INFLUENCE OF FUZZY PARTITION GRANULARITY ON FUZZY RULE-BASED SYSTEM BEHAVIOUR *

Oscar Cordón
Dept. Computer Science & A. I.
E.T.S. de Ingeniería Informática
University of Granada
18071 - Granada
e-mail: ocordon@decsai.ugr.es

Francisco Herrera
Dept. Computer Science & A. I.
E.T.S. de Ingeniería Informática
University of Granada
18071 - Granada
e-mail: herrera@decsai.ugr.es

Pedro Villar
Dept. Computer Languages & Syst.
Univ. of Vigo - E.T.S. Informática
Campus Universitario As Lagoas
32004 - Ourense
e-mail: pvillar@uvigo.es

Summary

In this contribution, we will analyse the importance of the fuzzy partition granularity for linguistic variables in the design of Fuzzy Rule-Based Systems.

Keywords: Fuzzy Rule-Based Systems, Data Base, Fuzzy Partition Granularity, Rule Base Learning.

1 INTRODUCTION

A Fuzzy Rule-Based System (FRBS) presents two main components: 1) the Inference System, which puts into effect the fuzzy inference process needed to obtain an output from the FRBS when an input is specified (see [1, 10]), and 2) the Knowledge Base (KB) representing the known knowledge about the problem being solved, composed of the Rule Base (RB) constituted by the collection of fuzzy rules, and of the Data Base (DB) containing the membership functions asociated to the linguistic variables.

The design of the first component has been widely analyzed in the specialized literature, however, the design of the second component seems to be a more difficult decision because the composition of the KB depends directly on the problem being solved. With the aim of solving this problem, in the last few years, many approaches have been presented to automatically learn the RB from numerical information (input-output data pairs representing the system behaviour) taking as a base different techniques, such as: adhoc data covering algoritmhs [12], gradient descent algorithms [9], Clustering algorithms [13], Neural Networks [11] and Genetic Algorithms [2, 6].

However, there is scarce information about the way to derive the DB, and most of the RB learning methods need of the existence of a previous definition for it (although some of them are able to learn both the definitions of the DB and the RB, they do not show usually good behaviour due to the large learning problem complexity). A very common way to proceed involves considering uniform fuzzy partitions with the same number of terms for all the linguistic variables. Triangular or trapezoidal-shaped membership functions are usually considered due to their simplicity.

Therefore, this operation mode (to obtain a previous uniform DB definition and to automatically learn then an RB) makes the DB have a significant influence on the FRBS performance. This is why some authors try to refine the preliminary DB definition considered once the RB have been derived. To put this into effect, a tuning process considering the whole KB obtained (the preliminary DB and the derived RB) is used a posteriori, to adjust the membership function parameters to improve the FRBS behaviour (for some examples of these kinds of methods, based on Neural Networks and Genetic Algorithms, refer to [2, 7, 9]). Nevertheless, the tuning process usually only adjusts the membership function shapes and not the number of linguistic terms in each fuzzy partition, which remains fixed from the begining of the design process.

Although at first sight, the selection of the fuzzy partition granularity does not seem to be a DB design task as important as others, such as the choice of the membership function shapes for the linguistic terms, it is of significant importance. It plays an important role in many characteristics of the FRBS, such as the *accuracy* in Fuzzy Modelling or the *smoothness* in Fuzzy Control.

The granularity of the input variables specifies the maximum number of fuzzy rules that may compose the RB, thus having a strong influence on aspects such us:

- The *complexity of the rule learning*, a very large

*This research has been supported by CICYT TIC96-0778

number of possible fuzzy rules makes it more complex.

- The *interpretability of the FRBS*, a desirable characteristic in some problems, such us in Fuzzy Linguistic Modelling, which is difficult to achieve when the RB presents a large number of rules.

- The *accuracy of the FRBS*, which is directly related to the granularity of the fuzzy input and output spaces.

The aim of this contribution is to analyse the influence of the fuzzy partition granularity on the FRBS performance. To be precise, we will try to give an answer to the question: *is it a good operation mode to consider uniform fuzzy partitions with the same number of labels for all the linguistic variables?*.

To do so, we will work with different RB learning methods and we will compare its behavior when considering DBs with a different number of linguistic terms for each linguistic variable. The membership functions considered will always be triangular-shaped, symmetrical and uniformly distributed, thus making the granularity of the fuzzy partitions be the only parameter of the DB having influence on the learnt RB and, consequently, on the final FRBS behaviour.

2 FRBSs WITH THE SAME NUMBER OF LABELS FOR EACH VARIABLE

As we have said, the DB is normally defined by choosing an equal number of linguistic terms for all the variables, and by considering uniform fuzzy partitions in the variable universe of discourse for these labels. This choice is not guided by any specific characteristic of the problem, neither by any general rule.

We are going to consider three automatic learning methods of FRBSs:

- Two ad-hoc data covering learning methods:

  - The fuzzy rule generation method of Wang and Mendel [12] (it will be denoted **WM**).

  - One adaptation of the fuzzy clasification rule generation method presented in [8], that makes the process able to deal with rules with fuzzy consequent, which can be found in [3]. This method considers the n-dimensional table representation for the RB to generate. On every cell of this table the subset of the input-output data pairs belonging to the corresponding input fuzzy subspace is considered. The consequent associated to the rule will be the output variable label that maximizes some covering criterion over the training set. This method will be denoted **EGM** (Explorative Generation Method).

- The third method is called Descriptive Mogul [2] (it will be denoted **D-Mogul**). It uses an ad-hoc generation process plus a genetic selection process based on the MOGUL methodology [5]. There is also a third step, a genetic tuning process, that will not be used in order to compare the three methods in the same conditions, that is to say, maintaining the initial uniform fuzzy partitions.

An electrical network distribution problem in northern Spain [4] will be used as benchmark. The system tries to estimate the lenght of the low voltage line installed in a determined village. The problem has three variables, two input variables (the population and the radius of the village), and one output variable (the lenght of the installed line). We will design different FRBSs according to the learning methods and the fuzzy partitions. To compare their behaviour, we have randomly divided the set of data pairs into two subsets comprising 396 and 99 examples, denoted *training set* and *test set* respectively. The former will be considered by the three learning methods to derive de RB composition, while the later will be used to evaluate the prediction ability of the generated fuzzy models.

In this part of the study, the three learning methods were run considering all the posibilities, with the constraint that all the variables should have the same number of labels. We used the interval {3-9} as posible values for the number of linguistic terms. Therefore each method was run seven times. In order to compare their behaviour, we will use the mean square error of the FRBS over the training set (MSE_{tra}) and the test set (MSE_{tst}). The results of the study are shown in table 1 (where the best MSE_{tst} value found in each method appears in bold type). The analysis of these results lead us to point out two main conclusions:

- The learning methods design the FRBS with best behaviour (best MSE_{test}) using a different value for the number of variable labels (six for WM and D-Mogul and eight for EGM).

- The differences among the FRBS obtained are significant enough to justify the fact that the fuzzy partition granularity is an important task, that must be adequately analysed before starting the RB learning.

On the other hand, it is interesting to observe that an excessively high number of labels can create an overlearning problem. Particulary, considering the WM and D-Mogul methods, the FRBSs with best MSE_{tra} use nine labels. However, the value of the MSE_{tst} in both cases is significantly worse than the value obtained with the FRBS with six labels.

Table 1: Results with the same number of labels.

		WM	EGM	D-Mogul
3	MSE_{tra}	594276,3	322227,6	186172,7
	MSE_{tst}	626566,8	293986,9	162589,4
4	MSE_{tra}	301732,0	292714,5	200628,4
	MSE_{tst}	270747,4	270349,8	180553,0
5	MSE_{tra}	298446,0	329726,2	166484,8
	MSE_{tst}	282058,1	306325,7	170550,1
6	MSE_{tra}	239563,0	317516,6	161810,5
	MSE_{tst}	**194842,8**	311065,8	**157403,3**
7	MSE_{tra}	222622,7	267923,9	167621,1
	MSE_{tst}	240018,2	249523,8	207597,6
8	MSE_{tra}	241716,7	199421,3	149415,4
	MSE_{tst}	216651,6	**180000,4**	168025,1
9	MSE_{tra}	197613,4	201272,8	148068,6
	MSE_{tst}	283645,5	224805,7	205396,9

According to these conclusions, an interesting operation mode to design an FRBS is to run the learning method as many times as posible values of number of labels considered, maintaining this value equal for all the variables. Working in this way, we could find the FRBS with best behaviour only with seven runs (3-9 labels). The cost of this study is relatively low, although it should be considered that some kinds of methods have a run time that grows exponentially with the number of labels.

In the following Section, we will consider that the variables can have a different number of labels.

3 FRBSs WITH ANY NUMER OF LABELS IN EACH VARIABLE

In this part, we are going to analyse the FRBS obtained by considering all the posibilities of label number values for each individual variable. The study has been achieved with the ad-hoc data covering methods (WM, EGM), because we pretend to find the granularity with best behaviour and both are deterministic methods. Trying this study with methods that can give a different FRBS in different runs (probabilistic ones as D-Mogul) is complicated, because we could find the best granularity for one initial seed, but we could not state that we have found the absolute optimal granularity for this problem and learning method. Moreover, these non deterministic methods require a long run time. In our case, considering the said interval (3-9 labels), it would be needed $\mathbf{7^n}$ runs, being **n** the number of variables of the problem and **7** the number of posible values of labels.

The best results of this study are shown in table 2. We can observe that the fuzzy partition granularity that shows the best results (both MSE_{tst} and MSE_{tra}), is different in the two methods. The reason is that each method uses in a different form the information contained in the DB during the learning process. Therefore, we can draw the following consequence:

> *There is not an "universal" fuzzy partition granularity, that can be used with good behaviour for all the FRBS learning method.*

Besides, if we compare these results with the ones obtained in the previous part, we can see that the best MSE found (both MSE_{tst} and MSE_{tra}), is much better than the best MSE values found for the FRBS with the same number of labels in each variable. The improvement is so important as to justify a search for the optimal granularity, or at least, for a very good one.

Table 2: Best results with any number of labels.

		WM	EGM
Best Result in MSE_{tra}	N. Lab.	6 9 9	8 8 6
	MSE_{tra}	186904,3	192498,2
	MSE_{tst}	264896,5	167731,5
Best Result in MSE_{tst}	N. Lab.	9 6 9	7 6 7
	MSE_{tra}	202370,9	210983,0
	MSE_{tst}	146355,0	152412,4

Obviously, the problem of this search is the high number of runs needed. Particulary, in one problem with five input variables and one output variable, we would have to run **117649** times (7^6, considering the said interval) the learning method. Of course, if the learning method is inherently slow, the search problem is highly time consuming, and can be almost impossible. Besides, as discussed in this Section, the study would have a relative effect if it is achieved with a non deterministic method. However, with a deterministic method as fast as WM, the exhaustive study can be performed in a reasonable time; very fast if the problem has a few variables, with a high profit.

4 CONCLUSIONS

As an important conclusion, we can say:

> *The choice of the fuzzy partition granularity is an important task for the FRBS design, that should be considered since the begining of the design process.*

This conclusion has been proven by the main results obtained with the experiments achieved, and already comented in the previous sections:

- There is not a number of labels that, used in all the variables, allow us to design the FRBS with best behaviour in all the FRBS learning methods.
- There is neither an "absolute" fuzzy partition granularity that generates the FRBS with best behaviour considering a different number of labels in each variable, for all the FRBS learning methods.
- The improvement obtained using a good fuzzy partition granularity is very important.

Considering these results, especially the fact that each learning method has a different behaviour with the same granularity, we can assert the next conclusion:

> *The design of a good DB, in the aspect of fuzzy partition granularity, depends not only on the problem, but also on the RB learning method considered.*

Therefore, assuming that the fuzzy partition granularity is an important choice, it can be interesting to work with efficient methods in order to find the optimal granularity for one problem and one learning method. Nowadays, there are many search techniques that could find the optimal solution, or at least, a very good granularity, with a relatively low cost. Our future work will focus on the objective to design an efficient granularity search method.

References

[1] Cordón, O., Herrera, F., Peregrín, A., Applicability of the fuzzy operators in the design of fuzzy logic controllers, *Fuzzy Sets and Systems* 86(1), 1997, pp. 15-41.

[2] Cordón, O., Herrera, F., A three-stage evolutionary process for learning descriptive and approximative fuzzy logic controller knowledge bases from examples, *International Journal of Approximate Reasoning* 17(4), 1997, pp. 369-407.

[3] Cordón, O., Herrera, F. A Proposal for Improving the Accuracy of Linguistic Modeling, Technical Report DECSAI-98113, Dept. Of Computing Science and A.I., University of Granada, May 1998.

[4] Cordón, O., Herrera, F., Sánchez, A., Solving Electrical Distribution Problems Using Hybrid Evolutionary Data Analysis Techniques,*Applied Intelligence* 10, 1999, pp. 5-24.

[5] Cordón, O., del Jesus, M.J., Herrera, F., Lozano, A., MOGUL: A Methodology to Obtain Genetic Fuzzy Rule-Based Systems Under the Iterative Rule Learning, *International Journal of Intelligent Systems* 14(9), 1999.

[6] González, A., Pérez, R., SLAVE: A Genetic Learning System Based on an Iterative Approach, *IEEE Transactions on Fuzzy Systems* 7(2), 1999, pp. 176-191.

[7] Herrera, F., Lozano, M., Verdegay, J.L., Tuning fuzzy controllers by genetic algorithms, *International Journal of Approximate Reasoning* 12, 1995, pp. 299-315.

[8] Ishibuchi, H., Nozaki, K., Tanaka, H., Distributed representation of fuzzy rules and its application to pattern classification, *Fuzzy Sets and Systems* 52, 1992, pp. 21-32.

[9] Jang, J.R., ANFIS: adaptive-network-based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics* 23(3), 1993, pp. 665-684.

[10] Kiszka, J., Kochanska, M., Sliwinska, D., The influence of some fuzzy implication operators on the accuracy of a fuzzy model - Parts I and II, *Fuzzy Sets and Systems* 15, 1985, pp. 111-128, 223-240.

[11] Takagi, H., Suzuki, N., Koda, T., Kojima, Y., Neural networks designed on approximate reasoning architecture and their applications, *IEEE Transactions on Neural Networks* 3(5), 1992, pp. 752-760.

[12] Wang, L.X., Mendel, J.M., Generating fuzzy rules by learning from examples, *IEEE Transactions on Systems, Man, and Cybernetics* 22, 1992, pp. 1414-1427.

[13] Yoshinari, Y., Pedrycz, W., Hirota, K., Construction of fuzzy models through clustering techniques, *Fuzzy Sets and Systems* 54, 1993, pp. 157-165.