente granularidad [13]. En tales situaciones necesitamos herramientas que nos permitan gestionar la información lingüística multi-granular. En [13] se propuso un MLD 2-tuplas multi-granular basado en el concepto de jerarquía lingüística. Una *Jerarquía Lingüística, LH*, es un conjunto de niveles $l(t, n(t))$, donde cada nivel t es un conjunto de términos lingüísticos con una granularidad $n(t)$ diferente [13]. Los niveles están ordenados por granularidad, de manera que podemos definir un nivel a partir del anterior: $l(t, n(t)) \rightarrow l(t + 1, 2 \cdot n(t) - 1)$. En [13] se definió una familia de funciones de transformación entre etiquetas de diferentes niveles. Para establecer el modelo computacional seleccionamos un nivel que usamos para uniformizar la información y así podemos usar los operadores definidos en el modelo 2-tuplas.

3 DISEÑO DEL SISTEMA

A continuación presentamos el sistema de asistencia al tratamiento y prevención de las lumbalgias crónicas basado en tecnologías web [26], para prestar asesoramiento y ayuda a los pacientes en los tratamientos de sus patologías, con el objetivo de mejorar su calidad de vida. Para ello se actuará en tres frentes distintos:

1. Se desarrollará una base de datos multimedia con ejemplos de los tratamientos a desarrollar en cada patología.

2. Se desarrollará un sistema de recomendaciones de tratamiento personalizado [1], que usará como base de información la anteriormente mencionada y una base de datos de perfiles de usuario.

3. Se desarrollará una herramienta web para estudiar la evolución de los pacientes y la adherencia a los tratamientos asignados. Esta herramienta, nos permitirá poner en valor el sistema diseñado, así como la eficacia de la actividad desarrollada por el paciente. En definitiva, será una herramienta para evaluar la calidad de vida de los pacientes mediante técnicas de visualización de información, como por ejemplo los gráficos radiales, que fácilmente permiten actualizar la evolución de los pacientes de modo cualitativo [4]. Esta herramienta también nos permitirá realimentar el sistema, modificando los perfiles de usuarios a raíz de la evolución observada por los pacientes.

En la figura 1, podemos ver una representación del esquema de funcionamiento del sistema propuesto.

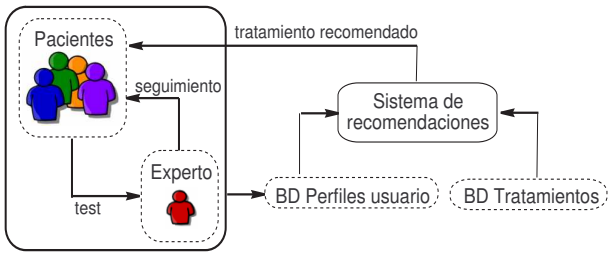


Figura 1: Diseño del sistema.

Para representar los distintos conceptos que tienen que ser valorados en los distintos frentes, usamos diferentes conjuntos de etiquetas (S_1, S_2, ...) elegidos de una LH. Vamos a distinguir tres conceptos: *grado de importancia* de cada subgrupo de diagnóstico para la aplicación de un tratamiento o para establecer el perfil de un paciente en función de la patología presentada a raíz de los tests realizados (etiqueta de S_1), *grado de relevancia* de un tratamiento para un paciente (etiqueta de S_2), y *grado de similitud* entre las patologías de dos pacientes (etiqueta de S_3). En concreto, usaremos etiquetas seleccionadas de una LH de 2 niveles (de 5 y 9 etiquetas), con $S_1 = S^5$, $S_2 = S^9$ y $S_3 = S^9$.

3.1 BASE DE DATOS DE TRATAMIENTOS

Los tratamientos son los items que va a recomendar nuestro sistema, así que un aspecto clave es obtener una representación adecuada de los mismos. Como cada tratamiento es adecuado para un subgrupo de diagnóstico con una patología determinada, usaremos dichos subgrupos para representar el tratamiento correspondiente. Inicialmente, hemos establecido tres subgrupos de diagnósticos, pues hay que tener en cuenta que partimos de pacientes con lumbalgia mecánica crónica, y por tanto, con un diagnóstico previo en el que se han analizado muchas otras características. Los tres subgrupos son los siguientes:

- Debilidad Muscular.
- Inestabilidad Lumbar.
- Alteraciones Biomecánicas.

A cada uno de los subgrupos le corresponde un tratamiento específico que consta inicialmente de 10 ejercicios definidos y de recomendaciones en cuanto a las actividades comunes, tales como manejo de cargas, sentarse, levantarse o higiene postural. Sin embargo, estos 10 ejercicios, se pueden intercambiar entre los tres subgrupos para componer un programa personalizado para cada paciente. Por ello, para cada tratamiento vamos a usar un vector de características, VT [17] de 3 elementos (uno por cada subgrupo), donde en cada posición se almacena una 2-tupla $\in S_1$ que indica el grado de importancia de cada uno de los subgrupos para la aplicación del tratamiento.

3.2 ESTABLECIMIENTO DE PERFILES DE USUARIO

En esta fase, el objetivo es desarrollar una base de datos que identifique el estado patológico de cada paciente en base al cual los tratamientos se irán personalizando. El paciente se someterá a una serie de tests [7], cuyos resultados serán analizados por los expertos para establecer la patología que presenta y así formar su perfil. Dicha patología vendrá dada por una serie de características, que son las mismas que hemos usado para la representación de los tratamientos, es decir, que a raíz de los resultados de los tests, los expertos

valoran la pertenencia de la patología del paciente a cada uno de los tres grupos de diagnóstico. Por tanto, también usamos un vector, VP [17], de 3 elementos donde en cada posición se almacena una 2-tupla $\in S_1$ que indica el grado de importancia del correspondiente grupo de diagnóstico para la patología presentada por el paciente.

Este perfil se complementará durante la fase de realimentación, en la que los propios pacientes irán suministrando evaluaciones sobre los tratamientos que el sistema les va proponiendo. Es decir, en función de la patología del paciente, el sistema propone un tratamiento que el paciente realiza y posteriormente evalúa. Pero también un experto analizará la evolución de los pacientes y evaluará el tratamiento aplicado. Ambas evaluaciones de paciente y experto, etiquetas de S_2, se van agregando mediante el operador $\overline{x}^e$ definido en [12].

3.3 MÉTODO DE GENERACIÓN DE RECOMENDACIONES

Como método de generación de recomendaciones vamos a usar un enfoque híbrido. Ante un tratamiento X, medimos su similitud con las patologías de cada uno de los pacientes y seleccionamos los que hayan tenido una similitud superior a un umbral lingüístico $\alpha \in S_1$. Al representar tanto los tratamientos como las patologías de los pacientes mediante el modelo vectorial, para el cálculo de la similitud usaremos la medida del coseno [17], pero modificada para que nos permita trabajar con 2-tuplas:

$$\sigma_l(V_1, V_2) =$$

$$\Delta\left(\frac{\sum_{k=1}^{n}(\Delta^{-1}(v_{1k}, \alpha_{v1k}) \times \Delta^{-1}(v_{2k}, \alpha_{v2k}))}{\sqrt{\sum_{k=1}^{n}(\Delta^{-1}(v_{1k}, \alpha_{v1k}))^2} \times \sqrt{\sum_{k=1}^{n}(\Delta^{-1}(v_{2k}, \alpha_{v2k}))^2}}\right) \quad (1)$$

A continuación, vamos a considerar las evaluaciones que sobre el tratamiento X hayan proporcionado pacientes con patologías similares a los seleccionados en la fase anterior, y las que hayan suministrado los expertos tras analizar la evolución de los pacientes que hayan trabajado X. Para medir la similitud entre perfiles de usuario, también usamos σ_l, pero considerando que el resultado obtenido en este caso es una etiqueta de S_3, puesto que estamos midiendo la similitud ente perfiles. Debemos también establecer un umbral lingüístico $\beta \in S_3$, que nos indique cuándo hemos encontrado una alta similitud entre perfiles de usuario.

Por último, para calcular la relevancia con la que recomendar X a los usuarios seleccionados (etiqueta de S_2), agregamos el valor obtenido en el primer paso (previamente transformado a etiqueta de S_2 mediante las funciones definidas en [13]), así como la información referente a evaluaciones tanto de pacientes como de expertos. Para ello, usamos el operador de agregación de 2-tuplas $\overline{x}^w$ definido en [12], dándole un mayor peso al valor calculado en el primer paso y a las evaluaciones proporcionadas por los expertos. Aquellos usuarios para los que se haya obtenido una relevancia estimada por encima de un determinado umbral $\gamma \in S_2$, se les recomienda el tratamiento X.

4 CONCLUSIONES

Hemos presentado el diseño de un sistema de recomendaciones lingüístico difuso de asistencia al tratamiento y prevención de las lumbalgias crónicas. El sistema analiza los perfiles de los pacientes, y en función de las patologías que presenten, va recomendando tratamientos y/o ejercicios para que los puedan realizar en sus hogares o puestos de trabajo, y así ayudarles a prevenir las lumbalgias crónicas. El sistema permitirá realizar programas de prevención específicos para la espalda y promover actitudes y hábitos de vida saludables entre la población.

Como principal trabajo futuro, nos proponemos la implantación del sistema y su posterior validación como instrumento para el seguimiento y control de los pacientes con lumbalgia y, otros posibles que pudieran beneficiarse de estas tecnologías. El estudio se realizará sobre una población entre 18 y 65 años, con lumbalgia crónica que vivan en Granada.

Agradecimientos

Este trabajo ha sido desarrollado con la financiación de los proyectos SAINFOWEB (00602) y FUZZYLING (TIN2007-61079).

Referencias

[1] G. Adomavicius, A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17, pág. 734-749, 2005.

[2] GBJ. Anderson. Epidemiological features of chronic low back pain. *Lancet*, 354, pág. 581-85, 1999.

[3] D. Ben-Arieh, C. Zhifeng. Linguistic labels aggregation and consensus measure for autocratic decision-making using group recommendations. *IEEE Transactions on Systems, Man, and Cybernetics Part A-Systems and Humans*, 36(3), pág. 558-568, 2006.

[4] K. Börner, C. Chen, K. Boyack. Visualizing Knowledge Domains. *Blaise Cronin (Ed.), Annual Review of Information Science & Technology*, 37, pág. 179-255, 2003.

[5] R. Burke. Hybrid Web Recommender Systems. P. Brusilovsky, A. Kobsa, and W. Nejdl: The Adaptive Web. *LNCS 4321*, pág. 377-408, 2007.

[6] Y. Cao, Y. Li. An intelligent fuzzy-based recommendation system for consumer electronic products. *Expert Systems with Applications*, 33, pág. 230-240, 2007.

[7] A. Cuesta Vargas. Efectividad de la fisioterapia basada en la evidencia con la carrera acuática sobre la lumbalgia crónica mecánica inespecífica. *Tesis doctoral, Universidad de Málaga*, 2007.

[8] F. García Pérez, S. Alcántara Bumbiedro. Importancia del ejercicio físico en el tratamiento del dolor lumbar inespecífico. *Rehabilitación*, 37, pág. 323-32, 2003.

[9] JH. Gracey, SM. McDonough, GD. Baxter. Physiotherapy management of low back pain. A survey of current practice in Northern Ireland, *Spine*, 27, pág. 406-11, 2002.

[10] U. Hanani, B. Shapira, P. Shoval. Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction*, 11, pág. 203-259, 2001.

[11] F. Herrera, E. Herrera-Viedma, J.L. Verdegay. Direct approach processes in group decision making using linguistic OWA operators. *Fuzzy Sets and Systems*, 79, pág. 175-190, 1996.

[12] F. Herrera, L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6), pág. 746-752, 2000.

[13] F. Herrera, L. Martínez. A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 31(2), pág. 227-234, 2001.

[14] E. Herrera-Viedma, A. G. López-Herrera, M. Luque, C. Porcel. A Fuzzy Linguistic IRS Model Based on a 2-Tuple Fuzzy Linguistic Approach. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems* , 15(2), pág. 225-250, 2007.

[15] E. Herrera-Viedma, G. Pasi, A. G. López-Herrera, C. Porcel. Evaluating the information quality of web sites: A methodology based on fuzzy computing with words. *J. of the American Society for Information Science and Technology*, 57(4), pág. 538-549, 2006.

[16] KJ. Kim, H. Ahn. A recommender system using GA K-means clustering in an online shopping market. *Expert systems with applications*, 34, pág. 1200-1209, 2008.

[17] R.R. Korfhage. *Information Storage and Retrieval*. New York: Wiley Computer Publishing, 1997.

[18] http://www.lastfm.es/

[19] G. Lekakos, P. Caravelas. A hybrid approach for movie recommendation. *Multimedia tools and applications*, 36, pág. 55-70, 2008.

[20] S. Loh, F. Lorenzi, R. Saldaña, D. Licthnow. A Tourism Recommender System based on Collaboration and Text Analysis. *Information Technology & Tourism*, 6, pág. 157165, 2004.

[21] I. Miralles. Prevención del dolor lumbar: Efectividad de la Escuela de Columna. *Rev Soc Esp Dolor*, 8, pág. 14-21, 2001 .

[22] http://www.movielens.org/login

[23] http://www.mystrands.com/

[24] AC. Papageorgiou, PR. Croft, S. Ferry, MI. Jayson, AJ. Silman. Estimating the prevalence of low back in the general population. Evidence from the South Manchester Back Pain Survey. *Spine*, 20(17), pág. 1889-94, 1995.

[25] P. Reisnick, H.R. Varian. Recommender Systems. *Special issue of Comm. of the ACM*, 40(3), pág. 56-59, 1997.

[26] J. Schiller. *Mobile Communications*. Addison Wesley, 2003 (2. edition).

[27] MG. Soukup, B. Glomsröd, JH. Lönn, K. Bö, S. Larsen. The Effect of a Mensendieck Exercise Program as Secondary Prophylaxis for Recurrent Low Back Pain: A Randomized, Controlled Trial With 12-Month Follow-up. *Spine*, 24(15), pág. 1585-1592, 1999.

[28] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences*, *part I* 8, pág. 199-249, 1975. *Part II* 8, pág. 301-357, 1975. *Part III* 9, pág. 43-80, 1975.

ANÁLISIS, FILTRADO Y CLASIFICACIÓN DE SENTENCIAS CONDICIONALES EN DOCUMENTOS CIENTÍFICOS

Cristina Puente [1] **José A. Olivas** [2]

[1] Dept. de Sistemas Informáticos, ETS de Ingeniería, Universidad Pontificia Comillas, 28015-Madrid.
cpuente2@upcomillas.es.com
[2] Dept. de Tecnologías y Sistemas de la Información, Universidad de Castilla La-Mancha, 13002-Ciudad Real.
joseangel.olivas@uclm.es

Resumen

En este trabajo se presenta un análisis comparativo de estructuras condicionales entre textos convencionales y textos científicos. Para su elaboración se ha utilizado un prototipo capaz de extraer y clasificar determinadas sentencias condicionales en documentos de texto. Además, se proponen unas guías para la aplicación de técnicas de Soft-Computing para aprovechar lo mejor posible las características semánticas de las sentencias seleccionadas.

Palabras Clave: Búsqueda Web, relaciones causales, deducción, Soft-computing.

1 INTRODUCCIÓN

Como es sabido, la mayoría de los motores de búsqueda en Internet se basan en indexaciones lexicográficas, es decir, buscan las palabras que el usuario introduce en una consulta, dentro de los documentos almacenados en la Web que están a su alcance. El principal problema de este enfoque es que se pierde la capacidad de recuperar documentos conceptualmente relacionados con la petición original, por ejemplo a través de sinónimos o antónimos, o de otros elementos conceptualmente relacionados.
Como el propio profesor Zadeh propone [7], se puede ir un paso más adelante y tratar de dotar a los motores de búsqueda de capacidades deductivas, es decir, que sean capaces de responder a preguntas como ¿Cuál es la tercera provincia más extensa de la península ibérica?.
Un intento de abordar este reto puede pasar por el análisis de la causalidad. Este problema, tema central de este estudio, ha sido una cuestión bastante discutida en los últimos años. El profesor Zadeh defiende [8] que no se puede establecer una definición genérica de la causalidad, puesto que depende de múltiples factores, y no se puede englobar dentro de la lógica clásica ni dentro de las teorías probabilísticas. Como ejemplo supóngase que se tiene una tienda de abrigos y que se pretende aumentar las ventas; con este fin se aumenta el presupuesto de marketing en un 20%. Seis meses después las ventas aumentan un 10%, ¿se debe este aumento de ventas al aumento de presupuesto de marketing?¿en qué medida?.
Como en casos similares al del ejemplo, no se puede establecer una definición genérica de causalidad, lo que no quiere decir que no se pueda tratar de forma efectiva, sin necesitar una formulación general, rigurosa y operacional, ya que los factores a tratar dentro de la causalidad son más complejos que las situaciones en las que se aplica la teoría de probabilidades.
Existen dos tipos básicos de causalidad: por un lado la denominada causalidad hacia delante (forward causality) expresada de la forma "¿Cuáles son los efectos provocados por una causa concreta?", y la causalidad inversa, expresada de la forma "¿Cuáles han sido las causas que han generado como consecuencia un efecto determinado?". En un contexto concreto, la causalidad hacia delante es más fácil de tratar que la causalidad inversa, puesto que lo que se suele conocer es el desencadenante de la acción. En la causalidad inversa puede haber múltiples factores que hayan provocado un hecho, por lo que es bastante más complejo de analizar y de tratar.
En el ejemplo de ventas anteriormente expuesto, se sabe que las ventas han subido, y uno de los posibles factores, pero ¿existen otros factores que hayan influido en este hecho?. Y si los hay ¿qué peso tiene cada uno de los factores en la subida de las ventas?. En los ejemplos como el caso anterior cuando la causa es provocada por una serie de factores, alguno de ellos desconocido, el problema está en la ponderación de los mismos. En general, este problema no tiene una solución rigurosa y precisa, sino que más bien se basa en la percepción, juicio y criterio de establecimiento.
Ésta es la principal causa por la cual no se puede precisar una definición global, exacta y rigurosa de la causalidad, puesto que depende en gran medida del criterio y la

ponderación de determinados factores, y dependiendo el caso, unos serán conocidos y otros no tanto.
Puesto que el objetivo del presente trabajo es realizar una aportación que mejore la eficiencia en las búsquedas a través de Internet, el análisis de la causalidad está centrado en la extracción de este tipo de sentencias dentro de un determinado documento para su posterior tratamiento.
Si el análisis se centra en los aspectos semánticos, se obtendrán diversas oraciones condicionales que por su contenido no pueden ser tratadas como causales, como se puede ver en el ejemplo: "Si apruebo, me dejo de llamar Francisco", donde el hecho de que la persona se llame Francisco no lo causa el aprobado.
Para que una sentencia pueda ser catalogada como causal tiene que cumplir al menos tres condiciones:

- La causa debe ser anterior al efecto.
- Siempre que suceda la causa debe producirse el efecto.
- Causa y efecto deben estar próximos en el espacio.

No todas las sentencias condicionales cumplen estas tres premisas, por lo que teniendo esto en cuenta, y puesto que el análisis sintáctico es bastante menos complejo de tratar que el semántico, en este trabajo se presenta un análisis comparativo en cuanto a la extracción y clasificación de oraciones condicionales entre textos de ámbito general, con multitud de metáforas y giros lingüísticos, y artículos de tipo científico, mucho más técnicos y específicos. Para llevar a cabo esta operación, se ha utilizado un prototipo de clasificación y extracción de sentencias condicionales, cuyo funcionamiento se describe brevemente a continuación.

2 PROTOTIPO PARA LA DETECCIÓN Y CLASIFICACIÓN DE ESTRUCTURAS CONDICIONALES

En [1], se presentó un prototipo programado mediante el analizador morfológico *Flex*[1] y el lenguaje de programación *C*, que permite la extracción de una serie de estructuras condicionales previamente definidas. Para su implementación, fue necesario efectuar un estudio sintáctico acerca de la forma en la que se podía presentar una sentencia condicional dentro de un texto [3]. Según se observó había dos tipos de estructuras dentro de las sentencias condicionales. Por un lado estaban aquellas sentencias en las que la aparición de determinados verbos marcaban la forma condicional de la oración, y por otro nos encontramos con otro tipo de estructuras en las que la función condicional estaba denotada por conjunciones en algunos casos y por adverbios en otros.

[1] http://www.gnu.org.

Este estudio dio lugar a una serie de estructuras etiquetadas en cuanto a la forma en la que se presentan las sentencias condicionales, denotando en cada caso los tiempos verbales que deben aparecer, tanto en el antecedente como en el consecuente.
Puesto que el número de estructuras analizadas era bastante elevado, se decidió implementar un subconjunto como primer paso para la creación de un prototipo capaz de detectarlas y clasificarlas. En concreto se decidió analizar las estructuras del idioma inglés pertenecientes al condicional en su forma clásica *if x then y*, correspondientes al *first*, *second* y *third conditional*, lo que dio lugar a las 20 estructuras definidas en la figura 1, y que sirvieron de entrada y base a los procesos de detección y clasificación.

e1: if + present simple + present simple.
e2: if + present simple + future simple.
e3: if + present simple + may/might (para expresar posibilidad).
e4: if + present simple + must/should.
e5: if + present continuous + future simple.
e6: if + past simple + would + infinitivo.
e7: if + past simple + might/could.
e8: if + past continuous +would + infinitive (o would be + gerund).
e9: if + past perfect +would + infinitivo.
e10: if + past perfect + would have(perfect conditional) + past participle.
e11: if + past perfect + might/could have + past participle.
e12: if + past perfect + perfect conditional continuous.
e13: if + past perfect continuous + perfect condicional.
e14: if + past perfect + conditional simple.
e15: if + present simple + imperative.
e16: if + past perfect + would + be + gerund.
e23: for this reason, as a result.
e24: due to, owing to.
e26: provided that.
e31: have something to do, a lot to do.
e34: so that, in order that.
e38: although, even though.
e43: in case that, in order that.
e44: on condition that, supossing that.

Figura 1: Estructuras condicionales implementadas.

2.1. SUBSISTEMA DE DETECCIÓN Y SUBSISTEMA DE CLASIFICACIÓN

El algoritmo de detección es el que se encarga del filtrado de las sentencias condicionales cuyo planteamiento está dentro de las 20 estructuras mencionadas que se han seleccionado para la creación de este prototipo (figura 1).

Para llevar a cabo el análisis de las distintas unidades sintácticas que componen una sentencia, se ha construido un autómata de estados finitos (figura 2), que permita desarrollar esta tarea. *Flex* establece de una manera sencilla la definición de cada uno de los estados por los que debe pasar el autómata junto con las acciones a efectuar en cada uno.

Para implementar el algoritmo de detección sólo han sido necesarios dos estados, el inicial y el condicional. El proceso comienza en el estado inicial, donde va procesando las unidades sintácticas que le van llegando. Si el *token* a procesar coincide con la conjunción condicional, el autómata pasará al estado condicional (*detect*) etiquetando la sentencia como tal. En caso contrario el proceso seguirá analizando unidades sintácticas desde el estado inicial, almacenándolas en un buffer, a esperas de que le llegue una conjunción condicional que le señale la naturaleza de la oración. Si el proceso recibe como unidad sintáctica un punto y final desechará la oración en la que se encuentra, vaciando el buffer donde ha sido almacenada.

Si la unidad sintáctica analizada coincide con una conjunción condicional, el autómata deberá determinar la posición de este *token* dentro de la sentencia para establecer si el consecuente de la oración ha sido procesado ya y se encuentra almacenado, o si está aún por analizar, tanto antecedente como consecuente. Para completar éste proceso, se ha creado un vector con las formas verbales más comunes en el idioma inglés, de tal manera que el algoritmo va activando determinadas posiciones de este vector, en función de los verbos y los tiempos verbales que va detectando.

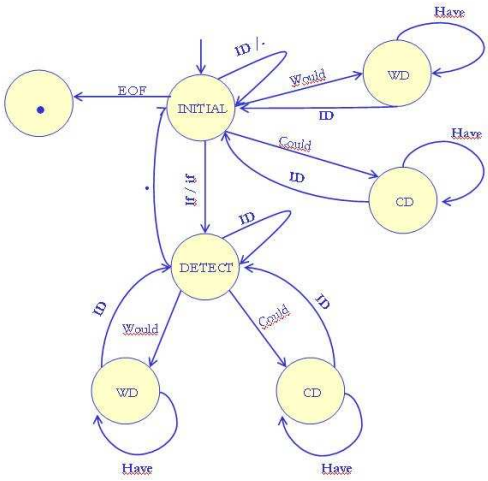


Figura 2: Autómata de detección y clasificación..

Una vez que el algoritmo de detección ha concluido el análisis de todas las unidades sintácticas que componen una sentencia condicional, procederá a efectuar la clasificación de la misma. El algoritmo generado para este prototipo permite la distinción del tipo de condicional que se ha detectado en función de la forma verbal y la formación de la estructura sintáctica, tomando como referencia las estructuras condicionales mencionadas.

Puesto que el proceso de clasificación está basado en el proceso de detección, se ha aprovechado el autómata creado, efectuando los ajustes necesarios para realizar ambas tareas.

Se han añadido dos estados nuevos, *cw* y *wd*, cuya misión es comprobar las formas verbales compuestas, ya que *Flex* delimita los tokens por espacios. De esta manera, si la forma verbal a procesar se tratara de *would have*, el autómata se desplazaría al estado *wd* al detectar *would* y quedaría a la espera de un verbo que pudiera formar un tiempo compuesto como *be* o *have*. Si el siguiente token analizado se correspondiera con *have* el proceso activaría la posición asociada a *would have* en el vector de verbos en lugar de *would*, y retornaría al estado anterior.

Una vez ampliado el autómata, se procedió a revisar las estructuras sintácticas que fundamentan la aplicación. Estudiando la composición de cada una, se observó que casi siempre aparecía o bien en el antecedente de la sentencia o bien en el consecuente, una forma verbal que marcaba la clasificación dentro de una estructura u otra.

También se percibió que dicha forma verbal generalmente estaba compuesta por alguno de los verbos utilizados en el vector de formas verbales implementado para el algoritmo de detección, por lo que para clasificar una sentencia en un tipo de estructura determinado bastará con recorrer dicho vector analizando las formas verbales encontradas a lo largo de la oración. Por ejemplo si tuviéramos la oración: *If the weather is fine, we'll go to the beach*, el vector de formas verbales registraría *is* en la posición 2 y *'ll* en la posición 21. Al recorrer el vector y localizar la posición 2 y 21 activada el proceso determinaría que se trata de una estructura de tipo 2 con la forma if + present simple + future simple como se puede ver a continuación:

>Número de linea:1 If the weather is fine we 'll go to the beach.
***CLASIFICACIÓN: FIRST CONDITIONAL ESTRUCTURA 2*

3 ANÁLISIS DE LOS RESULTADOS

En [1], se planteó un modelo de pruebas para evaluar la fiabilidad en la extracción y clasificación del prototipo creado, así como para contrastar los resultados en función del tipo de lenguaje empleado. De ésta manera, utilizando como entrada los textos que aparecen en la tabla 1, se puede observar que el porcentaje de sentencias detectadas es bastante elevado, independientemente del número de páginas que compongan el documento. De igual forma el porcentaje de estructuras que el programa ha sido capaz de clasificar (ver figura 1 para su sintaxis), supera ampliamente el cincuenta por ciento en todos los casos salvo en caso de las noticias de la sección del *Herald Tribune*, que exhibe un lenguaje más complejo, cargado

de metáforas y giros lingüísticos que el resto de los documentos analizados.

Tabla 1: Estudio comparativo con textos de distintos ámbitos.

Titulo del documento	Número de Páginas	Nº de Palabras	Nº E. Detectadas	Nº E. Desechadas	Nª E. Clasificadas	Nª E. No Clasificadas	% Clasificadas
Devil wears	198	139.883	358	41	218	140	60,9
The Historian	358	245.333	602	253	433	169	71,9
Harry Potter	288	150.666	345	136	251	94	72,7
Lord of the	258	188.262	608	152	489	119	79,9
Evangelios	145	88.232	216	1	161	55	74,5
Noticias Ciencia	7	3.069	5	0	5	0	100
Noticias Internacional	11	4.940	8	0	4	4	50
Guión Kill Bill	249	42.618	78	6	65	13	83,4
Discurso Madre Teresa	6	4.063	15	4	12	3	80
Discurso Luther King	15	8.628	22	5	21	1	95,4

Tabla 2: Distribución de estructuras condicionales en textos de distintos ámbitos.

Titulo del documento	e2	e3	e4	e6	e7	e8	e9	e10	e11	e12	e13	e16	e23	e24	e26	e31	e34	e38	e43	e44	tot
Devil wears Prada	37	8	7	36	35	21	0	2	0	2	0	2	4	10	0	2	8	122	19	0	315
The Historian	106	7	24	69	81	47	2	9	16	6	1	3	3	2	0	7	108	220	7	0	718
Harry Potter	49	3	6	37	46	42	0	6	3	5	0	4	0	12	0	1	35	61	8	1	319
Lord of the Rings	201	31	55	51	53	34	0	10	0	6	0	4	0	0	0	2	25	11	2	0	485
Evangelios	110	3	13	6	6	6	0	5	0	0	0	0	0	0	0	0	11	1	0	0	161
Noticias Ciencia	2	0	0	0	2	1	0	0	0	0	0	0	1	2	0	0	0	0	0	0	8
Noticias Internacional	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	5
Guión Kill Bill	34	0	1	5	4	3	0	1	0	0	0	1	0	0	0	0	2	3	1	0	55
Discurso Madre Teresa	7	0	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
Discurso Luther King	10	3	6	1	0	0	0	0	0	0	0	0	0	2	0	0	2	2	0	0	26

Como se puede apreciar en la tabla 2, la distribución de estructuras condicionales clasificadas es bastante uniforme de forma global, apareciendo sentencias pertenecientes tanto al *first*, *second* o *third* condicional junto con sus variantes.

Para centrar más este estudio en un ámbito menos generalista, se decidió seleccionar una serie de documentos de tipo científico pertenecientes a diferentes disciplinas como física, química, matemáticas, etc., con un modelo de redacción mucho más riguroso y estricto que en las pruebas anteriores. Los resultados, como se puede ver en la tabla 3, presentan un mayor volumen de estructuras detectadas y una recesión importante en cuanto a las desechadas, lo que indica a priori que el tratamiento lingüístico se ajusta más a la norma que en los casos anteriores. En cuanto al porcentaje de clasificación, ha sido algo menor que en los textos de la tabla 1, sobre todo en el caso del texto de Pierre y Marie Curie, lo que se puede deber al tipo de lenguaje un tanto desfasado que utiliza.

La diferencia fundamental a la hora de analizar textos científicos, aparece en la distribución de las sentencias clasificadas (tabla 4), existiendo una gran discrepancia entre las estructuras *e4* y *e8* particularmente, ya que en los textos de ámbito general aparecen un número significativo de sentencias de este tipo, y en el caso de textos de tipo científico el número se ve reducido considerablemente, probablemente porque estas sentencias se utilizan en la formulación de conjeturas. Por otro lado, las proposiciones con una formulación más categórica utilizando el presente simple, como la del ejemplo planteado, son las que tienen mayor presencia dentro de la tabla 4.

***CONSECUENTE: A fuzzy set is empty if **ANTECEDENTE: and only if its membership function is identically zero on X.*
>Número de linea:29 -- A fuzzy set is empty if and only if its membership function is identically zero on X--
CLASIFICACIÓN: FIRST CONDITIONAL ESTRUCTURA 2.

Las estructuras del *second* y *third conditional*, tienen una presencia bastante escasa en los textos analizados, lo que señala una diferencia importante con respecto a los textos anteriores, muy a tener en cuenta para futuras investigaciones.

Tabla 3: Estudio comparativo con textos científicos.

Titulo del documento	Número de Páginas	Nº de Palabras	Nº E. Detectadas	Nº E. Desechadas	Nª E. Clasificadas	Nª E. No Clasificadas	% Clasificadas
Artículo A. Einstein	9	3.746	14	1	11	3	78,6
Artículo L. A. Zadeh 1965	16	5.855	33	0	29	4	87,9
Artículo R. B. Myerson	17	7.144	66	1	43	23	65,2
Artículo A. Flemming	10	4.341	20	0	14	6	70,0
Artículo P. y M. Curie	2	913	5	0	2	3	40,0
Artículo J.A. Olivas, A. Sobrino	7	3.302	14	0	10	4	71,4
Artículo J. Stasheff	14	6.261	18	0	11	7	61,1
Quantum Cosmology	8	4.509	24	0	18	6	75,0

Tabla 4: Distribución de estructuras condicionales en textos científicos.

Titulo del documento	e2	e3	e4	e6	e7	e8	e9	e10	e11	e12	e13	e16	e23	e24	e26	e31	e34	e38	e43	e44	tot
Artículo A. Einstein	1	1	0	0	0	0	0	0	0	0	0	0	2	1	0	0	2	4	0	0	11
Artículo L. A. Zadeh 1965	24	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	29
Artículo R. B. Myerson	10	7	8	2	3	1	0	1	0	0	0	0	0	0	0	0	8	3	0	0	43
Artículo A. Flemming	4	1	0	1	0	0	0	0	0	0	0	0	0	3	0	0	4	1	0	0	14
Artículo P. y M. Curie	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2
Artículo J.A. Olivas, A. Sobrino	2	0	0	0	0	1	0	0	0	0	0	0	1	3	0	0	1	2	0	0	10
Artículo J. Stasheff	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	4	2	0	11
Quantum Cosmology	6	1	1	4	1	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18

4 APLICACIÓN DE TÉCNICAS BORROSAS

En el trabajo del profesor Trillas [2], se sientan las bases de lo que puede ser un análisis de una sentencia condicional desde la lógica borrosa. En este trabajo se plantea el problema de la validez de un razonamiento en base al análisis del condicional, para lo cual a través de un ejemplo de una frase muy compleja de Vila-Matas analiza el efecto que tendría sobre la sentencia el análisis sobre la base de distintas implicaciones borrosas.

Tomando como base este trabajo se está trabajando principalmente en dos líneas. La primera trata la asignación de valores de verdad borrosos a los antecedentes y consecuentes (en la línea de Trillas) para poder desencadenar mecanismos de inferencia, representando las sentencias detectadas mediante *Protoforms* [6], [7].

La segunda línea, pretende establecer algún tipo de índice de relación borroso entre los conceptos que subyacen al antecedente y el consecuente de las oraciones clasificadas, en base al tipo de condicional gramatical y la implicación borrosa que pudiera asociársele, para aportar un parámetro útil más (junto a otros como sinonimia, hiperonimia, etc.) en las relaciones entre conceptos, con el fin de mejorar conceptualmente las búsqueda en la Web.

5 CONCLUSIONES

Las pruebas ejecutadas y los resultados obtenidos muestran tres conclusiones fundamentalmente. La primera es que la aparición de un mayor número de estructuras condicionales no está supeditado al volumen de texto analizado, sino al estilo literario y a la forma de expresión en la que está trasmitido como se ha mostrado en la tabla 1 y en la tabla3. La segunda refleja que los resultados obtenidos son coherentes con los tipos de estructuras estudiadas y se ajustan a los patrones definidos en las mismas, lo que explica a su vez el alto porcentaje de estructuras clasificadas frente a las detectadas.
Por ultimo, queda demostrado mediante la tabla 4 que el lenguaje científico utiliza una expresión basada en formulaciones en tiempo presente mayoritariamente, desechando casi por completo las estructuras del *second* y del *third* condicional.

Agradecimientos

Proyecto PAC06-0059 SCAIWEB, JCCM, España, y proyecto TIN2007-67494 F-META, MEC-FEDER, España.

Referencias

[1] C. Puente, J. A. Olivas, *Análisis y detección de determinadas estructuras condicionales en documentos de texto.* II Simposio sobre lógica Fuzzy y Soft Computing, LFSC'07,Zaragoza, Pág. 89-96, 2007

[2] E. Trillas. *Lógica borrosa y narrativa: un párrafo de Vila–Matas*, XII Congreso español sobre tecnologías y lógica fuzzy, Jaén, 23-34, 2004.

[3] J. A. Olivas; C. Puente, A. Tejado. *Searching for causal relations in text documents for ontological application*. Proc. of the Int. Conf. on Artificial Intelligence IC-AI'2005, vol. II, CSREA Press, 463 – 468, 2005.

[4] J. A. Olivas; P. Garcés; , F. P. Romero,. *An application of the FIS-CRM model to the FISS metasearcher: Using fuzzy synonymy and fuzzy generality for representing concepts in documents*. International Journal of Approximate Reasoning (Soft Computing in Recognition and Search) 34, Pág. 201 – 219, 2003.

[5] L. A. Zadeh, *From Search Engines to Question Answering Systems: The Problems of World Knowledge, Relevance, Deduction and Precisiation. In Fuzzy Logic and the Semantic Web*, edited by Elie Sanchez, Elsevier 2006.

[6] L. A. Zadeh. *A Prototype-Centered Approach to Adding Deduction Capability to Search Engines -- The Concept of Protoform*, Berkeley Initiative in Soft Computing, Submitted to the BISC mailing list 19th Dec 2001.

[7] L. A. Zadeh. *From search engines to Question-Answering System: The need for new tools*, E. Menasalvas, J. Segovia, P.S. Szczepaniak (Eds.): Proceedings of the Atlantic Web Intelligence Conference - AWIC'2003. LNCS, Springer, 15 – 17, 2003.

[8] L. A. Zadeh. *Precisiated Natural Language (PNL)*. AI Magazine, Vol. 25, No. 3, 74-91, 2004.

Filtrado de información mediante prototipos borrosos y perfiles basados en criterios de calidad de datos

Francisco P. Romero [1] **, Ismael Caballero** [2]**, Jose A. Olivas** [1]**, Eugenio Verbo** [2]

[1] Departamento de Tecnologías y Sistemas de la Información. Grupo SMILE.
Universidad de Castilla – La Mancha
Paseo de la Universidad, 4, 13071, Ciudad Real
{FranciscoP.Romero, JoseAngel.Olivas}@uclm.es
[2] Centro Mixto de I+D Indra UCLM
Indra Software Labs
Ronda de Toledo,s/n, 13004, Ciudad Real
{icaballerom, emverbo}@indra.es

Resumen

Las nuevas formas de manipulación y acceso de la información que han aparecido por el uso de Internet, están generando grandes volúmenes de datos en constante crecimiento. Las herramientas clásicas de gestión utilizadas por los usuarios fracasan en el objetivo de satisfacer sus necesidades de información. La utilización de técnicas de filtrado basadas en categorías conceptuales se ha mostrado como una estrategia útil en estos entornos, aunque fracasa relativamente en las situaciones en que los datos que se recogen son incompletos o incorrectos como sucede en entornos de uso tan general como el correo electrónico. En este trabajo es presentada una aproximación a la resolución de estos problemas mediante la utilización sinérgica del filtrado de información mediante categorías definidas mediante prototipos conceptuales junto con la aplicación de perfiles basados en criterios de calidad de datos.

Palabras Clave: Filtrado de Información, Prototipos Borrosos, Calidad de Datos.

1 INTRODUCCIÓN

Los recientes avances en las tecnologías software y de redes han llevado a las organizaciones a desarrollar nuevos sistemas de información basados en Internet, como una forma de acercarse lo máximo posible a sus clientes, a sus trabajadores, y en general, a cualquier persona implicada en la actividad organizacional. Para que esta actividad sea lo más fructífera posible, las compañías consideran necesario recopilar, guardar y mantener grandes volúmenes de documentos electrónicos. De hecho, son cada vez más las que se conciencian de que los datos son, junto con sus empleados, uno de sus activos más importantes [9]. Es tal el interés por recoger datos, que muchas organizaciones están digitalizando su información histórica para que pueda ser procesada junto con los datos nuevos que se van generando en el día a día. Además, dada la tendencia a la globalización es frecuente que estos datos se encuentren distribuidos entre distintos puntos [2], lo que incrementa la dificultad de la gestión de estos datos. Pero a la hora de desarrollar un trabajo, por ejemplo, recuperación de información, búsqueda de contenidos o categorización [4], manejar tal cantidad de datos supone sacrificar en rendimiento de los Sistemas de Información. Como este sacrificio puede disminuir la satisfacción de los usuarios con los sistemas, se hace necesario filtrar convenientemente las colecciones de datos que se van a procesar.

Actualmente las técnicas de soft-computing se presentan como una herramienta eficiente para resolver estos problemas [14] [7] en sus diferentes tareas [4][6]. En concreto técnicas como el filtrado basado en categorías [19], suponen un incremento de rendimiento considerable en los procesos de tratamiento de documentos repositorios dinámicos. La utilización de la lógica borrosa, y en concreto de los prototipos borrosos [16], con el objetivo de representar las categorías del modelo de filtrado permite crear un modelo más entendible y actualizable para realizar las tareas de filtrado de información. Para poder paliar carencias de los modelos semánticos y borrosos de filtrado en repositorios heterogéneos es necesario introducir en el proceso políticas de calidad de datos definidas por los propios usuarios para los correspondientes contextos [3], ya que la calidad de los datos es el principal criterio para la discriminación de los datos que no son útiles a los usuarios en un contexto. De esta manera, se logrará una mayor eficacia en la gestión del repositorio de documentos.

El presente trabajo se organiza de la siguiente manera. En el apartado 2 se explicarán los criterios de calidad de datos que van a ser utilizados para el filtrado de los documentos. En el apartado 3 se explicará el modelo de filtrado que aúna la utilización de estos criterios con la utilización de prototipos deformables borrosos. En el apartado 4 se presentará el proceso de tratamiento de nuevos documentos. Para concluir, en los últimos apartados, se describirán los experimentos realizados y las conclusiones y el trabajo futuro a realizar.

2 CRITERIOS DE CALIDAD DE DATOS PARA EL FILTRADO DE OCUMENTOS.

Los marcos de trabajo para la gestión de la calidad de los datos la calidad de datos (en adelante DQ), proporcionan a los usuarios un conjunto de técnicas y herramientas fundamentales para evaluar la calidad de los contenidos de un documento. Conociendo el nivel global de calidad de datos de un documento, un usuario o un sistema pueden decidir si utilizarlo o descartarlo.

Para determinar si el documento es útil se analizarán si los datos contenidos para comprobar presentan un nivel adecuado de calidad estudiado con respecto a ciertas dimensiones de calidad de datos. Teniendo en cuenta que la "adecuación para el uso" depende de la tarea y del rol del agente o stakeholders que maneje las fuentes y los documentos, es necesario identificar a partir de los requisitos de usuario, aquellas dimensiones que mejor representan las necesidades de calidad de datos para cada uno de los tipos de usuarios. La noción de gestión lleva aparejada intrínsecamente la idea de medición [18], por lo que es preciso establecer los mecanismos adecuados de medición. Dada la naturaleza de los datos, la realización de esta medición implica tener en cuenta una serie de aspectos tal y como se describe en [5] a la hora de definir las correspondientes medidas, como puedan ser la subjetividad o el método de medición. Los aspectos propuestos están alineados con el estándar de medición de software ISO/IEC 15939 (ISO/IEC 2000), lo que hace que los términos específicos de calidad de datos puedan ser fácilmente utilizados por desarrolladores que ya utilicen este estándar.

En este trabajo el modelo de calidad de datos utilizado para evaluar los documentos entrantes al sistema de filtrado consta de las tres siguientes dimensiones:

1. **Compleción:** El documento posee todos sus ítems completos y con información.
2. **Fiabilidad:** La fuente de la información, bien sea el autor u otro tipo de fuente identificada es fiable en cuanto a sus contenidos.
3. **Oportunidad:** La información recibida es adecuada al contexto temporal en el que es ubicado el tema sobre el que trata.

Con el fin de medir la completitud se calcula la proporción, en primer lugar entre los ítems que contiene el documento y los que debería tener (en documentos XML se valida con respecto a su esquema XSD), y posteriormente se valida cuales de los que posee poseen valor. La fórmula sería la siguiente:

$$completitud(d) = \frac{items_{completos}}{items_{totales}} \quad (1)$$

De esta manera se medirá la proporción entre los ítems completos y los ítems que teóricamente tendría que tener el documento según su esquema.

La fiabilidad del documento debe evaluarse de forma cualitativa con respecto a los datos que se dispongan de él, por ejemplo, a partir del autor. Se necesita disponer por lo tanto de una calificación mediante etiquetas lingüísticas de la fiabilidad de estos autores (Tabla 1). Estos son metadatos que complementan el significado del dato en la dimensión de calidad elegida.Si no se dispone de autor esta dimensión no puede ser medida.

Tabla 1: Fiabilidad por autores

Autor	**Fiabilidad**
Tim Harford	Alta
Collen McCullough	Media
Dan Brown	Baja

La oportunidad del documento se medirá a partir de la diferencia existente entre una fecha establecida (dependiente del contexto de la tarea) y la fecha en que se generó el documento. Si no se cuenta con esta última fecha, la oportunidad no podrá ser medida. Los grados de oportunidad posibles se identificarán mediante etiquetas lingüísticas (muy oportuno, oportuno, poco oportuno). Diferentes categorías pueden tener diferentes criterios de oportunidad en sus temas.

A partir de estas tres dimensiones es necesario establecer el nivel de calidad de datos del documento en cuestión, para ello se realizará una agregación de los tres valores penalizando aquellos documentos en los que no se pueda medir alguna de estas dimensiones.

Dado que la percepción de la calidad es algo subjetivo e impreciso se hace necesario el uso de reglas borrosas para medir/determinar/evaluar adecuadamente la calidad de los datos. Las variables lingüísticas de entrada serían las dimensiones de calidad de datos y la variable de salida sería la calidad de datos del documento con un conjunto de etiquetas como el siguiente: muy alta, alta, normal, baja, muy baja. La inferencia sobre estas reglas será de tipo Mamdani [12].

3. MODELO DE FILTRADO DE INFORMACIÓN

El modelo de filtrado que se presenta en este trabajo está basado en una estructura basada en categorías conceptuales. Para construir esta estructura partiendo de un conjunto de documentos no estructurados, es necesario seguir los siguientes pasos:

1. Calcular la calidad de datos de todos los documentos y descartar aquellos cuyas mediciones no estén dentro de los rangos de aceptación establecidos en los requisitos de calidad de datos de los usuarios.que no superen un umbral mínimo de calidad.
2. Preproceso lingüístico: Selección de las palabras que van a representar conceptualmente a las categorías de la estructura.
3. Representación de los documentos en base a los conceptos tratados en sus contenidos. Para ello se utiliza el modelo FIS-CRM [13].
4. Agrupación de los documentos conceptualmente similares mediante clustering [15].
5. Extracción de conceptos claves.
6. Creación de una base de conocimiento utilizando Categorías Prototípicas Borrosas.

3.1. ESTRUCTURA BASADA EN CATEGORÍAS

La estructura básica de filtrado se basa en las diferentes categorías extraídas mediante procesos de clustering. Estas categorías comprenden una serie de documentos conceptualmente similares. Las categorías estarán organizadas en una jerarquía que reflejará los diferentes niveles de especificidad tratados en los conceptos. Cada uno de esos documentos puede estar clasificado en diferentes categorías sobre las cuales poseerán un grado de pertenencia.

Cada categoría de la estructura vendrá representada por una Categoría Prototípica Borrosa, que se podrá definir mediante una representación formal a partir de números borrosos y una definición parámetrica realizada mediante marcos que en este caso obtendrán pares de acción-probabilidad, los cuales ayudarán en la toma de decisiones de gestión sobre los documentos entrantes.

3.2. CONCEPTOS CLAVE

Cada agrupación posee un vocabulario propio obtenido a partir de los conceptos reflejados en los documentos que contiene. Debido a que un porcentaje reducido de los conceptos definen a la categoría es necesario establecer mecanismos para identificar los grupos de conceptos claves para posteriormente utilizarlos en el filtrado. Para la extracción de los conceptos clave que definen el contenido de la agrupación se define una medida de la importancia de éste dentro de la agrupación teniendo en cuenta el número de documentos de la categoría en los que aparece, su peso dentro estos documentos y también su aparición dentro del resto de agrupaciones de la estructura. El peso del concepto i en la categoría C (w_i^c) vendrá dado en función de la de sus peso en los documentos de la categoría (w_{ij}), la cantidad normalizada de documentos en los que aparece (*docs(i,C)*) y un factor logarítmico que determine su aparición en otras categorías:

$$w_i^C = \sum_{j \in C} w_{ij} \times \left(1 + \frac{docs(i,C)}{|C|}\right) \times Ln\left(\frac{|cats|}{cats(C)} + 1\right) \quad (2)$$

Para su posterior utilización este valor se normaliza dividiéndolo por la media de los valores obtenidos en la categoría. A partir del estudio de la distribución de estos valores entre los diferentes conceptos se pueden determinar los tres niveles de importancia que se van a utilizar en este modelo. (Tabla 2).

Tabla 2: Niveles de relevancia de los conceptos.

NOMBRE	DESCRIPCIÓN
Nivel 1	Términos con mayor importancia (cuarto superior)
Nivel 2	Términos con importancia suficiente para ser utilizables (2º cuarto superior)
Nivel 3	Término de nula importancia (debajo de la media)

Los conceptos de nivel 1 serán los que se utilicen para identificar la categoría dentro de la estructura, es decir, conformarán la etiqueta de la agrupación. Los conceptos de nivel 1 y de nivel 2 serán los utilizados en el proceso de filtrado.

3.3 PROTOTIPICALIDAD DE LOS DOCUMENTOS

El grado de prototipicalidad de un objeto dentro de un conjunto depende positivamente de la semejanza hacia otros objetos del su categoría, y su diferencia con los objetos del resto de categorías. En este caso, también se ve influido por su nivel de calidad de datos ya que si el documento no tiene un nivel adecuado de calidad entonces los resultados podrían verse falseados al utilizar documentos que no deben ser tenidos en cuenta. Estos factores se calcularán de la siguiente forma:

- **Grado de compatibilidad:** El grado de compatibilidad de un documento con el cluster al que pertenece será la media de similitud entre el documento y el resto de los documentos presentes en la agrupación. (Formula 3)

$$comp(d_i, C) = \frac{\sum_{j \in C - d_i} sim(d_i, d_j)}{|C| - 1} \quad (3)$$

- **Grado de incompatibilidad:** El grado de incompatibilidad de un documento con el resto de clusters vendrá dado por la media de la diferencia entre la similitud media (cohesión) de los documentos de su propio cluster y su similitud con los documentos del resto de los cluster.

$$\overline{comp}(d_i,\overline{C}) = \frac{\sum_{c\notin C}\left(\frac{\sum_{i\in C} sim(d_i,d_j)}{|C|} - \frac{\sum_{h\in c} sim(d_i,d_h)}{|c|}\right)}{|\overline{C}|} \quad (4)$$

- **Coeficiente agregado de calidad de datos** calculado a partir de la agregación de los valores de las tres dimensiones de calidad de datos descritas en la sección segunda.

Agregando estos tres valores se obtendrá el grado final de prototipicalidad del documento dentro de la agrupación.

3.4. REPRESENTACIÓN FORMAL DE LOS PROTOTIPOS.

Con el fin de representar los números borrosos correspondientes a las Categorías Prototípicas Borrosas es necesario determinar o bien el número de "prototipos" que se quieren tener en el modelo de predicción o bien el porcentaje de similitud mínimo para que los elementos que lo posean puedan formar parte del mismo grupo. En el caso de estar obteniendo prototipos borrosos según [21] sería suficiente con dividir el grupo en tres grupos, pero al estar trabajando con Categorías Prototípicas Deformables Borrosas el número de subdivisiones es indeterminado. Esta aproximación indefinida es mucho más cercana a la realidad, sobre todo en el ámbito documental donde existe una gran diversidad de colecciones y agrupaciones dentro de ellas cuya representación suele ser compleja.

Para el caso de los prototipos de agrupaciones documentales se procederá en dos fases. En la primera fase se ejecutará el clustering jerárquico hasta obtener un cierto umbral de similitud (media de los valores). Luego se agregarán los valores de cada uno de los grupos con lo que obtendremos el centro de cada uno de los prototipos obtenidos.

Teóricamente, estos grados obtenidos vendrían a ser los centros de los números borrosos triangulares que representarían a los prototipos, pero esto depende del número de prototipos obtenidos. Si como resultado de este proceso el número de prototipos obtenidos es mayor que un número razonable entonces se procede a fusionar aquellos números borrosos que se consideren muy solapados.

Esta labor de fusión debe realizarse de forma automática ejecutando una vez más el algoritmo de clustering jerárquico sobre los centros de los números borrosos obtenidos.

Cada uno de los grupos obtenidos tras la realización de esta segunda agrupación supondrá la construcción de un número borroso diferente y por lo tanto de un prototipo. Cada uno de esos números borrosos, será triangular si tiene un solo elemento que lo defina o trapezoidal en el caso de que venga definido por más de un elemento. En la Figura 1 se pueden observar un ejemplo de esta segunda fase de agrupamiento en la que se pasa de 10 prototipos a una representación agregada de únicamente tres.

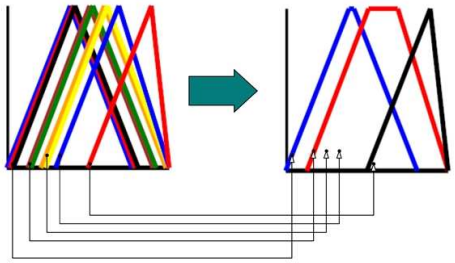


Figura 1: Obtención de los números borrosos.

El marco de definición de cada prototipo servirá para determinar las acciones especiales a realizar con cada documento entrante según su afinidad a la categoría en la que sea clasificada. A cada uno de los prototipos le corresponderá una definición paramétrica de acciones.

Los marcos están conformados por pares acción-probabilidad. Las acciones que se pueden ejecutar ante la llegada de un documento entrante son las siguientes: leer inmediatamente, clasificar, enfatizar, poner en espera, descartar, etc. Las probabilidades de acción o bien vendrán dadas de forma automática o definidas por el usuario mediante variables lingüísticas como las siguientes: seguro (100%), muy probable (75%), probable (50%), poco probable (25%), no realizable (0%). En la tabla 3 se puede observar un ejemplo de esta definición paramétrica

Tabla 3: Definición paramétrica de los prototipos.

Parámetro	**Prototipo 1**	**Prototipo 2**	**Prototipo 3**
Enfatización	No	Probable	Seguro
Eliminación	Probable	No	No
Puesta en Espera	Seguro	Probable	No
Clasificar	Probable	Seguro	Seguro

3. FILTRADO DE DOCUMENTOS ENTRANTES

El proceso de tratamiento de un documento entrante a un repositorio debe ser transparente el usuario. Este proceso se realizará en las siguientes etapas.

1. Obtención del grado de calidad de datos del nuevo documento a partir de las dimensiones anteriormente. Si el documento no supera un umbral mínimo de calidad de datos el documento no será procesado.
2. Preproceso lingüístico del documento, eliminación de stop words y extracción de raíces léxicas.
3. Representación de los conceptos reflejados en el contenido del documento.
4. Comparación conceptual entre cada una de las categorías y el numero documento. Para ello se utilizan, por un lado los conceptos existentes en el documento y por otro los conceptos relevantes y subrelevantes (notados por R y S respectivamente) de cada categoría. De esta manera se obtendrá un valor X de similitud conceptual entre documento y categoría. El cálculo se realizará mediante la fórmula 5:

$$X = \frac{Ln\left(\sum_{i \in d,c} w_i^d * w_i^c\right)}{Ln\left(\left|D\right| * \left(\left|(R+S)\right| - \left|(R \cup S) \cap D\right|\right)\right)} \quad (5)$$

5. Agregación al valor de similitud conceptual el valor de calidad de datos obtenido para el documento. Obtención de un valor X' que supone una similitud enriquecida.
6. Determinación de la afinidad a las diferentes Categorías Prototípicas Borrosas (Figura 2).

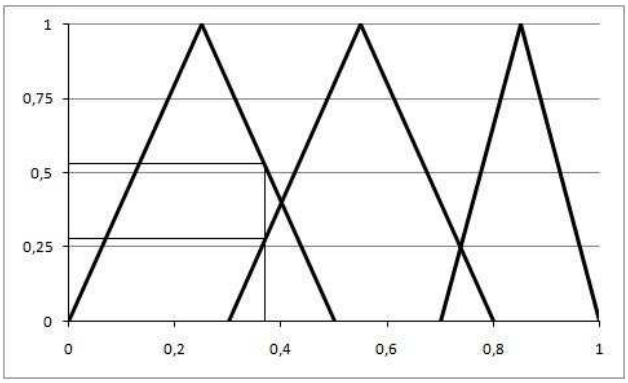


Figura 2: Afinidad calculada de un documento (X'=0,37) a los números borrosos de una categoría (μ_1=0,53, μ_2=0,28).

7. Determinar el prototipo correspondiente al documento mediante la modificación de los prototipos originales. Este proceso se realiza mediante una combinación lineal modificada de éstos mediante el grado de afinidad a los prototipos y los parámetros que los definen.

4. EXPERIMENTACIÓN

Para la evaluación del proceso de filtrado de información se ha utilizado la colección Reuters-21756 [10], que es una colección de noticias que han sido formateadas y preparadas para su utilización en la evaluación de sistemas de categorización de documentos [10]. Con el fin de evaluar el modelo de filtrado es necesario trabajar con dos subconjuntos de documentos, uno de entrenamiento del sistema y otro de evaluación. En los trabajos existentes en la literatura se utiliza normalmente divisiones estándar de la colección, principalmente la división ModApte [1].

El experimento se ha realizado en dos etapas principales, el entrenamiento del sistema y la evaluación en sí. El entrenamiento del sistema se ha realizado en los siguientes pasos:

1. Extraer los documentos de Reuters que cumplen la división ModApté y que corresponden a la subdivisión de entrenamiento.
2. Realizar la medición de los criterios de calidad de datos correspondientes sobre los documentos de entrenamiento.
3. Descartar aquellos que no superen una calidad mínima.
4. Construir la estructura borrosa y conceptual necesaria para el filtrado por categorías. Se utiliza para ello la clasificación dada por la propia colección de Reuters (clasificación ideal).
    a. Extracción de conceptos relevantes.
    b. Determinación de los grados de pertenencia de los documentos en base a sus conceptos y a su calidad.

La fase de evaluación del modelo de filtrado se ha ejecutado en los siguientes pasos para cada documento:

1. Cálculo de las dimensiones de calidad de datos para el documento entrante.
2. Eliminación del documento si no cumple los criterios mínimos de calidad.
3. Comparación del documento entrante con las categorías existentes en el modelo. Obtención del grado de comparación.
4. Calcular el grado de compatibilidad con los prototipos y determinar que acciones realizar.
5. Comprobar que las acciones corresponden a las teóricamente establecidas.

Una vez filtrados y clasificados todos los documentos de la colección de evaluación, se comparan las clases obtenidas con las clases definidas de forma teórica. De esta forma se pueden calcular los siguientes parámetros:

- ni: Número de documentos que se han clasificado en la categoría i.
- ti: Número de documentos que teóricamente deberían haberse clasificado en la categoría i.
- ci: Número de documentos que se han clasificado correctamente en la categoría i.

Una vez obtenidos estos parámetros se pueden calcular las dos métricas fundamentales en sistemas de recuperación [17] de información que son:

- Precisión: Proporción entre los documentos clasificados correctamente dentro de la categoría y el total de elementos clasificados en ella.

$$p_i = \frac{c_i}{n_i} \quad (6)$$

- Exhaustividad (recall): Proporción entre el número de documentos clasificados correctamente dentro de la categoría j y el total de elementos de los que dispone teóricamente la categoría.

$$r_i = \frac{c_i}{t_i} \quad (7)$$

Dada la relación inversa entre precisión y exhaustividad, se establece a su vez la medida F [17] que combina en su formulación la precisión y el exhaustividad (formula 8). Los valores más altos indicarán un mejor resultado

$$F(i, j) = \frac{2 * r_{ij} * p_{ij}}{r_{ij} + p_{ij}} \quad (8)$$

En la Tabla 4 se puede observar los valores de precisión, exhaustividad y métrica F obtenidas en el experimento para las diferentes categorías del modelo.

Tabla 4:Resultados del experimento.

categoría	p	r	F
acq	94%	95%	94%
coffee	69%	89%	77%
crude	85%	73%	79%
earn	99%	90%	94%
grain	93%	86%	89%
interest	71%	83%	76%
money-fx	87%	62%	72%
ship	76%	67%	71%
trade	89%	76%	82%
TOTAL			83%

La medición de la precisión y exhaustividad en el experimento ha arrojado valores suficientemente altos lo cual nos lleva a considerar a las técnicas elegidas como herramientas eficientes para llevar a cabo tareas de filtrado de información.

A su vez, el resultado ofrecido en cuanto a la métrica F (F = 83%) es bastante satisfactorio, si se compara con el obtenido sin utilizar criterios de calidad de datos (F = 80%).

5. CONCLUSIÓN Y TRABAJO FUTUROS

En este trabajo se ha presentado una aproximación al problema del filtrado de información en base a Categorías Prototípicas Borrosas. La creación de una estructura jerárquica y borrosa conformada por diferentes agrupaciones de documentos que definen mediante conceptos a estas categorías proporciona una estrategia eficiente de filtrado de documentos ante una actualización constante de los contenidos. Al análisis conceptual de los documentos se le ha incorporado un imprescindible análisis en base a criterios de calidad de datos que proporcionan la seguridad de que la información filtrada va a ser suficientemente relevante para el usuario.

Como primera ampliación futura de este trabajo se realizarán nuevos experimentos utilizando colecciones de mayor volumen como puede ser RCV1[11]. A su vez se incluirán en el modelo de calidad de datos otros aspectos relacionados con la relevancia como la exactitud del documento, la presentación, la comprensión, etc, ampliándose de esta manera los criterios a utilizar en el filtrado.

En el apartado teórico se ve necesario la definición completa de operadores de agregación (OWA [20]) para la formalización de las diferentes funciones de agregación que se utilizan en el proceso de filtrado.

Dentro de un enfoque más práctico se persigue la aplicación de estas técnicas en entornos reales como puede ser el filtrado de documentos clínicos según el formato CDA [8].

Agradecimientos

La investigación realizada está financiada por el proyecto TIN2007-67494 del Ministerio de Educación y Ciencia, el proyecto PAC06-0059 y PBI06-0024 de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1] C. Apté, F. Damerau, S.M. Weis: "Automated Learning of Decision Rules for Text Categorization", ACM Transactions on Information Systems, 1994.

[2] C. Batini. M. Scannapieco "Data Quality: Concepts, Methodologies and Techniques". Berlin, Springer-Verlag Berlin Heidelberg, 2006.

[3] C. Bizer, R. Cyganiak,. "The TriQL.P Browser: Filtering Information using Context-, Content- and Rating-Based Trust Policies". 4th International Semantic Web Conference, Galway, Ireland, 2005.

[4] G. Bordogna; M. Pagani, M.; G. Pasi: "A dynamical Hierarchical fuzzy clustering algorithm for document filtering", Soft Computing in Web Information Retrieval: Models and Applications, Series Studies in fuzziness and Soft Computing , Springer Verlag, Vol. 197, pp. 1-23, 2006.

[5] I. Caballero, E. M. Verbo, "A Data Quality Measurement Information Model based on ISO/IEC 15939". 12th International Conference on Information Quality, MIT, Cambridge, MA, 2007.

[6] Guha, R., R. McCool,. "Semantic Search". World Wide Web, Budapest (Hungary), ACM, 2003.

[7] E. Herrera-Viedma, G. Pasi. "Fuzzy approaches to access information on the Web: recent developments and research trends", Proc. of the Third Conference of the EUSFLAT, pp. 25-31, 2003.

[8] H.S. Kim, T. Tran, "A Clinical Document Architecture (CDA) to Generate Clinical Documents within a Hospital Information System for E-Healthcare Services" Proceedings of the Sixth IEEE International Conference on Computer and Information Technology (CIT'06) - Volume 00 IEEE Computer Society: 254 2006

[9] Y. W. Lee, L. L. Pipino, "Journey to Data Quality" Cambridge, MA, USA, Massachussets Institute of Technology, 2006.

[10].D.D. Lewis, "Reuters-21578 text categorization text collection 1.0" . http://www.research.att.com/~lewis.

[11] D. D. Lewis "Reuters collection RCV1" http://www.daviddlewis.com/resources/testcollections/rcv1/.

[12] E.H. Mamdani; S. Assilian; "An experiment in linguistic synthesis of fuzzy controllers". International Journal Man-Machine. Studies, vol. 7, pp. 1–13, 1975.

[13] J.A. Olivas; P.J. Garcés; F.P. Romero: "An application of the FIS-CRM model to the FISS metasearcher: Using fuzzy synonymy and fuzzy generality for representing concepts in documents". Int. Jour. of Approx. Reasoning (Soft Computing in Recognition and Search) Vol. 34, pp. 201-219, 2003.

[14] G. Pasi,. "Flexible information retrieval: some research trends", Mathware and Soft Computing Vol 9, pp 107-121, 2002.

[15] F.P. Romero; J.A. Olivas; P.J. Garcés: "A soft Approach to Hybrid Models for Document Clustering". Proceedings of the Information Processing and Management of Uncertainty in Knowledge-based Systems IPMU'06, Paris Les Cordeliers, France. Vol 1, pp. 1040 - 1045, 2006.

[16] F.P. Romero; J.A. Olivas; P.J. Garcés: "Inference Based On Fuzzy Deformable Prototypes For Information Filtering In Dynamic Web Repositories", Proceedings of the FUZZ-IEEE 2007., Londres, GB. 2007.

[17] C. J. van Rijsbergen. "Information Retrieval". Butterworths, London, 2nd edition, 1979.

[18] G. Shankaranarayanan, Y. Cai. "Supporting data quality management in decision-making." Decision Support Systems Vol. 42, nº 1: pp. 302-317, 2006.

[19] L.H. Ungar, D.P. Foster. "Clustering Methods for Collaborative Filtering". Proceedings of the Workshop on Recommendation Systems, AAAI Press, Menlo Park California, 1998.

[20] Yager, R.: "On Ordered Weithted Averaging Aggregation Operations in Multicriteria Decision making". IEEE Transactions on Systems, Man and Cybernetics Vol. 18, pp. 183-190, 1988.

[21] L. A. Zadeh, "A note on prototype set theory and fuzzy sets". Cognition 12, pp. 291- 297, 1982.

Arquitectura para la integración de esquemas relacionales difusos basada en ontologías: una aplicación para la web

Ignacio J. Blanco[1] **Carmen Martínez-Cruz**[2] **M. Amparo Vila**[1]

[1] Dpto. Ciencias de la Computación e I. A., Universidad de Granada
C/ Periodista Daniel Saucedo Aranda S/N, 18071, Granada
{iblanco,vila}@decsai.ugr.es

[2] Dpto. Informática, Universidad de Jaén
Campus de Las Lagunillas, 035-A3, 23071, Jaén
cmcruz@ujaen.es

Resumen

En el presente trabajo, se plantea una arquitectura que propone la incorporación de una ontología como interfaz para la representación de esquemas de Bases de Datos Difusas (BDD) sobre un SGBD (Sistema Gestor de Bases de Datos) cualquiera. Esta arquitectura especifica la implantación de dichos esquemas en diversas implementaciones relacionales, teniendo en cuenta las particularidades de cada implementación planteada.

Palabras Clave: Ontología, Base de datos Relacional, Web Semántica.

1 Introducción

Las bases de datos relacionales se han instaurado como herramientas imprescindibles para el almacenamiento y consulta de grandes cantidades de información. Sin embargo, los métodos de consulta requieren cada vez una mayor expresividad y potencia de lenguaje.

Las ampliaciones difusas del modelo relacional han permitido la representación de conceptos imprecisos dentro de una base de datos relacional [9] proporcionando, en algunos casos, estructuras de catálogo relacional adicionales para su almacenamiento [10].

Estas estructuras de catálogo son manipuladas de forma automática por extensiones del sublenguaje DML del lenguaje relacional SQL, las cuales incorporan una forma unificada para la definición de tablas que incluyan columnas definidas sobre dominios imprecisos.

[1]El desarrollo de este trabajo ha sido financiado parcialmente por los Proyectos TIN2006-15041-C04-01 y P06-TIC-01433

La extensión de una implementación del estándar de lenguaje SQL está ligada a un sistema relacional específico, lo cual hace imposible su uso en otras implementaciones relacionales. Por ejemplo, la extensión *FSQL* [6] está ligada al sistema relacional implementado por Oracle®, e implementada en el lenguaje de programación provisto dentro del propio sistema relacional (Oracle® PL/SQL) [2].

Sin embargo, el hecho de la inexistencia de una extensión del sublenguaje DML para otras implementaciones relacionales no implica que no puedan ser usadas para almacenar la misma información imprecisa que se almacenaría en una base de datos relacional extendida con la funcionalidad FSQL [4].

Nuestra aproximación se basa en una representación de la información imprecisa usando una ontología que actúe de interfaz entre el usuario y el SGBD que almacenará en última instancia la información. Esta ontología se divide en dos [3]: una donde se representa la estructura de los datos en el modelo relacional extendido para datos difusos (*Ontología del Catálogo*), y otra donde se representan las estructuras relacionales correspondientes al esquema de BDD que se define (*Ontología del Esquema*). De esta manera, la semántica de la información imprecisa se traslada a la ontología, dejando la tarea del almacenamiento de la instancia relacional al sistema de gestión de la base de datos concreto.

En el presente trabajo, presentamos una arquitectura para implantar la información contenida en una ontología del esquema en una base de datos relacional. Dicha arquitectura no dependerá de la implementación relacional en la que se implante el esquema relacional.

En el apartado 2, se describe brevemente la ontología en la que se basa la arquitectura que se propone y se añade un ejemplo ilustrativo del uso de dicha ontología, y como sería representada utilizando un lenguaje web. En el apartado 3, se examinan las posibilidades de integración de diversos tipos de Sistemas

Gestores de Bases de Datos Relacionales, con distintas características, en la arquitectura propuesta. En el apartado 4, describiremos las ventajas de la propuesta incidiendo sobre su aplicabilidad en la web. Para concluir, expondremos las conclusiones del presente trabajo y las líneas de desarrollo que surgen a partir del mismo.

2 Ontología para la representación de datos difusos

Las ontologías [8, 12, 14] han aparecido en el campo de la inteligencia artificial como uno de los mecanismos de modelado de datos mas utilizados en la actualidad. El gran éxito de las mismas se basa fundamentalmente en su aplicación a la web.

Debido a la ingente cantidad de información que se encuentra descrita en la web, se ha visto la necesitad de utilizar herramientas que permitan dotar de semantica a la información que se encuentra en la misma. Las ontologías, han proporcionado dicha semántica, dado que el conocimiento que representan puede ser modelado utilizando lenguajes web específicos (OWL [1], RDFS[15]) y de esta forma el código puede ser inscrustado en dicha web (bien en el mismo documento, bien en el sitio web como otro documento)[4, 7].

En [3], se describe el uso de una ontología para la representación y el manejo de la información imprecisa que puede ser almacenada en una base de datos relacional. La representación de dicha información, requiere que la ontología modele el catálogo de datos siguiendo un estándar (ANSI SQL) y la extensión de datos difusos propuesta en GEFRED [11, 10]. Este modelado deberá permitir la representación del esquema de BD Difusas (BDD) y la representación de los valores que se incluirán en dicho esquema. Esta realidad que está claramente separada en un SGBD cualquiera (catálogo y espacio de tablas), requiere del establecimiento de dos subontologías en esta propuesta:

- **Ontología del Catálogo**: ontología basada en la planteada por Calero [5] y extendida para contener todos los tipos de datos difusos representables en una base de datos de tipo FSQL. Esta ontología permitirá describir completamente el esquema de la base de datos que se desee representar mediante la instanciación de la misma. A su vez, en la ontología habrá metadatos definidos (datos que representan la estructura de la información), concretamente metaclases, para permitir que las instancias de las mismas, se conviertan en clases, que a su vez puedan ser instanciables, es decir, para posteriormente poder definir tuplas a través de la ontología (descrita a continuación). El esquema del catálogo esta descrito con detalle en [4] y [3].
- **Ontología del Esquema de BDD**: el esquema definido en la ontología del catálogo se transformará en una nueva ontología consistente en nuevas clases que se corresponderán con tablas, y sus atributos que se corresponderán con propiedades. Dicha ontología, generada, entre otros procedimientos, gracias a la definición de las metaclases, permitirá la definición de las tuplas mediante la instanciación de la misma. Por tanto, cada ontología del esquema de datos, representará un esquema concreto de BD, independiente del SGBD donde vaya a ser representado.

2.1 Ejemplo

Para ver con más claridad esta propuesta nos plantearíamos la definición del esquema que aparece en la figura 1 en forma de ontología:

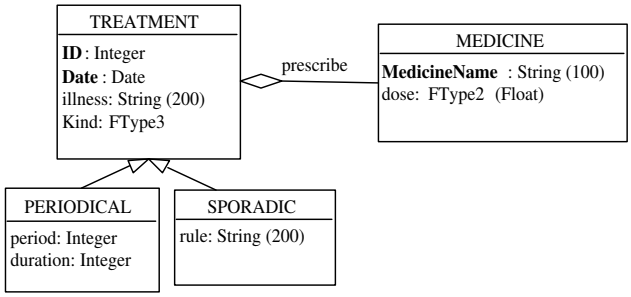


Figura 1: Arquitectura integrada

Dicho esquema sería definido sobre la ontología del catálogo, descrita en [3], mediante la instanciación de las clases *BaseTable, FuzzyColumn, UniqueColumn, BaseColumn* (para definir tablas y columnas difusas, únicas o normales respectivamente), *PrimaryKey, ReferencialConstraint* (para definir restricciones de clave primaria o ajena) y *FType1, FType2,FType3, Float, Numeric, etc.* (para definir los tipos de datos difusos o cl'asicos). En las figuras 2 y 3 aparecen ejemplos del código en OWL que definiría partes del esquema propuesto de ejemplo.

Sin embargo, esto es sólamente el esquema. La definición de datos, requeríría que las intancias de tablas, que anteriormente han sido definidas (instancias de *BaseTable*), fueran tratadas como clases. Esta formaría la *Ontología del Esquema de BDD*, y un ejemplo de instanciación de dicha ontología, lo proponemos en la figura 4. La definición de esta ontología del esquema implica la creación explícita de las restricciones sobre los datos previamente definidas como instancias

```
DEFINICION DE TABLAS

<fdtscho :BaseTable rdf:ID="Treatment">
  < fdtscho :hasColumns rdf :resource="# ID">
  < fdtscho :hasColumns rdf :resource="#Date">
  < fdtscho :hasColumns rdf :resource="#Illness">
  < fdtscho :hasColumns rdf :resource="#Kind">
  < fdtscho :isConstrained rdf:resource="# PK_Treatment">
</ fdtscho : BaseTable rdf :ID="Treatment">

<fdtscho :BaseTable rdf:ID="Medicine">
  < fdtscho :hasColumns rdf :resource="#Name">
  < fdtscho :hasColumns rdf :resource="#Dose">
</ fdtscho : BaseTable rdf :ID="Treatment">
.....

DEFINICION DE COLUMNAS

<fdtscho :FuzzyColumn rdf :ID="Dose">
   < fdtscho :nameCol rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #string">Dose
   </ fdtscho :nameCol >
   < fdtscho :hasFuzzyType >
      < fdtscho :FType2 rdf:ID=" FDT_Medicine_Dose">
        < fdtscho :margin rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #float">0.5
        </ fdtscho :margin>
        < fdtscho :much rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #float" >3.0
        </ fdtscho :much>
       < fdtscho :hasNumericType rdf:resource="#DT_Medicine_Dose"/>
      </ fdtscho :FType2 >
   </ fdtscho :hasFuzzyType >
</ fdtscho :FuzzyColumn >

<fdtscho :BaseColumn rdf :ID="illness">
   < fdtscho :hasDataType rdf:resource="#DT_ illness"/>
   < fdtscho :nameCol rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #string"
   >Illness</ fdtscho :nameCol >
 </ fdtscho :BaseColumn >

 < fdtscho :UniqueColumn rdf :ID="Periodical_ TreatmentID ">
   < fdtscho :hasReferences rdf :resource="# FKPeriodicalTreatment "/>
   < fdtscho :ordinalPos rdf :datatype ="http:// www.w3.org/2001/ XMLSchema #string"
      >0</ fdtscho :ordinalPos >
   < fdtscho :nameCol rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #string"
      >ID</ fdtscho :nameCol >
   < fdtscho :hasUniqueConst rdf :resource="# PKPeriodicalTreatment "/>
   < fdtscho :hasDataType rdf:resource="#DT_ID"/>
  </ fdtscho :UniqueColumn >
....
```

Figura 2: Código OWL: definición tablas y columnas

en la ontología del catálogo. De esta forma, por ejemplo, el atributo dose de Medicamento solo podrá contener datos de *FType2*, y el atributo *ID* y *Date* de *Treatment_Sporadic*, será tomado de aquellos definidos previamente en *Treatment* (dado que es clave ajena). Como podemos ver ambas ontología, estan interconectadas y son dependientes, igual que pasa en los SGBDs con la definición de esquemas en el catálogo y como tablas.

3 Arquitectura

Dependiendo del tipo de SGBD Relacional (SGBDR) en el que se quiera implantar la base de datos representada por la ontología, será necesario interponer una serie de fases que traduzcan las acciones recogidas de forma implícita en la ontología a las sentencias propias de un SGBDR.

```
DEFINICION DE RESTRICCIONES

< fdtscho :PrimaryKey rdf :ID=" PKTreatment ">
      < fdtscho :hasUniqueCol rdf :resource="#ID"/>
      < fdtscho :hasUniqueCol rdf :resource="#Date"/>
</ fdtscho :PrimaryKey >
....
< fdtscho :ReferentialConstraint rdf :ID=" FK _ periodicalTreatment ">
  < fdtscho :referencedUnique rdf :resource="# PK _Treatment"/>
  < fdtscho :ObjectName rdf : datatype =
    "http:// www.w3.org /2001/ XMLSchema #string">
     FK _ preriodicalT </ fdtscho :ObjectName >
  < fdtscho :referencedColumn rdf :resource="# PeriodicalT _ID">
  < fdtscho :referencedColumn rdf :resource="# PeriodicalT _Date">
  < fdtscho :constraintTable rdf :resource="#Periodical_Treatment">
</ fdtscho :ReferentialConstraint >
...
```

Figura 3: Código OWL: Definición de restricciones de PK y FK

3.1 SGBDR + FSQL

La extensión difusa *FSQL* del lenguaje relacional SQL permite la creación y manejo de estructuras relacionales capaces de contener atributos con dominio difusos.

Un sistema con estas características incorpora sus propias estructuras de catálogo para la representación de los dominios difusos y para la creación de relaciones con datos difusos. Además, incorpora el conjunto de funciones necesarias para gestionar dichas sentencias FSQL. Por ello, parece que carece de sentido la implantación de la *ontología para la representación de datos difusos* dado que la definición de datos difusos en este caso se podría realizar utilizando directamente las sentencias FSQL, sin entrar en detalles de representación.

Sin embargo, para implantar las estructuras relacionales contenidas en la *ontología del esquema de datos difusos* es necesario generar las sentencias FSQL que creen dichas estructuras en la base de datos, empleando un módulo de traducción como el observado en la figura 5. Con esta arquitectura, traducimos la ontología al lenguaje FSQL y la sentencia traducida se envía a la base de datos para su procesamiento (que incluye traducción a SQL con capacidades funcionales) y ejecución.

3.2 SGBDR con capacidades funcionales

La implantación de un esquema de BDD en un SGBDR que carezca de la implementación *FSQL* pasa por la creación de las estructuras de catálogo relacional extendido de forma difusa, que permitirán almacenar la información referente a la *ontología para la representación de datos difusos*. Incorporando las relaciones de dicho catálogo, almacenamos la información

```
DEFINICION DE TABLAS

 <Treatment rdf:ID="1">
    <ID rdf:resource="# DT_ID ">
    <date rdf:resource="#DT_Date">
    <illness rdf:resource="# FDT _Illness">
    <kind rdf:resource="# FDT _Kind">
 </Treatment >

 <Medicine rdf:ID="2">
    <Name rdf:resource="#DT_Name">
    <dose rdf:resource="# FDT _Dose">
 </Medicine >
.....

DEFINICION DE TIPOS DE DATOS

  <fdtscho :Intervalar rdf:ID=" FDT _Dose">
     < fdtscho :a rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #float">2
     </ fdtscho :a>
     < fdtscho :b rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #float">4
     </ fdtscho :b>
 </fdtscho :Intervalar >

  <fdtscho :Simple rdf:ID=" FDT_Kind">
     < fdtscho :degree rdf:datatype ="http:// www.w3.org/2001/ XMLSchema #float">0.2
     </ fdtscho :degree>
     < fdtscho :discreteID rdf:resource: "# DF _Blando "
     </ fdtscho :discreteID >
 </fdtscho :Simple>
....
```

Figura 4: Código OWL: Definición de Tuplas

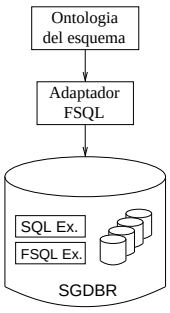


Figura 5: Arquitectura de integración con un SGBDR con capacidades FSQL

referente a los dominios y datos difusos contenidos en la base de datos, haciendo posible el acceso a la información desde las funciones y procedimientos que se creen a tal efecto pero sin depender de la ontología.

Sin embargo, toda la funcionalidad que incorpora la implementación del lenguaje *FSQL*, vista en el apartado 3.1, tendrá que ser proporcionada en este caso por una serie de bloques funcionales implementados en el lenguaje procedimental que proporcione la implementación relacional dada (Oracle® PL/SQL, PG/PLSQL).

Si bien proponemos un paso intermedio para la representación en *FSQL* de toda operación realizada a través de la ontología, en el caso de este tipo de sistemas, será necesaria la traducción de la sentencia al lenguaje SQL (el cual incluirá llamadas a funciones que resuelvan las cuestiones difusas) puesto que el sistema carece de la capacidad de compilar y ejecutar sentencias *FSQL*. El procesamiento de la sentencia FSQL tiene que ser realizado a nivel del *Adaptador a SQL con funciones*, mostrado en la figura 6, el cual traduce la sentencia FSQL a una sentencia SQL con llamadas a funciones que tendrán que estar almacenadas en la base de datos para su llamada cuando se ejecute la sentencia SQL traducida.

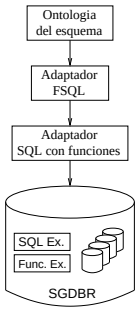


Figura 6: Arquitectura de integración con un SGBDR con capacidades funcionales

3.3 SGBDR sin capacidades funcionales

En este último caso, en el que se implanta el esquema de base de datos contenido en un SGBDR que no incorpora capacidades funcionales (es decir, que carece de un lenguaje de programación que le permite la creación de bloques de sentencias para la operación con datos), la base de datos actúa como mero recipiente de datos. De este modo, todas las fases del tratamiento de la sentencia *FSQL* (traducción y procesamiento) tendrán que ser implementadas en un lenguaje de programación externo al propio SGBDR.

Este es el caso de SGBDR como MySQL®, en el que se puede proveer el tratamiento de la sentencia *FSQL* mediante una serie de elementos funcionales programados en Java. Lo cual ralentizaría la ejecución de cualquier operación que requiriera de esta funcionalidad.

En un sistema de este tipo, el elemento denominado *Módulo funcional*, que se muestra en la figura 7, tendrá que realizar la interpretación de la consulta FSQL y, mediante diversas consultas SQL a la base de datos, obtener conjuntos de datos relacionales (no difusos) a los que aplicará funciones para proporcionar la semántica difusa a estos datos y poder operar con ellos.

3.4 Arquitectura unificada

Reuniendo todas las posibilidades de SGBDR vistas en los apartados anteriores, la arquitectura unificada

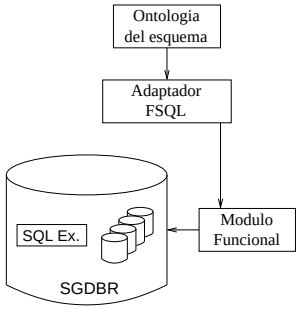


Figura 7: Arquitectura de integración con un SGBDR sin capacidades funcionales

quedaría como se muestra en la figura 8.

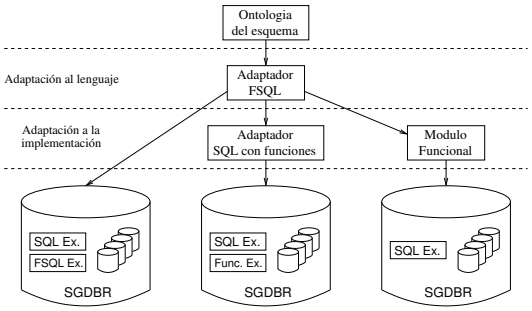


Figura 8: Arquitectura integrada

En dicha arquitectura, la ontología actuaría como interfaz abstracta para la definición de datos difusos o clásicos, independiente de cualquier SGBD que actue como recipiente de la información. Dicha ontología proporcionaría tanto la definición de los datos de la BDD, como cualquier petición de manipulación sobre los mismos, sin tener en cuenta las particularidades que cada SGBDD tiene en su representación.

Además, en la arquitectura separamos por niveles todos los aspectos relativos al lenguaje de consulta de datos difusos en el que nos basamos (como primer paso para la operación sobre la base de datos) de los aspectos específicos relacionados con la implementación concreta.

4 Aplicaciones en la web

Es frecuente encontrar páginas web cuyo objetivo es la exportación del contenido de bases de datos a Internet. Sin embargo, la mayoría de ellas proveen de un interfaz guiado de consulta a la base de datos a la que acceden. De este modo, el esquema de la base de datos subyacente tiene que ser entendido mediante la interpretación de la estructura de la página web que permite el acceso a los datos. Sin embargo, la mayoría estos interfaces (back-end) a una base de datos (también llamadas páginas web dinámicas), suelen estar carentes de semántica, que sólo se obtendrá una vez lanzada la consulta y obtenida la información[4]. Una ontología, anexada a estos formularios, que modele la información de la estructura de datos que dicha página está representando, permitiría que la interfaz pudiera ser indexada semánticamente y recuperada en cualquier consulta realizada a través de un buscador.

Sin embargo, es posible encontrar también páginas que, únicamente, proveen de acceso a una base de datos permitiendo al usuario consultar cualquier cosa en virtud de su capacidad de acceso y su conocimiento de los lenguajes de consulta. Es el ejemplo de un interfaz de consulta como el provisto por Oracle® iSQLPlus[13]. Este tipo de páginas no incorporan una descripción del esquema de base de datos subyacente, aunque se conozca el lenguaje de consulta.

De este modo, cuando un usuario busque una base de datos que le provea de cierta información, pasará por alto una página que contenga un interfaz de este tipo, puesto que no aporta información ninguna sobre lo que está buscando.

La posibilidad de asociar una ontología que describa la información (como la que se muestra en la figura 4) que hay en la BD subyacente a un interfaz web de este tipo, como se muestra en la figura 9, permitiría aportar semántica a dichas páginas, y así que pudieran estar incluidas en el resultado de la búsqueda realizada por un usuario.

Figura 9: Ontología de descripción de esquema de BD accesible por interfaz de consulta genérico

Por último, no podemos olvidar que la especificación explícita de esquemas de BD (difusas) a través de la web, permitirá compartir información entre fuentes de datos heterogéneos. Así, nos encontraremos con la posibilidad de combinar datos definidos utilizando es-

quemas de bases de datos, sino como esquemas de xml, ontologías, folksonomías, etc.

5 Conclusiones y trabajos futuros

El uso de esta ontología tiene las siguientes ventajas:

- hace posible la creación de esquemas de bases de datos relacionales que incluyan atributos de tratamiento difuso sin que el usuario tenga necesidad de conocer profundamente el lenguaje de manejo de dichos datos, y
- hace posible la replicación de esquemas de bases de datos relacionales en varios SGBDR.

El planteamiento de una arquitectura que traslade estos esquemas de base de datos relacional a implementaciones relacionales específicas (Oracle®, PostgreSQL®, MySQL®) sentando las bases para la incorporación de nuevas implementaciones supone un avance en el desarrollo de un sistema unificado de acceso a bases de datos relacionales que almacenan datos imprecisos.

Sin embargo, esta arquitectura tiene otras posibilidades que habrán de ser estudiadas. Al igual que puede ser usada para la replicación de esquemas en varios SGBDR, puede ser modificada para permitir la distribución de esquemas de bases de datos o de instancias.

Por otro lado, al igual que se ha usado la ontología para la representación del esquema de base de datos, puede usarse la misma para el manejo de la instancia del mismo, tanto a la hora de incorporar instancia (inserción), como de eliminar instancia (borrado) o de modificarla (actualización). Además, esta arquitectura, junto con la ontología, constituyen un interfaz para la consulta de datos sin que sea necesario el conocimiento del lenguaje de consulta relacional.

Referencias

[1] Grigoris Antoniou and Frank van Harmelen. *Handbook on Ontologies in Information Systems*, chapter Web Ontology Language: OWL, pages 67–92. Springer-Verlag, 2003.

[2] I. Blanco, C. Martinez-Cruz, J.M. Serrano, and M.A. Vila. A first approach to multipurpose relational database server. *Mathware and Soft Computing*, 12(2-3):129–153, 2005.

[3] I. Blanco, C. Martínez-Cruz, and M. A. Vila. *Handbook of Research on Web Information Systems Quality*, chapter Looking for Information in Fuzzy Relational Databases accessible via the Web, pages Chapter XVIII,300–324. Idea Group Reference, 2007.

[4] Ignacio J. Blanco, M. Amparo Vila, and Carmen Martinez-Cruz. The use of ontologies for representing database schemas of fuzzy information. *International Journal of Intelligent Systems*, 23(4):419–445, February 2008.

[5] Coral Calero and Mario Piattini. *An Ontological Approach to SQL:2003*, chapter Ontological Engineering: Principles, Methods, Tools and Languages, pages 49–102. Springer-Verlag, 2006.

[6] R. A. Carrasco, M. A. Vila, and J. Galindo. Fsql: a flexible query language for data mining. *Enterprise information systems IV*, pages 68–74, 2003.

[7] Tim Finin, James Mayfield, Anupam Joshi, R. Scott Cost, and Clay Fink. Information retrieval and the semantic web. In *HICSS '05: Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) - Track 4*, page 113.1, Washington, DC, USA, 2005. IEEE Computer Society.

[8] A. Gómez-Pérez, M. Férnandez-López, and O. Corcho-García. *Ontological Engineering*. Springer-Verlag New York, Inc., 2003.

[9] Zogmin Ma. *Fuzzy Database Modeling of Imprecise and Uncertain Engineering Information*. Springer, 2006.

[10] J. M. Medina, O. Pons, and M. A. Vila. Gefred. a generalized model of fuzzy relational databases. *Information Sciences*, 76(1-2):87–109, 1994.

[11] J. M. Medina, M. A. Vila, J. C. Cubero, and O. Pons. Towards the implementation of a generalized fuzzy relational database model. *Fuzzy Sets and Systems*, 75:273–289, 1995.

[12] Natalya F. Noy. *Handbook on Ontologies*, chapter Tools for Mapping and Merging Ontologies, pages 366–384. Springer, 2004.

[13] Oracle. Isqlplus web enviroment. http://150.214.108.124/isqlplus, January 2007.

[14] S. Staab and R. Studer. *Handbook on Ontologies*. Springer, 2004.

[15] W3C Recommendation, http://www.w3.org/RDF/. *Resource Description Framework (RDF)*, february 1999.

ALGORITMO DE RANKING DINÁMICO BASADO EN REALIMENTACIÓN IMPLÍCITA

Javier de la Mata **José A. Olivas** **Jesús Serrano-Guerrero**

Departamento de Tecnologías y Sistemas de la Información. Grupo SMILE.
Universidad de Castilla – La Mancha
Paseo de la Universidad, 4, 13071, Ciudad Real
{Javier.mata, JoseAngel.Olivas, Jesus.Serrano}@uclm.es

Resumen

Este trabajo presenta un algoritmo de ranking dinámico basado en la realimentación implícita del usuario. El sistema expande la consulta original incluyendo términos relacionados con los diferentes significados de los términos originales. Cuando el usuario selecciona un documento, el sistema favorece los documentos recuperados con consultas pertenecientes al mismo dominio semántico. De esta forma, la interacción del usuario contribuye a mejorar los resultados.

Palabras Clave: Meta-buscador, retroalimentación implícita, expansión de consulta.

1 INTRODUCCIÓN

La finalidad de los sistemas de búsqueda es recuperar la mayor cantidad posible de documentos que sean relevantes a una consulta del usuario y reducir en la medida de lo posible la cantidad de documentos no relevantes. La entrada al sistema de Recuperación de Información es una consulta en lenguaje natural normalmente formada por un conjunto de palabras clave. Esta forma de actuación suele proporcionar resultados adecuados en aquellos casos en los que se utiliza un número suficiente de términos en la consulta.

La recuperación de información en Internet requiere una estrategia de búsqueda adecuada, que represente de forma adecuada la necesidad de información del usuario. Pero es habitual que el usuario no aporte suficiente información al sistema de búsqueda para permitir afinar la búsqueda. Según el estudio realizado por Jansen [6] en torno al 62% de las consultas están formadas por 1 o 2 términos. Cuando se utilizan pocos términos aparecen problemas debido a los múltiples significados que puede tener una palabra, es decir, la polisemia. Así mismo, aparece otro inconveniente debido a las diferentes palabras utilizadas por los usuarios para describir el mismo concepto, conocido como problema del vocabulario [5].

Existen diferentes técnicas que intentan resolver las carencias de información proporcionada por el usuario. La más utilizada es la denominada expansión de consulta [4]. La técnica de expansión de consulta (o expansión de términos) es el proceso de incorporar en la consulta original términos relacionados bien de forma manual, automática o interactiva. Esta técnica construye nuevas consultas utilizando sinónimos de palabras, formas morfológicas o el reajuste de pesos de la consulta original. Mediante estas nuevas consultas se consigue aumentar el valor de exhaustividad ya que se recuperan más documentos, pero a expensas de una reducción en la precisión.

Para la realización de la expansión de consultas se necesita alguna fuente de conocimiento de la que obtener los nuevos términos que se incluyen en las consultas. Algunos sistemas recurren a bases de conocimiento previamente construidas como tesauros u ontologías de las que obtienen los términos relacionados semánticamente. Uno de los tesauros más utilizados en recuperación de información es WordNet [7].

Otros sistemas requieren la intervención del usuario en el proceso de búsqueda para indicar la relevancia de los documentos recuperados. El sistema utiliza la realimentación aportada por el usuario para extraer un conjunto de términos representativos de los documentos seleccionados y construir nuevas consultas utilizando estos términos. Esta técnica se denomina retroalimentación de relevancia [9]. El algoritmo más nombrado en la literatura que implementa esta técnica es

el algoritmo de Rocchio [8]. Una variación de esta técnica es la pseudo retroalimentación de relevancia, que no requiere de la intervención del usuario al asumir que los primeros documentos del listado de resultados son los más relevantes, y por tanto utiliza estos documentos para construir nuevas consultas.

Otros sistemas utilizan la interacción del usuario con el buscador como indicador de los documentos relevantes. Estos sistemas consideran como documentos relevantes aquellos que son consultados por los usuarios. Por tanto es necesario un proceso de monitorización del comportamiento del usuario. Este mecanismo se denomina realimentación por relevancia implícita.

La evaluación no es una tarea sencilla cuando no se dispone de un sistema de indexado de los documentos. Esta situación se da en los meta buscadores. Estas herramientas utilizan otros buscadores para obtener el conjunto de documentos. Su funcionamiento es muy parecido a los buscadores tradicionales de cara al usuario, pero en realidad tienen que ofrecer una evaluación basada en los resultados devueltos por otros buscadores. Además, a menudo utilizan diferentes motores de búsqueda, por lo que es necesario un proceso de mezcla previo.

Normalmente se suele aceptar que cuanto más frecuentemente aparece un documento en los listados de resultados obtenidos de diferentes fuentes para una misma consulta más probable será que ese documento sea relevante. Por tanto diferentes motores de búsqueda tienden a recuperar similares documentos relevantes para una consulta pero suelen diferir en los documentos no relevantes. Este fenómeno se denomina “efecto coro” [11]. En la práctica, la relevancia de un documento se refuerza cuando varios métodos de recuperación sugieren que ese elemento es relevante a una consulta.

En el marco del meta buscador GUMSe hemos desarrollado un algoritmo de evaluación basado en la realimentación que recibe del usuario. Cuando el sistema conoce los documentos que el usuario revisa es posible reajustar el ranking, permitiendo escalar posiciones a los documentos semánticamente relacionados con los documentos seleccionados.

Este trabajo está estructurado de la siguiente forma: la sección 2 presenta la arquitectura y funcionamiento general del meta buscador GUMSe. El algoritmo de evaluación se detalla en la sección 3. En la sección 4 se muestran unos resultados experimentales. Finalmente se termina con algunos apuntes del algoritmo propuesto.

2 ARQUITECTURA Y PROCCESO DE BÚSQUEDA EN GUMSe

Los meta buscadores proporcionan una forma de acceder a múltiples motores de búsqueda mostrando un único entorno de interacción para el usuario. Aprovechando las ventajas que proporcionan, hemos desarrollado GUMSe [3] como una plataforma de prueba que nos permita probar nuevas ideas aplicables al proceso de búsqueda en la Web. La arquitectura de GUMSe (Fig. 1) está compuesta de un conjunto de agentes que se comunican entre sí mediante una red. Cada uno de ellos tiene una función específica, lo que permite incorporar nuevos mecanismos en cada uno de ellos sin que afecte al resto.

El proceso de búsqueda seguido en este sistema se puede descomponer en una serie de fases coordinadas por al agente Search Broker. En cada una de estas fases se va refinando sucesivamente la carga semántica de la consulta para permitir expandir los resultados de la misma con documentos relacionados conceptualmente.

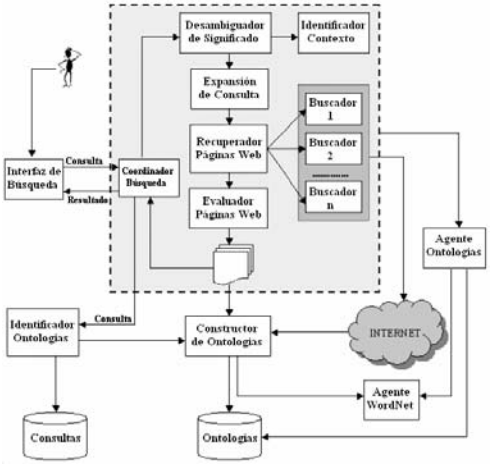


Figura 1. Arquitectura del metabuscador GUMSe

La primera etapa se denomina refinamiento del significado. Se trata de obtener la máxima información para reducir los efectos negativos producidos por la polisemia. Este fenómeno provoca una desviación en el cálculo de la relevancia, debido a los múltiples significados que pueden tener algunas palabras. Para tratar de solucionar este problema se utilizará el concepto de acepción o significado para permitir enfocar la búsqueda y mejorar los resultados. Se trata de obtener de forma manual o automática el significado que mejor se adapte a las intenciones de búsqueda del usuario.

La siguiente etapa es la expansión de consulta. En esta fase se utiliza la información adquirida previamente para aportar una mayor carga semántica al conjunto de consultas.

El siguiente paso, denominado recolección de la Web, consiste en lanzar las distintas consultas a una serie de motores de búsqueda para obtener una colección de

páginas entre las que se encontrarán los documentos más relevantes.

Por último se evalúan los resultados. Se tiene que construir un listado ordenado de documentos utilizando un único criterio de relevancia. En esta fase se enmarca el presente trabajo.

3 RANKING DINÁMICO

Los usuarios que utilizan los buscadores Web suelen utilizar un número reducido de términos en las consultas que utilizan. Existen trabajos que indican que los usuarios noveles utilizan 1 o 2 términos como mucho en un alto porcentaje de consultas [6]. Esta situación crea problemas de ambigüedad debido a la utilización de términos con varios significados, sobre todo cuando se utilizan pocos términos en la consulta.

En el meta buscador GUMSe se utilizan diferentes mecanismos de desambiguación para experimentar en que medida ayudan a mejorar los resultados de la búsqueda. Hasta ahora, los mecanismos se empleaban en el proceso previo a la búsqueda, bien mediante la colaboración interactiva del usuario (desambiguación interactiva) o de forma automática. Este último mecanismo utilizaba el historial de consultas desambiguadas de forma interactiva para desambiguar términos. Para ello se utilizaba la información suministrada por usuarios anteriores del sistema [1].

En este trabajo se presenta un nuevo mecanismo de evaluación de documentos que utiliza la interacción del usuario como indicador de la relevancia de los documentos. Se ha diseñado para resolver las situaciones en las que no es posible desambiguar correctamente como por ejemplo consultas con un solo término polisémico. Estas situaciones se producen cuando el usuario no utiliza la desambiguación interactiva y no se tiene suficiente información para desambiguar de forma automática.

Cuando un usuario pincha en un documento del listado, se considera que ese documento está relacionado con las necesidades de información del usuario. La información utilizada para recuperar dicho documento se utiliza para orientar los resultados de la búsqueda hacia el dominio semántico relacionado con el documento seleccionado. Para incorporar este mecanismo es necesario capturar qué documentos del listado de documentos son revisados por el usuario. El dominio semántico al que pertenece un documento está determinado por el conjunto de consultas utilizadas para recuperar ese documento. En la etapa previa de expansión de consultas el sistema genera un conjunto de consultas formado por un número fijo de consultas para cada posible acepción de cada término de la consulta original. Se considera una consulta Q introducida por el usuario formada por n términos polisémicos:

$$Q = \{T_1, T_2, \ldots, T_n\} \text{ para } n >= 1$$

Cada término T_i tiene 1 o varios significados denotados de la siguiente forma:

$$T_i = \{S_1, S_2, S_3, \ldots, S_m\} \quad m >= 1$$

Por tanto, el número total de significados que forman la consulta original es:

$$NS = \sum_{i=0}^{n} Significados(T_i) \tag{1}$$

A partir de cada significado o acepción se genera un número fijo de consultas mediante la técnica de expansión de consultas. Teniendo en cuenta que cada término de la consulta T_i tiene uno o más significados, el número de nuevas consultas generadas viene dado por la siguiente fórmula:

$$NQ = MaxQ * NS \tag{2}$$

MaxQ es una constante que indica el número de consultas generadas para cada significado de un término. Actualmente MaxQ vale 10.

Para cada significado diferente de cada término se generan MaxQ consultas nuevas. Estas consultas se construyen añadiendo nuevos términos relacionados con ese significado a la consulta. Estos términos relacionados se extraen de WordNet. El término original se puede incluir o no, dependiendo del tipo de consulta. Solamente se expande cada vez utilizando un significado concreto de un término de la consulta. El resto de términos de la consulta original se incluyen tal cual. De esta forma, cada consulta potencia un significado concreto de alguno de los términos. Por tanto, los documentos obtenidos con esa consulta tendrán mayor relación con ese significado concreto. Esta información se utilizará después para potenciar los documentos obtenidos con las consultas de los documentos que el usuario revise.

Cuando se han construido las NQ nuevas consultas se envían al buscador correspondiente junto con la consulta original. De cada consulta se obtiene un listado de documentos formado por un número fijo de documentos (actualmente 50) para las nuevas consultas generadas y el doble (100) para la consulta original.

El siguiente paso es extraer de cada consulta los datos de cada documento. Se obtienen el título, la dirección URL, la posición en el listado y con que consulta se recuperó. Con esos datos, se procede a calcular la relevancia general

de cada documento. La relevancia se calcula según la siguiente formula:

$$RInicial\ (D) = \sqrt[\beta]{DQ} * \sum_{i=1}^{DQ} \sqrt[\alpha]{STG_i} * (MDR - P_i) * CSM_i \quad (3)$$

Donde DQ es el número de consultas en las que aparece el documento D en la respuesta, P_i es la posición que ocupa D en el listado de resultados de la consulta i (siendo la mejor posición la posición 1), CSM_i es la medida de similitud conceptual de la consulta i y STG_i es el grado de confianza que se tiene en el buscador que realiza la consulta i. El parámetro α se utiliza para suavizar la dependencia de los resultados del buscador, sin dejar de tenerlo en cuenta. En nuestros experimentos toma el valor 2.

El valor CSM es el Grado de Similitud Conceptual e indica la afinidad semántica de la nueva consulta generada con la original. Este valor depende de los términos que se utilizan en la nueva consulta. Cuanto más relacionados estén con los términos originales mayor será su valor. Por ejemplo, utilizar sinónimos en la nueva consulta implica un alto valor de CSM para esa consulta.

También se calcula el grado de afinidad semántica de cada documento con cada uno de los posibles significados. Si un mismo documento se obtiene en consultas pertenecientes a diferentes dominios semánticos se calcula para cada uno de ellos el grado de pertenencia. Para cada documento se calcula un vector de significados, que refleja las veces que el documento aparece en las consultas relacionadas con dicho significado. El sistema calcula el grado de pertenencia normalizado del documento para cada significado de la siguiente forma:

$$VS_i(D) = \frac{QDom_i * \sum_{j=0}^{NQDom} (MaxD - P_j) * CSM_j}{MaxQ * \sum_{j=0}^{NQDom} CSM_j} \quad (4)$$

$$0 \le i \le NS$$

Donde QDom es el número de consultas relacionadas con el i-ésimo significado en las que aparece el documento. MaxD es el máximo número de documentos que devuelve la consulta (actualmente 50). P_j es la posición del documento en la consulta j y CSM_j es la similitud de la j-esima consulta con la original introducida por el usuario.

La longitud del vector de significados es igual a la suma del número de significados de cada término de la consulta original (veáse fórmula 1) más uno. La posición cero indica el grado de pertenencia con la consulta original. Este elemento se considera para tener en cuenta los documentos que no están relacionados con los significados contemplados en WordNet.

Se utiliza la realimentación proporcionada por el usuario para determinar que significados se ajustan mejor a sus necesidades de información. Conociendo que documentos revisa el usuario, se puede conocer que consulta o consultas se utilizaron para obtener ese documento y por tanto conocer que significado potenciaba esa consulta.

El sistema utiliza un vector de documentos seleccionados para recalcular la posición de los documentos en el ranking. Inicialmente todas las posiciones de este vector valen cero. Conforme el usuario va seleccionando documentos, el vector de documentos seleccionados almacena la media de los vectores de significados de los documentos seleccionados.

$$VDS_i = \frac{\sum_{j=0}^{NDSel} VS_i(D_j)}{NDSel} \quad (5)$$

$$0 \le i \le NS$$

NDSel es el número de documentos que el usuario ha revisado de la consulta actual. Cuando el usuario selecciona un documento, el sistema recalcula el vector de conceptos utilizando la fórmula anterior. Esta información se utiliza para recalcular el ranking de cada documento. Una vez que se ha seleccionado algún documento se recalcula el ranking para cada documento. En el cálculo del nuevo ranking se tiene en cuenta el ranking inicial y la información almacenada en el vector de documentos seleccionados. Para recalcular el ranking del documento D se aplica la siguiente fórmula:

$$R(D) = \left(1 - \frac{NDSel}{MaxDSel}\right) * RInicial(D) + \left(\frac{NDSel}{MaxDSel}\right) * \sqrt{DQ} * \sum_{i=0}^{NS} VS_i(D) * VDS_i \quad (6)$$

MaxDSel es un valor configurable que determina cuantos documentos tiene que revisar el usuario para dejar de tener en cuenta el ranking inicial. DQ es el número de consultas en las que aparece el documento D. Una vez calculado el Ranking dinámico para cada documento, se procede a normalizar el nuevo ranking de cada documento.

4 EXPERIMENTO

En esta sección se muestra un experimento realizado utilizando el algoritmo de ranking dinámico. La consulta elegida para el mismo es “apache”. Es una consulta con un solo término, el cual tiene 3 significados posibles en WordNet:

1. Tribu de indios americana.

2. Gángster parisino
3. Lengua de la tribu apache.

Se ha elegido esta consulta por la peculiaridad del término Apache. WordNet presenta algunas deficiencias [2] como fuente de conocimiento para el proceso de búsqueda. Una buena base de conocimiento sobre el que se asienta el proceso de expansión de consulta conlleva una mejora en los resultados. Sin embargo, WordNet todavía adolece de carencias, como por ejemplo la no inclusión de significados relacionados con las nuevas tecnologías. Por ejemplo, WordNet no reconoce Apache como un servidor Web muy popular ni tampoco como un modelo de helicóptero de ataque.

Además, en este caso dos de los significados de Apache están íntimamente relacionados (el primer y segundo significado). Este es otro de los problemas de WordNet, denominado granularidad de los significados. Es difícil determinar que significados se pueden incluir dentro de uno o separarlos en varios significados. En este experimento, muchos de los documentos recuperados con consultas relacionadas con el primer significado también aparecían en consultas relacionadas con el tercer significado y viceversa. Este fenómeno contribuye a penalizar a los documentos recuperados únicamente con consultas relacionadas con el segundo significado debido a que la granularidad es muy fina. Una granularidad excesivamente fina como la de WordNet puede ser contraproducente en los sistemas de recuperación de información, por lo que a veces es útil agrupar significados similares en uno solo [10].

Tabla 1: Variación del ranking con 1 a 6 documentos seleccionados para el significado "Gángster parisino"

Pos	1 Doc	2 Doc	3 Doc	4 Doc	5 Doc	6 Doc
1	1	1	1	1	37	37
2	2	2	2	2	27	27
3	3	3	3	3	67	67
4	4	4	4	4	73	73
5	5	5	5	27	82	82
6	6	6	6	37	90	161
7	7	7	7	67	161	176
8	8	8	27	5	176	197
9	9	9	8	6	197	198
10	10	27	37	73	198	242
11	11	10	67	82	242	377
12	12	11	9	90	377	648
13	13	37	73	7	166	90
14	14	12	82	161	183	166
15	15	13	10	176	205	205
16	16	67	90	197	206	206
17	27	14	11	198	207	207
18	17	15	25	166	648	661
19	18	25	12	242	258	183
20	19	16	161	183	393	258

El proceso de búsqueda comienza cuando el usuario introduce la consulta. A partir de la consulta original se construyen 10 nuevas consultas por cada significado, utilizándose por tanto 31 consultas incluyendo la original. De la consulta original se obtuvo un listado de 100 documentos, mientras que del resto el listado estaba formado como máximo por 50 documentos, aunque algunas consultas devolvieron menos documentos. Se obtuvieron un total de 1527 documentos, de los cuales 234 aparecían en más de un consulta.

En la tabla 1 se muestran la variación del ranking cuando se han seleccionado de 1 a 6 documentos relacionados con el significado "Gángster parisino". Cada documento se identifica mediante la posición del ranking inicial. El valor de MaxDSel es 6 para este experimento, lo que conlleva que en el último caso sólo se tiene en cuenta en el ranking el grado de pertenencia al significado favorecido (en este caso únicamente el segundo significado). En la tabla 2 se muestran que documentos fueron seleccionados. Para cada documento se muestra su vector de significados, pudiéndose apreciar que los seis documentos están fuertemente relacionados con el segundo significado de WordNet. Conforme se aumenta el número de documentos seleccionados se potencia el significado correcto y van apareciendo en las primeras posiciones del ranking los documentos relacionados con ese significado de forma progresiva.

Tabla 2: Vector de significados de cada documento seleccionado y variación del vector de documentos seleccionados para el significado "Gángster parisino".

Núm. de Docs. Seleccionados	Documentos Seleccionados	Vector de Doc. Seleccionados
1	27 -> (0, 0, 0.75, 0)	(0, 0, 0.75, 0)
2	37 -> (0, 0, 0.96, 0)	(0, 0, 0.86, 0)
3	67 -> (0, 0, 0.86, 0)	(0, 0, 0.86, 0)
4	73 -> (0, 0, 0.77, 0)	(0, 0, 0.84, 0)
5	82 -> (0, 0, 0.72, 0)	(0, 0, 0.81, 0)
6	90 -> (0, 0, 0.70, 0)	(0, 0, 0.79, 0)

Los documentos relacionados con el segundo significado aparecen en malas posiciones en el ranking inicial debido a que los otros dos significados de "apache" están muy relacionados entre sí, lo que favorece que aparezca el mismo documento en consultas relacionadas con ambos significados. El ranking inicial favorece los documentos que aparecen en los resultados de varias consultas. Como puede verse, conforme se seleccionan documentos el ranking va incluyendo en sus primeras posiciones documentos relacionados con el significado correcto.

En la tabla 3 se muestra otro ejemplo en el que se han seleccionado otros seis documentos. La tabla 4 muestra que documentos han sido revisados, así como su vector de significados. Puede apreciarse que esos documentos han aparecido en consultas relacionadas con el significado primero y tercero de "apache", así como en la consulta original (posición 0 del vector). En este caso, la aparición de nuevos documentos en el ranking no es tan drástica como en el caso anterior debido a que las primeras

posiciones del ranking inicial incluyen documentos relacionados con los significados de los documentos seleccionados.

Tabla 3: Variación del ranking con 1 a 6 documentos para los documentos seleccionados en la tabla 4.

Pos.	1 Doc.	2 Doc.	3 Doc.	4 Doc.	5 Doc.	6 Doc.
1	1	1	1	1	1	1
2	2	2	2	2	2	4
3	3	3	3	4	4	9
4	4	4	4	3	3	2
5	5	5	5	5	9	12
6	6	6	6	6	6	8
7	7	7	8	8	8	76
8	8	8	7	9	5	42
9	9	9	9	7	12	103
10	10	10	10	12	7	99
11	11	11	12	10	10	84
12	12	12	11	11	16	112
13	14	13	13	16	11	35
14	13	14	16	13	42	95
15	15	16	14	42	76	16
16	16	15	15	35	103	6
17	19	19	19	76	99	3
18	18	18	18	84	35	159
19	20	20	35	99	84	160
20	17	22	22	103	13	254

Tabla 4: Vector de significados de cada documento seleccionado y variación del vector de documentos seleccionados.

Núm. de Docs. Seleccionados	Documentos Seleccionados	Vector de Doc. Seleccionados
1	2 -> (0.04, 0.32, 0, 0.44)	(0.04, 0.32, 0, 0.44)
2	4 -> (0.15, 0.58, 0. 0.13)	(0.10, 0.45, 0, 0.28)
3	9 -> (0, 0.68, 0, 0.17)	(0.06, 0.53, 0, 0.25)
4	13 -> (0.18, 0.50, 0, 0)	(0.09, 0.52, 0, 0.19)
5	33 -> (0, 0.35, 0, 0.22)	(0.07, 0.49, 0, 0.19)
6	35 -> (0, 0.55, 0, 0.24)	(0.06, 0.50, 0, 0.20)

5 CONCLUSIONES

En este trabajo se ha presentado un algoritmo de evaluación dinámico basado en la realimentación implícita proporcionada por el usuario del sistema. Este algoritmo se ha desarrollado para mejorar los resultados de la búsqueda en aquellos casos en los que el usuario utiliza términos ambiguos en su consulta y decide no utilizar mecanismos de desambiguación o cuando el sistema no tiene suficiente información para establecer una desambiguación automática.

Todavía quedan muchos aspectos por pulir, siendo el principal el diseño de la forma de interacción del sistema con el usuario, bien cambiando el listado cada vez que se seleccione un documento o hacerlo en cada nueva pantalla de resultados. También se pretende profundizar en un futuro en el cálculo de la similitud de las nuevas consultas generadas con la original.

Agradecimientos

La investigación realizada está financiada por el proyecto TIN2007-67494 del Ministerio de Educación y Ciencia, el proyecto PAC06-0059 de la Junta de Comunidades de Castilla-La Mancha.

Referencias

[1] de la Mata, J.; Olivas, J. A.; Serrano-Guerrero, J.: Desambiguación de Significados en GUMSe basada en el refinamiento del sentido. Actas del XII Congreso Español Sobre Tecnologías y Lógica Fuzzy, Pág 315-320, 2004.

[2] de la Mata, J.; Olivas, J. A.; Serrano-Guerrero, J.: The Gaps of the Thesaurus Wordnet Used in Information Retrieval. EUROCAST 2005, LNCS 3643, Pág. 229 – 234, 2005.

[3] de la Mata, J.; Olivas, J. A.; Serrano-Guerrero, J.: Overview of an Agent Based Search Engine Architecture. In Proceedings of the International Conference on Artificial Intelligence IC-AI'04, Vol. I., Pág. 62 – 67. 2004

[4] Efthimiadis, E.N.: Query Expansion. In Williams, M.E., (ed.): Annual Review of Information Systems and Technology (ARIST), Vol.31, Pág 121-187, 1996.

[5] Furnas, G. W., Landauer, T. K., Gomez, L. M. ; Dumais, S. T. : The vocabulary problem in human-system communication: An analysis and a solution. Comm. ACM, 30, 11, Pág. 964-971, 1987.

[6] Jansen, B. J., Spink, A., and Saracevic, T.: Real life, real users, and real needs: A study and analysis of user queries on the web. Information Processing and Management. 36(2), Pág. 207-227, 2000.

[7]Miller, G.: WORDNET: An Online Lexical Database, International Journal of Lexicography, 3(4), 1990.

[8] Rocchio, J.: Relevance feedback information retrieval. In Gerard Salton (ed.): The Smart Retrieval System - Experiments in Automatic Document Processing, Pág. 313–323, 1971.

[9]Salton, G., and Buckley, C. Improving retrieval performance by relevance feedback. Journal of the American Society for Information Science. 41. 4. Pág. 288-297, 1990.

[10] Snow, R.; Prakash, S.; Jurafsky, D.; Ng, A. Y.: Learning to Merge Word Senses. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Pág. 1005-1014, 2007.

[11] Vogt, C. C.: Adaptive Combination of Evidence for Information Retrieval. PhD thesis, University of California, San Diego, 1999.

BUDI: Plataforma para la búsqueda difusa de información en repositorios de documentos

Jesus Serrano-Guerrero, Francisco P. Romero, Javier de la Mata, Jose A. Olivas
Departamento de Tecnologías y Sistemas de la Información, Universidad de Castilla-La Mancha,
Paseo de la Universidad 4, 13071 Ciudad Real
{Jesus.Serrano, FranciscoP.Romero, Javier.Mata, JoseAngel.Olivas}@uclm.es

Resumen

La necesidad de cualquier usuario cuando consulta una base de datos documental, bien sea departamental o general como la Web, es obtener resultados relevantes y que le proporcionen el conocimiento necesario para poder cubrir sus requerimientos. Para ello es necesario que el sistema sea capaz realizar son una serie de procesos previos y diferenciados para la correcta extracción de información. La plataforma BUDI surge como respuesta a esta imperiosa necesidad del mundo actual con el objetivo de mejorar los resultados de los procesos de búsqueda actuales sobre repositorios de documentos mediante la utilización de herramientas para la extracción y uso de conocimiento procedente de tesauros y ontologías, modelado lingüístico borroso y algoritmos de clustering.

Palabras Clave: Lógica borrosa, metabuscador, recuperación de información, web semántica.

1. INTRODUCCIÓN

La búsqueda en la Web se ha convertido en una tarea compleja e ineficiente debido al gran volumen de datos que hay en la red y al crecimiento exponencial de la información que se produce. Esto se hace patente al ver como cada vez que un usuario realiza una búsqueda, éste recibe una información donde la mayor parte es irrelevante. El problema radica en que los principales buscadores de la red basan los criterios de búsqueda en aspectos léxicos con respecto a la consulta. El mismo problema se plantea en las medianas y grandes organizaciones (y cada vez más también en las pymes) que almacenan ingentes volúmenes de información provocando una gran dificultad para obtener una información concreta y precisa en un momento dado.

Este trabajo se presenta como evolución de los realizados durante los últimos años por el grupo de investigación SMILE (Soft Management of Internet and e-learning) [10] en este ámbito. Cabe destacar las aportaciones en cuanto a la definición del modelo de representación conceptual de documentos FIS-CRM y su utilización en FISS [9], fzMail [7] una herramienta de gestión de correo electrónico mediante técnicas de soft-computing y GUMSe [11] un meta-buscador basado en una plataforma de agentes que utiliza Wordnet para realizar procesos de tratamiento de las consultas.

BUDI es un proyecto que está empezando a fraguarse. Su objetivo principal es el desarrollo de una herramienta que realice búsquedas conceptuales que pueda trabajar no sólo sobre buscadores Web sino también sobre cualquier tipo de base de datos documental o motor de indexación (Terrier [8], SMART [2], etc.). El producto final será un componente software reutilizable en aplicativos o sistemas de información "verticales" en diferentes entornos (jurídicos, sanitarios, parlamentarios, etc.). De esta forma BUDI se integrará con sistemas empresariales que almacenan información de negocio interactuando, por ejemplo, con sistemas de gestión documental, bases de datos, sistemas Data Warehouse, etc., permitiendo la localización rápida y precisa de información. La estructura de BUDI se construirá mediante dos componentes principales (Figura 1):

1. **El núcleo operativo del sistema.** Estará formado por componentes capaces de capaz de indexar, realizar búsquedas, ordenar o clasificar respuestas y mostrar resultados. Todas estas tareas las realizará mediante técnicas de lógica borrosa, tratamiento semántico y procesamiento de lenguaje natural.
2. **Base de Conocimiento Semántica:** Para realizar las tareas de procesamiento semántico de consultas y resultados necesitará contar con un sistema constituido por diferentes conjuntos de términos, sinonimias, ontologías, etc. que identifiquen distintas disciplinas del saber (información jurídica, sanitaria, etc.). Este sistema tendrá una gestión única e independiente.

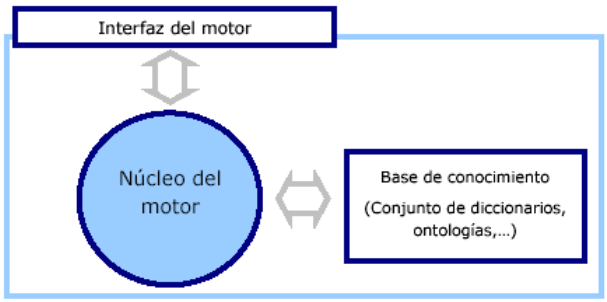


Figura 1: Arquitectura básica de BUDI.

La filosofía de funcionamiento de BUDI se basa en aplicar tecnologías de Soft-Computing para diseñar soluciones de búsqueda basadas en:

- Extracción masiva de información (Data-Mining).
- Clustering de los documentos recuperados basados en grupos temáticos, que pueden estar o no predefinidos [13].
- Reglas de asociación borrosa que ayudan al usuario a encontrar nuevos términos utilizables en su consulta.
- Representación de los documentos usando extensiones del modelo espacio vectorial estándar [9].
- Metabuscadores: Fusión de resultados ofrecidos por otros buscadores y/o motores de búsqueda específicos.[4]

La estructura del artículo es la siguiente. En el apartado 2 se explicará de forma genérica los diferentes módulos que conforman la arquitectura de BUDI. En el apartado 3 se detallarán las técnicas utilizadas para el procesamiento de consultas, mientras que el apartado 4 será el destinado a la explicación de los mecanismos de evaluación, ranking, clustering que se realizan sobre los resultados ofrecidos por los motores de búsqueda. El apartado 5 servirá para describir el módulo de gestión de recursos semánticos. Finalmente en el apartado 6 se presentarán las conclusiones y las líneas de desarrollo de esta plataforma.

2. ARQUITECTURA DE BUDI

La Plataforma BUDI está concebida para incorporar un valor añadido a la búsqueda en diferentes entornos documentales mediante la utilización de diversas técnicas de lógica borrosa. Su diseño está orientado a su aplicabilidad global no limitándose a su interacción con buscadores Web o bases de datos documentales. El objetivo es poder intercambiar con facilidad cualquiera de los componentes que conforman BUDI para poder adaptar el motor a diferentes escenarios, modificar por eficiencia un componente para aplicaciones concretas o añadir nuevas funcionalidades al motor.

La funcionalidad del sistema se organiza en módulos diferentes, agrupados por los tipos de tareas que deben realizar. De esta manera, se cuenta con una serie módulos operativos que se ocuparán de efectuar las tareas clave del sistema como la transformación de las consultas y la evaluación de resultados. Estos módulos interactúan con los elementos clave como el motor de búsqueda y la aplicación cliente que presenta los resultados al usuario. También existen una serie de módulos de gestión o soporte que se ocuparán de las tareas internas propias del sistema como la manipulación de los recursos semánticos y la configuración del sistema.

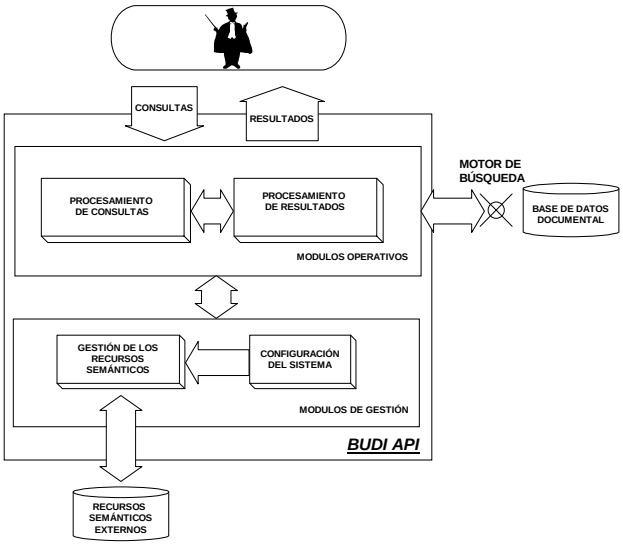


Figura 2: Arquitectura detallada de BUDI.

La lista de módulos de la aplicación es la siguiente:

- **Módulo de procesamiento de consultas:** Comprenderá los diferentes mecanismos de procesamiento de la consulta realizada por el usuario tales como la desambiguación de las partículas de búsqueda, la expansión de consultas, etc.
- **Módulo de procesamiento de resultados:** Se ocupa de, a partir de los resultados ofrecidos por un motor de búsqueda ante una serie de consultas, interpretar la relevancia de los documentos recuperados con respecto a las necesidades del usuario de forma que se presenten de una manera ordenada.
- **Módulos de gestión de la información semántica:** Da soporte a los dos módulos anteriores. Su función es gestionar los recursos semánticos externos, como tesauros, diccionarios y ontologías, que suministren la información semántica necesaria para realizar las tareas de búsqueda difusa.
- **Módulo de Configuración.** La configurabilidad del sistema es un objetivo clave en la plataforma por lo que existe un módulo específicamente dedicado a su gestión mediante la utilización de ficheros XML.

3. PROCESADOR DE CONSULTAS

El módulo de procesamiento de consultas se ocupa de transformar la consulta recibida con el objetivo de conseguir una mayor calidad de los resultados explotando toda la información semántica que tiene disponible.

Una de las principales tareas que realiza es generar nuevas consultas a partir de las palabras relacionadas semánticamente con las existentes en la consulta introducida por el usuario. Las nuevas consultas tienen un grado de compatibilidad con la consulta original obtenido a partir de los correspondientes grados de relación de los términos implicados. Previamente para cada término y sus posibles conceptos se realiza un proceso de desambiguación de los mismos. Esto intenta reducir al mínimo el número de búsquedas y resultados no útiles y por consiguiente mejorar la calidad de los resultados.

Las consultas a realizar se expresarán en un lenguaje genérico basado en XML. Dado que cada motor de búsqueda a utilizar tiene sus propias características y operadores de búsqueda concretos será necesario transformar estas consultas genéricas en peticiones específicas a cada motor. También la conexión a cada uno de estos motores será específica y diseñada a partir de conectores los cuales servirán para ejecutar la consulta y devolver los resultados.

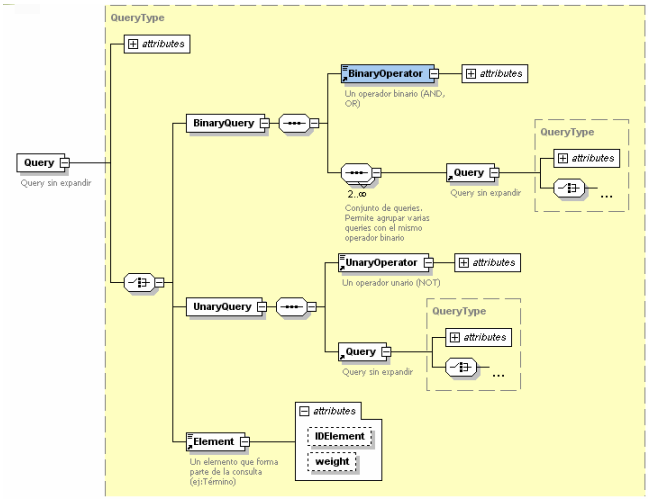


Figura 3: Esquema de las consultas a recibir.

3.1. RECEPCIÓN DE CONSULTAS

La aplicación cliente enviará la consulta que ha realizado el usuario a la API para que esta la procese y devuelva los resultados generados. Además de la consulta podrá ir información adicional que pueda ofrecer la aplicación cliente (idioma, etc.).

Dentro de esta fase de recepción se incluye en análisis de los operadores que hay en la consulta, los cuales vendrán definidos por la aplicación cliente. Los más básicos son los siguientes:

- Operadores lógicos: Este conjunto de operadores está constituido por los operadores clásicos boléanos "AND", "OR", "NOT".
- Operadores de exactitud: Este tipo de operadores se utiliza para indicar que lo que se busca es literal, y normalmente suele implementarse mediante la utilización de las comillas ‘”’.
- Operadores de truncamiento: Se aplican una serie de caracteres especiales en la consulta que se utilizan como comodines: *.$, ?, etc.
- Operadores de campo: Muchos buscadores permiten especificar mediante palabras clave distintos comportamientos o filtros a la hora de realizar la búsqueda.

3.1. DESAMBIGUACIÓN DE TÉRMINOS

El objetivo de la fase de desambiguación de términos es refinar la búsqueda para evitar la polisemia en las búsquedas basadas en palabras clave, que es una de las técnicas principalmente utilizada en los buscadores actualmente. Este inconveniente nace a raíz de los múltiples significados que pueden tomar los distintos términos que intervienen en una consulta, provocando una desviación en la relevancia de los resultados obtenidos [5].

Para tratar de solucionar este problema se utilizará el concepto de acepción o significado para permitir enfocar la búsqueda y mejorar los resultados. Se necesita por lo tanto un proceso automático de identificación de la acepción concreta basada en los términos participantes de la consulta. Esta forma de resolución se puede dar en caso de poseer una ontología que permita definir el contexto, así como una serie de términos relacionados semánticamente [15]. Por ejemplo, ante la consulta “altitud montaña más alta Europa” en caso de poseer una ontología sobre geografía se puede determinar la acepción correcta y poder evitarse así la devolución de resultados incorrectos o tener que interactuar con el usuario para refinar el significado de los términos.

3.1. EXPANSIÓN DE CONSULTAS

El objetivo es construir un conjunto de consultas que traten de obtener documentos que no sería posible recuperar mediante la consulta original, pero que están relacionados con ella. El propósito es obtener en las primeras posiciones del ranking documentos relacionados conceptualmente pero que utilizando la consulta original no se recuperan o lo hacen en posiciones de relevancia muy bajas. En definitiva se pretende diversificar la búsqueda probando con nuevas consultas.

A partir de la consulta original se crean N consultas expandidas a partir de las relaciones semánticas que pueden ser obtenidas del módulo gestor de recursos semánticos. Dado que ciertos términos pueden estar muy relacionados, este proceso de expansión está íntimamente relacionado con el proceso anterior de desambiguación de significados. Para cada una de las consultas creadas se calcula un factor para cada una de las consultas en función de los términos utilizados y su similitud con los originales. Este factor se denomina Medida de Similitud Conceptual (CSM o *Conceptual Similarity Measure*) y se utilizará para la posterior fase de ranking.

3.2 EJECUTOR DE CONSULTAS

Su misión es enviar las consultas a los buscadores utilizados y obtener el conjunto de documentos que posteriormente serán evaluados. Para ello es preciso adaptar las consultas expandidas al lenguaje de consulta concreto utilizado por cada buscador. Esta tarea de traducción la realiza este componente. Debido a que distintos buscadores utilizan distintos operadores, se ha decidido realizar las consultas expandidas en una representación en XML, que posteriormente es traducida de acuerdo con las características de cada buscador. Este componente incorpora los conectores a cada uno de los buscadores incorporados.

4. PROCESADOR DE RESULTADOS

Habitualmente, las técnicas de expansión de consulta aumentan el factor conocido como recall sin que lleve asociado un aumento del factor llamado precisión. Debido a este inconveniente, no basta con un mecanismo de refinamiento de las consultas para obtener un conjunto de documentos más relevantes para el usuario. Se requiere un posterior procesamiento de los distintos resultados obtenidos que implica la integración y reajuste de la relevancia con respecto a la búsqueda del usuario. Por este motivo, se incorporan mecanismos de evaluación que permitan aumentar la precisión de la búsqueda, reduciendo la importancia de los documentos alejados conceptualmente.

El objetivo de esta fase es la ordenación de un conjunto de documentos procedentes de la fase de procesado de consultas con respecto a la consulta introducida por el usuario. El procesamiento de los resultados de las consultas realizadas se realizará en las siguientes fases.

1. **Recepción de los documentos, consultas y motores de búsqueda utilizados:** La entrada a esta fase son los documentos resultantes, las consultas ejecutadas y los motores de búsqueda utilizados. Con estos datos pueden conocerse por cada motor de búsqueda qué consultas se han hecho y qué documentos fueron devueltos.
2. **Preprocesado de los documentos:** Los datos anteriores son analizados con la finalidad de detectar posibles documentos que han sido devueltos por distintos buscadores o por el mismo buscador pero por distintas consultas, y que lógicamente podrían ser tomados como documentos distintos cuando realmente son el mismo documento.
3. **Ranking:** Una vez analizados los datos anteriores hay que dar una puntuación a los documentos resultantes de la ejecución de la consulta. Este valor debe ser proporcional a la relevancia de cada documento a la consulta realizada. Como factores claves para determinar este grado de relevancia se determinan los siguientes:

   - La afinidad semántica de la consulta utilizada para recuperarlo: Se supone que cuanto mayor sea la Medida de Similitud Conceptual (CSM) de la consulta utilizada para recuperar un documento más relevante será ese documento.
   - La posición del documento en el listado de resultados: Los buscadores tienen sus propios mecanismos de evaluación, por lo que se asume que cuanto mejor situados estén en el listado más relevantes serán. Esto es posible en buscadores Internet que ya utilizan sus propios algoritmos de ranking como Google con Page Rank [14].
   - La valoración del buscador utilizado: Se tiene en cuenta la fiabilidad de los resultados proporcionados por cada buscador. Se supone que cuanto mayor número de documentos se le exigen a un buscador en detrimento del resto (Grado de Confianza o STG) más relevantes son sus resultados.
   - Efecto “coro”: Se basa en que habitualmente distintos buscadores suelen recuperar los mismos documentos relevantes ante una misma consulta, mientras que difieren en los no relevantes. Basándonos en este principio, se aceptan como documentos más relevantes aquellos que aparecen con más frecuencia en las respuestas obtenidas a partir de diferentes consultas.

   Para la implementación de este ranking se han realizado dos aproximaciones. Una de ellas está basada en el Sistema de Votación de Borda (*Borda Count*) [4], al cual se le han realizado ciertas modificaciones con el fin de tener en cuenta todos los factores definidos como clave además de reflejar ciertas evoluciones existentes en la literatura basadas en lógica borrosa como los aparecidos en [1] y [12]. Por otro lado, se ha seguido la especificación del algoritmo rápido de ranking utilizado en [11] que permite mediante una única fórmula la extracción del grado de relevancia de un documento según los factores explicados.

4. **Formateado de resultados:** Una vez se ha establecido el orden definitivo con que los documentos van a ser devueltos al usuario, se pasa a representar los documentos con un formato determinado que será el formato de salida que recibirá la aplicación cliente. Este formato vendrá condicionado por los datos aportados por los distintos motores de búsqueda, por ejemplo, el nombre del documento, el autor, la fecha, un resumen, etc.

5. **Ordenación:** Los datos que se generan como salida del algoritmo de ranking tienen un cierto grado que de relevancia con respecto a la consulta del usuario pero no forman una lista ordenada sino un conjunto, así pues con respecto a ese grado de relevancia se crea una lista ordenada de documentos. De esta manera, queda confeccionar una lista ordenada según la puntuación de cada resultado. Para este fin se ha usado el método de ordenación Quicksort.

6. **Retorno de resultados:** El resultado final de este módulo es una lista ordenada de documentos, que pueden tener distintas características en función de los motores de búsqueda de donde provienen los documentos. Así, no es lo mismo, un documento que provenga de un buscador de Internet que un documento proveniente de una base de datos documental. El primero tendrá características como un snippet y una URL, mientras que un documento de una base de datos documental tendrá datos relativos al autor, la fecha de creación, etc. Esta lista de documentos es devuelta a la aplicación cliente para que la gestione o visualice como crea conveniente.

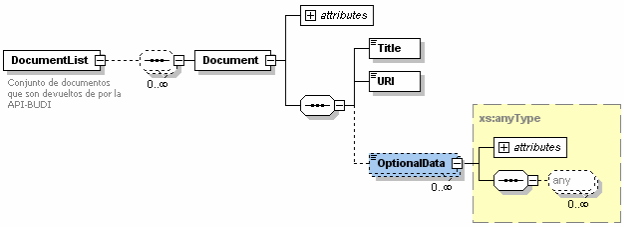


Figura 4: Estructura genérica de resultados.

5. GESTOR DE RECURSOS SEMÁNTICOS

La principal funcionalidad de este módulo es ofrecer los mecanismos de gestión de los recursos semánticos externos que facilitan la información semántica necesaria para llevar a cabo funcionalidades como la desambiguación, la expansión de la consulta o la evaluación de los resultados.

Existirán tres tipos principales de recursos:

- **Listas de palabras.** Son listas de palabras que pueden ser tratadas de forma negativa como las stop-words o de forma positiva como los lemarios.
- **Tesauros:** Serán gestionados para obtener todos los datos semánticos disponibles dado un término. Se pueden realizar cálculos de los grados de relación entre palabras.
- **Ontologías:** Se deberán poder gestionar diferentes ontologías, las relaciones entre sus partes y posibles grados de relación entre sus partes. Estas ontologías pueden ser genéricas, en algún formato específico o construidas ad-hoc de forma automática [3] o por expertos.

El módulo estará organizado en tres capas (Figura 5). La primera capa es la de servicios, que permitirá a los módulos operativos satisfacer las diferentes necesidades de información semántica que existan en sus procesos. La segunda capa será la de gestión u orquestación que permitirá traducir esas peticiones en diferentes consultas a los recursos externos que sean accesibles. Por último, existirá una capa de conexión que permitirá la consulta a diferentes recursos semánticos mediante conectores específicos.

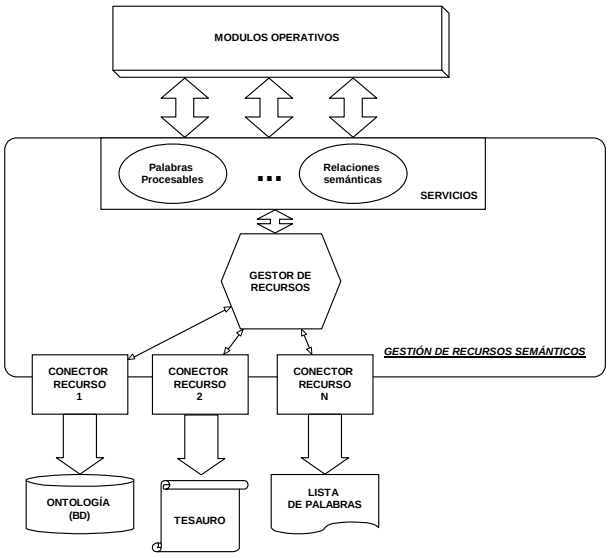


Figura 5: Arquitectura del Gestor de Recursos Semánticos.

6. CONCLUSIÓN Y TRABAJO FUTUROS

En este trabajo se ha presentado la plataforma BUDI, la cual está orientada a la aplicación de técnicas de soft-computing en el proceso de búsqueda en repositorios de información textual, tanto en Internet, como en repositorios departamentales (bases de datos XML, etc.).

El siguiente y más importante paso será estudiar la forma de someter la herramienta a un proceso de evaluación estandarizado. Nuestra intención es poder cuantificar con una colección estándar de documentos y consultas los resultados, como por ejemplo TREC [6]. TREC aporta una colección de documentos y consultas estandarizados sobre las que se realizan una serie de pruebas del rendimiento de los SRI. Estos resultados nos servirán como indicadores para conocer la eficacia de los algoritmos implementados, aportando información valiosa para poder mejorar el sistema, así como para comparar nuestro sistema con otros SRI.

Agradecimientos

La investigación realizada está financiada por el proyecto TIN2007-67494 del Ministerio de Educación y Ciencia, el proyecto PAC06-0059 de la Junta de Comunidades de Castilla-La Mancha y el proyecto FIT-340001-2007-4 financiado por el Ministerio de Industria, Turismo, y Comercio.

Referencias

[1] A. Díaz, E. Raghavan, “On Fuzzy Result Merging for Metasearch”, Proceedings of the IEEE Fuzzy Systems Conference 2007.

[2] C. Buckley, G. Salton; J.Allan, and A. Singhal. "Automatic Query Expansion Using SMART: (TREC-3)," in D. K. Harman, editor, The Third Text Retrieval Conference (TREC-3), Gaithersburg, MD, 1995. U.S. Dept. of Commerce, National Institute of Standards and Technology. 1995.

[3] D. Widyantoro, J. Yen.“Incorporating fuzzy ontology of term relations in a search engine”, Proceedings of the BISC Int. Workshop on Fuzzy Logic and the Internet, pp. 155-160, 2001.

[4] de Borda, Mémoire sur les élections au scrutin, in Histoire de l’Academie Royale de Sciences, Paris, 1781.

[5] E. Agirre, G. Rigau. “Word sense disambiguation using Conceptual Density”. In Proceedings of the 16th conference on Computational linguistics Vol. 1, pp. 16-22., Copenhagen, Denmark, 1996.

[6] E.M. Voorhees; D. Harman: “Overview of the Sixth Text REtrieval Conference (TREC-6)”. TREC 1997, Gaithersburg, Maryland, USA, pp. 1-24, 1997.

[7] F. P. Romero, J. A. Olivas; P. Garcés. “FzMail: Using FIS-CRM for e-mail classification” Journal of Advanced Computational Intelligence and Intelligent Informatics, JACIII. Fuji Technology Press Ltd, Japan, Vol.11, No.1 pp. 40-50, 2007.

[8] I. Ounis G. Amati, V. Plachouras. B. He, C. Macdonald, C. Lioma. “Terrier: A high performance and scalable information retrieval platform”. In: Proceedings of the twenty-ninth annual international ACM/SIGIR conference on research and development in information retrieval workshop on open source information retrieval (OSIR 2006), ACM Press, Seattle, Washington.

[9] J. A. Olivas, P. Garcés, F.P. Romero: “An application of the FIS-CRM model to the FISS metasearcher: Using fuzzy synonymy and fuzzy generality for representing concepts in documents”. Int. Jour. of Approx. Reasoning (Soft Computing in Recognition and Search) Vol. 34, pp. 201-219, 2003.

[9] J. A. Olivas, J. de la Mata.;J. Serrano-Guerrero. (2004): “Ontology Constructor Agent for improving Web Search with GUMSe”. In Proc. of Information Processing and Management of Uncertainty in Knowledge –Based Systems, IPMU’04. Perugia, Italia.

[10] J. A. Olivas; J. de la Mata.; J. Serrano-Guerrero.; P. Garcés; F. P. Romero. “Desarrollo de motores inteligentes de búsqueda en Internet en el marco del grupo de investigación SMILe-ORETO”. Olivas, J. A., Sobrino, A. (eds.): “Recuperación de información textual, Text Information Retrieval”. Universidade de Santiago de Compostela Santiago de Compostela, España, 2006.

[11] J. de la Mata, J. A. Olivas, J. Serrano-Guerrero. “Combinación de Mecanismos de Evaluación en el Metabuscador GUMSe”. Actas del II Simposio sobre Lógica Fuzzy y Soft Computing, 2007.

[12] J.L. García-Lapresta, M. Martínez-Panero, “A fuzzy Borda Count in multiperson decision making”, in Trzaskalik, T., Michnik, J. (eds), Multiple Objective and Goal Programming. Recent Developments, Advances in Soft Computing, Springer, 2002, pp. 46-60.

[13] K. Lagus; T. Honkela; S. Kaski; T. Kohonen. “WEBSOM for textual data mining”. Artificial Intelligence, pp. 16-22, 1999.

[14] L. Page; S. Brin; R. Motowani; T. Winograd. “The PageRank Citation Ranking: Bringing Order to the Web”, Stanford Digital Library working paper, 1997.

[15] P. Garcés; J. A. Olivas; F. P. Romero. “Concept-matching IR systems versus Word-matching IR systems: Considering fuzzy interrelations for indexing web pages”. Journal of the American Society for Information Science and Technology JASIST, Wiley, USA. Vol. 57, nº 4, pp. 564 - 576, 2006.

Modelos de Utilidad para Sistemas de Recuperación de Información Estructurada basados en Modelos Gráficos Probabilísticos

Luis M. de Campos Juan M. Fernández-Luna Juan F. Huete Carlos Martín-Dancausa

Departamento de Ciencias de la Computación e Inteligencia Artificial
E.T.S.I. Informática y de Telecomunicación, Universidad de Granada,
18071 – Granada, Spain
{lci,jmfluna,jhg,cmdanca}@decsai.ugr.es

Resumen

En este trabajo se proponen nuevos modelos de utilidad para el sistema de recuperación de información estructurado Garnata, que está basado en redes bayesianas y diagramas de influencia. La evaluación de los nuevos modelos de utilidad se ha realizado empleando la colección documental en XML de INEX.

Palabras Clave: Recuperación de información, XML, redes bayesianas, diagramas de influencia, utilidad

1 Introducción

Garnata [6] es un Sistema de recuperación de información estructurada para documentos XML, basado en modelos gráficos probabilísticos [10, 11]. Al ser estructurado, es capaz de devolver al usuario componentes de los documentos, no solo los documentos completos, para ajustarse lo mejor posible a las necesidades del usuario que realiza la consulta [1].

La forma en que el modelo de recuperación implementado en Garnata calcula la relevancia de cada componente o unidad estructural de un documento es combinando dos tipos de información diferentes. Por un lado la especificidad de la unidad respecto de la consulta: cuantos más términos de la unidad aparezcan en la consulta, más relevante resulta la primera, es decir más claramente la unidad trata solamente (de al menos una parte) del tema de la consulta y no de otros. Por otro lado, la exhaustividad de la unidad respecto de la consulta: cuantos más términos de la consulta emparejan con los de la unidad, más relevante resulta esta última, o sea más claramente la unidad abarca completamente el tema de la consulta.

Estas dos dimensiones sobre la relevancia de una unidad respecto de una consulta se calculan de forma diferente. Para calcular la especificidad, se calcula la probabilidad de relevancia de cada unidad mediante un proceso de inferencia en la red bayesiana que representa la colección de documentos estructurada (véase la figura 1). La exhaustividad se obtiene definiendo la utilidad de cada unidad en función de la proporción de términos de la consulta que aparecen en dicha unidad. Entonces transformando la red bayesiana en un diagrama de influencia (véase la figura 2), se calcula la utilidad esperada resolviendo el diagrama de influencia, cuyo resultado consiste en multiplicar la probabilidad de relevancia por la utilidad de la unidad.

Nuestro propósito en este trabajo es modificar el funcionamiento del sistema definiendo la utilidad de una forma diferente, de tal modo que se penalicen de manera más acusada aquellas unidades que no contengan la mayor parte de los términos de la consulta. Definiendo un modelo paramétrico, es posible ajustar el grado de utilidad para que el modelo tenga un comportamiento más parecido a un AND estricto (si no todos o casi todos los términos de la consulta están en la unidad considerada, ésta no será apenas relevante) o a un AND menos estricto.

La evaluación de los nuevos modelos de utilidad propuestos para el sistema Garnata se ha realizado empleando la colección documental INEX (INitiative for the Evaluation of XML Retrieval).

2 Modelos Gráficos Probabilísticos en el Sistema Garnata

El sistema Garnata está basado en un diagrama de influencia y en la red bayesiana subyacente correspondiente. En esta sección describiremos esos dos modelos y cómo usarlos para recuperar componentes de documentos de una colección documental mediante inferencia probabilística (ver las referencias [3, 4] para más detalles).

2.1 La Red Bayesiana

Consideramos tres clases diferentes de entidades asociadas a una colección de documentos estructurados, que se representan por medio de tres clases de variables aleatorias: *términos de indexación*, *unidades estructurales básicas*, y *unidades estructurales complejas*, que a su vez se representan en la red bayesiana mediantes los correspondientes *nodos*. Los nodos término forman el conjunto $\mathcal{T} = \{T_1, T_2, \ldots, T_l\}$; $\mathcal{U}_b = \{B_1, B_2, \ldots, B_m\}$ es el conjunto de unidades estructurales básicas, aquellos componentes de los documentos que únicamente contienen términos, mientras que $\mathcal{U}_c = \{S_1, S_2, \ldots, S_n\}$ es el conjunto de las unidades estructurales complejas, que están compuestas de otras unidades, básicas o complejas. Aquellas unidades que contengan tanto términos como otras unidades se consideran como unidades complejas, donde el texto asociado se asigna a una nueva unidad básica denominada *virtual*, véase el ejemplo de la Figura 1[1]. El conjunto de todas las unidades estructurales es por tanto $\mathcal{U} = \mathcal{U}_b \cup \mathcal{U}_c$.

Las variables aleatorias binarias asociadas con cada nodo T, B o S toman sus valores en los conjuntos $\{t^-, t^+\}$, $\{b^-, b^+\}$ or $\{s^-, s^+\}$ (el término/unidad es no relevante o es relevante), respectivamente. Una unidad es relevante para una consulta si satisface la necesidad de información del usuario expresada mediante esa consulta. Un término es relevante en tanto en cuanto el usuario cree que aparecerá en unidades/documentos relevantes.

Con respecto a los arcos, existirá un arco desde cada nodo (ya sea término o unidad estructural) hasta la unidad estructural en la que está contenido. La estructura jerárquica del modelo determina que cada unidad estructural $U \in \mathcal{U}$ tiene *sólo una* unidad estructural como hija: la única unidad estructural que contiene a U (excepto para los nodos hoja, o sea los documentos completos, que no tienen ningún nodo hijo). Denotaremos $U_{hi(U)}$ al único nodo hijo asociado con el nodo U (con $U_{hi(U)} = \text{null}$ si U es un nodo hoja).

Los valores numéricos de las probabilidades condicionales necesarias para especificar completamente la red bayesiana, $p(b^+|pa(B))$ y $p(s^+|pa(S))$, para cada nodo en $\mathcal{U}_b$ y $\mathcal{U}_c$, respectivamente, y para cada configuración $pa(X)$ de los correspondientes conjuntos de padres $Pa(X)$, se especifican mediante el modelo canónico propuesto en [2], que incorpora un procedimiento de inferencia muy eficiente. Tales probabilidades se definen como sigue:

$$\forall B \in \mathcal{U}_b, \quad p(b^+|pa(B)) = \sum_{T \in R(pa(B))} w(T, B)\,, \quad (1)$$

$$\forall S \in \mathcal{U}_c, \quad p(s^+|pa(S)) = \sum_{U \in R(pa(S))} w(U, S)\,, \quad (2)$$

donde $w(T, B)$ es un peso asociado a cada término T perteneciente a la unidad básica B y $w(U, S)$ es un peso que mide la importancia de la unidad U dentro de S. En todo caso $R(pa(U))$ es el subconjunto de padres de U (términos para B, y unidades básicas o complejas para S) instanciados a relevantes en la configuración $pa(U)$, es decir, $R(pa(B)) = \{T \in Pa(B) \,|\, t^+ \in pa(B)\}$ y $R(pa(S)) = \{U \in Pa(S) \,|\, u^+ \in pa(S)\}$. Estos pesos pueden definirse de cualquier modo, siempre que $w(T, B) \geq 0$, $w(U, S) \geq 0$, $\sum_{T \in Pa(B)} w(T, B) \leq 1$ y $\sum_{U \in Pa(S)} w(U, S) \leq 1$.

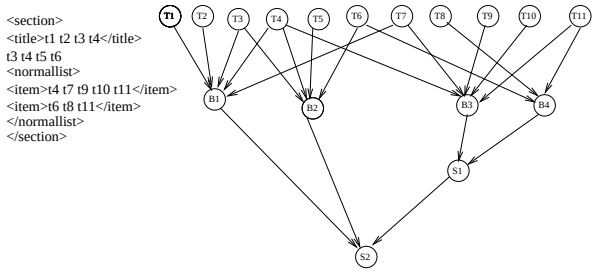


Figura 1: Documento XML y la correspondiente red bayesiana. Ti representan términos índice; la unidad básica B1 corresponde a la marca `<title>`, y B3 y B4 a la marca `<item>`; las unidades complejas S1 y S2 corresponden a las marcas `<normallist>` y `<section>` respectivamente; B2 es una unidad virtual usada para almacenar el texto que hay dentro de S2 que no está incluido en ninguna otra unidad contenida en S2.

2.2 El Diagrama de Influencia

La red bayesiana se extiende incluyendo nodos de decisión, representando las posibles alternativas disponibles para el decisor, y nodos de utilidad, transformándola de este modo en un diagrama de influencia. Para cada unidad estructural $U_i \in \mathcal{U}$, R_i representa la variable de decisión relacionada con mostrar o no U_i al usuario (con valores r_i^+ y r_i^-, que significan 'recuperar U_i' y 'no recuperar U_i', respectivamente), y el nodo utilidad V_i mide el valor de utilidad de la correspondiente decisión. Consideraremos también un *nodo de utilidad global* Σ que representa la utilidad conjunta del modelo completo (asumimos un comportamiento aditivo del modelo).

Además de los arcos entre nodos presentes en la red bayesiana, se incluye un conjunto de arcos dirigidos

[1] Obviamente este tipo de unidad virtual es no recuperable, es sólo un formalismo para distinguir claramente entre unidades que contienen sólo texto y unidades contenedoras (que contienen sólo otras unidades).

hacia los nodos de utilidad, que se emplean para indicar qué variables tienen una influencia directa en la utilidad de cada decisión. Para representar que la función de utilidad de V_i obviamente depende de la decisión tomada y del valor de relevancia de la unidad estructural considerada, se usan arcos desde cada nodo unidad estructural U_i y cada nodo de decisión R_i hasta el nodo utilidad V_i. Además, incluimos también arcos desde $U_{hi(U_i)}$ hasta V_i, que representan que la utilidad de la decisión relativa a recuperar la unidad U_i también depende de la relevancia de la unidad que la contiene (aunque para aquellas unidades U donde $U_{hi(U)} = \text{null}$, este arco no existe). Las funciones de utilidad asociadas a cada nodo de utilidad V_i son por tanto $v(r_i, u_i, u_{hi(U_i)})$, con $r_i \in \{r_i^-, r_i^+\}$, $u_i \in \{u_i^-, u_i^+\}$, y $u_{hi(U_i)} \in \{u^-_{hi(U_i)}, u^+_{hi(U_i)}\}$.

Finalmente, el nodo de utilidad global Σ tiene a todos los nodos de utilidad V_i como padres. Estos arcos representan el hecho de que la utilidad conjunta del modelo dependerá de los valores individuales de utilidad de cada unidad estructural. La Figura 2 representa el diagrama de influencia correspondiente al ejemplo anterior.

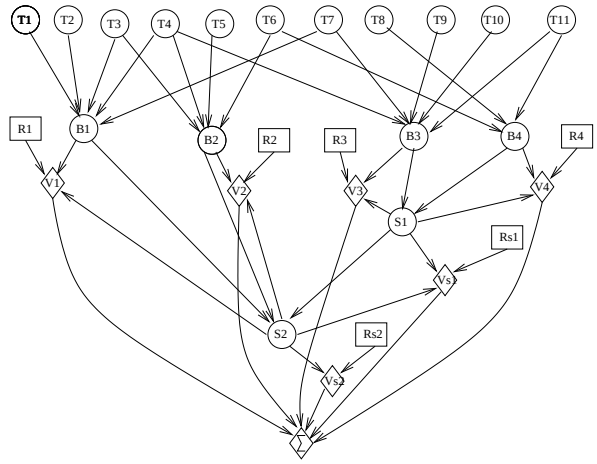


Figura 2: Diagrama de influencia para el ejemplo de la Figura 1.

2.3 Inferencia y Toma de Decisiones

Nuestro objetivo es, dada una consulta, calcular la utilidad esperada de recuperar cada unidad estructural, y entonces proporcionar una ordenación de esas unidades en orden decreciente de utilidad esperada. Sea $\mathcal{Q} \subseteq \mathcal{T}$ el conjunto de términos empleados para expresar la consulta. Cada término $T_i \in \mathcal{Q}$ se instanciará a su valor t_i^+; sea q la correspondiente configuración de las variables de $\mathcal{Q}$. Queremos calcular la utilidad esperada de cada decisión dado q. Como hemos supuesto una utilidad global aditiva, y las diferentes variables de decisión no están directamente enlazadas unas con otras, podemos procesar cada una de forma independiente. La utilidad esperada de recuperar cada unidad U_i puede calcularse mediante:

$$EU(r_i^+ \mid q) = \sum_{\substack{u_i \in \{u_i^-, u_i^+\} \\ u_{hi(U_i)} \in \left\{u^-_{hi(U_i)}, u^+_{hi(U_i)}\right\}}} v(r_i^+, u_i, u_{hi(U_i)})\, p(u_i, u_{hi(U_i)}|q) \tag{3}$$

Aunque las probabilidades a posteriori bidimensionales $p(u_i, u_{hi(U_i)}|q)$ en la ecuación (3) podrían calcularse de forma exacta [4], ésto es mucho más costoso que calcular las probabilidades a posteriori unidimensionales $p(u_i|q)$, las cuales pueden obtenerse de forma muy eficiente debido a las características del modelo canónico empleado para definir las probabilidades condicionales y a la topología de la red. Por tanto, aproximamos las probabilidades bidimensionales mediante $p(u_i, u_{hi(U_i)}|q) = p(u_i|q) \times p(u_{hi(U_i)}|q)$. El cálculo de las probabilidades unidimensionales se basa en las siguientes fórmulas exactas, que dan lugar a un algoritmo de cálculo de las mismas muy eficiente (de complejidad lineal en el tamaño del subgrafo inducido por los nodos término de la consulta y sus descendientes) [3, 4]:

$$\forall B \in \mathcal{U}_b,\ \ p(b^+|q) = \sum_{T \in Pa(B) \setminus Q} w(T,B)\, p(t^+) + \sum_{T \in Pa(B) \cap R(q)} w(T,B)$$

$$\forall S \in \mathcal{U}_c,\ \ p(s^+|q) = \sum_{U \in Pa(S)} w(U,S)\, p(u^+|q)\,.$$

Estas probabilidades representan el componente de especificidad de cada unidad estructural: cuantos más términos de los que indexen una unidad U_i también pertenezcan a Q, más probable es U_i.

2.4 Valores de Utilidad

Puesto que en este trabajo estamos centrados en la componente de utilidad, no especificaremos cómo se definen los parámetros del modelo correspondientes a los pesos de los términos $w(T,B)$ y las unidades $w(U,S)$ en las expresiones de las probabilidades condicionales, y nos centraremos en las utilidades necesarias para calcular la utilidad esperada de recuperar una unidad estructural, $EU(r_i^+ \mid q)$, o sea $v(r_i^+, u_i, u_{hi(U_i)})$. En Garnata estas utilidades se forman con un componente que depende de la unidad implicada y otro componente independiente de la unidad y dependiente únicamente de cuál de las cuatro configuraciones, $(u_i^-, u^-_{hi(U_i)})$, $(u_i^-, u^+_{hi(U_i)})$, $(u_i^+, u^-_{hi(U_i)})$ o

$(u_i^+, u_{hi(U_i)}^+)$, se está considerando:

$$v(r_i^+, u_i, u_{hi(U_i)}) = nidf_Q(U_i) \times v(u_i, u_{hi(U_i)}) \quad (4)$$

con $v(u_i^-, u_{hi(U_i)}^-) = v^{--}$, $v(u_i^-, u_{hi(U_i)}^+) = v^{-+}$, $v(u_i^+, u_{hi(U_i)}^-) = v^{+-}$ y $v(u_i^+, u_{hi(U_i)}^+) = v^{++}$.

La parte que depende de la unidad implicada se define como la suma de las frecuencias inversas (idfs) de aquellos términos contenidos en U_i que también pertenecen a la consulta Q, normalizadas por la suma de las frecuencias inversas de los términos contenidos en la consulta: una unidad U_i será más útil (más exhaustiva), con respecto a la consulta Q, cuantos más términos de la consulta también pertenezcan a U_i:

$$nidf_Q(U_i) = \frac{\sum_{T \in An(U_i) \cap Q} idf(T)}{\sum_{T \in Q} idf(T)} \quad (5)$$

$An(U_i)$ en la ecuación anterior representa el conjunto de términos que contiene (directa o indirectamente) la unidad estructural U_i.

3 Nuevos Modelos de Utilidad

Como puede apreciarse en la ecuación 5, la utilidad o exhaustividad de una unidad estructural U_i respecto de una consulta Q aumenta de manera lineal conforme más términos de la consulta aparecen en la unidad (hasta alcanzar su máximo de 1 si todos los términos de la consulta aparecen en la unidad). En nuestra experiencia con el sistema en diferentes aplicaciones [5, 7] hemos observado que este crecimiento lineal, cuando se combina con las probabilidades calculadas a partir de la red bayesiana (que miden la especificidad), puede ocasionar que unidades estructurales pequeñas, que sólo emparejan con una fracción de los términos de la consulta, resulten preferibles a otras unidades mayores, pero que contienen más términos de la consulta. En muchos casos este comportamiento no es el deseado, pues probablemente un usuario que utiliza diversos términos de consulta espera encontrar la mayor parte de esos términos en las unidades relevantes obtenidas como respuesta a su consulta. Por este motivo creemos que resulta interesante plantearse otros modelos de utilidad que primen de manera más clara (de forma no lineal) la aparición de más términos de la consulta.

Concretamente, en este trabajo proponemos un modelo de utilidad no lineal paramétrico, de forma que a mayor valor del parámetro, mayor número de términos de la consulta debe contener una unidad estructural para tener un grado de utilidad elevado. Una forma de conseguir este comportamiento es utilizando la siguiente transformación:

$$nidf_{Q,n}(U_i) = nidf_Q(U_i) \frac{e^{(nidf_Q(U_i))^n} - 1}{e - 1} \quad (6)$$

De este modo, cuando $n = 0$ tenemos que $nidf_{Q,0}(U_i) = nidf_Q(U_i)$, es decir reproducimos el modelo anterior, y cuanto mayor sea el valor del parámetro entero n, obtenemos un comportamiento más semejante al de un operador AND estricto. En la figura 3 podemos observar diferentes gráficas de la función $x\frac{e^{x^n}-1}{e-1}$ para diferentes valores de n.

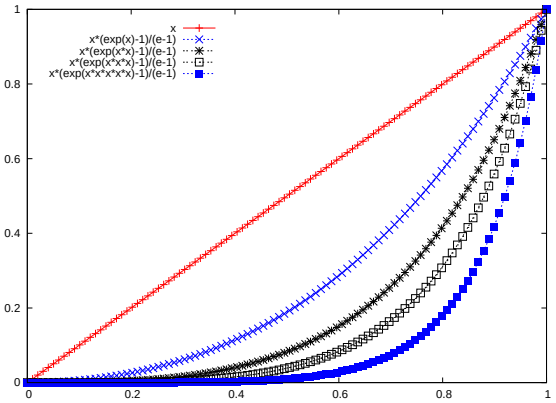


Figura 3: Función $x\frac{e^{x^n}-1}{e-1}$, para $n = 0, 1, 2, 3, 5$.

4 Experimentación

La evaluación de los nuevos modelos de utilidad propuestos se realizará con la colección documental utilizada en las dos últimas ediciones del "ad hoc track" de INEX (INitiative for the Evaluation of XML Retrieval), que es una versión en XML de la versión inglesa de Wikipedia a principios de 2006 [8], que contiene 659338 artículos. En la edición de 2007 se generaron 130 consultas, de las cuales se usaron las 99 que tenían asociados juicios de relevancia [9].

Las medidas de rendimiento usadas en cada una de las tres tareas del "ad hoc track" de INEX fueron las siguientes [9]:

- Precisión interpolada (iP) a diferentes valores de exhaustividad (0.00, 0.01, 0.05 y 0.10) así como la precisión media (MAiP) sobre los 101 puntos estándares de exhaustividad, para la tarea "focused"[2].
- Precisión media generalizada (MAgP) y precisión generalizada (gP) en diferentes rangos (5, 10, 25 y

[2]Las unidades estructurales que se muestran al usuario no pueden solaparse.

50), para la tareas "relevant in context"[3] y "best in context"[4].

En las figuras 4, 5 y 6 podemos observar la evolución de estas medidas de rendimiento para los distintos valores del parámetro n.

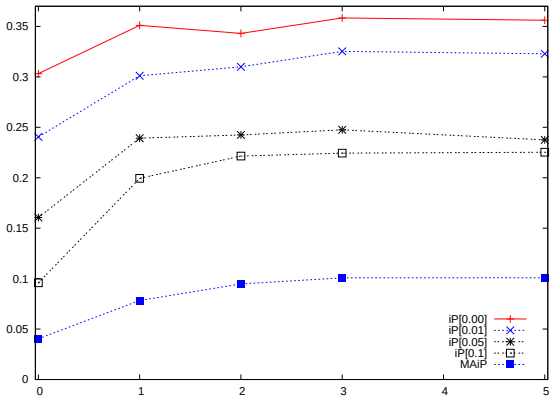


Figura 4: Resultados de las diferentes medidas de rendimiento en la tarea "focused", para $n = 0, 1, 2, 3, 5$.

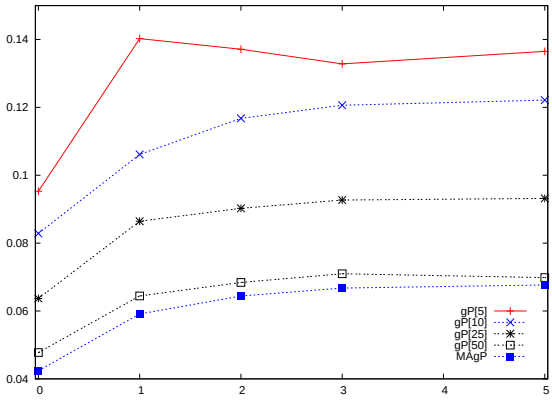


Figura 5: Resultados de las diferentes medidas de rendimiento en la tarea "relevant in context", para $n = 0, 1, 2, 3, 5$.

En las tablas 1, 2 y 3 podemos también observar los resultados en términos de las posiciones relativas que se obtendrían comparando nuestros resultados para las diferentes medidas de rendimiento con los resultados obtenidos en INEX-2007.

[3]Además de evitar solapamientos, las unidades estructurales se deben de mostrar agrupadas por documentos.

[4]Solo se muestra una única unidad estructural por cada documento, aquella unidad que constituye el mejor punto de inicio para leer el documento.

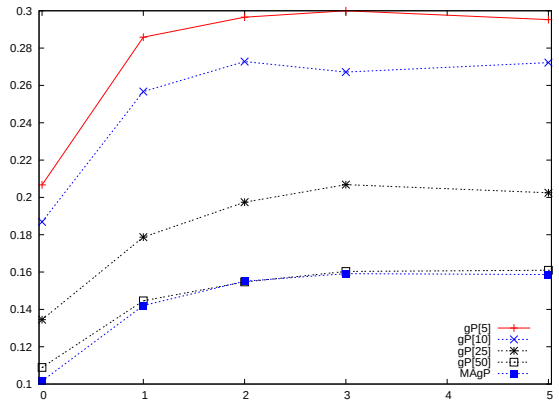


Figura 6: Resultados de las diferentes medidas de rendimiento en la tarea "best in context", para $n = 0, 1, 2, 3, 5$.

n	0	1	2	3	5
iP[0.0]	61/79	51/79	53/79	47/79	47/79
iP[0.01]	67/79	53/79	51/79	41/79	42/79
iP[0.05]	68/79	52/79	48/79	45/79	53/79
iP[0.1]	69/79	50/79	40/79	39/79	39/79
MAiP	68/79	49/79	37/79	33/79	33/79

Tabla 1: Posiciones relativas en el ranking de INEX para las diferentes medidas de rendimiento en la tarea "focused", para $n = 0, 1, 2, 3, 5$.

Tanto en las figuras como en las tablas puede apreciarse claramente un salto cuantitativo importante cuando se emplea el modelo paramétrico con $n > 0$, obteniéndose los mejores resultados con valor de n entre 3 y 5.

5 Conclusiones

En este trabajo hemos estudiado nuevos modelos de utilidad para el sistema de recuperación de información estructurada Garnata.

La evaluación de los modelos propuestos sobre la colección documental INEX pone de manifiesto la im-

n	0	1	2	3	5
gP[5]	51/66	32/66	34/66	34/66	34/66
gP[10]	51/66	39/66	37/66	32/66	29/66
gP[25]	48/66	33/66	31/66	29/66	28/66
gP[50]	49/66	36/66	32/66	29/66	31/66
MAgP	44/66	40/66	37/66	31/66	31/66

Tabla 2: Posiciones relativas en el ranking de INEX para las diferentes medidas de rendimiento en la tarea "relevant in context", para $n = 0, 1, 2, 3, 5$.

n	0	1	2	3	5
gP[5]	43/71	14/71	12/71	12/71	12/71
gP[10]	44/71	15/71	10/71	13/71	10/71
gP[25]	48/71	34/71	23/71	13/71	17/71
gP[50]	47/71	31/71	24/71	18/71	17/71
MAgP	45/71	29/71	24/71	21/71	21/71

Tabla 3: Posiciones relativas en el ranking de INEX para las diferentes medidas de rendimiento en la tarea "best in context", para $n = 0, 1, 2, 3, 5$.

portancia que una adecuada elección de tales modelos tiene desde el punto de vista de la efectividad de la recuperación. El nuevo modelo utilizado en este trabajo obtiene unos resultados experimentales considerablemente mejores que los conseguidos sin él.

Agradecimientos. Este trabajo ha sido financiado conjuntamente por la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía, el Ministerio de Educación y Ciencia y el programa Consolider Ingenio 2010, mediante los proyectos TIC-276, TIN2005-02516 y CSD2007-00018, respectivamente.

Referencias

[1] Y. Chiaramella. Information retrieval and structured documents. *Lect. Notes Comput. Sc.*, **1980**:291–314, 2001.

[2] L.M. de Campos, J.M Fernández-Luna and J.F. Huete. The BNR model: foundations and performance of a Bayesian network-based retrieval model. *Int. J. Appr. Reason.*, **34**: 265–285, 2003.

[3] L.M. de Campos, J.M. Fernández-Luna and J.F. Huete. Using context information in structured document retrieval: An approach using Influence diagrams. *Inform. Process. Manag.*, **40**(5): 829–847, 2004.

[4] L.M. de Campos, J.M. Fernández-Luna and J.F. Huete. Improving the context-based influence diagram for structured retrieval. *Lect. Notes Comput. Sc.*, **3408**: 215–229, 2005.

[5] L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, C. Martín, A.E. Romero. An information retrieval system for parliamentary documents. *Bayesian Networks: A Practical Guide to Applications*, O. Pourret, P. Naim, B. Marcot (Eds.), 203–223, Wiley, 2008.

[6] L.M. de Campos, J.M. Fernández-Luna, J.F. Huete and A.E. Romero. Garnata: An information retrieval system for structured documents based on probabilistic graphical models. Proceedings of the Eleventh International Conference of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), 1024–1031, 2006.

[7] L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, A.E. Romero. Influence Diagrams and Structured Retrieval: Garnata implementing the SID and CID models at INEX'06. *Lect. Notes Comput. Sc.*, **4518**:165–177, 2007.

[8] L. Denoyer, P. Gallinari. The Wikipedia XML Corpus. *SIGIR Forum*, **40**:64–69, 2006.

[9] N. Fuhr, M. Lalmas, A. Trotman (Eds.). Pre-Proceedings of INEX'2007. Disponible en http://inex.is.informatik.uni-duisburg.de/2007/inex07/pdf/2007-preproceedings.pdf

[10] F.V. Jensen. *Bayesian Networks and Decision Graphs*, Springer Verlag, 2001.

[11] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan and Kaufmann, San Mateo, 1988.

USANDO INFORMACIÓN DE SEGUNDA MANO EN UN SISTEMA DE RECOMENDACIÓN COLABORATIVO

L.M. de CAMPOS[1] **J.M. FERNÁNDEZ-LUNA**[1] **J.F. HUETE**[1] **M.A. RUEDA-MORALES**[1]

[1] Dpto. de Ciencias de la Computación e I.A., E.T.S.I. Informática, Universidad de Granada, 18071 – España
{lci,jmfluna,jhg,mrueda}@decsai.ugr.es

Resumen

La construcción de Sistemas de Recomendación Colaborativos ha recibido una considerable atención en los últimos años. El problema de estos sistemas se presenta en aquellos productos para los que tenemos poca información, ya que las predicciones suelen ser erroneas. En este trabajo presentamos una idea novedosa con la que vamos a obtener información de calidad a partir de las Bases de Datos, y que nos va a permitir mejorar las predicciónes del sistema. Para demostrar su valía, vamos a usar un sistema probabilístico que presentamos en este trabajo, y otro no probabilístico ajeno a nosotros.

Palabras Clave: Sistemas de Recomendación Colaborativos, Redes Bayesianas, vecindario, información de segunda mano.

1 INTRODUCCIÓN

Tomar decisiones sobre situaciones como ir a cenar a un restaurante, ver una película o planear unas vacaciones pueden ser tareas bastante complicadas por una razón principal: la gran cantidad de opciones de restaurantes, películas y destinos vacacionales que existen. Los Sistemas de Recomendación (SR) surgen para intentar paliar las dificultades de tratar con tantas opciones. En términos generales, los SR producen sugerencias (recomendaciones) sobre productos (o acciones) dentro de un determinado dominio en el cuál está interesado el usuario. En concreto, en este trabajo nos centraremos en películas como productos susceptibles de ser recomendados.

Hay muchos tipos de SR [11, 4] dependiendo de la información que se use para realizar la recomendación:

- Basados en contenido: Se recomiendan productos similares a aquellos que el usuario activo ha votado positivamente en el pasado.
- Colaborativos: Se recomiendan productos que otros usuarios con gustos similares al usuario activo consideran buenos.
- Híbridos: La recomendación se realiza combinando las aproximaciones colaborativa y basada en contenido.

En este artículo nos centraremos en la variante colaborativa. Uno de los problemas de este tipo de sistemas surge cuando el usuario del sistema quiere información sobre un producto determinado y las personas con gustos similares a él no se la pueden ofrecer (no han visto la película, no han cenado en el restaurante...). En este caso el sistema ofrecerá una valoración del producto para el usuario que, seguramente, no va a ser de calidad. Una posible solución a este problema es utilizar información, que podríamos llamar de segunda mano, a la hora de recomendar. Por ejemplo, supongamos la siguiente situación: Yo pregunto a mis amigos por una determinada película pero ninguno, o pocos, de éstos la han visto. Sin embargo, ocurre que éstos tienen amigos, con gustos parecidos a ellos, que sí que la han visto. Parece obvio pensar que, si esos amigos le dicen (ya que tienen sus mismos gustos) si le va a gustar o no, él podría decirme a mi (ya que tenemos los gustos parecidos) si me va a gustar o no. Este es el objetivo de nuestro trabajo: obtener nueva información, y de calidad, para mejorar las recomendaciones del sistema.

El resto del artículo se divide de la siguiente forma: Después de una breve introducción a los SR Colaborativos, en la Sección 2, describiremos los modelos con los que vamos a experimentar en este trabajo. En la sección 3 expondremos cómo vamos a obtener la información de segunda mano. La sección 4 presenta los resultados obtenidos mediante la ejecución de distin-

tos experimentos y, para finalizar, en la sección 5 se ofrecen algunas conclusiones y posible trabajo futuro.

2 SISTEMAS DE RECOMENDACIÓN COLABORATIVOS

Podemos encontrar distintos tipos de SR dependiendo de los recursos utilizados para determinar el vecindario, así como para determinar el voto propuesto por el sistema. En [8] se hace un análisis de distintas variantes de algoritmos colaborativos basados en vecindario tales como similitud vector-coseno, correlación de Pearson y métodos estadísticos Bayesianos. En [12, 10], se usan algoritmos de aprendizaje sobre Redes Bayesianas para aprender las distribuciones de probabilidad.

En este trabajos nos centraremos en dos de ellos: El primero basado en vecindario y el segundo, un modelo probabilístico basado en Redes Bayesianas.

2.1 Promedio de desviaciones de la media del vecindario[2]

Partiendo del modelo propuesto por Grouplens [3], Promedio de desviaciones de la media del vecindario (Weighted Average of Deviations from the neighbor's mean) es un SR Colaborativo que usa un algoritmo basado en el vecindario:

$$voto_{a,i} = \overline{r}_a + \frac{\sum_{u=1}^{n}(r_{u,i} - \overline{r}_u) \cdot sim_{a,u}}{\sum_{u=1}^{n} sim_{a,u}}$$

donde $voto_{a,i}$ es la predicción de voto para el usuario activo a del item i, n es el número de usuarios similares (tamaño del vecindario) y $sim_{a,u}$ es la medida de similitud entre el usuario activo y el usuario similar u.

Para medir la similitud entre usuarios, usado también como la base de pesos en diferentes sistemas colaborativos, usa el Coeficiente de Correlación de Pearson (CCP): Sea U el conjunto de usuarios y U_a y U_b dos usuarios de dicho conjunto, el CCP para ellos se calcula de la forma:

$$CCP(U_a, U_b) = \frac{\sum_j (r_{a,j} - \overline{r}_a)(r_{b,j} - \overline{r}_b)}{\sqrt{\sum_j (r_{a,j} - \overline{r}_a)^2 \sum_j (r_{b,j} - \overline{r}_b)^2}} \quad (1)$$

donde las sumas sobre j se hacen sobre aquellos items para los cuales los usuarios U_a y U_b tienen votos, $I(U_a) \cap I(U_b)$ (donde $I(U)$ es el conjunto de items votados por el usuario U en el conjunto de datos). Si no hay items comunes en el historial de voto de U_a y U_b, entonces $CPP(U_a, U_b) = 0$ por defecto. Además, $\overline{r}_a$ es el voto medio para el usuario U_a, esto es,

$$\overline{r}_a = \frac{1}{|I(U_a)|} \sum_{I_k \in I(U_a)} r_{a,k}.$$

El valor final de similitud se obtendrá aplicando un factor de corrección que devalue aquellos valores de CCP que se hayan obtenido con menos de 50 películas en común, esto es:

$$sim(U_a, U_b) = CCP(U_a, U_b) \cdot FC$$

con

$$FC = \begin{cases} 1 & \text{si } n > 50 \\ \frac{n}{50} & \text{en otro caso} \end{cases} \quad (2)$$

siendo n el número de películas en común.

Para finalizar la exposición del modelo, decir que, para cada predicción de voto sobre un item, se utilizan solo los usuarios similares que han visto dicho item, esto es, para cada usuario, su vecindario varía dependiendo del item que se evalue.

2.2 SR Colaborativo mediante Redes Bayesianas

En esta sección se presenta cómo se pueden modelar las relaciones entre usuarios (y su uso a la hora de predecir el voto) mediante Redes Bayesianas (RB): Sea U un conjunto de nodos en la RB representando al conjunto de usuarios, cada variable usuario U_a representa la distribución de probabilidad asociada a su patrón de voto. Si asumimos que los votos son números enteros en una escala del 1 hasta $\#r$, los estados de cada nodo corresponden con el conjunto $\{1, 2, \ldots, \#r\}$.

Para facilitar la presencia de relaciones de dependencia entre individuos en el modelo (evitando una topología de red compleja), proponemos el uso de un nuevo conjunto de nodos V que identifiquen los votos colaborativos. Hay un nodo colaborativo por cada usuario del sistema, esto es, $V = V_1, V_2, \ldots, V_n$. Estos nodos se usarán para estimar las distribuciones de probabilidad de los votos de usuario y, por lo tanto, tomarán sus valores del mismo dominio de voto que U.

2.2.1 Aprendizaje

El conjunto de padres de una variable V_a en el grafo, $Pa(V_a)$, se aprenderá de la base de datos de votos y contendrá aquellas variables usuario, $U_b \in \mathcal{U}$, donde U_a y U_b deben tener una gran similitud entre sus gustos. Así, dado un valor de similitud, el conjunto

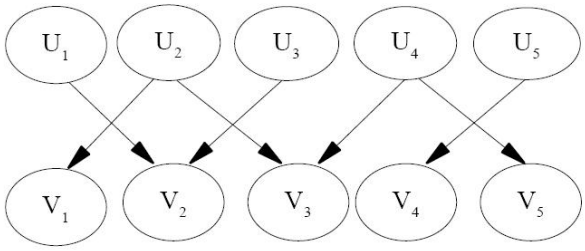


Figura 1: Topología del Sistema de Recomendación

$Pa(V_a)$ puede ser obtenido usando un umbral o considerando solo las primeras p variables en el ranking. Para medir la similitud entre usuarios, proponemos una combinación de dos criterios diferentes pero complementarios: correlación de voto entre items comunes (ver ecuación 1) y grado de solapamiento, esto es:

$$sim(U_a, U_b) = abs(CCP(U_a, U_b)) \cdot D(U_a, U_b)$$

Notar que con el uso del valor absoluto para CCP estamos diciendo que tanto los usuarios similares como los usuarios con gustos opuestos (complementarios) pueden ayudar el la predicción del voto final para un usuario activo. Esta es una de las diferencias con respecto al primer modelo, ya que allí se obtienen los usuarios similares a partir de aquellos que tienen una correlación positiva ($CCP > 0$).

El segundo criterio trata de penalizar aquellos vecinos altamente correlados que están basados en muy pocos items en común. Por consiguiente, tenemos que tener en cuenta el número de items que tanto U_a como U_b han votado en común, esto es, su grado de solapamiento. En esta situación, consideramos que cuanto más grande sea la probabilidad de que un usuario U_a vote un item que ha sido votado por U_b, mayor será la calidad de U_b como padre de U_a[1]. Este criterio se define mediante la siguiente expresión:

$$D(U_a, U_b) = \frac{|I(U_a) \cap I(U_b)|}{|I(U_b)|}$$

2.2.2 Estimando las distribuciones de probabilidad condicionadas

Para completar la especificación del modelo, tenemos que fijar los valores numéricos para las probabilidades condicionadas. Previamente, vamos a introducir alguna notación. Así, dada una variable X_i, usamos letras minúsculas para denotar la realización de la variable, esto es, $x_{i,j}$ significa que la variable X_i toma el j^{th} valor. Escribimos $Pr(x_{i,j}|pa(X_i))$ para $P(X_i = x_{i,j}|pa(X_i))$, siendo $pa(X_i)$ una configuración del conjunto de padres de X_i, $Pa(X_i)$, o, a veces, $Pr(X)$ para denotar la distribución de probabilidad.

Debemos distinguir entre nodos de usuario, U, y nodos colaborativos, V. Los primeros, al ser nodos raíz en el grafo, almacenan distribuciones de probabilidad marginadas mientras que los segundos, almacenan un conjunto de distribuciones de probabilidad condicionadas. Se definen de la forma:

- Para cada nodo usuario U_k necesitamos fijar la distribución de probabilidad apriori sobre su patrón de voto. En este trabajo proponemos éstas sean equiprobables, esto es, que $Pr(U_{k,s} = 0.2)$ si $s \in \{1, 2, \cdots, 5\}$.
- Para un nodo dado V_i debemos definir para cada configuración $pa(V_i)$ la distribución de probabilidad condicionada $Pr(V_{i,j}|pa(V_i))$. En estos casos, puede ser bastante complicado fijar y almacenar las tablas de probabilidades condicionadas (con un tamaño exponencial al número de padres). Por ello, proponemos el uso de un modelo canónico similar al presentado en [1], para representar las probabilidades condicionadas, que nos va a permitir diseñar un procedimiento de inferencia muy eficiente. Así, para un nodo X_i, definimos esas probabilidades de la forma:

$$Pr(x_{i,j}|pa(X_i)) = \sum_{Y_k \in Pa(X_i)} w(y_{k,l}, x_{i,j}) \quad (3)$$

  donde $y_{k,l}$ es el valor que la variable Y_k toma en la configuración $pa(X_i)$ y $w(y_{k,l}, x_{i,j})$ son pesos que miden cómo el l^{th} valor de la variable Y_k describe el j^{th} estado del nodo X_i, con $w(y_{k,\bullet}, x_{i,j}) \geq 0$ y $\sum_{Y_k \in Pa(X_i)} w(y_{k,\bullet}, x_{i,j}) \leq 1$.

El problema ahora viene en estimar los pesos dados por aquellos usuarios U_b con gustos similares, esto es, $U_b \in Pa(V_a)$. Al no conocer esos pesos, tenemos que estimarlos de la base de datos de entrenamiento. Siguiendo las ideas de [1], consideramos que $w(u_{b,t}, v_{a,s})$, esto es, el efecto que el voto t dado por el usuario U_b tiene en la predicción del voto de V_a se calcula por medio de:

$$w(u_{b,t}, v_{a,s}) = \frac{N(u_{b,t}, v_{a,s}) \cdot sim(U_a, U_b)}{N(u_{b,t}) \cdot \sum_{U_k \in Pa(V_a)} sim(U_a, U_k)}$$

con $1 \leq t, s \leq r$, siendo $N(u_{b,t}, v_{a,s})$ el número de items del conjunto $I(U_a) \cap I(U_b)$ que han sido votados con el valor t por el usuario U_b y que también han sido votados con el valor s por el usuario U_a, $N(u_{b,t})$

[1] Notar que hemos usado la penalización propuesta en la Ecuación 2, obteniendo peores resultados que usando esta propuesta.

el número de items del conjunto $I(U_a) \cap I(U_b)$ que han sido votados con el valor t por el usuario U_b y $sim(U_a, U_b)$ el valor de similitud entre el usuario U_a y U_b.

2.2.3 Inferencia

Dado un usario activo U_a y una película I_k en el sistema (votada por otros usuarios), el problema está en predecir el voto que el usuario U_a dará a la película. Desde un punto de vista probabilístico, tenemos que calcular[2] $Pr(v_{a,s}|I_k)$, $1 \leq s \leq \#r$. En primer lugar vamos a ver cómo incluir la evidencia, I_k, en la RB. Dada la película, tendremos que cambiar nuestra certeza en las distribuciones de voto para cualquier usuario que haya votado previamente a la película (nodos U en nuestro modelo), esto es,

$$Pr(U_i = s|I_k) = \begin{cases} 1 & \text{si } U_i \textit{ votó } s, \\ 0 & \text{si } U_i \textit{ votó } k \neq s. \end{cases}$$

Cuando el usuario no ha visto la película, la distribución de probabilidad aposteriori se corresponde con la apriori.

Una vez la evidencia ha sido incluida en el sistema, el problema es predecir el voto del usuario para dicho item I_k. Esos valores de probabilidad se calculan propagando la evidencia a través de la red.

Usando las ventajas de los modelos canónicos usados para expresar las distribuciones de probabilidad condicionada (ecuación 3), estas probabilidades finales pueden calcularse eficientemente [1] de la forma:

$$Pr(v_{a,s}|ev) = \sum_{j=1}^{m_{V_s}} \sum_{k=1}^{l_{U_j}} w(u_{j,k}, v_{a,s}) \cdot Pr(u_{j,k}|ev).$$

siendo m_{V_a} el número de padres de V_a, U_j un nodo en $Pa(V_a)$ y l_{U_j} el número de estados que toma U_j.

2.2.4 Obtención del voto

Una vez calculada la distribución de probabilidad aposteriori $Pr(V_{a,s}|ev)$, debemos decidir el voto que el sistema recomienda al usuario. En este trabajo estudiamos dos alternativas para la obtención del voto:

MP : Seleccionamos el voto con mayor probabilidad aposteriori:

$$\text{voto} = \arg_s \max\{Pr(V_a = s|ev)\}$$

DP : La idea es medir sólo la porción de evidencia que cada voto posible recibe gracias a la instanciación del item. Esta porción de evidencia puede ser definida fácilmente considerando la diferencia entre los valores de probabilidad apriori (sin evidencia) y aposteriori, y devolviendo el voto con el que obtenemos la máxima diferencia:

$$\text{voto} = \arg_s \max\{Pr(V_a = s|ev) - Pr(V_a = s)\}$$

3 OBTENCIÓN DE INFORMACIÓN DE SEGUNDA MANO

Los SR Colaborativos identifican grupos de personas con gustos similares al usuario activo y le recomienda aquellos productos que les hayan gustado. El problema surge cuando el usuario del sistema quiere información sobre un producto determinado y las personas con gustos similares a él no se la pueden ofrecer. En este caso el sistema ofrecerá una valoración del producto para el usuario que, seguramente, no va a ser de calidad. Ahora bien, ¿qué pasaría si la información de sus vecinos podemos obtenerla a partir de sus propios usuarios similares? O dicho de otra forma, ¿mejoraría la predicción de voto de un usuario a una película si hacemos uso de información de segunda mano?

Para probar nuestra idea modificamos la forma en la que los modelos obtienen el voto:

- Para el modelo promedio (sección 2.1):

  $$voto_{a,i} = \overline{r}_a + \frac{\sum_{u=1}^{n} (\hat{r}_{u,i} - \overline{r}_u) \cdot sim_{a,u}}{\sum_{u=1}^{n} sim_{a,u}}$$

  donde $\hat{r}_{u,i}$ es el voto del usuario si ha visto la película o, en caso contrario, el voto predicho por el modelo promedio si no la ha visto, esto es $\hat{r}_{u,i} = voto_{u,i}$.

- Para el modelo probabilístico (sección 2.2) modificamos la distribución de probabilidad aposteriori de los posibles votos de la forma:

  $$Pr(v_{a,s}|ev) = \sum_{j=1}^{m_{V_s}} \sum_{k=1}^{l_{U_j}} w(u_{j,k}, v_{a,s}) \cdot \hat{Pr}(u_{j,k}|ev).$$

  siendo

  $$\hat{Pr}(U_i = s|I_k) = \begin{cases} 1 & \text{si } U_i \leftarrow s, \\ 0 & \text{si } U_i \leftarrow k \neq s. \end{cases}$$

  donde $U_i \leftarrow s$ representa que, o bien el usuario U_i votó a I_k con el valor s, o bien, que el sistema probabilístico recomendó el voto s para el usuario utilizando los criterios de la sección 2.2.4.

[2] Con la misma filosofía, el sistema podría ser usado para recomendar a un usuario U_a aquellas películas que tienen una gran probabilidad de que le gusten. Esto podría hacerse instanciando el conjunto de películas no vistas por el usuario U_a.

4 ANÁLISIS EMPÍRICO

Esta sección presenta los resultados experimentales obtenidos mediante la prueba del sistema. Los conjuntos de datos utilizados en nuestros experimentos se han obtenido de la Base de Datos Movielens[3], que contiene 1682 películas y 943 usuarios que aportan su voto a las películas que han visto, dando lugar a 100000 votos en una escala del 1 al 5. Para las ejecuciones hemos usado los conjuntos de datos U1.base y U1.test hasta U5.base y U5.test que dividen la colección en 80% para entrenamiento y 20% para test respectivamente.

Para medir la capacidad de recomendación usamos el error medio absoluto (MAE) [9] que mide la desviación media absoluta entre el voto real y el voto predicho. Así mismo,hemos establecido el tamaño del vecindario en 30, esto es, para cada usuario hemos calculado los 30 usuarios más cercanos a sus gustos con los que realizar las predicciones.

La experimentación se ha realizado evaluando los dos modelos presentados en este trabajo: *modelo probabilístico*[4] (P) (sección 2.2) y *promedio de desviaciones de la media del vecindario* (A) (sección 2.1).

Los resultados obtenidos lanzando los sistemas con los conjuntos de datos originales obtenidos por ambos modelos se pueden ver en la tabla 1. Las ejecuciones con estos datos podemos considerarlas como *Baseline*(B). Podemos observar cómo el modelo promedio (BA) obtiene los mejores resultados y, en el modelo probabilístico (BP), los mejores resultados los obtenemos usando el criterio DP.

Tabla 1: Resultados Baseline.

	U_1	U_2	U_3	U_4	U_5
BP+MP	0.845	0.816	0.812	0.814	0.822
BP+DP	0.817	0.792	0.797	0.794	0.812
BA	0.745	0.737	0.731	0.723	0.734

La idea de este trabajo es la de usar información de segunda mano para mejorar la predicción de los sistemas colaborativos. Sin embargo, no todos los votos que predice el sistema tienen el mismo grado de utilidad. Por ejemplo, el valor que puede tener un voto obtenido a partir de 15 vecinos no puede ser el mismo que el obtenido a partir de uno solo. En este último caso, probablemente estamos introduciendo ruido en las predicciones. Es por esto que pretendemos introducir sólo información de calidad. Una alternativa para definir la calidad, es el estudiar el acierto del sistema calculando sólo los votos en los que el número de padres con el que se obtiene sea mayor que un cierto umbral. A la vista de los resultados de la Tabla 2, hemos considerado votos de calidad aquellos que se pueden obtener usando la información proporcionada por al menos el 40% de su vecindario ya que obtienen unos buenos resultados de MAE junto con una cobertura razonable (siendo la cobertura el porcentaje de datos que evaluamos con ese umbral).

Tabla 2: Calidad para U1.

% Padres	16	26	33	40
MAE_P	0.796	0.766	0.750	0.734
MAE_A	0.699	0.685	0.676	0.670
$Cobertura_P(\%)$	83.99	70.56	60.88	51.17
$Cobertura_A(\%)$	79.07	63.52	53.42	44.22

Para poder determinar el comportamiento del sistema y poder medir la ganancia que obtenemos mediante la inclusión de los nuevos votos, hemos reducido el conjunto de votos reales que se utilizan al recomendar. Así, para calcular el voto del usuario activo sobre una película, hemos eliminado aleatoriamente el 50% de los votos de calidad de sus padres. Las ejecuciones con estos datos podemos considerarlas como *Reducida* (R). En la Tabla 3 podemos ver el número de votos eliminados para cada subconjunto de datos ($E_{P,A}$). Hemos de tener en cuenta que cada vez que quitamos un voto, éste deja de utilizarse en todas las predicciones en las que entraba a formar parte. En la Tabla 4 podemos ver cómo varían los tamaños de los conjuntos de padres con los que realizamos las predicciones una vez eliminados los votos. Los resultados que proporciona el sistema los podemos ver en la Tabla 5 y como es de esperar, los valores de acierto empeoran con respecto al baseline para todas las ejecuciones de ambos modelos.

Tabla 3: Votos eliminados e Insertados.

	U_1	U_2	U_3	U_4	U_5
$E_{P,A}$	24114	23954	24001	24218	24347
I_P	45967	42523	44072	45688	47890
I_A	33487	34695	35716	36690	36819

Tabla 4: Evolución Vecindario en U1.

Pa	≤ 5	$6-12$	$13-18$	$19-24$	≥ 25
BP	4022	6570	5130	3512	766
RP	9618	9107	1254	21	0
BA	5203	6810	4750	2762	475
RA	10796	8214	984	6	0

[3] http://www.movielens.org

[4] Notar que para el modelo probabilístico, hemos añadido MP o DP a la identificación del experimento dependiendo de cómo se obtenga el voto (ver sección 2.2.4).

Tabla 5: Resultados Reducido.

	U_1	U_2	U_3	U_4	U_5
RP+MP	0.872	0.838	0.834	0.834	0.842
RP+DP	0.866	0.863	0.844	0.835	0.856
RA	0.784	0.812	0.775	0.771	0.777

Finalmente, la Tabla 6 presenta los resultados obtenidos cuando usamos información de segunda mano (ver Sección 3) (C), en la cuál introducimos todos los votos de calidad que ambos modelos predicen. Podemos ver en número de votos de calidad que introducimos en cada conjunto de datos en la Tabla 3 (I_P y I_A). Como podemos observar, los resultados son prometedores para ambos modelos, ya que el añadir los votos de calidad propuestos por ambos sistemas hace que sus valores de MAE rocen el valor del baseline y, en algunos casos, lo superen, lo cuál certifica la valía de nuestra idea.

Tabla 6: Resultados con votos de calidad.

	U_1	U_2	U_3	U_4	U_5
CP+MP	0.846	0.818	0.812	0.815	0.821
CP+DP	0.818	0.794	0.787	0.790	0.805
CA	0.741	0.738	0.732	0.725	0.733

5 CONCLUSIONES

Hemos presentado una idea novedosa para aplicar a los SR Colaborativos de tal forma que podemos obtener información indirecta de las Bases de Datos de votos haciendo uso de las predicciones que hace el propio sistema. Para ello hemos tenido que disminuir el número de votos de dichas Bases de Datos ya que, tras distintas experimentaciones, hemos comprobado que en aquellos usuarios en los que sus usuarios similares no le dan información, no es posible obtenerla indirectamente. Por este motivo, y como trabajo futuro, nos planteamos utilizar distintas Bases de Datos como NextFlix para probar nuestra propuesta, con el fin de poder mejorar las predicciones en conjunto del sistema a partir de los datos en su totalidad. Además, haremos un estudio más exhaustivo del conjunto de padres de los usuarios, dividiéndolos en intervalos y estudiando el comportamiento del sistema en cada uno de ellos.

Agradecimientos

Trabajo respaldado por el Ministerio de Educación y Ciencia y la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía bajo los proyectos TIN2005-02516 y TIC-276, respectivamente.

Referencias

[1] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete. A collaborative recommender system based on probabilistic inference from fuzzy observations. *Fuzzy Sets and Systems*, doi: 10.1016/j.fss.2008.01.016, 2008.

[2] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. An Algorithmic Framework for Performing Collaborative Filtering. *SIGIR.*, ACM 1-58113-096-1/99/0007, 1999.

[3] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. T. Riedl. Applying collaborative filtering to Usenet news. *Communications of the ACM*, 40(3):77-87, 1997.

[4] S. Kangas. Collaborative filtering and recommendation systems. *VTT Information Technology*, Research Report TTE4-2001-35, 2002.

[5] V. Robles, P. Larrañaga, J.M. Peña, O. Marbán, J. Crespo, and M.S. Pérez. Collaborative filtering using interval estimation naive Bayes. *Lecture Notes in Artificial Intelligence*, 2663:46–53, 2003.

[6] K. Miyahara and J. Pazzani. Collaborative filtering with the simple Bayesian classifier. *Proc. of the Pacific Rim Int. Conf. on Artif. Intell.*, 679–689, 2000.

[7] D. Heckerman, C. Chickering, D.Meek, C. Rounthwaite and C.Kadie. Dependency networks for inference, collaborative filtering and data visualization. *Jour. Mach. Learning Res. 1*, pp. 49-75, 2000.

[8] J.S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. *Proc. 14th Conference on Uncertainty in Artificial Intelligence*, 43–52, 1998.

[9] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22(1):5-53, 2004.

[10] C.J. Butz. Exploiting contextual independencies in web search and user profiling. *Proc. of World Congress on Computational Intelligence*, 1051–1056, 2002.

[11] P. Resnick and H.R. Varian. Recommender Systems. *Communications of the ACM*, 40(3):56-58, 1997.

[12] S.N. Schiaffino and A. Amandi. User profiling with case-based reasoning and Bayesian network. *Proc. of the Iberoamerican Conf. of Artificial Intelligence*, 12–21, 2000.

EXTRACTING FUZZY RULES FROM RECURRENT NEURAL NETWORKS TO INFER REGULAR GRAMMARS

Carlos J. Mantas

Department of Computer Science and Artificial Intelligence,
University of Granada, 18071 Granada, Spain
cmantas@decsai.ugr.es

Abstract

Recurrent neural networks can be trained to recognize strings of a regular language *L*. Finite state automata can be extracted from these networks. In this work, a method is defined for describing with fuzzy rules the transition function of an automaton. These fuzzy rules are obtained from the weights and biases from a recurrent neural network.

Keywords: Recurrent Neural Networks, Fuzzy Rules, Finite State Automata.

1 INTRODUCTION

In [3] Chomsky classified the languages in terms of generative grammars. The type of simplest grammar are regular grammars. The languages that are described with such grammars are called regular languages, which can be recognized by Finite State Automata (FSA) [6]. These formal machines are useful for designing and verifying digital circuits, carrying out the lexical analysis of a programming language, exploring and acting on large text files,…

On the other hand, in the early 1990s, Elman [4] introduced his well-known simple recurrent network. The connection between finite state machines and neural networks was again established. He compared the internal activations of the networks to the states of a finite state machine. Since this moment, an abundance of articles has been written on Recurrent Neural Networks (RNNs) and their connection with the state machines. Many references and revision can be found in [7].

Several articles describe techniques for transforming RNNs into FSA [1, 2, 5, 11, 12, 13]. In these works, the continuous state space is approximated by a finite set of states and the dynamics of the RNN is mapped to transitions among this discrete set of states. Normally, this process is based on dividing the output of each state neuron of the RNN into q intervals or looking for clusters in the state space. Since many partition values of q are available or many clusters can be found, many FSA can be extracted. How to obtain the best finite state automaton?.

In [9] fuzzy rules are extracted from a RNN to extract the associated automaton. In this case, the way of obtaining rules is inspired by [8] where fuzzy rules are obtained from Multilayer Feedforward Neural Networks (MFNN). The improvement of this extraction method motivated the work presented in [10], where the fuzzy rules extracted from a MFNN use the system inputs directly. This fact facilitates the analysis of acquired knowledge by the network. This last method is utilized in this paper.

In [10], MFNN are transformed into Takagi-Sugeno-Kang Fuzzy Rule Based Systems (TSK FRBSs) [14]. This fact allows to understand the action of a hidden neuron by analyzing two simple fuzzy rules. With this new tool, the extraction of FSA from RNNs can be faced in a simple and different way.

The transitions of the automaton contained into a RNN can be represented with fuzzy rules. We can find every state of the automaton by analyzing these transitions from the initial state. Finally, the automaton can be minimized with standard algorithms [6].

The paper is structured as follows. The initial sections introduce TSK Fuzzy Rule Based Systems, extraction of TSK fuzzy systems from neural networks, first-order RNNs and FSA. Next, the relations between FSA and RNNs are exposed. In Section 6, the method to extract FSA from RNNs is presented and it is shown with an example in Section 7.

2 ZERO-ORDER TSK FUZZY RULE BASED SYSTEMS

Commonly, the zero-order TSK FRBSs [14] have the rule structure that is introduced in the following definition.

Definition 1.
A FRBS with the following rule structure is named as zero-order TSK FRBS:
*R_k: If x_1 is B_1 * x_2 is B_2 * … * x_n is B_n then $y_k = p_k$,*

where x_i are the input features of problem, B_i are the linguistic labels of associated fuzzy sets, $p_k \in \mathbb{R}$ and y is the output variable.

This zero-order TSK FRBS computes its output according to the following definition.

Definition 2.
The output y of a FRBS with m zero-order TSK rules is computed as the weighted average of individual rule outputs y_i (i= 1,..,m) as follows:

$$y = \frac{\sum_{i=1}^{m} y_i \cdot g_i}{\sum_{i=1}^{m} g_i},$$

where $g_i = T(B_1(x_1), ..., B_n(x_n))$ is the matching degree between the IF-part of rule and the current system inputs, T is a t-norm (usually, the product operator) and $\mathbf{x} = (x_1, x_2, ... ,x_n)$ is the input of system.

In the following Sections, the fuzzy rules will be implemented using the product operator as t-norm *T*.

3 EXTRACTION OF TSK FUZZY RULES FROM NEURAL NETWORKS

Definition 3.
The action of hidden neuron j of a MFNN follows the given expression:

$$h_j(\mathbf{x}) = sigm\left(\sum_{i=1}^{n}(x_i \cdot w_{ij}) + b_j\right), \qquad (1)$$

where b_j is the bias of hidden neuron j, w_{ij} is the weight of connection from the input x_i to the hidden neuron j, "sigm" is the sigmoid activation function and $\mathbf{x} = (x_1, x_2, ... ,x_n)$ is the input of system.

The expression given in Def. 3 can be rewritten as:

$$h_j(\mathbf{x}) = sigm\left(\sum_{i=1}^{n}(x_i \cdot w_{ij} + b_{ij})\right), \qquad (2)$$

where $\sum_{i=1}^{n} b_{ij} = b_j$. This translation will be useful in the fuzzy rules extraction process.

The action of the h_j hidden neuron (Def. 3) is equivalent to the action of following zero-order TSK FRBS with two rules [10]:

R_1) **If** not A_j **then** $h^1_j(\mathbf{x}) = 0$
R_2) **If** A_j **then** $h^2_j(\mathbf{x}) = 1$.

where

- "not A_j" $\equiv x_1$ is $sigm((-1)\cdot(x\cdot w_{1j} + b_{1j})) \wedge \ldots \wedge x_n$ is $sigm((-1)\cdot(x\cdot w_{nj} + b_{nj}))$
  $\equiv x_1$ is $(1 - sigm(x\cdot w_{1j} + b_{1j})) \wedge \ldots \wedge x_n$ is $(1 - sigm(x\cdot w_{nj} + b_{nj}))$
  $\equiv x_1$ is not $(sigm(x\cdot w_{1j} + b_{1j})) \wedge \ldots \wedge x_n$ is not $(sigm(x\cdot w_{nj} + b_{nj}))$.

- "A_j" $\equiv x_1$ is $(sigm(x\cdot w_{1j} + b_{1j}))$ $\wedge \ldots \wedge x_n$ is $(sigm(x\cdot w_{nj} + b_{nj}))$.

- The aggregation $\wedge$ is implemented with the product operator.

As the function

$$sigm(x\cdot w_{ij} + b_{ij}) = sigm(w_{ij}\cdot(x + (b_{ij}/w_{ij})))$$

is centered in the value $(-b_{ij}/w_{ij})$ and its slope is proportional to w_{ij}, the propositions of the fuzzy rules can be linguistically interpreted:

- When $w_{ij} > 0.0$, the proposition "x_i is $sigm(w_{ij}\cdot(x+(b_{ij}/w_{ij})))$" can be linguistically interpreted as
  "x is approximately greater than $(-b_{ij}/w_{ij})$"
  $\equiv$ "x is $\approx>$ $(-b_{ij}/w_{ij})$" (Fig. 1).

- When $w_{ij} < 0.0$, the proposition "x_i is $sigm(w_{ij}\cdot(x+(b_{ij}/w_{ij})))$" can be linguistically interpreted as
  "x is not approximately greater than $(-b_{ij}/w_{ij})$"
  $\equiv$ "x is approximately smaller than $(-b_{ij}/w_{ij})$"
  $\equiv$ "x is $\approx<$ $(-b_{ij}/w_{ij})$" (Fig. 2).

The degree of uncertainty established by the term "approximately" is determined by the slope of the sigmoid function. If the factor $|w_{ij}|$ is high, then the sigmoid function has a sharp slope and the defined fuzzy set is practically crisp. If $|w_{ij}|$ is small, then the fuzzy set has a high uncertainty. This class of propositions will be avoided by using the simplification process of the fuzzy systems.

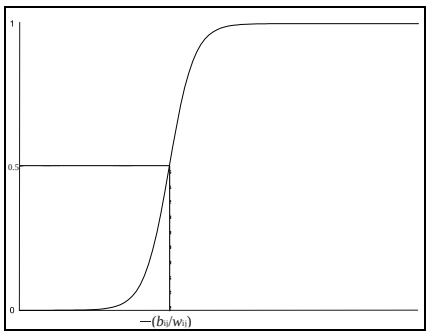


Figure 1: "x is approximately greater than $(-b_{ij}/w_{ij})$".

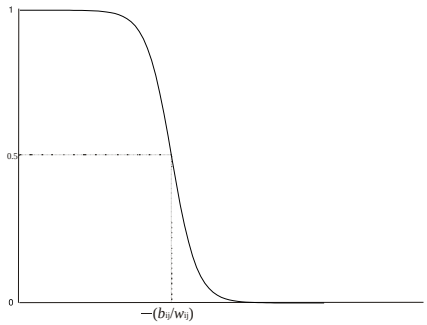


Figure 2: "x is approximately smaller than $(-b_{ij}/w_{ij})$".

3.1. SIMPLIFICATION OF THE TSK FUZZY SYSTEM

The number of fuzzy propositions can be reduced in the system obtained from a neural network. The idea is to take advantage of the distribution of the b_j bias of a neuron among the b_{ij} variables.

- The membership functions with small slope ($|w_{ij}|$ is small) produce propositions with high uncertainty. Their membership degrees can be almost 0.0 or 1.0 in the input attribute domain by adjusting the b_{ij} values. For example, if the domain of the variable x_i is $[a_1,a_2]$, the proposition "x_i is *sigm* $(x \cdot w_{ij} + b_{ij})$" ($w_{ij} \geq 0.0$) can have degrees almost 1.0 when $b_{ij} = (8 - a_1 \cdot w_{ij})$ because *sigm*(+8) ≈1.0. With this value for b_{ij}, the membership degrees of this proposition belong to [8, 8+ $(a_2-a_1) \cdot w_{ij}$]. A similar reasoning can be done with $w_{ij} \leq 0.0$ or *sigm*(−8).

- The terms of the fuzzy propositions with degrees almost 0.0 or 1.0 can be substituted by constant values equal to the mean value of their domains. In the previous example, the proposition "x_i is *sigm* $(x \cdot w_{ij} + b_{ij})$" is replaced by "$x_i$ is *sigm* $(8 + [(a_2-a_1)/2] \cdot w_{ij})$"

- The fuzzy propositions with constant values can be discarded of the rules. In order to make this, the original b_j bias of the hidden neuron is modified. In the previous example, $b_j^{new} = b_j + (8 + [(a_2-a_1)/2] \cdot w_{ij})$. The fuzzy propositions are recalculated by using this new bias.

4 FIRST-ORDER RECURRENT NEURAL NETWORKS

Let us suppose without loss of generality that:

a) The net has m state neurons (S_1, ..., S_m).

b) It has one binary input $I \in \{0,1\}$

c) The RNN's dynamics is given by

$$S_j^{t+1} = sigm\left(\sum_{i=1}^{m}(S_i^t \cdot w_{ij}) + I^t \cdot v_j + b_j\right), \quad (3)$$

$j=1,\ldots,m$

where b_j is the bias of the S_j neuron, I^t is the symbol in the position t of an binary string, , w_{ij} is the weight of connection from the S_i neuron to the S_j neuron, v_j is the weight of connection from the I input to the S_j neuron and *sigm* is the sigmoid activation function.

In order to facilitate the future rule extraction process, eq. 3 can be rewritten as:

$$S_j^{t+1} = sigm\left(\sum_{i=1}^{m}(S_i^t \cdot w_{ij} + b_{ij}) + (I^t \cdot v_j + b_{0j})\right), \quad (4)$$

$j=1,\ldots,m$, where $\sum_{i=0}^{m} b_{ij} = b_j$.

d) The network initial conditions are $S_j^1=0$ for $j=1,\ldots,m$.

e) The neuron S_1 is considered as output neuron, if $S_1^{end} \leq 0.5$ the input string is accepted, otherwise it is rejected.

5 FINITE STATE AUTOMATA AND THEIR RELATION WITH RNNs

A *finite state automaton* is a 5-tuple M= (Q,X,δ,q_0,F), where:

a) Q is a finite set of states.
b) X is the finite input alphabet.
c) δ:Q×X→Q is the transition function.
d) q_0 is the initial state.
e) F⊆Q is a set of accepting states.

We say that a string u is accepted by a finite state automaton M iff the state that is reached after u has been read by M is an accepting state. Regular languages can be recognized by FSA. That is, with each regular language L there is an associated finite state automaton M, such that a string u is accepted by M iff $u \in L$.

For a full discussion of what a regular language is and what other classes of languages there are, interested readers are referred to [6].

5.1 RELATIONS BETWEEN FIRST-ORDER RNNs AND FSA

When a RNN is trained to recognize strings of a regular language, several relations appear between the RNN and the automaton associated with the language. Mainly, the state neurons (S_1, ..., S_m) of the RNN encode the finite set of states Q of the automaton, whereas its transition function is established by the biases and weights of the RNN. With more detail, these relations are:

a) Each state $q \in Q$ of the automaton is represented in the network by m values in [0,1]. These values are the output of the state neurons, that is, $q=(S_1^t, ..., S_m^t) \in Q=[0,1]^m$.

b) The input I of the RNN belongs to the finite input alphabet of the automaton, that is, $I \in \{0,1\}=X$

c) The δ transition function of the automaton
$$\delta : Q \times X \rightarrow Q$$
is represented in the RNN as:
$$\delta((S_1^t, ..., S_m^t), I^t) = (S_1^{t+1}, ..., S_m^{t+1}). \quad (5)$$
This function δ is equivalent to m functions δ_j (j=1,…,m):
$$\delta_j: Q \times X \rightarrow (0,1)$$
$$\delta_j ((S_1^t, ..., S_m^t), I^t) = S_j^{t+1}. \quad (6)$$
Each function δ_j is implemented in the RNN by the action of the S_j state neuron. This action is determined by the weights and biases of the neuron. It is interesting to notice the similarity between the expression (6) and the equation 4.

d) The q_0 initial state of the automaton is encoded with the initial conditions of the RNN, that is
$$(S_1^1=0, ..., S_m^1=0) \equiv q_0$$

e) The set of accepting states $F \subseteq Q$ is encoded with the output neuron S_1, that is,
$(S_1, ..., S_m) \in F \Leftrightarrow S_1 \leq 0.5$

6 EXTRACTING FSA FROM RNNs

If the action of each S_j state neuron (that represents to the function δ_j) is understood by means of a TSK fuzzy system, we obtain a description of the transition function δ of the automaton. According to Section 3, this fact is possible. That is:

The action of the S_j neuron defined in eq. 4 is equivalent to the following TSK FRBS with two rules:

R$_1$) **If** not A_j **then** $S_j^{t+1} = 0$
R$_2$) **If** A_j **then** $S_j^{t+1} = 1$.

where

- "not A_j" ≡ [S_1^t is not $sigm(x \cdot w_{1j} + b_{1j}) \wedge \ldots \wedge S_m^t$ is not $sigm(x \cdot w_{mj} + b_{mj})$] $\wedge$ [I^t is not $sigm(x \cdot v_j + b_{0j})$]
- "A_j " ≡ [S_1^t is $sigm(x \cdot w_{1j} + b_{1j}) \wedge \ldots \wedge S_m^t$ is $sigm(x \cdot w_{mj} + b_{mj})$] $\wedge$ [I^t is $sigm(x \cdot v_j + b_{0j})$]

With this FRBS, we have a fuzzy explanation of the action of the S_j neuron and, according Section 5.1, we have also a comprehensible description of the partial transition δ_j.

If this process is repeated for each S_j state neuron (j=1,…,m), we obtain a description of the transition function δ of the automaton. Thus, starting from the initial state ($S_j^1=0$ for j=1,…,m) it is easy to get the total structure of the automaton. We must only analyze the transitions for each input symbol and each new state.

Next, the steps of this process are detailed:

Input: first-order RNN trained for recognizing a regular language L.
Output: Minimal finite state automaton that accepts the strings of the regular language L.

Step 1) For each S_j state neuron (j=1,…,m) of the RNN:

- Step 1.1) To extract a TSK FRBS with two rules from the neuron. This FRBS represents the function δ_j.
- Step 1.2) To simplify the fuzzy propositions of the system.
- Step 1.3) To rewrite the fuzzy rules by considering discrete the inputs of the fuzzy system. These inputs are the output values of the S_j state neurons and the input I in the time t.

Step 2) Starting from the initial state ($S_1^1=0$, ..., $S_m^1=0$), to analyze the structure of the automaton for each input and state by using the discrete description of the δ_j transition functions.

Step 3) To minimize the obtained automaton by making use of a standard minimization algorithm [6].

7 EXAMPLE OF EXTRACTION

In order to show as FSA can be extracted from first-order RNNs, Tomita's 4th grammar has been chosen [15]. This grammar generates the language composed by the binary strings not containing "000" as a substring. Its automaton is illustrated in Figure 3.

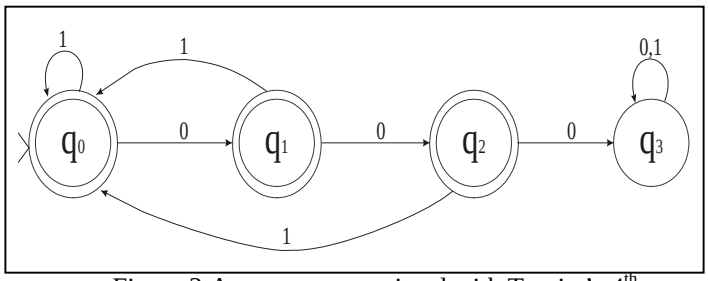


Figure 3:Automaton associated with Tomita's 4th grammar

In [9], a RNN with three state neurons (S_1, S_2, S_3) was trained to recognize the language that is generated by Tomita's 4th grammar. This trained RNN will be used for showing an example of extraction of FSA from RNNs.

This RNN was trained using the *real time recurrent learning* algorithm [16]. More details about the training conditions can be found in [9].

After training, the RNN that classifies all the strings in the data set had the following dynamics [9]:

$S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 + S_2^t \cdot 8.4 + S_3^t \cdot 0.2 + I^t \cdot (-6) - 4.6)$.
$S_2^{t+1} = sigm\,(S_1^t \cdot 0.2 + S_2^t \cdot (-0.2) + S_3^t \cdot 4.5 + I^t \cdot (-3) - 3.2)$.
$S_3^{t+1} = sigm\,(S_1^t \cdot 0 + S_2^t \cdot 0 + S_3^t \cdot 0 + I^t \cdot (-8) + 3.8)$.

7.1. APPLYING THE ALGORITHM

Step 1
For each state neuron (S_1, S_2, S_3):

State neuron S_3

Step 1.1) From the neuron S_3,

$$S_3^{t+1} = sigm\,(S_1^t \cdot 0 + S_2^t \cdot 0 + S_3^t \cdot 0 + I^t \cdot (-8) + 3.8)$$
$$= sigm\,(S_1^t \cdot 0 + S_2^t \cdot 0 + S_3^t \cdot 0 + [(-8) \cdot (I^t - 0.475)])$$

the following TSK fuzzy system can be extracted which describes the function δ_3:

If I^t is $sigm((-8)\cdot(x-0.475))$ **then** $S_3^{t+1} = 1$
If I^t is not $sigm((-8)\cdot(x-0.475))$ **then** $S_3^{t+1} = 0$

Which can be linguistically interpreted as:

If I^t is $\approx< 0.475$ **then** $S_3^{t+1} = 1$
If I^t is $\approx> 0.475$ **then** $S_3^{t+1} = 0$

Step 1.2) In this case, there is not simplification process.

Step 1.3) As $I^t \in \{0,1\}$, the system can be approximated by the following discrete system:

If $I^t = 0$ **then** $S_3^{t+1} = 1$
If $I^t = 1$ **then** $S_3^{t+1} = 0$

Intuitively, this discrete system indicates that the neuron S_3 is only activated when the input '0' is presented.

State neuron S_2

Step 1.1) From the neuron S_2,

$$S_2^{t+1} = sigm\,(S_1^t \cdot 0.2 + S_2^t \cdot (-0.2) + S_3^t \cdot 4.5 + I^t \cdot (-3) - 3.2)$$

we can see that the weights w_{12} and w_{22} produce membership functions with small slope. Therefore, we will use the bias value for being able to replace these membership functions by constant values in the future simplification process:

$$S_2^{t+1} = sigm\,((S_1^t \cdot 0.2 + 8) + (S_2^t \cdot (-0.2) - 8) + (S_3^t \cdot 4.5) + (I^t \cdot (-3) - 3.2)$$
$$= sigm\,([0.2\cdot(S_1^t+40)] + [(-0.2)\cdot(S_2^t+40)] + [4.5\cdot S_3^t] + [(-3)\cdot(I^t+1.07)]).$$

The following TSK fuzzy system can be extracted from the previous neuron S_2 which describes the function δ_2:

If S_1^t is $sigm(0.2\cdot(x+40))$ $\wedge$
S_2^t is $sigm((-0.2)\cdot(x+40))$ $\wedge$
S_3^t is $sigm(4.5\cdot x)$ $\wedge$
I^t is $sigm((-3)\cdot(x+1.07))$
then $S_2^{t+1} = 1$

If S_1^t is not $sigm(0.2\cdot(x+40))$ $\wedge$
S_2^t is not $sigm((-0.2)\cdot(x+40))$ $\wedge$
S_3^t is not $sigm(4.5\cdot x)$ $\wedge$
I^t is not $sigm((-3)\cdot(x+1.07))$
then $S_2^{t+1} = 0$

Step 1.2) As the proposition "S_1^t is $sigm(0.2\cdot(x+40))$" is practically equal to 1.0 when $S_1^t \in (0,1)$ and the proposition "S_2^t is $sigm((-0.2)\cdot(x+40))$" is practically equal to 0.0 when $S_2^t \in (0,1)$, we can apply the simplification process on these propositions.

As $(0.2\cdot(x+40)) \in [8,8.2]$ and $((-0.2)\cdot(x+40)) \in [-8.2, -8]$, we can replace the terms of the previous propositions by the mean value of these domains. We obtain the propositions "S_1^t is $sigm(8.1)$" and "S_2^t is $sigm(-8.1)$".

With these new propositions the TSK fuzzy system is:

If S_1^t is *sigm*(8.1) $\wedge$ S_2^t is *sigm*(–8.1)
$\wedge$ S_3^t is *sigm*(4.5·x) $\wedge$ I^t is *sigm*((–3)·(x+1.07))
then $S_2^{t+1} = 1$

If S_1^t is not *sigm*(8.1) $\wedge$ S_2^t is not *sigm*(–8.1)
$\wedge$ S_3^t is not *sigm*(4.5·x)
$\wedge$ I^t isnot *sigm*((–3)·(x+1.07))
then $S_2^{t+1} = 0$

Transforming this fuzzy system into neuron, we have:

S_2^{t+1} = *sigm* ([8.1]+ [–8.1]+[4.5·S_3^t]+[(–3)·(I^t+1.07)])
= *sigm* ([4.5·S_3^t]+[(–3)·(I^t+1.07)]).

Step 1.3) As $I^t \in \{0,1\}$, the action of the neuron S_2 can be rewritten as:

R_1) When I^t =0, S_2^{t+1} = *sigm* (4.5·S_3^t+(–3.21)) = *sigm* (4.5·(S_3^t–0.71)).
R_2) When I^t =1, S_2^{t+1} = *sigm* (4.5·S_3^t+(–6.21)) = *sigm* (4.5·(S_3^t–1.38)).

Hence, we can extract a new fuzzy description of the neuron S_2:

R_1) **When** I^t =0
If S_3^t is *sigm*(4.5·(x–0.71)) **then** $S_2^{t+1} = 1$
If S_3^t is not *sigm*(4.5·(x–0.71)) **then** $S_2^{t+1} = 0$

R_2) **When** I^t =1
If S_3^t is *sigm*(4.5·(x–1.38)) **then** $S_2^{t+1} = 1$
If S_3^t is not *sigm*(4.5·(x–1.38)) **then** $S_2^{t+1} = 0$

Which can be linguistically interpreted as:

R_1) **When** I^t =0
If S_3^t is ≈> 0.71 **then** $S_2^{t+1} = 1$
If S_3^t is ≈< 0.71 **then** $S_2^{t+1} = 0$

R_2) **When** I^t =1
If S_3^t is ≈> 1.38 **then** $S_2^{t+1} = 1$
If S_3^t is ≈< 1.38 **then** $S_2^{t+1} = 0$

Considering discrete the output values of the neuron S_2 in {0,1}, the proposition "S_3^t is ≈> 1.38" is always false and "S_3^t is ≈< 1.38" is always true. Therefore, rule R_2 can be simplified.

The fuzzy system can be translated into the discrete following:

R_1) **When** I^t =0
If $S_3^t = 1$ **then** $S_2^{t+1} = 1$
If $S_3^t = 0$ **then** $S_2^{t+1} = 0$

R_2) **When** I^t =1 $S_2^{t+1} = 0$

Intuitively, this discrete system indicates that the neuron S_2 is activated when the input '0' is presented and S_3 is active (input '0' was presented in the time (t–1)).

State neuron S_1

Step 1.1) From the neuron S_1,
S_1^{t+1} = *sigm*(S_1^t·15.2 + S_2^t ·8.4 + S_3^t ·0.2 + I^t ·(–6)–4.6)

we can see that the weight w_{31} produces a membership function with small slope. Therefore, we will use the bias value for being able to replace this membership function by a constant value in the future simplification process:

S_1^{t+1} = *sigm* ((S_1^t ·15.2) + (S_2^t ·8.4) + (S_3^t ·0.2 +8) + (I^t ·(–6) – 4.6 –8)
= *sigm* ((S_1^t ·15.2) + (S_2^t ·8.4) + (S_3^t ·0.2 +8) + (I^t ·(–6) – 12.6)
= *sigm* ([15.2·S_1^t]+ [8.4·S_2^t]+ [0.2·(S_3^t+40)]+ [(–6)·(I^t+2.1)]).

The following TSK fuzzy system can be extracted from the previous neuron S_1 which describes the function δ_1:

If S_1^t is *sigm*(15.2·x) $\wedge$ S_2^t is *sigm*(8.4·x)
$\wedge$ S_3^t is *sigm*(0.2·(x+40))
$\wedge$ I^t is *sigm*((–6)·(x+2.1))
then $S_1^{t+1} = 1$

If S_1^t is not *sigm*(15.2·x) $\wedge$ S_2^t is not *sigm*(8.4·x)
$\wedge$ S_3^t is not *sigm*(0.2·(x+40))
$\wedge$ I^t is not *sigm*((–6)·(x+2.1))
then $S_1^{t+1} = 0$

Step 1.2) As the proposition " S_3^t is *sigm*(0.2·(x+40))" is practically equal to 1.0 when $S_3^t \in (0,1)$, we can apply the simplification process on this proposition.

As (0.2·(x+40))∈[8,8.2], we can replace the term of the previous proposition by the mean value of this domain. We obtain the propositions "S_3^t is *sigm*(8.1)". With this new proposition the TSK fuzzy system is:

If S_1^t is *sigm*(15.2·x) $\wedge$ S_2^t is *sigm*(8.4·x)
$\wedge$ S_3^t is *sigm*(8.1) $\wedge$ I^t is *sigm*((–6)·(x+2.1))
then $S_1^{t+1} = 1$

If S_1^t is not *sigm*(15.2·x) $\wedge$ S_2^t is not *sigm*(8.4·x)
$\wedge$ S_3^t is not *sigm*(8.1)
$\wedge$ I^t is not *sigm*((–6)·(x+2.1))
then $S_1^{t+1} = 0$

Transforming this fuzzy system into neuron, we have:
$S_1^{t+1} = sigm\,([15.2 \cdot S_1^t] + [8.4 \cdot S_2^t] + [8.1] + [(-6) \cdot (I^t + 2.1)])$
$= sigm\,(S_1^t \cdot 15.2 + S_2^t \cdot 8.4 + I^t \cdot (-6) - 4.5)$.

Step 1.3) As $I^t \in \{0,1\}$, the action of the neuron S_1 can be rewritten as:

R_1) When $I^t = 0$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 + S_2^t \cdot 8.4 - 4.5)$.
R_2) When $I^t = 1$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 + S_2^t \cdot 8.4 - 10.5)$.

Now, we can analyze the behavior of this system when the output of neuron S_1^t or the neuron S_2^t is discrete in {0,1}. The final discrete system is the same in an independent way of the chosen neuron. We are going to consider the neuron S_2^t.

If we suppose that $S_2^t \in \{0,1\}$ the action of the neuron S_1 is:

$R_{1.1}$) When $I^t = 0$ and $S_2^t = 0$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 - 4.5) = sigm\,(15.2 \cdot (S_1^t - 0.3))$.
$R_{1.2}$) When $I^t = 0$ and $S_2^t = 1$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 + 3.9) = sigm\,(15.2 \cdot (S_1^t + 0.26))$.
$R_{2.1}$) When $I^t = 1$ and $S_2^t = 0$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 - 10.5) = sigm\,(15.2 \cdot (S_1^t - 0.69))$.
$R_{2.2}$) When $I^t = 1$ and $S_2^t = 1$, $S_1^{t+1} = sigm\,(S_1^t \cdot 15.2 - 2.1) = sigm\,(15.2 \cdot (S_1^t - 0.14))$.

From this action we can extract the following fuzzy rules:

$R_{1.1}$) **When** $I^t = 0$ **and** $S_2^t = 0$,
If S_1^t is $\approx>$ 0.3 **then** $S_1^{t+1} = 1$
If S_1^t is $\approx<$ 0.3 **then** $S_1^{t+1} = 0$
$R_{1.2}$) **When** $I^t = 0$ **and** $S_2^t = 1$,
If S_1^t is $\approx>$ (−0.26) **then** $S_1^{t+1} = 1$
If S_1^t is $\approx<$ (−0.26) **then** $S_1^{t+1} = 0$
$R_{2.1}$) **When** $I^t = 1$ **and** $S_2^t = 0$,
If S_1^t is $\approx>$ 0.69 **then** $S_1^{t+1} = 1$
If S_1^t is $\approx<$ 0.69 **then** $S_1^{t+1} = 0$
$R_{2.2}$) **When** $I^t = 1$ **and** $S_2^t = 1$,
If S_1^t is $\approx>$ 0.14 **then** $S_1^{t+1} = 1$
If S_1^t is $\approx<$ 0.14 **then** $S_1^{t+1} = 0$

If this system is analyzed by considering that $S_1^t \in \{0,1\}$, the rule $R_{1.2}$ is simplified. We obtain the following discrete system:

$R_{1.1}$) **When** $I^t = 0$ **and** $S_2^t = 0$,
If $S_1^t = 1$ **then** $S_1^{t+1} = 1$
If $S_1^t = 0$ **then** $S_1^{t+1} = 0$
$R_{1.2}$) **When** $I^t = 0$ **and** $S_2^t = 1$,
then $S_1^{t+1} = 1$
$R_{2.1}$) **When** $I^t = 1$ **and** $S_2^t = 0$,
If $S_1^t = 1$ **then** $S_1^{t+1} = 1$
If $S_1^t = 0$ **then** $S_1^{t+1} = 0$
$R_{2.2}$) **When** $I^t = 1$ **and** $S_2^t = 1$,
If $S_1^t = 1$ **then** $S_1^{t+1} = 1$
If $S_1^t = 0$ **then** $S_1^{t+1} = 0$

This discrete system is easily comprehensible. It can be summarized so:

When $I^t = 0$ **and** $S_2^t = 1$, $S_1^{t+1} = 1$
Otherwise, $S_1^{t+1} = S_1^t$

Intuitively, this discrete system indicates that the neuron S_1 with output '0' only changes to '1' when the input '0' is presented and S_2 is active (substring '00' is being presented), that is, when substring '000' appears. On the other hand, the neuron S_1 with output '1' (substring '000' was presented) remains with this value.

Step 2

In the previous steps, we have found discrete systems that describe the functions δ_1, δ_2 and δ_3. With all these functions δ_j, a description of the δ transition function of the automaton is obtained. Starting from the initial state ($S_1^1=0$, $S_2^1=0$, $S_3^1=0$), we can visit every state of the automaton. We must only analyze each transition for every input and state.

With the previous functions δ_j, we get the automaton that is illustrated in Figure 4. It corresponds with the not minimal one that recognizes the language generated by Tomita's 4th grammar.

Step 3

When the automaton of Fig. 4 is minimized, the minimal one associated with the Tomita's 4th grammar (Figure 3) is achieved.

In this way, we have shown with an example as the presented method can extract FSA from RNNs trained with strings of a regular language.

8 CONCLUSIONS

Several Methods to understand feedforward artificial neural networks by means of fuzzy rule based systems have been recently published. These works have been used to obtain a new procedure for extracting FSA from RNNs that infer regular grammars.

The similarities between RNNs and FSA are analyzed. The part of the RNN that is related with the transition function of the automaton is interpreted with fuzzy rules. Thus, we can understand the transitions among states of

the automaton and it is easy to get its structure from the initial state.

This study allows to extract FSA from trained RNNs. Besides, it is not necessary a search process in the output domain of the neurons to find the states of the automaton. This is an advantage of the presented method opposite other previously published methods to extract FSA from RNNs.

Acknowledgements

This work has been supported by the Project TIN2007-60199.

References

[1] Alquézar R., Sanfeliu A., "An algebraic framework to represent finite state machines in single-layer recurrent neural networks", Neural computation, vol. 7, no5, p. 931-949, 1995

[2] Rafael C. Carrasco , Mikel L. Forcada, "Simple Strategies to Encode Tree Automata in Sigmoid Recursive Neural Networks", IEEE Transactions on Knowledge and Data Engineering, v.13 n.2, p.148-156, March 2001

[3] Chomsky, Noam, "Three Models for the Description of Language,", IRE Transactions on Information Theory, Vol. 2 No. 2, p. 113-123, 1956.

[4] Elman J.L., "Finding structure in time", Cognitive Science, 14(2), p. 179-211, 1990.

[5] C. Lee Giles, C.B. Miller, D. Chen, H.H. Chen, G.Z. Sun, and Y.C. Lee, "Learning and extracting finite state automata with second-order recurrent neural networks", Neural Computation, 4(3), p. 393-405, 1992.

[6] Hopcroft J. & Ullman J.D., "Introduction to Automata Theory, Languages, and Computation", Reading, MA: Addison-Wesley, 1979.

[7] H. Jacobsson, "Rule Extraction from Recurrent Neural Networks: A Taxonomy and Review", Neural Computation, Vol. 17, Number 6, p. 1223-1263, 2005.

[8] E. Kolman & M. Margaliot, "Are neural networks white boxes?", IEEE Transactions on Neural Networks, vol.16, 4, pp. 844-852, 2005.

[9] E. Kolman & M. Margaliot, "Extracting Symbolic Knowledge from Recurrent Neural Networks- A Fuzzy Logic Approach", to be published in Fuzzy Sets and Systems, 2008 (available online in http://www.eng.tau.ac.il/~michaelm/).

[10] C.J. Mantas & J.M. Puche "Artificial Neural Networks are zero-order TSK Fuzzy Systems", to be published in IEEE Transactions on Fuzzy Systems, 2008.

[11] Christian W. Omlin , C. Lee Giles "Constructing deterministic finite-state automata in recurrent neural networks", Journal of the ACM (JACM), v.43 n.6, p.937-972, Nov. 1996.

[12] Christian W. Omlin, C. Lee Giles, "Rule Revision With Recurrent Neural Networks", IEEE Transactions on Knowledge and Data Engineering, v.8 n.1, p.183-188, February 1996.

[13] Omlin, C.W., Thornber, K.K. and Giles, C.L., "Fuzzy finite-state automata can be deterministically encoded in recurrent neural networks", IEEE Transactions on Fuzzy Systems, v5, p. 76-89, 1998.

[14] T. Takagi & M. Sugeno, "Fuzzy identification of systems and its application to modeling and control", IEEE Transactions on Systems, Man and Cybernetics, 15 (1), p. 116-132, 1985.

[15] Tomita M., "Dynamic construction of finite-state automata from examples using hillclimbing", in 'Proceedings of Fourth Annual Cognitive Science Conference', Ann Arbor, MI, p. 105–108, 1982.

[16] R.J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks", Neural Computation, vol. 1, p. 270-280,1989.

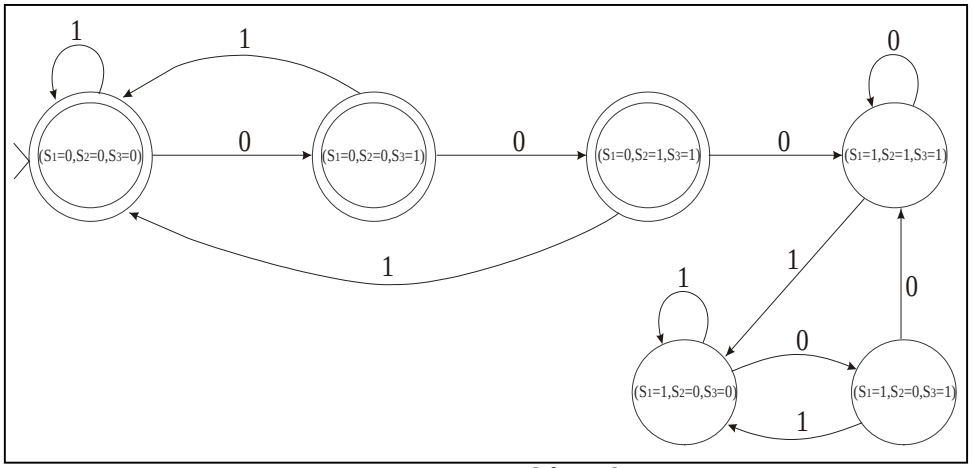


Figure 4: Automaton extracted from the RNN

UTILIZACIÓN DE UN SISTEMA BASADO EN REGLAS DIFUSAS PARA LA APLICACIÓN DE OPERADORES EN UN ALGORITMO COOPERATIVO-COMPETITIVO

M. Dolores Pérez Godoy A. Jesús Rivera Rivas M. José del Jesus Díaz F. José Berlanga Rivera

Departamento de Informática, Universidad de Jaén, {lperez, arivera, mjjesus, berlanga}@ujaen.es

Resumen

En este trabajo se presenta CO^2RBFN, un algoritmo bioinspirado para el diseño de redes neuronales, concretamente Redes de Funciones de Base Radial (RBFNs). El método de aprendizaje está basado en la programación evolutiva con un enfoque cooperativo-competitivo en el que cada individuo representa una neurona y la población al completo la red. Destaca en el mismo la utilización de un Sistema Basado en Reglas Difusas (SBRD) que representa conocimiento experto sobre los operadores a aplicar a una neurona en base a sus características. El algoritmo se ha aplicado a la resolución de problemas de clasificación.

Palabras Clave:RBFNs, SBRDs, Clasificación, Cooperativo-Competitivo.

1 INTRODUCCIÓN

Las Redes de Funciones de Base Radial (RBFNs) constituyen uno de los paradigmas más populares dentro del campo de las Redes Neuronales. Este modelo de cómputo ha demostrado su solvencia a la hora de abordar problemas como aproximación de funciones [7], clasificación [4] o predicción de series temporales [20]. Las Funciones de Base Radial (RBFs) se usaron inicialmente en interpolación numérica y aproximación funcional [17]. A finales de los 80 tienen lugar las primeras investigaciones de redes neuronales basadas en RBFs [3][11].

Son muchas las características que despiertan el interés por este tipo de redes entre las cuales cabe destacar: su topología simple con tan solo una capa oculta, su capacidad de ser un aproximador universal [13] o la analogía entre la salida de estas redes y los campos receptivos localizados, encontrados en estructuras biológicas cerebrales.

Básicamente una salida de una RBFN ofrece una suma ponderada, por unos pesos, de las respuestas de cada una de las neuronas/RBFs que conforman la red. Una salida o respuesta de una RBF tiene un carácter local y viene dada en función de un centro y un radio. Así pues, el objetivo a la hora de diseñar una RBFN es determinar el centro y el radio que caracteriza cada RBF así como su peso para cada salida de la red.

Las técnicas empleadas en el diseño de RBFNs son muy diversas. El algoritmo típico de diseño de RBFNs tiene dos etapas. En la primera etapa se determinan los centros y los radios de las RBFs, mientras que en la segunda etapa se calculan sus pesos. Para determinar los centros y los radios se pueden emplear técnicas de clustering [14]. En la segunda etapa y para calcular los pesos se pueden utilizar algoritmos como el LMS [21], SVD [6], etc.

El diseño de una RBFN también se puede abordar desde el punto de vista de la computación evolutiva [1]. En la mayoría de estas propuestas evolutivas, un individuo representa una RBFN completa. De esta manera, los operadores evolutivos actúan sobre los individuos añadiendo RBFs, eliminándolas o modificándolas.

Sin embargo y según el trabajo de Potter [16] la computación evolutiva tradicional presenta ciertos problemas relacionados con la evaluación de subcomponentes independientes en el diseño de ciertos modelos. Así se propone la Coevolución Cooperativa [16] para extender el modelo evolutivo tradicional y conseguir un entorno de diseño en el que los individuos de la población representen una parte de la solución evolucionando en paralelo. Ahora los individuos no sólo compiten por sobrevivir sino que deben cooperar para alcanzar una solución.

Los autores ya han desarrollado un importante trabajo en diseño híbrido de RBFNs para la aproximación de funciones y predicción de series temporales [18].

Manteniendo el enfoque coopertivo-competitivo pero cambiando operadores, evaluación de individuos y esquema evolutivo, actualmente trabajamos en el desarrollo de una nueva propuesta para el diseño de RBFNs orientadas a resolver problemas de clasificación [15]. En el trabajo actual se propone un marco de evolución cooperativo-competitivo donde además se utiliza un sistema basado en reglas difusas para decidir la aplicación de los operadores.

La organización del resto de este trabajo se describe a continuación. En la sección 2 se introducen las RBFNs y se explica cómo se aplican a problemas de clasificación de patrones. En la sección 3 se presenta el algoritmo propuesto y sus resultados se muestran en la sección 4. Por último se explican las conclusiones alcanzadas y las líneas de trabajo futuro en la sección 5.

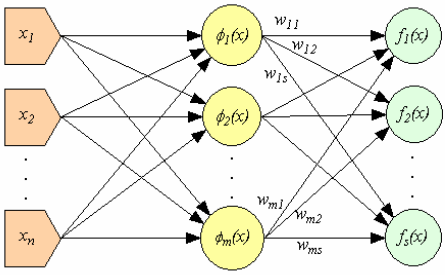


Figura 1: Topología de una RBFN

2 REDES DE FUNCIONES DE BASE RADIAL Y SU APLICACIÓN A PROBLEMAS DE CLASIFICACIÓN

Una Red de Funciones de Base Radial es un tipo de red neuronal hacia delante con tres capas: la capa de entrada con *n* nodos, una capa oculta con *m* neuronas o RBFs, y una capa de salida con uno o varios nodos, ver Figura 1. Las *m* neuronas/RBFs de la capa oculta ofrecen una activación simétrica radial $\phi_i: R^n \rightarrow R$, que puede tomar diferentes formas aunque la más común es la función gausiana que viene dada por la expresión: $\phi_i(\vec{x}) = \phi_i(e^{-(\|\vec{x}-\vec{c}_i\|/d_i)^2})$, donde $\vec{c}_i \in R^n$ es el centro de la función base ϕ_i, $d_i \in R$ es el radio y como $\|\ \|$ suele utilizarse la norma euclídea en R^n. Los nodos de salida implementan la siguiente ecuación:

$$f_j(\vec{x}) = \sum_{i=1}^{m} w_{ij}\phi_i(\vec{x}) \qquad (1)$$

En un escenario de clasificación la RBFN debe establecer una relación entre un espacio de entrada X^n hacia un conjunto finito de clases *Y* con *k* clases. De esta manera un típico conjunto de entrenamiento *S* sería:

$$S = \{(\vec{x}_u, y_u) \mid x_u \in X^n,\ y \in Y,\quad u = 1,..., p\} \qquad (2)$$

donde $\vec{x}_u$ es el vector de características e y_u es la clase a la que pertenece. Normalmente, en un problema de clasificación, el número de salidas de la RBFN se corresponde con el número de clases (*k*). Para entrenar la RBFN la pertenencia a la y_u se codifica en un vector binario $\vec{z}_u \in \{0,1\}^k$ a través de la relación $\vec{z}_u^i = 1$ sii $y_u = i$, y $\vec{z}_u^i = 0$ *en otro caso*. La clase obtenida por la red, ante una determinada entrada, será la salida de la red con un mayor grado de activación.

En la bibliografía especializada se pueden encontrar distintos métodos evolutivos [4][9][12] para el diseño de RBFNs orientadas al problema de clasificación. Sin embargo, la mayoría de las aproximaciones existentes trabajan con métodos evolutivos típicos donde un individuo representa una red completa, por lo que pueden tener problemas de alto costo computacional y convergencia prematura a mínimos locales. Como se ha comentado anteriormente estos problemas pueden mitigarse utilizando técnicas cooperativas-competitivas donde un individuo representa una neurona o RBF tal y como se propone en este trabajo. Hasta ahora en la bibliografía no son muchos los trabajos [15][18][19][20] que implementan este tipo de paradigma, debido sobre todo a la dificultad que entraña definir los parámetros de cooperación y competición entre los individuos.

3 CO^{2}RBFN: UN ALGORITMO HÍBRIDO COOPERATIVO-COMPETITIVO PARA EL DISEÑO DE RBFNs

Se propone un enfoque cooperativo-competitivo híbrido dentro de un marco evolutivo, para la resolución de problemas de clasificación. En esta propuesta cada individuo de la población representa una función base y la población entera es la responsable de la solución final. Se tiene un entorno en el que los individuos cooperan para alcanzar la solución definitiva, no obstante, también compiten por la supervivencia, dado que si el trabajo de un individuo no es bueno dicho individuo será eliminado.

Este escenario tiene en cuenta la respuesta local de las neuronas y la interpretabilidad de esta clase de redes, lo cual establece una importante guía de diseño. Para medir la asignación de crédito de un individuo, se proponen tres factores que evalúan el papel de cada RBF en la red. El algoritmo usa un Sistema Basado en Reglas Difusas para decidir la probabilidad de aplicación de los operadores a las RBFs.

El algoritmo puede trabajar tanto con atributos lineales (reales y enteros) como nominales. La distancia utilizada es la distancia HDVM [22]:

$$HVDM(x, y) = \sqrt{\sum_{a=1}^{n} d_a^2(x_a, y_a)} \qquad (3)$$

donde *n* es el número de atributos. La función $d_a^2(x,y)$ devuelve la distancia entre los valores *x* e *y* del atributo *a* y se define como:

$$d_a(x,y) = \begin{cases} 1 \quad si\ x\ o\ y\ son\ \ desconocidos \\ VDM_a(x,y) \quad si\ a\ \ es\ \ \text{nominal} \\ \text{distancia euclídea} \quad si\ \ a\ \ es\ \ lineal \end{cases} \tag{4}$$

Las etapas principales de CO^{2}RBFN se explican en las siguientes subsecciones. A continuación se muestra su pseudocódigo:

```
1. Inicialización de las RBFs
2. Entrenamiento de las RBFs
3. Evaluación de las RBFs
4. Aplicación de los operadores
   a las RBFs
5. Sustitución de las RBFs que
   han sido eliminadas
6. Selección de las mejores
   RBFs
7. Salto al paso 2 si no se
   verifica la condición de
   Parada
```

Algoritmo 1. Principales etapas de CO^{2}RBFN

3.1. INICIALIZACIÓN DE LA RED

El proceso para definir la red inicial es simple. Las neuronas se colocan aleatoriamente intentado cubrir las diferentes clases de los patrones del conjunto de entrenamiento. El número de RBFs está especificado como un parámetro (es el tamaño de la población, m).
El centro de cada RBF, $\vec{c}_i$, se inicializa a partir de un patrón del conjunto de entrenamiento, teniendo en cuenta que las RBFs se distribuyan equitativamente entre las diferentes clases existentes. El radio, d_i, se inicializa a la mitad de la media de las distancias entre los centros. Finalmente los pesos, w_{ij} , se ponen a cero.

3.2. ENTRENAMIENTO DE LAS RBFs

Durante esta etapa se entrenan los pesos de las RBFs. El propósito de entrenar los pesos es el de explotar la información local extraída del comportamiento de las RBFs. Dicho entrenamiento se realiza mediante la técnica LMS [21].

3.3. EVALUACIÓN DE LAS RBFs

Se requiere un mecanismo de asignación de crédito para poder evaluar el papel de cada función base dentro del entorno cooperativo-competitivo. Para este propósito se consideran tres parámetros, a_i, e_i, o_i para cada RBF ϕ_i.
La contribución, a_i, de la RBF ϕ_i, $i=1...m$, a la salida de la RBFN, se determina considerando su peso, w_i, y el número de patrones de entrenamiento dentro de su radio, pi_i . Se penaliza una RBF con poco peso y pocos patrones dentro de su radio:

$$a_i = \begin{cases} |w_i| & if\ \ pi_i > q \\ |w_i| * (pi_i / q) & en\ otro\ caso \end{cases} \tag{5}$$

donde q es la media de los valores pi_i menos su desviación típica.
La medida de error, e_i, para cada RBF ϕ_i, se obtiene contando los patrones que estando dentro de su radio están mal clasificados:

$$e_i = \frac{pibc_i}{pi_i} \tag{6}$$

donde $pibc_i$ y pi_i son el número de patrones mal clasificados y el número total de patrones, dentro de su radio, respectivamente.
El solapamiento de la RBF, ϕ_i, con otras RBFs se cuantifica usando el parámetro o_i. Este parámetro se calcula partiendo de la base de la metodología *fitness sharing* [5], que intenta mantener la diversidad en la población. El factor se expresa como:

$$o_i = \sum_{j=1}^{m} o_{ij} \tag{7}$$

donde o_{ij} mide el solapamiento entre las RBF ϕ_i y ϕ_j, $j=1...m$.

3.4. APLICACIÓN DE LOS OPERADORES A LAS RBFs

En el algoritmo existen cuatro operadores que se pueden aplicar a las RBFs:

- Operador Elimina: es un operador que elimina una RBF.
- Operador Mutación Aleatoria: es un operador que modifica el centro y el radio de la RBF aleatoriamente. El radio se modifica con una probabilidad inversamente proporcional al número de características en el problema de clasificación (n). La modificación, en caso de llevarse a cabo, puede oscilar en un porcentaje entre un 5% y un 50% sobre el radio antiguo. En cuanto a la modificación del centro, el número de coordenadas que se van a mutar se obtiene aleatoriamente entre un 1% y un 25 % del total de características del conjunto de entrenamiento. Si la coordenada a mutar tiene un valor real se le provoca un incremento o decremento aleatorio de un valor entre un 5% y un 50% del radio. Por el contrario, si tiene un valor nominal se produce un cambio a uno de los valores nominales posibles, con una probabilidad inversamente proporcional a la distancia entre el valor actual y los posibles valores de cambio.
- Operador Mutación Informada: este operador también puede modificar el radio y algunas coordenadas del centro de la RBF usando información de su entorno. El centro se modifica como sigue:

$$c'_{ij} = c_{ij} \pm h \quad \forall j = 1 \ldots n \tag{8}$$

El incremento o decremento del centro antiguo se decide mediante un número aleatorio h ($0.05 \cdot di \leq h \leq 0.5 \cdot di$). El centro se modifica para aproximar la neurona al centro de los patrones de entrenamiento que están dentro del radio de la RBF y pertenecen a su misma clase.

El objetivo de modificar el radio es que el mayor número posible de patrones que pertenecen a la misma clase que la RBF, queden dentro de su radio. De esta forma el radio se modifica de la siguiente forma:

$$\begin{cases} d' = d + h & si\ u \leq D + A \\ d' = d - h & en\ otro\ caso \end{cases} \quad (9)$$

donde h, es un número aleatorio $(0.05 \cdot d_i \leq h \leq 0.5 \cdot d_i)$, D es el número de patrones que no pertenecen a la clase de la RBF dividido por el número de patrones que pertenecen a la clase, ambas cantidades medidas dentro de su radio. A, es el número de patrones que pertenecen a la misma clase que la RBF dividido por el número de patrones que no pertenecen a la misma clase, ambas cantidades medidas dentro de dos veces el radio de la RBF.

- Operador Nulo: no se realiza ninguna acción sobre la RBF.

Estos operadores se aplican a la población entera de RBFs. La probabilidad de elegir un operador para una RBF dada viene determinada por un sistema difuso tipo Mamdani [10], cuyas entradas son los parámetos a_i, e_i y o_i, éstos determinan la asignación de crédito de cada RBF. Estas entradas se consideran como variables lingüísticas va_i, ve_i y vo_i , y las salidas son $p_{elimina}$, p_{ma}, p_{mi} y p_n, que representan la probabilidad de aplicar los operadores Elimina, Mutación Aleatoria, Mutación Informada y Nulo, respectivamente. La Figura 1 muestra las funciones de pertenencia para las etiquetas lingüísticas de las entradas y las salidas respectivamente. El número de etiquetas lingüísticas se ha determinado empíricamente, con centros y bases directamente relacionados con su significado. La Tabla 1 muestra las reglas usadas para relacionar los antecedentes y los consecuentes. El bajo número de reglas permite que el sistema difuso diseñado sea muy simple.

En el diseño de reglas se tiene en cuenta el hecho de que una RBF es peor si su contribución (a_i) es baja, su error (e_i) es alto y su solapamiento (o_i) también es alto. En el lado opuesto, una RBF es mejor cuando su contribución es alta y su error y solapamiento son bajos. De esta forma, la probabilidad de eliminar una RBF se incrementa cuando ésta tiene un mal comportamiento; la probabilidad de mutar va aumentando a medida que el comportamiento es mejor. El operador nulo se aplicará cuando el comportamiento de una RBF sea muy bueno.

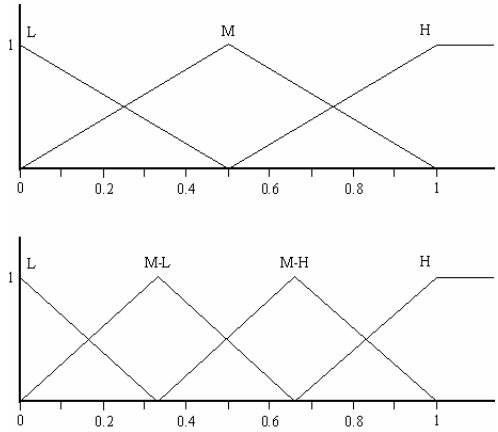


Figura 1. Arriba: etiquetas lingüísticas para las variables de entrada. Abajo: etiquetas lingüísticas de las variables de salida.

Tabla 1: Base de reglas utilizada.

	Antecedentes			Consecuentes			
	v_a	v_e	v_o	$P_{elimina}$	P_{mr}	P_{mi}	p_{nulo}
R1	L			M-H	M-H	L	L
R2	M			M-L	M-H	M-L	M-L
R3	H			L	M-H	M-H	M-H
R4		L		L	M-H	M-H	M-H
R5		M		M-L	M-H	M-L	M-L
R6		H		M-H	M-H	L	L
R7			L	L	M-H	M-H	M-H
R8			M	M-L	M-H	M-L	M-L
R9			H	M-H	M-H	L	L

3.5. INTRODUCCIÓN DE NUEVAS RBFs

En este punto el algoritmo sustituye las RBFs eliminadas por otras nuevas. Existen dos alternativas para situar las nuevas RBFs, ambas con probabilidad de 0.5: en un patrón mal clasificado que esté fuera de cualquier RBF o en un patrón obtenido de forma aleatoria.

3.6. SELECCIÓN DE LAS MEJORES RBFNs

Después de aplicar los operadores de mutación aparecen nuevas RBFs y éstas son comparadas con sus padres para determinar cuáles tienen un mejor comportamiento en la red. Las mejores RBFs serán las elegidas para formar parte de la nueva población.

4 RESULTADOS EXPERIMENTALES

Para la experimentación se han utilizado cuatro bases de datos obtenidas de UCI Repository of Machine Learning Database: Hepatitis, Sonar, Wbcd y Wine.

La Tabla 2 muestra las características de las bases de datos.

El estudio se ha realizado usando ten-fold cross validación, es decir, diez particiones para entrenamiento y test, el 90% de los datos para entrenamiento y el 10% para test. Para cada base de datos se obtiene la media de las diez particiones. El número de generaciones se fija a 200.

Tabla 2: Características de las bases de datos

Bases de Datos	*#Instancias*	*#Atributos Lineales*	*#Atributos Nominales*	*#Clases*
Hepatitis	155	6	13	2
Sonar	208	60	0	2
Wbcd	699	9	0	2
Wine	178	13	0	3

Las tablas de la 3 a la 6 muestran los resultados obtenidos con CO^{2}RBFN y con otros métodos. Los otros métodos con los que se compara pertenecen a distintos paradigmas dentro del campo del aprendizaje automático:

- C4.5: algoritmo que genera reglas de clasificación en forma de árboles de decisión a partir de un conjunto de ejemplos dado. Implementación obtenida de la herramienta Keel [8].
- GP-COACH: propuesta evolutiva que utiliza la Programación Genética para aprender SBRDs compactos y precisos en problemas con alta dimensionalidad y que sigue el enfoque cooperativo-competitivo de representación de reglas [2]. Implementación propia.
- GenéticoRBFN: algoritmo genético para el diseño de RBFNs basado en el esquema Pittsburgh. En dicho esquema un individuo es una red completa. Implementación propia.
- MultiObjetivoRBFN: algoritmo para el diseño de RBFNs con enfoque multiobjetivo, basado en el algoritmo NSGA2 (un individuo representa una red). Implementación propia.
- Perceptrón Back: algoritmo para el diseño de Redes Perceptrón Multicapa que utiliza el algoritmo Backpropagation para el aprendizaje. Implementación obtenida de la herramienta Keel [8].
- Perceptrón Grad: algoritmo para el diseño de Redes Perceptrón Multicapa que utiliza el algoritmo de Gradiente Conjugado para el aprendizaje. Implementación obtenida de la herramienta Keel [8].

El análisis de los resultados muestra que CO^{2}RBFN obtiene RBFNs con una estructura simple y los resultados son comparables, incluso superiores, a los obtenidos con otros métodos. La simplicidad de la red obtenida implica una buena interpretabilidad, lo cual es una característica importante en los problemas de clasificación.

Tabla 3: Resultados obtenidos con Hepatitis

Algoritmo	*#nodos/ reglas*	*Test (%)*
C4.5	7.1	89.647
GP-COACH	9.033	80.539
GenéticoRBFN	7.5	86.711
MultiObjetivoRNFN	8	87.479
Perceptrón Back.	30	71.817
Perceptrón Grad.	10	79.461
CO^{2}RBFN	8	**87.399**

Tabla 4: Resultados obtenidos con Sonar

Algoritmo	*#nodos/ reglas*	*Test (%)*
C4.5	14.3	71.071
GP-COACH	13.767	65.651
GenéticoRBFN	7.68	73.305
MultiObjetivoRNFN	6	73.814
Perceptrón Back.	30	67.762
Perceptrón Grad.	10	73.524
CO^{2}RBFN	8	**75.086**

Tabla 5: Resultados obtenidos con Wbcd

Algoritmo	*#nodos/ reglas*	*Test (%)*
C4.5	12.4	94.995
GP-COACH	8.433	94.896
GenéticoRBFN	6.24	96.713
MultiObjetivoRNFN	8	96.969
Perceptrón Back.	30	87.722
Perceptrón Grad.	10	94.985
CO^{2}RBFN	5	**97.083**

Tabla 6: Resultados obtenidos con Wine

Algoritmo	*#nodos/ reglas*	*Test (%)*
C4.5	5.1	94.902
GP-COACH	7.767	93.584
GenéticoRBFN	10.44	95.275
MultiObjetivoRNFN	12	95.643
Perceptrón Back.	30	93.301
Perceptrón Grad.	10	96.111
CO^{2}RBFN	7	**96.739**

5 CONCLUSIONES

Este trabajo presenta un algoritmo bioinspirado para diseñar redes de funciones de base radial. En él se mantiene una población de RBFs que cooperan para lograr una solución final y que compiten por su supervivencia. El comportamiento (asignación de crédito) de cada función base dentro de la red completa se mide en base a tres factores: la contribución, a_i, de la RBF a la salida de la red, el error, e_i, de la RBF y el solapamiento de la RBF con otras RBFs.

Se utilizan cuatro operadores para poder aplicar a una RBF dada en el proceso evolutivo: Elimina, Mutación Aleatoria, Mutación Informada y Nulo. Con las dos modalidades de mutación se consigue un equilibrio adecuado entre las cualidades de explotación y exploración que todo algoritmo evolutivo debe poseer. Por un lado la mutación informada utiliza información del entorno de una neurona para modificar ésta de forma que se adapte óptimamente a su entorno. La mutación aleatoria promueve modificaciones de forma que se favorezca la exploración del entorno y se huya de óptimos locales.

La aplicación de dichos operadores es determinada mediante un Sistema basado en Reglas Difusas. Las entradas a dicho sistema son los parámetros a_i, e_i, y o_i, usados para medir la asignación de crédito y las salidas son las probabilidades de aplicación de los distintos operadores.

El algoritmo propuesto se ha evaluado con bases de datos conocidas y los resultados obtenidos son comparables e incluso superiores a los obtenidos con otros métodos.
Como líneas de trabajo futuro se estudiará la aplicación del algoritmo a bases de datos no balanceadas.

Agradecimientos

Este trabajo ha sido subvencionado por los proyectos españoles TIN2005-04386-C05-03 y TIN2007-60587.

Referencias

[1] Bäck, T.; Hammel, U.; Schwefel, H. Evolutionary computation: comments on the history and current state. *IEEE Trans. on Evolutionary Computation*, v.1 n.1 April. Pág. 3-17, 1997.

[2] Berlanga, F.J. Del Jesus, M.J.; Herrera, F. A Novel Genetic Cooperative-Competitive Fuzzy Rule Based Learning Method using Genetic Programming for High Dimensional Problems. *3rd International Workshop on Genetic and Evolving Fuzzy Systems (GEFS08)*. WittenBommerholz (Germany) Pág. 101-106, 2008.

[3] Broomhead, D.; Lowe, D. Multivariable functional interpolation and adaptive networks. *Complex System*. v 2. Pág. 321-355, 1988.

[4] Buchtala, O.; Klimek, M.; Sick, B. Evolutionary optimization of radial basis function classifiers for data mining applications. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, Vol 35, n. 5, Oct. Pág. 928-947, 2005.

[5] Goldberg, D.; Richardson J. Genetic algorithms with sharing for multimodal function optimization. In Grefenstette (ed.), *Proc. of the 2nd International Conference on Genetic Algorithms*. Lawrence Erlbaum Associates., Pág. 41-49, 1987.

[6] Golub, G.; Van Loan, C. Matrix computations. J. *Hopkins University Press*, 3rd ed., 1996.

[7] González, J., Rojas, I., Ortega, J., Pomares, H., Fernández, J., Fco, A. Multiobjective evolutionary optimization of the size, shape, and position parameters of radial basis function networks for function approximation. *IEEE Transactions on Neural Networks*, Volume 14, Issue 6, Pág. 1478-1495, 2003.

[8] KEEL: Knowledge Extraction based on Evolutionary Learning (Spanish National Project TIC2002-04036-C05 and TIN2005-08386-C05). http://www.keel.es.

[9] Lacerda, E; Carvalho, A.; Braga A.; Ludermir T. Evolutionary Radial Functions for Credit Assessment. *Applied Intelligence* 22. Springer Netherlands. Pág. 167-181, 2005.

[10] Mandani, E.; Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man-Machine Stud.*, v. 7, n. 1, Pág. 1-13, 1975.

[11] Moody, J.; Darken, C, J. Fast learning in networks of locally-tuned processing units. *Neural Computing*. vol 1. Pág. 281-294, 1989.

[12] Neruda, R.; Kudová, P. Learning methods for radial basis function networks. *Future Generation Computer Systems*, Volume 21, Issue 7, pp 1131-1142, 2005.

[13] Park, J.; Sandberg, I. Universal approximation using radial-basis function networks. *Neural Comput.*, vol. 3, Pág. 246-257, 1991.

[14] Pedrycz, W. Conditional fuzzy clustering in the design of radial basis function neural networks. *IEEE Transactions on Neural Networks* 9 (4), Pág. 601-612, 1998.

[15] Pérez-Godoy, M; Rivera, A.J.; Jesus, M.J.; Rojas, I. CoEvRBFN: an approach to solving the classification problem with a hybrid cooperative-coevolutive algorithm. *9th International Work-Conference on Artificial Neural Networks (IWANN07)*, LNCS 4507, Pág. 324-332, 2007.

[16] Potter, M.; De Jong, K. Cooperative Coevolution: an architecture for evolving coadapted subcomponents". *Evolutionary Computation*, 8(1), Pág. 1-29, 2000.

[17] Powell, M. Radial basis functions for multivariable interpolation: A review. In *IMA. Conf. on Algorithms for the approximation of functions and data*, Pág. 143-167, 1985.

[18] Rivera, A.J.; Rojas, I.; Ortega, J.; del Jesús, M.J.; A new hybrid methodology for cooperative-coevolutionary optimization of radial basis function networks. *Soft Computing*. ISSN 1432-7643. D.O.I: http://dx.doi.org/10.1007/s00500-006-0128-9, 2006.

[19] Topchy A.; Lebedko, O.; Miagkikh V.; Kasabov N.; Adaptive training of radial basis function networks based on co-operative evolution and evolutionary programming. *Prog. in connectionist-based information syst*. N. Kasabov et al (eds), Springer, Pág. 253-258, 1998.

[20] Whitehead, B; Choate, T. Cooperative-competitive genetic evolution of Radial Basis Function centers and widths for time series prediction. *IEEE Trans. on Neural Networks*, Vol.7, No.4, July, Pág.869-880, 1996.

[21] Widrow, B.; Lehr, M.A. 30 Years of adaptive neural networks: perceptron, madaline and backpropagation. *Proceedings of the IEEE*, v. 78 n. 9, sept. Pág. 1415-1442, 1990.

[22] Wilson, D. Randall, and Tony R. Martinez. Improved Heterogeneous Distance Functions. *Journal of Artificial Intelligence Research*, vol. 6, no. 1, Pág. 1-34, 1997.

Clasificación del Cerdo Ibérico Utilizando Espectroscopía Infrarroja y Redes Neuronales

A. Guillén, F. García del Moral, G. Rubio, L. J. Herrera, I. Rojas

Resumen

Los espectros visible y cercano al infrarrojo consisten en los tonos y combinaciones de bandas de las moléculas de absorción fundamentales en la región infrarroja. El análisis de estos espectros puede resultar difícil pues que las bandas vibran y pueden solaparse. Sin embargo, la información que puede extraerse de ellos puede resultar muy beneficiosa para la industria alimenticia tanto a nivel de consumidores como de productores y distribuidores. Este beneficio vendría dado por la posibilidad de clasificar el tipo de carne mediante una prueba no invasiva que indique la calidad de la carne, su procedencia, etc. Este trabajo aplica diversas técnicas de clasificación combinadas con selección de variables de modo que sea posible escoger que longitudes de onda son las más significativas a la hora de determinar si el tipo de carne es de cerdo ibérico o de cerdo blanco. Tal y como muestran los experimentos, los resultados son prometedores e invitan a profundizar en el campo de la clasificación de objetos en base a su espectroscopía.

Palabras Clave: RBFNN, Información mutua, clasificación, NIRS .

1 Introducción

La metodología basada en la espectroscopía del infrarrojo cercano (NIRS) se fundamenta en las correlaciones que pueden existir entre la estructura del elemento y el espectro que se mide como respuesta a la irradiación de una luz. El espectro generado muestra como se han absorbido ciertas longitudes de onda en función de la vibración de las partículas. El espectro puede mostrarse muy complejo siendo difícil de interpretar o intentar asociar ciertas propiedades del elemento medido, de modo que si es posible realizar una correcta interpretación del espectro, ésta puede proporcionar información útil sobre la muestra [3].

En la industria alimentaria está aplicando esta técnica dado que estos métodos son rápidos, baratos, portables y no invasivos, de modo que la pieza no sufre ninguna alteración ni modificación que pudiera estropear sus cualidades o dejar de hacerla apta para su consumo [11, 13].

Dentro de la industria alimentaria de la carne, varios autores [10, 3, 9] sugieren la necesidad de desarrollar nuevas metodologías de aplicación de esta técnica para poder identificar las distintas calidades de la carne de modo que sea posible identificar si la carne que el consumidor va a comprar es tan buena como ha sido etiquetada. Dentro de este marco, este trabajo hace una aproximación a la combinación de las disciplinas del *Soft Computing* y NIRS. El objetivo es obtener modelos que, dado un conjunto de espectros, sean capaces de identificar que longitudes de onda son las más relevantes y, en base a ellas, poder realizar una clasificación de nuevas muestras que se realicen. Las muestras de carne con las que se ha realizado el trabajo son de dos tipos de cerdos: blanco e ibérico. Dado que hay una gran diferencia entre ambas carnes, tanto a nivel de precio como de calidad, es especialmente útil poder distinguir de qué tipo es una muestra de carne a través de su NIRS. El resto del trabajo está estructurado como se describe a continuación: la Sección 2 describirá los distintos paradigmas del *Soft Computing* aplicados a este problema, la Sección 3 mostrará los resultados de dicha aplicación y, finalmente, en la Sección 4 se describirán las conclusiones.

2 Análisis de Datos y Técnicas para Clasificación

Esta sección describe las distintas técnicas usadas para analizar y clasificar las muestras obtenidas de las medidas realizadas en el laboratorio: Información Mutua (Mutual Information, MI), Redes Neuronales de Funciones de Base Radial (Radial Basis Function Neural Networks, RBFNNs) y Máquinas de Vectores Soporte (Support Vector Machines, SVMs).

2.1 Selección de longitudes de onda mediante información mutua

Una instancia de un cerdo está compuesta por la concatenación de 4 medidas correspondientes al lado magro y veteado de las carrilladas izquierda y derecha. Dado que cada espectro tiene 2151 valores de reflectancia, cada cerdo tiene 8604 variables que lo representan. Para poder identificar cuales de esas variables son representativas para poder realizar una clasificación, se utilizará el concepto de Información Mutua (*Mutual Information*, MI), también conocido como entropía cruzada. Sea $X = \{\vec{x}_k\}$ e $Y = \{y_k\}$ para $k = 1...n$ entonces la MI entre X e Y puede ser definida como la cantidad de información que el grupo de variables X proporcionan sobre la variable de salida Y y se expresa como:

$$I(X,Y) = H(Y) - H(Y|X). \tag{1}$$

donde $H(Y)$ es la entropía de la variable Y, la cual mide la incertidumbre sobre Y. En el caso de tener variables continuas y, siguiendo la formulación de Shannon, puede definirse como:

$$H(Y) = -\int \mu_Y(y) \log \mu_Y(y) dy, \tag{2}$$

donde $\mu_Y(y)$ es la función de densidad marginal. Esta función puede definirse a su vez como la unión de la función de densidad de probabilidad $\mu_{X,Y}$ de X con Y, es decir:

$$\mu_Y(y) = \int \mu_{X,Y}(x,y) dx. \tag{3}$$

La entropía condicional $H(Y|X)$, que mide la incertidumbre de Y dado un X conocido, se define en el caso continuo como:

$$-\int \mu_X(x) \int \mu_Y(y|X=x) \log \mu_Y(y|X=x) dy dx. \tag{4}$$

La conclusión que se obtiene es que la información mutua $I(X,Y)$ es el grado de disminución de la incertidumbre sobre Y una vez que X es conocido. Gracias a las propiedades de la entropía y la información mutua, ésta última puede definirse también como:

$$I(X,Y) = H(X) + H(Y) - H(X|Y), \tag{5}$$

lo cual nos lleva a:

$$I(X,Y) = \int \mu_{X,Y}(x,y) \log \frac{\mu_{X,Y}(x,y)}{\mu_X(x)\mu_Y(y)} dx dy. \tag{6}$$

Esto nos indica que para calcular la información mutua entre dos grupos de variables pertenecientes a X, solo es necesario estimar la función de densidad de probabilidad conjunta entre X y Y. Aunque para realizar esta estimación se pueden emplear diversas técnicas como se describe en [2], la que se utiliza para la inicialización de los individuos se basa en el algoritmo de los K vecinos más cercanos (*K-Nearest Neighbors*, KNN) debido a que posee ciertas ventajas sobre los otros métodos basados en histogramas o en *kernels* [8].

2.2 Redes Neuronales de Funciones de Base Radial

2.2.1 Descripción

Dependiendo de la función de activación que posea una neurona y del número de capas ocultas, se pueden distinguir diversos tipos de redes neuronales. En concreto, Broomhead and Lowe presentaron en [1] las RBFNNs. Este tipo de redes han sido aplicadas a un gran número de problemas de clasificación y regresión, obteniendo resultados satisfactorios [7, 4]. Una RBFNN es una red donde las entradas están conectadas a todas las neuronas de la única capa oculta de procesamiento, donde cada una de ellas implementa una función de activación Gausiana definida como:

$$\phi(\vec{x}; \vec{c}_i, r_i) = \exp\left(\frac{-||\vec{c}_i - \vec{x}||^2}{r_i^2}\right), i = 1, ..., m \tag{7}$$

donde $\vec{x}$ es un vector de entrada, $\vec{c}_i$ es el centro y r_i es el radio de la i-ésima RBF.

La capa de salida realiza la suma ponderada de todas las salidas de la red neuronal:

$$\mathcal{F}(\vec{x}; C, R, \Omega) = \sum_{i=1}^{m} \phi(\vec{x}; \vec{c}_i, r_i) \cdot \Omega_i \tag{8}$$

donde Ω_i son los pesos de salida que regulan la contribución de cada una de las neuronas de la capa oculta a la salida final de la red. Este conjunto de pesos puede ser calculado de forma óptima mediante la resolución de un sistema de ecuaciones lineal. La arquitectura de una RBFNN se muestra en la Figura 1. Para obtener la salida final de la red en términos de clasificación, se redondea el valor continuo generado por la red al valor entero más cercano, asignado un 0 para el cerdo blanco y un 1 para el ibérico.

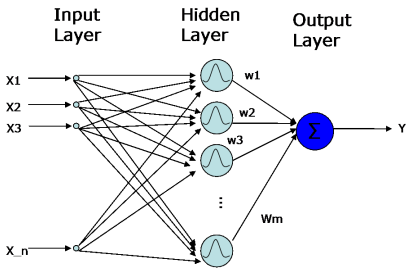


Figura 1: Arquitectura de una RBFNN.

2.2.2 Metodología de diseo

Una de las metodologías de diseo de una RBFNN consiste en una secuencia de pasos donde la inicialización de los centros de las RBFs de la capa oculta es lo primero que debe hacerse y que, tradicionalmente, se realiza usando algoritmos de clustering [6, 5, 12].

El algoritmo utilizado para disear las RBFNNs es una adaptación del presentado en [6] en la que se ha modificado el modo en el que la partición de los vectores de entrada entre los centros de los clusteres es realizada. Tal y como hacen los algoritmos de clustering clásicos, se define una función de distorsión basada en una partición posibilística que debe ser minimizada. La función de distorsión mide la combinación de las distintas funciones de pertenencia considerando la modificación de la distancia Euclidea mediante un parámetro de supervisión, quedando definida como:

$$J_h(U,C,T,W;X) = \sum_{k=1}^{n}\sum_{i=1}^{m}(u_{ik}^{h_f} + t_{ik}^{h_p})D_{ikW}^2 \qquad (9)$$

sujeta a las restricciones: $\sum_{i=1}^{m} u_{ik} = 1 \ \forall k = 1...n$ y $\sum_{k=1}^{n} t_{ik} = 1 \ \forall i = 1...m$.

La posición final de los centros es alcanzada mediante un proceso de optimización alterna mediante la iteración de Picard donde la función de distorsión se va minimizando utilizando expresiones para calcular los centros y las funciones de pertenencia sucesivamente.

2.3 Máquinas de Vectores Soporte

Los métodos de kernel se basan todos en el trabajo de Vapnik de generalización del clasificador hiperplanar para casos no lineales. El clasificador hiperplanar binario es una técnica de clasificación definida de la siguiente forma: dadas un conjunto de ejemplos (x_i, y_i), $i = 1, \ldots, N$, donde $x \in \mathbb{R}R^d$ e $y_i \in \{-1, 1\}$ (etiqueta de conjunto), se busca resolver:

$$\begin{array}{rcl} \min\limits_{w\in\mathbb{R}R^d,\ b\in\mathbb{R}R} \tau(w,b) & = & \frac{1}{2}\|w\|^2, \\ \textit{subject to} & & \\ y_i(\langle w, x\rangle + b) & \geq & 1, \quad \forall i = 1, \ldots, n. \end{array} \qquad (10)$$

Con esto se persigue obtener el hiperplano definido por (w, b) que separe las clases con mayor distancia a las mismas (con margen máximo). Este problema puede no tener solución, por lo que se relaja la condición permitiendo que se clasifiquen mal algunos ejemplos pero controlando la proporción error/precisión mediante el parámetro $C > 0$ quedando el problema como:

$$\begin{array}{rcl} \min\limits_{w\in\mathbb{R}R^d,\ b,\zeta_i\in\mathbb{R}R} \tau(w,b,\zeta) & = & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\zeta_i, \\ \textit{subject to} & & \\ \zeta_i & \geq & 0 \\ y_i(\langle w, x\rangle + b) & \geq & 1 - zeta_i, \quad \forall i = 1, \ldots, N. \end{array} \qquad (11)$$

Este es un problema de optimización de programación cuadrática resoluble mediante el método de los multiplicadores de Lagrange. Aplicando el método de los multiplicadores de Lagrange obtenemos el problema en forma dual:

$$\begin{array}{rcl} \max\limits_{\alpha_i\in\mathbb{R}R} J(\alpha) & = & \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j \langle \mathbf{x_i}, \mathbf{x_j}\rangle, \\ \textit{subject to} & & \\ \sum_{i=1}^{n}\alpha_i y_i & = & 0 \\ \alpha_i & \in & [0, C], \quad \forall i = 1, \ldots, N. \end{array} \qquad (12)$$

Resolviendo este problema puede formularse el clasificador como:

$$f(x) = \sum_{i=1}^{N}\alpha_i y_i \langle \mathbf{x_i}, \mathbf{x}\rangle + \mathbf{b} \qquad (13)$$

Como puede observarse, todo queda formulado en término de productos escalares de los vectores de entrada. Es notable que sólo los puntos de la entrada que estén en el margen tienen $\alpha_i > 0$ por lo que el resto puede ignorarse en el modelo final. Es decir, estamos

ante un modelo disperso (*sparse*), lo cual es una gran ventaja.

Las Máquinas de Vectores de Soporte (SVM) se basan en aprovechar esto para definir una función kernel $k(x, x') = \langle \Phi(x), \Phi(x) \rangle$ donde Φ es una función que mapea los puntos del espacio de entrada a un espacio de características de mayor dimensionalidad en el cual se define el hiperplano. Este proceso se denomina *Kernel Trick*. El mapeado Φ no tiene por que definirse explícitamente, ya que puede utilizarse como kernel cualquier función k que cumpla el teorema de Mercer, ejemplos de funciones que cumplen con este son los denominados kernel lineal, polinomial y Gaussiano (también RBF):

$$k(x_i, x_j) = \langle x_i, x_j \rangle. \tag{14}$$

$$k(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d, \ d \in \mathbb{R}N \tag{15}$$

$$k(x_i, x_j) = \exp\left[-\frac{1}{\sigma^2}\|x_i - x_j\|^2\right]. \tag{16}$$

Es notable que algunas funciones kernel utilizadas en la literatura no cumplen el teorema de Mercer, como el kernel sigmoide (también MLP, por su similitud a las redes neuronales MLP con función de activación sigmoide):

$$k(x_i, x_j) = \tanh\left(s\langle x_i, x_j \rangle + t\right). \tag{17}$$

El *Kernel Trick* se ha utilizado para obtener extensiones no lineales de otros algoritmos, cómo por ejemplo el Análisis de Componentes Principales.

3 Resultados experimentales y Discusión

Las longitudes de onda fueron obtenidas mediante un espectrofotoradiómetro FieldSpec Pro JR con un rango de 350 a 2500 nm. El espectrofotoradiómetro se aplicó sobre ambos lados (veteado y graso) del músculo masetero izquierdo y derecho de la carrillada, obteniendo por tanto cuatro medidas para cada cerdo. Dado el alto coste de la carne de cerdo ibérico, solo se dispuso de 15 ejemplares de estos, haciendo las mismas medidas sobre otros 15 ejemplares de cerdo blanco.

En primer lugar, se realizó un ranking de las longitudes de onda de modo que se pudieran identificar las más importantes. Tras esto, se disearon los distintos modelos clasificadores utilizando la variable más relevante según el ranking, luego las dos más relevantes,

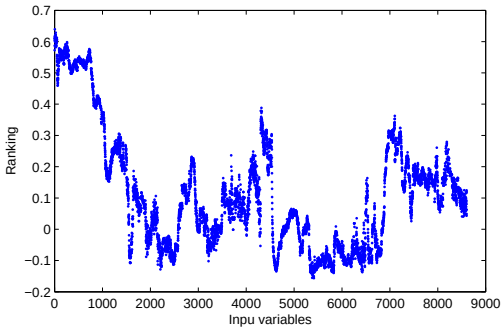


Figura 2: Ranking de las variables medidas.

y así sucesivamente. En la Figura 2 se muestran los resultados de los valores obtenidos al calcular la información mutua entre una longitud de onda y los valores de salida. El eje X en la Figura se corresponde con la concatenación de los 4 espectros medidos para cada cerdo en el siguiente orden: carrillada derecha lado magro, carrillada derecha lado veteado, carrillada izquierda lado magro, carrillada izquierda lado veteado. De modo que, por ejemplo, la variable 3101 se corresponde con la reflectancia obtenida para la longitud de onda de 1050nm del lado veteado de la carrillada derecha. Como muestra la Figura 2, las longitudes de onda más pequeas del lado magro de las carrilladas muestran más diferencias entre las dos especies.

Dado el reducido número de muestras (30), los clasificadores fueron entrenados utilizando todos los conjuntos de datos dejando para test 2 muestras. Esto significa que se emplearon 435 conjuntos de test compuestos por 28 muestras cada uno. En la Tabla 1 se muestran los resultados obtenidos para las RBFNNs y las SVMs. Los datos en las columnas muestran el número total de clasificaciones erróneas realizadas a lo largo de toda la batería de experimentos y el porcentaje final de aciertos.

Las RBFNNs generadas con el algoritmo FPCFA obtuvieron mejores resultados que aquellas generadas con el FCM y que las SVM. Esto es debido a que el FPCFA incorpora un mecanismo de supervisión que le permite centrarse en la clasificación de una especie concreta, mejorando de este modo los resultados. Conforme se aumenta el número de variables que se utilizan con entrada en los clasificadores, los resultados de clasificación empeoran. Esto es bastante razonable puesto que las RBFNN son modelos locales y conforme aumenta el número de variables, aumenta exponencialmente el espacio que deben cubrir y no hay sucifientes muestras que permitan llevar a cabo un entrenamiento adecuado. Las SVMs también proporcionaron unos resultados bastante satisfactorios superando los resulta-

dos obtenidos por las RBFNN que fueron generadas usando el algoritmo FCM.

Tabla 1: Clasificaciones erróneas y porcentaje de aciertos de los distintos modelos utilizando diferentes números de neuronas.

#Variables	RBFNN(FCM)	SVM	RBFNN(FPCFA)
1	15 - 97.8%	8 - 98.8%	2 - 99.7%
2	9 - 98.6%	3 - 99.6%	2 - 99.7%
3	6 - 99.3%	2 - 99.7%	3 - 99.6%

4 Conclusiones

Este trabajo ha presentado la combinación de diversas técnicas de soft computing con el objetivo de ser capaces de clasificar una muestra de carne en dos tipos: cerdo ibérico o cerdo blanco. Tal y como muestran los resultados, las técnicas aplicadas son capaces de clasificar con bastante precisión los distintos tipos de especímenes, animando la investigación en el diseo de arquitecturas específicas que permitan realizar la clasificación de forma automática al medir el espectro infrarrojo. El empleo de la teoría de la información mutua para determinar qué longitudes de onda son las más significativas se ha mostrado muy efectivo de modo que los modelos empleados para la clasificación, SVMs y RBFNNs, han sido capaces de realizar una clasificación precisa.

Este trabajo ha sido financiado por el Poyecto del CICYT TIN2007-60587 y el Proyecto P07-TIC-02768 de la Junta de Andalucia.

Referencias

[1] D. S. Broomhead and D. Lowe. Multivariate functional interpolation and adaptive networks. *Complex Systems*, 2:321–355, 1988.

[2] B.V. Bonnlander and A.S. Weigend. Selecting input variables using mutual information and nonparametric density estimation. In *Proc. of the ISANN*, Taiwan, 2004.

[3] D. Cozzolino and I. Murray. Identification of animal meat muscles by visible and near infrared reflectance spectroscopy. *Lebensmittel-Wissenschaft und-Technologie*, 37:447–452, 2004.

[4] A. Guillén, I. Rojas, J. González, H. Pomares, L. J. Herrera, and F. Fernández. Multiobjective RBFNNs Designer for Function Approximation: An Application for Mineral Reduction. *Lecture Notes in Computer Science*, 4221:511–520, 2006.

[5] A. Guillén, I. Rojas, J. González, H. Pomares, L.J. Herrera, O. Valenzuela, and A. Prieto. A Possibilistic Approach to RBFN Centers Initialization. *Lecture Notes in Computer Science*, 3642:174–183, 2005.

[6] A. Guillén, I. Rojas, J. González, H. Pomares, L.J. Herrera, O. Valenzuela, and A. Prieto. Improving Clustering Technique for Functional Approximation Problem Using Fuzzy Logic: ICFA algorithm. *Lecture Notes in Computer Science*, 3512:272–280, June 2005.

[7] A. Guillén, I. Rojas, E. Ros, and L.J. Herrera. Using Fuzzy Clustering Technique for Function Approximation to Approximate ECG Signals. *Lecture Notes in Computer Science*, 3562:538–547, June 2005.

[8] A. Kraskov, H. Stgbauer, and P. Grassberger. Estimating mutual information. *Physics Review*, June 2004.

[9] V. Ortiz-Somovilla, F. Espaa-Espaa, A.J. Gaitán-Jurado, J. Prez-Aparicio, and E.J. De Pedro-Sanz. Proximate analysis of homogenized and minced mass of pork sausages by NIRS. *Food Chemistry*, 101:1031–1040, 2007.

[10] V. Ortiz-Somovilla, F. Espaa-Espaa, E.J. De Pedro-Sanz, and A.J. Gaitán-Jurado. Meat mixture detection in Iberian pork sausages. *Meat Science*, 71:490–497, 2005.

[11] H.J. Swatland. Physical measurements of meat quality: optical measurements, pros and cons. *Meat Science*, 36:251–259, 1994.

[12] Z. Uykan, C. Gzelis, M. E. Celebei, and H. N. Koivo. Analysis of Input–Output Clustering for Determining Centers of RBFN. *IEEE Transactions on Neural Networks*, 11(4):851–858, July 2000.

[13] J.J. Xia, E.P. Berg, J.W. Lee, and G. Yao. Characterizing beef muscles with optical scattering and absorption coefficients in VIS-NIR region. *Meat Science*, 75:78–83, 2007.

MODELADO RECURSIVO DE SERIES TEMPORALES UTILIZANDO MODELOS TSK AVANZADOS

L.J. Herrera[1] **H. Pomares**[1] **I. Rojas**[1] **A. Guillén**[2]
G. Rubio[1]

[1] Depto Arquitectura y Tecnología de Computadores, Universidad de Granada
[2] Depto Informática, Universidad de Jaén

Resumen

Los problemas de predicción de series temporales se pueden ver como un problema de aproximación funcional modificado utilizando datos de entrada/salida, en el que las entradas se obtienen utilizando valores de la serie en momentos anteriores. Así, el modelo obtenido normalmente predice el valor de la serie en el momento $(t + h)$ utilizando momentos de tiempo anteriores $(t - \tau_1), (t - \tau_2), \ldots, (t - \tau_n)$. Sin embargo, el aprendizaje de un modelo para predicción de series temporales a largo plazo puede verse como un problema mucho más complicado, puesto que puede utilizar sus propias salidas como valores de entrada para predecir valores sucesivos (predicción recursiva). El trabajo presentado en este trabajo muestra la aptitud del modelo neuro-difuso TSK interpretable TaSe para resolver el problema de predicción de series temporales a largo plazo utilizando predicción recursiva. Este trabajo además introduce algunas técnicas que incrementan el rendimiento de este tipo de modelo avanzado de predicción a un paso (y en general de cualquier modelo de predicción a un paso), cuando se utiliza recursivamente para predicción de series temporales a largo plazo.

1 INTRODUCCIóN

Hay un gran número de metodologías que tratan el problema de la predicción de series temporales [12], [7]. Las Redes Neuronales Artificiales, Sistemas Difusos, Máquinas de Vectores Soporte, Modelos AR-ARMA, etc. están dentro de los paradigmas que se han aplicado a este problema. En el caso de la predicción de series temporales a largo plazo, se pueden tomar dos tendencias a la hora de predecir los datos: predicción directa y predicción recursiva [13]. La predicción directa implica la construcción de modelos distintos para cada horizonte de predicción que se necesite. Por otra parte, la predicción recursiva solo utiliza un modelo para predecir todos los horizontes necesarios, pero con el inconveniente de que el error cometido en los horizontes cercanos se transmite a los horizontes más lejanos, al usarse las mismas salidas del modelo como entradas para predecir los siguientes datos desconocidos. Este trabajo se centra en la predicción recursiva, mostrando cómo se puede usar un modelo TSK para construir el modelo recursivo, y presentando una técnica general que mejora la precisión en la predicción para este tipo de problemas.

La metodología TSK neuro-difusa concreta que se utiliza en este trabajo es el modelo TaSe [6]. Ésta es una metodología de aprendizaje difusa avanzada que proporciona un modelo interpretable de aproximación dado un conjunto de datos de entrada y salida. Este modelo elimina, al menos parcialmente, los dos inconvenientes principales del entrenamiento de sistemas difusos tipo TSK basados en grid: la maldición de la dimensionalidad y la pérdida de interpretabilidad.

El resto del trabajo se organiza como sigue: la sección 2 describe brevemente el modelo TaSe y una aplicación a un modelo de predicción de series a un paso. La sección 3 presenta brevemente un punto de vista modificado de la predicción de series temporales a largo plazo. La sección 4 presenta una metodología de aprendizaje modificada, mostrando su aplicación para el modelo TaSe, para mejorar la predicción recursiva de series temporales a largo plazo. Finalmente la sección 5 finaliza el trabajo.

2 APRENDIZAJE DE UN MODELO TaSe RECURSIVO PARA UNA SERIE TEMPORAL

Un sistema difuso TSK consta de un conjunto de K reglas IF-THEN que típicamente tienen la forma [11]

$$R^k : IF\ x_1\ es\ A_1^k\ y\ \ldots\ y\ x_n\ es\ A_n^k \quad THENy = Y^k(\vec{x}) \tag{1}$$

donde los A_i^k son conjuntos difusos caracterizados por funciones de pertenencia $A_i^k(x_i)$, x_i son las variables de entrada y $Y^k(\vec{x})$ es un polinomio en cualquier orden función de las variables de entrada.

Los modelos difusos TSK con particionamiento basado en grid presentan dos problemas principales a la hora de tratar con problemas de modelado de datos: la maldición de la dimensionalidad, que provoca un crecimiento exponencial en el número de reglas [1] y la pérdida de la interpretabilidad de los modelos locales en el entrenamiento global del modelo.

El modelo difuso TSK TaSe (basado en la Serie de Taylor) intenta minimizar estos problemas gracias a la utilización de [6]:

- consecuentes de orden alto que permiten reducir el incremento exponencial en el número de reglas necesarias para alcanzar la precisión dada y
- una estructura de los consecuentes y de los antecedentes que permite entrenar a los modelos locales (reglas), dándoles la expresión del desarrollo en serie de Taylor de la salida de la función en torno al centro de cada regla, a la vez que se optimiza el modelo global según el enfoque de mínimos cuadrados (LSE).

Esta última propiedad es posible gracias a dos características de la estructura del modelo TaSe:

1) Los consecuentes de las reglas tienen una forma general de polinomio centrado en un punto que admite la serie de Taylor. El teorema de Taylor establece que si una función continua $f(\vec{x})$ definida en un intervalo $(c-r, c+r)$ es l veces derivable, ésta se puede aproximar en torno al punto $\vec{x} = \vec{c}$ como su expansión en serie de Taylor

$$f(\vec{x}) \simeq f(\vec{c}) + (\vec{x}-\vec{c})^T [\nabla(\vec{c})]_{i=1\ldots n} + \frac{1}{2}(\vec{x}-\vec{c})^T W(\vec{x}-\vec{c}) + \ldots \tag{2}$$

donde W es una matriz triangular de dimensiones $n \times n$ con las derivadas parciales de segundo orden. La serie de Taylor abre una puerta a la aproximación de cualquier función a través de polinomios, esto es, a través de la adición de un cierto número de funciones sencillas. Utilizando la formulación general del modelo TSK, podemos utilizar consecuentes $Y^k(\vec{x})$ que tienen la forma centrada simulando la ecuación 2.

2) Los antecedentes de las reglas siguen una estructura que permite dicha interpretabilidad. Esto es una condición doble. Primero, la salida del modelo debe ser continua y derivable, de forma que el teorema de Taylor se pueda aplicar (ver punto anterior). Segundo, el grado de solapamiento de todas las reglas debe desvanecerse en cada centro de regla. De esta forma conseguimos que cada que cada punto $\vec{c}^k$ en el espacio n-dimensional identificado por el centro de una regla k, quede afectado tan solo por el consecuente $Y^k(\vec{x})$ de su regla respectiva en la salida global del sistema. Además, los coeficientes del polinomio $Y^k(\vec{x})$ contendrán la información del valor de $f(\vec{x})$ en $\vec{c}^k$, así como su comportamiento en torno a dicho punto (sus derivadas).

Esto se consigue en los sistemas basados en grid mediante la utilización de un tipo especial de función de pertenencia que conserve las propiedades de continuidad y derivabilidad de la salida. Una base de funciones de pertenencia locales ordenadas de orden ν (OLMF de orden ν), resumidamente es un conjunto de funciones de pertenencia (F.P.) locales que recorren un dominio de una variable tales que, el centro de una F.P. coincide con los extremos de las F.P. adyacentes y todas las F.P. que forman la base son continuas y ν-veces derivables y tienen ν-derivada igual a cero en sus centros y extremos. Al utilizar bases de funciones OLMF de orden ν, obtenemos reglas difusas cuyos consecuentes (de hasta orden ν) son la expansión en serie de Taylor de la salida del modelo en torno a los centros de las reglas [2], [6].

El modelo TaSe es un sistema difuso TSK modificado con un tipo específico de antecedentes de reglas y consecuentes de reglas basados en la serie de Taylor. Este sistema está por tanto dotado de ambas capacidades, la capacidad aproximativa de las reglas TSK y las ventajas de interpretabilidad de los sistemas difusos tradicionales, puesto que cada consecuente puede verse como la expansión en serie de Taylor en torno a los centros de las reglas. En el trabajo [6] se proporciona también una metodología automática para realizar tanto la estimación de los parámetros como la identificación de la estructura del sistema difuso TaSe a partir de un conjunto de datos de entrada/salida. Ahora presentamos la aplicación del modelo TaSe a un problema de predicción de series a un paso, antes de pasar a ver la predicción a largo plazo utilizando modelos recursivos.

Metodo		RMSE de Test
Modelo Auto Regresivo		0.19
RNN de Correlación en Cascada		0.06
RNN Back-Prop.		0.02
Polinomio 6 orden		0.04
Método predictivo lineal		0.55
Sistema difuso y algoritmo genético)	5FP	0.0492
	7FP	0.0422
	9 FP	0.037
ANFIS		0.007
Red RBF clásica		0.0114
PG-RBF		0.0030
Modelo TaSe con 27 reglas		0.0013
GBFS tradicional con 192 reglas		0.0041
Red neuro-difusa 9 reglas		0.0239

Tabla 1: Comparación del error de aproximación para diferentes métodos para la serie temporal Mackey-Glass con paso de predicción igual a 6. Ver [10] y tablas 4.3 y 6.3

2.1 Aplicación a la predicción a un paso de una serie temporal

La serie temporal caótica Mackey-Glass [8] es un benchmark muy conocido para el modelado de sistemas, que ha sido ampliamente utilizado en la bibliografía de este tipo de problemas. Para realizar comparaciones justas con trabajos previos, se han escogido los parámetros de forma que los vectores de entrada/salida para el modelo tienen el formato:

$$[x\,(t-18)\,, x\,(t-12)\,, x\,(t-6)\,, x\,(t)\,; x\,(t+6)] \quad (3)$$

Se generaron 1000 muestras para este estudio, de las que se han seleccionado 500 para entrenamiento y validación (escogiendo aleatoriamente 400 y 100 respectivamente), y las últimas 500 para test.

El modelo TaSe óptimo obtenido por el algoritmo de aprendizaje [6] obtiene un modelo que con tan solo 27 reglas obtiene resultados de predicción muy buenos en comparación con otros paradigmas presentes en la literatura que también trataron este problema. Lo resultados muestran para este ejemplo la conveniencia del modelo TaSe, no solo debido a sus propiedades de interpretabilidad y transparencia y a la interacción entre el modelo global y los modelos locales, sino también debido a su alto grado de precisión para problemas de aproximación funcional y predicción de series temporales [6].

3 Predicción a Largo Plazo

Como se ha mencionado, hay dos tendencias principales al tratar predicción de series temporales a largo plazo: predicción directa y predicción recursiva [13]. Debido a las características de cada enfoque, puede esperarse que la predicción directa proporcione un mejor rendimiento en precisión de la predicción, puesto que utiliza un modelo específico para cada horizonte deseado. Sin embargo, tiene dos inconvenientes principales: por un lado la comprensión del comportamiento de la serie temporal se pierde cuando se utilizan diferentes modelos para cada horizonte de predicción. Por otro lado, para horizontes de predicción a muy largo plazo, como ocurre en algunos casos [4], un enfoque de predicción directa aparentemente es inútil al requerirse un número excesivamente grande de modelos. Aún para horizontes largos, cuando el ruido de la serie es considerable, un enfoque de predicción directa puede ser deseable para obtener mejores predicciones. Pero para niveles de ruido aceptables, la predicción recursiva puede obtener buenas predicciones utilizando un solo modelo, que es globalmente más comprensible.

3.1 Predicción recursiva

Ahora el problema de la predicción recursiva se estudia más en profundidad. Para problemas de predicción a largo plazo, el objetivo podría ser expresado como: "teniendo una secuencia cortada de datos, intentar predecir el resto de la secuencia hasta que se alcance un cierto momento". Así, los datos de entrenamiento para este modelo serían los de la ecuación 3, pero en realidad, la aplicación del modelo para predecir la serie completa debería utilizar iterativamente los datos de entrenamiento

$$\hat{x}\,(t+6) = f\,(\hat{x}\,(t-18)\,,\ \hat{x}\,(t-12)\,,\ \hat{x}\,(t-6)\,,\ \hat{x}\,(t)) \quad (4)$$

y en principio, esto no es para lo que se ha entrenado nuestro modelo inicial.

Supóngase un ejemplo sencillo de serie temporal $x(t)_{t=1,\ldots,4} = \{1,\ 2.5,\ 3.5,\ 4\}$. Un modelo que intenta predecir $\hat{x}(t+1) = f(x(t))$ utilizaría como entrenamiento los pares de datos [1, 2.5], [2.5, 3.5], [3.5,4]. Así, utilizando mínimos cuadrados, un modelo lineal sencillo que utiliza un solo parámetro constante $\hat{x}(t+1) = a * x(t)$ resolvería la ecuación

$$[1, 2.5, 3.5]' * a = [2.5, 3.5, 4]'. \quad (5)$$

La solución es $a = 1.29$ y el MSE $= 0.6$. Pero de hecho, el error de predicción a largo plazo, comenzando con

el valor inicial $x(t_0) = 1$ y prediciendo los siguientes valores utilizando la salida del modelo recursivamente es MSE = 2.72 (la predicción "a largo plazo" sería [1.29, 1.29^2, 1.29^3]). Así pues, la manera correcta de entrenar el modelo recursivo a largo plazo sería

$$\left[1, 1 * a, 1 * a^2\right]' * a = [2.5, 3.5, 4]' \qquad (6)$$

cuya solución es $a = 1.67$ y el error de predicción MSE a "largo plazo" es 0.54. No obstante, incluso para este ejemplo sencillo, esto es un problema de optimización no lineal, con una complejidad muy alta, y por tanto las técnicas de modelado de aproximación funcional tradicionales no funcionan aquí.

4 Mejorando la Predicción a Largo Plazo Utilizando Modelos de Predicción a Un Paso

En la sección anterior se ha visto en un ejemplo sencillo, la tarea recursiva tan compleja que sería modelar directamente un problema de predicción a largo plazo a partir de un conjunto de entrada/salida. Como se ha visto en [13], en principio el modelo de predicción a un paso puede utilizarse directamente para realizar dicha tarea. Esta sección presenta una metodología general para mejorar los predictores a un paso cuando se utilizan como predictores recursivos.

4.1 Utilizando el modelo difuso TSK TaSe

Considérese la serie temporal Mackey-Glass, y el modelo TaSe óptimo obtenido por el algoritmo de entrenamiento [6]. Partamos de que utilizamos este modelo predictor a un paso para realizar predicciones a largo plazo sobre la serie temporal Mackey-Glass. Suponiendo que se conoce una parte inicial de la serie, y que el objetivo es predecir el resto utilizando este modelo (al menos se necesitarían un total de 23 (18+5) ejemplos como parte conocida inicial de la serie para predecir el resto de 500 valores). Así, a partir de este trozo inicial, las salidas estimadas se usarían iterativamente como entradas nuevas (ver ecuación 4). La figura 1 muestra la predicción a largo plazo (ecuación 4), con $\hat{x}(t) = x(t)$ para los primeros valores $t = 1, \ldots, 23$, para el conjunto de test. El RMSE obtenido para el conjunto de entrenamiento es igual a 0.030, y para el conjunto de test es 0.031. Estos resultados muestran que se obtiene una solución bastante buena para este problema de predicción a largo plazo, utilizando recursivamente un predictor a un paso. Gracias a la precisión del modelo TaSe, se obtiene una secuencia de salidas muy similar a los datos originales.

Sin embargo, se presenta ahora una modificación en el proceso de entrenamiento que permitirá obtener mejores resultados. En [3], Schoelkopf y otros añadieron pequeñas modificaciones en cada muestra de entrenamiento, expandiendo el conjunto de entrenamiento para un problema de clasificación de dígitos escritos a mano. De forma similar, el conjunto de datos de entrenamiento de la serie temporal para este caso podría expandirse de forma que pequeñas modificaciones de la serie se considerasen también en el entrenamiento. Sería posible por ejemplo, añadir un nuevo conjunto de ejemplos obtenidos a partir de la serie original, calculando la serie $x(t + 0.5)$. Sin embargo, éste es un problema de interpolación y esta tarea puede ser en algunos casos tan difícil o más como el problema de predicción original. Con todo, para la serie Mackey-Glass, hemos obtenido una muy buena aproximación a un paso para la serie (RMSE = 0.0011). Así, utilizando el mismo modelo original, el conjunto de datos de entrenamiento se ha expandido para considerar una nueva sección de la serie temporal, que es una pequeña modificación de la original y que, de hecho, presenta el mismo comportamiento. El conjunto expandido de muestras de entrenamiento consta ahora por tanto de la serie original $x(t), t = 1, \ldots, 500$ más un nuevo trozo ficticio de serie $\hat{x}, t = 1, \ldots, 478$, donde el $\hat{x}$ ha sido obtenido utilizando el modelo predictivo TaSe.

Para el reentrenamiento del modelo TaSe utilizando el conjunto de entrenamiento extendido, nótese que tan solo se han actualizado los coeficientes en los consecuentes de las reglas. La estructura y parámetros en los antecedentes se quedan como están, con el objetivo de evitar un coste computacional excesivo y para conservar la simplicidad del enfoque propuesto. La figura 2 muestra la predicción a largo plazo del conjunto de test, obtenida con el modelo TaSe reentrenado utilizando el conjunto de entrenamiento extendido. El RMSE obtenido para la predicción a largo plazo del conjunto de entrenamiento es de 0.015 y para el conjunto de test es 0.019. Intentos sucesivos de reextender el conjunto de entrenamiento, repitiendo el mismo proceso (reextendiendo el conjunto de entrenamiento, utilizando la salida predicha a un paso para el modelo reentrenado) no produjo mejoras significativas para la predicción a largo plazo para este ejemplo.

Los resultados obtenidos muestran que este proceso puede verse como un proceso de eliminación de ruido en una serie temporal. Si se encuentra un modelo a un paso suficientemente preciso, y que es capaz de eliminar parcialmente el ruido de la serie, la salida de dicho modelo puede utilizarse como nuevos datos de entrenamiento, con el objetivo de mejorar los resultados de predicción recursiva a largo plazo [5].

Como segundo ejemplo, considérese ahora una serie

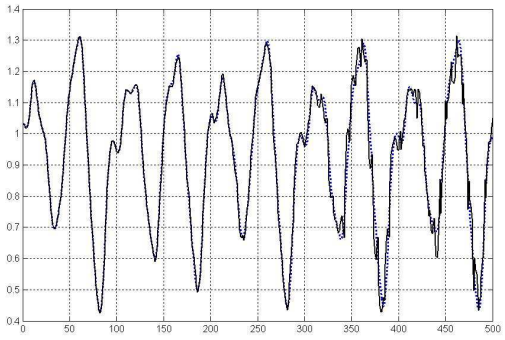


Figura 1: Predicción a largo plazo original (línea punteada) y aproximada (línea sólida) para el conjunto de test utilizando el modelo TaSe. Error de test RMSE = 0.031.

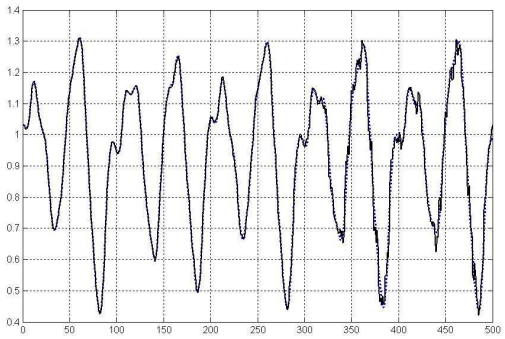


Figura 2: Predicción a largo plazo original (línea punteada) y aproximada (línea sólida) para el conjunto de test utilizando el conjunto de entrenamiento expandido. RMSE de test = 0.019

temporal distinta, el flujo mensual en pies cúbicos por segundo del río Snake, tomada de [14]. Nótese cómo para este caso, la serie presenta un alto nivel de ruido, aunque con un patrón bien reconocible en su comportamiento.

De nuevo para esta serie, se tomaron 1000 muestras, 500 para entrenamiento y validación (tomando aleatoriamente 400 y 100 respectivamente) y 500 para test. En este caso los valores tomados por $h, \tau_0, \ldots, \tau_n$ en la ecuación $\hat{x}(t+h) = f(x(t-\tau_0), x(t-\tau_1), \ldots, x(t-\tau_n))$ son [3, 0, 3, 6, 9] (ver la referencia [9] sobre cómo encontrar valores adecuados para estos parámetros dada una serie temporal). El modelo TaSe óptimo para este problema de predicción a un paso tiene tan solo dos reglas con RMSE de entrenamiento = 878 y RMSE de test = 890. Nótese que puesto que la misma medida del error se ha utilizado para este ejemplo (la raíz del error cuadrático medio, RMSE), los errores en esta serie ruidosa no normalizada son mucho mayores que para el ejemplo anterior. La predicción a largo plazo para los conjuntos de entrenamiento y test, proporcionan un RMSE total igual a 1769 para el conjunto de entrenamiento y 1235 par el conjunto de test. La figura 3 muestra la predicción a largo plazo para el flujo mensual del río Snake, utilizando el predictor TaSe a un paso. Puede observarse que la capacidad de predicción a largo plazo del modelo a un paso se desvanece en un rango corto de tiempo.

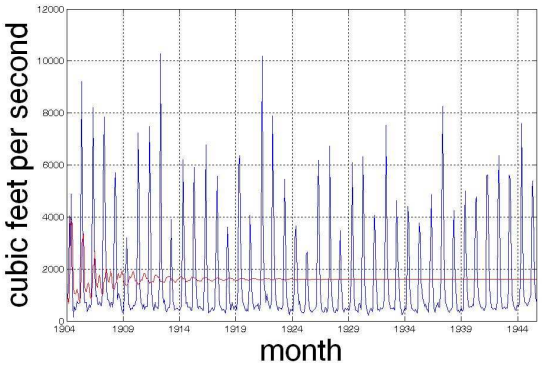


Figura 3: Predicción a largo plazo original (línea punteada) y aproximada (línea sólida) para el conjunto de entrenamiento. RMSE de entrenamiento = 1769

De nuevo para este ejemplo de serie temporal, los datos de entrenamiento se ampliaron con la salida del predictor TaSe a un paso, considerado como un nuevo trozo de la serie temporal. Este trozo ficticio se añadió a los datos de entrenamiento, conservando la estructura óptima anterior del modelo TaSe y tan solo reobteniendo de nuevo los consecuentes óptimos de las reglas. Tras una iteración, la predicción a largo plazo mejoró hasta un RMSE de entrenamiento de 1618 y 1060 de test. La figura 4 muestra la predicción a largo plazo para los conjuntos de test y entrenamiento, utilizando el conjunto de datos extendido de entrenamiento para la serie del flujo del río Snake. Se observa claramente la mejora en la predicción que se obtiene en la predicción de la serie a largo plazo.

5 Conclusiones

Este trabajo presenta una técnica general para mejorar los resultados de predicción recursiva para predicción de series temporales a largo plazo utilizando el modelo TSK TaSe. Para introducir y comprobar la eficacia del método, se ha aplicado a dos ejemplos de series temporales, la serie artificial bien conocida de Mackey-Glass y la serie del flujo de un río. En general, cuando se realiza una predicción a largo plazo de

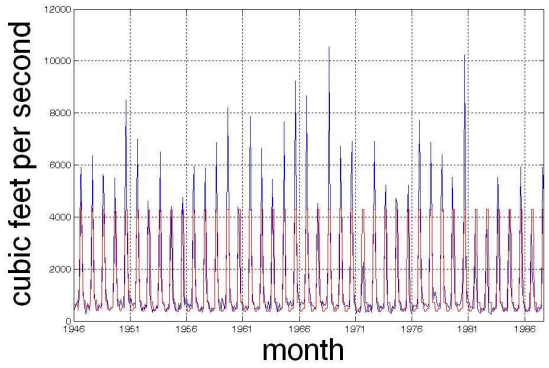


Figura 4: Original (dotted line) and approximated (solid line) long term prediction for the test data set using the expanded training data set for the river flow problem. Test RMSE = 1060

una serie temporal dada, con un modelo de predicción cualquiera, se obtiene una copia sin ruido de la serie original. Las simulaciones presentadas han mostrado cómo el rendimiento de un predictor a un paso puede mejorarse cuando se utiliza dicha salida para realizar predicción recursiva a largo plazo. Esta mejora se obtiene utilizando la predicción a un paso sin ruido como una nueva parte ficticia de la serie temporal, que se añade a los datos de entrenamiento, y reentrenando el modelo con este conjunto de datos de entrenamiento extendido. En las simulaciones, en ambos casos ha mejorado su predicción recursiva a largo plazo para ambas series temporales.

Agradecimientos

Este trabajo ha sido financiado por el Proyecto del CICYT TIN2007-60587 y el Proyecto P07-TIC-02768 de la Junta de Andalucía.

Referencias

[1] Bellman, R., Adaptive Control Processes: A Guided Tour,(Princeton University Press, 1961)

[2] Bikdash, M.: A Highly Interpretable Form of Sugeno Inference Systems. IEEE Trans. Fuz. Syst.. **7(6)** (1999) 686–696

[3] Decoste, D., Schoelkopf, B.: Training Invariant Support Vector Machines. Machine Learning, **46**, (2002) 161–190

[4] EUNITE Competition 2003: Prediction of product quality in glass manufacturing, www.eunite.org (2003)

[5] L.J. Herrera and H. Pomares, I. Rojas, A. Guillén, A. Prieto, O. Valenzuela, "Recursive Prediction for Long Term Time Series Forecasting Using Advanced Models", Neurocomputing, 70 (2007) 2870–2880

[6] Herrera, L.J., Pomares, H., Rojas, I., Valenzuela, O., Prieto, A.: TaSe, a Taylor Series Based Fuzzy System Model that Combines Interpretability and Accuracy. Fuzzy Sets and Systems, **153** (2005) 403–427

[7] A. Lendasse and E. Oja and O. Simula and M. Verleysen: "Time Series Prediction Competition: The CATS Benchmark", Proceedings of the IEEE International Joint Conference on Neural Networks, (2004) 1615–1620

[8] Mackey, M.C., Glass,L.: Oscillation and chaos in physcological control systems. Science. **197** (1977) 287–289

[9] McNames, J., Suykens, J.A.K., Vandewalle, J.: Winning Entry of the K. U. Leuven Time Series Prediction Competition

[10] Rojas, I., Pomares, H., Bernier, J.L., Ortega, J., Pino, B., Pelayo, F.J., Prieto, A.: Time Series analysis using normalized PG-RBF network with regression weights. NeuroComputing, **42**, (2002) 267–285

[11] Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modelling and control. IEEE Trans. Syst. Man and Cyber. **15** (1985) 116–132

[12] Weigend, AS., Gershenfeld, NA.: Time Series Prediction: Forecasting the Future and Understanding the Past, Addison-Wesley (1993)

[13] Yongnan Ji, Jin Hao, Nima Reyhani, Amaury Lendasse: Direct and Recursive Prediction of Time Series Using Mutual Information Selection. IWANN 2005, Lecture Notes in Computer Science. New York: Springer, **3512** (2005) 1010–1017

[14] http://www-personal.buseco.monash.edu.au/ hyndman/TSDL/hydrology.html

UN ALGORITMO PARA LA OBTENCIÓN DE LOS EXTREMOS DE LAS MEDIDAS k-ADITIVAS

Pedro Miranda[1] **Elías Fernández-Combarro**[2]

[1] Departamento de Estadística e Investigación Operativa, Universidad Complutense de Madrid, pmiranda@mat.ucm.es
[2] Centro de Inteligencia Artificial, Universidad de Oviedo, elias@aic.uniovi.es

Resumen

El conjunto de medidas k-aditivas es un politopo cuya estructura es muy compleja. En este trabajo se propone un algoritmo de obtención de los vértices de este politopo. El algoritmo se basa en añadir sucesivamente las restricciones adicionales que tiene la familia de medidas k-aditivas como subpolitopo del de las medidas difusas generales, obteniendo en cada paso los vértices del politopo intermedio.

Palabras Clave: Politopos, vértices, medidas difusas, medidas k-aditivas, adyacencia.

1 INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA

Las medidas difusas (también conocidas como capacidades o medidas no aditivas) son una generalización de las distribuciones de probabilidad. Más concretamente, son medidas en las que el axioma de aditividad se ha relajado a una condición de monotonía. Esta extensión es necesaria en muchas situaciones prácticas, en las que la aditividad es demasiado restrictiva.

Así, en el marco de la Teoría de la Decisión, los modelos axiomáticos basados en probabilidades, como por ejemplo los de von Neumann y Morgenstern [33] o Anscombe y Aumann [2], pueden dar lugar a resultados inconsistentes con las preferencias del decisor debido a una *aversión al riesgo* del mismo, como ocurre en las paradojas de Ellsberg [9] y Allais [1]. Sin embargo, los modelos basados en medidas difusas [3, 29] son capaces de interpretar y modelar estas situaciones. Por otra parte, el análisis y uso de las medidas difusas se ha visto impulsado en los últimos años con varias representaciones equivalentes de una medida difusa [12], que se obtienen a partir de transformaciones lineales invertibles aplicadas sobre la medida.

Las medidas difusas también han sido aplicadas con éxito tanto en Teoría de Decisión Multicriterio como en Teoría de Juegos. En el primer caso, las medidas difusas son capaces de introducir vetos y favores en el sistema de evaluación [12], así como interacciones entre los distintos criterios [13]. En la Teoría de Juegos Cooperativos, las medidas difusas representan la fuerza de las distintas coaliciones de jugadores; además están relacionadas con el valor de Shapley [30], como se ha demostrado en [11]. Otros campos en los que se han aplicado las medidas difusas son la Combinatoria [28], las funciones Pseudo-Booleanas [15], etc.

Sin embargo, a pesar de todas estas ventajas, la aplicación en la práctica de las medidas difusas queda limitada por el incremento en la complejidad de la medida. En el caso de un espacio finito de cardinal n, sólo necesitamos conocer $n-1$ valores para determinar completamente una probabilidad, mientras que para definir una medida difusa sobre el mismo referencial se necesitan 2^n-2 coeficientes. Este incremento exponencial de la complejidad es el verdadero *talón de Aquiles* de las medidas difusas.

Para tratar de reducir la complejidad de las medidas difusas se han definido una serie de subfamilias en las que se añade una restricción adicional que permita reducir el número de coeficientes pero que al mismo tiempo mantiene una gran riqueza en términos de modelización. Ejemplos de subfamilias son las λ-medidas [32], las medidas k-intolerantes [17, 18], las p-simétricas [24], las descomponibles [7], etc. Nosotros trataremos en este trabajo con las medidas k-aditivas definidas por Grabisch en [11].

Consideremos una situación que puede modelarse mediante una medida k-aditiva (posiblemente a través de la integral de Choquet [4]); en este sentido, una caracterización axiomática de un modelo de este tipo puede

verse en [23]. El siguiente paso ha de ser la identificación de la correspondiente medida.

Para ello, supondremos que tenemos una información muestral disponible. Usando la terminología de Teoría de la Decisión, esta información consiste en una serie de objetos para los que conocemos su evaluación en cada criterio, así como su evaluación global. Asumiremos que esta información es de tipo numérico; en caso contrario, debería aplicarse previamente alguna de las herramientas que permiten transformar datos ordinales en numéricos como los métodos MACBETH [8] o TOMASO [19]. La medida (que tal vez no es única) que buscamos será aquélla que se ajuste mejor a estos datos.

Si el criterio que se toma es el de mínimos cuadrados, existen varias técnicas para resolver el problema [14, 34]. En un trabajo anterior [5], hemos propuesto un algoritmo de aprendizaje basado en algoritmos genéticos [16]; en dicho algoritmo, el operador de cruce utilizado era la combinación convexa. Esto era posible porque el conjunto de medidas k-aditivas es un poliedro convexo; este operador resulta ser lógico y sencillo, y permite reducir el gasto computacional. Sin embargo, tiene el problema de que la región de búsqueda se reduce en cada iteración. Esto se resuelve mediante un operador de mutación, pero aún así la región de búsqueda queda determinada por la población inicial. Por lo tanto, es necesario que esta población inicial sea lo suficientemente amplia como para que contenga la medida que se busca, y obviamente, la mejor opción es tomar como población inicial el conjunto de vértices del poliedro.

El objetivo de este trabajo es dar un algoritmo para determinar estos puntos extremos.

2 MEDIDAS k-ADITIVAS

Sea $X = \{1, ..., n\}$. Denotaremos los subconjuntos de X por $A, B, ...$ El conjunto de partes de X se denotará por $\mathcal{P}(X)$.

Definición 1 *[31] Una* **medida difusa** *sobre $(X, \mathcal{P}(X))$ es una aplicación $\mu : \mathcal{P}(X) \to [0, 1]$ que verifica*

- $\mu(\emptyset) = 0$, $\mu(X) = 1$ *(condiciones de frontera).*
- $\forall A, B \in \mathcal{P}(X)$, *si* $A \subseteq B$, *entonces* $\mu(A) \leq \mu(B)$ *(monotonía).*

Denotaremos al conjunto de todas las medidas difusas sobre X por $\mathcal{FM}(X)$. Nótese que $\mathcal{FM}(X)$ es un politopo convexo definido por las distintas restricciones de monotonía y, por tanto, queda determinado por sus vértices.

Definición 2 *Una medida difusa es* $\{0, 1\}$**-valuada** *si sólo toma los valores 0 y 1.*

Nótese que para una medida $\{0, 1\}$-valuada μ, existen subconjuntos A que satisfacen la siguiente condición:

$$\begin{array}{l} \mu(A) = 1, \\ \mu(B) = 1, \quad \forall B \supseteq A, \\ \mu(C) = 0, \quad \forall C \subset A. \end{array} \tag{1}$$

Un subconjunto en estas condiciones se llama **subconjunto minimal** para μ. Es fácil demostrar que una medida $\{0, 1\}$-valuada queda completamente definida a través de sus conjuntos minimales. Si una medida difusa tiene como familia de subconjuntos minimales la familia $\mathbf{C}$, la denotaremos $\mu_{\mathbf{C}}$.

Los vértices de $\mathcal{FM}(X)$ han sido obtenidos por Radojevic en [27, Proposición 2]:

Proposición 1 *El conjunto de las medidas* $\{0, 1\}$*-valuadas es el conjunto de vértices de* $\mathcal{FM}(X)$.

Hay otras funciones de conjunto que pueden utilizarse para representar una medida difusa. En este trabajo necesitaremos la llamada transformada de Möbius.

Definición 3 *[28] Sea μ una función de conjunto (no necesariamente una medida difusa) sobre X verificando $\mu(\emptyset) = 0$. La* **transformada o inversa de Möbius** *de μ es otra función de conjunto dada por*

$$m(A) := \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B), \, \forall A \subseteq X. \tag{2}$$

El valor $m(A)$ representa la fuerza del conjunto A en todas las coaliciones en las que participa. Nótese que la inversa de Möbius puede tomar valores negativos.

Como habíamos dicho en la introducción, para determinar una medida difusa son necesarios $2^n - 2$ coeficientes. El número de coeficientes crece exponencialmente con n, así como la complejidad del problema de aprendizaje. Esto reduce considerablemente la aplicación en situaciones reales de las medidas difusas. Es por ello que se han definido distintas subfamilias; de todas ellas, aquí usaremos las medidas k-aditivas.

Definición 4 *[11] Una medida difusa μ es* k**-aditiva** *si su transformada de Möbius se anula para todos los $A \subseteq X$ tales que $|A| > k$ y existe al menos un subconjunto A con k elementos tal que $m(A) \neq 0$.*

Así, una probabilidad no es más que una medida 1-aditiva. Las medidas k-aditivas son un término medio entre las probabilidades y las medidas difusas generales. Para definir una medida k-aditiva, el número de

coeficientes que son necesarios se reduce a

$$\sum_{i=1}^{k} \binom{n}{i}. \tag{3}$$

Más resultados sobre las medidas k-aditivas pueden verse en [13]. Denotaremos al conjunto de todas las medidas k'-aditivas con $k' \leq k$ por $\mathcal{FM}^k(X)$; nótese que $\mathcal{FM}^k(X)$ es un politopo.

Contra lo que cabría esperar, hemos demostrado en [22] que hay extremos de $\mathcal{FM}^k(X)$ que no son medidas $\{0,1\}$-valuadas. Más concretamente, se puede demostrar lo siguiente:

Proposición 2 *Sea un referencial X finito. Entonces los vértices de $\mathcal{FM}^1(X)$ y $\mathcal{FM}^2(X)$ son las medidas $\{0,1\}$-valuadas que están en estos subconjuntos.*

Teorema 1 *Si $|X| > 3$, entonces hay vértices de $\mathcal{FM}^k(X)$, con $3 \leq k \leq n-1$, que no son medidas $\{0,1\}$-valuadas.*

Estos extremos provienen de combinaciones convexas de medidas $\{0,1\}$-valuadas adyacentes en las que se producen compensaciones de las correspondientes transformadas de Möbius.

3 ADYACENCIA EN $\mathcal{FM}(X)$

En esta sección vamos a tratar el problema de caracterizar las condiciones en las que dos medidas $\{0,1\}$-valuadas son adyacentes. Para ello, nótese que dos vértices en un politopo en $\mathbb{R}^p$ son adyacentes si el segmento que los une es una arista, es decir, si el punto medio del segmento no puede ponerse como combinación convexa de otros vértices. Una forma de comprobar si dos vértices son adyacentes es ver si la matriz B de las restricciones que ambos extremos verifican con igualdad tiene rango $p-1$. En el caso de $\mathcal{FM}(X)$, estas restricciones son

$$\mu(\emptyset) = 0, \mu(X) = 1.$$

$$\mu(A) - \mu(A \backslash x_i) \geq 0 \quad \forall A \subseteq X, x_i \in A. \tag{4}$$

Esto proporciona un algoritmo para determinar la adyacencia, pero es muy lento y costoso. De hecho, se ha demostrado en [20, 26] que el problema de determinar la adyacencia en los vértices de un politopo es, en ocasiones, un problema NP-completo (véase [10] para una definición de problema NP-completo y nociones similares). Sin embargo, en esta sección daremos una condición necesaria y suficiente para la adyacencia en $\mathcal{FM}(X)$ que puede usarse para demostrar que la adyacencia puede determinarse en un tiempo polinomial en el número de conjuntos minimales de los vértices.

Desde un punto de vista formal, el conjunto $\mathcal{FM}(X)$ es un subpolitopo del politopo envolvente convexa de las funciones booleanas. Por otra parte, la adyacencia en este último politopo es trivial, ya que dos vértices son adyacentes en él si y sólo si su distancia de Hamming es 1. Sin embargo, no podemos usar este resultado para estudiar la adyacencia de $\mathcal{FM}(X)$; de hecho, puede demostrarse que medidas que están muy alejadas desde el punto de vista de la distancia de Hamming como funciones booleanas (y que por tanto no son adyacentes en dicho politopo), son adyacentes en $\mathcal{FM}(X)$.

Lema 1 *Si $\mu_1, \mu_2 \in \mathcal{FM}(X)$, $\mu_1 \wedge \mu_2$ y $\mu_1 \vee \mu_2$ verifican*

$$\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 = \frac{1}{2}(\mu_1 \wedge \mu_2) + \frac{1}{2}(\mu_1 \vee \mu_2) \tag{5}$$

y $\mu_1 \wedge \mu_2 \leq \mu_1, \mu_2 \leq \mu_1 \vee \mu_2$.

Como consecuencia directa se tiene:

Corolario 1 *Si μ_1 y μ_2 son adyacentes en $\mathcal{FM}(X)$, entonces $\mu_1 \leq \mu_2$ ó $\mu_2 \leq \mu_1$.*

Sin embargo, ésta no es una condición suficiente. De hecho, si dos vértices μ_1 y μ_2 no son comparables, por el Lema 1 se tiene que su supremo y su ínfimo no son adyacentes, a pesar de que estas dos medidas sí son comparables.

Definición 5 *[6] Sea $\mathbf{C}$ una colección de subconjuntos de X, y μ un vértice de $\mathcal{FM}(X)$ cuya familia de subconjuntos minimales es $\mathbf{D}$. Decimos que μ es* **$\mathbf{C}-$descomponible** *si existe una partición de $\mathbf{D}$ en dos subconjuntos no vacíos, $\mathbf{A}$ y $\mathbf{B}$, tales que $\mathbf{A} \not\subseteq \mathbf{C}$ y $\mathbf{B} \not\subseteq \mathbf{C}$, y de forma que dados $A \in \mathbf{A}$ y $B \in \mathbf{B}$, existe $C \in \mathbf{C}$ tal que $C \subseteq A \cup B$.*

Teorema 2 *[6] Sean $\mu, \mu_{\mathbf{C}}$ dos vértices de $\mathcal{FM}(X)$ tales que $\mu > \mu_{\mathbf{C}}$. Entonces, μ y $\mu_{\mathbf{C}}$ son adyacentes en $\mathcal{FM}(X)$ si y sólo si μ no es $\mathbf{C}$-descomponible.*

A partir de este resultado se puede demostrar:

Teorema 3 *[6] Sean μ y μ' dos extremos de $\mathcal{FM}(X)$. Entonces, se puede estudiar si son adyacentes en tiempo polinomial en el número de subconjuntos minimales de μ más el número de subconjuntos minimales de μ'.*

4 OBTENCIÓN DE LOS VÉRTICES DE LAS MEDIDAS k-ADITIVAS

Consideremos el conjunto $\mathcal{FM}(X)$. Desde un punto de vista geométrico, el conjunto $\mathcal{FM}^k(X)$ es un subpolitopo de $\mathcal{FM}(X)$ al que hemos añadido nuevas

restricciones; estas nuevas restricciones son $m(A) = 0, \forall |A| > k$.

Sea $|A| > k$ y consideremos la nueva restricción $m(A) = 0$. Entonces el conjunto de vértices del politopo de las medidas difusas que verifican además esta restricción está formado por los vértices $\{0,1\}$-valuados que verifican dicha restricción y los puntos de corte del hiperplano $m(A) = 0$ con las aristas del poliedro $\mathcal{FM}(X)$ [25]. Como conocemos por los resultados de la sección anterior cuáles son los vértices adyacentes de $\mathcal{FM}(X)$, conocemos también las aristas. Por tanto, el procedimiento anterior nos permite obtener los vértices del nuevo politopo.

De esta manera, podemos reiterar el procedimiento anterior con el nuevo politopo y con otra restricción, y así sucesivamente. Así se obtienen los vértices de $\mathcal{FM}^k(X)$. Lo que hace lento y costoso a este algoritmo es el tener que evaluar en cada paso la adyacencia de los nuevos vértices, ya que en este caso no se tienen resultados que caractericen la adyacencia similares a los de la sección anterior. Es por ello que estamos obligados a tratar con la definición general de adyacencia, con todo lo que ello supone a efectos de gasto computacional.

Nótese que del algoritmo que acabamos de describir se desprende una caracterización de los extremos de las medidas k-aditivas (y de los politopos intermedios). Sea $\mathbf{A}$ una colección de subconjuntos de X. Denotemos por $\mathcal{FM}_{\mathbf{A}}(X)$ el subconjunto de medidas de $\mathcal{FM}(X)$ cuyo Möbius se anula en todos los conjuntos de la colección $\mathbf{A}$. Se tiene el siguiente resultado:

Teorema 4 *Sean $\mathbf{A}$ una colección de subconjuntos de X (eventualmente vacía), B un subconjunto de X que no está en $\mathbf{A}$ y μ una medida de $\mathcal{FM}_{\mathbf{A}\cup\{B\}}(X)$ que no es un extremo de $\mathcal{FM}_{\mathbf{A}}(X)$. Entonces μ es un extremo de $\mathcal{FM}_{\mathbf{A}\cup\{B\}}(X)$ si y sólo si existen μ_1 y μ_2 dos extremos adyacentes de $\mathcal{FM}_{\mathbf{A}}(X)$ que no están en $\mathcal{FM}_{\mathbf{A}\cup\{B\}}(X)$ y $\lambda \in (0,1)$ tales que $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$.*

Como consecuencia de este resultado, en el caso de las medidas $(n-1)$-aditivas se tiene lo siguiente:

Teorema 5 *Sea μ una medida $(n-1)$-aditiva que no es $\{0,1\}$-valuada. Entonces μ es un vértice de $\mathcal{FM}^{n-1}(X)$ si y sólo si*

$$\mu = \lambda\mu_1 + (1-\lambda)\mu_2,$$

donde $\lambda \in (0,1)$ y μ_1 y μ_2 son dos vértices adyacentes de $\mathcal{FM}(X)$ que no están es $\mathcal{FM}^{n-1}(X)$.

Esto hace que en este caso particular podamos obtener todos los vértices de $\mathcal{FM}^{n-1}(X)$ sin más que conocer la adyacencia entre vértices de las medidas generales. Sin embargo, esto no es cierto en general; para otros grados de aditividad es posible encontrar vértices que no son $\{0,1\}$-valuados que no pueden ponerse como una combinación convexa de exactamente dos vértices $\{0,1\}$-valuados.

Ejemplo 1 *Consideremos $|X| = 5$. Sean μ_1 la medida difusa cuyos elementos minimales son los unipuntuales y μ_2 la medida difusa cuyos elementos minimales son los subconjuntos de 4 elementos. Por otra parte, sea μ_3 la medida difusa cuyos minimales son $\{(1,2),(1,3),(2,4),(3,5),(4,5)\}$.*

A partir de los resultados de la sección anterior puede comprobarse que μ_1 y μ_2 son vértices adyacentes de $\mathcal{FM}(X)$. Análogamente, también son adyacentes μ_2 y μ_3.

Por otra parte, a partir de la transformada de Möbius, puede verse que

$$\frac{4}{5}\mu_1 + \frac{1}{5}\mu_2, \frac{4}{5}\mu_3 + \frac{1}{5}\mu_2 \in \mathcal{FM}^4(X). \qquad (6)$$

Por tanto, estas medidas son extremos de $\mathcal{FM}^4(X)$ que no son $\{0,1\}$-valuadas. Denotemos a estas medidas μ y μ' respectivamente. Se puede comprobar que estas dos medidas son adyacentes como vértices de $\mathcal{FM}^4(X)$. Además,

$$\frac{1}{4}\mu + \frac{3}{4}\mu' \in \mathcal{FM}^3(X). \qquad (7)$$

Por lo dicho anteriormente, este es un extremo de las medidas 3-aditivas, y necesita de tres extremos de $\mathcal{FM}(X)$ para obtener una combinación convexa que nos conduzca hasta él.

La siguiente tabla muestra el número de extremos de las medidas k-aditivas para distintos valores de k y n.

Tabla 1: Vértices de las medidas k-aditivas.

$n\backslash k$	1	2	3	4	5
1	1				
2	2	4			
3	3	9	18		
4	4	16	303	166	
5	5	25	?	407201	7579

El número de medidas 3-aditivas cuando $|X| = 5$ no ha podido ser hallado debido a que hay un gran número de politopos intermedios y el proceso de paso de uno a otro es muy costoso.

A partir de los resultados de la tabla puede observarse que parece que el número de extremos k-aditivos que no son $\{0,1\}$-valuados es mucho mayor que el número

de vértices $\{0,1\}$-valuados. Esto confirma la enorme complejidad que tiene esta familia de medidas difusas frente a su sencillez en términos de interpretación.

5 Conclusiones y problemas abiertos

En este trabajo hemos planteado un algoritmo para calcular los vértices de las medidas k-aditivas. Este algoritmo sigue una serie de pasos iterativos que permite ir perfilando el conjunto $\mathcal{FM}^k(X)$. También hemos caracterizado los extremos de este conjunto.

El primer problema abierto que surge es el de caracterizar de alguna manera la adyacencia en los politopos intermedios. Esto nos permitiría reducir en gran medida el coste computacional del procedimiento.

A partir de los resultados obtenidos, puede observarse que, al igual que en el caso de las medidas difusas generales, el conjunto de extremos es demasiado grande para poder ser considerado como población inicial. Por ello, es necesario buscar otra población inicial alternativa. Una opción es considerar una población seleccionada al azar. Un primer paso en esta dirección para el caso general puede encontrarse en [21].

Agradecimientos

Esta investigación ha sido financiada parcialmente por la beca MEC y FEDER TIN2007-61273, y por los proyectos MTM2007-61193 y CAM-UCM910707.

Referencias

[1] M. Allais. Le comportement de l'homme rationnel devant le risque: critique des postulats de l'école américaine. *Econometrica*, (21):503–546, 1953. In French.

[2] F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, (34):199–205, 1963.

[3] A. Chateauneuf. Modelling attitudes towards uncertainty and risk through the use of Choquet integral. *Annals of Operations Research*, (52):3–20, 1994.

[4] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, (5):131–295, 1953.

[5] E. F. Combarro and P. Miranda. Identification of fuzzy measures from sample data with genetic algorithms. *Computers and Operations Research*, 33(10):3046–3066, 2006.

[6] E. F. Combarro and P. Miranda. On the polytope of non-additve measures. *Fuzzy Sets and Systems*, to appear.

[7] D. Dubois and H. Prade. A class of fuzzy measures based on triangular norms. *Int. J. General Systems*, 8:43–61, 1982.

[8] C. A. Bana e Costa and J.-C. Vansnick. MACBETH- An interative path towards the construction of cardinal value function. *International Transactions in Operational Research*, 1(4):489–500, 1994.

[9] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quart. J. Econom.*, (75):643–669, 1961.

[10] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory NP-Completeness*. Mc Graw Hill, New York, 1979.

[11] M. Grabisch. k-order additive discrete fuzzy measures. In *Proceedings of 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, pages 1345–1350, Granada (Spain), 1996.

[12] M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5:587–607, 1997.

[13] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, (92):167–189, 1997.

[14] M. Grabisch and J.-M. Nicolas. Classification by fuzzy integral-performance and tests. *Fuzzy Sets and Systems, Special Issue on Pattern Recognition*, (65):255–271, 1994.

[15] P. L. Hammer and R. Holzman. On approximations of pseudo-boolean functions. *Zeitschrift für Operations Research. Mathematical Methods of Operations Research*, (36):3–21, 1992.

[16] J. H. Holland. *Adaptation in natural and artificial systems*. Ann Arbor: The University of Michigan Press, 1975.

[17] J.-L. Marichal. k-intolerant capacities and Choquet integrals. In *Proceedings of Tenth International Conference of Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*, pages 601–608, Perugia (Italy), July 2004.

[18] J.-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155(3):771–791, 2004.

[19] J.-L. Marichal, P. Meyer, and M. Roubens. Sorting multiattribute alternatives : the TOMASO method. *Computer and Operations Resarch*, 32(4):861–877, 2005.

[20] T. Matsui. NP-completeness of non-adjacency relations on some 0-1-polytopes. In *Lecture Notes in Operations Research, 1*, pages 249–258. 1995.

[21] P. Miranda and E.F. Combarro. On the structure of some families of fuzzy measures. *IEEE Transactions on Fuzzy Systems*, 15(6):1068–1081, 2007.

[22] P. Miranda, E.F. Combarro, and P. Gil. Extreme points of some families of non-additive measures. *European Journal of Operational Research*, 33(10):3046–3066, 2006.

[23] P. Miranda and M. Grabisch. Characterizing k-additive fuzzy measures. In *Proceedings of Eighth International Conference of Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*, pages 1063–1070, Madrid (Spain), July 2000.

[24] P. Miranda and M. Grabisch. p-symmetric fuzzy measures. In *Proceedings of Ninth International Conference of Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*, pages 545–552, Annecy (France), July 2002.

[25] K. G. Murty. Adjacency on convex polyhedra. *SIAM Review*, 13(3):377–386, 1971.

[26] C. Papadimitriou. The adjacency relation on the travelling salesman polytope is NP-complete. *Mathematical Programming*, 14(1), 1978.

[27] D. Radojevic. The logical representation of the discrete Choquet integral. *Belgian Journal of Operations Research, Statistics and Computer Science*, 38(2–3):67–89, 1998.

[28] G. C. Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, (2):340–368, 1964.

[29] D. Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, (97(2)):255–261, 1986.

[30] L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the theory of Games*, volume II of *Annals of Mathematics Studies*, pages 307–317. Princeton University Press, 1953.

[31] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[32] M. Sugeno and T. Terano. A model of learning based on fuzzy information. *Kybernetes*, (6):157–166, 1977.

[33] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press, New Jersey (USA), 1944.

[34] J. Wang Z. Wang, K. Xu and G. Klir. Using genetic algorithms to determine nonnegative monotone set functions for information fusion in environments with random perturbation. *International Journal of Intelligent Systems*, 14:949–962, 1999.

EL RIESGO FINANCIERO DE UNA CARTERA MODELADA COMO CONJUNTO ALEATORIO

Ignacio Cascos

Dpto. Estadística, Universidad Carlos III de Madrid, ignacio.cascos@uc3m.es

Resumen

Representamos la ganancia asociada a una inversión como un conjunto aleatorio en el espacio Euclídeo d-dimensional y estudiamos cómo cuantificar el riesgo de la misma. Se justificará que medir este riesgo es equivalente a medir el riesgo de las *situaciones extremas* descritas por el conjunto aleatorio.

Palabras Clave: Conjunto aleatorio. Medida de riesgo. Riesgo coherente.

1 INTRODUCCIÓN

Es habitual modelar los rendimientos de una acción o cartera de valores a lo largo de un período de tiempo fijado de antemano como una variable aleatoria ξ. En el momento de efectuar una inversión, su ganacia (futura) es desconocida, por lo que puede considerarse como un valor aleatorio. La variable representará, por tanto, la ganancia que conlleva cierta inversión y el riesgo de dicha inversión puede cuantificarse utilizando una *medida de riesgo*.

Una *medida de riesgo univariante* asocia a cada variable aleatoria esencialmente acotada un valor real que cuantifica el riesgo de la inversión cuya ganancia está descrita por la variable. Podemos utilizar las medidas de riesgo para tomar decisiones sobre futuras inversiones o para diseñar estrategias de inversión.

La medida de riesgo univariante más utilizada es el *Valor en Riesgo* (VaR) que, para cualquier nivel $\alpha \in (0,1)$ fijado de antemano, se define como la menor cantidad de capital extra que necesita una compañía para que la probabilidad de que incurra en bancarrota sea exactamente α. Es decir, el valor en riesgo de una variable aleatoria ξ es el opuesto de su mayor cuantil de nivel α,

$$\mathrm{VaR}_\alpha(\xi) = -\inf\{t : \Pr(\xi \leq t) > \alpha\}\,.$$

El valor en riesgo tiene propiedades interesantes, es *equivariante por traslaciones*, *monotóno* y *homogéneo*. Satisface por tanto:

R1 $\mathrm{R}(\xi + x) = \mathrm{R}(\xi) - x$ para $x \in \mathbb{R}$. Si la inversión tiene una parte de beneficio garantizado, debemos descontar ese beneficio sobre la medida de riesgo;

R2 $\mathrm{R}(\eta) \leq \mathrm{R}(\xi)$ para $\xi \leq \eta$ c.s. Si una inversión es mejor que otra (c.s.), su riesgo será menor;

R3 $\mathrm{R}(\lambda\xi) = \lambda\mathrm{R}(\xi)$ para $\lambda > 0$. Al expresar el beneficio en una divisa distinta de la original, su riesgo se reescalará.

A cualquier funcional sobre un espacio de variables aleatorias esencialmente acotadas que satisface **R1**, **R2** y **R3** le llamamos *medida de riesgo univariante*

Artzner et al. [1] observaron que el Valor en Riesgo no es *coherente* con la máxima financiera de que la diversificación produce una reducción del riesgo de una cartera de inversiones. Así, la combinación de distintas clases de activos o instrumentos (diversificación), compensa un posible descenso en el precio de alguno de ellos, especialmente si los activos están poco correlacionados. Por lo tanto, al combinar distintos activos incurrimos en un riesgo menor. La modelización matemática de una combinación de activos se realiza mediante la suma de las variables aleatorias que modelan sus ganancias. De tal modo que una *medida de riesgo coherente*, además de las propiedades anteriores debe ser *subaditiva*. Es decir, el riesgo de la suma de dos variables aleatorias debe ser menor o igual que la suma de sus riesgos individuales.

R4 $\mathrm{R}(\xi+\eta) \leq \mathrm{R}(\xi)+\mathrm{R}(\eta)$. La diversificación reduce el riesgo.

Una medida de riesgo univariante que satisface **R4** es una *medida de riesgo coherente univariante.*

La media de riesgo coherente más utilizada en la práctica es el *déficil esperado* (ES) que se define para un nivel $\alpha \in (0, 1]$ como

$$\mathrm{ES}_\alpha(\xi) = -\frac{1}{\alpha}\int_0^\alpha F_\xi^{-1}(t)\mathrm{d}t\,,$$

donde $F_\xi^{-1}(t) = \inf\{t : \Pr(\xi \leq t) > \alpha\}$ es la función cuantil de la variable aleatoria ξ.

Otra medida de riesgo coherente es el *mínimo esperado* (EM) definida para un nivel $n \geq 1$ como

$$\mathrm{EM}_n(\xi) = -\mathbb{E}\min\{\xi_1, \xi_2, \ldots, \xi_n\}\,,$$

donde $\xi_1, \xi_2, \ldots, \xi_n$ son n copias independientes de ξ.

De un modo más general, es posible modelar el beneficio asociado a una inversión como un vector aleatorio. Cada uno de los componentes de dicho vector haría referencia a cada uno de los distintos activos que componen la cartera. Jouini et al. [5] son los primeros que proponen medir el riesgo de los vectores aleatorios con medidas de riesgo que toman como valores subconjuntos de $\mathbb{R}^d$. No obstante, las únicas familias de riesgos que construyen toman como valor traslaciones de un cono fijado de antemano, por lo que toda la información sobre el riesgo estará contenida en el vértice del cono trasladado.

Más adelante, Cascos y Molchanov [3] describen medidas de riesgo para vectores aleatorios que toman valores en un cono genérico dotado de una relación de orden (para poder comparar riesgos) y con las operaciones de la suma y producto por escalares, para poder operar con los riesgos. El espacio Euclídeo d-dimensional con la relación de orden componente a componente es el espacio más sencillo (distinto de la recta real) que tiene estas propiedades. Lamentablemente, demuestran que si una medida de riesgo coherente para vectores aleatorios toma valores en $\mathbb{R}^d$ es necesariamente marginalizada, en el sentido de que cada uno de sus compontes será una medida de riesgo coherente evaluada sobre cada una de las componentes del vector aleatorio. Desde un punto de vista más constructivo, se pueden considerar otras estructuras más complejas (e.g. familias de subconjuntos de $\mathbb{R}^d$) que dan lugar a medidas de riesgo más *interesantes.*

A lo largo del presente trabajo, modelaremos una cartera como un conjunto aleatorio en el espacio Euclídeo d-dimensional. Es decir, cada resultado posible de la inversión será un conjunto de puntos en $\mathbb{R}^d$. Imaginemos que estamos considerando la posibilidad de firmar un contrato de futuro que nos faculta para adquirir activos de varios tipos distintos a un precio que varía según el número de activos de cada tipo de compremos. El beneficio de esta inversión vendrá determinado por el precio de los activos que será un conjunto y por el valor (aleatorio) de los mismos en la fecha de vencimiento del contrato.

Una ligera modificación de la definición de medida de riesgo para una cartera modelada como un vector aleatorio dada por Cascos y Molchanov [3] hace que sea válidad para conjuntos aleatorios.

Obsérvese que en el caso univariante, un conjunto aleatorio cerrado y convexo es un intervalo aleatorio y parece natural medir el riesgo de ese intervalo aleatorio como el riesgo de la peor situación descrita por el mismo, es decir, como el riesgo de la variable aleatoria definida como el extremo inferior del intervalo. No obstante, la falta de un orden total en el espacio Euclídeo multivariante no nos permite generalizar esta situación cuando el conjunto aleatorio toma como valores subconjuntos de $\mathbb{R}^d$, lo que constituye una motivación para el interés de este trabajo.

En la Sección 2 se describirán brevemente algunos conceptos matemáticos que necesitaremos a lo largo del trabajo. En la Sección 3 se definirán las nuevas medidas de riesgo que estudiaremos. A continuación, en la Sección 4 justificaremos que para medir el riesgo de un conjunto aleatorio hay que considerar las situaciones extremas descritas por el mismo y finalmente en la Sección 5 propondremos dos medidas de riesgo para conjuntos aleatorios. Una de ellas está inspirada en las regiones centrales del zonoide de Koshevoy y Mosler [6] y la otra en las regiones centrales de la envolvente convexa esperada de Cascos [2].

2 PRELIMINARES

Para que el término riesgo cobre sentido en un contexto financiero, el beneficio asociado a una inversión debe ser aleatorio, por lo tanto una cartera debe ser siempre un elemento aleatorio. Identificaremos una *cartera-conjunto* con un *conjunto aleatorio compacto y convexo* en $\mathbb{R}^d$.

Antes de introducir los conjuntos aleatorios, repasaremos algunos conceptos necesarios. Dados A, B dos subconjuntos de $\mathbb{R}^d$, su *suma de Minkowski (o elemento a elemento)* se define como $A \oplus B = \{a+b : a \in A,\, b \in B\}$. Dado $\lambda \in \mathbb{R}$, el *producto de un conjunto por un escalar* se define como $\lambda A = \{\lambda a : a \in A\}$. Escribiremos $-A$ para referirnos al *conjunto reflejado* de A, i.e., $-A = \{-a : a \in A\}$. Finalmente, dado $A \subset \mathbb{R}^d$, su *función soporte* en $u \in \mathbb{R}^d$ vendrá dada por $h(A, u) = \sup\{\langle x, u\rangle : x \in A\}$, donde $\langle\cdot,\cdot\rangle$ es el producto escalar. La función soporte de un conjunto compacto y convexo es continua y lo caracteriza.

Un *conjunto aleatorio compacto y convexo* $\mathbb{R}^d$ es una aplicación medible de un espacio de probabilidad en la familia de conjuntos compactos y convexos de $\mathbb{R}^d$ dotada de la topología generada por la distancia de Hausdorff y la σ-algebra de Borel correspondiente. Decimos que un conjunto aleatorio X es esencialmente acotado si la variable aleatoria $\sup\{\|x\| : x \in X\}$ lo es. Denotamos como $\mathcal{L}_\infty^d$ al conjunto de todos los conjuntos aleatorios compactos y convexos de $\mathbb{R}^d$ en el espacio de probabilidad $(\Omega, \mathfrak{F}, \Pr)$.

Un vector aleatorio ξ es una *selección* de un conjunto aleatorio X si $\xi \in X$ *c.s.* Todo conjunto aleatorio no vacío X tiene, al menos, una selección y si al menos una selección de X tiene primer momento finito, podemos definir la *esperanza de un conjunto aleatorio*, denotada $\mathbb{E}X$, como el conjunto de esperanzas de todas las selecciones integrables. Si el conjunto aleatorio es convexo (como es el caso), su esperanza será convexa y, como es esencialmente acotado, además será un conjunto compacto. Alternativamente, la esperanza de un conjunto aleatorio puede darse implícitamente por su función soporte como

$$h(\mathbb{E}X, u) = \mathbb{E}h(X, u) \quad \text{para todo } u \in \mathbb{R}^d .$$

Para poder comparar dos carteras-conjunto, necesitamos un orden parcial para subconjuntos de $\mathbb{R}^d$. Dados $A, B \subset \mathbb{R}^d$, escribimos $B \preceq A$ si $A \oplus \mathbb{R}^d_+ \subseteq B \oplus \mathbb{R}^d_+$. Dados dos carteras-conjunto X e Y, si $Y \preceq X$ c.s. entonces X es preferida a Y en el sentido de que si escogemos X, incurrimos en un menor riesgo.

3 MEDIDA DE RIESGO DE UNA CARTERA-CONJUNTO

Trabajaremos con riesgos en la familia $\mathbb{C}_-$ formada por todos los subconjuntos cerrados y convexos de $\mathbb{R}^d$ que son *inferiores*. Un conjunto $A \subset \mathbb{R}^d$ es inferior si $A \oplus \mathbb{R}^d_- = A$.

Una *medida de riesgo de una cartera-conjunto* cuantificará el riesgo de un conjunto aleatorio (cartera-conjunto) como un elemento de $\mathbb{C}_-$. Mediante la relación de inclusión, podemos comparar dos riesgos (dos elementos de $\mathbb{C}_-$) y podemos operar con los riesgos utilizando la suma de Minkowski y el producto por escalares positivos.

Definición Un funcional $\rho : \mathcal{L}_\infty^d \mapsto \mathbb{C}_-$ decimos que es una *medida de riesgo coherente de una cartera-conjunto* si $\rho(\{0\}) = \mathbb{R}^d_-$ y cumple las siguientes condiciones

RC1 Equivarianza por traslaciones: $\rho(X \oplus \{x\}) = \rho(X) \oplus \{-x\}$ para todo $x \in \mathbb{R}^d$;

RC2 Monotonía: $\rho(X) \subseteq \rho(Y)$ si $Y \preceq X$ c.s.;

RC3 Homogeneidad: $\rho(tX) = t\rho(X)$ para todo $t > 0$ y $X \in \mathcal{L}_\infty^d$;

RC4 Subaditividad: $\rho(X \oplus Y) \subseteq \rho(X) \oplus \rho(Y)$ para todo $X, Y \in \mathcal{L}_\infty^d$.

Si ρ sólamente cumple **RC1**, **RC2** y **RC3** decimos que se trata una *medida de riesgo de una cartera-conjunto.*

Variables aleatorias Tal como se ha visto en la Sección 1, la definición de medida de riesgo (coherente) univariante satisface propiedades análogas a las anteriores, sustituyendo la suma de Minkowski por la adición de números reales y las relaciones de orden que se utilizan para comparar riesgos y carteras $\subseteq$ y $\preceq$ por la relación $\leq$ sobre la recta real.

Vectores aleatorios La definición de medida de riesgo (coherente) para un vector aleatorio propuesta por Cascos y Molchanov [3] plantea riesgos definidos en un cono genérico y la comparación de vectores aleatorios según la relación menor o igual componente a componente. El cono $\mathbb{C}_-$ constituye una familia de riesgos particularmente interesante en este caso. Los riesgos que toman valores en esta familia se estudian con detalle en Cascos y Molchanov [4].

3.1 CARTERAS UNIVARIANTES

Dada una variable aleatoria ξ, definimos el conjunto aleatorio unipuntual $\{\xi\}$. Si nos restringimos a conjuntos unipuntuales de la recta real, la relación $\preceq$ se convierte en la relación menor o igual sobre $\mathbb{R}$. Además, la familia $\mathbb{C}_-$ estará compuesta por intervalos del tipo $(-\infty, x]$ y la relación de contenido $\subseteq$ entre ellos se convertirá también en la relación menor o igual entre sus extremos superiores. En consecuencia, la medida de riesgo coherente para carteras-conjunto ρ, se convierte en $\rho(\{\xi\}) = (-\infty, \mathrm{R}(\xi)]$, donde R es una medida de riesgo coherente para carteras univariantes.

Un conjunto aleatorio aleatorio compacto y convexo en $\mathbb{R}$ es obviamente un intervalo aleatorio $X = [\min X, \max X]$. La relación de orden $\preceq$ sólo afecta a $\min X$, por lo tanto si ρ es una medida de riesgo coherente para carteras-conjunto, tenemos que $\rho(X) = (-\infty, \mathrm{R}(\min X)]$, donde R es una medida de riesgo coherente univariante.

3.2 CARTERAS MULTIVARIANTES UNIPUNTUALES

A partir de un vector aleatorio ξ en $\mathbb{R}^d$, podemos construir el conjunto aleatorio (unipuntual) $\{\xi\}$. La relación $\preceq$ entre conjuntos aleatorios se convierte en

la relación menor o igual componente a componente, la suma de Minkowski en la suma de puntos de $\mathbb{R}^d$ y el producto por un escalar positivo es la operación externa del espacio vectorial $\mathbb{R}^d$ sobre $\mathbb{R}$.

4 CONJUNTOS DE ACEPTACIÓN

Un asunto crucial a la hora de cuantificar el riesgo de una cartera es decidir qué carteras deterministas son aceptables. Para esto, en primer lugar se toma la propiedad de monotonía **RC2** y se consideran aceptables todas las carteras deterministas que son preferibles a la $\{0\}$, de un modo más general, son aceptables todas aquellas carteras cuyo riesgo es menor que el de $\{0\}$.

En la literatura clásica sobre riesgos financieros las caracterizacines de medidas de riesgo a partir de resultados de dualidad juegan un papel fundamental. El esquema habitual que siguen estos resultados consiste en construir un conjunto de carteras que consideramos *aceptables*, es decir, con un riesgo asumible. Si una medida de riesgo es equivariante afín, puede escribirse a partir de la cantidad beneficio extra garantizado necesario para que una cartera sea aceptable. Finalmente, si el conjunto de carteras aceptables tiene *buenas* propiedades, es decir, si se trata de un cono convexo, disponemos de resultados de dualidad que nos permiten caracterizar dichos riesgos.

Los resultados de dualidad anteriormente mencionados se satisfacen en espacios de Banach, pero la familia de conjuntos compactos y convexos de $\mathbb{R}^d$ no lo es. De tal modo que para construir resultados de dualidad en el nuevo contexto, en lugar de considerar estos conjuntos, consideramos la función soporte que los caracteriza. Dicha función soporte es continua y podemos restringirla a la esfera unidad S^{d-1} en $\mathbb{R}^d$ y aprovechar que el conjunto de funciones continuas en S^{d-1} con la distancia del supremo es un espacio de Banach. La distancia del supremo entre funciones soporte es la distancia de Hausdorff entre los conjuntos compactos y convexos que determinan.

4.1 CONO RESTRINGIDO

En primer lugar consideraremos riesgos definidos exclusivamente en el cono restringido

$$\mathbb{C}_{\mathrm{r}} = \{\{x\} \oplus \mathbb{R}^d_- : x \in \mathbb{R}^d\}\,.$$

Las carteras-conjunto aceptables serán aquellas pertenecientes al *conjunto de aceptación restringido*, que es el conjunto de carteras-conjunto cuyo riesgo es menor que el de la cartera determinista $\{0\}$,

$$\mathcal{A}_{\mathrm{r}} = \{X \in \mathcal{L}^d_\infty : \rho(X) \subseteq \mathbb{R}^d_-\}\,.$$

Si ρ es una medida de riesgo coherente, el conjunto $\mathcal{A}_{\mathrm{r}}$ es un cono convexo y además, aprovechando que trabajamos exclusivamente con riesgos en $\mathbb{C}_{\mathrm{r}}$, el riesgo asociado a una cartera-conjunto puede escribirse como

$$\rho(X) = \mathbf{min}\{x : X + x \in \mathcal{A}_{\mathrm{r}}\} \oplus \mathbb{R}^d_-\,,$$

donde **min** es el mínimo componente a componente.

Si en lugar de considerar un conjunto aleatorio como tal, consideramos la función soporte (aleatoria) que lo define, podemos estudiar el conjunto dual del de funciones soporte (aleatorias) aceptables. Obsérvese que la relación de preferencia establecida entre conjuntos aleatorios (adaptada ahora a sus funciones soporte) juega un papel clave en la construcción de este conjunto.

Reconstruyendo el conjunto de aceptación a partir de su dual, podemos demostrar que si ρ es coherente,

$$\rho(X) = \{(\mathrm{R}_1(\min X_1), \ldots, \mathrm{R}_d(\min X_d))\} \oplus \mathbb{R}^d_-\,, \quad (1)$$

donde X_i es el intervalo aleatorio dado por la proyección de X en la i-ésima coordenada y $\mathrm{R}_1, \ldots, \mathrm{R}_d$ son d medidas de riesgo univariantes coherentes. Por tanto, si e_i es el i-ésimo vector de la base canónica en $\mathbb{R}^d$, $\min X_i = -h(X, -e_i)$. En consecuencia, el riesgo del conjunto aleatorio X será el mismo que el del conjunto aleatorio unipuntual constituido por los mínimos de las proyecciones de X en cada una de las coordenadas $\{\mathbf{min}X\}$. Nótese que si consideramos medidas de riesgo que toman valores en todo $\mathbb{C}_-$, el riesgo de $\{\mathbf{min}X\}$ no es necesarimanete una traslación de $\mathbb{R}^d_-$.

En general, podemos demostrar que si ρ es una medida de riesgo coherente de una cartera-conjunto, tenemos que

$$h(\rho(X), e_i) = \mathrm{R}(-h(X, -e_i)) = \mathrm{R}(\min X_i)\,, \quad (2)$$

donde R es una medida de riesgo coherente univariante. La clave para obtener este resultado es que las direcciones descritas por los $-e_i$ son las direcciones extremas que determinan la relación de orden $\preceq$ entre las carteras.

4.2 SITUACIONES EXTREMAS

Se ha comprobado que para evaluar el riesgo de una cartera-conjunto estamos considerando las peores situaciones descritas por el conjunto aleatorio. Es decir, evaluamos el riesgo de la función soporte del conjunto en las direcciones $\{-e_1, -e_2, \ldots, -e_d\}$. El motivo por el que ocurre esto es porque éstas son las direcciones extremas consideradas por la relación de orden $\preceq$ entre subconjuntos de $\mathbb{R}^d$.

Sería posible cambiar **RC2** y construir nuevas relaciones de orden entre subconjuntos compactos y con-

vexos de $\mathbb{R}^d$. Otra relación natural sería considerar el orden según el cual X es preferible a Y si $Y \oplus \mathbb{R}^d_- \subseteq X \oplus \mathbb{R}^d_-$ c.s. Con ese orden obtendríamos relaciones análogas a (1) y (2) con el máximo en lugar del mínimo.

5 EJEMPLOS PARTICULARES DE MEDIDAS DE RIESGO

A continuación, describiremos dos medidas de riesgo para carteras-conjunto. La primera está inspirada en las regiones del zonoide de Koshevoy y Mosler [6] y está relacionada con el déficit esperado y la segunda está inspirada en las regiones de la envolvente convexa esperada de Cascos [2] y está relacionada con el mínimo esperado.

5.1 RIESGO DEL ZONOIDE (ρ_{ZD})

Para un conjunto aleatorio X en $\mathbb{R}^d$, construimos el conjunto aleatorio $\tilde{X}$ en $\mathbb{R}^{d+1}$ añadiendo una componente extra (en la primera coordenada) que es constante e igual a 1, $\tilde{X} = (1, X)$. Para $\alpha \in (0, 1]$, definimos

$$\mathrm{ZD}^{\alpha}(X) = \frac{1}{\alpha}\mathrm{proj}_{\alpha}\mathbb{E}\mathrm{co}\big(\{0\} \cup \tilde{X}\big)\,,$$

donde para cada $C \subset \mathbb{R}^{d+1}$, $\mathrm{proj}_{\alpha}(C)$ es la proyección en las d últimas coordenadas de la intersección del hiperplano $\{(\alpha, x) : x \in \mathbb{R}^d\}$ con C.

El funcional

$$\rho^{\alpha}_{\mathrm{ZD}}(X) = -\big(\mathrm{ZD}^{\alpha}(X) \oplus \mathbb{R}^d_+\big)$$

será una medida de riesgo coherente para carteras-conjunto.

Esta medida de riesgo está relacionada con el déficit esparado. En concreto, en el caso unidimensional ($d = 1$), tenemos

$$\rho^{\alpha}_{\mathrm{ZD}}(X) = \big(-\infty, \mathrm{ES}_{\alpha}(\min X)\big]\,,$$

donde $\min X$ es la variable aleatoria definida como el extremo inferior del intervalo aleatorio X. En general, en cualquier dimensión, tenemos que para $u \in \mathbb{R}^d_+$ se cumple

$$h(\rho^{\alpha}_{\mathrm{ZD}}(X), u) = \mathrm{ES}_{\alpha}\big(-h(X, -u)\big)\,.$$

5.2 RIESGO DE LA ENVOLVENTE CONVEXA ESPERADA (ρ_{E})

Para cualquier $n \geq 1$, definimos

$$\rho^{n}_{\mathrm{E}}(X) = -\big(\mathbb{E}\mathrm{co}(X_1 \cup X_2 \cup \ldots \cup X_n) \oplus \mathbb{R}^d_+\big)\,,$$

donde $X_1, X_2, \ldots, X_n$ son n copias independientes del conjunto aleatorio X y co representa la envolvente convexa. El funcional ρ^n_{E} es una medida de riesgo coherente para carteras-conjunto que está relacionada con el mínimo esperado. En el caso unidimensional tenemos ahora que

$$\rho^{n}_{\mathrm{E}}(X) = \big(-\infty, \mathrm{EM}_n(\min X)\big]\,,$$

y en general, si $d \geq 1$ y $u \in \mathbb{R}^d_+$ se cumple

$$h(\rho^{n}_{\mathrm{E}}(X), u) = \mathrm{EM}_n\big(-h(X, -u)\big)\,.$$

Agradecimientos

El autor desea agradecer el apoyo financiero del proyecto MTM2005-02254 del Ministerio de Educación y Ciencia.

Referencias

[1] Artzner, P., Delbaen, F., Eber, J.M. y Heath, D. (1999) Coherent measures of risk. *Math. Finance*, 9:203–228.

[2] Cascos, I. (2007) Depth functions based on a number of observations of a random vector. WP 07-29. Statistics and Econometrics Series, Universidad Carlos III de Madrid. Disponible en `http://econpapers.repec.org/paper/ctewsrepe/`

[3] Cascos, I. y Molchanov, I. (2007) Multivariate risks and depth-trimmed regions. *Finance Stoch.*, 11:373–397.

[4] Cascos, I. y Molchanov, I. (2008) Multivariate spectral risk measures. *En preparación.*

[5] Jouini, E., Meddeb, M. y Touzi, N. (2004) Vector-valued coherent risk measures. *Finance Stoch.*, 8:531–552.

[6] Koshevoy, G. y Mosler, K. (1997) Zonoid trimming for multivariate distributions. *Ann. Statist.*, 25:1998–2017.

REGIONES CENTRALES BASADAS EN FAMILIAS DE SUCESOS DIFUSOS

Pedro Terán

Grupo Decisión Multicriterio Zaragoza. Facultad de Ciencias Económicas y Empresariales.
Universidad de Zaragoza. Gran Vía, 2. E-50005 Zaragoza. teran@unizar.es

Resumen

Presentamos una definición abstracta de centralidad en una familia de distribuciones multivariantes. Las definiciones y su interpretación están basadas en la Teoría de Conjuntos Difusos. Intuitivamente, un punto se considera central, con respecto a una familia de sucesos difusos de referencia, si "pertenece solamente a sucesos que pueden tener alta probabilidad". El concepto de centralidad es difuso a su vez y da lugar a un conjunto difuso de puntos centrales, que está muy relacionado con la noción de profundidad estadística, de creciente importancia en Análisis Multivariante.

Palabras Clave: Suceso difuso, Centralidad, Profundidad estadística, Probabilidad, Capacidad de Choquet, Conjunto aleatorio.

1 INTRODUCCIÓN

En este trabajo, presentamos una perspectiva unificadora, basada en conceptos difusos, de las medidas de tendencia central en el ámbito del Análisis Multivariante. La idea básica es que la noción de *centralidad* de un punto en una distribución de probabilidad (o, más generalmente, en una familia de distribuciones) es difusa. Consideramos que un punto es central en la medida en que los sucesos difusos a los que pertenece tienen alta probabilidad. Eligiendo adecuadamente la familia de sucesos difusos de referencia, se obtienen diferentes definiciones que incluyen muchos conceptos conocidos.

El ámbito general de la definición no se circunscribe a distribuciones de probabilidad, sino que puede aplicarse igualmente a familias de distribuciones. Esto incluye varios tipos de modelos de probabilidades imprecisas, como las capacidades de Choquet y los conjuntos aleatorios.

En este trabajo presentamos las propiedades básicas de las *regiones centrales* y α*-centrales* así definidas, así como múltiples ejemplos de conceptos que caen bajo esta definición general: medidas de tendencia central en el caso multivariante, esperanzas y medianas de conjuntos aleatorios, y funciones de profundidad.

2 PRELIMINARES

Llamamos *suceso difuso* a una aplicación medible definida en un espacio medible Ω y con valores en $[0,1]$. La familia de todos los sucesos difusos es un σ-anillo y cualquier probabilidad definida en Ω puede extenderse a sucesos difusos [14] como

$$P(A) = \int A \mathrm{d}P.$$

Aunque esta extensión no es única, sí es ampliamente la más utilizada en la práctica.

Una *capacidad* (o *capacidad de Choquet*) es una aplicación ν de la σ-algebra de Borel de $\mathbf{R}^d$ al intervalo $[0,1]$ con las siguientes propiedades:

a) $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$, $\nu(A) \leq \nu(B)$ si $A \subset B$,

b) $\nu(C_n) \searrow \nu(C)$ si $C_n \searrow C$ y C_n, C son cerrados,

c) $\nu(A_n) \nearrow \nu(A)$ si $A_n \nearrow A$.

Esta es la definición dada en [5]; existe una definición más general (véase, p.ej., [8]) en la cual conjuntos compactos reemplazan a los cerrados.

Una capacidad ν se llama *submodular* si

$$\nu(A \cup B) + \nu(A \cap B) \leq \nu(A) + \nu(B)$$

para cualesquiera sucesos A, B.

De forma análoga a la probabilidad de un suceso difuso, podemos definir su capacidad reemplazando la integral de Lebesgue por una integral de Choquet:

$$\nu(A) = \int A \mathrm{d}\nu = \int_0^1 \nu(A_\alpha)\mathrm{d}\alpha.$$

3 DEFINICIONES Y EJEMPLOS

Los elementos básicos de la definición son:

(I) Una familia de distribuciones $\mathcal{Q}$.

(II) Una familia de sucesos difusos $\mathcal{A}$ (sucesos de referencia).

Ejemplos incluidos en (I) son:

(a) Una variable aleatoria.

(b) Un modelo estadístico, paramétrico o no paramétrico.

(c) Probabilidades superiores e inferiores.

(d) Una capacidad de Choquet.

(e) Un conjunto aleatorio.

(f) Un entorno de una probabilidad.

(g) Otras familias definidas en función de una probabilidad.

Supondremos en lo sucesivo que la familia $\mathcal{Q}$ es compacta en la topología débil.

1. Definición. La *región central* de $\mathcal{Q}$ respecto a $\mathcal{A}$ se define como

$$\mathbf{C}_{\mathcal{A}}(\mathcal{Q}) = \{x \in \mathbf{R}^d \mid \forall A \in \mathcal{A}, \sup_{P\in\mathcal{Q}} P(A) \geq A(x)\}.$$

La interpretación intuitiva de esta definición es la siguiente: consideramos que un punto x es central en $\mathcal{Q}$ cuando "si x 'pertenece' a un suceso de referencia, necesariamente este tiene probabilidad 'alta' para al menos una distribución de la familia $\mathcal{Q}$".

2. Observación. Si la familia $\mathcal{A}$ es cerrada bajo complementación con el complemento $c(\alpha) = 1 - \alpha$, entonces se tiene que

$$\mathbf{C}_{\mathcal{A}}(\mathcal{Q}) = \{x \in \mathbf{R}^d \mid \forall A \in \mathcal{A}, \sup_{P\in\mathcal{Q}} P(A) \geq A(x) \geq \inf_{P\in\mathcal{Q}} P(A)\}$$

y por tanto la definición es intuitivamente más fuerte ya que se tiene también que "si x 'no pertenece' a un suceso de referencia, entonces su probabilidad 'es baja' para al menos una distribución de la familia $\mathcal{Q}$".

3. Observación. Se pueden normalizar los sucesos difusos de $\mathcal{A}$ sin que $\mathbf{C}_{\mathcal{A}}(\mathcal{Q})$ varíe, por lo que sólo es necesario considerar familias de sucesos difusos cuyo supremo es 1.

Esta definición de centralidad es nítida y podemos preferir considerar el concepto como gradual, con puntos más o menos centrales en la familia de distribuciones dada. Una forma de conseguirlo es "flexibilizar" la definición, formalizando la noción intuitiva como sigue:

4. Definición. Para cualquier $\alpha \in (0, 1]$, la *región α-central* de $\mathcal{Q}$ respecto a $\mathcal{A}$ se define como

$$\mathbf{C}^{\alpha}_{\mathcal{A}}(\mathcal{Q}) = \{x \in \mathbf{R}^d \mid \forall A \in \mathcal{A}, \sup_{P\in\mathcal{Q}} P(A) \geq \alpha A(x)\}.$$

Podemos, por tanto, definir un conjunto difuso de puntos centrales como

$$C(x) = \sup\{\alpha \in (0, 1] \mid x \in \mathbf{C}^{\alpha}_{\mathcal{A}}(\mathcal{Q})\},$$

cuyos α-cortes van a ser las regiones α-centrales.

5. Proposición. Sea I la implicación difusa de Goguen. Entonces,

$$C(x) = \inf_{A\in\mathcal{A}} I(A(x), \sup_{P\in\mathcal{Q}} P(A)) = \inf_{A\in\mathcal{A}} \sup_{P\in\mathcal{Q}} I(A(x), P(A)).$$

6. Propiedades básicas. Si los elementos de $\mathcal{A}$ son inferiormente semicontinuos, entonces cada $\mathbf{C}^{\alpha}_{\mathcal{A}}(\mathcal{Q})$ es cerrado.

Si los elementos de $\mathcal{A}$ son convexos, entonces cada $\mathbf{C}^{\alpha}_{\mathcal{A}}(\mathcal{Q})$ es convexo.

Si $\mathcal{A}$ es invariante por transformaciones afines, entonces cada $\mathbf{C}^{\alpha}_{\mathcal{A}}(\mathcal{Q})$ es afínmente equivariante.

7. Ejemplos (caso unidimensional). Supongamos que $\mathcal{Q}$ está formada únicamente por la distribución inducida por una variable aleatoria ξ. Entonces, para distintas elecciones de $\mathcal{A}$ y $\alpha \in (0, 1]$, obtenemos que $\mathbf{C}^{\alpha}_{\mathcal{A}}(\{P_\xi\})$ es:

(1) La esperanza de ξ (si existe, en otro caso la región central es vacía).

(2) La mediana de ξ.

(3) La moda de ξ (si ξ es discreta).

8. Centralidad y profundidad. Nuestra definición de centralidad está muy relacionada con las *funciones*

de profundidad definidas en Análisis Multivariante [6]. Muchas funciones de profundidad en la literatura son casos particulares de conjuntos difusos de puntos centrales.

9. Ejemplos (caso multivariante). Sea ξ un vector aleatorio. Entonces, para distintas elecciones de $\mathcal{A}$ y $\mathcal{Q}$, obtenemos que $C(x)$ es:

(1) La profundidad semiespacial de x [13].

(2) La profundidad simplicial de x [7].

(3) La profundidad zonoidal de x [9].

(4) La profundidad, por *convex hull peeling*, de x [4].

(5) La profundidad, por la envolvente convexa esperada, de x [1].

(6) La profundidad, por el orden simétrico, de x [3].

(7) La probabilidad del suceso $\{\xi = x\}$ (si ξ es discreta).

En particular, en el caso multivariante se obtienen las *medianas generalizadas* (puntos de máxima profundidad) asociados a las anteriores funciones de profundidad; además de la esperanza y la moda, como en el caso unidimensional.

4 REGIONES CENTRALES PARA CAPACIDADES DE CHOQUET Y CONJUNTOS ALEATORIOS

En esta sección mostramos cómo extender las regiones centrales a modelos de probabilidades imprecisas que están implícitamente asociados a familias de distribuciones.

En particular, para cada capacidad se puede definir su *núcleo* o conjunto de probabilidades dominadas por ella; y para un conjunto aleatorio compacto, su conjunto de *distribuciones seleccionables*. Tenemos, por tanto,

$$\operatorname{core}\nu = \{P \text{ probabilidad } \mid P \leq \nu\},$$

$$\operatorname{Sel}(X) = \{P_\xi \mid \xi \in X \text{ casi seguro}\}.$$

De este modo, podemos definir

$$\mathbf{C}^\alpha_{\mathcal{A}}(\nu) = \mathbf{C}^\alpha_{\mathcal{A}}(\operatorname{core}\nu),$$

$$\mathbf{C}^\alpha_{\mathcal{A}}(X) = \mathbf{C}^\alpha_{\mathcal{A}}(\operatorname{Sel}(X)).$$

Hay que notar que para cada conjunto aleatorio compacto X, se tiene también una capacidad ν_X determinada por las relaciones

$$\nu_X(A) = P(X \cap A \neq \emptyset\}, \quad A \text{ compacto.}$$

10. Proposición. Sea ν una capacidad submodular, entonces

$$\mathbf{C}^\alpha_{\mathcal{A}}(\nu) = \{x \in \mathbf{R}^d \mid \forall A \in \mathcal{A}, \nu(A) \geq \alpha A(x)\}$$

y, para cualquier $x \in \mathbf{R}^d$,

$$C(x) = \inf_{A \in \mathcal{A}} I(A(x), \nu(A)).$$

Aparentemente, $\mathbf{C}^\alpha_{\mathcal{A}}(X)$ podría no estar bien definida, en el sentido de que la familia de distribuciones $\operatorname{Sel}(X)$ puede no ser cerrada. La siguiente proposición establece que la definición dada es correcta y consistente con la definición para capacidades.

11. Proposición. Sea X un conjunto aleatorio compacto, entonces

$$\mathbf{C}^\alpha_{\mathcal{A}}(X) = \mathbf{C}^\alpha_{\mathcal{A}}(\operatorname{cl}\operatorname{Sel}(X)) = \mathbf{C}^\alpha_{\mathcal{A}}(\nu_X).$$

12. Ejemplos (conjuntos aleatorios). Si X es un conjunto aleatorio compacto, entonces para diferentes elecciones de $\mathcal{A}$ y $\alpha \in (0, 1]$, se cumple que $\mathbf{C}^\alpha_{\mathcal{A}}(X)$ es:

(1) La esperanza de Aumann de X (si Ω es no atómico; en el caso general, la envolvente convexa de la esperanza de Aumann).

(2) La esperanza de Herer de X.

(3) La esperanza radio-vector de X.

(4) La esperanza de Vorob'ev de X.

(5) La mediana de Vorob'ev de X.

5 COMENTARIOS FINALES

Aparentemente, Mosler [9] fue el primero en utilizar, en la definición de la profundidad zonoidal, la condición $x/y \geq \alpha$ como relajación gradual de la condición $x \geq y$ que permite obtener, de una definición de mediana generalizada, una función de profundidad estadística.

Cascos y López-Díaz [2] dieron una definición más general de profundidad como sigue: para una distribución multivariante P, y dada cualquier familia de funciones medibles acotadas $\mathcal{F}$,

$$D^\alpha_{\mathcal{F}}(P) = \{x \mid \exists Q \leq \alpha^{-1}P \mid \forall f \in \mathcal{F}, f(x) \leq \int f \mathrm{d}Q\},$$

$$d_{\mathcal{F}}(x) = \sup\{\alpha \mid x \in D^\alpha_{\mathcal{F}}(P)\}.$$

Esta definición, inspirada por la teoría de órdenes estocásticos, está muy relacionada con la nuestra (y la ha inspirado), para la elección especial

$$\mathcal{Q} = \{Q \mid Q \leq \alpha^{-1}P\}.$$

La principal diferencia es que requiere la existencia de una distribución Q para la que todas las desgualdades se cumplan simultneamente. Eso hace que sea más difícil de estudiar analíticamente.

Teniendo en cuenta que cada función $f \in \mathcal{F}$ puede reemplazarse por $A = f/\|f\|_\infty + .5$, que es un suceso difuso, sin que la familia $D^\alpha_{\mathcal{F}}(P)$ cambie, puede comprobarse que una familia general $\mathcal{F}$ no alcanza mayor generalidad que una familia de sucesos difusos.

Hemos estudiado la definición general de $\mathbf{C}^\alpha_{\mathcal{A}}(\mathcal{Q})$ en [11]. En concreto, [11, Section 7] compara exhaustivamente la definición de Cascos y López-Díaz con la nuestra. Ambas resultan ser equivalentes, en el caso particular $\mathcal{Q} = \{Q \mid Q \leq \alpha^{-1}P\}$ estudiado por ellos, si las funciones en $\mathcal{F}$ son superiormente semicontinuas.

En el trabajo [10] analizamos la relación del concepto de profundidad estadística con las definiciones clásicas de medidas de posición para conjuntos aleatorios.

En este trabajo hemos mostrado la interpretación natural del concepto de región central asociada a una familia de distribuciones, en términos de la Teoría de Conjuntos Difusos:

-Las funciones integrables aparecen como sucesos difusos.

-Sus integrales aparecen como probabilidades de esos sucesos.

-La condición gradual $x/y \geq \alpha$ aparece en la forma de la implicación difusa de Goguen.

-El concepto de *centralidad* de un punto se modela mediante el conjunto difuso C de puntos centrales.

Un estudio más profundo de esta interpretación, y de otras relaciones entre conjuntos difusos y funciones de profundidad estadística, se presentará en [12].

Agradecimientos

Investigación apoyada por el *Ministerio de Educación y Ciencia* (MTM2005-02254 y TSI2005-02511) y el *Gobierno de Aragón* (PM2007-034).

Referencias

[1] I. Cascos (2007). The expected convex hull trimmed regions of a sample. *Comput. Statist.* **22**, 557–569.

[2] I. Cascos, M. López-Díaz (2005). Integral trimmed regions. *J. Multivar. Anal.* **96**, 404–424.

[3] I. Cascos-Fernández, M. López-Díaz (2004). The zonoid hull of a multivariate probability distribution and its trimmed regions. En: M. López-Díaz, M. Á. Gil, P. Grzegorzewski, O. Hryniewicz, J. Lawry, editores. *Soft methodology and random information systems*, 423–430. Springer, Berlin.

[4] P. J. Huber (1972). Robust statistics: a review. *Ann. Math. Statist.* **43**, 1041–1067.

[5] P. J. Huber, V. Strassen (1973). Minimax tests and the Neyman-Pearson lemma for capacities. *Ann. Statist.* **1**, 251–263.

[6] R. Liu, R. Serfling, D. Souvaine, editores (2006). *Data depth: robust multivariate analysis, computational geometry and applications*. American Mathematical Society, Providence.

[7] R. Y. Liu (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18**, 405–414.

[8] I. Molchanov (2005). *Theory of random sets*. Springer, London.

[9] K. Mosler (2002). *Multivariate dispersion, central regions and depth. The lift zonoid approach*. Lect. Notes Statist. **165**. Springer, Berlin.

[10] P. Terán. Classical location estimators for random sets in the light of statistical depth. *Enviado para su publicación.*

[11] P. Terán. Integral central regions and integral depth. *Enviado para su publicación.*

[12] P. Terán. Statistical depth and fuzzy sets. *En preparación.*

[13] J. W. Tukey (1975). Mathematics and the picturing of data. *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974)*, 523–531. Canad. Math. Congress, Montreal.

[14] L. A. Zadeh (1968). Probability measures of fuzzy events. *J. Math. Anal. Appl.* **23**, 421–427.

Control estadístico de procesos imprecisos

Ana Belén Ramos Guajardo[1] **Ángela Blanco Fernández**[2]

[1] European Centre for Soft Computing, anabelen.ramos@softcomputing.es
[2] Departamento de Estadística e I.O., Universidad de Oviedo, blancoangela.uo@uniovi.es

Resumen

En este trabajo se presenta un método de control estadístico de procesos utilizando técnicas de contraste de hipótesis en el contexto de datos imprecisos. Con el fin de analizar la estabilidad de un proceso modelado a través de una variable aleatoria difusa, el estudio se centrará en el valor medio (difuso) y la variabilidad (valor real) de dicha variable.

Palabras Clave: Control estadístico de procesos, variable aleatoria difusa, test de hipótesis.

1 Introducción

El control estadístico de procesos (CEP), introducido por Shewart (ver [11]), es una metodología de trabajo destinada a asegurar y controlar que un proceso responde de forma estable, predecible y dentro de las tolerancias establecidas y a elaborar criterios de decisión para la detección de desviaciones y el inicio de acciones correctivas (ver, por ejemplo, [12]).

Las características que definen la calidad de un conjunto de valores dados están relacionadas con determinadas magnitudes medibles que se denominan variables de control y que deben tener valores bien definidos y varianzas limitadas.

Las causas de variabilidad pueden ser clasificadas en dos grupos. El primero de ellos se basa en las causas que se asumen como parte inherente al proceso y no necesitan ser estudiadas pues el coste que esto supondría no tendrá una repercusión proporcional en la mejora de la calidad. Se denominan causas no asignables o causas comunes. En el segundo grupo se encuentran las causas que provocan variaciones sobrevenidas, sistemáticas o aleatorias, cuya magnitud es crítica para el resultado de la actividad y que, por tanto, deben ser investigadas. Se denominan causas asignables o causas especiales.

El objetivo del CEP es analizar la estabilidad de un proceso, basado en el control de las causas asignables anteriormente descritas. Un método para conseguir este fin consiste en confirmar la hipótesis de que el valor medio μ y la desviación estándar σ de una variable X, la cual caracteriza el funcionamiento del proceso, permanecen constantes en el transcurso del tiempo. En definitiva, se trata de realizar un test de hipótesis sobre el valor de la media y la variabilidad del proceso.

Se pretende, por tanto, establecer un criterio estadístico que nos permita verificar las dos hipótesis siguientes:

- La media no ha cambiado: $\mu_0 = \mu_i$.
- La variabilidad no ha cambiado: $\sigma_0 = \sigma_i$.

Dada una variable de control X, los contrastes de hipótesis anteriores se llevarán a cabo haciendo uso de estadísticos basados en la media y en la varianza de esa variable.

En general el control se realizará mediante la extracción periódica de una muestra de n elementos, es decir, la realización de n medidas suficientemente próximas en el tiempo (en la escala temporal del proceso) para garantizar que las mismas se realizan en un estado dado del mismo.

El valor de la media y la varianza del proceso pueden ser conocidos a priori, por nuestro conocimiento histórico del mismo, o estimarse a partir de un conjunto inicial de k muestras.

Gran parte de los problemas reales en los que interviene el control estadístico de procesos son modelados

mediante variables aleatorias reales (ver, por ejemplo, [10]). Sin embargo, en muchas ocasiones los modelos clásicos no son suficientes ya que no recogen la imprecisión que a veces subyace en la realidad y que se puede expresar a través de los llamados datos experimentales imprecisos o valores difusos.

En este trabajo el objetivo es analizar la posibilidad de imitar los métodos clásicos cuando la variable de control es una variable aleatoria difusa (VAD) y estudiar ejemplos reales que responden a esta situación.

2 Preliminares

Sea $\mathbb{R}^{\mathbf{p}}$ el espacio Euclídeo p-dimensional con la norma usual $\|\cdot\|$. Se denota por $\mathcal{K}_c(\mathbb{R}^{\mathbf{p}})$ la clase de subconjuntos no vacíos, compactos y convexos de $\mathbb{R}^{\mathbf{p}}$ y por $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$ la clase de conjuntos difusos compactos y convexos de $\mathbb{R}^{\mathbf{p}}$, es decir, $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}}) = \{U : \mathbb{R}^{\mathbf{p}} \to [0,1] \mid U_\alpha \in \mathcal{K}_c(\mathbb{R}^{\mathbf{p}}) \quad \forall \alpha \in [0,1]\}$ (donde U_α es el α-nivel del conjunto difuso U para todo $\alpha \in (0,1]$, y U_0 es la clausura del soporte de U).

El espacio $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$ puede dotarse de una estructura semilineal por medio de la *suma* y el *producto por escalares* definidos aplicando el principio de extensión de Zadeh [13]. Para cada α-nivel, esta aritmética coincide con la suma de Minkowski y el producto por escalares (es decir, $(U \oplus V)_\alpha = U_\alpha + V_\alpha = \{u+v \,|\, u \in U_\alpha, v \in V_\alpha\}$ y $(\lambda U)_\alpha = \lambda U_\alpha = \{\lambda\, u \,|\, u \in U_\alpha\}$ para todo $U, V \in \mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$, $\lambda \in \mathbb{R}$ y $\alpha \in [0,1]$). También es útil considerar la *diferencia de Hukuhara* de $U, V \in \mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$, $U -_H V$, definida (si existe) como el elemento $W \in \mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$ tal que $U = V \oplus W$.

Otro elemento fundamental a tener en cuenta es la función soporte (ver, por ejemplo, Klement *et al.* [5]). La *aplicación función soporte* se define como

$$s : \mathcal{F}_c(\mathbb{R}^{\mathbf{p}}) \longrightarrow \mathcal{L}(\mathbb{S}^{\mathbf{p-1}} \times [0,1])$$

con $s(U)$ la función soporte de U, s_U, tal que

$$s_U(u,\alpha) = \sup_{w \in U_\alpha} \langle u, w\rangle, \quad u \in \mathbb{S}^{\mathbf{p-1}},\ \alpha \in [0,1],$$

donde $\langle\cdot,\cdot\rangle$ es el producto interno usual en $\mathbb{R}^{\mathbf{p}}$ y $\mathbb{S}^{\mathbf{p-1}}$ denota la esfera unidad en $\mathbb{R}^{\mathbf{p}}$, es decir, $\mathbb{S}^{\mathbf{p-1}} = \{u \in \mathbb{R}^{\mathbf{p}} | \, \|u\| = 1\}$.

La función soporte preserva la estructura semilineal de $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$, es decir, dados $U, V \in \mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$, $\lambda > \mathbf{0}$, se verifica $s_{U+V} = s_U + s_V$, y $s_{\lambda U} = \lambda s_U$. Además, si exite la diferencia de Hukuhara $U -_H V$, entonces $s_{U -_H V} = s_U - s_V$.

La función soporte permite identificar isométricamente el espacio $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$ con un cono convexo del espacio de funciones integrables Lebesgue $\mathcal{L}(\mathbb{S}^{\mathbf{p-1}} \times [0,1])$ (ver, por ejemplo, [5] and [3]) sobre la base de diferentes métricas. Por ejemplo, Körner and Näther [7] consideran una familia de métricas generalizadas D_K en $\mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$, dadas por

$$[D_K(U,V)]^2 = \int_{(\mathbb{S}^{(\mathrm{p-1})} \times [0,1])^2} (s_U(u,\alpha) - s_V(u,\alpha))(s_U(v,\beta) - s_V(v,\beta)) dK(u,\alpha,v,\beta)$$

donde K es un núcleo simétrico y definido positivo.

La familia de métricas D_K representa la distancia tipo L_2 genérica en el espacio de Banach $\mathcal{L}(\mathbb{S}^{\mathbf{p-1}} \times [0,1])$, lo que permite expresar D_K en términos del producto escalar (o la norma asociada) en este espacio como

$$D_K^2(U,V) = \langle s_U - s_V, s_U - s_V\rangle_K = \|s_U - s_V\|_K^2.$$

Dado el espacio de probabilidad $(\Omega, \mathcal{A}, \mathcal{P})$, una *variable aleatoria difusa* (VAD) es una aplicación D_K-medible Borel $\mathcal{X} : \Omega \longrightarrow \mathcal{F}_c(\mathbb{R}^{\mathbf{p}})$. Esta definición es equivalente al concepto clásico introducido por Puri & Ralescu [8] (ver [2] y [7]).

Si $\|\mathcal{X}_0\| = \sup_{x \in \mathcal{X}_0} \|x\| \in \mathcal{L}^1(\,\Omega\,, \mathcal{A}, \mathcal{P})$, se puede definir el *valor medio difuso* de $\mathcal{X}$, $E(\mathcal{X})$, como el único conjunto difuso tal que $(E(\mathcal{X}))_\alpha$ se identifica con la integral de Aumman del conjunto aleatorio $\mathcal{X}_\alpha$, para todo $\alpha \in [0.1]$ (ver [8]). En el caso particular $p = 1$, se cumple que $(E(\mathcal{X}))_\alpha = [E(\inf(\mathcal{X}_\alpha), E(\sup(\mathcal{X}_\alpha)]$ para todo $\alpha \in [0,1]$.

Además, si $E(\|\mathcal{X}_0\|^{\mathbf{2}}) < \infty$, se puede definir la *varianza* de $\mathcal{X}$ (ver [7]) como el valor real

$$\sigma^2_{\mathcal{X}} = E\Big(D_K^2(\mathcal{X}, E(\mathcal{X}))\Big).$$

Dada una muestra aleatoria simple $\{\mathcal{X}_i\}_{i=1}^n$ obtenida a partir de $\mathcal{X}$, se considera la *media muestral* de $\mathcal{X}$, dada por $\overline{\mathcal{X}}_n = (\mathcal{X}_1 + \ldots + \mathcal{X}_n)/n$, como el estimador difuso de $E(\mathcal{X})$, y la *varianza muestral* definida por $\widehat{\sigma}^2_{\mathcal{X}} = \frac{1}{n}\sum_{i=1}^n D_K^2(\mathcal{X}_i, \overline{\mathcal{X}}_n)$ como estimador de $\sigma^2_{\mathcal{X}}$.

Para ilustrar el comportamiento del método propuesto, se analizará un caso real cuyos datos son *números difusos triangulares*. Éstos conforman un caso particular de conjuntos difusos, con $p = 1$, y se denotan por $A = Tri(A^I, A^m, A^S)$, donde $A^I, A^m, A^S \in \mathbb{R}$ son tales que A^I y A^S son el ínfimo y supremo del 0-corte, y $[A^m, A^m]$ es el 1-corte de A (ver Figura 1).

3 Procedimiento para el control estadístico de procesos imprecisos

Un proceso con datos imprecisos viene determinado por una VAD $\mathcal{X}$ que actúa como variable de control

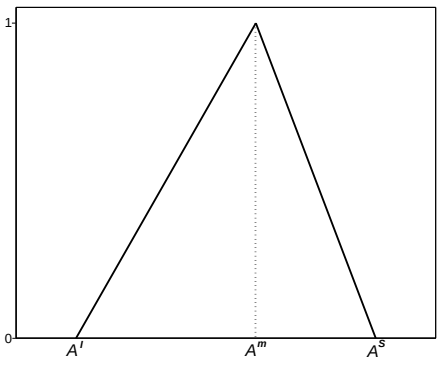


Figura 1: Número difuso triangular A.

caracterizando el funcionamiento del proceso a lo largo del tiempo.

El valor medio y la varianza de $\mathcal{X}$ pueden ser conocidos de antemano (bien porque estén prefijados por un experto o bien por el conocimiento histórico del problema). En otro caso, como etapa preliminar al análisis de la estabilidad del proceso, dichos valores deben ser estimados a partir de una muestra que se supone tomada en condiciones de estabilidad y que se denomina *muestra de control*.

Una vez fijadas las condiciones de estabilidad, resumidas en los valores de referencia $E_0(\mathcal{X})$ y σ_0^2, el procedimiento de control consiste en extraer una muestra del proceso cada cierto periodo de tiempo. El método de extracción de las muestras es acumulativo, es decir, la muestra extraída en el intervalo de tiempo $[0,t]$ es la unión de la muestras tomadas en los intervalos $[0,s]$ y $[s,t]$, donde $s<t$.

Para cada periodo $[0,t]$, se analiza si los correspondientes momentos muestrales se mantienen 'próximos' a los de referencia, con un cierto nivel de confianza. En este trabajo se propone el uso de técnicas de contraste de hipótesis para llevar a cabo este análisis.

Dada una muestra del proceso $\{\mathcal{X}_i\}_{i=1}^n$ obtenida a partir de $\mathcal{X}$ en el intervalo de tiempo $[0,t]$, se contrastan simultáneamente las siguientes hipótesis:

$$H_0^1 : E(\mathcal{X}) = E_0(\mathcal{X}) \text{ frente a } H_1^1 : E(\mathcal{X}) \neq E_0(\mathcal{X}) \quad (1)$$

$$H_0^2 : \sigma_{\mathcal{X}}^2 = \sigma_0^2 \text{ frente a } H_1^2 : \sigma_{\mathcal{X}}^2 \neq \sigma_0^2 \quad (2)$$

La resolución de ambos contrastes ha sido obtenida en [4] y [9], respectivamente, mediante el empleo de técnicas bootstrap utilizando los siguientes estadísticos de contraste:

$$T_1 = \frac{D_K^2(\overline{\mathcal{X}}_n, E(\mathcal{X}))}{\widehat{\sigma}_{\mathcal{X}}^2} \quad (3)$$

$$T_2 = \frac{\sqrt{n}\left(\widehat{\sigma}_{\mathcal{X}}^2 - \sigma_{\mathcal{X}}^2\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(D_K^2\left(\mathcal{X}_i, \overline{\mathcal{X}}_n\right) - \widehat{\sigma}_{\mathcal{X}}^2\right)^2}} \quad (4)$$

Dado un nivel de significación β, las regiones críticas asociadas a (1) y (2) son de la forma $\{T_1 > c_\beta^1\}$ y $\{|T_2| > c_\beta^2\}$, respectivamente.

3.1 Resolución bootstrap

A diferencia de los métodos desarrollados en los contrastes clásicos para la media y la varianza de una variable aleatoria real, en los que generalmente se conoce la distribución exacta del estadístico, en el contexto difuso actualmente sólo se conocen aproximaciones asintóticas a tal distribución (ver, por ejemplo, [6] y [7]).

Las técnicas bootstrap que se emplean en [4] y [9] para la resolución de estos contrastes no necesitan del conocimiento de la distribución exacta o límite del estadístico y además mejoran, en general, la velocidad de convergencia en comparación con las técnicas asintóticas.

Dada $\mathcal{X}$ la VAD control de un proceso y una muestra aleatoria simple $\{\mathcal{X}_i\}_{i=1}^n$ a partir de $\mathcal{X}$, el procedimiento bootstrap para la resolución de (1) (análogamente para (2)), se puede resumir en los siguientes pasos:

i) Calcular el valor del estadístico T_1 para esa muestra.

ii) Obtener una muestra bootstrap, $\{\mathcal{X}_i^*\}_{i=1}^n$ a partir de la anterior y calcular el valor del estadístico bootstrap
$$T_1^* = \frac{D_K^2(\overline{\mathcal{X}}_n, \overline{\mathcal{X}}_n^*)}{\widehat{\sigma}_{\mathcal{X}^*}^2}$$

iii) Repetir ii) un elevado número B de veces, obteniendo el conjunto de estimadores bootstrap $\{T_1^* \dots T_B^*\}$.

iv) Aproximar el p-valor del contraste (1) por la proporción de valores en $\{T_1^* \dots T_B^*\}$ mayores que T_1.

Observación 3.1 Para el contraste (2), por la forma de su región crítica, en el paso iv) se aproxima el p-valor por la proporción de valores en $\{T_1^* \dots T_B^*\}$ cuyo valor absoluto sea mayor que el de T_2.

3.2 Procedimiento de control del proceso

Dado un proceso modelado a través de la VAD $\mathcal{X}$ el objetivo es analizar su estabilidad a lo largo del tiempo.

Para ello se propone un método de control basado en contrastes de hipótesis para la media y la varianza de $\mathcal{X}$, del siguiente modo:

1. Fijar los valores de referencia $E_0(\mathcal{X})$ y σ_0^2 bajo condiciones de estabilidad del proceso.
   - 1.i) Prefijados por un experto o conocidos históricamente.
   - 1.ii) Estimados a partir de una muestra de control $\{\mathcal{X}_i\}_{i=1}^{k}$ que se asume obtenida bajo estabilidad. Calculados su media y varianza muestrales, se define $E_0(\mathcal{X}) = \overline{\mathcal{X}}$ y $\sigma_0^2 = \widehat{\sigma}^2$.
2. Control del proceso en el intervalo de tiempo $[0, t]$:
   - 2.i) Extraer una m.a.s. de $\mathcal{X}$ en ese tiempo, $\{\mathcal{X}_i\}_{i=1}^{n_t}$.
   - 2.ii) Resolver los contrastes (1) y (2) mediante el algoritmo bootstrap propuesto en la Sección 3.1.
   - 2.iii) Si el p-valor obtenido en ambos contrastes supera el nivel de significación deseado, se puede concluir que el proceso se mantiene estable en el tiempo $[0, t]$ y se continua el control en 3. En otro caso, si alguna de las dos hipótesis nulas se rechaza, se debe detener el control estadístico y analizar las causas de inestabilidad que provocaron ese rechazo.
3. Control del proceso en el intervalo de tiempo $[0, s]$, $s > t$:
   - 3.i) Extraer una m.a.s. de $\mathcal{X}$ en el tiempo $[t, s]$, $\{\mathcal{X}_i\}_{i=n_{t+1}}^{n_s}$. Tomar como m.a.s. en $[0, s]$ la unión de ésta con la obtenida en 2.i), denotada por $\{\mathcal{X}_i\}_{i=1}^{n_s}$.
   - 3.ii) Considerar $t = s$ y volver al paso 2.ii).

4 Ejemplo ilustrativo

Con el objeto de ilustrar en la práctica el procedimiento propuesto en las secciones precedentes, se muestra su aplicación sobre un ejemplo real.

Se pretende analizar la variación de la concentración ($\mu m/cm^3$) de partículas PM10 en el aire en la ciudad de Madrid entre los años 2003 y 2007. Los datos se encuentran disponibles en *www.mambiente.munimadrid.es*.

El Ministerio de Medio Ambiente español define las partículas PM10 como aquellas partículas de polvo, cenizas, hollín, metal, cemento o polen, dispersas en la atmósfera, cuyo diámetro varía entre 2.5 y 10 μm.

Se considera el proceso consistente en la medición de la concentración de esas partículas semanalmente, modelado a través de una VAD triangular $\mathcal{X}$ cuyos datos $\mathcal{X}_i = (X_i^I, X_i^m, X_i^S)$ vienen dados por la concentración mínima y máxima de partículas PM10 alcanzada durante la semana (X_i^I y X_i^S, respectivamente) y la concentración media semanal (X_i^m).

Se toma como muestra control los valores semanales correspondientes al año 2002, obteniendo como valores de referencia del proceso $E_0(\mathcal{X}) = (23.1385, 38.25, 57.1819)$ y $\sigma_0^2 = 90.9729$. Aunque los datos pertenecen a una serie temporal, la correlación no es relevante y se puede suponer la independencia entre las observaciones.

Se comienza el análisis de la estabilidad del proceso a partir de la primera semana de 2003, ejecutando el algoritmo de control cada cuatro semanas.

Como resultado, a un nivel de significación $\beta = 0.05$, se puede admitir la estabilidad del proceso hasta el período de tiempo [0,52 semanas] (que corresponde a enero-diciembre de 2003), en el que el p-valor asociado al contraste (2) es 0.0376. Por tanto, se debe detener el control y analizar las causas del rechazo de H_0^2.

Comparando los valores muestrales para ese intervalo de tiempo, $\overline{\mathcal{X}} = (23.6713, 41.2192, 62.8533)$ y $\widehat{\sigma}_{\mathcal{X}}^2 = 165.6341$, con los correspondientes valores de referencia del proceso, se puede observar un aumento significativo de ambas magnitudes. En la Figura 2 se muestra la evolución del proceso entre enero de 2002 y diciembre de 2003. Los datos se representan mediante líneas verticales cuya longitud corresponde a la amplitud del 0-nivel (el intervalo $[X_i^I, X_i^S]$) y la cruz al valor que define el 1-nivel (X_i^m). Así mismo, se muestran las medias difusas de referencia (año 2002, color azul) y la correspondiente al año 2003 (color rojo), pudiendo comprobarse el aumento notado anteriormente.

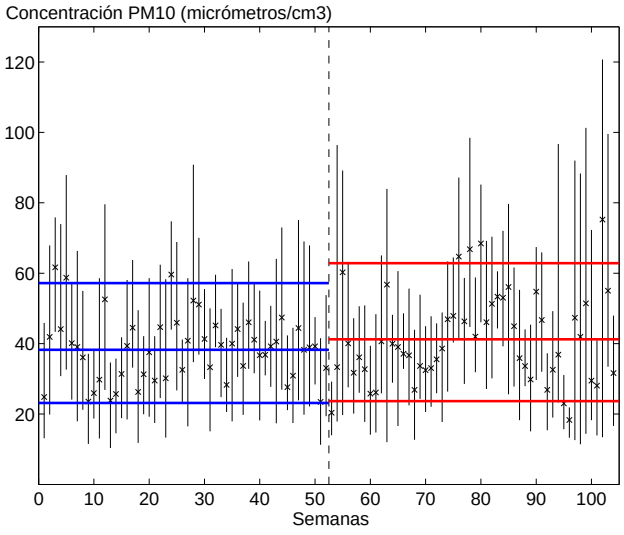


Figura 2: Control del proceso 2002-2003.

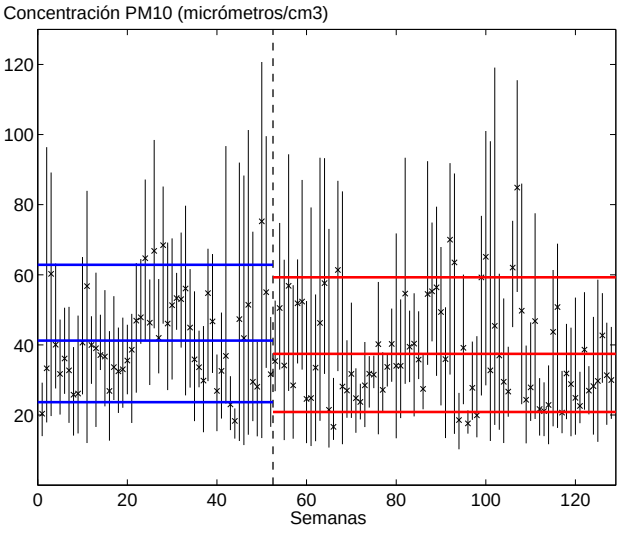


Figura 3: Control del proceso 2004-Junio 2005.

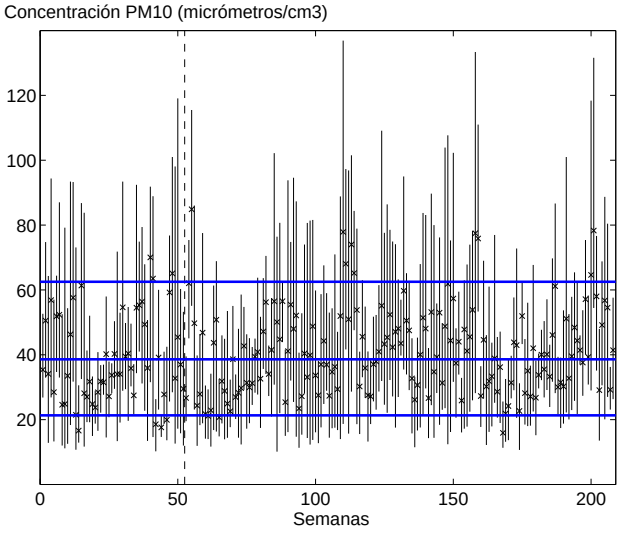


Figura 4: Control del proceso 2004-2007.

En este punto se propone modificar las condiciones de estabilidad, considerando como muestra control la correspondiente al año 2003, con valores de referencia $E_0(\mathcal{X}) = (23.6713,41.2192,62.8533)$ y $\sigma_0^2 = 165.6341$, y comenzar un nuevo análisis de la estabilidad del proceso a partir de la primera semana de 2004.

Los resultados obtenidos desde ese momento permiten asegurar que el proceso se mantiene estable hasta el periodo de tiempo [0,76 semanas] (hasta junio de 2005), en el que el p-valor asociado al test (1) resulta 0.037.

En este caso, la comparación de los valores de referencia con los valores muestrales correspondientes a este periodo, $\overline{\mathcal{X}} = (20.8711, 37.4488, 59.2908)$ y $\widehat{\sigma}^2_{\mathcal{X}} = 201.5077$, indica una disminución de la concentración media de partículas PM10, y un aumento de la variabilidad (aunque ésta no suficientemente elevada para conducir al rechazo de H_0^2). En la Figura 3 se puede observar gráficamente este comportamiento.

Procediendo de forma análoga, se considera como nueva situación de estabilidad la correspondiente al año 2004, con $E_0(\mathcal{X}) = (21.3087,38.5988,62.5117)$ y $\sigma_0^2 = 186.6727$, y se ejecuta el control del proceso desde enero de 2005.

A partir de los resultados obtenidos, se observa que el proceso se mantiene estable ya hasta el final del estudio. La Figura 4 muestra la evolución del proceso en este periodo, donde el valor medio y la varianza de referencia no son rechazados al nivel de confianza del 95%.

Por tanto, se puede concluir que la concentración de partículas PM10 en el aire de la ciudad de Madrid hasta diciembre de 2007 no ha sufrido una variación significativa con respecto a los valores del año 2004.

5 Conclusiones

El método desarrollado en este trabajo permite una extensión de los estudios de control estadístico de procesos al contexto impreciso. Los desarrollos propuestos para el análisis de la estabilidad de un proceso con datos imprecisos se basan en el control de la media difusa y la varianza real de los datos a través de contrastes de hipótesis resueltos mediante técnicas bootstrap.

Además, para ilustrar la utilidad práctica de este método, se ha desarrollado sobre un ejemplo real y se han analizado los resultados obtenidos.

Agradecimientos

Las investigaciones realizadas en este trabajo han sido financiadas parcialmente por el Ministerio Español de Ciencia e Innovación a través del proyecto MTM2006-07501, por el Gobierno del Principado de Asturias mediante el Plan PCTI 2006-2009 (beca BP06-010), y por el centro de investigación internacional European Center for Soft Computing.

Referencias

[1] C. Bertoluzza, N. Corral, A. Salas, "On a new class of distances between fuzzy numbers, " *Mathware & Soft Computing* **2** 71–84, 1995.

[2] A. Colubi, J.S. Domínguez-Menchero, M. López-Díaz, R. Ralescu. " A $D_E[0,1]$ representation of random upper semicontinuous functions," *Proc. Amer. Math. Soc.* vol. 130 pp. 3237–3242, 2002.

[3] P. Diamond, P. Kloeden, "Metric Spaces of Fuzzy Sets," World Scientific, Singapore, 1994.

[4] G. Gonzlez-Rodrguez, A. Colubi, M. Montenegro, M. A. Gil, "Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data", *Fuzzy Sets and Systems*, vol. 157 (19), pp. 2608–2613, 2006.

[5] E. Klement, M.L. Puri, D.A. Ralescu,D.A. "Limit theorems for fuzzy random variables", *Proc. Roy. Soc. London Ser. A 1832* pp. 171–182, 1986.

[6] R. Körner, "An asymptotic α-test for the expectation of random fuzzy variables," *J. Stat. Plann. Inference*, vol. 83, pp. 331–346, 2000.

[7] R. Körner, W. Näther, " On the variance of random fuzzy variables, " In: Bertoluzza, C., Gil, M.A. and Ralescu, D.A. (Eds.) *Statistical Modeling, Analysis and Management of Fuzzy Data*, Physica-Verlag, Heidelberg, pp. 22–39, 2002.

[8] M.L.Puri,D.A. Ralescu, "Fuzzy random variables," *J. Math. Anal. Appl.*, vol. 114, pp. 409–422, 1986.

[9] A.B. Ramos, G. González-Rodríguez, M.A. Gil, A. Colubi. "One-sample bootstrap tests for the variance of a fuzzy random variable". *Submitted*.

[10] B. Ramírez, G. Runger, "Quantitative techniques to evaluate process stability", *Quality Engineering*, vol. 18, pp. 53–68, 2006.

[11] W.A. Shewhart, *Statistical Method from the Viewpoint of Quality Control*, Dover Publications, New York, 1939, 1986.

[12] W.H. Woodall, D.C. Montgomery, "Research issues and ideas in statistical process control", *Journal of Quality Technology*, vol. 31, 4, pp. 376–387, 1999.

[13] L.A.Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, Part 1.", *Inform. Sci.*, vol. 8, pp. 199–249; Part 2. *Inform. Sci.* vol. 8, pp. 301–353; Part 3. *Inform. Sci.*, vol. 9, pp. 43–80, 1975.