ESTYLF'22

ACTAS

XXI Congreso Español sobre Tecnologías y Lógica Fuzzy

ESTYLF 2022
TOLEDO

5-7 Septiembre 2022

Universidad de Castilla-La Mancha

Actas del XXI Congreso de Tecnologías y Lógica Fuzzy (ESTYLF'22)

Escuela de Ingeniería Industrial y Aeroespacial

Campus de la Fábrica de Armas

Toledo, España

Septiembre 5-7, 2022

Publicador por:
Grupo de Investigación de Sistemas Inteligentes ORETO
Departamento de Tecnologías y Sistemas de Información
Escuela de Ingeniería Industrial y Aeroespacial
Universidad de Castilla-La Mancha
http://eventos.uclm.es/go/estylf22

ISBN: 978-84-09-42613-3

Credits:
Portada: Luis Jiménez Jiménez & Juan Moreno García
Logo: Javier Alonso Albusac Jiménez
LaTeX editor: Juan Moreno García y Luis Jiménez Linares
Actas generadas con LaTeX's 'confproc' package, version 0.8 (optional: by V. Verfaille)

Bienvenida

En nombre del comité organizador os damos la bienvenida al XXI Congreso Español sobre Tecnologías y Lógica Fuzzy que se celebra en la antigua Fábrica de Armas en el Campus de Toledo de la Universidad de Castilla-La Mancha. Toledo, ciudad Patrimonio de la Humanidad, acoge este evento con gran ilusión deseando que se alcancen todos los objetivos esperados. En esta edición se pretende crear un foro de discusión acerca del estado actual y futuro de las tecnologías y la lógica Fuzzy en el marco general de la Inteligencia Artificial, reflexionando sobre las líneas de trabajo que nos puedan conducir al desarrollo de proyectos novedosos y útiles para la sociedad. Finalmente queremos dar las gracias a nuestros patrocinadores, sin su soporte no podríamos haber organizado este evento. Esperamos que disfrutéis de las mesas redondas, plenarias y sesiones organizadas, y, por supuesto, del programa social que con mucha ilusión hemos diseñado. Bienvenidos todos al ESTYL2022.

Presidencia

Juan Moreno García
Luis Magdalena
Susana Montes

Comité Local

Julio Muñoz
David Muñoz-Valero
Luis Rodríguez-Benítez
Luis Jiménez-Linares
Ester del Castillo Herrera

Comité de Programa

Jesús	Alcala-Fdez	UGR
Cristina	Alcalde	UPV/EHU
Jose M.	Alonso	USC
Sergio	Alonso	UGR
Edurne	Barrenechea	UPNA
Senén	Barro	USC
Fernando	Bobillo	UIZAR
Alberto	Bugarín-Diz	USC
Humberto	Bustince	UPNA
Francisco Javier	Cabrerizo	UGR
José M.	Cadenas	UM
Tomasa	Calvo	UAH
Pablo	Carmona	UEX
Juan Luis	Castro	UGR
Pablo	Cordero	UMA
Óscar	Cordón	UGR
Maria Eugenia	Cornejo Piñero	UCA
Susana	Cubillo	UPM
Rocio	De Andres	USAL
Maria José	del Jesús	UJAEN
Miguel	Delgado	UGR
Susana	Díaz	UNIOVI
Jorge	Elorza	UNAV
Javier	Fernandez	UPNA
Mikel	Galar	UPN
José Luis	García-Lapresta	UVA
Luis	Garmendia	UCM
Lluis	Godo	IIIA - CSIC
Daniel	Gómez	UCM
Antonio	Gonzalez	UGR
Luis	Jimenez Linares	UCLM
Maria Teresa	Lamata	UGR

Bonifacio	Llamazares	UVA
Carlos	Lopez-Molina	UPN
Luis	Magdalena	UPM
Maria J.	Martin-Bautista	UGR
Luis	Martínez	UJAEN
Sebastià	Massanet	UIB
Francisco	Mata	UJAEN
Jesús	Medina	UCA
Jose M.	Molina	UC3M
Javier	Montero	UCM
Susana	Montes	UNIOVI
Juan	Moreno García	UCLM
Manuel	Ojeda-Aciego	UMA
Jose Angel	Olivas	UCLM
Antonio	Peregrin	UHU
Hector	Pomares	UGR
Ana	Pradera	URJC
Jordi	Recasens	UPC
Juan Vicente	Riera	UIB
Adolfo	Rodríguez	UNILEON
Rosa	Rodríguez	UJAEN
Francisco P.	Romero	UCLM
Daniel	Sanchez	UGR
Luciano	Sanchez	UNIOVI
José Antonio	Sanz Delgado	UPN
Jesús	Serrano-Guerrero	UCLM
Miguel-Angel	Sicilia	UAH
Vicenc	Torra	U. Skövde
Aida	Valls	URV
José Luis	Verdegay	UGR
Rosa María	Rodríguez	UJAEN
Javier Alonso	Albusac	UCLM

PROGRAMA

1 Lunes 5 de Septiembre.

1 Sesión 1 Aula A:Toma de Decisiones

9 Sesión 1 Aula B:Fundamentos

17 Sesión 2 Aula A:Procesamiento del Lenguaje Natural

Metodología lingüística de toma de decisiones multi criterio para la evaluación de servicios turísticos basada en la valoración de las opiniones del cliente

1st Itzcóatl Bueno
Facultad de Estadística
Universidad Complutense de Madrid
Madrid, España
izcoal8@gmail.com

2st Ramón A. Carrasco
Dpto. de Organización de Empresas y Marketing
Universidad Complutense de Madrid
Madrid, España
ramoncar@ucm.es

3st Carlos Porcel
Dpto. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
cporcel@decsai.ugr.es

4st Enrique Herrera-Viedma
Dpto. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
viedma@decsai.ugr.es

5st Luis G. Pérez
Dpto. de Informática
Universidad de Jaén
Jaén, España
lgonzaga@ujaen.es

6st Francisco Mata
Dpto. de Informática
Universidad de Jaén
Jaén, España
fmata@ujaen.es

***Abstract*—En este trabajo se presenta un resumen del trabajo titulado "A linguistic multi-criteria decision making methodology for the evaluation of tourist services considering customer opinion value" en el que se propone una metodología que integra múltiples técnicas de toma de decisiones y con la que se pretende obtener un ranking de hoteles a partir de las opiniones de sus clientes. La idea es obtener un valor para cada cliente que se calcula mediante el modelo RFH (*Recency - Frequency - Helpfulness*). La información sobre los usuarios se extrae de redes sociales y se gestiona y agrega utilizando el enfoque lingüístico difuso multi granular basado en 2-tuplas. Además, se verificó la funcionalidad de la metodología presentando un caso de negocio aplicándola sobre datos extraídos de TripAdvisor.**

***Index Terms*—Modelado lingüístico difuso, valor de la opinión de clientes, toma de decisión multi criterio, evaluación de servicios turísticos**

I. Introduction

La expansión de las nuevas tecnologías y uso intensivo de aplicaciones en línea provoca que cada vez más los clientes dejen sus comentarios en foros y redes sociales para valorar productos y servicios experimentados. Este tipo de comentarios de los clientes tiene un gran impacto en el proceso de compra o selección de servicios por parte del consumidor, que tiene acceso a una infinidad de opiniones de otros usuarios que han experimentado los productos y servicios con anterioridad.

Estas opiniones se pueden publicar en diferentes formatos. Las encuestas de satisfacción son una herramienta habitual para obtener una valoración global del producto y/o sus características. Sin embargo, no todas las características que influyen en la satisfacción del cliente pueden identificarse mediante cuestionarios o encuestas. Por ello, otro formato cada vez más utilizado a la hora de proporcionar opiniones consiste en que los usuarios las expresen en lenguaje natural. Otro aspecto a considerar es que hay usuarios más influyentes que otros y existen varios modelos para obtener el valor percibido del opinante. Uno de los más relevantes es el modelo RFM (*Recency - Frequency - Monetary*) que valora a un cliente en función de la recurrencia, la frecuencia y el valor monetario de la compra. El modelo RFH es la adaptación del anterior sustituyendo el valor monetario por la utilidad percibida del opinante. Además, se añade el problema de la incertidumbre presente en estas situaciones, que en numerosos ámbitos se ha solucionado con la aplicación de un modelado lingüístico difuso.

Trabajar con tal variedad de fuentes, formatos y consideraciones de forma conjunta aporta mayor valor añadido, pero deriva en dificultades a la hora de agregar todos esos datos y extraer información útil. Para paliar estos problemas, en [1] se presenta una metodología que integra las distintas fuentes de datos mediante el modelado lingüístico multi granular basado en 2-tuplas y calcula un valor sobre los clientes basado en el modelo RFH, que permite cuantificar la influencia de las opiniones de cada uno. La versatilidad de la metodología presentada en [1] permite que se pueda usar en cualquier dominio lo que, además de considerarse una propuesta innovadora, la convierte en altamente aplicable en numerosos casos de negocio.

El resto del resumen presenta la metodología en la sección II, caso de uso y discusión de sus resultados en la sección III y las conclusiones y trabajos futuros en la sección IV.

Trabajo financiado a través de los proyectos PID2019-103880RB-I00/AEI/10.13039/501100011033 y TIN2016-75850-R

II. Metodología

En la Fig. 1 se presenta la visión global de la metodología presentada en el trabajo. Para centrar un caso de negocio específico se optó por el sector turístico, en concreto para obtener un raking de hoteles, por la importancia que tiene y también por la gran cantidad de datos disponibles en línea.

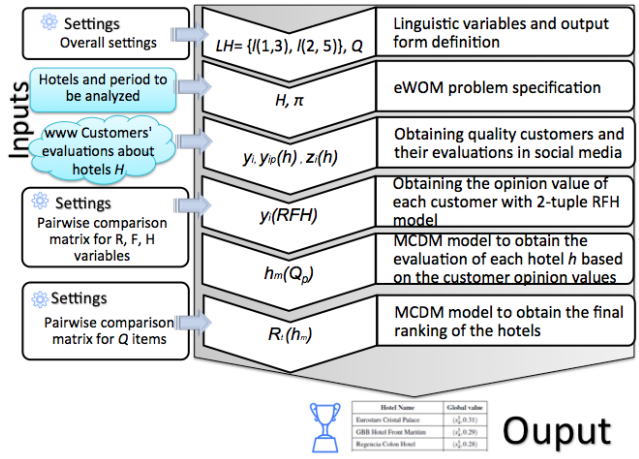


Fig. 1. Esquema global de la metodología.

Antes de proceder con la aplicación de la metodología es necesario obtener los datos de las revisiones proporcionadas y cuestionarios rellenados sobre los hoteles en cuestión. En la fase 1 se tienen en cuenta los ajustes generales, como los tipos de escala o datos recibidos, para establecer la jerarquía lingüística a usar y las preguntas de los cuestionarios recibidos. La fase 2 introduce un conjunto de hoteles sobre los que las preguntas del paso anterior han sido medidas; además, se define un período de análisis que reduce el conjunto de hoteles a los que han recibido opiniones en dicho período. En la fase 3, las evaluaciones proporcionadas por los creadores de opiniones sobre los hoteles filtrados se introducen en las métricas seleccionadas lo que genera: información relacionada con un revisor, valoración que el revisor suministra para cada pregunta en el cuestionario de un hotel dado y su opinión sobre el hotel expresada en lenguaje natural. A partir de los pasos previos, en la fase 4 se obtienen 3 medidas: días transcurridos desde que el revisor publicó su última revisión, frecuencia de publicación de opiniones durante el período analizado y utilidad percibida por otros creadores de opiniones; además, se parametriza la matriz de preferencias en relación a estas tres dimensiones para ponderar lo que es más relevante a la hora de evaluar un revisor. Después se aplica el modelo RFH basado en 2-tuplas para generar como salida de esta fase una valoración global del cliente. En la fase 5 se calcula el valor promedio de las valoraciones del cliente sobre las preguntas de la encuesta para cada hotel. Por último, se usa como entrada la valoración media de cada pregunta en el cuestionario y las opiniones en lenguaje natural que recibe cada hotel para parametrizar una matriz de preferencias sobre la importancia de los aspectos a los que se refiere cada pregunta con respecto a las demás. Entonces se aplica el modelo AHP (*Analytic Hierarchy Process*) sobre dichas entradas y la configuración establecida para obtener el ranking final de hoteles. Gracias a la aplicación del modelo AHP se puede parametrizar la importancia de cada dimensión del modelo RFH para obtener un valor del cliente según la especificación del ámbito de aplicación, así como la importancia del criterio valorado para cada hotel para obtener su valoración global.

III. Caso de uso y discusión de resultados

En el trabajo se incluye un caso de uso en el que se detalla la aplicación de la metodología sobre un data set obtenido a partir del rastreo de revisiones de hoteles de TripAdvisor, usando la API proporcionada por la propia compañía para la extracción de información.

La aplicación de la metodología permite obtener un ranking de los hoteles incluidos en el estudio, pero no basándose directamente en las características de los establecimientos, sino considerando las experiencias personales de sus clientes obtenidas de diferentes fuentes y con distintas escalas o etiquetas lingüísticas. En este sentido, otro punto fuerte de la propuesta es que los datos procedentes de varios sitios web como por ejemplo TripAdvisor, Booking, Expedia, etc. pueden ser combinados con distintas formas de valorar la satisfacción del cliente, adaptándose de manera sencilla con la inclusión de más niveles en la jerarquía.

IV. Conclusiones y trabajos futuros

En este trabajo se ha desarrollado una metodología que permite obtener un ranking en un dominio concreto a partir de las opiniones personales de los clientes, sean cuestionarios o reseñas, lo que mejora la fiabilidad de dicho ranking. Se verifica su funcionalidad sobre un caso de uso real con datos obtenidos de TripAdvisor para obtener un ranking de hoteles analizados.Se calcula un valor de las opiniones de los clientes usando el modelo RFH basado en 2-tuplas e incorporando un modelado lingüístico difuso multi granular que permite adaptarse a datos provenientes de diversas fuentes o distintas escalas sin pérdida de información. Este valor del cliente sirve como ponderación a la hora de tener una valoración media de cada hotel en función de los usuarios que han compartido su experiencia en él. A partir del conjunto de hoteles con la valoración media obtenida, el ranking final se genera aplicando el esquema AHP.

Como trabajos futuros se plantean la incorporación de técnicas más precisas para la extracción de opiniones o la aplicación de la metodología en otros sectores distintos del turístico.

References

[1] I. Bueno, R. A. Carrasco, C. Porcel, G.Kou and E. Herrera-Viedma, "A linguistic multi-criteria decision making methodology for the evaluation of tourist services considering customer opinion value," Applied Soft Computing Journal, vol. 101, 107045, 2021.

Consensus Reaching with Large Group of Citizens

Luis G. Pérez
Department of Computer Science
University of Jaén
Jaén, Spain
lgonzaga@ujaen.es

Francisco Mata
Department of Computer Science
University of Jaén
Jaén, Spain
fmata@ujaen.es

Carlos Porcel
Department of Computer Science
and Artificial Intelligence
University of Granada
Granada, Spain
cporcel@decsai.ugr.es

***Abstract*—Smart cities and Smart Governance are being developed with the aim of improving the life of their citizens. To do so, it is crucial to boost the communication among citizens and the government or administration. Citizen participation based systems grounded in consensus reaching processes play a key role since these systems can achieve a minimum level of agreement before making a decision. However, these solutions faces some issues that should be addressed. One of them is the efficiency of the systems when they are dealing with many citizens. To mitigate this flaw, our approach will use Social Network Analysis to create clusters of citizens to improve the process.**

***Index Terms*—Smart cities, Smart Governance, Consensus Reaching Processes, Social Network Analysis**

I. Introduction

There is no unique definition of the term *Smart Cities*. However, all of them agree that the information and communications technologies play a fundamental role in the development of Smart Cities. These technologies improve the citizens' quality of life, and ensure a sustainable economic, social and environmental development [1].

In the literature, we can find the features that Cities should meet to be considered *Smart* [2]: e-Administration, open data and big data, sensorization and connectivity, security and privacy, transparent decision making and citizen participation.

One of the fields that has undergone drastic changes because of digital and telecommunication technologies in Smart Cities, is their management. Typically, management was performed by a reduced group of experts, politicians, etc. However, Smart cities have introduced the term Smart Governance to describe the management of the city by using digital and telecommunication technologies. Smart Governance should be participatory and involves both citizens as well as other agents presented in the city (public, private, civil, etc.).

Citizen participation systems based on consensus for decision-making problems [3]–[5] can play a key role in order to improve and make easier their participation. As aforementioned in [5], with many citizens, these systems may not be suitable since they may suffer from efficiency problems.

In this contribution, we address this drawback modifying the consensus model scheme presented in [3] by using clustering techniques. With these clusters, we can identify groups of citizens that have already reached a suitable level of consensus and, therefore, they do not need to participate in the consensus process. This approach is organized as follows: in section II, our proposal is presented and in section III we will describe the conclusions and future works.

This result has been supported under the research Action 1 with code EI_TIC1_2021 by the University of Jaén

II. Proposal

In this section we have summarized the main phases of our consensus reaching process. These steps are:

A. Gather citizens' preferences

Each citizen e_i provides its preferences on the set of possible alternatives in $X = \{x_1, \ldots, x_n\}$ obtaining a fuzzy preference relation P_i.

B. Gather trust information and compute citizens' importance

In order to do so, we will use a fuzzy sociometric matrix and the steps explained in [6] to gather the trust information.

Since some citizens may not provide a trust value on a specific citizen, that matrix obtained may be incomplete. In [6] a method was proposed to infer, based on transitivity, the unknown trust values from a citizen to another. At the end of this process, we will obtain the sociometric matrix.

Finally, once we have computed this matrix, we can calculate the weight of citizen e_k as follows:

$$\lambda_k = \frac{C(e_k)}{\sum_{i=1}^{m} C(e_k)}$$

Where $C(e_k)$ is the relative node in-degree centrality index [7] associated with an citizen e_k:

$$C(e_k) = \frac{1}{m-1} \sum_{i=1, j \neq k}^{m} s_{jk}$$

and s_{jk} is the trust degree that citizen e_j assigns to citizen e_j.

C. Compute global consensus degree

The objective of this step is to know if the consensus has been reached. If it has been accomplished, the process can move to the selection phase. Several consensus measurement methods have been suggested [8], [9]. We will follow the calculation suggested in [6]. The steps we should carry out are:

1) For each pair of citizens e_i, e_j $(i < j)$ a similarity matrix SM_{ij} is computed.

2) A consensus matrix CM is computed by aggregating similarity matrix, taking into account the importance weights $w_{ij} \in [0, 1]$ associated with each pair of citizens $(e_i, e_j), i < j$. This importance is obtained from the trust social network.
3) Consensus degree is computed at three different levels: level of pairs of alternatives (cp^{lk}), level of alternatives (ca^l) and level of preference relation obtaining the overall consensus degree cr.

This overall consensus degree cr is compared with a global consensus threshold $\mu_G \in [0, 1]$. If $cr \geq \mu_G$, then the consensus process ends and the group moves on to the selection process; otherwise, we need to perform the next steps.

D. Create clusters

Dividing experts into clusters makes the process more efficient [10]. In [11] the agglomerative hierarchical clustering method is proposed in order to divide citizens into small groups. In this article, the algorithm starts with t clusters where each cluster has one citizen e_i. In each step, it merges the clusters that are the closest. For each new cluster distribution, the consensus level is computed. The optimal solution is the cluster distribution $C = \{C_1, \ldots, C_m\}$ with highest consensus. In aforementioned article, a similarity-trust measure is used. In this contribution, we only take into account the trust information.

E. Compute collective preference relation

The collective preference relation P_C is computed for each pair of alternatives by aggregating citizens' preference relations taking into account the importance of each citizen.

F. Compute overall cluster preference relation and cluster importance for each cluster

To compute the overall cluster preference relation P_{C_j}, each cluster is processed as an isolated group of citizens. The importance of each cluster is the aggregation of the importance of each citizen of the cluster.

G. Compute proximity among collective preference and overall cluster preference relations

A proximity matrix PP_i is computed by measuring the distance between each cluster overall preference relation P_{C_j} and collective preference relation P_C.

H. Identify alternatives and cluster that need to change their opinions

To identify these clusters, we will use a method similar to the one shown in [3] but instead of dealing with experts, we deal with clusters and cluster preference relations.

I. Compute cluster consensus degree and identify groups that needs to improve their consensus degree

This step is only calculated for clusters that were not selected in the previous step but whose consensus degree is not acceptable (i.e., the consensus degree within the cluster needs to be computed and compared with a threshold μ_C).

J. Advice generation process for citizens

Once we have reached this step, we face three kind of clusters: clusters that need to change their overall cluster opinion, clusters that does not need to change their overall cluster opinion but whose inner consensus degree is not suitable and finally, and those ones that does not need a consensus reaching process.

For the first two types of clusters, the advice generation process is a similar one the presented in [3] where each user's opinions are compared to the collective preference relation to decide if they need to modify their judgments. In our approach, for the first category of clusters, we will compare with the collective preference relation P_C, while for the second one, we will use the cluster overall preference relation P_{C_j}.

III. Conclusions and future works

This consensus reaching model faces the problem of dealing with large amount of citizens by creating clusters. Thanks to these clusters, we can limit the consensus reaching process to the clusters that needs to change their opinion. Moreover, in the future, this model can be easily improve to identify the most influential citizen of a clusters, their leaders; and leaders could be used to improve and reduce the amount of advices generated to reach the required consensus.

References

[1] R. D. Orejon-Sanchez, D. Crespo-Garcia, J. R. Andres-Diaz, and A. Gago-Calderon, "Smart cities' development in spain: A comparison of technical and social indicators with reference to european cities," *Sustainable Cities and Society*, vol. 81, 2022.

[2] H. Scholl and M. Scholl, "Smart governance: A roadmap for research and practice," *iConference 2014 Proceedings*, 03 2014.

[3] F. Mata, L. Martinez, and E. Herrera-Viedma, "An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 2, pp. 279–290, 2009.

[4] I. Palomares, L. Martínez, and F. Herrera, "A consensus model to detect and manage noncooperative behaviors in large-scale group decision making," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 3, pp. 516–530, 2014.

[5] F. Mata, A. Verdejo, L. G. Pérez, and C. Porcel, "A preliminary study of a citizen participation system based on consensus for decision-making processes," in *54th Hawaii International Conference on System Sciences, HICSS 2021, Kauai, Hawaii, USA, January 5, 2021*, pp. 1–10, ScholarSpace, 2021.

[6] H. Zhang, I. Palomares-Carrascosa, Y. Dong, and W. Wang, "Managing non-cooperative behaviors in consensus-based multiple attribute group decision making: An approach based on social network analysis," *Knowledge-Based Systems*, vol. 162, pp. 29–45, 2018.

[7] J. Scott and P. Carrington, *The SAGE Handbook of Social Network Analysis*. SAGE Publications, Inc., 2014.

[8] F. Chiclana, J. García, M. J. del Moral, and E. Herrera-Viedma, "A statistical comparative study of different similarity measures of consensus in group decision making," *Information Sciences*, vol. 221, pp. 110–123, 2013.

[9] T. Gonzalez-Arteaga, de Andres Calle R., and F. Chiclana, "A new measure of consensus with fuzzy preference relations: The correlation consensus degree," *Knowledge-Based Systems*, vol. 107, pp. 104–116, 2016.

[10] N. Perony, R. Pfitzner, I. Scholtes, C. Tessone, and F. Schweitzer, "Enhancing consensus under opinion bias by means of hierarchical decision making," *Advances in Complex Systems*, vol. 16, 03 2014.

[11] W. S. A. W. Ahlim, N. H. Kamis, S. A. S. Ahmad, and F. Chiclana, "Similarity–trust network for clustering-based consensus group decision-making model," *International Journal of Intelligent Systems*, vol. 37, no. 4, pp. 2758–2773, 2022.

El Método Lexicográfico para Problemas de Programación Lineal Borrosa Multi-objetivo

Laura Romero
Departamento de Matemática
Universidad de Cienfuegos
Cienfuegos, Cuba
lromero@ucf.edu.cu

Boris Pérez-Cañedo*
Departamento de Matemática
Universidad de Cienfuegos
Cienfuegos, Cuba
bpcanedo@gmail.com

José Luis Verdegay
Departamento de Ciencias de la Computación e IA
Universidad de Granada
Granada, España
verdegay@decsai.ugr.es

***Resumen*—El artículo resume resultados recientes sobre el uso de un método lexicográfico para resolver problemas de programación lineal borrosa multi-objetivo. Los elementos básicos del método se discuten brevemente y se da un ejemplo para ilustrar su funcionamiento.**

***Index Terms*—programación lineal borrosa, número borroso, orden lexicográfico, optimización multi-objetivo**

I. Introducción

La optimización matemática es una herramienta valiosa para la toma de decisiones. Sin embargo, la toma de decisiones no es una tarea sencilla. De hecho, está plagada de incertidumbres del mundo real que surgen en las etapas de análisis, modelización y solución de cualquier problema. Esto es particularmente cierto en el área de los sistemas automatizados de decisión (SAD), donde se atribuye a estos sistemas alguna forma de comportamiento humano similar al de un experto para llevar a cabo tareas complejas que involucran información imprecisa, múltiples objetivos en conflicto, varios tipos de datos, etc.

La optimización borrosa (OB) iniciada por Bellman y Zadeh [1] se ha convertido en una técnica fundamental para la investigación de operaciones, la inteligencia artificial y los SAD. Uno de los modelos de OB que ha sido objeto de investigación reciente es el modelo de programación lineal con números borrosos (NBs) como parámetros y/o variables de decisión (problema de PLB en lo adelante) [2]. Un enfoque generalizado para resolver tales problemas de PLB es transformarlos en problemas convencionales (no borrosos) mediante funciones lineales de ordenación [2]. Sin embargo, las funciones lineales de ordenación pueden transformar dos NBs diferentes en el mismo número real; en consecuencia, usarlas para la PLB no garantiza soluciones con valores objetivos óptimos únicos en el caso uni-objetivo ni soluciones óptimas de Pareto en el caso multi-objetivo [3], [4]. Para evitar estas deficiencias, se pueden usar múltiples índices lexicográficamente para comparar NBs [3], [5]. Sin embargo, hasta hace poco tiempo, no existía un método general para manejar las restricciones de desigualdad borrosas utilizando criterios lexicográficos de ordenación. En [6], los autores propusieron dicho método para el caso uni-objetivo, y luego se usó con un método de escalarización en [3] y [4] para resolver problemas de programación de proyectos y asignación de atraques, respectivamente.

*Autor para correspondencia. Investigación respaldada por los proyectos PID2020-112754GB-I0, MCIN/AEI y B-TIC-640-UGR20 FEDER/Junta de Andalucía.

En este artículo, resumimos los elementos básicos del método lexicográfico para la programación lineal borrosa multi-objetivo (PLBM) presentado en [3]. En aras de la simplicidad, restringimos la discusión posterior a los NBs triangulares (NBTs). La sección II presenta algunas definiciones básicas y el planteamiento del problema. La sección III presenta un método lexicográfico para la PLB en el caso uni-objetivo. En la Sección IV se describe un método ϵ-restringido para resolver el caso multi-objetivo. En la Sección V se presenta un ejemplo ilustrativo. Las conclusiones y líneas futuras de investigación son presentadas en la Sección VI.

II. Definiciones Básicas y Planteamiento del Problema

En este artículo, un NBT se denota como de costumbre mediante la tupla $\tilde{a} = (a, b, c)$, y la igualdad de dos NBTs se define elemento a elemento. El conjunto de todos los NBTs se denota por $\mathcal{T}(\mathbb{R})$. La suma, la multiplicación de dos NBTs y la multiplicación por un escalar se denotan por $\oplus$, $\otimes$ y $\times$, respectivamente.

Definición 1 ([3]): Sea $\tilde{a} \in \mathcal{T}(\mathbb{R})$ arbitrario y f_t ($t = 1, 2, 3$) funciones lineales de los parámetros de $\tilde{a}$ con matriz de coeficientes no singular. Además, sea $\leq_{lex}$ la relación de orden lexicográfica en $\mathbb{R}^3$. Para cualquier $\tilde{a}, \tilde{b} \in \mathcal{T}(\mathbb{R})$, la desigualdad estricta $\tilde{a} \prec \tilde{b}$ se cumple ssi $(f_t(\tilde{a}))_{t=\overline{1,3}} <_{lex} (f_t(\tilde{b}))_{t=\overline{1,3}}$; la desigualdad débil $\tilde{a} \preceq \tilde{b}$ se cumple ssi $(f_k(\tilde{a}))_{t=\overline{1,3}} <_{lex} (f_t(\tilde{b}))_{t=\overline{1,3}}$ o $(f_k(\tilde{a}))_{t=\overline{1,3}} = (f_t(\tilde{b}))_{t=\overline{1,3}}$. El modelo matemático de la PLBM es

$$\begin{aligned} \text{mín } & (\tilde{z}_k(\tilde{x})) = \left(\sum_{j \in J} \tilde{c}_{jk} \otimes \tilde{x}_j \right) \\ & \text{para } k \in K := \{1, 2, \ldots, p \geq 2\}, \\ \text{s.a. } & \sum_{j \in J} \tilde{a}_{ij} \otimes \tilde{x}_j \preceq \tilde{b}_i \quad \text{para } i \in I := \{1, 2, \ldots, m\}, \\ & \tilde{x}_j \in \mathcal{T}(\mathbb{R}) \quad \text{para } j \in J := \{1, 2, \ldots, n\}. \end{aligned} \tag{1}$$

Definición 2 ([3]): Una solución factible $\tilde{x}$ del problema (1) se dice que es una solución borrosa óptima de Pareto si, de acuerdo con la Definición 1, no existe otra solución factible

$\tilde{y}$ tal que $\tilde{z}_k(\tilde{y}) \preceq \tilde{z}_k(\tilde{x})$, para $k \in K$, con al menos una desigualdad estricta.

III. Método Lexicográfico para el Caso Uni-objetivo

Supongamos $p = 1$ en el problema (1). Aplicando la Definición 1, obtenemos

$$\begin{aligned} &\text{lex mín } (f_t(\tilde{z}(\tilde{x})))_{t=\overline{1,3}} \\ &\text{s.a. } \left(f_t\left(\sum_{j\in J}\tilde{a}_{ij}\otimes\tilde{x}_j\right)\right)_{t=\overline{1,3}} \\ &\qquad \leq_{lex} \left(f_t\left(\tilde{b}_i\right)\right)_{t=\overline{1,3}} \quad \text{para } i\in I, \\ &\qquad \tilde{x}_j \in \mathcal{T}(\mathbb{R}) \quad \text{para } j\in J. \end{aligned} \tag{2}$$

Para manejar las restricciones lexicográficas, el problema (2) se transforma en el problema de optimización lexicográfica (3), en el que ϵ y L son constantes positivas pequeña y grande, respectivamente.

$$\begin{aligned} &\text{lex mín } (f_t(\tilde{z}(\tilde{x})))_{t=\overline{1,3}} \\ &\text{s.a. } -L\sum_{r=1}^{t-1} y_{ir} + \epsilon y_{it} \\ &\qquad \leq f_t\left(\tilde{b}_i\right) - f_t\left(\sum_{j\in J}\tilde{a}_{ij}\otimes\tilde{x}_j\right), \\ &\qquad f_t\left(\tilde{b}_i\right) - f_t\left(\sum_{j\in J}\tilde{a}_{ij}\otimes\tilde{x}_j\right) \leq L y_{it}, \\ &\qquad y_{it}\in\{0,1\} \quad \text{para } t=1,2,3 \text{ y } i\in I, \\ &\qquad \tilde{x}_j \in \mathcal{T}(\mathbb{R}) \quad \text{para } j\in J. \end{aligned} \tag{3}$$

IV. Método ϵ-restringido Borroso

La siguiente escalarización del problema (1) se propuso en [3] donde se demostró que produce soluciones borrosas óptimas de Pareto, siempre que se resuelva con el método lexicográfico de la Sección III.

$$\begin{aligned} &\text{mín } \tilde{z} \\ &\text{s.a. } \tilde{z}\oplus\sum_{k\in K\backslash\{q\}}\lambda_k\times\tilde{s}_k^+ \\ &\qquad = \tilde{z}_q(\tilde{x})\oplus\sum_{k\in K\backslash\{q\}}\lambda_k\times\tilde{s}_k^-\oplus(-M,0,M), \\ &\qquad \tilde{z}_k(\tilde{x})\oplus\tilde{s}_k^+ = \tilde{\epsilon}_k\oplus\tilde{s}_k^- \quad \text{para } k\in K\backslash\{q\}, \\ &\qquad \tilde{s}_k^- \preceq \tilde{s}_k^+, \tilde{s}_k^-\geq 0, \tilde{s}_k^+\geq 0 \quad \text{para } k\in K\backslash\{q\}, \\ &\qquad \sum_{j\in J}\tilde{a}_{ij}\otimes\tilde{x}_j \preceq \tilde{b}_i \quad \text{para } i\in I, \\ &\qquad \tilde{z},\tilde{s}_k^-,\tilde{s}_k^+,\tilde{x}_j\in\mathcal{T}(\mathbb{R}) \quad \text{para } j\in J \text{ y } k\in K\backslash\{q\}. \end{aligned} \tag{4}$$

En el problema (4), M es una constante grande. De manera similar al método convencional (no borroso) ϵ-restringido, las soluciones borrosas óptimas de Pareto del problema (1) se obtienen resolviendo el problema (4) proporcionando valores para λ_k y $\tilde{\epsilon}_k$.

V. Ejemplo

La Tabla I muestra soluciones óptimas de Pareto del problema biobjetivo

$$\begin{aligned} &\text{mín } (7,10,11)\otimes\tilde{x}_1\oplus(8,10,13)\otimes\tilde{x}_2, \\ &\text{mín } (2,3,4)\otimes\tilde{x}_1\oplus(4,7,12)\otimes\tilde{x}_2, \\ &\text{s.a. } (1,2,4)\otimes\tilde{x}_1\oplus(2,8,10)\otimes\tilde{x}_2 = (3,22,36), \\ &\qquad (2,3,6)\otimes\tilde{x}_1\oplus(4,10,15)\otimes\tilde{x}_2 = (6,29,54), \\ &\qquad \tilde{x}_1,\tilde{x}_2 \text{ son NBTs no negativos,} \end{aligned}$$

obtenidas utilizando el método ϵ-restringido borroso con el criterio lexicográfico $f_1(\tilde{a}) = b$, $f_2(\tilde{a}) = a-c$ y $f_3(\tilde{a}) = a+c$. Puede verse que las soluciones están empatadas con respecto al primer índice de comparación; por lo tanto, para distinguir una solución de otra, debe recurrirse a los otros índices y, en este caso, el segundo índice es suficiente para concluir que ninguna solución domina a otra.

Tabla I
Soluciones óptimas de Pareto

$\tilde{x}_1$	$(0,3,3,1)$	$(0,3,3,4)$	$(0,3,3,5)$
$\tilde{x}_2$	$(1{,}5,2,2{,}36)$	$(1{,}5,2,2{,}24)$	$(1{,}5,2,2{,}2)$
1º objetivo	$(12,50,64{,}78)$	$(12,50,66{,}52)$	$(12,50,67{,}1)$
2º objetivo	$(6,23,40{,}72)$	$(6,23,40{,}48)$	$(6,23,40{,}4)$

VI. Conclusiones

Los NBs se utilizan en la OB para representar datos numéricos aproximados. El uso de funciones lineales de ordenación es un enfoque generalizado para comparar los valores borrosos de las funciones objetivo y seleccionar soluciones "óptimas". Sin embargo, estas funciones no siempre disciernen entre NBs diferentes. El método para PLBM propuesto en [3] garantiza soluciones borrosas óptimas de Pareto. Este método utiliza criterios lexicográficos para comparar NBs y evita las limitaciones de las funciones de lineales de ordenación.

Los problemas de OB abordados en este artículo se plantearon (como suele hacerse) fuera del contexto en el que tiene lugar la toma de decisiones. En el futuro, sería interesante investigar el uso de la OB (particularmente, la PLBM) en la modelización y solución de problemas dependientes del contexto.

Referencias

[1] R. E. Bellman, and L. A. Zadeh, "Decision-making in fuzzy environment," Manag. Sci., vol. 17, pp. 141–164, 1970.

[2] R. Ghanbari, K. Ghorbani-Moghadam, N. Mahdavi-Amiri, and B. De Baets, "Fuzzy linear programming problems: models and solutions," Soft Comput., vol. 24, pp. 10043–10073, 2020.

[3] B. Pérez-Cañedo, J. L. Verdegay, and R. Miranda Pérez, "An epsilon-constraint method for fully fuzzy multiobjective linear programming," Int. J. Intell. Syst., vol. 35, pp. 600–624, 2020.

[4] B. Pérez-Cañedo, J. L. Verdegay, A. Rosete, and E. R. Concepción-Morales, "A multi-objective berth allocation problem in fuzzy environment," Neurocomput., in press.

[5] B. Pérez-Cañedo, J. L. Verdegay, E. R. Concepción-Morales, and A. Rosete, "Lexicographic methods for fuzzy linear programming," Mathematics, vol. 8, pp. 1–21, 2020.

[6] B. Pérez-Cañedo, and E. R. Concepción-Morales, "A method to find the unique optimal fuzzy value of fully fuzzy linear programming problems with inequality constraints having unrestricted L-R fuzzy parameters and decision variables," Expert Syst. Appl., vol. 123, pp. 256–269, 2019.

Sistema médico basado en lógica difusa para la extracción de patrones en un entorno distribuido: dirigido hacia el diagnóstico y a la co-morbilidad.

1st Carlos Fernandez-Basso
Dpt de Ciencias de la Computación e Inteligencia Artificial Universidad de Granada, Granada, España
cjferba@decsai.ugr.es

2nd Karel Gutiérrez-Batista
Dpt de Ciencias de la Computación e Inteligencia Artificial Universidad de Granada, Granada, España
karel@decsai.ugr.es

3rd Roberto Morcillo-Jiménez
Dpt de Ciencias de la Computación e Inteligencia Artificial Universidad de Granada, Granada, España
robermorji@decsai.ugr.es

4th Maria-Amparo Vila
Dpt de Ciencias de la Computación e Inteligencia Artificial Universidad de Granada, Granada, España
avila@decsai.ugr.es

5th Maria J. Martin-Bautista
Dpt de Ciencias de la Computación e Inteligencia Artificial Universidad de Granada, Granada, España
mbautis@decsai.ugr.es

***Abstract*—En nuestro trabajo [1] abordamos la extracción de conocimiento a partir de historias clínicas utilizando técnicas de minería de datos [2] junto con reglas de asociación [3] en conjunción con la lógica difusa en un entorno distribuido. A parte de la extracción del conocimiento [4] nos centramos en la transformación y tratamiento de los datos para mejorar la modelización de los mismos. Técnicas como la lógica difusa [5], unidas con la extracción de conocimiento mediante reglas de asociación, ayudan a mejorar la interpretabilidad de los datos y a tratarlos con la incertidumbre inherente de los datos del mundo real. Para ello, proponemos un sistema que automáticamente preprocesa la base de datos, transformando y adaptando los datos para el proceso de minería de datos y enriqueciendo los datos para generar patrones más interesantes. Realizamos la fuzzificación de la base de datos médicos para representar y analizar los datos médicos del mundo real con su incertidumbre inherente, así cómo descubrir interrelaciones y patrones entre diferentes características como son el diagnóstico y el alta hospitalaria. Finalmente, el sistema visualiza los resultados obtenidos para facilitar la interpretabilidad de la información extraída, mejorando significativamente la comprensión de los datos médicos dirigidos al usuario final, para conseguir tomar decisiones adecuadas.**

La figura 1 representa el sistema completo que sigue nuestro trabajo. Se puede dividir en tres grandes bloques. En el primero, se realiza la recogida de datos. Estos datos pueden ser agregados a partir de fuentes externas de otras áreas, servicios del hospital e incluso sensores del propio centro hospitalario, consiguiendo de esta manera que, nuestro sistema recoja los datos y los fusione con los datos históricos del paciente para tenerlos todos modelados en función del mismo. De esta manera se consigue ver la trazabilidad del paciente y sus respectivos diagnósticos en cada una de sus visitas.

En el siguiente bloque, los datos se almacenan en una arquitectura Big Data [6] que permite utilizar grandes conjuntos de datos procedentes de fuentes heterogéneas. Para ello, se han utilizado bases de datos no relacionales [7], [8] porque proporcionan una gran flexibilidad para almacenar datos de fuentes heterogéneas y grandes volúmenes. Por otro lado, se han utilizado herramientas de procesamiento distribuido como Apache Spark [9] para gestionar, procesar y analizar estos datos masivos de forma eficiente. Finalmente en el último bloque, se realizará la extracción de conocimiento de la base de datos procesada. Este bloque utiliza un algoritmo de reglas de asociación difusas en Big Data [10] que nos permitirá extraer reglas de asociación utilizando Spark. Además, la interfaz de usuario implementará algunas de las herramientas más utilizadas para la visualización de reglas de asociación para facilitar la comprensión e interpretación de los resultados a los usuarios finales. El sistema procesa estos datos a través de tres módulos: preprocesamiento, enriquecimiento y fuzzificación. En nuestro trabajo [1], se han tomado los datos de dos hospitales españoles: el hospital clínico de Granada y el hospital Costa del Sol de Marbella. Todos estos datos tienen que fusionarse utilizando el historial del paciente y adaptando los campos para que estos datos puedan fusionarse y conseguir obtener los modelos relaconales. En nuestro trabajo, se presentan varios módulos de preprocesamiento. En ellos, algunas variables han sido transformadas para modelar mejor la información que contienen. Además, en algunos casos, se han utilizado datos externos para enriquecer los datos contenidos en las variables del conjunto de datos. Por último, se ha llevado a cabo un proceso de tratamiento de la incertidumbre utilizando lógica difusa. En este proceso se ha utilizado un novedoso proceso de fuzzificación que automáticamente o guiado por el conocimiento de los expertos puede utilizar etiquetas difusas para modelar mejor la información y permitir resultados más interpretativos para los usuarios finales.

En el módulo de preprocesamiento se realiza un preprocesamiento clásico, eliminando de la base de datos los códigos que no contienen información útil, normalizando algunos valores y transformando algunas variables factoriales. En este mismo módulo, posteriormente, se ha llevado a cabo un tratamiento más dirigido hacía los datos médicos. Variables como los códigos de historia, los códigos postales o las fechas se han transformado en datos que representan mejor la información. Por ejemplo, los códigos postales se han transformado en datos de municipios,

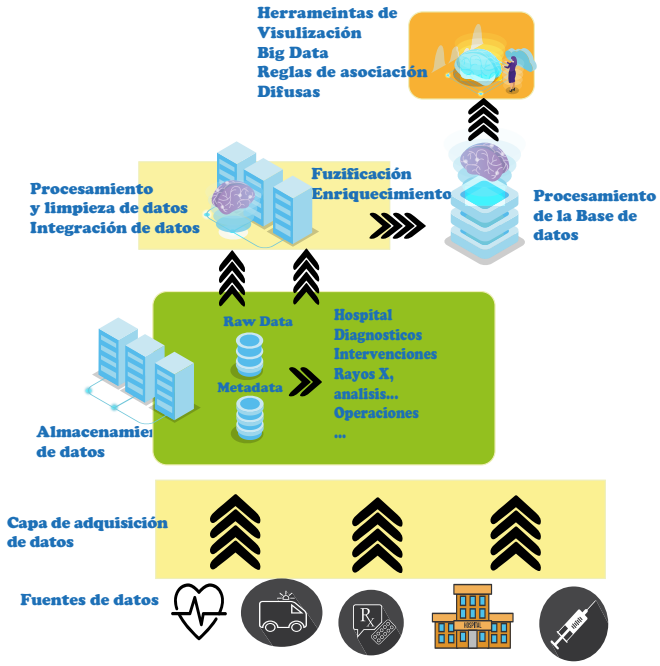


Fig. 1. Proceso general de nuestro trabajo.

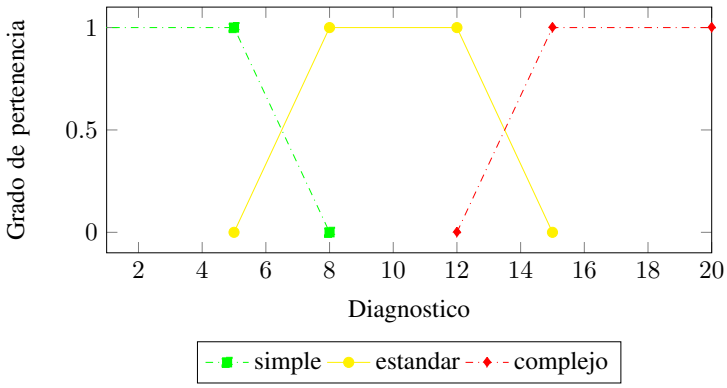


Fig. 2. Distribución de las etiquetas difusas en función de la variable *diagnostico*.

ciudades y países y las fechas por su mes, año y día de la semana. Además, las fechas se han procesado con la edad del paciente y los datos relacionados con su estancia en el hospital, así como la hora de ingreso, la estancia en la Unidad de Cuidados Intensivos (UCI), etc. Seguidamente hemos creado un módulo de enriquecimiento de datos en el que hemos añadido información a los códigos asociados a cada uno de estos campos, modelando su estructura para añadir la información que se encuentra en las taxonomías y bases de datos externas en las que se descodifican. Para este enriquecimiento, se han creado una serie de procesos para extraer la información contenida en la codificación de la variable y se ha modelado esta información para añadirla a la base de datos. Esto ocurre porque las variables codificadas, como los diagnósticos, tienen una codificación en forma de árbol que incluye otros niveles. Además, en el nivel inferior, encontramos información interesante como sinónimos, alteraciones o problemas relacionados con la enfermedad. Por lo tanto, nuestro preprocesamiento para cada diagnóstico generará hasta 5 niveles diferentes de la enfermedad.

En la parte final del procesamiento, hemos desarrollado una novedosa metodología de fuzzificación de características para mejorar la interpretabilidad de los datos. Muchos de los datos procedentes del sistema son difíciles de representar e interpretar por los usuarios finales, ya que se trata de un valor continuo con medidas que a menudo son complejas de comprender. La fuzzificación de estos datos puede mejorar los resultados encontrados por los algoritmos de minería y, al mismo tiempo, aumentar la interpretabilidad de los resultados obtenidos. Proponemos un algoritmo de fuzzificación [1] que permite un tratamiento automático de los valores de los datos según su distribución, en función de las variables del conjunto de datos o de la información proporcionada por un experto. Para ello, hemos desarrollado un algoritmo distribuido en Spark siguiendo la filosofía MapReduce. Esto nos permite procesar grandes cantidades de datos, como es el caso de los datos almacenados por los diferentes hospitales del sistema sanitario español. El objetivo es extraer patrones entre los diagnósticos y las características de los pacientes para estudiar la comorbilidad [11] y mejorar el conocimiento de la información en nuestro sistema. Esto puede mejorar tareas como la obtención de un diagnóstico o la prevención de un posible diagnóstico relacionado con factores del paciente como la obesidad, el tabaquismo, la diabetes etc, consiguiendo agrupar los diagnósticos como simples, estandar o complejos (Fig 2).

REFERENCES

[1] C. Fernandez-Basso, K. Gutiérrez-Batista, R. Morcillo-Jiménez, M.-A. Vila, and M. J. Martin-Bautista, "A fuzzy-based medical system for pattern mining in a distributed environment: Application to diagnostic and co-morbidity," *Applied Soft Computing*, vol. 122, p. 108870, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494622002538

[2] J. Davis and A. Clark, "Data preprocessing for anomaly based network intrusion detection: A review," *Computers and Security*, vol. 30, no. 6-7, pp. 353–375, 2011.

[3] P. Y. TAŞER, K. U. BİRANT, and D. Birant, "Multitask-based association rule mining," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 28, no. 2, pp. 933–955, 2020.

[4] C. Ordonez, N. Ezquerra, and C. A. Santana, "Constraining and summarizing association rules in medical data," *Knowledge and information systems*, vol. 9, no. 3, pp. 1–2, 2006.

[5] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy sets and systems*, vol. 1, no. 1, pp. 3–28, 1978.

[6] A. Beam and I. Kohane, "Big data and machine learning in health care," *JAMA - Journal of the American Medical Association*, vol. 319, no. 13, pp. 1317–1318, 2018.

[7] M. D. Ruiz, J. Gomez-Romero, C. Fernandez-Basso, and M. J. Martin-Bautista, "Big data architecture for building energy management systems," *IEEE Transactions on Industrial Informatics*, 2021.

[8] J. Yoon, D. Jeong, C.-H. Kang, and S. Lee, "Forensic investigation framework for the document store nosql dbms: Mongodb as a case study," *Digital Investigation*, vol. 17, pp. 53–65, 2016.

[9] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, and I. Stoica, "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing," in *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*. USENIX Association, 2012, pp. 2–2.

[10] C. Fernandez-Basso, M. D. Ruiz, and M. J. Martin-Bautista, "Spark solutions for discovering fuzzy association rules in big data," *International Journal of Approximate Reasoning*, vol. 137, pp. 94–112, 2021.

[11] A. R. Feinstein, "The pre-therapeutic classification of co-morbidity in chronic disease," *Journal of Chronic Diseases*, vol. 23, no. 7, pp. 455–468, 1970.

Conexiones de Galois relacionales difusas entre digrafos transitivos difusos

I.P. Cabrera, P. Cordero, E. Muñoz-Velasco, M. Ojeda-Aciego
Departamento de Matemática Aplicada
Universidad de Málaga
Málaga, Spain
Email: ipcabrera@uma.es

***Resumen*—Presentamos una versión difusa de la noción de conexión de Galois relacional entre grafos dirigidos transitivos difusos (digrafos T-difusos) en el entorno específico en el que el álgebra subyacente de valores de verdad es un álgebra de Heyting completa. Los componentes de dicha conexión de Galois difusa son relaciones difusas que satisfacen ciertas propiedades razonables expresadas en términos de lo que denominamos *full powering*. Además, proporcionamos una condición necesaria y suficiente bajo la cual es posible construir un adjunto a la derecha para una relación difusa dada entre un digrafo T-difuso y un conjunto no estructurado.**

I. Introduction

Desde su introducción en 1944 por Ore [18], la noción de conexión de Galois desempeña un papel clave en una amplia gama de líneas de investigación en matemáticas y ciencias de la computación [8]. Para ilustrar el interés actual por las conexiones de Galois, mencionamos sólo algunos estudios muy recientes que ilustran algunas novedosas áreas de aplicación. Oh y Kim [17] consideran las conexiones de Galois y sus duales para desarrollar una visión topológica de los retículos de conceptos sobre retículos residuados completos. Fernández-Alonso y Magaña [9] estudian las conexiones de Galois entre los retículos de prerradicales de dos anillos A y B inducidos por un par de funtores adjuntos entre las categorías A-Mod y B-Mod. Madrid et al. [16] proponen una definición alternativa de operador de aproximación basada en operadores de cierre e interior obtenidos a partir de una conexión de Galois isótona. Horváth et al. [15] caracterizan los grupos de invariancia de conjuntos de funciones booleanas como cierres de Galois de una conexión de Galois adecuada.

Los estudios anteriores tienen en común que consideran conexiones de Galois entre conjuntos con diferentes niveles de estructuración. Un tema de investigación reciente en esta línea es el problema de la construcción de conexiones de Galois [13]. En pocas palabras, el problema se puede describir de la siguiente manera: dada una función $f\colon A \to B$ entre diferentes estructuras (por ejemplo, el dominio A puede ser un retículo y el codominio B un conjunto sin estructura), se busca establecer las condiciones necesarias y suficientes bajo las cuales es posible dotar a B de una estructura deseada y construir una función $g\colon B \to A$ tal que el par (f, g) sea una conexión de Galois. Es importante señalar que el hecho de que A y B no tengan necesariamente la misma estructura excluye la aplicación del teorema del functor adjunto de Freyd. En el primer trabajo en esta dirección de investigación [12] se consideraron dos casos en los que el dominio A era un conjunto o bien parcialmente ordenado o bien preordenado.

No es de extrañar que las conexiones de Galois también se hayan introducido en el ámbito de la teoría de conjuntos difusos. Los grados de libertad disponibles suelen dan lugar a diferentes posibilidades de generalización, y en el caso de una conexión de Galois difusa se dispone de distintas definiciones no equivalentes [1], [14], [19]. En trabajos anteriores [2], [3] exploramos el mencionado problema de construcción en varios entornos difusos, extendiendo satisfactoriamente el problema a las conexiones de Galois entre un dominio difuso A y un rango difuso B, aunque sus componentes estaban formadas por funciones nítidas, con lo que no se obtiene una noción verdaderamente difusa de conexión de Galois.

Más recientemente, el enfoque presentado en [4] nos devolvió al caso nítido al considerar las conexiones de Galois cuyo dominio y rango son sólo conjuntos dotados de relaciones arbitrarias y cuyas componentes son relaciones (propias), dando lugar a lo que llamamos *conexiones de Galois relacionales*. Los estudios posteriores se centraron en los casos en los que el dominio A tiene la estructura de un digrafo transitivo [5] o de un digrafo transitivo difuso [6], estudiando las conexiones de Galois cuyas componentes izquierda y derecha son relaciones nítidas que satisfacen ciertas propiedades razonables expresadas en términos de lo que denominamos *full powering*.

En este trabajo, nos centramos en el caso particular en el que el álgebra subyacente de valores de verdad es un álgebra de Heyting completa, y exponemos por primera vez una noción adecuada de *conexión de Galois relacional difusa* entre digrafos transitivos difusos, siendo ahora ambas componentes relaciones difusas. De forma similar a [4]–[6], nuestra noción se apoya en una de las diversas definiciones estándar equivalentes en el caso nítido, a saber, que las componentes de la conexión de Galois son antítonas y sus composiciones inflacionarias. Se demuestra que esta noción de conexión de Galois relacional difusa hereda la mayoría de las caracterizaciones equivalentes de la noción de conexión de Galois nítida. Para esta noción de conexión de Galois relacional difusa, hemos caracterizado la existencia de un adjunto a la derecha para una relación difusa dada entre un digrafo T-difuso y un conjunto no estructurado.

En cuanto a la noción particular de conexión de Galois relacional difusa propuesta, cabe mencionar que hemos estudiado las propiedades mínimas necesarias para caracterizar esta noción en términos de una condición de Galois natural [10]. Estas propiedades resultan estar relacionadas con el marco de las funciones difusas perfectas [7].

Como trabajo futuro, por un lado, exploraremos más a fondo la relación con las funciones difusas perfectas de alguno de nuestros trabajos anteriores [3]; por otro lado, vale la pena investigar la posibilidad de abandonar la restricción de Heyting para la caracterización de la hipótesis general para la caracterización de la existencia y la construcción del adjunto a la derecha.

REFERENCIAS

[1] R. Bělohlávek. Fuzzy Galois connections. *Mathematical Logic Quarterly*, 45(4):497–504, 1999.

[2] I. Cabrera, P. Cordero, F. Garcia-Pardo, M. Ojeda-Aciego, and B. De Baets. On the construction of adjunctions between a fuzzy preposet and an unstructured set. *Fuzzy Sets and Systems*, 320:81–92, 2017.

[3] I. Cabrera, P. Cordero, F. Garcia-Pardo, M. Ojeda-Aciego, and B. De Baets. Galois connections between a fuzzy preordered structure and a general fuzzy structure. *IEEE Transactions on Fuzzy Systems*, 26(3):1274–1287, 2018.

[4] I. Cabrera, P. Cordero, E. Muñoz-Velasco, and M. Ojeda-Aciego. A relational extension of Galois connections. In *Proc. of Intl Conf on Formal Concept Analysis (ICFCA)*, Lecture Notes in Artificial Intelligence, vol 11511, pages 290–303, 2019.

[5] I. Cabrera, P. Cordero, E. Muñoz-Velasco, M. Ojeda-Aciego, and B. De Baets. Relational Galois connections between transitive digraphs: characterization and construction. *Information Sciences*, 519:439–450, 2020.

[6] I. Cabrera, P. Cordero, E. Muñoz-Velasco, M. Ojeda-Aciego, and B. De Baets. Relational Galois connections between transitive fuzzy digraphs. *Mathematical Methods in the Applied Sciences*, 43(9):5673–5680, 2020.

[7] M. Demirci. Fuzzy functions and their applications. *J. of Mathematical Analysis and its Applications*, 252:495–517, 2000.

[8] K. Denecke, M. Erné, and S. L. Wismath. *Galois connections and applications*, volume 565. Springer, 2004.

[9] R. Fernández-Alonso and J. Magaña. Galois connections between lattices of preradicals induced by ring epimorphisms. *Journal of Algebra and Its Applications*, 19(3):2050045, 2020.

[10] F. García-Pardo, I. Cabrera, P. Cordero, and M. Ojeda-Aciego. On Galois connections and soft computing. *Lect. Notes in Computer Science*, 7903:224–235, 2013.

[11] F. García-Pardo, I. Cabrera, P. Cordero, and M. Ojeda-Aciego. On closure systems and adjunctions between fuzzy preordered sets. *Lect. Notes in Computer Science*, 9113:114–127, 2015.

[12] F. García-Pardo, I. Cabrera, P. Cordero, M. Ojeda-Aciego, and F. Rodríguez. On the definition of suitable orderings to generate adjunctions over an unstructured codomain. *Information Sciences*, 286:173–187, 2014.

[13] F. García-Pardo, I. Cabrera, P. Cordero, M. Ojeda-Aciego, and F. Rodríguez. On the existence of isotone Galois connections between preorders. *Lect. Notes in Computer Science*, 8478:67–79, 2014.

[14] G. Georgescu and A. Popescu. Non-commutative fuzzy structures and pairs of weak negations. *Fuzzy Sets and Systems*, 143(1):129–155, 2004.

[15] E. Horváth, R. Pöschel, and S. Reichard. Invariance groups of functions and related Galois connections. *Contributions to Algebra and Geometry*, 2020.

[16] N. Madrid, J. Medina, and E. Ramírez-Poussa. Rough sets based on Galois connections. *Int. J. of Applied Mathematics and Computer Science*, 30(2), 2020.

[17] J.-M. Oh and Y.-C. Kim. Fuzzy Galois connections on Alexandrov L-topologies. *Journal of Intelligent & Fuzzy Systems*, 2020. To appear.

[18] Ø. Ore. Galois connexions. *Transactions of the American Mathematical Society*, 55:493–513, 1944.

[19] W. Yao and L.-X. Lu. Fuzzy Galois connections on fuzzy posets. *Mathematical Logic Quarterly*, 55(1):105–112, 2009.

Sobre la agregación de T-subgrupos*

S. Ardanza-Trevijano
Dpto de Física y Matemática Aplicada
Fac. Ciencias
Universidad de Navarra
Pamplona, España
sardanza@unav.es

M.J. Chasco
Dpto de Física y Matemática Aplicada
Fac. Ciencias
Universidad de Navarra
Pamplona, España
mjchasco@unav.es

J. Elorza
Dpto de Física y Matemática Aplicada
Fac. Ciencias
Universidad de Navarra
Pamplona, España
jelorza@unav.es

M. de Natividade
Dpto de Física y Matemática Aplicada
Fac. Ciencias
Universidad de Navarra
Pamplona, España
mnatividade@external.unav.es

F.J. Talavera
Dpto de Física y Matemática Aplicada
Fac. Ciencias
Universidad de Navarra
Pamplona, España
ftalaveraan@alumni.unav.es

***Resumen*—Se recopilan dos posibles definiciones de agregación de subgrupos difusos: en conjuntos y en productos. Posteriormente, se estudian las condiciones que deben cumplir el grupo sobre el que se definen los subgrupos difusos y el operador de agregación para asegurar la preservación de la estructura de min-subgrupo en conjuntos. Se muestra que la condición de preservación de T-subgrupos en productos es más fuerte que la de preservación en conjuntos.**

***Index Terms*—Operador de agregación, t-norma, T-subgrupo, agregación de T-subgrupos, preservación de la estructura de T-subgrupo.**

I. Introducción

El estudio de la preservación de estructuras difusas mediante agregaciones comenzó con los artículos de Ovchinnikov, Fodor y Roubens ([2], [3]). Posteriormente, Saminger, Mesiar y Bodenhofer introdujeron en [6] otra definición para la agregación de estructuras difusas, con un dominio mayor. Desde entonces, ambas definiciones se han empleado en distintos trabajos.

En el artículo de Pedraza, Rodríguez-López and Valero [4] se recogen ambas notaciones y se establece la definición de agregación en conjuntos y en productos para distinguirlas. Además se demuestra que la estructura de cuasimétrica difusa se preserva en conjuntos pero no en productos confirmando que ambas definiciones no son equivalentes.

Teniendo en cuenta esta distinción, nuestro trabajo se centra en el estudio de la preservación de la estructura de min-subgrupo difuso en conjuntos. Veremos que la estructura del grupo ambiente juega un papel importante. También estudiamos la relación entre la preservación de T-subgrupos bajo agregación en conjuntos y en productos. Este estudio completa el realizado en [1] por algunos de los autores de este trabajo acerca de la preservación de T-subgrupos en conjuntos bajo operadores de agregación binarios.

Una versión extendida de este trabajo se ha enviado a la revista Fuzzy Sets and Systems y se encuentra actualmente en fase de revisión.

II. Preliminares

En esta sección introducimos las definiciones y los resultados que se utilizarán más adelante.

Definición 1. *Una norma triangular, t-norma para abreviar, es una operación binaria T en el intervalo unidad $[0,1]$ tal que, para todo $a,b,c \in [0,1]$, se satisfacen las siguientes propiedades:*

$T1$. $T(a,b) = T(b,a)$. *(Conmutatividad)*
$T2$. $T(a,T(b,c)) = T(T(a,b),c)$. *(Asociatividad)*
$T3$. $T(a,b) \leq T(a,c)$ *whenever* $b \leq c$. *(Monotonía)*
$T4$. $T(a,1) = a$. *(Condiciones de contorno)*

Rosenfeld introdujo la noción de subgrupo difuso en [5] empleando la t-norma del mínimo (T_M para abreviar). Frecuentemente, esta definición se completa añadiendo un axioma de normalización ($G1$) y utilizando una t-norma genérica como vemos a continuación.

Definición 2. *Sea G un grupo, $\mu : G \to [0,1]$ un subconjunto difuso de G y T una t-norma. Se dice que μ es T-subgrupo difuso de G si:*

$G1$. $\mu(e) = 1$ *donde $e \in G$ denota el elemento neutro.*
$G2$. $\mu(x) = \mu(x^{-1})\ \forall x \in G$
$G3$. $\mu(xy) \geq T(\mu(x),\mu(y))\ \forall x,y \in G$

Definición 3. *Sea $\boldsymbol{A} : \bigcup_{n\in\mathbb{N}}[0,1]^n \to [0,1]$ una función. $\boldsymbol{A}$ se dice operador de agregación o función de agregación si:*

A1. *Para todo $\boldsymbol{x} = (x_1,\ldots,x_n)$, $\boldsymbol{y} = (y_1,\ldots,y_n) \in [0,1]^n$ tal que $x_i \leq y_i\ \forall i \in \{1,\ldots,n\}$, se tiene que $\boldsymbol{A}(x_1,\ldots,x_n) \leq \boldsymbol{A}(y_1,\ldots,y_n)$*
A2. $\boldsymbol{A}(\boldsymbol{0}) = \boldsymbol{A}(0,\ldots,0) = 0$ *y* $\boldsymbol{A}(\boldsymbol{1}) = A(1,\ldots,1) = 1$.
A3. $\boldsymbol{A}(x) = x$ *para todo* $x \in [0,1]$

Cada operador $\boldsymbol{A}$ representa una colección de operadores de agregación n-ésimos $A_{(n)} : [0,1]^n \to [0,1]$ cumpliendo A1 y A2 si $n \geq 2$ y también A3 en el caso $n = 1$. Utilizaremos A

en vez de $A_{(n)}$ de ahora en adelante cuando no haya lugar a confusión.

Definición 4. *Sea $\boldsymbol{A} : \bigcup_{n\in\mathbb{N}}[0,1]^n \to [0,1]$, un operador de agregación, T una t-norma y n T-subgrupos difusos $\mu_1,\ldots,\mu_n$ de un grupo G.*

Denotamos por $\boldsymbol{\mu}$ a la aplicación $\boldsymbol{\mu} : G \to [0,1]^n$, con $\boldsymbol{\mu}(x) = (\mu_1(x),\ldots,\mu_n(x))$ siendo x un elemento de G. Del mismo modo, $\tilde{\boldsymbol{\mu}}$ denota una aplicación $\tilde{\boldsymbol{\mu}} : \prod_{i=1}^n G \to [0,1]^n$ con $\tilde{\boldsymbol{\mu}}(\boldsymbol{x}) = (\mu_1(x_1),\ldots,\mu_n(x_n))$ siendo $\boldsymbol{x} = (x_1,\ldots,x_n)$ un elemento de $\prod_{i=1}^n G$.

*Definimos la **agregación de subgrupos difusos en conjuntos** como $A \circ \boldsymbol{\mu}$, donde:*

$$A \circ \boldsymbol{\mu}(x) = A(\mu_1(x),\ldots,\mu_n(x))$$

*De la misma manera, definimos la **agregación de subgrupos difusos en productos** como $A \circ \tilde{\boldsymbol{\mu}}$, donde:*

$$A \circ \tilde{\boldsymbol{\mu}}(\boldsymbol{x}) = A(\mu_1(x_1),\ldots,\mu_n(x_n))$$

Si $\mu_1 = \ldots = \mu_n$ usamos el término autoagregación de subgrupos difusos (en conjuntos o en productos).

Con esta terminología decimos que :

***A preserva la estructura de T-subgrupo en conjuntos** si, y sólo si $A \circ \boldsymbol{\mu}$ es T-subgrupo para cualquier $\boldsymbol{\mu}$ de la forma definida arriba.*

***A preserva la estructura de T-subgrupo en productos** si, y sólo si $A \circ \tilde{\boldsymbol{\mu}}$ es T-subgrupo para cualquier $\tilde{\boldsymbol{\mu}}$ de la forma definida arriba.*

III. Agregación de T-subgrupos

En este apartado se recogen los resultados obtenidos acerca de la preservación de la estructura de min-subgrupo bajo operadores de agregación en conjuntos. Además, se muestra que la preservación de T-subgrupos en productos implica su preservación en conjuntos pero no al revés.

La siguiente proposición indica que el cumplimiento de la condición $G3$ es clave para que $A \circ \boldsymbol{\mu}$ y $A \circ \tilde{\boldsymbol{\mu}}$ sean subgrupos difusos.

Proposición 5. *Sea G un grupo no trivial, $\boldsymbol{A} : \bigcup_{n\in\mathbb{N}}[0,1]^n \to [0,1]$ un operador de agregación, y $\mu_1,\ldots,\mu_n$ T-subgrupos, entonces $A \circ \boldsymbol{\mu}$ y $A \circ \tilde{\boldsymbol{\mu}}$ satisfacen los axiomas de subgrupo difuso $G1$ y $G2$.*

De ahora en adelante nos centraremos en la preservación de la propiedad G3. Dados dos subgrupos difusos μ y η, en [1] se define el conjunto $\mathcal{L}(\mu,\eta) = \{x,y \in G \mid \mu(x) > \mu(y) \text{ and } \eta(y) > \eta(x)\}$ que jugará un papel importante. Se puede extender esta definición a n subgrupos difusos de la siguiente manera:

Definición 6. *Sea G un grupo y $\mu_1,\ldots,\mu_n \in [0,1]^G$ subgrupos difusos para una t-norma T. Podemos definir el conjunto:*

$$\mathcal{L}(\mu_1,\ldots,\mu_n) := \{x,y \in G \mid \exists i_0,j_0 \in \{1,\ldots,n\} \text{ con } \mu_{i_0}(x) > \mu_{i_0}(y) \text{ and } \mu_{j_0}(y) > \mu_{j_0}(x)\}$$

Teorema 7. *Sea G un grupo no trivial, $\mu_1,\ldots,\mu_n$ min-subgrupos de G y $\boldsymbol{\mu} = (\mu_1,\ldots\mu_n) : G \to [0,1]^n$. Entonces $\mathcal{L}(\mu_1,\ldots\mu_n) = \emptyset$ si, y sólo si, para todo operador de agregación $\boldsymbol{A}$, $A \circ \boldsymbol{\mu}$ es min-subgrupo.*

Teorema 8. *Sea G un grupo no trivial. Las siguientes afirmaciones son equivalentes:*

(i) *$\mathcal{L}(\mu,\eta) = \emptyset$ para todo $\mu,\eta \in [0,1]^G$ min-subgrupos.*
(ii) *$\mathcal{L}(\mu_1,\ldots,\mu_n) = \emptyset$ para todo $\mu_1,\ldots,\mu_n \in [0,1]^G$ min-subgrupos y para todo $n \geq 2$.*
(iii) *El retículo de subgrupos de G es una cadena.*
(iv) *Toda función de agregación preserva la estructura de min-subgrupo en conjuntos.*
(v) *Toda función no decreciente $F : [0,1]^n \to [0,1]$ preserva la estructura de min-subgrupo en conjuntos, es decir, $F \circ \boldsymbol{\mu}$ es min-subgrupo para cualquier $\boldsymbol{\mu}$.*

Proposición 9. *Sea G un grupo no trivial, T una t-norma, y $\boldsymbol{A}$ y un operador de agregación. Si $\boldsymbol{A}$ preserva la estructura de T-subgrupo en productos, entonces $\boldsymbol{A}$ también la preserva en conjuntos.*

El siguiente ejemplo muestra que el recíproco de la proposición anterior no es cierto.

Ejemplo 10. *Consideremos $\mathbb{Z}_p$, el grupo cíclico de orden primo p, T_M la t-norma del mínimo y T_P la t-norma del producto como operador de agregación. Se puede comprobar que la estructura de subgrupo difuso se conserva siempre en conjuntos y no es difícil encontrar dos subgrupos difusos cuya agregación en productos no sea subgrupo difuso.*

IV. Conclusiones

Hemos presentado dos formas de agregar subgrupos difusos y, concretamente, hemos estudiado la agregación de min-subgrupos en conjuntos. En esta línea, hemos demostrado que toda función de agregación preserva la estructura de T_M-subgrupo en conjuntos si, y sólo si, el retículo de subgrupos del grupo G sobre el que definen los subgrupos difusos es una cadena. Hemos mostrado también que si un operador de agregación preserva la estructura de T-subgrupo en productos, también lo hace en conjuntos y que la implicación contraria no se cumple.

Agradecimientos

Los autores agradecen las ayudas del Ministerio de Economía y Competitividad (Grant PGC 2018-098623-B-I00).

Referencias

[1] Bejines, C., Chasco, M., and Elorza, J. Aggregation of fuzzy subgroups. *Fuzzy Sets and Systems 418* (2021), 170–184.
[2] Fodor, J. C., and Roubens, M. Aggregation and scoring procedures in multicriteria decision making methods. In *[1992 Proceedings] IEEE International Conference on Fuzzy Systems* (1992), IEEE, pp. 1261–1267.
[3] Ovchinnikov, S. Social choice and lukasiewicz logic. *Fuzzy Sets and Systems 43*, 3 (1991), 275–289.
[4] Pedraza, T., Rodríguez-López, J., and Valero, Ó. Aggregation of fuzzy quasi-metrics. *Information Sciences* (2021), 362–389.
[5] Rosenfeld, A. Fuzzy groups. *Journal of mathematical analysis and applications 35*, 3 (1971), 512–517.
[6] Saminger, S., Mesiar, R., and Bodenhofer, U. Domination of aggregation operators and preservation of transitivity. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10*, supp01 (2002), 11–35.

Sobre funciones de implicación basadas en F-cadenas

Isabel Aguilo*†, Sebastia Massanet*†, Juan Vicente Riera*†
* *Grupo de investigación en Soft Computing, Procesamiento de Imágenes y Agregación (SCOPIA)*
Departamento de Ciencias Matemáticas e Informática
Universitat de les Illes Balears, 07122 Palma, España
†*Instituto de Investigación Sanitaria de las Islas Baleares (IdISBa), 07010 Palma, España*
{isabel.aguilo, s.massanet, jvicente.riera}@uib.es

***Abstract*—En la literatura, se han propuesto una gran cantidad de métodos de construcción de funciones de implicación borrosas. Una de las estrategias más utilizadas es la construcción de nuevas funciones de implicación borrosas a partir de otras implicaciones. En esta línea, recientemente se introdujo un nuevo método de construcción usando las conocidas como F-cadenas a partir de una función de implicación dada. En este trabajo se generaliza el método anterior mediante el uso de toda una familia de funciones de implicación borrosas. Una vez introducido el método, se estudian las propiedades adicionales de la función de implicación borrosa generada en función de las propiedades adicionales de las implicaciones consideradas para generarla. Se concluye que dicho método de construcción preserva importantes propiedades adicionales y generaliza varios métodos de construcción existentes.**

***Index Terms*—Función de implicación borrosa, función de agregación, F-cadena**

I. Resumen

Uno de las principales temas de investigación en el campo de la lógica borrosa y del razonamiento aproximado es cómo modelar los condicionales borrosos y realizar inferencias borrosas. Comúnmente, los condicionales borrosos se modelan utilizando relaciones difusas funcionalmente expresables a través de la llamadas funciones de implicación borrosas. La definición más aceptada de dichos operadores es la siguiente:

*Definición 1.1 (***[1]):** Un operador binario $I : [0,1]\times[0,1] \to [0,1]$ es una *función de implicación borrosa*, o una *implicación borrosa*, si satisface:

- (I1) $I(x,z) \geq I(y,z)$ cuando $x \leq y$, para todo $z \in [0,1]$.
- (I2) $I(x,y) \leq I(x,z)$ cuando $y \leq z$, para todo $x \in [0,1]$.
- (I3) $I(0,0) = I(1,1) = 1$ e $I(1,0) = 0$.

Este tipo de operadores son imprescindibles en razonamiento aproximado, pero, además, tienen muchas aplicaciones en otros ámbitos como pueden ser la construcción de medidas de subconjuntos difusos, el procesamiento de imágenes, la teoría de integración basada en medidas monótonas, etc. (ver [1], [2], [4] y las referencias internas). Motivados por la gran cantidad de posibles aplicaciones, las investigaciones sobre este tipo operadores no sólo se han centrado en su aspecto eminentemente práctico, sino también desde el punto de vista estrictamente teórico (ver por ejemplo los monográficos [1], [2]). En este estudio teórico, uno de los principales temas relacionados con las funciones de implicación borrosas es encontrar nuevos métodos de construcción. Este interés radica en la necesidad de tener diferentes familias de funciones de implicación borrosas con diferentes propiedades adicionales para poder elegir la familia de estos operadores que tenga las propiedades adicionales necesarias para poder modelar los condicionales borrosos en cada aplicación (ver [7]). Entre los diferentes métodos de construcción que generan nuevas implicaciones borrosas a partir de implicaciones dadas, cabe mencionar:

- los métodos clásicos como la conjugación, la reciprocidad, las contrapositivizaciones superior, inferior y media, y las combinaciones máxima, mínima o convexa de funciones de implicación borrosas (ver [1]),
- los métodos de construcción más recientes que han sido introducidos y estudiados por muchos autores como son algunos tipos nuevos de contrapositivizaciones, métodos de generación de umbral horizontal y vertical, etc. (ver [5] para una recopilación de estos métodos).

Recientemente, en [3] se introduce un nuevo método de construcción basado en F-cadenas. El concepto de F-cadena se propone en [6] de la siguiente manera.

*Definición 1.2 (***[6]):** Sea $F : [0,1]^n \to [0,1]$ una función de agregación, y sea $\mathbf{c} : [0,1] \to [0,1]^n$ una función creciente tal que $\mathbf{c}(0) = (0,\ldots,0)$, $\mathbf{c}(1) = (1,\ldots,1)$ verificando $F(\mathbf{c}(t)) = t$ para todo $t \in [0,1]$. Entonces $\mathbf{c}$ se llama F-cadena.

Usando este concepto, en [3], se investiga un método de construcción basado en el siguiente resultado:

*Teorema 1.1 (***[6]):** Sea $F : [0,1]^n \to [0,1]$ una función de agregación, $\mathbf{c} : [0,1] \to [0,1]^n$ una F-cadena, y sea I una función de implicación borrosa. Entonces, la función dada por $I_{F,\mathbf{c}} : [0,1]^2 \to [0,1]$ dada por

$$I_{F,\mathbf{c}}(x,y) = F(I(c_1(x),c_1(y)),\ldots,I(c_n(x),c_n(y)))$$

para todo $x, y \in [0,1]$ es una función de implicación borrosa.

Este trabajo es parte del proyecto de I+D+i PID2020-113870GB-I00 - "Desarrollo de herramientas de Soft Computing para la Ayuda al Diagnóstico Clínico y a la Gestión de Emergencias (HESOCODICE)", financiado por MCIN/AEI/10.13039/501100011033/.

En [6], los autores demuestran que si la función de implicación elegida verifica el principio de neutralidad por la izquierda (**NP**), es decir, $I(1, y) = y$ para todo $y \in [0, 1]$ o el principio de identidad (**IP**), es decir, $I(x, x) = 1$ para todo $x \in [0, 1]$, la función $I_{F,\mathbf{c}}$ también lo verifica. Sin embargo, para otras propiedades clásicas como puede ser la propiedad de orden (**OP**), dada por $I(x, y) = 1$ si y sólo si $x \leq y$, o la contraposición con respecto a la negación clásica (**CP**(N_C)), dada por $I(x, y) = I(1 - y, 1 - x)$ para todo $x, y \in [0, 1]$, la F-cadena debe de verificar algunas propiedades adicionales como, por ejemplo, ser simétrica o estrictamente creciente.

En este trabajo se propone generalizar el método de construcción propuesto en [6] y concretado en el teorema 1.1. De esta forma, se tiene el siguiente resultado:

Teorema 1.2: Sea $F : [0, 1]^n \to [0, 1]$ una función de agregación, $\mathbf{c} : [0, 1] \to [0, 1]^n$ una F-cadena, y $I_1, \ldots, I_n : [0, 1]^2 \to [0, 1]$ n funciones de implicación borrosas. Entonces, la función $(I_1, \ldots, I_n)_{F,\mathbf{c})} : [0, 1]^2 \to [0, 1]$ dada por

$$(I_1, \ldots, I_n)_{F,\mathbf{c}}(x, y) = F(I_1(c_1(x), c_1(y)), \ldots, I_n(c_n(x), c_n(y)))$$

para todo $x, y \in [0, 1]$ es una función de implicación borrosa. En particular, es evidente que si todas las funciones de implicación borrosas consideradas son la misma, se recupera el método establecido en el teorema 1.1. Además, se demuestra que este método de construcción preserva siempre algunas propiedades adicionales de las funciones de implicación borrosas consideradas inicialmente. Por ejemplo, cuando la familia de funciones de implicación borrosas $I_1, \cdots, I_n$ satisfacen (**NP**) o (**IP**) la función de implicación borrosa $(I_1, \ldots, I_n)_{F,\mathbf{c}}$ también las satisface.

En cambio, otras propiedades adicionales requieren para su preservación determinadas propiedades de la función de agregación y/o de la F-cadena como se puede observar en el siguiente resultado. En el mismo, se utilizan, por un lado, el concepto de función de agregación dual propia, esto es, una función de agregación $F : [0, 1]^n \to [0, 1]$ que satisface $F(1 - x_1, \ldots, 1 - x_n) = 1 - F(x_1, \ldots, x_n)$ para todos $x_1, \ldots, x_n \in [0, 1]$; y, por otro lado, el concepto de función de agregación que no tiene multiplicadores de 1, esto es, una función de agregación $F : [0, 1]^n \to [0, 1]$ que satisface $F(x_1, \ldots, x_n) = 1$ si, y solo si, $x_1 = \ldots = x_n = 1$.

Proposición 1.1: Sea $F : [0, 1]^n \to [0, 1]$ una función de agregación, $\mathbf{c} : [0, 1] \to [0, 1]^n$ una F-cadena e $I_1, \ldots, I_n : [0, 1]^2 \to [0, 1]$ una familia de funciones de implicación borrosas. Las siguientes afirmaciones son ciertas:

(i) Si la negación natural de $I_1, \ldots, I_n \in \mathcal{I}$ es la negación clásica N_C, esto es, $I_i(x, 0) = 1 - x$ para todo $x \in [0, 1]$ y $1 \leq i \leq n$, y si F es una función de agregación dual propia, entonces $(I_1, \ldots, I_n)_{F,\mathbf{c}}(x, 0) = 1 - x$ para todo $x \in [0, 1]$.

(ii) Si $I_1, \ldots, I_n$ satisfacen (**CP**(N_C)) para todo $1 \leq i \leq n$, F es una función de agregación dual propia y la F-cadena $\mathbf{c}$ satisface la condición de simetria

$$\mathbf{c}(1 - t) + \mathbf{c}(t) = (1, \ldots, 1)$$

para todo $t \in [0, 1]$, entonces $(I_1, \ldots, I_n)_{F,\mathbf{c}}$ satisface (**CP**(N_C)).

(iii) Si $I_1, \ldots, I_n$ satisfacen (**OP**) para todo $1 \leq i \leq n$, la F-cadena $\mathbf{c}$ es estrictamente creciente en cada componente y F no tiene multiplicadores de 1, entonces $(I_1, \ldots, I_n)_{F,\mathbf{c}}$ también satisface (**OP**).

(iv) Si $I_1, \ldots, I_n$ satisfacen (**CB**) para todo $1 \leq i \leq n$, esto es, $I_i(1, y) \geq y$ para todo $y \in [0, 1]$, F satisface $F(x_1, \ldots, x_n) \geq \max(x_1, \ldots, x_n)$ para todos $x_1, \ldots, x_n \in [0, 1]$ y la F-cadena $\mathbf{c}$ satisface $c_i(x) = x$ para todo $1 \leq i \leq n$, entonces $(I_1, \ldots, I_n)_{F,\mathbf{c}}$ satisface (**CB**).

(v) Si $I_1, \ldots, I_n$ satisfacen $I_i(x, y) = 1 \Leftrightarrow x = 0$ o $y = 1$ para todos $1 \leq i \leq n$, F no tiene multiplicadores de 1 y la F-cadena $\mathbf{c}$ es estrictamente creciente y satisface que $\mathbf{c}(x) \in (0, 1)$ cuando $x \in (0, 1)$, entonces $(I_1, \ldots, I_n)_{F,\mathbf{c}}(x, y) = 1 \Leftrightarrow x = 0$ o $y = 1$.

Aunque la preservación de propiedades adicionales es una característica importante de cualquier método de construcción de funciones de implicación borrosas, el factor más importante para proponer el método de construcción descrito en el teorema 1.2 es el hecho de que generaliza un gran número de métodos de construcción de funciones de implicación borrosas. Es decir, si consideramos funciones de agregación, F-cadenas y funciones de implicación borrosas adecuadas, se recuperan los siguientes métodos de construcción:

- Los métodos de generación de umbral horizontal y vertical.
- Las combinaciones máxima, mínima y convexa.
- La conjugación.
- La contrapositivación superior e inferior.

Estos resultados permitirán analizar los anteriores métodos de construcción de funciones de implicación borrosas de forma conjunta, simplificando futuras investigaciones sobre ellos.

REFERENCIAS

[1] M. Baczyński, B. Jayaram, Fuzzy Implications. Studies in Fuzziness and Soft Computing, vol. 231. Springer, Berlin Heidelberg, 2008.

[2] M. Baczyński, B. Jayaram, S. Massanet, and J. Torrens, Fuzzy implications: Past, present, and future, in Springer Handbook of Computational Intelligence, J. Kacprzyk and W. Pedrycz, Eds. Springer Berlin Heidelberg, 2015, pp. 183–202.

[3] R. Mesiar, A. Kolesárová, Construction of Fuzzy Implication Functions Based on F-chains. In: Halaš, R., Gagolewski, M., Mesiar, R. (eds) New Trends in Aggregation Theory. AGOP 2019. Advances in Intelligent Systems and Computing, vol 981. Springer, Cham.

[4] M. Mas, M. Monserrat, J. Torrens, E. Trillas, A survey on fuzzy implication functions, IEEE Transactions on Fuzzy Systems, 15(6), 1107–1121, 2007.

[5] S. Massanet, J. Torrens. An overview of construction methods of fuzzy implications. In M. Baczynski, G. Beliakov, H. Bustince, and A. Pradera, editors, Advances in Fuzzy Implication Functions, volume 300 of Studies in Fuzziness and Soft Computing, pages 1–30. Springer Berlin Heidelberg, 2013.

[6] L. Jin, R. Mesiar, M. Kalina, S. Borkotokey, Generalized phi-transformation of aggregation functions, Fuzzy Sets and Systems, 372, 124–141, 2019.

[7] E. Trillas, M. Mas, M. Monserrat, J. Torrens, On the representation of fuzzy rules, International Journal of Approximate Reasoning, 48, 583–597, 2008.

On Representations by Levels and Fuzzy Sets

Daniel Sánchez
Dept. Computer Science and A.I.
University of Granada, Spain
daniel@decsai.ugr.es

***Abstract*—We discuss about the relation between two different theories: representations by levels and fuzzy sets. Though both are intended to represent gradualness, the latter allows to do so only for sets, whilst the former allows to represent gradual objects of any kind (elements, properties, etc.).**

***Index Terms*—Representations by levels, fuzzy sets**

I. Introduction

As pointed out in [1], fuzzy sets are sets affected by *gradualness* in membership. The extension and importance of the work in Fuzzy Set Theory (FST) has led to the idea that FST is *the* theory of gradualness, that is, if something is gradual, we can represent it by using fuzzy sets.

However, in recent years, other theories *different from FST* have been proposed for the same purpose, particularly the Theory of Representations by Levels (RLs) [2] and the X-mu approach [3]. These theories are very similar between them (with subtle but important differences) and also to the idea of Gradual Sets [4], though the latter are considered by its authors as new notions within FST.

We discuss here about the relation between FST and RLs.

II. Representations by levels

A. Representation

RLs consider that gradualness is represented by a set of *levels* represented by real numbers in $(0,1]$. These levels represent degrees of "relaxation" in the definition of a *gradual object*, which can be defined as an assignment of individual crisp objects to levels. Level 1 represent "no relaxation", that is, being maximally strict in the definition of the gradual object. On the other end of the scale, level 0 represents "no restriction" (maximum relaxation), and hence is never used as a level in practice. Intermediate values represent a degree of relaxation proportional to its distance to 1 (eq. to 0).

Let us consider a category of crisp objects $\mathbb{C}$. For instance, $\mathbb{C}$ may be the set of real numbers, the set of all algorithms written in a given programming language, or the power set of a crisp set X. Then the category of gradual objects of $\mathbb{C}$, $G_{\mathbb{C}}$, is characterized by all functions of the form

$$\rho : (0,1] \to \mathbb{C} \tag{1}$$

This work has been partially supported by the Spanish Ministry of Science, Innovation and Universities and the European Regional Development Fund - ERDF (Fondo Europeo de Desarrollo Regional - FEDER) under project PGC2018-096156-B-I00 "Recuperación y Descripción de Imágenes mediante Lenguaje Natural usando Técnicas de Aprendizaje Profundo y Computación Flexible".

In the particular case $\mathbb{C} = 2^X$ that we have proposed before, $G_{\mathbb{C}}$ is the set of all gradual subsets of X, that is, all subsets of X affected by gradualness. Using RLs, the elements of $G_{\mathbb{C}}$ are represented by means of functions of the form $\rho : (0,1] \to 2^X$.

This allows us to see a first difference with FST, which employs functions of the form $\mu : X \to [0,1]$ for the same purpose: FST extends the range of the indicator function of a set from $\{0,1\}$ (for crisp sets) to $[0,1]$ (for fuzzy sets), and hence the resulting objects is *always a set affected by gradualness*. On the contrary, we can use RLs to define categories of gradual objects other than sets. For instance, the category of gradual real numbers is characterized by all functions of the form $\rho : (0,1] \to \mathbb{R}$, that assign a single real number to each level. On the contrary, the so-called *fuzzy (real) numbers* are in fact (fuzzy) sets of numbers, particularly intervals affected by graduality [1].

A very useful relation between FST and RLs is that every fuzzy subset A of X can be represented as an RL by considering $\rho_A : (0,1] \to 2^X$ defined by $\rho_A(\alpha) = A_\alpha$, where A_α is the α-cut of A. However, let us remark that not every RL can be seen as a fuzzy set since sets in different levels are not required to be nested, and hence *fuzzy sets can be seen as particular types of RLs*.

To conclude, we have seen that FST is useful for representing gradualness in sets, but not in other categories. We have also seen that, even in the case of sets, there are sets affected by gradualness that can be represented by using RLs and not by fuzzy sets (those assignments that are not nested with respect to levels). We shall discuss about the semantics of such RLs that are not fuzzy sets later. We shall also provide a different view of fuzzy sets from the perspective of RLs.

B. Operations

We have seen in the previous section that a *gradual object* of type or category $\mathbb{C}$ is an assignment of individual objects from $\mathbb{C}$ to levels in $(0,1]$. In practice, it is usual to consider a finite, nonempty set of levels $\Lambda \subset (0,1]$ with $\Lambda = \{\alpha_1, \ldots, \alpha_m\}$ and $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_m > \alpha_{m+1} = 0$, representing the set of levels that are relevant and distinguishable in a certain problem, particularly by humans, since we consider that gradualness arises from the human mind. Humans are able to distinguish a finite amount of degrees of a property when asked, and computers are able to represent finite sets of levels only, so this is a reasonable practical assumption. In such case, we shall represent a RL as a pair (Λ, ρ). However,

the function ρ extends to any level in $(0, 1]$ as follows: given $\alpha_i > \alpha > \alpha_{i+1}$, it is $\rho(\alpha) = \rho(\alpha_i)$.

How to extend operations, predicates and algorithms defined on objects from $\mathbb{C}$ to objects in $G_{\mathbb{C}}$? In a simple and unique way: working in each level independently. Hence, operations with gradual objects are transformed into the collection of the corresponding ordinary operations on objects from $\mathbb{C}$ in each level. This is computationally feasible when a finite set of levels Λ is considered.

In the particular case of sets, this levelwise treatment is applied to operations like union, intersection, complement and difference, among others, as well as to predicates like inclusion or equality.

A remarkable property of levelwise operations (apart from being unique and direct) is that they keep the same properties as their counterparts in $\mathbb{C}$. Particularly, RLs of gradual sets satisfy all Boolean properties and hence keep the Boolean algebra structure, something that it is not possible with fuzzy sets. This holds also for gradual real numbers represented by RLs, that keep their field structure, etc.

The cost in every case is that operations are no longer functionally expressible. This also explains why nestedness of levels is lost in the case of sets when the complement operation is involved in a chain of operations (a chain that defines the semantics of the result from that of the operands). Finally, in the case of RLs on 2^X for some X, measures like cardinality, etc. are extended by performing the measures in each level independently. Notice that using this idea, the cardinality of an RL defined on 2^X with X finite, is an RL-natural number, that is, an assignment of individual natural numbers to levels.

C. Measures

The information contained in an RL can be *summarized* by means of measures defined on $(0, 1]$, among others. For instance, let (Λ, ρ) be an RL defined on $\mathbb{C}$ and P a logical condition or event defined on $\mathbb{C}$. Then the measure of (Λ, ρ) associated to P is[1]

$$\nu^P((\Lambda, \rho)) = \sum_{\alpha_i \in \Lambda | P(\rho(\alpha_i))} (\alpha_i - \alpha_{i+1}) \tag{2}$$

As an example, let us consider $\mathbb{C} = 2^X$ with $X = \{x_1, \ldots, x_n\}$ and $P_i(\rho(\alpha))$ being "$x_i \in \rho(\alpha)$". Then $\nu^{P_i}((\Lambda, \rho))$ is the probability that in a level α taken at random from $(0, 1]$ it is $x_i \in \rho(\alpha)$.

This particular case allows us to offer a particular view of fuzzy sets from the perspective of the theory of RLs applied to subsets of a given set X: fuzzy sets are collections of measures of membership obtained from RLs. For our example we can define:

$$\nu^{P_1}((\Lambda, \rho))/x_1 + \cdots + \nu^{P_n}((\Lambda, \rho))/x_n \tag{3}$$

which is a fuzzy set defined by the set of measures $\{\nu^{P_1}, \ldots, \nu^{P_n}\}$. It is easy to show that when (Λ, ρ) is obtained from a fuzzy set as the corresponding assignment of α-cuts to levels, there is a one-to-one correspondence between the RL and the fuzzy set defined in (3). If no predicate is specified and we are working with RLs of sets, we shall call ν to the membership function defined in (3). Hence, we can see that given a fuzzy set μ and an RL given by $\rho(\alpha) = \mu_\alpha$, it is $\mu = \nu$. When a RL defined on 2^X does not satisfy nestedness, this one-to-one correspondence is lost. Hence, different RLs can yield the same fuzzy set as summary.

[1] If the set of levels considered is not finite, the measure can be defined by means of the corresponding integral. We shall only consider the finite case here for simplicity.

III. RL-SYSTEMS

Fuzzy sets and RLs complement very well in what we call RL-systems. RLs defining gradual subsets of X can be given as input to a system by means of fuzzy sets (for instance, corresponding to linguistic labels), and fuzzy sets can be employed for summarizing the information contained in the RLs within the system (losing information, but allowing an easy view to those familiar with FST).

Through the use of measures, it is possible to see that some of the classical operations in FST arise intuitively from individual operations on RLs using RL-systems. This arises from recent, unpublished results as the following: let us consider two fuzzy sets μ_A and μ_B and the corresponding functions ρ_A and ρ_B obtained by assigning the α-cut to each level $\alpha \in (0, 1]$. Then:

- The fuzzy summary of the complement of ρ_A by levels for every x is $1 - \mu_A(x)$.
- For every t-norm between the minimum and Lukasiewicz (that are those used in practice), it is always possible to define two RLs ${\rho'}_A$ and ${\rho'}_B$ such that $\nu'_A = \mu_A$, $\nu'_B = \mu_B$, and the summary ν of the intersection by levels of ${}'_A$ and ${\rho'}_B$ is the intersection of μ_A and μ_B as given by the t-norm chosen. The same can be proven for any t-conorm between the maximum and Lukasiewicz's.
- The sigma-count cardinality of a fuzzy set μ is the expected value of the probability measure that assigns to each natural number k the probability that the α-cut of μ at a level $\alpha \in (0, 1]$ taken at random has exactly k elements.
- Any RL on 2^X with X finite can be obtained by performing set operations on a finite set of RLs corresponding to α-cuts of fuzzy sets defined on X.

IV. CONCLUSIONS

Further novel results are omitted here due to lack of space. We consider that RL-systems, that have been already used in many published practical applications, can provide a significant step forward in the treatment of graduality in AI.

REFERENCES

[1] D. Dubois and H. Prade, "Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets," *Fuzzy Sets Syst.*, vol. 192, pp. 3–24, 2012.
[2] D. Sánchez, M. Delgado, M. Vila, and J. Chamorro-Martínez, "On a non-nested level-based representation of fuzziness," *Fuzzy Sets and Systems*, vol. 192, no. 1, pp. 159–175, 2012.
[3] T. P. Martin, "The X-mu representation of fuzzy sets," *Soft Computing*, vol. 19, pp. 1497–1509, 2015.
[4] D. Dubois and H. Prade, "Gradual elements in a fuzzy set," *Soft Computing*, vol. 12, pp. 165–175, 2008.

Un algoritmo de aprendizaje de medidas difusas supervisado para la combinación de clasificadores

Mikel Uriz
Neuraptic AI
Calle de Juan de Mariana 15
25045, Madrid
uriz@neuraptic.ai

Daniel Paternain, Humberto Bustince, Mikel Galar
Institute of Smart Cities (ISC)
Universidad Pública de Navarra (UPNA)
Campus de Arrosadia, 31006 Pamplona
{daniel.paternain,bustince,mikel.galar}@unavarra.es

***Index Terms*—Medida difusa, integral de Choquet, agregación, *ensemble*, clasificación.**

I. Resumen

La fusión o agregación de información hace referencia a la necesidad de reducir un conjunto de datos en un único valor que recoja y resuma la información proporcionada por cada fuente de datos [1]–[4]. La principal herramienta matemática para la fusión de información son las funciones de agregación.

Definition 1: [1]–[4]Decimos que la función $f : [0,1]^N \to [0,1]$ es una función de agregación si satisface las condiciones de contorno $f(0,\dots,0) = 0$ y $f(1,\dots,1) = 1$, y la monotonía creciente, si $x_i \leq y_i$ para todo $i \in \{1,\dots,N\}$, entonces $f(x_1,\dots,x_N) \leq f(y_1,\dots,y_N)$.

La elección de la función de agregación a la elegir depende totalmente de la aplicación. Existen situaciones en las que es necesario tener en consideración las relaciones o interacciones entre las fuentes de información a agregar. Una forma de hacer esto es mediante el uso de integrales difusas basadas en medidas difusas [5], como las integrales de Choquet [6] o las de Sugeno [7], las cuales son capaces de capturar dichas interacciones. La desventaja de este método es definir los 2^N parámetros de una medida difusa (siendo N el número de datos a agregar).

Definition 2: Sea $\mathcal{N} = \{1,\dots,N\}$. Una medida difusa discreta es una función $m : 2^{\mathcal{N}} \to [0,1]$ que satisface las condiciones de contorno $m(\emptyset) = 0$, $m(\mathcal{N}) = 1$ y la monotonía con respecto a la inclusión, $m(A) \leq m(B)$ siempre que $A \subset B$ para todo $A, B \subseteq \mathcal{N}$.

Definition 3: Dada una medida difusa $m : 2^{\mathcal{N}} \to [0,1]$, la integral discreta de Choquet viene dada por

$$C_m(x_1,\dots,x_N) = \sum_{i=1}^{N} (x_{\sigma(i)} - x_{\sigma(i-1)}) m(\{\sigma(i),\dots,\sigma(N)\}) \tag{1}$$

donde $\sigma : \mathcal{N} \to \mathcal{N}$ es una permutación tal que $x_{\sigma(1)} \leq \dots \leq x_{\sigma(N)}$ y $x_{\sigma(0)} = 0$ por convención.

En *machine learning*, los *ensembles* son combinaciones de modelos o clasificadores que generan predicciones diferentes para cada ejemplo. En el caso concreto de la clasificación supervisada, las salidas de los modelos son, generalmente, probabilidades de que el ejemplo pertenezca a cada una de las posibles clases del problema. Así, los clasificadores del *ensemble* generarán diferentes probabilidades para cada clase, que tendrán que ser combinados para llegar a una predicción final única. Es en este punto en el que las agregaciones basadas en medidas difusas pueden aportar un valor significativo, siempre que las medidas hayan sido definidas de manera adecuada.

El aprendizaje inteligente de medidas difusas se puede afrontar principalmente de dos maneras, supervisada y no supervisada. Los métodos de aprendizaje supervisados aprenden a partir de un conjunto de pares (x, y), donde x es el conjunto de entradas a agregar e y es la salida esperada del modelo. De esta manera, se puede definir un proceso de optimización para ajustar la medida difusa a dicho conjunto de entrenamiento. Evidentemente, la optimización se basa en la minimización de una función de coste, siendo el error cuadrático media (MSE) la función de coste más utilizada en la literatura [8]–[11]. En el caso de los métodos no supervisados, se basan en la predicción de la medida difusa en base a una heurística determinada [12]–[14].

Tras un análisis exhaustivo de la literatura relativa a los métodos de aprendizaje de medidas supervisados, hemos detectado varios problemas que pueden ser potencialmente mejorados en la combinación de clasificadores. El principal de ellos es el hecho de que la función de coste que gobierne el aprendizaje sea el MSE, incluso aunque el problema que se esté tratando de resolver sea un problema de clasificación y no de regresión. En el mundo del *machine learning* es bien sabido que existen otras funciones de coste que son más adecuadas para tratar problemas de clasificación que el MSE, como el *cross-entropy* (CE) [15]. De este modelo, nuestra principal propuesta consiste en la sustitución del MSE por otras funciones de coste más adecuadas para problemas de clasificación.

Evidentemente, la sustitución de la función de coste tiene efectos importantes en el proceso de optimización. Por esta razón, nuestra propuesta de aprendizaje consiste en la utilización de los avances en el área del *deep learning*, como son las librerías de autoderivación como *Tensorflow* o *PyTorch*, para implementar un algoritmo de aprendizaje basado en descenso por gradiente. Esto supone una simplificación en el diseño del algoritmo debido a la computación automática de

gradientes. Además, nos permite la utilización de hardware específico, como la GPU, para acelerar el proceso de aprendizaje.

Teniendo en cuenta todo lo anterior, la novedad de este trabajo consiste en la modificación de la función de coste para el aprendizaje de medidas difusas de manera supervisada, la propuesta de un nuevo algoritmo de aprendizaje de medidas basado en descenso por gradiente que puede estar basado en cualquier función de coste y la definición de tres nuevas políticas de actualización para garantizar la monotonía y evitar bloqueos en el aprendizaje.

Para analizar las diferencias entre nuestra propuesta y otras propuestas de la literatura, hemos realizado un extenso estudio experimental en problemas de clasificación binarios y multiclase. En ambos problemas utilizaremos conjuntos de datos del mundo real para evaluar la potencial aplicabilidad del aprendizaje de la medida difusa. Los experimentos realizados se basan en la utilización de la integral de Choquet como función de agregación, ya que es una de las agregaciones basadas en medias difusas más conocidas y aplicadas. En total, hemos analizado 58 conjuntos de datos (29 binarios y 29 multiclase). Para la construcción de los *ensembles* hemos utilizado la técnica de *Bagging* [16] y los resultados los analizamos con los correspondientes test estadísticos no paramétricos [17].

Tras analizar los resultados obtenidos, podemos concluir que la utilización de otras funciones de coste como el CE en lugar del MSE nos permite obtener una ventaja en problemas de clasificación, tal y como nos planteábamos al comenzar el trabajo. En cuanto a las políticas de actualización propuestas, hemos comprobado que la más sencilla es también la que mejor resultados obtiene. Por último, nuestra propuesta basada en CE y con la mejor política de actualización ha sido comparada con otros métodos del estado del arte, obteniendo resultados equivalentes en problemas sencillos y mejores resultados contrastados en problemas complejos en los que es necesario aprender múltiples medidas difusas.

Acknowledgment

Mikel Uriz ha sido apoyado por CDTI y el Ministerio de Ciencia e Innovación bajo el proyecto Neotec 2021 (SNEO-20211147). Este trabajo también ha sido apoyado por el Ministerio de Ciencia e Innovación bajo el proyecto PID2019-108392GB-I00 (AEI/10.13039/501100011033).

References

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

[2] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap, *Aggregation Functions*. Cambridge University Press, 2009.

[3] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*. Springer, 2007.

[4] G. Beliakov, H. Bustince, and A. Pradera, *A Practical Guide to Averaging Functions*, 2nd ed. Springer, 2015.

[5] M. Grabisch, T. Murofushi, M. Sugeno, and J. Kacprzyk, *Fuzzy Measures and Integrals. Theory and Applications*. Physica Verlag, Berlin, 2000.

[6] G. Choquet, "Theory of capacities," *Ann. Inst. Fourier*, vol. 5, pp. 1953–1954, 1953.

[7] M. Sugeno, "Theory of fuzzy integrals and its applications," Ph.D. dissertation, Tokyo Institute of Technology, 1974.

[8] M. Grabisch, "A new algorithm for identifying fuzzy measures and its application to pattern recognition," in *Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symposium*, 1995, pp. 145–150.

[9] M. A. Islam, D. T. Anderson, A. J. Pinar, and T. C. Havens, "Data-driven compression and efficient learning of the choquet integral," *IEEE Transactions on Fuzzy Systems*, vol. 26, pp. 1908–1922, 2018.

[10] M. A. Islam, D. T. Anderson, A. Pinar, T. C. Havens, G. Scott, and J. M. Keller, "Enabling explainable fusion in deep learning with fuzzy integral neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 28, pp. 1291–1300, 2020.

[11] D. T. Anderson, J. M. Keller, and T. C. Havens, "Learning fuzzy-valued fuzzy measures for the fuzzy-valued sugeno fuzzy integral," in *Computational Intelligence for Knowledge-Based Systems Design*, E. Hüllermeier, R. Kruse, and F. Hoffmann, Eds. Springer Berlin Heidelberg, 2010, pp. 502–511.

[12] M. Uriz, D. Paternain, H. Bustince, and M. Galar, "An empirical study on supervised and unsupervised fuzzy measure construction methods in highly imbalanced classification," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2020, pp. 1–8.

[13] D. Štefka and M. Holeňa, "Dynamic classifier aggregation using interacion-sensitive fuzzy measures," *Fuzzy Sets and Systems*, vol. 270, pp. 25–52, 2015.

[14] M. Uriz, D. Paternain, H. Bustince, and M. Galar, "Unsupervised Fuzzy Measure Learning for Classifier Ensembles From Coalitions Performance," *IEEE Access*, vol. 8, pp. 52 288–52 305, 2020.

[15] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215–232, 1958.

[16] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, pp. 123–140, 1996.

[17] S. García, A. Fernández, J. Luengo, and F. Herrera, "Advanced non-parametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power," *Information Sciences*, vol. 180, pp. 2044–2064, 2010.

Modelos de Consenso basados en Mínimo Coste para Expresiones ELICIT

1º Diego García-Zamora
Universidad de Jaén
Jaén, España
dgzamora@ujaen.es

2º Álvaro Labella
Universidad de Jaén
Jaén, España
alabella@ujaen.es

3º Pedro Nuñez-Cacho
Universidad de Jaén
Jaén, España
pnunez@ujaen.es

4º Rosa M. Rodríguez
Universidad de Jaén
Jaén, España
rmrodrig@ujaen.es

5º Luis Martínez
Universidad de Jaén
Jaén, España
martin@ujaen.es

***Resumen*—Durante los últimos años, la adopción de políticas para fomentar la Economía Circular (EC) en las empresas y administraciones públicas se está convirtiendo en una necesidad. Sin embargo, en ocasiones la información disponible para implementar de forma efectiva tales medidas no es objetiva, sino de carácter subjetivo, y se requiere de conocimiento experto para tomar estas decisiones. Involucrar decisores humanos en los problemas de toma de decisión implica, no solo la gestión de los posibles conflictos, sino también facilitarles la expresión de sus opiniones mediante preferencias lingüísticas cercanas a su forma de pensar. En este trabajo presentamos un modelo de Consenso Integral basado en Mínimo Coste (MIMC) que, mediante el modelado de opiniones de expertos a partir de Expresiones Lingüísticas Comparativas con Translación Simbólica (ELICIT), persigue obtener una solución óptima de un problema de decisión minimizando el coste de modificar las valoraciones iniciales de los expertos y garantizando un nivel de acuerdo mínimo. Este modelo se aplicará para la selección de políticas de EC que debe adoptar una empresa textil en un futuro inmediato.**

***Palabras clave*—Consenso basado en mínimo coste, Información ELICIT, Enfoque lingüístico difuso, Economía Circular**

I. Introducción

La Economía circular (EC) diseña procesos de producción regenerativos y restaurativos, minimizando sus externalidades negativas y aprovechando al máximo las materias primas utilizadas [1]. Los principios de la EC impulsan, entre otras cuestiones, el empleo de energías renovables y la reducción de emisiones y residuos. En el proceso de implementación de un modelo de producción bajo los principios de la EC, se hace necesaria la toma de decisiones basadas en información que puede ser vaga e imprecisa. Por ello, emplear técnicas de decisión para la resolución de problemas de Toma de Decisión en Grupo (TDG) en contexto de incertidumbre, que involucran el conocimiento experto humano, resulta esencial para garantizar el éxito de la puesta en marcha de las medidas de circularidad más convenientes en cada situación. Por otro lado, hay que asegurarse de que los decisores estén de acuerdo con la solución a implantar, evitando de este modo conflictos que bloqueen las medidas propuestas como solución.

Los modelos de MIMC [2] se aplican en la resolución de problemas de TDG con el objetivo de obtener una solución consensuada que minimice el coste de modificar las preferencias de los expertos, asegurando que la distancia entre la opinión individual de cada experto y la colectiva sea inferior a un valor determinado y garantizando que el nivel de acuerdo entre ellos sobrepase un cierto umbral de consenso. Sin embargo, estos modelos se definieron inicialmente para tratar con valoraciones numéricas [2], ignorando el uso de información lingüística. Entre las distintas propuestas para el modelado de información lingüística [3], [4], el modelado ELICIT [5], que proporciona una representación basada en el Enfoque Lingüístico Difuso [6] que permite realizar cálculos precisos con valores lingüísticos sin perder interpretabilidad ni información.

En esta contribución, presentamos un modelo MIMC que integra información ELICIT y, además, lo aplicaremos a la resolución de un problema de TDG en EC.

II. Un modelo de MIMC para información ELICIT

El modelo lingüístico 2-tupla [3] permite manipular de forma precisa valores lingüísticos representándolos mediante un $(s_i, \alpha) \in \overline{S} := S \times [-0{,}5, 0{,}5[$, donde s_i es un término lingüístico perteneciente a cierto conjunto de términos lingüísticos $S = \{s_0, s_1, \dots s_g\}$ ($g \in \mathbb{N}$ par) y $\alpha \in [-0{,}5, 0{,}5[$ representa la desviación de la función de pertenencia difusa del término lingüístico 2-tupla respecto a la de s_i.

Los términos lingüísticos 2-tupla pueden transformarse unívocamente en valores numéricos en el intervalo $[0, g]$ mediante la aplicación $\Delta^{-1} : \overline{S} \to [0, g]$ dada por $\Delta_S^{-1}(s_i, \alpha) = i + \alpha,\ \forall\ (s_i, \alpha) \in \overline{S}$.

Aunque el modelado lingüístico 2-tupla permite manipular de forma precisa expresiones lingüísticas modeladas mediante números difusos, estas expresiones no pueden modelar la duda entre varios términos lingüísticos. Por esto, Labella et al. [5] propusieron la información ELICIT, un marco teórico basado en el Enfoque Lingüístico Difuso que permite conservar la precisión e interpretabilidad del modelo 2-tupla mejorando su expresividad a la hora de modelar la indecisión entre varios términos lingüísticos.

Formalmente, la información ELICIT se puede denotar mediante una expresión $[\overline{s}_i, \overline{s}_j]_{\gamma_1,\gamma_2}$, donde $\overline{s}_i, \overline{s}_j \in \overline{S}, i \leq j$ son dos valores 2-tupla. Además, la información ELICIT considera dos parámetros γ_1, γ_2 que garantizan que no haya pérdida de información al manipular estas expresiones. Cualquier Número Difuso Trapezoidal (NDTr) se puede representar unívocamente mediante una expresión ELICIT:

Proposition 1: Sea $\overline{\overline{S}}$ el conjunto de todos los valores ELICIT. Entonces, la aplicación ζ dada por:

$$\zeta : \mathcal{T} \to \overline{\overline{S}}$$
$$T(a,b,c,d) \to [\overline{s}_1, \overline{s}_2]_{\gamma_1,\gamma_2},$$

donde $\overline{s}_1 = \Delta_S(gb) \quad \gamma_1 = a - \text{máx}\left\{b - \frac{1}{g^2}, 0\right\}$

$\overline{s}_2 = \Delta_S(gc) \quad \gamma_2 = d - \text{mín}\left\{c + \frac{1}{g^2}, 1\right\},$

es una biyección.

Desde el punto de vista teórico, las expresiones ELICIT son generadas por una gramática libre de contexto que modela estructuras lingüísticas comparativas cercanas al lenguaje humano como *al menos malo*, *entre bueno y muy bueno* o *como mucho muy malo* [5]. Labella et al. propusieron el uso de la media ponderada difusa para agregar valores ELICIT así como una medida para calcular distancias entre este tipo de expresiones. Además, los valores ELICIT se pueden comparar mediante el cálculo de su magnitud [5].

Puesto que la manipulación de la información ELICIT se puede hacer a través de los NDTrs correspondientes, es posible definir modelos de MIMC para este tipo de información. Sean $O_1, O_2, ..., O_m \in \mathcal{M}_{n\times n}(\mathcal{T})^*$ los NDTrs obtenidos de las preferencias dadas por los expertos en forma de expresiones ELICIT y sean $T_1, T_2, ..., T_m \in \mathcal{M}_{n\times n}(\mathcal{T})^*$ sus opiniones modificadas. La función coste, la distancia máxima a la colectiva permitida y la medida de consenso propias de los modelos de MIMC se adaptan a información ELICIT como sigue:

$$\min_{T_1^{i,j},\ldots,T_m^{i,j}\in\mathcal{T}} \sum_{k=1}^{m}\sum_{i<j} c_k^{ij}\delta(T_k^{i,j}, O_k^{i,j})$$
$$s.t. \begin{cases} \overline{T}^{i,j} = A(T_1^{i,j}, T_2^{i,j}, \ldots, T_m^{i,j}), 1 \le i < j \le n, \\ \delta(T_k^{i,j}, \overline{T}^{i,j}) \le \varepsilon, 1 \le i < j \le n, k = 1, 2, ..., m, \\ 1 - \frac{1}{N}\sum_{k<l}\sum_{i<j}\frac{w_k+w_l}{m-1}\delta(T_k^{i,j}, T_l^{i,j}) \ge \mu_0, \end{cases}$$
(ELICIT-MIMC)

donde $c_k^{ij} \in [0,1] (\sum_{k=1}^{m}\sum_{i<j} c_k^{ij} = 1)$ modela el coste de mover la opinión del experto e_k respecto a la alternativa x_i sobre la x_j, $w_1, w_2, ..., w_m \in [0,1]$ $(\sum_{k=1}^{m} w_k = 1)$ son los pesos de cada experto, $N = \frac{n(n-1)}{2}$, $\mu_0 \in [0,1[$ es el umbral de consenso deseado, $\varepsilon \in]0,1]$ es la máxima desviación absoluta respecto a la colectiva permitida, $\delta : \mathcal{T}\times\mathcal{T} \to [0,1]$ es una distancia en $\mathcal{T}$ [7] y $A : \mathcal{T}^m \to \mathcal{T}$ es la media ponderada difusa.

III. Caso de estudio

Este caso de estudio se centra particularmente en la empresa textil *EasyClothes*, que desea aplicar políticas de EC para adaptar su proceso de producción a las nuevas medidas impuestas por el gobierno. La dirección consulta a cinco altos ejecutivos ($m = 5$) sobre qué aspecto de la empresa centrar sus políticas verdes durante los próximos dos años [8]. Se valoran cuatro posibles aspectos x_1: Inversión en energías renovables, x_2: Reducción del consumo de agua, x_3: Reducción de residuos sólidos y x_4: Reducción de emisiones de gases. Los ejecutivos dan sus opiniones empleando un conjunto de términos lingüísticos formado por S={Mucho Peor (MP), Peor (P), Ligeramente Peor (LP), Igual (I), Ligeramente Mejor (LM), Mejor (M), Mucho Mejor (MM)}.

Los valores colectivos en consenso para cada alternativa son:

$$x_1 : Entre\ (LP, -0{,}33)^{0{,}002} y\ (LP, 0{,}01)^{-0{,}002},$$
$$x_2 : Entre\ (I, 0{,}28)^{0{,}019}\ y\ (LM, 0{,}17)^{-0{,}032},$$
$$x_3 : Entre\ (LM, -0{,}34)^{0{,}003}\ y\ (M, -0{,}13)^{-0{,}053},$$
$$x_4 : Entre\ (LP, -0{,}15)^{-0{,}025}\ y\ (M, -0{,}5)^{0{,}035}.$$

Calculando la magnitud de cada expresión ELICIT se obtiene el siguiente ranking: $x_3 \succ x_2 \succ x_4 \succ x_1$. Por lo tanto, la alternativa seleccionada es x_3: Reducción de residuos sólidos.

IV. Conclusiones

Este trabajo presenta un modelo MIMC para información ELICIT que ha sido aplicado en la resolución de un problema de TDG relacionado con EC, permitiendo obtener soluciones en consenso con mínimo coste.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Economía y Competitividad a través del Proyecto Nacional Español PGC2018-099402-B-I00 y la beca Postdoctoral Ramón y Cajal (RYC-2017-21978), el proyecto FEDER-UJA 1380637 y ERDF, por el Ministerio de Ciencia, Innovación y Universidades a través de una beca de Formación de Profesorado Universitario (FPU2019/01203) y por la Junta de Andalucía, Plan Andaluz para la Investigación, Desarrollo e Innovación (POSTDOC 21-00461).

Referencias

[1] J. Górecki, P. Núñez-Cacho, F. Corpas-Iglesias, and V. Molina, "How to convince players in construction market? strategies for effective implementation of circular economy in construction sector," *Cogent Engineering*, vol. 6, no. 1, p. 1690760, 2019.

[2] R. M. Rodríguez, Á. Labella, B. Dutta, and L. Martínez, "Comprehensive minimum cost models for large scale group decision making with consistent fuzzy preference relations," *Knowledge-Based Systems*, vol. 215, p. 106780, 2021.

[3] L. Martínez, R. M. Rodríguez, and F. Herrera, *The 2-tuple Linguistic Model*. Springer International Publishing, 2015.

[4] R. M. Rodríguez, L. Martínez, V. Torra, Z. Xu, and F. Herrera, "Hesitant fuzzy sets: state of the art and future directions," *International Journal of Intelligent Systems*, vol. 29, no. 6, pp. 495–524, 2014.

[5] Á. Labella, R. M. Rodríguez, and L. Martínez, "Computing with comparative linguistic expressions and symbolic translation for decision making: ELICIT information," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 10, pp. 2510–2522, 2019.

[6] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, part i, ii, iii," *Information Sciences*, vol. 8,8,9, pp. 199 – 249, 301–357, 43–80, 1975.

[7] S. Heilpern, "Representation and application of fuzzy numbers," *Fuzzy sets and Systems*, vol. 91, no. 2, pp. 259–268, 1997.

[8] "Data set with 10 experts," https://sinbad2.ujaen.es/sites/default/files/2022-05/ESTYLF_preferencias.pdf, accessed May 14, **2022**.

Detección del cyberbullying: Un enfoque difuso

1st J. Angel Diaz-Garcia
Departamento de Ciencias de la Computación e IA
Universidad de Granada
Granada, España
0000-0002-9263-1402

2nd Carlos Fernandez-Basso
Departamento de Ciencias de la Computación e IA
Universidad de Granada
Granada, España
0000-0002-8809-8676

3rd Jesica Gómez-Sánchez
Departamento de Psicología Evolutiva y de la Educación
Universidad de Granada
Granada, España
0000-0002-0247-1012

4th Karel Gutiérrez-Batista
Departamento de Ciencias de la Computación e IA.
Universidad de Granada
Granada, España
0000-0003-2711-4625

5th M. Dolores Ruiz
Departamento de Ciencias de la Computación e IA
Universidad de Granada
Granada, España
0000-0003-1077-3173

6th Maria J. Martin-Bautista
Departamento de Ciencias de la Computación e IA
Universidad de Granada
Granada, España
0000-0002-6973-477X

I. Resumen

Las redes sociales y las nuevas tecnologías están presentes hoy en día en casi todos los aspectos de las relaciones sociales. Estas tecnologías, que tienen tantas ventajas en la vida cotidiana de la sociedad, pueden convertirse en un arma de doble filo. Entre otros aspectos negativos que pueden desgranarse del uso de las redes sociales, uno de los más reconocidos y también dañinos es el cyberbullying.

El cyberbullying es un problema presente tanto en niños como en adolescentes, y puede definirse como un comportamiento intencionado, violento, cruel y repetitivo contra otro individuo. Este problema, omnipresente en la sociedad actual, está incluso catalogado por el Centro de Control y Prevención de Enfermedades (CDC) de EE.UU. como una grave amenaza para la salud pública [1]. El cyberbullying continuado puede contribuir a serios problemas psicológicos llegando incluso en algunos casos al suicidio [3]. Ser capaces de detectar esta problemática de una manera precoz y eficaz es por tanto una tarea de especial interés para la sociedad, y en la cual los sistemas basados en Inteligencia Artificial (IA) pueden tomar un papel muy relevante.

En este trabajo, proponemos un estudio del cyberbullying basado en técnicas de PLN (Procesamiento de Lenguaje Natural) y Minería de Datos, como el clustering difuso y los modelos de representación vectorial. La premisa inicial es que en un mismo mensaje pueden aparecer varios tipos de cyberbullying y su modelización con un grado de pertenencia puede ofrecer un gran potencial. Para validar estas premisas en cuanto a nuestros resultados, nos hemos apoyado en una validación por parte de expertos en Psicología. Las principales contribuciones de este trabajo al estado del arte son:

- Ofrecemos ua propuesta para el estudio difuso del cyberbullying basado en que un mismo mensaje puede contener ataques de diferentes tipos o características.
- La propuesta de un enfoque automático para obtener grados de pertenencia a uno o varios tipos de cyberbullying a partir de *sentence embeddings*.

Normalmente, los trabajos presentes en la literatura sobre técnicas de IA para detectar el cyberbullying se basan en algoritmos de clasificación supervisada [9], [7]. Recientemente también encontramos una prometedora rama basada en deep learning [2]. Estos trabajos, aunque ofrecen buenos resultados en términos de precisión y detección, necesitan de grandes bases de datos previamente etiquetadas, y de un proceso de entrenamiento exhaustivo. Además los basados en deep learning adolecen de que son modelos de tipo caja negra, es decir, poco explicables. Es por ello, que encontramos un nicho en la detección del cyberbullying para sistemas como el propuesto en este paper, interpretables y no supervisados.

Nuestro sistema está basado en 3 etapas: preprocesado de datos, creación de sentence embeddigns, y por último una capa de análisis de datos difusa. Para la experimentación se ha utilizado un dataset [10] compuesto por 46017 tweets con 6 categorías entre las que se encuentra no cyberbullying y 5 tipos de cyberbullying: género,religión, raza, edad y otros.

1) Preprocesado: Se han convertido los textos a minúsculas, se han eliminado los signos de puntuación, se han descartado las palabras que contienen dígitos, se han eliminado las palabras vacías y los espacios en blanco. Por último, se han eliminado los enlaces.
2) Sentence Embeddings: Los algoritmos de minería de datos no pueden trabajar con texto en bruto, es por

ello, que son necesarias técnicas de representación textual. Una de las más famosas, son las representaciones vectoriales también conocidos como word embeddings. Los word embeddings permiten representar cada palabra mediante un vector denso de longitud fija. Cada vector contiene el significado semántico de cada palabra para un contexto específico. Los word embeddings pueden ser estáticos [6] o contextualizados [4] , según guarden o no guarden información del contexto en que una palabra aparece. En este paper, utilizamos un enfoque contextualizado pero basado en frases, es decir, generamos un vector para cada frase en lugar de por palabra. Para ello hemos hecho uso de Sentence-BERT (SBERT) [8]. Tras su aplicación tenemos un vector denso de 768 dimensiones.

3) Análisis difuso: Tras tener los vectores, agrupamos cada vector en su categoría correspondiente (no cyberbulliying o su tipo de cyberbullying). El primer paso para calcular los grados de pertenencia es reducir la dimensión de los vectores resultantes por categoría a una sola dimensión. Para ello hemos usado *Uniform Manifold Approximation and Projection* (UMAP), [5]. Tras esto, hemos realizado el cálculo de la media y la desviación típica para cada clase, lo que nos ofrece los grados de pertenencia a cada una de ellas.

Para la experimentación por un lado usamos fuzzy k-means sobre el dataset, y por otro hemos utilizado el enfoque difuso descrito anteriormente en el cual hemos calculado los centroides aplicando la media y desviación típica sobre cada grupo de embeddigns. Los centroides obtenidos y los resultados son de mejor calidad en el caso de nuestro enfoque que en el ofrecido por el fuzzy k-means. Concretamente en el caso del fuzzy k-means, había un solapamiento de centroides que hacia que el coeficiente de silueta de los clusters sea de peor calidad que si lo comparamos con el caso de nuestro enfoque difuso donde se produce un menor solapamiento entre los centroides y por consiguiente, sobre los clusters. También, hemos encontrado que se produce una redistribución de las representaciones de las clases de cyberbullying, ofreciendo el estudio un mejor ajuste a la realidad, ya que antes de nuestro modelado difuso teníamos que un mensaje correspondía a un tipo de ataque y a ningún otro, ahora, tenemos mensajes que por ejemplo tienen pertenencia a ataques de religión y edad por igual, lo que puede ayudar a analizar y entender mejor la problemática del cyberbullying.

El enfoque presentado mejora la explicabilidad del análisis del cyberbullying cuando se superponen diferentes tipos de ataque en un mismo documento. El sistema también enriquece el número de ejemplos de una determinada clase de ataque al estudiarlos en conjunción con otros tipos, lo que aporta un gran valor por ejemplo para el posterior entrenamiento de algoritmos de clasificación. Para futuras investigaciones, proponemos utilizar este enfoque de procesamiento para crear un sistema de ayuda a la decisión que permita detectar y resolver casos de acoso en diferentes entornos como grupos de clase o redes sociales.

AGRADECIMIENTOS

La investigación presentada en este trabajo fue parcialmente apoyada por la Junta de Andalucía y el programa operativo FEDER en el marco del proyecto BigDataMed (P18-RT-2947 y B-TIC-145-UGR18) y la subvención PLEC2021-007681 financiada por MCIN / AEI / 10.13039 / 501100011033 y por la Unión Europea NextGenerationEU / PRTR. Por último, el proyecto también está parcialmente apoyado por el Ministerio de Educación, Cultura y Deporte de España (FPU18/00150). Además, este trabajo ha sido parcialmente apoyado por el Ministerio de Universidades a través del programa Margarita Salas - NextGenerationEU financiado por la UE.

REFERENCES

[1] Centers for disease control and prevention: Technology and youth: Protecting your child from electronic aggression. *Injury Prevention & Control: Violence Prevention Retrieved April*, 24:2013, 2009.

[2] Monirah Abdullah Al-Ajlan and Mourad Ykhlef. Deep learning algorithm for cyberbullying detection. *Int. J. Adv. Comput. Sci. Appl*, 9(9):199–205, 2018.

[3] Sheri Bauman, Russell B Toomey, and Jenny L Walker. Associations among bullying, cyberbullying, and suicide in high school students. *Journal of adolescence*, 36(2):341–350, 2013.

[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[5] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[6] Tomás Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. In Yoshua Bengio and Yann LeCun, editors, *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*, 2013.

[7] Sayanta Paul and Sriparna Saha. Cyberbert: Bert for cyberbullying identification. *Multimedia Systems*, pages 1–8, 2020.

[8] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019.

[9] Hugo Rosa, Nádia Pereira, Ricardo Ribeiro, Paula Costa Ferreira, Joao Paulo Carvalho, Sofia Oliveira, Luísa Coheur, Paula Paulino, AM Veiga Simão, and Isabel Trancoso. Automatic cyberbullying detection: A systematic review. *Computers in Human Behavior*, 93:333–345, 2019.

[10] Jason Wang, Kaiqun Fu, and Chang-Tien Lu. Sosnet: A graph convolutional network approach to fine-grained cyberbullying detection. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 1699–1708. IEEE, 2020.

Fuzzy Logic and Referring Expression Generation: Semantics, quantification, and measures. What else?

Nicolás Marín, Gustavo Rivas-Gervilla, and Daniel Sánchez
Department of Computer Science and Artificial Intelligence
University of Granada
Email: {nicm, griger, daniel}@decsai.ugr.es

***Abstract*—Fuzzy Logic is a suitable tool to deal with natural language in computer systems. In recent times there have been relevant contributions of this technology in the field of Natural Language Generation, for example providing tools to model the semantics of linguistic terms, and building linguistic summaries and descriptions of interest to the user in relation to the data in a dataset. However, the use of this technology poses many challenges, among which it stands out to facilitate its adaptation to the desired application domain in a context-aware and user-friendly way. In this work, we reflect on these topics in relation to referring expression generation.**

***Index Terms*—Fuzzy Logic, Computing with Words, Computational Theory of Perceptions, Natural Language Generation**

I. Introduction

Since its appearance, Fuzzy Logic has been applied in different fields and for several applications. Some relevant contributions of Fuzzy Logic can be found within the Natural Language Generation (NLG) field. The relation between Fuzzy Logic and natural language is highlighted by means of the Computational Theory of Perceptions and Computing with Words paradigm, proposed by Lotfi A. Zadeh [13], [14].

The language that humans use to verbally communicate is vast and complex, and there exists a huge amount of aspects that should be taken into account in order to elaborate a model that represents this language in a computer. For example, we can consider the scene in Figure 1, where different black and white objects with different sizes can be found.

In this case, to refer to the set $F_1 \cup F_2 \cup G$, we can consider the expression "the four small black circles" or "among the small circles", the four black ones". We can consider different descriptions or expressions to refer to the same set of objects. In each of these expressions, several linguistic terms and labels are employed, and the semantics of them have to be represented in the computer.

Within the NLG community, some problems have been associated to the use of Fuzzy Logic to manage the complexity of these linguistic terms employed in the human communication [12]:

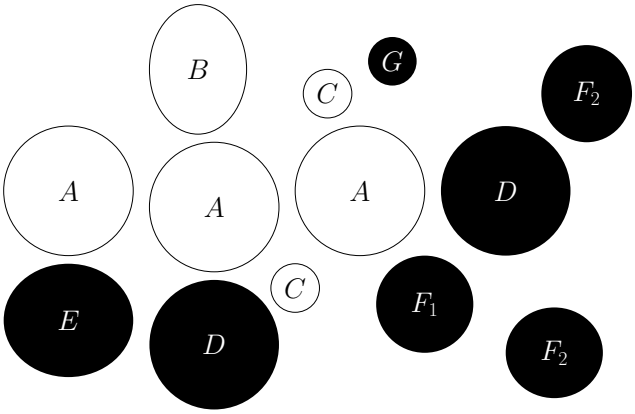


Fig. 1. An example scene that consists of a set of black and white geometric shapes of different sizes. The objects with the same color and dimensions are identified by the same letter that represents the set comprised of these objects.

- The complexity of defining the proper membership function to represent a certain linguistic term by means of a fuzzy set.
- The possibility of using different operators to define the same operation between fuzzy sets, e.g., in order to combine the semantics of various linguistic terms.

Despite these criticisms, Fuzzy Logic has been employed in the NLG field with important applications and relevant results, e.g. [4], [6], [11]. In the following, we illustrate the power of the use of Fuzzy Logic for the representation and management of natural language in a computer.

II. Semantics

In the previous expressions, there are several linguistic labels as "small" or "circle". These labels can be seen as vague concepts since, for example, each object in the scene can be considered small in a certain degree. In addition, some objects in Figure 1 are not perfect circles but ellipses, among them the objects in group F_2; these objects could be identified as circles in a certain degree according to their eccentricity.

From the perspective of Fuzzy Logic, linguistic labels can be modelled as fuzzy sets, where the membership degree of each element in the underlying domain represents to which extent the corresponding linguistic label can be applied to the element. Several techniques have been proposed to obtain these membership functions automatically [1], [3], [10].

Authors appear in alphabetical order. All authors have participated in conceptualization and research tasks.

This work has been partially supported by the Spanish Ministry of Science, Innovation and Universities and the European Regional Development Fund - ERDF (Fondo Europeo de Desarrollo Regional - FEDER) under project PGC2018-096156-B-I00 "Recuperación y Descripción de Imágenes mediante Lenguaje Natural usando Técnicas de Aprendizaje Profundo y Computación Flexible".

The correct definition of these membership functions is a crucial task for the performance of the system. This definition can be more complex when we consider context-aware properties like size, since the accomplishment degree of each object varies with the context in which the object is considered.

The enumerated works can be seen as a proof of the potential of Fuzzy Logic to overcome the first problem that the NLG community has attributed to it. However, some questions arise from these studies about the modelling of linguistic terms. **How to define methods to obtain membership functions for properties with different complexity? How does the order in which properties are considered affect those membership functions? How to guarantee coherence between membership functions corresponding to related or compound terms? Among many.**

In addition to the modelling of these linguistic labels, in the expression that human daily use other linguistic elements appear as, e.g., quantifiers.

III. Quantification

In the previous expressions, not only adjectives appear and contribute to the information provided, but quantifiers are also employed. The extent to which a certain quantifier applies to a set of objects is also a matter of degree. We can think of the expression "the seven black circles". Since the linguistic label "circle" is not fully accomplished by the objects in $E \cup F_2$, the truth value of this expression is gradual. In addition, other more imprecise quantifiers as "the majority" or "just a few" have a vague linguistic definition which has to be represented in the system. And there exist several types of generalized quantifiers [5].

Quantification can be employed in different ways. For instance, "the seven small objects" in an image can serve to point to the seven smallest objects when there are more than seven objects that are small to some degree. If there are only seven small objects, then the quantifier is employed to enforce the expression by means of overspecification. Another possibility is to employ quantification in combination with relational properties, as in "the object that has more triangles to the right than to the left", as explained in [9].

Again, many questions arise. **How to model the membership functions of quantifiers? How to choose the most appropriate quantifier in a certain context? How to guarantee coherence between the different quantifiers employed in a description? Among others.**

In any case, a proper calculation of the accomplishment degree of a certain quantified sentence needs to be defined, satisfying given properties [2].

IV. Measures

Finally, as we have shown at the beginning of this text, several descriptions can be generated to describe or refer to the same set of objects. Since we are working with linguistic labels, Fuzzy Logic brings the tools needed to compute the degree of specificity and referential success of a certain expression, two very related concepts.

Different users and applications will imply a different semantics of specificity and referential success. For example, we will require a more or less strict semantics of referential success depending on the type of system that we are building. Then, we need a wide catalog of measures to use the one that better accomplishes our objectives. In this line, several works to derive new measures of referential success and specificity has been proposed [7], [8]. **The development of proper specificity and referential success measures is an open research area, with open problems such as analysing the impact of different kind of properties, as context-aware or collective ones, in the specificity and referential success measurement.**

V. Conclusions

Fuzzy Logic has been shown to be an appropriate tool to manage the vagueness typical of linguistic terms, with different works that have addressed several problems and questions related with the generation of natural language and with the representation of linguistic terms.

According to the different questions that we have introduced previously and the vast challenge that representing and modelling the natural language as a whole is, many other problems remain open in this promising research field.

References

[1] M.-S. Chen and S.-W. Wang. Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets Syst.*, 103(2):239–254, 1999.

[2] M. Delgado, M. D. Ruiz, D. Sánchez, and M. A. Vila. Fuzzy quantification: a state of the art. *Fuzzy Sets Syst.*, 242:1–30, 2014.

[3] A. Homaifar and E. McCormick. Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *IEEE Trans. Fuzzy Syst.*, 3(2):129–139, 1995.

[4] J. Kacprzyk and S. Zadrozny. Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries, and natural-language generation. *IEEE Trans. Fuzzy Syst.*, 18(3):461–472, 2010.

[5] E. L. Keenan and J. Stavi. A semantic characterization of natural language determiners. *Linguist Philos.*, 9(3):253–326, 1986.

[6] N. Marín, G. Rivas-Gervilla, and D. Sánchez. Scene selection for teaching basic visual concepts in the refer4learning app. In *2017 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2017, Naples, Italy, July 9-12, 2017*, pages 1–6. IEEE, 2017.

[7] N. Marín, G. Rivas-Gervilla, D. Sánchez, and R. R. Yager. Specificity measures and referential success. *IEEE Trans. Fuzzy Syst.*, 26(2):859–868, 2018.

[8] N. Marín, G. Rivas-Gervilla, D. Sánchez, and R. R. Yager. Specificity measures based on fuzzy set similarity. *Fuzzy Sets Syst.*, 401:189–199, 2020.

[9] N. Marín, G. Rivas-Gervilla, and D. Sánchez. Fuzzy quantified relational properties for referring expression generation. WCCI 2022, Accepted.

[10] D. Özdemir and L. Akarun. A fuzzy algorithm for color quantization of images. *Pattern Recognit.*, 35(8):1785–1791, 2002.

[11] A. Ramos-Soto, A. Bugarín Diz, S. Barro, and J. Taboada. Linguistic descriptions for automatic generation of textual short-term weather forecasts on real prediction data. *IEEE Trans. Fuzzy Syst.*, 23(1):44–57, 2015.

[12] K. van Deemter. *Not Exactly: in Praise of Vagueness*. Oxford University Press, 2010.

[13] L. A. Zadeh. A new direction in AI: toward a computational theory of perceptions. *AI Mag.*, 22(1):73–84, 2001.

[14] L. A. Zadeh. From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. In D. Vanderveken, editor, *Logic, Thought and Action*, volume 2 of *Logic, Epistemology, and the Unity of Science*, pages 507–544. Springer, 2005.

Framework para la Descripción Automática de Procesos en Lenguaje Natural: Aplicación en un Proceso Asistencial Integrado de Estenosis Aórtica

Yago Fontenla-Seco*, Manuel Lama*, Violeta González-Salvado†, Carlos Peña-Gil†, Alberto Bugarín-Diz*
*Centro Singular de Investigación en Tecnologías Intelixentes (CiTIUS), Universidade de Santiago de Compostela
Santiago de Compostela, España
{yago.fontenla.seco, manuel.lama, alberto.bugarin.diz}@usc.es
†Departamento de Cardiología, Hospital Clínico Universitario de Santiago de Compostela, SERGAS IDIS CIBERCV
Santiago de Compostela, España
{violeta.gonzalez.salvado, carlos.pena.gil}@sergas.es

***Resumen*—Presentamos un resumen del trabajo titulado *A framework for the automatic description of healthcare processes in natural language: Application in an aortic stenosis care process* publicado originalmente en el *Journal of Biomedical Informatics* [1]. En él, presentamos un *framework* para la generación automática de descripciones en lenguaje natural de procesos en el dominio de la salud. El *framework* se basa en la arquitectura más empleada en sistemas *Data-to-text* (D2T) [2] y está respaldado por un modelo general capaz de manejar datos de procesos. Así, el *framework* es capaz de cuantificar características del proceso en el tiempo, extraer relaciones temporales y tiempos de espera entre actividades y sus posibles causas y comparar atributos del proceso entre grupos, entre otras características. También integra técnicas de cuantificación borrosa para representar información cuantitativa relevante para el proceso con cierto grado de incertidumbre y describir esta información en lenguaje natural utilizando términos imprecisos. Se presenta una aplicación real del *framework* en el ámbito de la cardiología: el Proceso Asistencial Integrado del Estenosis Aórtica del Hospital Clínico Universitario de Santiago de Compostela. La aplicación ha sido validada por quince expertos en Cardiología, mostrando resultados muy positivos que permiten concluir que el lenguaje natural es la modalidad más eficiente de trasladar información a profesionales médicos, que existe una preferencia clara por el lenguaje natural frente a representaciones gráficas y que las descripciones en lenguaje natural proporcionan información relevante y útil sobre el proceso.**

***Index Terms*—Minería de procesos, Generación de Lenguaje Natural, Cuantificación borrosa, Descripciones lingüísticas de datos.**

I. Introducción

Aquí se resume el trabajo publicado originalmente en el *Journal of Biomedical Informatics vol.128* [1]. Los procesos permiten a las organizaciones representar y estructurar las actividades que tienen lugar en ellas y sus correspondientes sistemas de información. En organizaciones del ámbito sanitario los procesos sirven para estructurar las actividades clínicas y organizativas que se llevan a cabo con el objetivo de diagnosticar, tratar y prevenir problemas en la salud de los pacientes. Debido a grandes presiones para mejorar su efectividad y eficiencia, aparece la necesidad de comprender y mejorar estos procesos [3], [4]. La minería de procesos es una disciplina que tiene como objetivo explotar registros de eventos de procesos [1]. Mediante el descubrimiento del modelo de proceso subyacente a un registro de eventos, es capaz de extraer información relevante que ayuda a comprender y mejorar un proceso. Sin embargo, por su propia naturaleza (dinámicos, complejos, *ad-hoc*, y multi-disciplinarios [5]) los procesos de salud son complejos (*procesos espagueti*) y muy difíciles de comprender por expertos médicos, ya que no están especializados en el campo de los procesos ni en la visualización de analíticas de forma gráfica [3], [6].

En este sentido, la Generación de Lenguaje Natural (GLN) es una buena aproximación para mejorar la comprensión de un proceso, en particular de procesos en el dominio de la sanidad. Esta disciplina tiene como objetivo proporcionar al usuario textos capaces de trasladar las características más relevantes de un conjunto de datos en una forma fácilmente consumible [7]. Del mismo modo, la incorporación de términos imprecisos es muy efectiva para la sumarización y comunicación de estas características relevantes [8]. Partiendo de estas ideas, se presenta un *framework* para la generación de descripciones cualitativas y cuantitativas de procesos en lenguaje natural que integra técnicas de minería de procesos para la extracción de características relevantes de un proceso con técnicas de lógica difusa para manejar la imprecisión inherente del lenguaje natural mediante el modelado de expresiones borrosas.

II. Propuesta

El *framework* se desarrolla en cuatro etapas: preprocesado, interpretación de datos, planificación del documento y reali-

Esta investigación está financiada por el Ministerio de Ciencia e Innovación del Gobierno de España (ayudas PRE2018-083719, PID2020-112623GB-I00 y PDC2021-121072-C21), el Ministerio de Educación, Universidades y Formación Profesional de la Xunta de Galicia y el programa ERDF/FEDER (ayudas ED431G2019/04 y ED431C2022/19). Todas las ayudas han sido cofinanciadas por el Fondo Europeo de Desarrollo Regional (programa EDRF/FEDER)

[1]Mientras un proceso se ejecuta, toda actividad realizada es registrada en un registro de eventos. Cada instancia individual del proceso se conoce como caso y registra las actividades y recursos involucrados en dicha ejecución.

zación lingüística. En la primera etapa, el registro original de eventos es preprocesado para obtener un registro de eventos apto para la aplicación de técnicas de minería de procesos. La etapa de interpretación de datos se divide en varias fases, primero, partiendo del registro de eventos preprocesado, se extrae el modelo de proceso subyacente mediante algoritmos de descubrimiento. Después se lleva a cabo un análisis del proceso: i) se realiza el *replay* del registro de eventos sobre el modelo descubierto, ii) información sobre las perspectivas [2] temporal y de control de flujo del proceso se suman a la perspectiva de casos extraída del registro de eventos y iii) se lleva a cabo una extracción de las características más relevantes del proceso. Esta operación permite representar el registro de eventos (enriquecido) de manera tabular, al que se pueden aplicar técnicas de análisis de datos, cuantificación borrosa y sumarización lingüística. Concretamente, se utiliza conocimiento experto del dominio y técnicas de análisis estadístico para guiar la generación de sentencias cuantificadas borrosas en la fase de planificación del documento, que posteriormente se utilizarán para generar las descripciones en lenguaje natural. Para la generación de estas sentencias, basadas en el concepto de resúmenes lingüísticos de datos [9] y de protoforma [10] proponemos una serie de extensiones de las protoformas de tipo-I y tipo-II, definiendo nuevas protoformas más complejas y enriqueciendo su semántica [11]:

- Contextualización temporal de características: "En el año 2019, *en algunos casos (38,8 %)*, el paciente tuvo admisión de emergencia en el proceso."
- Frecuencia y relaciones temporales entre actividades: "*En aproximadamente la mitad de los casos (56,28 %)* en los que el paciente fue intervenido mediante TAVI, al paciente se le realizó un TAC *después* de su sesión médico-quirúrgica."
- Conformidad: "Se espera que el tiempo de espera entre la sesión médico-quirúrgica de un paciente y su intervención sea *menor a 30 días*. En cambio, *en algo más de la mitad de los casos* (59,30 %), el paciente tuvo un tiempo de espera entre la sesión médico-quirúrgica y su intervención *superior a 30 días.*"
- Comparación entre grupos de pacientes: "En *aproximadamente la mitad de los casos (42.25 %)* en los que el paciente fue intervenido *a lo largo del año 2019*, el paciente tuvo un tiempo de espera entre la sesión médico-quirúrgica y su intervención menor a 30 días. En cambio, *en muchos casos (77,50 %)* en los que el paciente fue intervenido *a lo largo del año 2020*, el paciente tuvo un tiempo de espera entre la sesión médico-quirúrgica y su intervención menor a 30 días."

Para la fase de realización lingüística se adoptó una estrategia basada en plantillas dinámicas, conocimiento experto y la librería de realización lingüística SimpleNLG-ES [12].

[2]Dimensiones en las que se analiza un proceso: la perspectiva de flujo se centra en el orden en el que se ejecutan las actividades, la perspectiva de casos se centra en las propiedades de estos, la perspectiva de recursos se centra en los recursos que intervienen en el proceso y la perspectiva temporal concierne la frecuencia de ejecución de actividades, su duración y tiempos de espera.

III. Resultados

Siguiendo los estándares habituales en sistemas GLN, se realizó una evaluación intrínseca manual por parte de 15 profesionales médicos, mediante un cuestionario en el que valoraron el interés y la calidad de la información presentada en las descripciones, su facilidad de comprensión, su eficiencia a la hora de trasladar la información frente a la representación gráfica usada típicamente en minería de procesos, y la la utilidad e intención de las descripciones en su entorno laboral. Los resultados indicaron un muy alto grado de satisfacción. Sobre una escala Likert de 1 a 5, los expertos encontraron las descripciones muy interesantes (4.11) y fáciles de comprender (4.40) y mostraron una clara preferencia por las descripciones en lenguaje natural frente a las representaciones gráficas habituales (4.28). También consideraron este tipo de descripciones útiles (4.12), que les ayudan a entender mejor lo que ocurre en su trabajo (4.06) y a completar tareas más rápido (4.06) y mejoran la calidad del trabajo que realizan (3.69).

Por tanto, el *framework* propuesto en este trabajo es completo (capaz de manejar todas las fases de generación de las descripciones desde la entrada de un registro de eventos hasta la salida de descripciones en lenguaje natural) y capaz de manejar de manera imprecisa información de procesos del dominio sanitario y que las descripciones que se generan con él son comprensibles, útiles y ayudan a los expertos a realizar su trabajo mejor y más eficientemente.

Referencias

[1] Y. Fontenla-Seco, M. Lama, V. González-Salvado, C. Peña-Gil, A. Bugarín-Diz, A framework for the automatic description of healthcare processes in natural language: Application in an aortic stenosis integrated care process, Journal of Biomedical Informatics 128 (2022) 104033.

[2] E. Reiter, An Architecture for Data-to-Text Systems, in: Proc. 11th European Workshop on Natural Language Generation, ENLG '07, ACL, USA, 2007, pp. 97–104.

[3] R. Mans, W. Aalst, van der, R. Vanwersch, Process mining in healthcare: evaluating and exploiting operational healthcare processes, SpringerBriefs in Business Process Management,, Springer, Germany, 2015.

[4] R. S. Mans, M. H. Schonenberg, M. Song, W. M. P. van der Aalst, P. J. M. Bakker, Application of process mining in healthcare – a case study in a dutch hospital, in: Biomedical Engineering Systems and Technologies, Springer Berlin Heidelberg, 2009, pp. 425–438.

[5] P. Homayounfar, Process mining challenges in hospital information systems, 2012 Federated Conference on Computer Science and Information Systems (FedCSIS) (2012) 1135–1140.

[6] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: A literature review, Journal of Biomedical Informatics 61 (2016) 224–236.

[7] E. Reiter, R. Dale, Building Natural Language Generation Systems, Cambridge University Press, USA, 2000.

[8] A. Ramos-Soto, A. Bugarín, S. Barro, On the role of linguistic descriptions of data in the building of natural language generation systems, Fuzzy Sets and Systems 285 (2016) 31–51.

[9] R. R. Yager, A new approach to the summarization of data, Information Sciences 28 (1) (1982) 69–86.

[10] L. A. Zadeh, A prototype-centered approach to adding deduction capability to search engines-the concept of protoform, in: Proc. NAFIPS-FLINT 2002, 2002, pp. 523–525.

[11] A. Ramos-Soto, P. Martín-Rodilla, Enriching linguistic descriptions of data: A framework for composite protoforms, Fuzzy Sets Syst. 407 (2021) 1–26.

[12] A. Ramos-Soto, J. Janeiro-Gallardo, A. Bugarín, Adapting SimpleNLG to Spanish, in: 10th International Conference on NLG, Association for Computational Linguistics, 2017, pp. 142–146.

Método basado en word embeddings para la adaptación no supervisada de recetas de cocina

Andrea Morales-Garzón
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
amoralesg@ugr.es

Juan Gómez-Romero
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
jgomez@decsai.ugr.es

Maria J. Martin-Bautista
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
mbautis@decsai.ugr.es

***Resumen*—El estudio de los hábitos alimentarios y el estilo de vida de la población es indispensable para entender la ciencia de los alimentos, y así poder automatizar tareas cotidianas relacionadas con el mundo culinario. Una de las tareas más comunes, y a su vez de las más complejas de automatizar, consiste en generar recetas adaptadas en función de restricciones concretas del usuario, como alergias, intolerancias, o restricciones autoimpuestas por el tipo de dieta (p. ej. dietas vegetarianas o hipocalóricas). En este trabajo, proponemos un método para adaptar recetas a través de la sustitución de sus alimentos por otros que, manteniendo la coherencia de la receta, cumplan los requisitos del usuario. Para ello, se usa un modelo predictivo de lenguaje entrenado en el dominio alimenticio para obtener una representación textual de los alimentos. Con una función de distancia difusa mapeamos los ingredientes de la receta a una tabla externa de composición de alimentos, expandiendo así la información nutricional de los alimentos para poder encontrar sustitutos. Se han implementado tres tipos de adaptación de recetas distinguiendo entre la adaptación basada en preferencias y con/sin restricciones. Con un 95 % de confianza, el método obtiene adaptaciones de recetas de calidad gracias a su capacidad de aprender relaciones intrínsecas entre los alimentos.**

***Palabras Claves*—Mapeo de datos, Adaptación de recetas, Food computing, Word embedding, Procesamiento de lenguaje natural**

I. Introducción

Una nutrición saludable es esencial para el desarrollo de la vida humana, ya que nuestros hábitos alimenticios tienen impacto directo en nuestra salud y calidad de vida. Con las técnicas adecuadas, los datos alimenticios se pueden utilizar para mayor comprensión de esta área de estudio, especialmente desde un punto de vista centrado en hábitos saludables. En este contexto se introduce el concepto de food computing, que engloba las técnicas y métodos computacionales que usan datos procedentes de fuentes de datos alimenticias [1].

Una de las líneas principales en food computing se centra en estudiar combinaciones de alimentos para adaptar recetas a las necesidades específicas de la población. Estas necesidades pueden ser muy diversas: un usuario podría querer modificar una receta porque no tenga disponibilidad de un ingrediente concreto en su cocina, por alergia o intolerancia a algún alimento, por razones médicas, por tipo de dieta, o simplemente por sus preferencias sobre algunos ingredientes concretos.

La forma clásica de abordar el problema es con ontologías que recojan el concepto de ingrediente y permitan extraer entidades similares. Además, se han utilizado aproximaciones como razonamiento basado en casos para modificar recetas manteniendo la coherencia. Aproximaciones más recientes buscan relaciones complejas estudiando la co-ocurrencia de ingredientes en recetas para encontrar posibles sustituciones. También estudiando la frecuencia relativa con la que los ingredientes coinciden, técnicas basadas en grafos y usando pseudorecetas para obtener recetas sintéticas personalizadas [2].

Este trabajo, publicado en [2], considera textos de preparación de recetas para aprender estructuras complejas y relaciones intrínsecas entre alimentos. A diferencia de trabajos previos, centrados en el análisis a nivel de ingrediente, se usan las representaciones alimenticias aprendidas por el modelo para mapear los ingredientes a bases de conocimiento externas que nos permitan agregar información nutricional y tomar decisiones especializadas acerca de los ingredientes de la receta. Esta fusión de datos es realmente relevante en nuestro caso, ya que existe la dificultad añadida de trabajar de forma conjunta con datos de naturaleza heterogénea, como pueden ser datos nutricionales especializados o textos de expertos.

Las principales contribuciones del trabajo son las siguientes: (1) modelo de word embedding entrenado en el dominio específico en nutrición, (2) método automático para mapear recursos de origen alimenticio, (3) algoritmo no supervisado de adaptación de recetas a partir de restricciones de usuario.

II. Método propuesto

El método que proponemos realiza adaptaciones de recetas a nivel de ingrediente, permitiendo modificar los ingredientes de una receta por otros alimentos más adecuados según las especificaciones del usuario. En la publicación original [2] se puede consultar la descripción detallada de la metodología.

Word embedding. Utilizamos un word embedding para obtener representaciones numéricas de textos alimenticios. Este modelo nos permite representar vocabulario en un espacio semántico multidimensional donde las distancias y relaciones entre las palabras del vocabulario son significativas. Seleccionar el algoritmo de word embedding más adecuado no es trivial, por lo que realizamos un estudio previo de los más

utilizados (word2vec, fasttext y Glove) para detectar el más apropiado para la tarea. Concluimos que se comportan de forma similar, por lo que usamos Word2vec por su simplicidad.

Mapeo de ingredientes. Proponemos una medida de distancia difusa entre conjuntos para detectar alimentos equivalentes/similares a partir de sus descripciones. Cada descripción es un conjunto que contiene los tokens que la forman, representados por el vector obtenido con el modelo de lenguaje. Se calcula como la distancia euclidiana entre los vectores de tokens de ambos conjuntos. Así, consideramos la vaguedad del lenguaje y la diferencia del nivel de detalle de las descripciones. La métrica está detallada en la publicación original [2].

Adaptación de la receta. Se parte de una receta y de la lista de restricciones del usuario (si procede). Con el word embedding se obtienen representaciones de los ingredientes de la receta. Con la función de distancia se comparan con las representaciones de los alimentos (que cumplan las restricciones) de la tabla de composición de alimentos, trabajando así de forma unificada con los dos conjuntos de datos. Finalmente se devuelven las opciones más similares para sustituir los ingredientes de la receta. Consideramos tres formas de adaptar recetas: (1) adaptación sin restricciones, donde se sugieren alimentos alternativos por los que sustituir los originales; (2) adaptación basada en preferencias, donde se sugieren alimentos que encajen más con lo que le gustaría al usuario (p.ej., alimentos bajos en grasa); (3) adaptación con restricciones, en este caso dietas restrictivas: la vegetariana y la vegana.

Validación. Evaluar un método de adaptación de recetas es complejo, ya que es un proceso muy subjetivo y dependiente de factores externos como la cultura o los gustos personales. Por ello, realizamos una encuesta online donde los usuarios pueden evaluar la calidad de las adaptaciones generadas.

III. Datos

En total se utilizan tres conjuntos de datos: (1) un corpus de recetas del que extraemos los textos de preparación de recetas con los que entrenar el modelo de lenguaje. Está formado por 267,071 recetas obtenidas de páginas webs de cocina especializadas; (2) un segundo corpus de recetas con el que realizar la validación de la adaptación de recetas, con recetas procedentes de Food.com, diferentes de las utilizadas en el entrenamiento del modelo; y (3) una tabla de composición de alimentos con su información nutricional para extrapolar los datos nutricionales de los ingredientes de la receta a adaptar y obtener sustitutos. En concreto se ha utilizado la tabla CoFID, con 2913 alimentos y su información nutricional asociada.

IV. Experimentos y resultados

La experimentación, resultados y las visualizaciones y ejemplos obtenidos, se pueden ver en la publicación original [2].

Word embedding. Para entrenar el modelo se realiza un preprocesamiento clásico a los textos de preparación, siendo cada frase un dato de entrada. Los hiperparámetros del modelo se eligen a partir de los mejores resultados obtenidos en la experimentación. Las visualizaciones realizadas muestran que el modelo es capaz de aprender relaciones semánticas entre datos, ilustrando que la distancia entre unos grupos de alimentos y otros se ve influida por su uso conjunto en recetas.

Mapeo de ingredientes. Con la medida de distancia difusa se obtiene el mapeo entre los ingredientes de las recetas de Food.com y sus alimentos más similares en la tabla de composición de alimentos. Los resultados mostrados en [2] muestran la eficacia de la función difusa para identificar alimentos semejantes y alternativos.

Adaptación de la receta. Combinando el word embedding con la medida de distancia anterior podemos acceder a información especializada de los ingredientes y extrapolarlos a nivel de receta para adaptarla de forma personalizada. Es importante tener en cuenta que esta función nos permite extender la información nutricional de los ingredientes para comprobar si cumplen o no las restricciones del usuario y saber cuáles modificar. Cabe destacar los buenos resultados obtenidos con la evaluación de los usuarios teniendo en cuenta que el método propuesto es no supervisado. En la encuesta participaron 40 usuarios con conocimiento general sobre cocina. Con un 95 % de confianza, las puntuaciones corresponden a niveles de satisfacción altos por parte del usuario, siendo categorizadas la mayoría de las recetas con buenas puntuaciones.

V. Conclusiones y Trabajos Futuros

En este trabajo se muestra que utilizar modelos de word embedding para representar alimentos permite obtener representaciones de calidad, debido a su capacidad de extraer relaciones e información intrínseca directamente del texto. Estas representaciones son especialmente útiles para adaptar recetas cuando van acompañadas de una métrica adecuada. En concreto, proponemos una métrica difusa, que permite mapear con éxito los ingredientes de las recetas con un recurso externo, en este caso, una tabla de información nutricional de alimentos. Los resultados obtenidos con la validación no sólo ilustran el buen funcionamiento del método propuesto, sino que confirman la calidad de las representaciones numéricas de los alimentos, pudiendo aplicarse estas a otros problemas de Food Computing. Como trabajo futuro, pretendemos extender el método de adaptación para poder considerar distintos estilos de cocina del mundo, así como utilizar modelos de lenguaje más avanzados para aprender representaciones de alimentos más sofisticadas y dependientes del contexto.

Agradecimientos

Este trabajo ha sido financiado por la Junta de Andalucía y el Fondo Europeo de Desarrollo Regional (FEDER) por medio del proyecto BigDataMed (P18-RT-2947 y B-TIC-145-UGR18) y por la Consejería de Transformación, Industria y Conocimiento de la Junta de Andalucía a través del programa de ayudas para formación predoctoral 2021.

Referencias

[1] W. Min, S. Jiang, L. Liu, Y. Rui, and R. Jain, "A survey on food computing," *ACM Computing Surveys (CSUR)*, vol. 52, no. 5, pp. 1–36, 2019.

[2] A. Morales-Garzón, J. Gómez-Romero, and M. J. Martin-Bautista, "A word embedding-based method for unsupervised adaptation of cooking recipes," *IEEE Access*, vol. 9, pp. 27 389–27 404, 2021.

Aplicación visual para la recuperación de imágenes caracterizadas mediante aprendizaje profundo usando consultas basadas en la presencia de objetos y de relaciones entre ellos

Juan Miguel Medina
Ciencias de la Computación e I.A.
Universidad de Granada
18071 Granada, España
medina@decsai.ugr.es

Miguel Medina-Martínez
Universidad de Granada
18071 Granada, España
miguemedina@correo.ugr.es

Daniel Sánchez
Ciencias de la Computación e I.A.
Universidad de Granada
18071 Granada, España
daniel@decsai.ugr.es

***Resumen*—Este documento presenta una aplicación multiplataforma que permite elaborar de forma visual consultas en las que se describen los objetos que deben aparecer en una imagen, los colores dominantes presentes en dichos objetos y las relaciones proporcionales y espaciales entre los mismos. Estas consultas también se pueden expresar mediante una sintaxis basada en lenguaje natural. La ejecución de estas consultas sobre una base de datos, construida sobre Neo4J® recupera aquellas imágenes cuyas características se corresponden con los parámetros establecidos en dichas consultas.**

***Index Terms*—Recuperación de imágenes basada en contenido**

I. Introduction

La recuperación de imágenes mediante consultas de alto nivel semántico es un reto de la inteligencia artificial que se ha plasmado en diferentes propuestas como puede ser el servicio Google Imágenes. Este servicio permite hacer búsqueda de imágenes introduciendo palabras clave en su buscador (de igual forma que la herramienta de búsqueda de imágenes clásica). Aunque en general este servicio resulta efectivo, presenta algunas limitaciones debido al mecanismo que utiliza para indexar las imágenes en su base de datos. Para extraer e indexar los descriptores y etiquetas de una imagen, que luego serán utilizados por sus algoritmos de búsqueda, tiene en cuenta el nombre del archivo de la imagen, la descripción alternativa (ALT) que el diseñador web asocia a dicha imagen y el contexto de la página web en la que se encuentra insertada. Todo esto es muy efectivo para el enorme volumen de imágenes que este buscador indexa, pero limita la precisión en cuanto a los resultados obtenidos para búsquedas más elaboradas. Por ejemplo, si introducimos la búsqueda "moto negra a la derecha de coche rojo", nos devuelve imágenes de motos negras, de motos rojas, de coches rojos, de coches negros, pero en ningún caso una imagen que contenga una moto negra a la derecha de coche rojo, ni siquiera imágenes donde aparezcan un coche y una moto, del color que sea.

En [1] se presenta un algoritmo basado en el uso de técnicas de aprendizaje profundo para detectar y etiquetar objetos presentes en una imagen y el grado de pertenencia a una categoría (gato, casa, etc.). Además se obtienen los colores dominantes presentes en estos objetos y las relaciones de proporción y espaciales entre los mismos, también con grados asociados. Todo ello permite caracterizar semánticamente cada imagen mediante un grafo difuso en el que los nodos representan los objetos en ella detectados y a los que, a su vez, se asocian los nodos que representan sus colores dominantes. Por último, las relaciones de proporción y espaciales entre los objetos se representan mediante arcos entre ellos.

Con las imágenes caracterizadas y representadas de esta manera en un Sistema de Gestión de Bases de Datos orientado a Grafos, Neo4J® [2] en nuestro caso, se pueden acometer con éxito consultas del tipo de la indicada más arriba, e incluso más complejas. La aplicación multiplataforma que se presenta facilita la creación y ejecución de este tipo de consultas, proporcionando mecanismos visuales para elaborarlas y también una sintaxis basada en lenguaje natural para expresarlas.

II. Arquitectura y Funcionamiento del Sistema

La Figura 1 ilustra el sistema con el que interactúa la aplicación que se presenta. Consta de un *Repositorio de Imágenes* que contiene las imágenes que van a ser procesadas para su posterior consulta. Mediante el *Módulo de Extracción de Características* se obtiene cada imagen desde el *Repositorio de Imágenes* (paso (1) en la Fig. 1), se procesa mediante las técnicas descritas en [1] y, como resultado, se genera un grafo con los objetos, colores y relaciones detectados en la imagen, expresado mediante una sentencia Cypher® [3], cuya ejecución insertará dicho grafo en la base de datos Neo4J® [2] (paso (2) en la Fig. 1). Este módulo también podrá ser invocado desde la *Aplicación de Consulta de Imágenes* para enviarle una imagen o un lote de imágenes para que las procese y actualice el servidor Neo4J® con los grafos descriptores extraídos.

El servidor de bases de datos orientado a grafos Neo4j® proporciona la estructura de datos ideal para representar las

características de alto nivel semántico extraídas desde las imágenes por el *Módulo de Extracción de Características* y también para elaborar complejas consultas que involucren todas estas características con las que obtener las imágenes que las satisfacen. Puesto que este tipo de consultas involucran la búsqueda de imágenes que contengan determinados objetos (personas, animales, vehículos, etc), con colores definidos para dichos objetos, con posiciones definidas entre dichos objetos e incluso satisfaciendo unas determinadas proporciones entre los mismos, la posibilidad de elaborar un boceto gráfico que exprese dichas condiciones puede ayudar definir semánticamente estas consultas.

La *Aplicación de Consulta de Imágenes* proporciona una intuitiva interfaz para elaborar dicho boceto de consulta. En la Figura 1 se ilustra un ejemplo de boceto que representa la consulta: "Obtén imágenes con una moto negra a la izquierda de una moto negra a la izquierda de una persona azul a la izquierda de una persona azul". A partir de este boceto, la *Aplicación de Consulta de Imágenes* genera la sentencia de consulta en la sintaxis Cypher® y la envía para su ejecución al servidor Neo4J® (paso (3) en la Fig. 1), recuperando desde el mismo las URLs de las imágenes que satisfacen la consulta (paso (4) en la Fig. 1); por último, mediante esas URLs recupera desde el *Repositorio de Imágenes* las imágenes concordantes y las visualiza (paso (5) en la Fig. 1). El texto de la consulta ilustrada también se puede enviar directamente desde la aplicación, donde un procesador de lenguaje natural puede generar la sentencia del paso (3) y, a partir de ahí, obtener también las imágenes concordantes.

III. Conclusiones y trabajos futuros

Se ha descrito someramente una aplicación que permite elaborar consultas de alta capacidad semántica, tanto en modo visual, como mediante expresiones en lenguaje natural. El prototipo desarrollado se puede ejecutar como una aplicación de escritorio, aplicación web y aplicación móvil.

Como trabajo futuro: mejorar la interfaz de la aplicación e incluir una opción para describir las imágenes recuperadas para personas con deficiencias visuales. También avanzar en la integración de todos los componentes del sistema.

Acknowledgment

Este trabajo ha sido soportado parcialmente por el Ministerio de Ciencia, Innovación y Universidades y ERDF (Fondo Europeo de Desarrollo Regional - FEDER) mediante el projecto **PGC2018-096156-B-I00** *Recuperación y Descripción de Imágenes mediante Lenguaje Natural usando Técnicas de Aprendizaje Profundo y Computación Flexible*

Referencias

[1] J. Chamorro-Martínez, N. Marín, M. Mengíbar-Rodríguez, G. Rivas-Gervilla, and D. Sánchez, "Referring expression generation from images via deep learning object extraction and fuzzy graphs," in *Proceedings FUZZ-IEEE 2021*, 2021, pp. 1–6.

[2] Neo4j®, "Neo4j graph database," https://neo4j.com/product/neo4j-graph-database/.

[3] Neo4j, "Cypher query languaje," https://neo4j.com/docs/cypher-manual/current/.

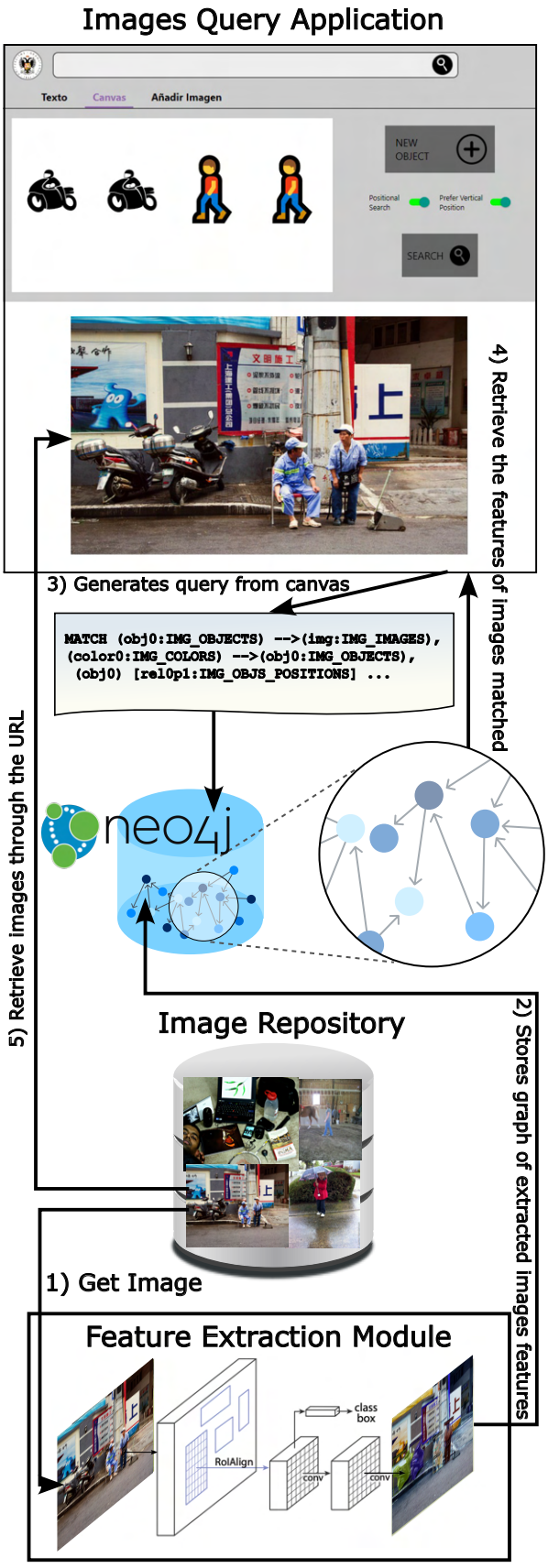


Figura 1. Arquitectura del sistema de caracterización y consulta de imágenes

Sistemas de recomendación evolutivos para objetos con estructuras complejas

Bartolome Ortiz-Viso
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain
bortiz@ugr.es

María J. Martín-Bautista
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain
mbautis@decsai.ugr.es

María-Amparo Vila
Department of Computer Science and Artificial Intelligence
University of Granada
Granada, Spain
vila@decsai.ugr.es

***Resumen*—Hoy en día, utilizamos ampliamente los sistemas de recomendación en una gran variedad de problemas. Estos sistemas de filtrado de información se han establecido como una de las principales herramientas que ayudan a los usuarios en situaciones cotidianas. Sin embargo, a medida que se desarrollan nuevas herramientas de procesamiento de la información, se plantean nuevos y desafiantes escenarios donde aplicar estos sistemas de recomendación en áreas como la salud, el ejercicio físico o el turismo.**

En estas situaciones estaremos frente a lo que se ha denominado como sistemas de recomendación complejos. En ellos deberemos balancear la información proveniente de de múltiples fuentes de información de ítems, usuarios o expertos, al mismo tiempo que diferenciamos entre restricciones o limitaciones y preferencias. Además, la recomendación final estará formada por un conjunto de elementos relacionados que deberá construirse.

En este resumen se presenta un modelo teórico de sistema de recomendación para este tipo de situaciones, basado en computación evolutiva y recomendación basada en contenido, publicado en [1].

***Palabras Clave*—Applied Computing, Complex Recommendation systems, Human-Computer Interaction, Information Retrieval, Recommendation systems.**

I. Introducción y trabajos relacionados

Los sistemas de recomendación constituyen hoy día una de las herramientas informáticas más presente en nuestras vidas. Nuestra forma de consumir productos de base tecnológica habitualmente requiere (seamos conscientes o no) de estos sistemas de filtrado de información. Podemos encontrar ejemplos de su uso en entretenimiento [4], comercio electrónico [5] o redes sociales [6].

Sin embargo, la mayoría de estas aplicaciones se encargan de recomendar objetos sencillos (libros, canciones, etc) frente a recomendaciones más complejas (dietas, rutas de viaje, fondos de inversión). Este es uno de los grandes retos actuales de la investigación en sistemas de recomendación [2], la recomendación de objetos complejos.

La complejidad de estas situaciones viene motivada por la naturaleza de las recomendaciones. Éstas a menudo serán un conjunto de objetos interrelacionados (a diferencia de uno solo), que optimizan un conjunto de objetivos (que pueden ser restrictivos o admitir un rango de valores) y que pueden tener o no una estructura asociada.

Como es natural, para que estos sistemas funcionen adecuadamente, necesitan de una gran cantidad de información adicional sobre los usuarios, los objetos que conforman la recomendación y sobre el contexto en el cual se desarrollan [7]. Toda esta información depende del área de recomendación. Así la recomendación de moda es habitual que incorpore información visual al proceso [8], la de turismo suele añadir información contextual [9], o la nutricional [10] información de la salud del usuario.

Muchos de estos sistemas, sin embargo, se enfocan más a la recomendación como una colección de objetos frente a una recomendación estructurada, donde la posición en la recomendación tiene un peso añadido.

Nuestro objetivo es, por tanto, ofrecer un modelo teórico que permita un acercamiento a la recomendación de objetos complejos estructurados. En este modelo utilizaremos los algoritmos evolutivos (usados con éxito en optimización en los sistemas de recomendación [3]) para la creación de objetos estructurados que verifican objetivos/restricciones mínimos. para posteriormente modificarlos en función de las preferencias de usuario [1].

II. Sistema de recomendación evolutivo para objetos estructurados

En esta sección presentaremos brevemente el sistema de recomendación propuesto en [1].

Nuestro sistema de recomendación tiene tres principales fuentes de información. Los dos primeros corresponden a elementos básicos de recomendación denotados por $\mathcal{I}_p$, el espacio de Items a recomendar y $\mathcal{U}_q$, el espacio de usuarios. Los elementos de cada uno de estos espacios pertenecen a espacio p,q-dimensional. Además, necesitamos una fuente de información adicional que contenga conocimiento experto, denotada $\mathcal{K}$, sobre nuestro problema.

Estos tres espacios, a su vez, codifican no sólo características de ítem o usuarios, si no también restricciones u objetivos que son necesarios alcanzar.

Los ítem de $\mathcal{I}_p$ no son los objetos finales a recomendar. Son los componentes del objeto que queremos recomendar.

Este trabajo fue financiado por la Unión Europea bajo el acuerdo de subvención No.816303 (Stance4Health)

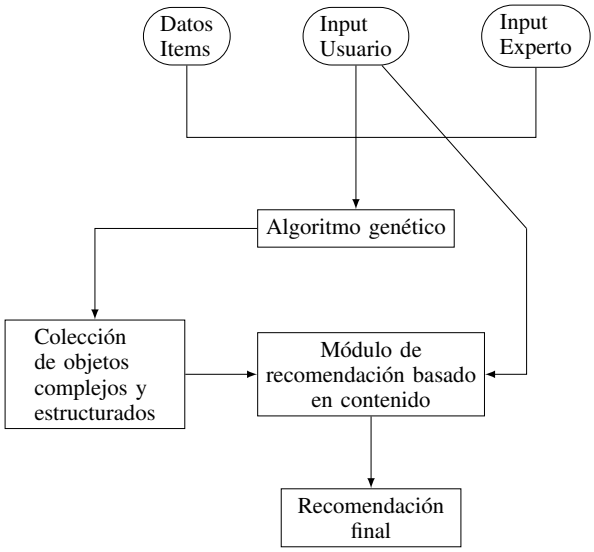


Figura 1. Esquema del modelo propuesto.

Es por ello que la restricción original más importante es la estructura de la recomendación final (incluida en $\mathcal{K}$), la cual se construirá con elementos más sencillos. Gracias a los algoritmos evolutivos obtendremos un conjunto de estas estructuras complejas sobre la cual actuar y recomendar. Nuestro primer objetivo es por tanto la generación de los ítems definitivos que formarán nuestra recomendación final : $\mathcal{L}(\mathcal{I},\mathcal{U},\mathcal{K})$. Esta forma representa conjuntos de elementos de $\mathcal{I}$, que dependen de las restricciones propuestas por el usuario, y de las características propuestas por otras fuentes de conocimiento.

Posteriormente, seleccionaremos, según su origen, aquellos parámetros que condicionan la estructura y sus componentes $\mathcal{E}_\mathcal{U},\mathcal{E}_\mathcal{K}$ y los parámetros que afectan al modelo generativo $\mathcal{V}_\mathcal{U},\mathcal{V}_\mathcal{K}$. Estos parámetros se utilizan para definir un problema de optimización basado en algoritmos genéticos de la forma:

$$\begin{cases} Min_{x\in\mathcal{H}}\|\mathcal{F}(x) = W_{\mathcal{V}_\mathcal{U}}f_\mathcal{U}(x) + W_{\mathcal{V}_\mathcal{K}}f_\mathcal{K}(x)\| \\ \text{donde}\mathcal{H} \subset \mathcal{L}(\mathcal{I},\mathcal{U},\mathcal{K}) \end{cases} \quad (1)$$

Con $f_{\mathcal{U}_K}$, $f_{\mathcal{V}_K}$ son funciones de ajuste a objetivos. Finalmente una vez obtenido un conjunto de objetos complejos que verifican las restricciones, pasamos a una segunda fase de modificación, donde ajustamos los objetos según preferencias del usuario. Cada elemento que compone nuestros ítem finales posee un conjunto de características, denotadas $i_1,\ldots,i_p$. Estas características pueden ir desde valores específicos valores específicos hasta etiquetas multivaluadas o booleanas (posee una cualidad o no posee una cualidad).

Este hecho nos permite transformar estos elementos en vectores de un espacio multidimensional sobre el que aplicar diferentes métricas. En nuestro caso, dado que este sistema está fundamentado principalmente características de los ítem, proponemos métricas como la similitud de coseno, para un enfoque basado en contenido.

III. Conclusiones y trabajo futuro.

Los sistemas de recomendación en espacios complejos suponen aún un escenario que plantea muchas vías de investigación. En particular, el articulo pone en relevancia que un acercamiento basado en algoritmos evolutivos y conocimiento especifico supone una vía atractiva para este tipo de problemas.

El enfoque mixto evolutivo-contenido, no solo puede ayudar a crear sistemas que sean capaces de combinar multitud de reglas y elementos para poder recomendar objetos complejos,si no que además pueden tener una estructura asociada. Este trabajo también influye en alguno de los problemas actuales de los sistemas de recomendación [2]. El primero es la necesidad de colecciones de objetos que recomendar. Estos dataset no son abundantes y constituyen una dificultad añadida para comparar sistemas.

El segundo más importante es la evaluación y adecuación de los sistemas en el tiempo. Nuestro trabajo parte de unas métricas fijas para ello. Sin embargo es necesario una mayor profundización en cómo evaluamos conjuntos de objetos, qué papel juega cada uno de ellos en la evaluación y cómo cambia esta última según interactúa el usuario en un determinado contexto. En este punto nace la necesidad de incorporar nuevas métricas, especialmente aquellas basadas en lógica difusa, que nos permitan acercarnos a una evaluación más acorde con las decisiones de un usuario.

Referencias

[1] Ortiz Viso, Bartolomé. " Evolutionary Approach in Recommendation Systems for Complex Structured Objects. " Fourteenth ACM Conference on Recommender Systems. 2020.

[2] Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Recommender Systems: Techniques, Applications, and Challenges." Recommender Systems Handbook. Springer, New York, NY, 2022. 1-35.

[3] Palomares, Iván, Hugo Alcaraz-Herrera, and Kao-Yi Shen. " F-EvoRecSys: An Extended Framework for Personalized Well-Being Recommendations Guided by Fuzzy Inference and Evolutionary Computing." International Journal of Fuzzy Systems (2022): 1-15.

[4] Schedl, Markus, et al. "Current challenges and visions in music recommender systems research." International Journal of Multimedia Information Retrieval 7.2 (2018): 95-116.

[5] Alamdari, Pegah Malekpour, et al. "A systematic study on the recommender systems in the E-commerce." IEEE Access 8 (2020): 115694-115716.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] Anelli, Vito Walter, et al. "Report on the 3rd workshop of knowledge-aware and conversational recommender systems (KARS/ComplexRec) at RecSys 2021." ACM SIGIR Forum. Vol. 55. No. 2. New York, NY, USA: ACM, 2022.

[8] Yin, Ruiping, et al. "Enhancing fashion recommendation with visual compatibility relationship." The World Wide Web Conference. 2019.

[9] Shen, Qijie, et al. "SAR-Net: A Scenario-Aware Ranking Network for Personalized Fair Recommendation in Hundreds of Travel Scenarios." Proceedings of the 30th ACM International Conference on Information and Knowledge Management. 2021.

[10] Chao, Wen-Yu, and Zachary Hass. "Choice-Based User Interface Design of a Smart Healthy Food Recommender System for Nudging Eating Behavior of Older Adult Patients with Newly Diagnosed Type II Diabetes." International Conference on Human-Computer Interaction. Springer, Cham, 2020.

Un estudio comparativo de series difusas para la evaluación de técnicas de deflexión de NEAs

1st J.M. Sánchez-Lozano
University Centre of Defence at the Spanish Air Force Academy, MDE-UPCT
30720 San Javier, Spain
juanmi.sanchez@cud.upct.es,
0000-0002-4615-8001

2nd M. Fernández-Martínez
University Centre of Defence at the Spanish Air Force Academy, MDE-UPCT
30720 San Javier, Spain
manuel.fernandez-martinez@cud.upct.es,
0000-0001-6291-9205

***Abstract*—Las extensiones de conjuntos difusos constituyen una de las principales áreas de investigación en el contexto de problemas relacionados con la inteligencia artificial. De hecho, su objetivo es abordar los problemas de toma de decisiones en el mundo real cuando la obtención de datos precisos no es una tarea sencilla. Es por ello por lo que recientemente se han introducido los conjuntos difusos esféricos como un paso más para modelar este tipo de problemas sobre la base del pensamiento humano. No obstante, dada su complejidad matemática, la necesidad de recurrir a números difusos esféricos en problemas de toma de decisiones que podrían ser abordados mediante series difusas tradicionales (triangulares) es una cuestión a analizar. En este estudio, se presenta un análisis comparativo realizado recientemente sobre ambos conjuntos difusos (series difusas esféricas vs. triangulares) en el que, a través de dos metodologías de toma de decisiones multi-criterio, AHP y TOPSIS, se evaluaron cuatro técnicas de desvío de asteroides (Kinetic Impactor, Ion-Beam Deflection, Enhanced Gravity Tractor y Laser Ablation). Finalmente se muestran los resultados de ambos estudios con el objetivo de comparar, por un lado, los órdenes de importancia de los criterios, y por otro, los rankings de alternativas obtenidos. Las conclusiones extraídas indican que la técnica Kinetic Impactor es la tecnología idónea, resaltando además que las series difusas esféricas presentan mayor sensibilidad ante variaciones en los pesos de los criterios.**

***Index Terms*—Números difusos triangulares, Números difusos esféricos, TOPSIS, AHP, Criterios, Asteroides**

I. Introducción

Las extensiones de conjuntos difusos a contextos más amplios constituyen una de las principales áreas de investigación en el contexto de problemas en inteligencia artificial. De esta manera, los conjuntos difusos esféricos han sido introducidos recientemente para modelar este tipo de problemas [1]. Alentados por ese tipo de conjuntos difusos, nos preguntamos en qué medida variaría una clasificación de alternativas, partiendo de metodologías de toma de decisiones multi-criterio, comparando series difusas triangulares (TFS) con las extensiones difusas más recientes tales como las series esféricas (SFS).

En este sentido, esta breve comunicación aborda esa pregunta al comparar las versiones con series difusas esféricas y triangulares de la metodología de toma de decisiones multi-criterio TOPSIS para clasificar un conjunto de alternativas involucradas en defensa planetaria. Específicamente, ilustramos esa comparación entre rankings por medio de un proceso de evaluación de 4 tecnologías de desvío de asteroides potencialmente peligrosos cercanos a la Tierra (Near Earth Asteroids, NEA) los cuales son objetos cuyas trayectorias orbitales pueden potencialmente cruzar la órbita de la Tierra. Las pequeñas perturbaciones que se pueden generar a consecuencia de este acercamiento pueden situarlos en una trayectoria de colisión con la de la Tierra. Por ello, realizar un seguimiento exhaustivo de estos asteroides es de gran interés por parte de la comunidad astronómica.

Partially funded by the research project PID2020-112754GBI00, financially supported by the Ministerio de Ciencia e Innovación (Spain).

II. Evaluación de técnicas de desvío de asteroides mediante series difusas

En este estudio, evaluamos un conjunto que consta de cuatro tecnologías de deflexión de NEAs: Kinetic Impactor (A_1), Ion-Beam Deflection (A_2), Enhanced Gravity Tractor (A_3) y Laser Ablation (A_4) con respecto a los siguientes criterios: tiempo de construcción (C_1), duración de la desviación activa (C_2), rotación del asteroide (C_3), composición del asteroide (C_4), estructura del asteroide (C_5), forma del asteroide (C_6), nivel de madurez de la tecnología (C_7) y riesgo de la misión (C_8).

En [2] y [3], se llevaron a cabo dos evaluaciones con respecto a dichas tecnologías de deflexión con respecto al conjunto de criterios antes mencionado por medio de las versiones TFN y SFN del enfoque AHP, respectivamente. A continuación, destacamos los principales resultados que proporcionó dicho análisis comparativo que involucró a dichos enfoques difusos. En primer lugar, en [2], el siguiente orden de preferencia para nuestro conjunto de criterios fue obtenido mediante la versión TFS de AHP:

$$C_1 > C_2 > C_7 > C_4 > C_5 > C_8 > C_6 > C_3.$$

Por otro lado, en [3] se proporcionó el siguiente orden de preferencia para ese conjunto de criterios tras de aplicar la versión SFS de AHP:

$$C_1 > C_2 > C_8 > C_4 > C_7 > C_5 > C_3 > C_6.$$

Así, los criterios C_7 (nivel de madurez de la tecnología) y C_4 (composición del asteroide) intercambiaron su orden de

importancia entre ambas versiones difusas de AHP. También es preciso resaltar que al criterio C_8 (riesgo de la misión) se le asignó un mayor nivel de importancia por parte de SFS AHP que por TFS AHP.

A continuación, comparamos las clasificaciones de las alternativas proporcionadas por las versiones SFS y TFS del procedimiento TOPSIS. TFS TOPSIS proporcionó el siguiente orden de preferencia para las cuatro tecnologías de deflexión:

$$A_1 > A_2 > A_3 > A_4,$$

mientras que SFS TOPSIS arrojó la siguiente clasificación para dichas alternativas:

$$A_1 > A_2 > A_4 > A_3.$$

Cabe señalar que la alternativa A_1 (Kinetic Impactor) aparece consolidada como la mejor opción para fines de deflexión de NEAs según los procedimientos SFS y TFS TOPSIS. Sin embargo, SFS TOPSIS intercambió las posiciones proporcionadas por TFS TOPSIS para las alternativas A_3 (Enhanced Gravity Tractor) y A_4 (Laser Ablation). Esto podría deberse a que SFS TOPSIS asignó un mayor nivel de importancia al criterio C_4 (composición del asteroide) que a C_7 (nivel de madurez de la tecnología).

REFERENCES

[1] F. K. Gündoğdu and C. Kahraman, "Spherical fuzzy sets and spherical fuzzy TOPSIS method," J Intell Fuzzy Syst, vol. 36, no. 1, pp. 337–352, 2019, doi: 10.3233/jifs-181401.

[2] J. M. Sánchez-Lozano, M. Fernández-Martínez, A. A. Saucedo-Fernández, and J. M. Trigo-Rodriguez, "Evaluation of NEA deflection techniques. A fuzzy Multi-Criteria Decision Making analysis for planetary defense," Acta Astronaut, vol. 176, pp. 383–397, 2020, doi: 10.1016/j.actaastro.2020.06.043.

[3] M. Fernández-Martínez and J. M. Sánchez-Lozano, "Assessment of Near-Earth Asteroid Deflection Techniques via Spherical Fuzzy Sets," Adv Astron, vol. 2021, pp. 1–12, 2021, doi: 10.1155/2021/6678056.

La modelización de la Polarización con Cadenas de Markov y su comparación con la medida de polarización difusa JDJ

1st Juan Antonio Guevara
Dpt. Estadística e Investigación Operativa
Universidad Complutense de Madrid
Madrid, España
juanguev@ucm.es

2nd Daniel Gómez
Dpt. Estadística e Investigación Operativa
Universidad Complutense de Madrid
Madrid, España
dagomez@estad.ucm.es

3rd Javier Castro
Dpt. Estadística e Investigación Operativa
Universidad Complutense de Madrid
Madrid, España
jcastroc@estad.ucm.es

4th Inmaculada Gutiérrez
Dpt. Estadística e Investigación Operativa
Universidad Complutense de Madrid
Madrid, España
inmaguti@ucm.es

5th José Manuel Robles
Dpt. Sociología Aplicada
Universidad Complutense de Madrid
Madrid, España
jmrobles@ucm.es

***Abstract*—En este estudio se aborda el problema del modelado de polarización con Cadenas de Markov (PMMC - *Polarization Modelling with Markov Chains*). Además, se compara esta probabilidad con la medida de polarización propuesta por Esteban y Ray y la medida de polarización difusa de Guevara et al. De esta manera, PMMC brinda la oportunidad de estudiar en profundidad cuál es el desempeño de estas medidas de polarización en condiciones específicas.**

***Index Terms*—Polarización, Cadenas de Markov, Medidas Difusas**

I. Introducción

En este trabajo se presenta un resumen del artículo titulado *A new approach to Polarization modeling using Markov Chains* [3] en el que se modeliza la polarización a través de cadenas de Markov y en el que se comparan las probabilidades de alcanzar una distribución polarizada con la medida de polarización borrosa JDJ propuesta en [2].

La medición de la polarización ha sido abordada desde diferentes perspectivas a lo largo de los años. Sin embargo, se debe dar un paso adelante siempre que en la literatura solo se hayan propuesto medidas para identificar los niveles de polarización en un instante de tiempo específico. En este sentido, las cadenas de markov pueden proporcionar los recursos adecuados para abordar esta tarea.

La medida de polarización difusa JDJ propuesta en [2] se explica a continuación. Sea X una variable unidimensional, y supongamos que X tiene dos posiciones extremas, siendo X_A y X_B μ_{X_A}, μ_{X_B} la funciones de grado de pertenencia, de modo que $\mu_{X_A}, \mu_{X_B} : N \longrightarrow [0,1]$ son funciones $\forall i \in N$. Se considera el riesgo de polarización entre dos individuos como la posibilidad de que estas dos situaciones:

1) Cómo de cerca i está del polo X_A y cómo de cerca j está de X_B.
2) Cómo de cerca i está del polo X_B y cómo de cerca j está de X_A.

Así, se tiene que:

$$JDJ(X) = \sum_{i,j \in N, i \leq j} \varphi\left(\phi(\mu_{X_A}(i), \mu_{X_B}(j)), (\phi(\mu_{X_B}(i), \mu_{X_A}(j))\right) \tag{1}$$

donde $\phi : [0,1]^2 \longrightarrow [0,1]$ es un operador de agregación de superposición y $\varphi : [0,1]^2 \longrightarrow [0,1]$ una función de agrupación.

II. Modelización de la Polarización con Cadenas de Markov

En esta sección se propone un modelo probabilístico novedoso para la Modelización de la Polarización con Cadenas de Markov (PMMC - *Polarization Modelling with Markov Chains* -).

A. Hipótesis.

1) ***Número de individuos para modelar***. Donde $N = \{1, , \ldots, k, \ldots, n\}$ es el conjunto completo.
2) ***Medición de actitud***. Sea $Z = \{1, ..., z\}$ una variable categórica para medir la actitud de N con z niveles. Se asume que Z tiene dos polos, 1 y z. Finalmente, sea Z_k la posición del individuo k en Z, siendo $k \in N$.
3) ***Independencia del comportamiento entre individuos***. Se propone especificar si $\forall k, l \in N$ con $k \neq l$, Z_k es independiente de Z_l o no.
4) ***Naturaleza de los polos***. Escenarios en los que. las personas están en posiciones extremas, el investigador tiene que decidir si la cadena termina - porque las actitudes extremas presumiblemente no cambiarán - o no.

5) ***Grado de inmovilidad***. Todo individuo presenta una probabilidad > 0 de permanecer en la misma posición.
6) ***Cambio de actitud***. Sea Z una variable categórica, Z_k el valor que el individuo k tiene en Z. Entonces, se tiene que $Z_k \pm d$ denota los posibles valores en Z que podría tomar el individuo k en la siguiente unidad de tiempo, siendo d las unidades a saltar en Z.
7) ***Simetría del cambio***. Indicar si las probabilidades de cambio son las mismas según: (a) *cambio de dirección* y (b) *cercanía a los polos*, es decir, si para un individuo $k \in N$, Z_k cambiará dependiendo de $\pm d$ y μ_{X_A}/μ_{X_B}.

B. Un nuevo modelo probabilístico

Estados en PMMC: Se denomina un estado en términos de PMMC a una distribución específica y estática de una población $N = \{1, , \ldots, k, \ldots, n\}$ a lo largo de la característica actitudinal Z donde $Z = \{1, \ldots, z\}$. Así, sea S el conjunto completo de estados con longitud $z^n = s$ donde $i \in S$ si y solo si $i = [Z_1, \ldots, Z_k, \ldots, Z_n]$. Asimismo, definimos como estados polarizados las distribuciones en las que una parte significativa de la población se sitúa en un polo de la variable y otra parte significativa de la población se sitúa en el otro polo.

Probabilidades de transición en PMMC: Dado $i, j \in S$ se define como P_{ij} a la probabilidad de llegar a una opinión poblacional j a partir de una opinión poblacional i. Además, i es un estado adyacente a j cuando presenta una probabilidad de transición > 0 de i a j. Los valores de P_{ij} dependen directamente de las reglas establecidas en las hipótesis propuesta anteriormente.

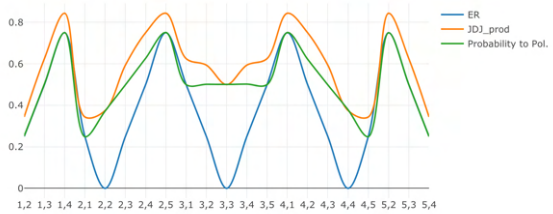
III. Experimento y resultados.

1) ***Número de individuos a modelar***. Se elige el caso más sencillo, siendo $n = 2$. Nótese que este ejemplo es equivale a una población divida en dos grupos de mismo tamaño con máxima homogeneidad intragrupal y máxima heterogeneidad intergrupal.
2) ***Medición de actitud***. Se considera una escala de categórica $Z = \{1, 2, 3, 4, 5\}$.
3) ***Independencia del comportamiento entre individuos***. Se asume que no hay independencia entre el comportamiento de los individuos.
4) ***Naturaleza de los polos***. Se asumen los valores de Z, 1 y 5 como posiciones absorbentes.
5) ***Grado de inmovilidad***. Todo individuo presenta una probabilidad de permanecer en su misma posición de $\alpha > 0$ donde $\alpha \in [0.1, 0.9]$ de 0.1 en 0.1.
6) ***Cambio de actitud***. Se asume que los individuos no cambian abruptamente sus actitudes. Un individuo k solo puede cambiar su opinión a valores adyacentes de la variable Z por unidad de tiempo, donde $d = 1$ siendo de Z_k a $Z_k \pm 1$ en el siguiente paso.
7) ***Simetría del cambio***. Se supone que las probabilidades de cambiar de posición son iguales en cualquier dirección y no se ven afectadas por la cercanía a los polos (μ_{X_A} o μ_{X_B}).

De acuerdo con lo anterior, se tiene que $S = \{[1, 1], [1, 2], [1, 3], ..., [3, 5], [4, 5], [5, 5]\}$ es el conjunto completo de estados, con longitud $s = z^n = 5^2$. Posteriormente, se calcula tantas matrices de transición $T = 25$ x 25 como diferentes valores de α donde $\alpha \in [0.1, 0.9]$. Se simulan las Cadenas de Markov para saber cuál es la probabilidad de que un estado no polarizado se polarice y cuántos pasos requiere el proceso para alcanzar estos estados específicos según cada valor de α. Primero, se consideran todos los estados a partir de S menos los absorbentes ($[1, 1]$, $[1, 5]$, $[5, 1]$ y $[5, 5]$) como estados iniciales. Después, se calcula la probabilidad frecuentista de que este estado alcance uno absorbente y se contabilizan cuántos pasos se requieren para que este estado los alcance. Para ello, se calculan 10000 iteraciones para cada estado inicial y valor de $\alpha \in [0.1, 0.9]$ por 0.1.

Así, se calculan los resultados medios de todas las 10000 y se observan las probabilidades y los pasos que podría tomar para alcanzar ambos estados polarizados $[1, 5]$ y $[5, 1]$.

Se calculan los valores de polarización de dos medidas de polarización, una nítida ER [1] y otra difusa JDJ donde $\phi =$ producto a lo largo de todos los estados transitorios del ejemplo propuesto. Además, en la figura mostrada a continuación se representan las probabilidades sumadas de alcanzar ambos polos para cada estado inicial. Se encuentra que ambas medidas siguen la misma tendencia que las probabilidades proporcionadas por PMMC. De hecho, las correlaciones entre las medidas de polarización y las probabilidades de PMMC son altas, siendo $r_{ER} = 0.759$ y $r_{JDJ} = 0.976$.



Los resultados han permitido conocer qué estados presentan un mayor riesgo de polarizarse en un futuro próximo. PMMC apoya la premisa de que las situaciones en las que toda la población mantiene la misma actitud no son iguales en términos de riesgo de polarización, mostrando similitud con la medida de polarización difusa JDJ [3]. Asimismo, ha encontrado un alto grado de correlación entre PMMC y las medidas de polarización, donde la medida difusa muestra mayor sintonía con el enfoque PMMC que la medida tradicional.

References

[1] Esteban, J.M.; Ray, D. On the measurement of polarization. *Econom. J. Econom. Soc.* **1994**, *62*, 819–851.
[2] Guevara, J.A.; Gómez, D.; Robles, J.M.; Montero, J. Measuring Polarization: A Fuzzy Set Theoretical Approach. In Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Lisbon, Portugal, 15–19 June 2020; Springer: Berlin, Germany, **2020**; pp. 510–522.

[3] Guevara, J.A.; Gómez, D.; Castro, J.; Gutiérrez, I.; Robles, J.M. A new approach to Polarization modeling using Markov Chains. In Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Milano, Italy, **In press**.

Análisis de conceptos formales bajo una visión intuicionista

Francisco Pérez-Gámez
Matemática Aplicada
Universidad de Málaga
Málaga, España
0000-0001-7518-1828

Pablo Cordero
Matemática Aplicada
Universidad de Málaga
Málaga, España
0000-0002-5506-6467

Manuel Enciso
Lenguajes y Ciencias de la Computación
Universidad de Málaga
Málaga, España
0000-0002-0531-4055

Ángel Mora
Matemática Aplicada
Universidad de Málaga
Málaga, España
0000-0003-4548-8030

Manuel Ojeda-Aciego
Matemática Aplicada
Universidad de Málaga
Málaga, España
0000-0002-6064-6984

***Index Terms*—Implicaciones, Información desconocida, Análisis de Conceptos Formales, lógica intuicionista.**

I. Introducción

Las lógicas trivaluadas se conocen como una extensión de la lógica clásica. En estas extensiones se añade un valor intermedio al conjunto de valores de verdad $\{F, T\}$. Este valor extra enriquece el poder de expresión de la lógica e induce una relación de orden considerando el valor intermedio localizado entre los otros dos valores.

Este es el caso de la lógica de Łukasiewicz [4] o la lógica de Kleene [3]. En cierto sentido, la lógica difusa también sigue esta idea, para ello, realiza una generalización donde introduce un conjunto de (infinitos) valores entre los dos valores de verdad. La principal diferencia entre las lógicas trivaluadas es el significado dado al tercer valor y a la negación.

Nosotros interpretaremos T como "tenemos información que prueba que es verdad" y F como " tenemos información que prueba que es mentira". Por último, el tercer valor de verdad puede ser interpretado como "desconocimiento", es decir, no tenemos ninguna información sobre si es verdad o falso. Por ejemplo, en un contexto con información de un hospital, la variable "estar embarazada" sigue esta interpretación ya que el desconocido podría ser visto como que la situación en la que no tenemos ninguna información.

Una revisión sobre el significado que puede tomar el desconocido se puede encontrar en [2]. En este artículo queremos extender la línea de investigación realizada en [5]. Para ello, presentamos una extensión graduada del marco algebraico que se encuentra en dicho artículo. Con esta extensión definimos lo que llamaremos un contexto formal parcial graduado. Finalmente, se presenta una conexión de Galois para capturar los conceptos de un contexto formal parcial graduado.

Apoyado por el ministerio español de ciencias,innovación y universidades (MCIU), Agencia Estatal de Investigación (AEI), Junta de Andalucía (JA), Universidad de Málaga (UMA) y Fondo Europeo de Desarrollo Regional (FEDER) a través de los proyectos PGC2018-095869-B-I00 , TIN2017-89023-P (MCIU/AEI/FEDER), y UMA2018-FEDERJA-001 (JA/UMA/FEDER).

II. Contextos formales parciales graduados

En esta sección vamos a extender el trabajo realizado en [5]. Por tanto, recomendamos su lectura puesto que los preliminares para este artículo se encuentran escritos en ese artículo. Consideramos el retículo residual $\mathbf{I} = ([0,1], \wedge, \vee, \otimes, \to, 0, 1)$.

Generalizamos el retículo $\underline{\mathbf{4}}$ considerando el retículo residual $\mathbf{I}^2 = \mathbf{I} \times \mathbf{I}$ donde el orden y las operaciones se extienden componente a componente.

Para un conjunto universal U los $\mathbf{4}$-conjuntos se extienden usando grados como $\mathbf{I}^2$-conjuntos, es decir, funciones de U a $\mathbf{I}^2$. De forma equivalente, un $\mathbf{I}^2$-conjunto $X\colon U \to \mathbf{I}^2$ puede ser como un ortopar paraconsistente graduado (X^+, X^-) tal que $X^+, X^-\colon U \to \mathbf{I}$ y $X(u) = (X^+(u), X^-(u))$ donde $X^+(u)$ significa *el grado en el que sabemos que u pertenece a X* y $X^-(u)$ significa *el grado en el que sabemos que u no pertenece a X*.

El conjunto de $\mathbf{I}^2$-conjuntos será denotado por $(\mathbf{I}^2)^U$. Las operaciones se pueden extender de $\mathbf{I}^2$ componente a componente: $(X \star Y)(u) = X(u) \star Y(u)$ para todo $u \in U$ y todo $\star \in \{\wedge, \vee, \otimes, \to\}$. También podemos considerar la relación de orden definida como $X_1 \sqsubseteq X_2$ si y solo si $X_1(u) \leq X_2(u)$ para todo $u \in U$, o, equivalentemente, si y solo si $X_1^+(u) \leq X_2^+(u)$ y $X_1^-(u) \leq X_2^-(u)$ para todo $u \in U$.

Se puede observar que $\underline{(\mathbf{I}^2)}^U = ((\mathbf{I}^2)^U, \wedge, \vee, \otimes, \to, \varepsilon, \dot{\imath})$ también es un retículo residual donde ε y $\dot{\imath}$ son los $\mathbf{I}^2$- conjuntos tales que $\varepsilon(u) = (0,0)$ y $\dot{\imath}(u) = (1,1)$ para todo $u \in U$. El $\mathbf{I}^2$-conjunto ε significa que no tenemos ninguna información sobre él, mientras que el conjunto $\dot{\imath}$ denota contradicción total, es decir, tener al mismo tiempo conocimiento de que todos los elementos pertenecen al conjunto y no pertenecen al conjunto.

Esto nos lleva a distinguir entre conjuntos consistentes e inconsistentes. Diremos que un $\mathbf{I}^2$-conjunto X en U es *consistente* si $X^+(u) \otimes X^-(u) = 0$ para todo $u \in U$, es decir, su imagen está contenida en $\mathbb{C} = \{(x_1, x_2) \in \mathbf{I}^2 \colon x_1 \otimes x_2 = 0\}$. Por otra parte, X es *inconsistente* si existe $u \in U$ tal que $X^+(u) \otimes X^-(u) > 0$.

El concepto de conjunto consistente es una generalización del conjunto difuso intuicionista [1] (que es bastante conocido), el cual, es una función $X \colon U \to [0,1]^2$ tal que, para todo $u \in U$, si $X(u) = (\alpha, \beta)$ entonces $\alpha + \beta \leq 1$. Es el caso particular en el que se considera el producto de Łukasiewicz ya que $0 = \alpha \otimes \beta = \text{máx}\{0, \alpha + \beta - 1\}$ es equivalente a $\alpha + \beta \leq 1$.

Es claro que $\mathbb{C}$ es un $\wedge$-subsemirretículo de $\underline{\mathbf{I}}^2$ que generaliza $\underline{\mathbf{3}}$, del mismo modo que $\underline{\mathbf{I}}^2$ generaliza $\underline{\mathbf{4}}$. El conjunto de conjuntos consistentes en U, denotado por $\mathbb{C}^U$, con su orden componente a componente es también un $\wedge$-semirretículo que será denotado por $\underline{\mathbb{C}}^U$.

Siguiendo el esquema realizado en el caso no graduado, consideraremos como equivalentes todos los elementos no consistentes y los identificaremos por $\dot{\imath}$. Para ello, usaremos el operador de clausura:

$$\mathcal{O} \colon (\mathbf{I}^2)^U \to (\mathbf{I}^2)^U \text{ donde } \mathcal{O}(X) = \begin{cases} X & \text{si } X \in \mathbb{C}^U, \\ \dot{\imath} & \text{en otro caso.} \end{cases}$$

Y denotaremos por $\dot{\mathbb{C}}^U$ su imagen, es decir, $\mathbb{C}^U \cup \{\dot{\imath}\}$. Como este conjunto es un sistema de clausura, sabemos que es un $\wedge$-subsemirretículo de $(\underline{\mathbf{I}}^2)^U$, y también que es un retículo completo que denotaremos por $\underline{\dot{\mathbb{C}}}^U$. Para distinguir el supremo de $(\underline{\mathbf{I}}^2)^U$ del de $\underline{\dot{\mathbb{C}}}^U$ usaremos los operadores $\vee$ y $\sqcup$ respectivamente. Es decir, dado $\{X_j \colon j \in J\} \subseteq \dot{\mathbb{C}}^U$ tenemos que $\bigsqcup_{j \in J} X_j = \mathcal{O}(\bigvee_{j \in J} X_j)$.

Ahora extendemos la definición de contexto formal parcial. Un contexto formal parcial graduado se define como una tupla $\mathbb{P} = (G, M, I)$ donde G y M son conjuntos cuyos elementos llamaremos *objetos* y *atributos* respectivamente, y I es un $\mathbb{C}$-conjunto en $G \times M$, es decir, $I \colon G \times M \to \mathbb{C}$. Por lo tanto, para cada $(g, m) \in G \times M$, el valor $I^+(g, m)$ es el grado en el que *sabemos que el objeto g tiene el atributo m* y el valor $I^-(g, m)$ es el grado en el que *sabemos que el objeto g no tiene el atributo m*. Dado un objeto y un atributo, podemos realizar operaciones que nos devuelva el conocimiento que tenemos sobre ese atributo o ese objeto: $I(g, .) \in \mathbb{C}^M$ y $I(., m) \in \mathbb{C}^G$ definidas del siguiente modo:

$$I(g, .)(x) = I(g, x) \text{ para todo } x \in M$$
$$I(., m)(x) = I(x, m) \text{ para todo } x \in G.$$

Para extender los resultados que tenemos sobre contextos formales parciales debemos considerar una nueva conexión de Galois.

Teorema 1: Consideremos un contexto formal parcial graduado $\mathbb{P} = (G, M, I)$, los operadores de derivación $(\ )^{\uparrow} \colon \mathbf{2}^G \to \dot{\mathbb{C}}^M$ y $(\ )^{\downarrow} \colon \dot{\mathbb{C}}^M \to \mathbf{2}^G$ definidos por

$$X^{\uparrow} = \bigwedge_{g \in X} I(g, \ )$$

$$Y^{\downarrow} = \{g \in G \colon Y \sqsubseteq I(g, \ )\}$$

forman una conexión de Galois entre los retículos $\underline{\mathbf{2}}^G$ y $\underline{\dot{\mathbb{C}}}^M$.

Con esta conexión de Galois podemos definir el retículo de conceptos de forma habitual. El elemento mínimo sería $(\varnothing, \dot{\imath})$ y para cualquier otro concepto (X, Y), si $Y(m) = (\alpha, \beta)$ el valor α es el grado en el que sabemos que todos los objetos en X tienen el atributo m, mientras que el valor β es el grado en el que sabemos que todos los objetos en X no tienen el atributo m (en términos matemáticos, α y β son, respectivamente, la cota inferior del grado en el que sabemos que cada objeto de X tiene o, respectivamente, no tiene el atributo m).

Para tener un marco de trabajo más general, podríamos considerar conjuntos difusos de objetos. Para ello, debemos introducir alguna notación adicional. De forma natural, de un contexto formal parcial graduado $\mathbb{P} = (G, M, I)$, podemos formar dos contextos formales difusos, $\mathbb{K}^+_{\mathbb{P}} = (G, M, I^+)$ y $\mathbb{K}^-_{\mathbb{P}} = (G, M, I^-)$. Los operadores de derivación en $\mathbb{K}^+_{\mathbb{P}}$ y $\mathbb{K}^-_{\mathbb{P}}$ serán denotados por $\oplus$ y $\ominus$ respectivamente. Por tanto, dado $X \in [0,1]^G$ los conjuntos difusos $X^{\oplus}, X^{\ominus} \in [0,1]^M$ se definen por

$$X^{\oplus}(m) = \bigwedge_{g \in G} \big(X(g) \to I^+(g, m)\big) \quad \text{y}$$
$$X^{\ominus}(m) = \bigwedge_{g \in G} \big(X(g) \to I^-(g, m)\big)$$

y, dado $Y \in [0,1]^M$, los conjuntos difusos $Y^{\oplus}, Y^{\ominus} \in [0,1]^G$ de forma análoga.

Podemos usar estos contextos formales para definir los operadores de derivación de un contexto parcial graduado como sigue:

Teorema 2: Dado un contexto parcial graduado $\mathbb{P} = (G, M, I)$, los operadores de derivación $(\ )^{\triangle} \colon [0,1]^G \to (\mathbf{I}^2)^M$ y $(\ )^{\triangledown} \colon (\mathbf{I}^2)^M \to [0,1]^G$ definidos por

$$X^{\triangle} = (X^{\oplus}, X^{\ominus}) \quad y \quad Y^{\triangledown} = (Y^+)^{\oplus} \cap (Y^-)^{\ominus}$$

forman una conexión de Galois entre los retículos $\underline{\mathbf{I}}^G$ y $\underline{\mathbf{I}^2}^M$. Este teorema, que es una extensión directa del teorema de conexión de Galois del artículo [5], define conceptos formales como pares de la conexión de Galois y estudia el correspondiente retículo de conceptos. Sin embargo, la principal diferencia entre este teorema y el anterior es que $X^{\triangle}$ podría ser inconsistente.

REFERENCIAS

[1] Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets and Systems **20**(1), 87–96 (1986)

[2] Ciucci, D., Dubois, D., Lawry, J.: Borderline vs. unknown: comparing three-valued representations of imperfect information. Intl J of Approximate Reasoning **55**(9), 1866–1889 (2014)

[3] Kleene, S.C.: Introduction to metamathematics. Van Nostrand (1952)

[4] Łukasiewicz, J.: O logice trojwartosciowej. Ruch Filozoficzny **5** (1920), English translation .ºn Three-Valued Logicïn Borkowski, L. (ed.) *Jan Łukasiewicz: Selected Works* (1970)

[5] Pérez-Gámez, F., Cordero, P., Enciso, M., Mora, A.: A new kind of implication to reason with unknown information. Lecture Notes in Computer Science **12733**, 74–90 (2021)

Implicaciones de atributos en el marco multiadjunto

M. Eugenia Cornejo
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
mariaeugenia.cornejo@uca.es

Jesús Medina
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
jesus.medina@uca.es

Francisco José Ocaña
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
franciscojose.ocana@uca.es

***Abstract*—Esta contribución continúa con el trabajo presentado en [4] y [8], aportando nuevas definiciones y propiedades relacionadas con las implicaciones de atributos en el marco de trabajo de los retículos de conceptos multiadjuntos, así como una aplicación al análisis forense digital.**

I. Introducción

El análisis de conceptos formales es una herramienta matemática, basada en la teoría de retículos, que nos permite extraer información significativa de bases de datos relacionales y visualizarla por medio de los conceptos y sus dependencias. Debido a que las bases de datos suelen contener una gran cantidad de información, el coste computacional asociado con el cálculo del retículo de conceptos es muy elevado, lo que dificulta considerar todo el contexto asociado a los datos y aplicar los procedimientos para formar la jerarquía conceptual de los mismos. Como una alternativa para obtener la información subyacente en el contexto, se presentan las implicaciones de atributos para proporcionar reglas que modelen el conjunto de datos. Las implicaciones de atributos han sido ampliamente estudiadas tanto en el caso clásico [3], [10] como en el caso difuso [1], [5]–[7], [11]. Sin embargo, aún no se han explorado plenamente las ventajas que proporciona el marco de trabajo de los retículos de conceptos multiadjuntos [9] en el tratamiento de las implicaciones de atributos. Este trabajo se dedica al estudio de la validez de las implicaciones de atributos en el entorno multiadjunto y su aplicación en el análisis forense digital, complementando los resultados presentados en [4] y [8].

II. Análisis de conceptos formales multiadjuntos

El análisis de conceptos formales permite calcular conceptos que describen la relación entre un conjunto de atributos y un conjunto de objetos, así como establecer una jerarquía entre ellos, obteniendo una estructura algebraica conocida como retículo de conceptos. Estas nociones han sido extendidas al marco multiadjunto [9] como se muestra a continuación.

Parcialmente financiado por los fondos FEDER 2014-2020 en colaboración con la Agencia Estatal de Investigación (AEI) a través del proyecto con referencia PID2019-108991GB-I00 y el proyecto financiado por la Junta de Andalucía y los fondos FEDER con referencia FEDER-UCA18-108612, y por la European Cooperation in Science & Technology (COST) Action CA17124.

Definición 1: Sean $(L_1, \preceq_1)$ y $(L_2, \preceq_2)$ dos retículos completos y $(P, \leq)$ un conjunto parcialmente ordenado.

- Un *marco multi-adjunto* es una tupla
$$(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$$
tal que $(\&_i, \swarrow^i, \nwarrow_i)$ es un triple adjunto con respecto a L_1, L_2, P, para todo $i \in \{1, \ldots, n\}$.
- Un *contexto* es una tupla (A, B, R, σ) tal que A y B son dos conjuntos no vacíos (interpretados como conjuntos de atributos y objetos, respectivamente), $R\colon A \times B \to P$ es una relación P-difusa y $\sigma\colon A \times B \to \{1, \ldots, n\}$ es una aplicación que asocia cualquier elemento en $A \times B$ con un determinado triple adjunto del marco.
- Un *concepto multiadjunto* es un par $\langle g, f\rangle$ satisfaciendo que $g \in L_2^B$, $f \in L_1^A$, $g^{\uparrow} = f$ y $f^{\downarrow} = g$, donde:
$$\begin{aligned} g^{\uparrow}(a) &= \inf\{R(a,b) \swarrow^{\sigma(a,b)} g(b) \mid b \in B\} \\ f^{\downarrow}(b) &= \inf\{R(a,b) \nwarrow_{\sigma(a,b)} f(a) \mid a \in A\} \end{aligned}$$
g y f se denominan extensión e intensión del concepto, respectivamente.
- El *retículo de conceptos multi-adjunto* asociado al marco y contexto considerado, viene dado por $(\mathcal{M}, \preceq)$ donde:
$$\mathcal{M} = \{\langle g, f\rangle \mid g \in L_2^B, f \in L_1^A, g^{\uparrow} = f, f^{\downarrow} = g\}$$
y cuyo orden viene definido por $\langle g_1, f_1\rangle \preceq \langle g_2, f_2\rangle$ si y solo si $g_1 \preceq_2 g_2$ (o equivalentemente, $f_2 \preceq_1 f_1$).

El conjunto de elementos supremo irreducibles del retículo de conceptos multiadjunto, denotado como $J_F(\mathcal{M})$, juega un papel fundamental en este trabajo, ya que permitirá simplificar el número de operaciones necesario para el cómputo de la validez de las implicaciones de atributos en el marco multiadjunto. En adelante, consideraremos un marco multiadjunto $(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$ y un intensificador [2] $*\colon L_2 \to L_2$, para definir el contexto (A, B^*, R, σ) asociado con los siguientes operadores de formación conceptos, donde $g \in L_2^B$, $f \in L_1^A$:

$$\begin{aligned} g^{\uparrow *}(a) &= \inf\{R(a,b) \swarrow^{\sigma(a,b)} g(b)^* \mid b \in B\} \\ f^{\downarrow}(b) &= \inf\{R(a,b) \nwarrow_{\sigma(a,b)} f(a) \mid a \in A\} \end{aligned}$$

En este caso, el retículo de conceptos multiadjunto $(\mathcal{M}_*, \preceq)$ se define de manera análoga a como se hizo la Definición 1.

III. Validez de las implicaciones de atributos en el marco multiadjunto

Antes de presentar la teoría desarrollada, conviene recordar la filosofía bajo la noción de implicación de atributos. Concretamente, una implicación de atributos indica que si un objeto satisface los atributos que aparecen en el antecedente de la regla entonces satisface los atributos dados en el consecuente. El significado e interpretación de estas implicaciones se ve reforzado por el cálculo de la validez de las mismas en el contexto deseado.

Definición 2: Dado un triple adjunto $(\&, \swarrow, \nwarrow)$ con respecto a L_1 y $f_1, f_2, f_3 \in L_1^A$, *el grado en el que una implicación* $f_2 \Leftarrow f_1$ *es válida en* f_3, se denota como $\|f_2 \Leftarrow f_1\|_{f_3}$, y se define como:

$$\|f_2 \Leftarrow f_1\|_{f_3} = S^1(f_2, f_3) \swarrow S^1(f_1, f_3)^*$$

donde $S^1(f_1, f_2) = \inf\{f_2(a) \nwarrow f_1(a) \mid a \in A\}$ y $*$ es un intensificador.

La noción de validez de una implicación difusa se extiende al marco de los retículos de conceptos multiadjuntos como se muestra a continuación.

Definición 3: Dados $f_1, f_2 \in L_1^A$, *el grado en el que una implicación* $f_2 \Leftarrow f_1$ *es válida en el retículo de conceptos multiadjuntos* $(\mathcal{M}, \preceq)$, se denota como $\|f_2 \Leftarrow f_1\|_{\mathcal{M}}$ y viene dado por:

$$\|f_2 \Leftarrow f_1\|_{\mathcal{M}} = \|f_2 \Leftarrow f_1\|_{Int(\mathcal{M}_*)}$$

donde $Int(\mathcal{M}_*)$ es el conjunto de intensiones de los conceptos del retículo $(\mathcal{M}_*, \preceq)$.

El siguiente teorema muestra que el grado en que una implicación difusa es válida en un retículo de conceptos multiadjunto se obtiene calculando solo el grado de validez en el conjunto de elementos supremo irreducibles del retículo. En consecuencia, no necesitamos calcular el grado de validez sobre todo el conjunto de intensiones de los conceptos del retículo, lo que implica una reducción importante en el número de cálculos a realizar.

Teorema 1: Sea $(\&, \swarrow, \nwarrow)$ un triple adjunto con respecto a L_1 satisfaciendo que $\top_1 \& y = y$ para cada $y \in L_1$. Dados $f_1, f_2 \in L_1^A$, entonces:

$$\|f_2 \Leftarrow f_1\|_{Int(\mathcal{M}_*)} = \|f_2 \Leftarrow f_1\|_{Int(J_F(\mathcal{M}_*))}$$

donde $Int(J_F(\mathcal{M}_*))$ denota al conjunto de intensiones de los elementos supremo irreducibles del retículo $(\mathcal{M}_*, \preceq)$.

IV. Aplicación al análisis forense digital

En esta sección se realiza un estudio inicial sobre la incidencia delictiva en la ciudad de Bilbao, España, durante los días que el club de fútbol local Athletic Club de Bilbao juega un partido en el estadio de San Mamés. Para ello, se ha considerado un conjunto de datos proporcionado por la Ertzaintza, policía autonómica del País Vasco, y se ha analizado la criminalidad en base a diferentes implicaciones de atributos, relacionadas con los delitos que aparecen en las jornadas de fútbol en la barrios cercanos al estadio. Para realizar el cálculo del grado de validez de las implicaciones de atributos se ha implementado un nuevo módulo en el software BintelMas. El contexto formal sobre el que se trabaja contiene como conjunto de atributos los tipos de delitos, como conjunto de objetos los días de partido y como intensificador la identidad. El triple adjunto usado para el cáculo de la validez es el de Lukasiewicz. De las implicaciones de atributos obtenidas, destacamos las siguientes:

Antecedente	{Hurto/0.07, Robo Violencia/0.17, Robo Intimidación/0.25 }
Consecuente	{Hurto/0.05, Robo Intimidación/0.01 }
Validez	1.0

Antecedente	{Estafa/0.6, Hurto/0.1, Amenazas/1.0, Robo Violencia/0.17 }
Consecuente	{Estafa/1.0, Hurto/0.05}
Validez	0.05

La primera implicación al presentar una grado de validez 1 nos permite asegurar con total certeza que cada vez que hay algún robo con intimidación, también habrá robo con violencia; mientras que la segunda al tener un grado de validez próximo a 0 hace que la descartemos a la hora de extraer información relevante. Nuestro objetivo es seguir analizando más implicaciones asociadas a este contexto, así como avanzar en el estudio teórico y aplicado de esta línea de investigación, con el fin de implementar un nuevo módulo que nos permita calcular bases de implicaciones.

References

[1] R. Bělohlávek, P. Cordero, M. Enciso, A. Mora, and V. Vychodil. Automated prover for attribute dependencies in data with grades. *International Journal of Approximate Reasoning*, 70:51 – 67, 2016.

[2] R. Bělohlávek and V. Vychodil. Formal concept analysis and linguistic hedges. *Int. J. General Systems*, 41(5):503–532, 2012.

[3] P. Cordero, M. Enciso, A. Mora, and M. Ojeda-Aciego. Computing left-minimal direct basis of implications. In M. Ojeda-Aciego and J. Outrata, editors, *CLA*, volume 1062 of *CEUR Workshop Proceedings*, pages 293–298. CEUR-WS.org, 2013.

[4] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Computing the validity of attribute implications in multi-adjoint concept lattices. In *International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE 2016*, volume II, pages 414–423, 2016.

[5] C. V. Glodeanu. Knowledge discovery in data sets with graded attributes. *International Journal of General Systems*, 45(2):232–249, 2016.

[6] J. Hill, H. Walkington, and D. France. Graduate attributes: implications for higher education practice and policy. *Journal of Geography in Higher Education*, 40(2):155–163, 2016.

[7] T. Kuhr and V. Vychodil. Fuzzy logic programming reduced to reasoning with attribute implications. *Fuzzy Sets and Systems*, 262:1 – 20, 2015. Theme: Logic and Computer Science.

[8] V. Liñeiro-Barea, J. Medina, and I. Medina-Bulo. Towards generating fuzzy rules via fuzzy formal concept analysis. In J. Kacprzyk, L. Koczy, and J. Medina, editors, *7th European Symposium on Computational Intelligence and Mathematices (ESCIM 2015)*, pages 60–65, 2015.

[9] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets and Systems*, 160(2):130–144, 2009.

[10] J. Rodríguez-Jiménez, P. Cordero, M. Enciso, and A. Mora. Negative attributes and implications in formal concept analysis. *Procedia Computer Science*, 31:758 – 765, 2014.

[11] V. Vychodil. Computing sets of graded attribute implications with witnessed non-redundancy. *Information Sciences*, 351:90 – 100, 2016.

Nonlinear Preferences in Consensus Reaching Processes. Extreme Values Amplifications

1st Diego García-Zamora
Departamento de Informática
Universidad de Jaén
Jaén, España
dgzamora@ujaen.es

2nd Álvaro Labella
Departamento de Informática
Universidad de Jaén
Jaén, España
alabella@ujaen.es

3rd Rosa M. Rodríguez
Departamento de Informática
Universidad de Jaén
Jaén, España
rmrodrig@ujaen.es

4st Luis Martínez
Departamento de Informática
Universidad de Jaén
Jaén, España
martin@ujaen.es

***Abstract*—This keywork summarizes the main contents of *Nonlinear preferences in group decision-making. Extreme Values Amplifications and Extreme Values Reductions* [1]. Consensus Reaching Processes (CRPs) play a key role to soften the discrepancies among the experts who take part in the resolution of a Group decision-making (GDM) problem. Even though, traditionally, such experts' opinions have been modeled by assuming that they follow a linear scale, several psychological studies show that better decisions are obtained if their opinions are remapped according to a nonlinear scale. Therefore, this contribution aims at extending such nonlinear preference modeling to CRPs. Here, the notions of Extreme Values Amplifications (EVAs) are introduced as those automorphisms in the interval $[0,1]$ which amplify the distances between the extreme values. Afterwards, their main properties are analyzed and, finally, they are applied to a classic CRP to show how these nonlinear scales impact the decision process.**

***Index Terms*—Non Linear Preferences, Group Decision Making, Consensus Reaching Process, Extreme Values Amplification**

I. Preliminaries

This section gives some basics about GDM and its CRPs.

A GDM problem is a decision situation in which two or more individuals have to choose a collective solution for a certain problem. Formally, the main elements in a GDM problem are:

- A set $X = \{X_1, X_2, ..., X_n\}$, $2 \leq n \in \mathbb{N}$, of alternatives,
- A set $E = \{e_1, e_2, ..., e_m\}$, where $2 \leq m \in \mathbb{N}$, of experts.

Here, we consider that experts provide their opinions by using Fuzzy Preference Relations (FPRs), i.e. matrices $P_1, P_2, ..., P_m \in \mathcal{M}_{n\times n}([0,1])$ such that $p_k^{ij} + p_k^{ji} = 1$ for all $i, j = 1, 2, ..., n$ and $k = 1, 2, ..., m$, where $p_k^{ij} \in [0,1]$ represents the degree of preference of the alternative X_i before the alternative X_j for the expert e_k.

A CRP is an iterative discussion process, usually coordinated by a moderator, which aims at dealing with the conflicts among experts' opinions which appear in a GDM situation [1], [2] (see Fig. 1).

The discussion process stops when a maximum number of rounds $MaxRounds \in \mathbb{N}$ is surpassed or a consensus degree $\mu \in [0,1]$ is reached.

II. Extreme Values Amplifications

In this section, we introduce the main novelty of the contribution, the notion of EVA, which provides nonlinear scales that amplify the distances between the more extreme values of the preferences given by the experts who take part in a GDM problem.

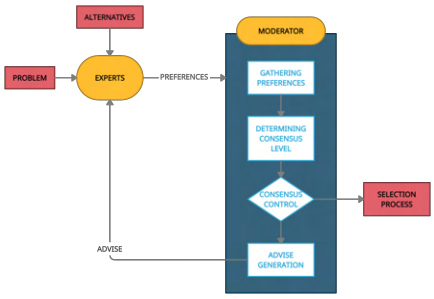


Fig. 1. General scheme for a CRP

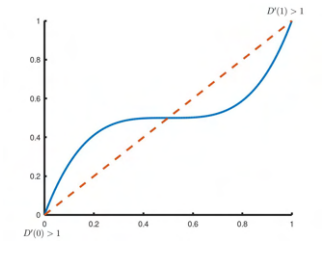


Fig. 2. Sketch of an EVA.

Definition 1 (Extreme Values Amplification [1]): Let $D : [0,1] \to [0,1]$ be a function satisfying:

1) D is an automorphism on the interval $[0,1]$,
2) D is a $\mathcal{C}^1$ function,
3) D satisfies $D(x) = 1 - D(1-x)\ \forall\ x \in [0,1]$,
4) $D'(0) > 1$ and $D'(1) > 1$,
5) D is twice differentiable near the extremes, concave close to 0, and convex close to 1.

D will be called then an Extreme Values Amplification (or EVA) on the interval $[0,1]$.

The functions satisfying the previous definition show the properties listed below.

1) They remap fuzzy preference relations onto fuzzy preference relations,

2) They amplify the distance between the extreme values, and reduce the distance between the intermediate ones,
3) They have a concrete geometrical pattern (see Fig. 2),
4) The amplification of distances is greater close to the extremes.

Some concrete examples of EVAs are:

Example 1: The family $s_\alpha : [0,1] \to [0,1]$

$$s_\alpha(x) := x - \alpha \cdot \sin(2\pi x - \pi) \ \forall \ x \in [0,1],$$

where $\alpha \in]0, 1/2\pi]$

Example 2: The family $m_\alpha : [0,1] \to [0,1]$

$$m_\alpha(x) := \begin{cases} \frac{1}{2} - \frac{1}{2}(1-2x)^\alpha & 0 \leq x < \frac{1}{2} \\ \frac{1}{2} + \frac{1}{2}(2x-1)^\alpha & \frac{1}{2} \leq x \leq 1 \end{cases},$$

where $\alpha > 1$.

III. Case study

Masthoff [3] and Delic [4] showed that, from a psychological point of view, nonlinear scales are more suitable for modeling experts' preferences. Here, we use the nonlinear scales provided by EVAs in the resolution of a CRP and provide a brief comparative analysis.

The supermarket franchise "EasyMarket" wants to select the most sustainable heating system for its 100 supermarkets. To do so, the company asks the maintenance managers $E = \{e_1, e_2, ..., e_{100}\}$ what kind of energy they prefer within X_1= *Natural gas*, X_2= *Geothermal*, X_3= *Aerothermal*, X_4= *Biomass*. To solve this decision situation, it has been used the consensus model proposed by Quesada et al. [2] ($\mu = 0.85$, $MaxRounds = 15$).

EVA	*Order of alternatives*	*Rounds*	*Consensus*
Classical	$X_4 \succ X_1 \sim X_2 \sim X_3$	10	0.85
$s_{0.08}$	$X_4 \succ X_1 \succ X_2 \sim X_3$	7	0.86
$s_{0.09}$	$X_4 \succ X_1 \succ X_2 \sim X_3$	7	0.87
m_2	$X_4 \succ X_1 \succ X_2 \sim X_3$	7	0.86
$m_{3.39}$	$X_4 \succ X_1 \succ X_2 \sim X_3$	5	0.87

TABLE I
Simulation results

Several simulations which use different EVAs have been carried out (see Table I and Fig. 3).

Fig. 3. MDS visualization of the simulations under different EVAs.

IV. Conclusions

When using the nonlinear scales provided by EVAs, the performance of the consensus model is improved. Whereas the classical version, which does not use EVAs, requires 10 discussion rounds to achieve a consensus degree of 0.85, when using EVAs the number of necessary rounds to reach a consensual solution is decreased and the final consensus degree is increased.

Acknowledgements

This work is partially supported by the Spanish Ministry of Economy and Competitiveness through the Spanish National Project PGC2018-099402-B-I00, and the Postdoctoral fellow Ramón y Cajal (RYC-2017-21978), the FEDER-UJA project 1380637 and ERDF, by the Spanish Ministry of Science, Innovation and Universities through a Formación de Profesorado Universitario (FPU2019/01203) grant and by the Junta de Andalucía, Andalusian Plan for Research, Development, and Innovation (POSTDOC 21-00461).

References

[1] D. García-Zamora, Á. Labella, R. M. Rodríguez, and L. Martínez, "Nonlinear preferences in group decision-making. Extreme Values Amplifications and Extreme Values Reductions," *International Journal of Intelligent Systems*, vol. 36, no. 11, pp. 6581–6612, Jul. 2021.
[2] F. J. Quesada, I. Palomares, and L. Martínez, "Managing experts behavior in large-scale consensus reaching processes with uninorm aggregation operators," *Applied Soft Computing*, vol. 35, pp. 873–887, 2015.
[3] J. Masthoff, "Group modeling: Selecting a sequence of television items to suit a group of viewers," *User Model. User-Adapt. Interact.*, vol. 14, pp. 37–85, 02 2004.
[4] A. Delic, F. Ricci, and J. Neidhardt, "Preference networks and non-linear preferences in group recommendations," in *IEEE/WIC/ACM International Conference on Web Intelligence*, 10 2019, pp. 403–407.

El uso de Metaheurísticas como Generadoras de Soluciones

Hanane El Raoui
Dept. of Management Science
Strathclyde Business School
Glasgow, Scotland
0000-0002-9079-3248

Marcelino Cabrera
Dept. Lenguajes y Sistemas Informáticos
Universidad de Granada
Granada, España
0000-0001-6090-5942

David A. Pelta
Dept. Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
0000-0002-7653-1452

Palabras clave: Metaheurísticas, Optimización, Diseño de Viajes Turísticos

En la mayoría de los problemas de toma de decisiones y/o de optimización del "mundo real", pueden existir una gran cantidad de criterios de decisión y objetivos que sean difíciles de incluir en la formulación de su modelo computacional. Aún en el caso que fuera posible, dicha inclusión puede aumentar la complejidad de dichos modelos hasta un punto en el que no puedan resolverse. Es necesario recordar que las decisiones finales (es decir, la selección de una solución) se toman a menudo basándose no sólo en los objetivos modelados, sino también en otros objetivos, sesgos y preferencias potencialmente subjetivos de los responsables de la toma de decisiones [1]. Otros algoritmos de optimización están orientados a encontrar soluciones a problemas con un único objetivo, o en el mejor caso, a encontrar conjuntos de soluciones no inferiores a los algoritmos multiobjetivo [2].

Si consideramos que existen parámetros y objetivos que son imposibles o al menos muy difíciles de modelar, se requiere el uso de metodologías de solución "no convencionales", no solo para optimizar los objetivos planteados, sino que pueden ser relevantes para explorar el espacio de las alternativas. Esto es así puesto que pueden existir soluciones sub-óptimas (en términos del modelo) pero perfectamente útiles desde otras perspectivas.

En este sentido, el rol de las metaheurísticas como "generadoras de soluciones" adquiere más importancia en dos sentidos: primero, porque sucesivas ejecuciones (o una ejecución única en el caso de técnicas basadas en una población) permiten la generación de un conjunto de soluciones potencialmente buenas, y segundo, si se dispone de solución de referencia, se puede definir un nuevo problema de optimización para generar soluciones con una calidad similar pero con una estructura completamente diferente.

El objetivo de esta contribución es mostrar como las metaheurísticas se pueden usar para generar soluciones, tomando como ejemplo una versión dinámica del problema de planificación de rutas turísticas. En la versión estática del problema tenemos un conjunto de puntos de interés (POI) en el que cada uno tiene un tiempo de visita y un nivel de interés, y el objetivo es encontrar un subconjunto de puntos que maximicen el nivel de interés en un tiempo determinado de visita. En la versión dinámica, el nivel de interés depende del momento en el que realizamos la visita al POI.

Este problema tiene dos elementos clave: la selección de los puntos de interés que son adecuados a los gustos de turista y en el diseño de itinerarios que permitan visitar dichos puntos o al menos un subconjunto de ellos. Mientras que el primer elemento está relacionado con los sistemas de recomendación el segundo de ellos se ajusta más a los problemas de optimización y decisión.

Para resolver el problema de diseño de rutas turísticas dependientes del tiempo (TTDP-TDS), comenzamos con un conjunto de N nodos o puntos de interés (POI). Cada nodo $i \in N$ tiene asociado una puntuación o un interés S_i y diversos factores de recomendación f_{it} que pondera el interés basándose en el momento en el que se realiza la visita al POI. También debemos tener en consideración el tiempo que utilizamos en desplazarnos de un nodo i a otro j (t_{ij}). Sin embargo, debido a que hay un tiempo máximo permitido, no todos los nodos del conjunto N pueden ser visitados.

La función objetivo maximiza la suma total de la puntuación obtenida.

$$Maximizar \quad \sum_{i \in N \setminus \{n_{ini}\}} \sum_{t \in T} S_i \cdot f_{it} \cdot y_{it}$$

La Figura 1 muestra un ejemplo de solución y cómo se realiza el cálculo de la función objetivo.

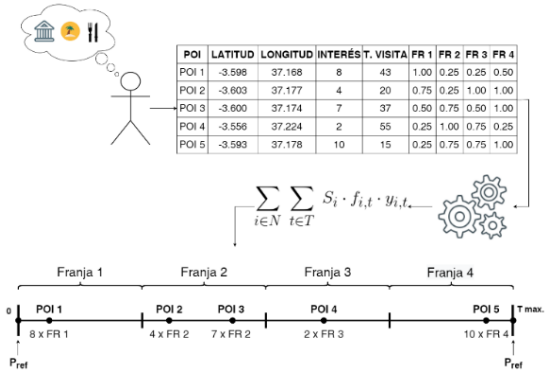


Fig. 1. Rutas de puntos con interés dependiente del tiempo.

Como ejemplo de metaheurística se utiliza un algoritmo evolutivo sin operador de cruce y un procedimiento GRASP para la generación de la población inicial. Fijando de forma aleatoria el POI inicial, creamos una lista de POI potenciales a visitar. De esta lista se selecciona un elemento aleatorio y se incluye en la solución. De esta forma repetimos el proceso mientras no se exceda el tiempo máximo de la ruta.

En cada generación, K padres se seleccionan de la población actual mediante una selección por ruleta y se incorporan a la población intermedia. Mediante los operadores de mutación disponibles se genera una solución hija de cada padre, y si es factible y mejor que la propuesta por el padre se incorpora a la población intermedia. Finalmente ordenamos la

población intermedia y escogemos los mejores elementos para sustituir a la población inicial.

Disponemos de cinco operadores de mutación: desplazamiento, intercambio, inversión, inserción y macro. Este último consiste en la selección de uno de los otros operadores dependiendo de una probabilidad proporcional a su éxito.

Con este algoritmo obtenemos un conjunto de 40 soluciones a las cuales se les calculan nuevos indicadores: *I* el interés de la ruta, *#POIs* el número de POIs visitados, *E* la eficiencia de la ruta, y T_{travel} el tiempo total de desplazamiento entre POIs. En la tabla siguiente se algunos valores de referencia que permiten observar la variabilidad de los indicadores propuestos.

	Min	Max	Mean	Std. Dev.
Nro. POIs	11	15	13.33	1.57
Interés	74.00	90.00	83.01	4.57
T. viaje	24.78	54.78	37.19	6.36
Eficiencia	84.77	92.60	89.18	1.60

Si tenemos en cuenta las preferencias de los usuarios para la evaluación de las soluciones podemos definir perfiles de visitante que nos permitan explorar las distintas soluciones con un punto de vista distinto en cada caso. Para explorar el problema definimos tres perfiles de usuarios:

1. Visita guiada: el turista quiere visitar el máximo de POIs de la ciudad maximizando el tiempo disponible, para lo que priorizamos las soluciones por el interés de la solución. $I \succeq_p \#POIs \succeq_p E \succeq_p T_{travel}$
2. Visita casual: un visitante que acude a la ciudad por otro motivo, pero que desea ver los sitios más interesantes de la ciudad, pero con el máximo tiempo libre durante la visita. En este caso maximizamos la efectividad de la visita. $E \succeq_p I \succeq_p \#POIs \succeq_p T_{travel}$
3. Visita de grupos grandes o con movilidad reducida: debido al tamaño del grupo o a los problemas de movilidad se considera una visita que minimice los tiempos de desplazamiento entre POIs (T_{travel}) $T_{travel} \succeq_p I \succeq_p \#POIs \succeq_p E$

Utilizando una metodología basada en intervalos y una distribución de posibilidad, se evalúan las 40 soluciones en términos de cada perfil, y para cada uno de ellos se determinan las 10 mejores rutas.

La Figura 2 muestra las 10 mejores soluciones para cada perfil, mostrando en la primera columna la solución de referencia y ordenando las columnas según la importancia del criterio. Las celdas en rojo muestran valores peores que los valores de referencia y en verde los que mejoran la solución. Utilizamos el coeficiente de similaridad de Jaccard para comparar dos soluciones $J(S_a, S_b)$.

Para el primer perfil, la solución de referencia $S^* = S_9$. Si nos fijamos en la segunda solución, S_4 solo tiene un valor de interés ligeramente inferior mientras que el resto de los factores son iguales. Sin embargo, S_4 es significativamente diferente de S^*, teniendo un valor de $J(S^*, S_4) = 0.58$. Las soluciones 5, 12, y 14 tienen el mismo valor de interés, pero son ligeramente peores en el resto de los factores.

Ref. Solution	Interest 90,00	POIs 15	Efficiency 90,57%	Travel time 33,83
9	0,00	0	0,00%	0,00
4	-1,50	0	0,00%	0,00
14	0,00	-1	-0,41%	1,09
12	0,00	-1	-1,24%	4,39
1	-3,50	0	-0,69%	1,64
5	0,00	0	-2,96%	10,70
16	-4,00	0	-0,82%	2,69
6	-3,50	0	-1,46%	5,25
17	-4,00	0	-1,34%	4,77
8	-3,50	0	-1,77%	5,27

(a) Profile 1

Ref. Solution	Efficiency 90,57%	Interest 90,00	POIs 15	Travel time 33,83
9	0,00%	0,00	0	0,00
4	0,00%	-1,50	0	0,00
37	2,03%	-10,00	-3	-9,05
1	-0,69%	-3,50	0	1,64
39	0,90%	-6,00	-3	-5,41
14	-0,41%	0,00	-1	1,09
11	-0,30%	-7,50	0	0,64
16	-0,82%	-4,00	0	2,69
6	-1,46%	-3,50	0	5,25
17	-1,34%	-4,00	0	4,77

(b) Profile 2

Ref. Solution	Travel time 33,83	Interest 90,00	POIs 15	Efficiency 90,57%
37	-9,05	-10,00	-3	2,03%
39	-5,41	-6,00	-3	0,90%
27	-4,36	-8,00	-3	0,35%
9	0,00	0,00	0	0,00%
4	0,00	-1,50	0	0,00%
26	-5,33	-12,00	-4	0,62%
11	0,64	-7,50	0	-0,30%
14	1,09	0,00	-1	-0,41%
1	1,64	-3,50	0	-0,69%
30	-3,13	-8,00	-4	0,00%

(c) Profile 3

Fig. 2. 10 mejores soluciones para los perfiles de usuario definidos.

El segundo perfil sigue primera la solución de referencia $S^* = S_9$, seguida de la S_4. Y como tercera solución aparece S_{37} que no está entre las 10 mejores del perfil 1. Esta solución mejora la eficiencia y el tiempo de desplazamiento, pero empeora su ranking debido al interés que es 10 unidades menor. La solución S_{39} también es interesante ya que mejora los mismos valores que S_{37}, pero sólo tiene un valor de interés 6 unides menor. Ambas soluciones solo visitan 12 POIs.

Finalmente, en el tercer perfil vemos que la solución de referencia ocupa la cuarta posición, siendo en este caso las soluciones S_{37}, S_{39}, y S_{27} las tres primeras. Todas mejoran el tiempo de desplazamiento y la eficiencia, sin llegar al 88% del interés de referencia (90), aunque visitan menos POIs (12).

Estos resultados muestran claramente la necesidad de tener un conjunto de soluciones en vez de concentrar nuestros esfuerzos en la obtención de una única solución óptima.

Como trabajo futuro, queremos explorar el enfoque "Modelado para Generar Alternativas" (MGA) para producir un conjunto de soluciones alternativas a nuestro problema, que ha sido utilizado con buenos resultados en [3,4] para generar diversos conjuntos de soluciones para el problema de la distribución de comida perecedera, conectando también el problema con el análisis de los conjuntos de soluciones con enfoques de toma de decisiones multicriterio.

AGRADECIMIENTOS

Los autores agradecen el apoyo de los proyectos PID2020-112754GB-I0, MCIN/AEI /10.13039/501100011033 y FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto (BTIC-640-UGR20).

REFERENCES

[1] G Gunalay, Y.; Yeomans, J.S. Simulation-optimization techniques for modelling to generate alternatives in waste management planning. J. Appl. Oper. Res. 2011, 3, 23–35.

[2] Janssen, J.; Krol, M.; Schielen, R.; Hoekstra, A. The effect of modelling quantified expert knowledge and uncertainty information on model-based decision making. Environ. Sci. Policy 2010, 13, 229–238.

[3] El Raoui, Hanane, Marcelino Cabrera-Cuevas, and David A. Pelta. "The Role of Metaheuristics as Solutions Generators." Symmetry 13.11 (2021): 2034.

[4] El Raoui, Hanane, et al. "On the Generation of Alternative Solutions for a Perishable Food Distribution Problem." Computational Intelligence Methodologies Applied to Sustainable Development Goals. Springer, Cham, 2022. 255-27

Estimación de Información Incompleta en Toma de Decisión en Grupo Mediante Computación Granular

Francisco Javier Cabrerizo
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
cabrerizo@decsai.ugr.es

Ignacio Javier Pérez
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
ijperez@decsai.ugr.es

Manuel Jesús Cobo
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
mjcobo@decsai.ugr.es

Sergio Alonso
dept. Lenguajes y Sistemas Informáticos
Universidad de Granada
Granada, España
zerjioi@ugr.es

María Ángeles Martínez
dept. Trabajo Social y Servicios Sociales
Universidad de Granada
Granada, España
mundodesilencio@ugr.es

Enrique Herrera-Viedma
dept. Ciencias de la Computación e IA
Universidad de Granada
Granada, España
viedma@decsai.ugr.es

***Resumen*—En este trabajo, se introduce el uso del concepto de granularidad de la información para estimar, con la máxima consistencia posible, los valores no proporcionados en una relación de preferencia difusa. Se emplea el concepto de relación de preferencia granular para representar cada valor no proporcionado mediante un gránulo de información (intervalo, en este caso) en lugar de como un valor numérico preciso. Esto proporciona la flexibilidad necesaria para estimar los valores faltantes de forma que la consistencia de la relación de preferencia difusa completa sea la máxima posible.**

***Palabras clave*—granularidad de la información, relación de preferencia difusa, consistencia, información incompleta**

I. Propuesta

La toma de decisión en grupo se define como la situación en la que una serie de personas, $P = \{p_1, p_2, \ldots, p_m\}$ $(m \geq 2)$, tienen que elegir de forma colectiva la mejor alternativa de entre un conjunto, $A = \{a_1, a_2, \ldots, a_n\}$ $(n \geq 2)$, de estas. Usualmente, se ha asumido que los participantes tienen el conocimiento necesario para hacer una distinción precisa del grado de preferencia de una alternativa sobre otra. Sin embargo, esta asunción puede ser irreal, sobre todo en aquellos procesos de decisión que involucran a muchas personas, las cuales tienen que considerar fuentes de información conflictivas a la hora de elegir entre diferentes alternativas. Por ejemplo, se ha comprobado que «el aumento de la intensidad del conflicto en una comparación multicriterio aumenta la probabilidad de que los responsables de la toma de decisión consideren dos alternativas como incomparables» [2]. Esto puede provocar que la información sea incompleta.

Este trabajo ha sido cofinanciado por el Programa Operativo FEDER 2014–2020 y por la Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía mediante el proyecto B-TIC-590-UGR20.

En el caso de las relaciones de preferencia difusas, se han desarrollado gran cantidad de métodos para estimar los valores faltantes. La mayoría de estos métodos se basan en criterios de consistencia y en los grados de preferencia proporcionados por la propia persona para evitar inconsistencias [6]. Estos métodos son los que mejores resultados obtienen.

Una nueva dirección de investigación, prometedora e innovadora, consiste en la construcción y conceptualización de modelos granulares, los cuales se pueden considerar como generalizaciones de los modelos numéricos existentes [5]. Un modelo granular se construye a un nivel de abstracción más alto y, de este modo, es capaz de hacer frente de mejor forma a las características esenciales del sistema que modela.

En este trabajo, versión resumida del artículo publicado en [1], se describe cómo generalizar los métodos numéricos de estimación de información incompleta a sus modelos granulares. En concreto, se presenta un modelo granular de estimación de información incompleta que tiene como homógolo numérico al desarrollado en [3]. Para ello, se introduce el concepto de distribución de la granularidad de la información como un factor clave para estimar los valores faltantes en una relación de preferencia difusa. Este concepto ya se ha aplicado con éxito para mejorar tanto el consenso como la consistencia en este tipo de problemas de decisión. A diferencia de los métodos numéricos existentes, se asume que los valores faltantes son gránulos de información en lugar de valores numéricos precisos. Por tanto, se introduce en la relación de preferencia un nivel de granularidad que da la flexibilidad necesaria para estimar los valores faltantes. Este concepto de granularidad de la información se emplea para optimizar (maximizar) un cierto criterio de optimización (consistencia). Es decir, los valores faltantes se estiman con el fin de aumentar lo máximo posible la consistencia de la relación de preferencia difusa.

La granularidad de la información se puede distribuir de diferentes formas. Sin embargo, por claridad de presentación, en este trabajo se usa una distribución uniforme, en la cual todos los valores estimados se tratan igual, sustituyéndose por intervalos de la misma longitud. Es decir, los gránulos de información consisten en intervalos. Por tanto, se emplea el símbolo $\mathbf{I}(R)$ para enfatizar que se usa una representación granular de la relación de preferencia difusa R, siendo, en este caso, $\mathbf{I}(\cdot)$ una familia de relaciones de preferencia intervalares asociadas a R. Así, se aprovecha el método de estimación de valores faltantes propuesto en [3] y se eleva a un nivel de abstracción superior, haciendo que los valores estimados sean gránulos de información (más generales y abstractos) para que el modelo se pueda construir en torno al marco conceptual proporcionado. Además, los intervalos se distribuyen simétricamente alrededor de los valores estimados.

En el modelo granular (intervalar) de relaciones de preferencia difusas, se debe considerar que los valores estimados se ajustan dentro de los límites que ofrece el nivel de granularidad admisible con el propósito de aumentar la consistencia de la relación de preferencia difusa. Esta mejora se produce a nivel de cada persona del grupo que participa en el proceso de toma de decisión. Esto se cuantifica mediante el siguiente índice de rendimiento:

$$Q = \frac{1}{m}\sum_{i=1}^{m} cl(R^i) \tag{1}$$

donde m es el número de personas que forman parte del grupo y $cl(R^i)$ representa el nivel de consistencia de la relación de prefencia difusa R^i proporcionada por la persona p_i, el cual se calcula según el método de estimación desarrollado en [3].

El problema de optimización consiste en maximizar el anterior índice de rendimiento. Esto se representa como:

$$\max_{R^1,R^2,\ldots,R^m \in \mathbf{I}(R)} Q \tag{2}$$

Esta tarea de optimización se realiza para cada relación de preferencia que es admisible según del nivel de granularidad de la información introducido. Dada la complejidad de esta tarea (el espacio de búsqueda es bastante grande ya que está compuesto de $\mathbf{I}(R)$), se requiere el uso de alguna técnica de optimización global avanzada. En este trabajo, se usa el algoritmo de optimización por enjambre de partículas [4]. Para esta tarea de optimización, esta técnica es adecuada ya que proporciona un considerable nivel de flexibilidad de optimización y no viene acompañada de una carga computacional excesiva.

En el algoritmo de optimización por enjambre de partículas, uno de los puntos más importantes es el establecimiento de la asociación entre la solución al problema y la representación de la partícula. En nuestro problema, cada partícula se modela mediante un vector cuyos componentes toman valores dentro del intervalo $[0, 1]$. En particular, si el grupo está formado por m personas, el vector se compone de $\sum_{i=1}^{m} \#VP^i$ componentes, siendo $\#VP^i$ el número de valores perdidos que hay en la relación de preferencia difusa R^i proporcionada por la persona p_i.

Tabla I
CONSISTENCIA OBTENIDA POR [3] Y EL MODELO PROPUESTO

	$n=5$	$n=10$	$n=15$	$n=20$
$m=5$	0.811	0.723	0.827	0.888
	0.853	**0.747**	**0.840**	**0.910**
$m=10$	0.743	0.655	0.677	0.777
	0.776	**0.682**	**0.698**	**0.801**
$m=15$	0.645	0.803	0.901	0.798
	0.688	**0.844**	**0.923**	**0.823**
$m=20$	0.754	0.771	0.697	0.866
	0.788	**0.803**	**0.727**	**0.891**

Supongamos un nivel de granularidad α, una relación de preferencia difusa incompleta R, y una entrada, r_{jk}, cuyo valor de preferencia no ha sido proporcionado. En esta entrada de $\boldsymbol{I}(R)$, el nivel de granularidad α implica un intervalo de valores posibles que se calcula mediante:

$$[d, e] = [\text{máx}(0, er_{jk} - \alpha/2), \text{mín}(er_{jk} + \alpha/2, 1)] \tag{3}$$

siendo er_{jk} el valor estimado por el método de estimación desarrollado en [3].

Para comprobar el rendimiento del modelo granular propuesto en este trabajo, se considera un conjunto de problemas de toma de decisión con diferente número de participantes (m) y de alternativas (n). Para ello, se generan de forma aleatoria relaciones de preferencia difusas incompletas y se aplica el método presentado en [3] y el modelo granular propuesto en este trabajo para estimar los valores perdidos.

La Tabla I refleja los resultados obtenidos por el método desarrollado en [3] (fuente normal) y los logrados por el modelo granular (negrita) en términos del nivel de consistencia. Se observa que el modelo granular obtiene relaciones de preferencia difusas completas con mejor consistencia.

En este trabajo, partiendo de relaciones de preferencia difusas incompletas, se ha presentado un marco algorítmico completo que da lugar a relaciones de preferencia granulares (intervalares) para estimar los valores faltantes. Se ha hecho hincapié en la motivación de usar gránulos de información para estimar los valores perdidos y obtener relaciones de preferencia difusas completas de mayor consistencia. Para una descripción más detallada de esta propuesta, consultar [1].

REFERENCIAS

[1] F. J. Cabrerizo, R. Al-Hmouz, A. Morfeq, M. A. Martínez, W. Pedrycz, and E. Herrera–Viedma, "Estimating incomplete information in group decision making: A framework of granular computing," Appl. Soft Comput., vol. 86, Jan. 2020, Art. no. 105930.

[2] S. Deparis, V. Mousseau, M. Öztürk, C. Pallier, and C. Huron, "When conflict induces the expression of incomplete preferences," Eur. J. Oper. Res., vol. 221, no. 3, pp. 602–593, Sep. 2012.

[3] E. Herrera-Viedma, F. Chiclana, F. Herrera, and S. Alonso, "Group decision–making model with incomplete fuzzy preference relations based on additive consistency," IEEE Trans. Syst., Man, Cybern. B. Cybern., vol. 37, no. 1, pp. 176–189, Feb. 2007.

[4] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in IEEE Int. Conf. on Neural Networks, Perth, Australia, 1995, pp. 1942–1948.

[5] W. Pedrycz, "From numeric models to granular system modeling," Fuzzy Inf. Eng., vol 7, no. 1, pp. 1–13, 2015.

[6] R. Ureña, F. Chiclana, J. A. Morente-Molinera, and E. Herrera-Viedma, "Managing incomplete preference relations in decision making: A review and future trends," Inf. Sci., vol. 302, pp. 14–32, May 2015.

Assisting users in decisions using fuzzy ontologies: application in the wine market

Ignacio Javier Pérez
Dept. of Computer Sciences and A. I.
University of Granada
Granada, Spain
ijperez@decsai.ugr.es

Juan Antonio Morente-Molinera
Dept. of Computer Sciences and A. I.
University of Granada
Granada, Spain
jamoren@decsai.ugr.es

Juan Miguel Tapia-García
Dept. of Quantitative Methods in E. and B.
University of Granada
Granada, Spain
jmtaga@ugr.es

Enrique Herrera-Viedma
Dept. of Computer Sciences and A. I.
University of Granada
Granada, Spain
viedma@decsai.ugr.es

Manuel Jesus Cobo
Dept. of Computer Sciences and A. I.
University of Granada
Granada, Spain
mjcobo@decsai.ugr.es

Francisco Javier Cabrerizo
Dept. of Computer Sciences and A. I.
University of Granada
Granada, Spain
cabrerizo@decsai.ugr.es

***Abstract*—Nowadays, wine has become a very popular item to purchase. There are a lot of brands and a lot of different types of wines that have different prices and characteristics. Since there is a lot of options, it is easy for buyers to feel lost among the high number of possibilities. Therefore, there is a need for computational tools that help buyers to decide which is the wine that better fits their necessities. In this article, a decision support system built over a fuzzy ontology has been designed for helping people to select a wine.**

***Index Terms*—Decision support systems; Fuzzy ontologies; Computing with words;**

I. Introduction

Wines are a popular item to purchase. Depending on the manufacture, they can have different characteristics. For instance, there are wines with different levels of alcohol, acidity, year, etc. Therefore, finding the perfect wine for a specific buyer is a quite difficult task. There is not a perfect wine for everybody since each person has different tastes and search for a specific experience. The high number of brands and wines that are available on the market make it difficult for the buyers to select the wine that better fit their necessities. They cannot handle by themselves the high amount of information about all the features, prices and brands. Therefore, there is a need for designing decision support systems (DSS) [1]–[4] that allow them to choose the wine that better fulfil their needs. This process must be carried out in an organized and fair way. The system should ask the buyers for parameters about what they need and provide them with a short list of wines that fulfil them. This way, buyers can decide which wine they should buy based on objective criteria. Thanks to this, they avoid being misled by the high amount of information. Also, they do not rely on criteria that may make the wrong choice.

This work was supported by the project B-TIC-590-UGR20 co-funded by the Programa Operativo FEDER 2014-2020 and the Regional Ministry of Economy, Knowledge, Enterprise and Universities (CECEU) of Andalusia.

In this paper (a summarized version of [4]), a novel DSS whose main purpose is to help buyers to choose a wine is designed. Users provide information to the system and they obtain several suggestions that they can use to select their most preferred wine. Fuzzy ontologies [5] (FOs) are used to store all the information in the system. For the DSS to work properly, the set of elements stored on the FO must be constantly updated. For this purpose, two different updating information processes are described.

II. A novel decision support system for assisting users in the selection of a wine

A. The Wines' Fuzzy Ontology

All the data that is used to support the users in their decision is stored in a FO [6]. Thanks to it, it is possible to have an overview of the wine that are present in the market in a concrete time. To provide up to date information to buyers, it is necessary to constantly update the FO. In this paper, two different ways of carrying out this update process are presented. The first one is to let wines' experts to do it without any computational intervention. This option is not a good one in cases like the tackled one where there is a high amount of information available. On the other hand, it is possible to design an automatic process that analyses certain Webpages to extract information about the new wines that keep appearing in the shops. For testing purposes, a real case example has been chosen: the Wine FO. Wine FO contains 623 individuals related to 6 different concepts: Year, Acidity, Alcohol, Price, Color, and Other users' opinions.

B. Decision support system design

This subsection describes how the designed DSS for choosing wines works. In Figure 1, this process is presented schematically.

1) **Preferences providing step**: Buyers provide their preferences according to the features that they want the

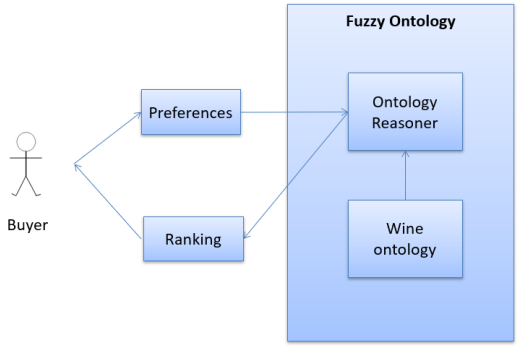


Fig. 1. Designed DSS scheme.

selected wine to have. They can choose which features they want to provide information for and which labels they want to associate to them. They can choose labels from the linguistic label sets that conform the FO to provide information to the system.

2) **FO search**: The preferences that the buyers have sent to the system are used for retrieving the individuals that better fulfil them. To carry out this task, at first, buyers determine the importance that each feature has to them by providing the set of weights in order to conform the user query. Then, the FO reasoner uses an aggregation operator to calculate the similarity that each individual of the FO has with the query. Depending on the type of relation, the similarity of the individual according to one of the concepts included in the query is calculated as a crisp relation or as a fuzzy relation.
3) **Feedback advice**: The users opinion system is called for the recommended wines. If one of the resulting wines have a low opinion value and a reasonable number of provided opinions, this issue is showed to the user as a warning. This way, users are alerted that the recommended wine did not fulfil other users expectations.
4) **Providing opinion**: After buying and testing the selected wine, the user is asked to carry out the feedback process. Its main purpose is to warn other users about problems with the selected choices.

C. Feedback process

In order to increase the reliability of the designed DSS, an user feedback process has been added to the system. Its main purpose is to aid users who are unsure about what wine they should choose. Thanks to this module, they can get benefit from the experience that other users that own their chosen wines had. The idea consists in allowing the users to provide an opinion about the wine after buying and trying it. In order to make this process as simple as possible and for avoiding users to lose a high amount of time, they only have to answer a single question. That question is: *did the wine fulfil your expectations?* The chosen wine will be recommended to people with similar tastes as the user who has already bought it. Therefore, the answer to this question will confirm if the DSS has done a good job recommending that specific wine according to that user type. A scheme of this process along with a small example is showed in Figure 2.

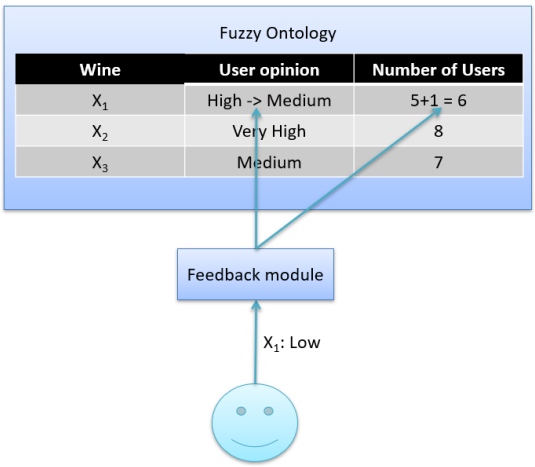


Fig. 2. Feedback process example.

D. Updating wine data procedure

The wine market is in a constant update. Wine companies keep creating new options. Therefore, in order for our DSS to be useful, information stored in the FO must be periodically updated. In this subsection, two different updating information methods are proposed. The first one takes automatically the information from the Internet. In the second one, the updating process is made manually by a wine expert.

Finally, in order to analyse and compare the behavior of the system, a discussion section is available in the extended version of the paper [4].

III. Conclusions

This article presents a DSS that help buyers to make decisions using an FO for generating recommendations. Buyers provide preferences to the system using LM and they receive the set of alternatives that better fulfil their needs according to the provided information.

References

[1] I. J. Pérez, R. Wikström, J. Mezei, C. Carlsson, and E. Herrera-Viedma, "A new consensus model for group decision making using fuzzy ontology," *Soft Computing*, vol. 17, no. 9, pp. 1617–1627, 2013.
[2] J. A. Morente-Molinera, I. J. Pérez, M. R. Ureña, and E. Herrera-Viedma, "Building and managing fuzzy ontologies with heterogeneous linguistic information," *Knowledge-Based Systems*, vol. 88, pp. 154–164, 2015.
[3] J. A. Morente-Molinera, I. J. Pérez, M. R. Ureña, and E. Herrera-Viedma, "Creating knowledge databases for storing and sharing people knowledge automatically using group decision making and fuzzy ontologies," *Information Sciences*, vol. 328, pp. 418–434, 2016.
[4] J. A. Morente-Molinera, F. J. Cabrerizo, S. Alonso, I. J. Pérez, and E. Herrera-Viedma, "Assisting users in decisions using fuzzy ontologies: Application in the wine market," *Mathematics*, vol. 8, no. 10, p. 1724, 2020.
[5] C. Carlsson, M. Brunelli, and J. Mezei, "Decision making with a fuzzy ontology," *Soft Computing*, vol. 16, no. 7, pp. 1143–1152, 2012.
[6] "The Fuzzy Wine Ontology." http://http://users.abo.fi/rowikstr/FuzzyWineOntology/, 2013. [Online; accessed 31-May-2022].

Convexity of Interval-valued Fuzzy Sets

1st Pedro Huidobro
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
huidobropedro@uniovi.es

2nd Pedro Alonso
Dept. Mathematics
University of Oviedo
Oviedo, Spain
palonso@uniovi.es

3th Vladimír Janiš
Dept. Mathematics
Matej Bel University
Banská Bystrica, Slovakia
vladimir.janis@umb.sk

4th Susana Montes
Dept. Statistics, O.R & M.E.
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—In this paper, we extend the idea of convexity for interval-valued fuzzy sets to cover some more comprehensive areas of imprecision. We show some of its attractive properties and study the conservation under the intersection and the cutworthy property. Finally, we propose a method to apply convexity to decision-making problems.**

***Index Terms*—Interval-valued fuzzy set, Epistemic interpretation, Intersection, Level set, Convexity, Decision-making**

I. Introduction

In a decision-making problem, it is common to consider at least three elements as the set of alternatives, the set of constrains and the utility function. In real-life, it could be difficult to take the exactly value, so in order to deal with that imprecision, interval-valued fuzzy sets (IVFS) can help. We also study convexity of IVFS in order to apply it to decision making problems. This communication is a summary of [4].

II. Basic concepts

Let us consider X as the referential set. An IVFS A on X is defined by $A : X \to L([0,1])$, where $L([0,1])$ is the family of closed intervals included in $[0,1]$, such that $A(x) = [\underline{A}(x), \overline{A}(x)]$. Obviously, $\underline{A}(x) \leq \overline{A}(x)\ \forall x \in X$. Let us denote the family of all IVFS on X by $IVFS(X)$.

A. Orders

As we are work with intervals, we need some interval orders for our study [3]. If $a = [\underline{a}, \overline{a}]$ and $b = [\underline{b}, \overline{b}]$ are two intervals in $L([0,1])$, we will say that a is lower than or equal to b if:

- Lattice order: $a \preceq_{Lo} b$ if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$, which is induced by the usual partial order in $\mathbb{R}^2$.
- Lexicographical order type 1: $a \preceq_{Lex1} b$ if $\underline{a} < \underline{b}$ or ($\underline{a} = \underline{b}$ and $\overline{a} \leq \overline{b}$).
- Lexicographical order type 2: $a \preceq_{Lex2} b$ if $\overline{a} < \overline{b}$ or ($\overline{a} = \overline{b}$ and $\underline{a} \leq \underline{b}$).
- The Xu and Yager order: $a \preceq_{YX} b$ if $\underline{a} + \overline{a} < \underline{b} + \overline{b}$ or ($\underline{a} + \overline{a} = \underline{b} + \overline{b}$ and $\overline{a} - \underline{a} \leq \overline{b} - \underline{b}$).

We also consider relations such as the interval dominance, Maximin, Maximax, Hurwicz or the weak order, but they are not orders, so we consider the previous ones are the most desirable. Eventually, the lexicographical orders and the Xu and Yager order belong to the family of admissible orders, which are total orders that refines the lattice order [2]. Bustince et al. [2] proposed a method for creating admissible orders from two continuous aggregations functions $\mathcal{A}$ and $\mathcal{B}$ such that $\mathcal{A}(s,t) = \mathcal{A}(u,v)$ and $\mathcal{B}(s,t) = \mathcal{B}(u,v)$ can only hold if $(s,t) = (u,v)$, $s,t,u,v \in [0,1]$.

This research has been partially supported by Spanish MINECO project PGC2018-098623-B-I00 (Pedro Huidobro and Susana Montes). Pedro Huidobro is also supported by the Severo Ochoa predoctoral grant program by the Principality of Asturias (PA-20-PF-BP19-169).

III. Operations for IVFS

A. Intersection

Given two elements A and B in $IVFS(X)$, it is said that A is contained in B if $A(x) \leq B(x)\ \forall x \in X$, for a chosen order in $L([0,1])$. Keeping the original idea of intersection for classical sets, Huidobro et al. [3] defined the o-intersection of A and B, two IVFS on X, denoted by $A \cap_o B$, as the greatest IVFS such that $A \cap_o B \subseteq_o A$ and $A \cap_o B \subseteq_o B$.

Using this definition, the chosen order plays a very important role. For instance, the Maximin, Maximax, Hurwicz or the weak order were discard because the intersection with them is not a unique IVFS, and with the interval dominance order, the result is just a point.

On the other hand, Huidobro et al. [4] proved that if A and B are sets in $IVFS(X)$, then, for any $x \in X$ we have that:

- for the lattice order:

$$A \cap_{Lo} B(x) = [\min\{\underline{A(x)}, \underline{B(x)}\}, \min\{\overline{A(x)}, \overline{B(x)}\}],$$

- for $\preceq_o$, an admissible order:

$$A \cap_o B(x) = \begin{cases} A(x) & \text{if } A(x) \preceq_o B(x), \\ B(x) & \text{if } B(x) \preceq_o A(x). \end{cases}$$

B. Union

They did a similar study with the union obtaining that [4]:

- for the lattice order:

$$A \cup_{Lo} B(x) = [\max\{\underline{A(x)}, \underline{B(x)}\}, \max\{\overline{A(x)}, \overline{B(x)}\}],$$

- for $\preceq_o$, an admissible order:

$$A \cup_o B(x) = \begin{cases} A(x) & \text{if } B(x) \preceq_o A(x), \\ B(x) & \text{if } A(x) \preceq_o B(x). \end{cases}$$

C. Level sets

One interesting concept of fuzzy sets and convexity is the α-cut or level set [5]. Here, we present a suitable definition of a level set for IVFS.

***Definition** 3.1:* [4] Let $\preceq_o$ be an order on $L([0,1])$. For any $A \in IVFS(X)$ and for any $[\alpha,\beta] \in L([0,1])$, we define the $[\alpha,\beta]$-level sets of A w.r.t. the order $\preceq_o$ as it follows:

$$A^o_{[\alpha,\beta]} = \{x \in X : [\alpha,\beta] \preceq_o A(x)\}.$$

With them, we were able to prove some properties and the decomposition theorem:

***Theorem** 3.1 (Decomposition Theorem):* [4] Let $\preceq_o$ be a total order in $L([0,1])$. For every $A \in IVFS(X)$, we have that

$$A = \underset{[\alpha,\beta]\in L([0,1])}{\cup_o} {}^o_{[\alpha,\beta]}A,$$

where ${}^o_{[\alpha,\beta]}A(x) = [\alpha,\beta]$ if $x \in A^o_{[\alpha,\beta]}$ and 0 otherwise.

IV. Convexity

We consider the following definition of convexity:

***Definition** 4.1:* [3] Let X be an ordered space and let $\preceq_o$ be an order in $L([0,1])$. An IVFS A on X is said to be o-convex (resp. strict o-convex), if for each $x < y < z$ in X the following inequalities are fulfilled: $A(x) \preceq_o A(y)$ or $A(z) \preceq_o A(y)$ (resp. $A(x) \prec_o A(y)$ or $A(z) \prec_o A(y)$)

With this definition of convexity, if A is a o-convex IVFS, then $A^o_{[\alpha,\beta]}$ are convex crisp sets for all $[\alpha,\beta] \in L([0,1])$, but the converse is true if $\preceq_0$ is a total order.

For total orders we have also obtained that the intersection of two o-convex (resp. strictly o-convex) IVFS is also o-convex (resp. strictly o-convex), whenever it is not empty.

Considering convex IVFS we can achieve good results about optimization:

***Theorem** 4.1:* [4] Let A be a convex IVFS over the ordered space X. Let $\preceq_o$ be an order on $L([0,1])$. If $x^* \in supp(A) = \{x \in X : [0,0] \prec_o A(x)\}$ is a strict local maximizer of $A(x)$, then it is also a global maximizer of $A(x)$ over $supp(A)$. The set of points at which $A(x)$ attains its global maximum over its support is a crisp convex set.

In addition, our results are improved using strict convexity:

***Theorem** 4.2:* [4] Let A be a strictly convex IVFS over the ordered space X. Let $\preceq_o$ be an order on $L([0,1])$.

i) If $x^* \in supp(A)$ is a local maximizer of $A(x)$, then it is also a global maximizer.
ii) $A(x)$ attains its maximum over $supp(A)$ at no more than one point.

V. Decision Making

In this section, a decision-making perspective is presented using the previous concepts. The idea is to follow the approach given by Bellman and Zadeh [1], where they consider that the constrains and the goals are connected in a symmetric way and the one showed by Yager and Basson [6], where they construct the decision as the intersection between the goals and constraints. Taking into account these points of view, Huidobro et al. [4] propose the following definition:

***Definition** 5.1:* [4] Let $X = \{x_1,,\ldots,x_n\}$ be the set of alternatives, $G_1,,\ldots G_p$ be the set of goals that can be expressed as IVFSs on the space of alternatives, and $C_1,\ldots,C_m$ be the set of constraints that can also be expressed as IVFSs on the space of alternatives. Let $\preceq_o$ be an order on $L([0,1])$. The goals and constraints are combined to form a decision D, which is an IVFS resulting from the intersection of the goals and the constrains. Thus, $D = G_1 \cap_o, \ldots \cap_o G_p \cap_o C_1 \cap_o \ldots \cap_o C_m$.

Combining this interpretation with Theorems 4.1 and 4.2, we obtain the following consequence:

***Corollary** 5.1:* [4] Let $\preceq_o$ be an order on $L([0,1])$, let $G_1,\ldots,G_p$ be the interval-valued fuzzy goals, $C_1,\ldots,C_m$ the interval-valued fuzzy constraints, and $D = G_1 \cap, \ldots \cap G_p \cap C_1 \cap, \ldots \cap C_m$ be the resulting decision.

- If the interval-valued fuzzy goals and the interval-valued fuzzy constraints are convex IVFS, then the resulting decision D is a convex IVFS and the set of maximizing decisions of the IVFS D is a convex crisp set.
- If the interval-valued fuzzy goals and the interval-valued fuzzy constraints are strictly convex IVFS, then the resulting decision D is a strictly convex IVFS and the set of maximizing decisions of D is a singleton or an empty set.

VI. Conclusions

In this paper, a definition of convexity for IVFS based on the order relation considered is presented. This definition satisfies properties as preservation of convexity under intersections and the cutworthy property. The importance of the order was shown, proving that we obtain good results with admissible orders. They also play an important role on the definition of convexity, achieving nice results as the decomposition theorem for IVFS, which characterizes an IVFS through its level sets. A new proposal about the use of interval-valued fuzzy sets and convexity to decision-making problems is present. It has to be mentioned that the order selected on $L([0,1])$ is closely linked to the preferred alternative.

References

[1] R. E. Bellman and L. A. Zadeh. Decision-making in a fuzzy environment. *Management Science*, 17(4):B–141–B–164, 1970.

[2] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar. Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems*, 220:69 – 77, 2013.

[3] P. Huidobro, P. Alonso, V. Janiš, and S. Montes. Orders preserving convexity under intersections for interval-valued fuzzy sets. In Marie-Jeanne Lesot, Susana Vieira, Marek Z. Reformat, João Paulo Carvalho, Anna Wilbik, Bernadette Bouchon-Meunier, and Ronald R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 493 – 505, Cham, 2020. Springer International Publishing.

[4] P. Huidobro, P. Alonso, V. Janiš, and S. Montes. Convexity and level sets for interval-valued fuzzy sets. *Fuzzy Optimization and Decision Making*, pages 1–28, 2022.

[5] G.J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall PTR, 1995.

[6] R. Yager and D. Basson. Decision making with fuzzy sets. *Decision Sciences*, 6(3):590–600, 1975.

Un estudio preliminar de relaciones de clausura difusas

Manuel Ojeda-Hernández
Dept. Matemática Aplicada
Universidad de Málaga
Málaga
manuojeda@uma.es

Inma P. Cabrera
Dept. Matemática Aplicada
Universidad de Málaga
Málaga
ipcabrera@uma.es

Pablo Cordero
Dept. Matemática Aplicada
Universidad de Málaga
Málaga
pcordero@uma.es

Emilio Muñoz-Velasco
Dept. Matemática Aplicada
Universidad de Málaga
Málaga
ejmunoz@uma.es

***Resumen*—Los operadores de clausura son elementos clave de las matemáticas tanto puras como aplicadas. Esta contribución trata la búsqueda de una definición de relación de clausura difusa que extienda de manera apropiada el concepto de operador de clausura en el marco de los retículos completos difusos. La condición que se busca extender es la correspondencia biyectiva con los sistemas de clausura difusos. Se parte de las definiciones existentes de relación de clausura difusa y se acotan las condiciones necesarias para la existencia de la biyección hasta que se encuentran las condiciones óptimas.**

***Index Terms*—Relación difusa, Operador de clausura, Retículo completo, Lógica difusa, Extensional**

I. Introducción

Este trabajo es un resumen del artículo [1], que está pendiente de publicación. El marco en el que se desarrolla es el de retículo completo difuso que fue definido por Bělohlávek en [2] con el nombre de conjunto $\mathbb{L}$-ordenado completamente retículo.

Aunque la definición de operador de clausura difuso parece ser la misma para la mayoría de autores, esto es, una aplicación clásica que satisface una versión difusa de ser inflacionaria, isótona e idempotente, la situación es muy distinta en el caso de los sistemas de clausura difusos. Hay distintas definiciones de esta noción en la bibliografía, por ejemplo, [3], [4].

El objetivo de este trabajo es definir el concepto de relación de clausura difusa de manera que esté en biyección con la noción de sistema de clausura difuso usada en [4].

Se comprueba que las relaciones de clausura difusas que se han definido anteriormente en la bibliografía como relaciones que son inflacionarias, isótonas e idempotentes, no están en correspondencia biyectiva con los sistemas de clausura difusos. Para que sí lo estén, es necesario introducir una definición más fuerte.

Por último, presentamos las conclusiones y las posibles líneas de trabajo futuro.

II. Preliminares

A lo largo del artículo, $\mathbb{L} = (L, \wedge, \vee, \otimes, \to, 0, 1)$ será un retículo residuado completo. Se dice que un conjunto no vacío A con una $\mathbb{L}$-relación binaria ρ, es un poset difuso si ρ es un *orden difuso*, es decir, si ρ es reflexiva, antisimétrica y transitiva.

Definición 1 ([2])**.** *Decimos que un poset difuso (A, ρ) es un retículo completo difuso si todo subconjunto difuso $X \in L^A$ tiene supremo e ínfimo.*

Concluimos esta sección con la definición de las estructuras de clausura en A.

Definición 2. *Dado un poset difuso $\mathbb{A} = (A, \rho)$, se dice que una aplicación* $\mathtt{c}\colon A \to A$ *es un* operador de clausura *si satisface las siguientes condiciones:*

1. $\rho(a, b) \leq \rho(\mathtt{c}(a), \mathtt{c}(b))$*, para todo* $a, b \in A$ *(isótona)*
2. $\rho(a, \mathtt{c}(a)) = 1$*, para todo* $a \in A$ *(inflacionaria)*
3. $\rho(\mathtt{c}(\mathtt{c}(a)), \mathtt{c}(a)) = 1$*, para todo* $a \in A$ *(idempotente)*

La definición de sistema de clausura que se va a usar en este trabajo es la siguiente, introducida originalmente en [4].

Definición 3. *Sea (A, ρ) un retículo completo difuso. Se dice que un conjunto $\mathcal{F} \subseteq A$ es un sistema de clausura si $\bigsqcap X \in \mathcal{F}$ para todo subconjunto difuso $X \in L^{\mathcal{F}}$.*

Cuando en lugar de considerar conjuntos clásicos tomamos conjuntos difusos, se puede definir un sistema de clausura difuso. La definición es la siguiente.

Definición 4. *Sea (A, ρ) un retículo completo difuso. Se dice que un conjunto difuso $\Psi \in L^A$ es un sistema de clausura difuso si su núcleo, $\Psi^{-1}(1)$, es un sistema de clausura y es el menor conjunto extensional que contiene a su núcleo.*

III. Relaciones de clausura difusas

Las relaciones de clausura difusas no son una novedad, estas estructuras se han usado anteriormente en la bibliografía. En [5], se definen en el marco de los conjuntos parcialmente ordenados de la siguiente manera.

Definición 5. *Sea $\mathbb{A} = (A, \rho)$ un poset difuso. Una relación difusa total $\kappa\colon A \times A \to L$ es* relación de clausura difusa *si verifica:*

- *κ es* inflacionaria*, es decir, $\rho_{\propto}(a, a^{\kappa}) = 1$ para todo $a \in A$.*

- κ *es* isótona*, es decir,* $\rho(a_1, a_2) \leq \rho_\propto(a_1^\kappa, a_2^\kappa)$ *para todo* $a_1, a_2 \in A$*.*
- κ *es* idempotente*, es decir,* $\rho_\propto(a^{\kappa\circ\kappa}, a^\kappa) = 1$ *para todo* $a \in A$*.*

donde $\rho_\propto(X, Y) = \bigwedge_{x,y\in A}(X(x) \otimes Y(y)) \to \rho(x, y)$*, para* $X, Y \in L^A$*.*

Teorema 6. *Sea* $\mathbb{A} = (A, \rho)$ *un retículo completo difuso. Entonces se verifica:*

1. *Si* κ *es una relación de clausura difusa, el conjunto difuso* Φ_κ *definido como* $\Phi_\kappa(a) = \rho_\propto(a^\kappa, a)$ *es un sistema de clausura difuso.*
2. *Si* Φ *es un sistema de clausura difuso, entonces la relación* $\kappa_\Phi \colon A \times A \to L$ *definida como* $\kappa_\Phi(a, b) = (\bigsqcap(a^\rho \otimes \Phi) \approx b)$ *es una relación de clausura difusa.*
3. *Si* Φ *es un sistema de clausura difuso, entonces* $\Phi = \Phi_{\kappa_\Phi}$*.*

Aunque la Definición 5 es la extensión natural de un operador de clausura al marco difuso y es interesante, la igualdad $\kappa_{\Phi_\kappa} = \kappa$ no se cumple en general. Por tanto, se necesita un concepto más fuerte para que la correspondencia sea biyectiva.

IV. Relaciones de clausura difusas fuertes

Para encontrar qué tipo de relaciones están en biyección con los sistemas de clausura difusos, se analizarán las propiedades de las relaciones κ_Φ definidas a partir de un sistema de clausura difuso $\Phi \in L^A$.

Nos centraremos en la extensionalidad, que ya era una propiedad fundamental en trabajos previos [4]. Una relación difusa $\kappa \colon A \times A \to L$ es extensional si lo es al considerarse como subconjunto difuso de $A \times A$, es decir, para todo $a_1, a_2, b_1, b_2 \in A$,

$$\kappa(a_1, b_1) \otimes (a_1 \approx a_2) \otimes (b_1 \approx b_2) \leq \kappa(a_2, b_2).$$

Es interesante destacar la relación difusa κ_Φ es extensional para todo $\Phi \in L^A$.

Estas propiedades apuntan a que las relaciones difusas en correspondencia biyectiva con los sistemas de clausura difusos han de ser extensionales. En general, esta condición no es suficiente para asegurar la biyección.

Si se añade una condición de minimalidad llegamos a la siguiente definición, que sí satisface las características deseadas.

Definición 7. *Sea* $\mathbb{A} = (A, \rho)$ *un retículo difuso completo. Se dice que una relación de clausura difusa* $\kappa \colon A \times A \to L$ *es fuerte si es minimal en el conjunto de las relaciones de clausura difusas extensionales, i.e.* κ *es extensional y, para toda relación difusa de clausura extensional* $\kappa_1 \colon A \times A \to L$*, se tiene que* $\kappa_1 \leq \kappa$ *implica* $\kappa_1 = \kappa$*.*

No toda relación de clausura difusa es una relación de clausura difusa fuerte.

El resultado principal de este artículo muestra que las relaciones de clausura difusas fuertes están en correspondencia biyectiva con los sistemas de clausura difusos.

Teorema 8. *Sea* $\mathbb{A} = (A, \rho)$ *un retículo completo difuso. Se tiene:*

1. *Si* κ *es una relación de clausura difusa fuerte, entonces* Φ_κ *es un sistema de clausura difuso.*
2. *Si* Φ *es un sistema de clausura difuso, entonces la relación difusa* κ_Φ *es una relación de clausura difusa fuerte.*
3. *Si* $\kappa \colon A \times A \to L$ *es una relación de clausura difusa fuerte, entonces* $\kappa_{\Phi_\kappa} = \kappa$*.*
4. *Si* Φ *es un sistema de clausura difuso, entonces* $\Phi_{\kappa_\Phi} = \Phi$*.*

V. Conclusiones y trabajo futuro

Este artículo presenta una extensión de las relaciones de clausura difusas. La definición introducida, llamada relación de clausura difusa fuerte, se ha mostrado que está en correspondencia biyectiva con los sistemas de clausura difusos. También hemos señalado que esta extensión está relacionada con las funciones difusas perfectas.

A corto plazo, trataremos de buscar otras extensiones de operadores de clausura que sean interesantes desde el punto de vista práctico pero no mantenga la biyección con los sistemas de clausura. Por otro lado, sería interesante estudiar la conexión entre κ_Φ y $\mu_\Phi(a, b) = m(a^\rho \otimes \Phi)(b)$, donde m es el conjunto difuso de mínimos, que tiene interés en ser considerado cuando $\mathbb{A} = (A, \rho)$ no tiene estructura de retículo.

Reconocimientos

Este trabajo ha sido parcialmente financiado por la beca FPU19/01467 (MCIU) y los proyectos de investigación nacionales con referencia TIN2017-89023-P (MCIU/AEI/FEDER) y PGC2018-095869-B-I00 y el proyecto de investigación autonómico UMA2018-FEDERJA-001 (JA/UMA/FEDER).

Referencias

[1] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco, "Fuzzy closure relations," *Fuzzy sets and systems*, To appear.
[2] R. Bělohlávek, "Concept lattices and order in fuzzy logic," *Annals of Pure and Applied Logic*, vol. 128, no. 1, pp. 277–298, 2004.
[3] R. Bělohlávek, "Fuzzy closure operators," *Journal of Mathematical Analysis and Applications*, vol. 262, pp. 473–489, 2001.
[4] M. Ojeda-Hernández, I. P. Cabrera, P. Cordero, and E. Muñoz-Velasco, "Closure systems as a fuzzy extension of meet-subsemilattices," in *Joint Proceedings of the 19th World Congress of the 12th Conference of the EUSFLAT*. Atlantis Press, 2021, pp. 40–47.
[5] I. P. Cabrera, P. Cordero, E. Muñoz-Velasco, M. Ojeda-Aciego, and B. De Baets, "Relational Galois connections between transitive digraphs: Characterization and construction," *Inf. Sci.*, vol. 519, pp. 439–450, 2020.

T-norms and t-conorms on a family of lattices

C. Bejines
dept. Matemática Aplicada
Universidad de Málaga
Málaga, Spain
cbejines@uma.es

***Abstract*—The article provided a complete classification of all t-norms on a family of lattices in terms of t-norms on discrete chains. Moreover, the cardinal of some classes on discrete chains was computed. Therefore, the number of t-norms on the family of lattices was obtained. Also, new results involving Archimedean and divisible t-norms were presented and we brang out dual results for t-conorms.**

***Index Terms*—T-norm, t-conorm, lattice.**

I. Introduction

T-norms and t-conorms are basic tools in the framework of Fuzzy Logic. They extend the conjunction and disjunction of classical sets and are suitable to define fuzzy algebraic structures. Firstly, they were defined on the interval $[0,1]$, but the need to work with incomparable elements is required in many contexts, so a more general algebraic structure is fundamental (see [6]). Consequently, they have been studied on bounded lattices.

One object of interest is to find which operators defined on bounded lattices are t-norms (or t-conorms), or at least to estimate the number of them. If the bounded lattice is not finite, the number of them is infinite. Also, finite lattices are used in the practice. For these reasons, we have focused on a family of finite lattices in the paper.

A pioneering article in this context was written by De Baets and Mesiar in 1999. They provided the number of t-norms on discrete chains up to a length of fourteen by computational methods (see [5] and table I).

n	P_n	n	P_n
1	2	7	13775
2	6	8	86417
3	22	9	590489
4	94	10	4446029
5	451	11	37869449
6	2386	12	382549464

TABLE I
NUMBERS OF T-NORMS ON DISCRETE CHAINS WITH $n+2$ ELEMENTS, THAT IS, n NON-TRIVIAL ELEMENTS.

Also, Bartušek and Navara studied conjunctions on finite chains using a computer program generating all t-norms (see [1]). A way to obtain new t-norms from ones already given is through the ordinal sum. However, the ordinal sum of t-norms on a finite lattice does not have to be a t-norm. In 2008, Saminger et al. studied the ordinal sum of t-norms and the horizontal sum of bounded lattices (see [10]). Recently, the number of t-norms on other families of lattices is found in [3].

It is worth noting that t-norms and t-conorms are not the only operators that are being investigated. Uninorms were recently defined on bounded lattices (see [7]) and characterized by means of t-norms and t-conorms (see [4]). More generally, aggregation functions and t-operators can be defined on bounded lattices (see [8]). All these operations are generalizations of the two concepts that we study in-depth in this paper: the t-norm and the t-conorm.

II. Main results

This extended abstract is based on [2]. In the article, we describe each and every one of the t-norms defined on a family of finite lattices in terms of the t-norms defined on the discrete chains. Some of them can be obtained using the ordinal sum, but most are not. The lattices (see Figure 1) resemble the horizontal sum of discrete chains (see [9]). Our family of lattices has the same elements as a horizontal sum of discrete chains, but the imposition of a new connection between two elements makes our lattice more complex to study.

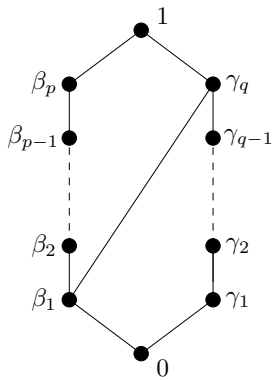


Fig. 1. The lattice L.

We denote by C_β and C_γ the chains $\{0,\beta_1,\beta_2,...,\beta_p,1\}$ and $\{0,\gamma_1,\gamma_2,...,\gamma_q,1\}$ respectively. Below, we present the six theorems which provide each t-norm in terms of discrete t-norms.

GACR 19-09967S, UMA2018-FEDERJA-001, POSTDOC-21-00875.

Theorem 1. Each t-norm T on L satisfying $T(\beta_1, \beta_1) = \beta_1$ and $T(\gamma_q, \gamma_q) = \beta_1$ is expressed as follows:

$$T(x,y) = \begin{cases} T'(x,y) & \text{if} \quad x,y \in C_\beta, \\ x \wedge y & \text{if} \quad x = 1 \text{ or } y = 1, \\ \beta_1 & \text{if} \quad (x,y) = (\beta_i, \gamma_q), \\ \beta_1 & \text{if} \quad (x,y) = (\gamma_q, \beta_i), \\ \beta_1 & \text{if} \quad x = y = \gamma_q, \\ 0 & \text{otherwise.} \end{cases}$$

where T' is a t-norm on C_β satisfying $T'(\beta_1, \beta_1) = \beta_1$.

Theorem 2. Each t-norm T on L satisfying $T(\beta_1, \beta_1) = \beta_1$ and $T(\gamma_q, \gamma_q) = \gamma_q$ is expressed as follows:

$$T(x,y) = \begin{cases} T_1(x,y) & \text{if} \quad x,y \in C_\beta, \\ T_2(x,y) & \text{if} \quad x,y \in C_\gamma, \\ \beta_1 & \text{if} \quad (x,y) = (\beta_i, \gamma_q), \\ \beta_1 & \text{if} \quad (x,y) = (\gamma_q, \beta_i), \\ 0 & \text{otherwise.} \end{cases}$$

where T_1 is a t-norm on C_β satisfying $T_1(\beta_1, \beta_1) = \beta_1$ and T_2 is a t-norm on C_γ satisfying $T_2(\gamma_q, \gamma_q) = \gamma_q$.

Theorem 3. Each t-norm T on L satisfying $T(\beta_1, \beta_1) = 0$ and $T(\gamma_q, \gamma_q) = \beta_1$ is expressed as follows: There is $k \in \{1, 2, ..., p\}$ such that

$$T(x,y) = \begin{cases} T_k(x,y) & \text{if} \quad x,y \in C_\beta, \\ x \wedge y & \text{if} \quad x = 1 \text{ or } y = 1, \\ \beta_1 & \text{if} \quad (x,y) = (\beta_i, \gamma_q), i > k, \\ \beta_1 & \text{if} \quad (x,y) = (\gamma_q, \beta_i), i > k, \\ \beta_1 & \text{if} \quad x = y = \gamma_q, \\ 0 & \text{otherwise.} \end{cases}$$

where T_k is a t-norm on C_β satisfying $T_k(\beta_1, \beta_p) = 0$ and $T_k(\beta_p, \beta_p) \leq \beta_k$.

Theorem 4. Each t-norm T on L satisfying $T(\beta_1, \beta_1) = 0$ and $T(\gamma_q, \gamma_q) \in C_\gamma \backslash \{\gamma_q\}$ is expressed as one of the following two formulas:

$$T(x,y) = \begin{cases} T_1(x,y) & \text{if} \quad x,y \in C_\beta, \\ T_2(x,y) & \text{if} \quad x,y \in C_\gamma, \\ 0 & \text{otherwise.} \end{cases}$$

where T_1 is a t-norm on C_β satisfying $T_1(\beta_1, \beta_1) = 0$ and T_2 is a t-norm on C_γ satisfying $T_2(\gamma_q, \gamma_q) < \gamma_q$, or, there is $k \in \{2, 3, ..., p\}$ such that

$$T(x,y) = \begin{cases} T_1(x,y) & \text{if} \quad x,y \in C_\beta, \\ T_2(x,y) & \text{if} \quad x,y \in C_\gamma, \\ \beta_1 & \text{if} \quad (x,y) = (\beta_i, \gamma_q), i \geq k, \\ \beta_1 & \text{if} \quad (x,y) = (\gamma_q, \beta_i), i \geq k, \\ 0 & \text{otherwise.} \end{cases}$$

where T_2 is a t-norm on C_γ satisfying $T_2(\gamma_q, \gamma_q) < \gamma_q$ and T_1 is a t-norm on C_β satisfying

1) $T_1(\beta_i, \beta_1) = 0$ for $i < k$.
2) If $T_1(\beta_k, \beta_k) = \beta_k$, then

$$T_1(\beta_1, \beta_p) = T_1(\beta_1, \beta_k) = \beta_1.$$

3) If $T_1(\beta_k, \beta_k) < \beta_k$, then

$$T_1(\beta_1, \beta_p) = T_1(\beta_1, \beta_k) = 0$$

and

$$T(\beta_p, \beta_p) \leq \beta_{k-1}.$$

Theorem 5. Each t-norm T on L satisfying $T(\beta_1, \beta_1) = 0$, $T(\gamma_q, \gamma_q) = \gamma_q$, and $T(\beta_1, \gamma_q) = \beta_1$ is expressed as follows:

$$T(x,y) = \begin{cases} T'(x,y) & \text{if} \quad x,y \in C_\gamma, \\ x \wedge y & \text{if} \quad x = 1 \text{ or } y = 1, \\ \beta_1 & \text{if} \quad (x,y) = (\beta_i, \gamma_q), \\ \beta_1 & \text{if} \quad (x,y) = (\gamma_q, \beta_i), \\ 0 & \text{otherwise.} \end{cases}$$

where T' is a t-norm on C_γ satisfying $T'(\gamma_q, \gamma_q) = \gamma_q$.

Theorem 6. Each t-norm T on L satisfying $T(\gamma_q, \gamma_q) = \gamma_q$ and $T(\beta_1, \gamma_q) = 0$ is expressed as follows:

$$T(x,y) = \begin{cases} T_1(x,y) & \text{if} \quad x,y \in C_\beta, \\ T_2(x,y) & \text{if} \quad x,y \in C_\gamma, \\ 0 & \text{otherwise.} \end{cases}$$

where T_1 is a t-norm on C_β satisfying $T_1(\beta_1, \beta_1) = 0$, and T_2 is a t-norm on C_γ satisfying $T_2(\gamma_q, \gamma_q) = \gamma_q$.

CONCLUSION

We have provided a complete classification and description of the t-norms defined on a family of lattices in terms of t-norms on discrete chains. Complementarily, the number of t-norms required on discrete chains is obtained by computational methods. We have done the same for Archimedean and divisible t-norms and the last part has highlighted that each result can be transferred to t-conorms (see published article [2]). Similar arguments could be applied to some more general lattices.

ACKNOWLEDGEMENTS

I wish to acknowledge Mirko Navara for his feedback and suggestions that have helped to present the results. This research was supported by the Ministry of Education of the Czech Republic under Project GACR 19-09967S. Also, I acknowledge the financial support of the Junta de Andalucía, Grants: UMA2018-FEDERJA-001 and POSTDOC-21-00875.

REFERENCES

[1] Bartušek, T., & Navara, M. (2002). Program for generating fuzzy logical operations and its use in mathematical proofs. *Kybernetika*, *38*, 235–244.
[2] Bejines, C. (2022). T-norms and t-conorms on a family of lattices. *Fuzzy Sets and Systems*, *439*, 55–79.
[3] Bejines, C., Bruteničová, M., Chasco, M., Elorza, J., & Janiš, V. (2021). The number of t-norms on some special lattices. *Fuzzy Sets and Systems*, *408*, 26–43.
[4] Çaylı, G. D. (2018a). A characterization of uninorms on bounded lattices by means of triangular norms and triangular conorms. *International Journal of General Systems*, *47*, 772–793.
[5] De Baets, B., & Mesiar, R. (1999). Triangular norms on product lattices. *Fuzzy Sets and Systems*, *104*, 61–75.
[6] Goguen, J. A. (1967). L-fuzzy sets. *Journal of mathematical analysis and applications*, *18*, 145–174.
[7] Karaçal, F., & Mesiar, R. (2015). Uninorms on bounded lattices. *Fuzzy Sets and Systems*, *261*, 33–43.
[8] Karaçal, F., & Mesiar, R. (2017). Aggregation functions on bounded lattices. *International Journal of General Systems*, *46*, 37–51.
[9] Saminger-Platz, S. (2006). On ordinal sums of triangular norms on bounded lattices. *Fuzzy Sets and Systems*, *157*, 1403–1416.
[10] Saminger-Platz, S., Klement, E. P., & Mesiar, R. (2008). On extensions of triangular norms on bounded lattices. *Indagationes Mathematicae*, *19*, 135–150.

Un modelo recíproco de Preferencia-Aversión

1st J. Tinguaro Rodríguez
Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid
Madrid, España
jtrodrig@mat.ucm.es

2nd Camilo Franco
Joint Research Centre (JRC)
Comisión Europea
Ispra, Italia
camilo.franco-de-los-rios@ec.europa.eu

3rd Javier Montero
Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid
Madrid, España
monty@mat.ucm.es

***Abstract*—Este trabajo presenta una instancia particular del modelo de Preferencia-Aversión (P-A), buscando combinar la rica capacidad expresiva de éste con la simplicidad operativa que exhiben las relaciones de preferencia recíprocas. El modelo P-A recíproco propuesto se construye entonces a partir de la conjunción de un par de estructuras recíprocas compatibles, una de carácter positivo y otra de carácter negativo, que permiten expresar, respectivamente, preferencia y ausencia de preferencia, y aversión y ausencia de aversión. Con unos requisitos de información inicial similares, la estructura P-A recíproca resultante de la combinación de estas estructuras recíprocas básicas y su posterior refinado ofrece una capacidad expresiva diferente, y en cierto modo más amplia, que la del modelo borroso estándar de representación de preferencias, y en particular admite la representación de una noción de conflicto o ambivalencia entre alternativas que constituye una novedad en relación a las capacidades expresivas del modelo estándar y recíproco.**

***Index Terms*—Estructuras de preferencia borrosas, modelo de Preferencia-Aversión, estructuras de preferencia recíprocas compatibles**

I. Introducción

Desde mediados del siglo pasado, las relaciones de preferencia han sido una herramienta central para la representación de las actitudes de los decisores y la descomposición de la complejidad de los problemas de toma de decisiones. Una característica principal de estas relaciones de preferencia es que pueden ser organizadas/descompuestas en términos de estructuras de preferencia, en las que coexisten diversas relaciones complementarias, cada una representando una noción preferencial separada, tal como preferencia estricta o indiferencia. Esto permite una especificación más rigurosa y clara de la semántica y la capacidad expresiva (i.e. la cantidad y naturaleza de las diferentes nociones preferenciales representables) de los modelos de toma de decisiones. En este sentido, las estructuras de preferencia con mayor capacidad expresiva permiten una representación más detallada y realista de los problemas de decisión, mientras que los modelos más simples ofrecen una mayor operatividad al requerir normalmente una menor cantidad de información inicial para comparar cada par de alternativas. Por ello, es importante desarrollar modelos de representación de preferencias que logren un equilibrio adecuado entre simplicidad y capacidad expresiva.

Este trabajo ha sido financiado parcialmente por el Gobierno de España (proyecto PGC2018-096509-B-100) y la Universidad Complutense de Madrid (grupo 910149).

Con este equilibrio en mente, el objetivo de este trabajo es proponer un modelo de representación de preferencias borrosas que permita representar un espectro de nociones preferenciales algo más amplio que el modelo estándar borroso [1] sin incurrir en una mayor necesidad de información inicial. En particular, nuestra propuesta se desarrolla como un caso particular del modelo bipolar de Preferencia-Aversión (P-A, ver [2], [3]), imponiendo la condición de reciprocidad sobre las estructuras básicas (positiva y negativa) de este modelo P-A. Esto conduce entonces a un modelo P-A recíproco, que se construye sobre la base de un par de estructuras recíprocas compatibles (ver [4]), mostrando respectivamente un carácter positivo (preferencia) y negativo (aversión). La estructura P-A recíproca completa surge entonces de la combinación conjuntiva de estas estructuras recíprocas básicas y su refinado posterior mediante la agrupación de relaciones que reflejen información equivalente.

II. Preliminares

Recordemos que una relación (débil) de preferencia borrosa se identifica con una relación binaria valuada $R : \mathbb{A}^2 \to [0, 1]$, que asigna, a cada par de alternativas a, b del conjunto de alternativas disponibles $\mathbb{A}$, el grado $R(a, b) \in [0, 1]$ en que a es al menos tan preferida como b. El modelo estándar de preferencias borrosas [1] permite entonces asociar a esta relación R una estructura de preferencia $R = \langle P, I, J\rangle$, donde P denota preferencia estricta, I indiferencia y J incomparabilidad. Para obtener esta descomposición este modelo precisa de la información inicial dada por las comparaciones $R(a, b)$ y $R^{-1}(a, b) = R(b, a)$ para cada par de alternativas $a, b \in \mathbb{A}$.

Por otro lado, el modelo de Preferencia-Aversión [2] considera dos relaciones débiles $R_+, R_- : \mathbb{A}^2 \to [0, 1]$, representando respectivamente dimensiones independientes de preferencia (de carácter positivo) y aversión (de carácter negativo), y asociando a cada una de ellas una estructura de preferencia estándar, respectivamente $R_+ = \langle P, I, J\rangle$ y $R_- = \langle Z, G, H\rangle$, donde las relaciones P, I, J mantienen el significado expuesto, mientras que Z, G, H constituyen sus contrapartes negativas –i.e. Z denota aversión estricta, G indiferencia en aversión y H incomparabilidad en aversión. La combinación de estas dos estructuras permite entonces expresar un amplio rango de hasta 10 nociones preferenciales. No obstante, el modelo P-A duplica las necesidades de información inicial del modelo estándar, requiriendo para cada 2 alternativas $a, b \in \mathbb{A}$ un par

de comparaciones positivas ($R_+(a,b)$ y $R_+^{-1}(a,b)$)) y un par de comparaciones negativas ($R_-(a,b)$ y $R_-^{-1}(a,b)$)).

Finalmente, una relación de preferencia recíproca es una relación borrosa $Q : \mathbb{A}^2 \to [0,1]$ que cumple la condición de reciprocidad $Q(a,b) + Q(b,a) = 1\ \forall a,b \in \mathbb{A}$. Esta restricción implica que el modelo recíproco requiere la mitad de información inicial que el modelo estándar, i.e. solamente una comparación $Q(a,b)$ para cada par de alternativas $a,b \in \mathbb{A}$. La condición de reciprocidad impide considerar las relaciones recíprocas como relaciones de preferencia débil, por lo que en principio no es posible aplicar sobre ellas la descomposición del modelo estándar. Sin embargo, como se muestra en [4], a una relación recíproca es posible asociarle (de manera compatible con el modelo estándar) una estructura de preferencia $Q = \langle P, L\rangle$, donde P sigue denotando preferencia estricta y L denota ausencia de preferencia, una relación simétrica que aglutina de manera indistinguible las nociones de indiferencia e incomparabilidad.

III. La estructura P-A recíproca

En aras de combinar la capacidad expresiva del modelo P-A con la simplicidad del modelo recíproco, nuestra propuesta consiste en imponer la condición de reciprocidad en ambas dimensiones, positiva y negativa, del modelo P-A. Esto es, consideraremos que la información de partida viene dada por dos relaciones borrosas $Q_+, Q_- : \mathbb{A}^2 \to [0,1]$ tales que $Q_+(a,b)+Q_+(b,a) = 1$ y $Q_-(a,b)+Q_-(b,a) = 1\ \forall a,b \in \mathbb{A}$. Nótese entonces que para cada par de alternativas $a,b \in \mathbb{A}$ la información inicial requerida consiste únicamente en una comparación positiva, $Q_+(a,b)$, y otra negativa, $Q_-(a,b)$.

Siguiendo el método de construcción de estructuras de preferencia recíproca presentado en [4], con cada una de estas relaciones Q_+, Q_- es posible asociar sendas estructuras recíprocas $Q_+ = \langle P, L_+\rangle$ y $Q_- = \langle Z, L_-\rangle$, con P y Z denotando como antes preferencia y aversión estrictas, y donde L_+ y L_- representan respectivamente ausencia de preferencia y ausencia de aversión. Por tanto, para cada par $a,b \in \mathbb{A}$ en cada dimensión pueden expresarse 3 situaciones relevantes desde el punto de vista decisional: preferencia estricta P, su inversa P^{-1} y ausencia de preferencia L_+, en el caso positivo; y aversión estricta Z, su inversa Z^{-1} y ausencia de aversión L_-, en el caso negativo.

La estructura de preferencia del modelo P-A recíproco surge entonces de la combinación conjuntiva de estas situaciones positivas y negativas. Esto es, las relaciones que componen la estructura P-A recíproca, denotada por R_{P-A}, se obtienen mediante la agregación conjuntiva de las relaciones en el cruce $\mathcal{R} = \{P, P^{-1}, L_+\} \times \{Z, Z^{-1}, L_-\}$. Esta conjunción puede ser implementada mediante diferentes operadores T, como t-normas o funciones overlap [5], aunque el uso de operadores no simétricos podría ser conveniente cuando sea necesario considerar diferentes contribuciones de las dimensiones positiva y negativa en este proceso de agregación.

Así pues, esta combinación de las estructuras recíprocas positiva y negativa conduce en principio a la obtención de 9 (3×3) relaciones, que no obstante es preciso refinar en tanto algunos pares de estas 9 relaciones están relacionados mediante la inversión de relaciones (esto es, de modo similar a como, e.g., P se relaciona con P^{-1}). En este sentido, la estructura P-A recíproca R_{P-A} ha de contener solo aquellas relaciones que expresen nociones preferenciales claramente distinguibles. Para formalizar este refinado, considérese la relación binaria $\sim$ sobre $\mathcal{R}$ dada para todo $A, B \in \mathcal{R}$ por

$$A \sim B \iff (A = B) \vee (A = B^{-1}) \tag{1}$$

Cualquier par de elementos de $\mathcal{R}$ relacionados mediante $\sim$ deben estar asociados a una única relación en R_{P-A}. En otras palabras, la estructura P-A recíproca ha de identificarse con el conjunto cociente de $\mathcal{R}$ por $\sim$. Es directo comprobar que $\sim$ constituye una relación de equivalencia sobre $\mathcal{R}$, y forma las siguientes 5 clases de equivalencia: $\{PZ^{-1}, P^{-1}Z\}, \{PL_-, P^{-1}L_-\}, \{L_+Z, L_+Z^{-1}\}$, $\{PZ, P^{-1}Z^{-1}\}$ y $\{L_+L_-\}$, donde e.g. $PZ(a,b) = T(P(a,b), Z(a,b))$, y similarmente para el resto de casos. Por tanto, es clara la validez de la siguiente definición:

$$R_{P-A} = \mathcal{R}/\sim = \langle PZ^{-1}, PL_-, L_+Z, PZ, L_+L_-\rangle \tag{2}$$

donde los representantes elegidos de las 5 clases de equivalencia son: preferencia fuerte PZ^{-1} (su inversa es aversión fuerte, i.e. $(PZ^{-1})^{-1} = P^{-1}Z$), semi-preferencia PL_- (su inversa es semi-preferencia inversa $P^{-1}L_-$), semi-aversión L_+Z (su inversa es semi-aversión inversa L_+Z^{-1}), ambivalencia PZ (su inversa es ambivalencia inversa $P^{-1}Z^{-1}$) y ausencia global de preferencia L_+L_- (simétrica).

IV. Conclusión

La estructura completa P-A recíproca aquí propuesta permite representar, con los mismos requisitos de información que el modelo borroso estándar, un rango de 5 nociones preferenciales, que se extiende desde preferencia fuerte hasta aversión fuerte pasando por semi-preferencia y semi-aversión, con dos estados inconclusivos de ausencia global de preferencia y ambivalencia. En concreto, esta ambivalencia expresa una situación conflictiva en que los aspectos positivos y negativos que surgen de la comparación de un par de alternativas se oponen entre sí, lo que constituye una noción preferencial novedosa en relación con los estados representables por las estructuras recíprocas y el modelo estándar borroso.

References

[1] J. Fodor and M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support. Dordrecht: Kluwer Academic Publishers, 1994.

[2] C. Franco, J. Montero and J.T. Rodríguez, "A fuzzy and bipolar approach to preference modeling with application to need and desire," Fuzzy Sets and Systems, vol. 214 pp. 20–34, 2013.

[3] C. Franco, J.T. Rodríguez and J. Montero, "Building the meaning of preference from logical paired structures," Knowledge-Based Systems, vol. 83, pp. 32–41, 2017.

[4] F. Castiblanco, C. Franco, J.T. Rodríguez and J. Montero, "A characterization of reciprocal fuzzy preference structures and its compatibility with standard fuzzy preference structures," Fuzzy Sets and Systems, vol. 422, pp. 48–67, 2021.

[5] de Miguel, L., Gómez, D., Rodríguez, J.T., Montero, J., Bustince, H., Dimuro, G.P., Sanz, J.A.: General overlap functions. Fuzzy Sets and Systems **372**, 81–96 (2019).

A Multi-Criteria Procedure Using Different Qualitative Scales

José Luis García-Lapresta
PRESAD Research Group
Dep. de Economía Aplicada
Universidad de Valladolid
Valladolid, Spain
lapresta@eco.uva.es

David Pérez-Román
PRESAD Research Group
Dep. de Organización de Empresas y CIM
Universidad de Valladolid
Valladolid, Spain
david@emp.uva.es

***Abstract*—In this contribution, we summarize the multi-criteria decision-making procedure introduced in Applied Soft Computing 106, 107279 (2021). We consider that a group of panelists evaluate several alternatives regarding different criteria, each one through a specific ordered qualitative scale (OQS). These OQSs are equipped with OPMs that collect the perceptions about the proximities between the terms of the scales by means of ordinal degrees of proximity. The weights assigned to the criteria are managed in an ordinal way by replicating the linguistic assessments obtained for each alternative in each criterion as many times as necessary until these replications reflect the proportions among weights. The linguistic assessments provided by panelists are compared with the highest terms of the corresponding OQSs. In order to aggregate the obtained ordinal degrees of proximity, a homogenization process is provided.**

***Index Terms*—multi-criteria decision-making, qualitative scales, ordinal proximity measures**

I. Introduction

We consider that each individual of a group of panelists assigns a linguistic term to every alternative in each criterion. These linguistic terms belong to an *ordered qualitative scale* (OQS) $\mathcal{L} = \{l_1, \ldots, l_g\}$, arranged from worst to best, i.e., $l_1 \prec \cdots \prec l_g$, with $g \geq 3$.

In García-Lapresta and Pérez-Román [3] we introduced the notion of *ordinal proximity measure* (OPM) in order to manage the perceptions between the linguistic terms of non-uniform OQSs[1] in a purely ordinal way.

An OPM is a mapping $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ that assigns an ordinal degree of proximity to each pair of linguistic terms of an OQS $\mathcal{L}$ satisfying four conditions:

1) Exhaustiveness: All the ordinal degrees of Δ should be used at least once.
2) Symmetry: The ordinal degree of proximity between two linguistic terms does not depend on the order of the comparison.
3) Maximum proximity: The maximum degree of proximity is only reached when comparing a linguistic term with itself.
4) Monotonicity: Given three linguistic terms arranged from the lowest to the highest, the ordinal proximity between the first and the second is higher than the ordinal proximity between the first and the third, and the ordinal proximity between the second and the third is higher than the ordinal proximity between the first and the third.

The mentioned ordinal degrees of proximity belong to a linear order $\Delta = \{\delta_1, \ldots, \delta_h\}$, with $\delta_1 \succ \cdots \succ \delta_h$, being δ_1 and δ_h the maximum and minimum degrees of proximity, respectively.

II. The procedure

García-Lapresta et al. [2] consider that a set of m panelists $P = \{p_1, \ldots, p_m\}$ evaluate a set of n alternatives $X = \{x_1, \ldots, x_n\}$ regarding a set of q criteria $C = \{c_1, \ldots, c_q\}$ through q OQSs $\mathcal{L}^k = \{l_1^k, \ldots, l_{g_k}^k\}$, $k = 1, \ldots, q$, equipped with *metrizable* OPMs (see García-Lapresta et al. [1]) $\pi^k : \mathcal{L}^k \times \mathcal{L}^k \longrightarrow \Delta^k$, where $\Delta^k = \{\delta_1^k, \ldots, \delta_{h_k}^k\}$.

Since criteria may be assessed through different OQSs equipped with the corresponding OPMs, a normalization process is needed. Two possible scenarios are possible:

1) If $h_1 = \cdots = h_q = h$, then, $\Delta^* = \Delta^1 = \cdots = \Delta^q = \{\delta_1, \ldots, \delta_h\}$.
2) Otherwise, let $\Delta^* = \{\delta_1^*, \ldots, \delta_{h_*}^*\}$ be the normalized set of ordinal degrees of proximity. Each Δ^k can be embedded into Δ^* through the mapping $\Gamma_k : \Delta^k \longrightarrow \Delta^*$ defined as $\Gamma_k(\delta_r^k) = \delta_{\gamma_k(r)}^*$, with $\gamma_k(r) = 1 + d_k \cdot (r-1)$ such that $\gamma_k(h_k) = h_*$. Thus, $\gamma_k(1) = 1, \gamma_k(2), \ldots, \gamma_k(h_k)$ are in arithmetic progression of difference d_k, $k = 1, \ldots, q$.

The multi-criteria decision-making procedure introduced in García-Lapresta et al. [2] is divided in different steps.

1) The assessments given by the panelists to the alternatives regarding all the criteria are replicated taking into account the importance of each criterion.
2) Then, the ordinal degree of proximity between each linguistic assessment and the highest possible assessment in the corresponding scale is calculated.

[1]The OQS {'poor', 'fair', 'good', 'very good', 'extremely good', 'excellent'} is not uniform if one may think that 'fair' is closer to 'good' than to 'poor'.

3) These ordinal degrees of proximity are normalized following the homogenization process.
4) Then, the alternatives are ranked from the medians of the normalized ordinal degrees of proximity taking into account an appropriate linear order on the set of feasible medians.
5) The procedure ends with a sequential tie-breaking method that provides the final ranking of the alternatives.

In García-Lapresta et al. [2], the mentioned procedure is applied to a real case study for new product development decision-making made from survey data of the Spain's third-largest international exporter in the food and beverage sector.

Acknowledgments: The authors are grateful to the Spanish *Agencia Estatal de Investigación* for its financial support (project PID2021-122506NB-I00).

REFERENCES

[1] García-Lapresta, J.L., González del Pozo, R., Pérez-Román, D.: Metrizable ordinal proximity measures and their aggregation, Information Sciences **448-449**, 149–163 (2018)
[2] García-Lapresta, J.L., Moreno-Albadalejo, P., Pérez-Román, D., Temprano-García, V.: A multi-criteria procedure in new product development using different qualitative scales. Applied Soft Computing **106**, 107279 (2021)
[3] García-Lapresta, J.L., Pérez-Román, D.: Ordinal proximity measures in the context of unbalanced qualitative scales and some applications to consensus and clustering. Applied Soft Computing **35**, 864–872 (2015)

Una versión alternativa del algoritmo de Chi, de aprendizaje de reglas difusas para clasificación, basada en la obtención de múltiples reglas sobre cada ejemplo

Leonardo Jara
Dpto. de Ciencias de la Computación e Inteligencia Artificial, Instituto Andaluz de Investigación (DaSCI) Universidad de Granada
Granada, España
0000-0003-1579-3018

Antonio González
Dpto. de Ciencias de la Computación e Inteligencia Artificial, Instituto Andaluz de Investigación (DaSCI) Universidad de Granada
Granada, España
0000-0001-8889-7593

Raúl Pérez
Dpto. de Ciencias de la Computación e Inteligencia Artificial, Instituto Andaluz de Investigación (DaSCI) Universidad de Granada
Granada, España
0000-0002-1355-1122

***Resumen*—El algoritmo de aprendizaje de reglas difusas para clasificación Chi ha sido frecuentemente utilizado por su simplicidad y eficiencia. Sin embargo este algoritmo puede tener problemas en algunos casos, al no cubrir con las reglas obtenidas suficientemente al conjunto de ejemplos. En este trabajo planteamos una modificación de dicho algoritmo basada en la idea de añadir más de una regla por cada ejemplo. Se proponen tres métodos alternativos y se experimenta sobre distintos conjuntos de ejemplos para estudiar en que casos es mejor la propuesta presentada.**

***Palabras Clave*—Reglas difusas, Clasificación, Aprendizaje, Inferencia difusa rápida.**

I. INTRODUCCION

Uno de los algoritmos de aprendizaje de reglas difusas más referenciado en la literatura es el algoritmo de Wang y Mendel [1], y su versión para clasificación conocida como algoritmo de Chi [2]. Este fue uno de los primeros algoritmos de aprendizaje de reglas difusas, y aunque no siempre es el mejor de todos los propuestos, es frecuentemente utilizado por su simplicidad y rapidez en la obtención de un conjunto de reglas difusas.

El algoritmo de Chi original propone sólo una regla para cada uno de los ejemplos de su conjunto de entrenamiento, aunque no todas estas reglas formarán parte del conjunto final de reglas, ya que se aplica una resolución de conflictos en la etapa final. En este trabajo, proponemos mejorar la precisión del algoritmo de Chi cambiando esta idea básica de proponer una regla para cada ejemplo, por una nueva, que consiste en proponer más de una regla difusa para cada ejemplo. La intención es mejorar la capacidad de precisión al obtener reglas difusas que cubren mejor el conjunto de ejemplos. Sin embargo este modelo puede provocar un aumento del número final de reglas que podría afectar al tiempo requerido para realizar la inferencia. En [3] se propone un modelo de inferencia eficiente, alternativo a la implementación tradicional de inferencia de reglas difusas, cuyo tiempo de razonamiento no depende del número de reglas. Así, utilizando este nuevo algoritmo de inferencia, la propuesta presentada en este trabajo podría ser una buena alternativa de mejora del algoritmo original para los casos estudiados en la parte experimental.

En la siguiente sección explicamos la estructura básica del algoritmo de Chi y los métodos propuestos para la generación de múltiples reglas sobre cada ejemplo, finalmente en la sección III se muestra la experimentación realizada con las conclusiones.

Trabajo cofinanciado por el Programa Nacional de Becas de Postgrado en el Extranjero "Don Carlos Antonio López, BECAL", otorgado por el Gobierno de la República del Paraguay y los proyectos RTI2018-098460-B-I00 y B-TIC-668-UGR20, incluyendo Fondos Europeos de Desarrollo Regional.

II. MÉTODOS PROPUESTOS

El algoritmo de Chi se propuso como una adaptación del algoritmo de Wang y Mendel [1] a problemas de clasificación. El esquema general de este algoritmo es el siguiente:

1. Suponemos definidos los dominios difusos de todas las variables lingüísticas.
2. A cada ejemplo se le asocia la regla difusa con la mejor adaptación a dicho ejemplo.
3. Se asigna un peso a cada regla.
4. Se selecciona el conjunto final de reglas, haciendo uso de los pesos y eliminando los posibles conflictos.

En el paso 2 del algoritmo de Chi, se asocia una regla a cada ejemplo e, en concreto se asocia la regla que mejor adaptación tenga con el ejemplo. Esta regla, llamada regla central al ejemplo en [3], es el punto central del conjunto de reglas que tiene adaptación con este ejemplo. Esto quiere decir que dicho conjunto de reglas, que llamaremos $N(e)$, incluye todas las reglas con adaptación mayor a cero.

Nuestra idea es seleccionar para cada ejemplo no una, sino varias reglas del conjunto $N(e)$. Sin embargo, no es conveniente tomar todas las reglas de este conjunto (o al menos

no en todos los casos) ya que el número de reglas obtenidas en $N(e)$ es 2^p siendo p el número de variables antecedentes con universo continuo y dominio difuso asociado, que puede ser muy grande en algunos casos.

Siguiendo la idea del algoritmo original, consideraremos que las reglas más relevantes para un ejemplo son las que tienen los valores más altos de adaptación con él. Así, si consideramos que $N(e)$ es un conjunto ordenado por la función de adaptación, las mejores reglas candidatas son las que aparecen en las primeras posiciones de ese conjunto ordenado. De esta forma es necesario ordenar el conjunto $N(e)$.

Pero obtener esta ordenación requiere un algoritmo poco eficiente, ya que en primer lugar es difícil generar todas las reglas, y en segundo lugar es aún más difícil ordenar el conjunto, en el caso de que el número de variables antecedentes con referencial continuo sea muy grande.

En lugar de generar completamente el conjunto $N(e)$ y luego ordenarlo en función del grado de adaptación con el ejemplo, utilizamos una estrategia de exploración de reglas (recursiva y basada en la idea de búsqueda en profundidad) partiendo de la regla central. Esta estrategia puede considerarse una buena heurística para recorrer las reglas en orden decreciente del grado de adaptación. El proceso que seguimos aquí es similar al propuesto en [3] para procesos de inferencia recursiva. A continuación exponemos los métodos propuestos de asignación de múltiples reglas para cada ejemplo.

- ***Ochi(k), Método basado en el orden***: Un modelo alternativo al algoritmo de Chi consiste en asociar a cada ejemplo las k primeras reglas del conjunto $N(e)$, siendo k un parámetro y usando la heurística anterior como método de selección de las reglas en N(e). Esta nueva versión se denomina $OChi(k)$. Así, si tomamos k=1, el modelo asigna al ejemplo de la regla con mejor adaptación, que es precisamente la regla central, y así se reproduce el algoritmo de Chi.
- ***UChi(t), Método basado en el umbral***: La estrategia cambia en este método con respecto al anterior, el orden no se realiza por las k primeras reglas del conjunto, realizamos un filtrado en $N(e)$ de todas las reglas que tenga un grado de adaptación superior a un umbral t, incluyendo a todos los antecedentes de la regla. Esta nueva versión se denomina $UChi(t)$.
- ***DChi(d), Método basado en la distancia*** : El método $DChi(d)$ comienza a explorar la regla central del ejemplo y a partir de ella explora las reglas que hemos ordenado heurísticamente de mejor a peor, ahora en lugar de explorar todas las reglas del conjunto N(e), la búsqueda se limita a una vecindad cercana a la regla central, y esa proximidad estará determinada por la distancia de Hamming y un parámetro d. Este parámetro mide la distancia entre dos cadenas de igual longitud a través del número de posiciones en las que los símbolos correspondientes son diferentes. Por lo tanto, la distancia entre dos reglas se define como el número de variables antecedentes que tienen una asignación diferente. Es decir que, si dos reglas difieren en la asignación a la misma variable antecedente, podemos decir que poseen una d = 1, si estas reglas difieren en 2 variables decimos que tenemos una d = 2 y así sucesivamente. En el caso de que ambas reglas sean iguales tenemos que d = 0.

III. ESTUDIO EXPERIMENTAL Y CONCLUSIONES

Analizamos el comportamiento del algoritmo original de Chi con los 3 métodos propuestos. Utilizamos 35 conjuntos de ejemplos, todos ellos extraídos del repositorio de aprendizaje automático UCI [4]. Estos conjuntos de datos también representan un amplio espectro de diferentes situaciones a considerar en el aprendizaje automático, tales como problemas binarios, problemas multiclase con un gran número de clases, problemas con pocas o muchas variables, problemas para los que se obtiene un número pequeño o grande de reglas, problemas con pocos o muchos ejemplos.

En el experimento estudiamos cuatro parámetros: la precisión en el conjunto de pruebas, el tiempo necesario para el aprendizaje, el tiempo necesario para realizar la inferencia en el conjunto de reglas obtenido y el porcentaje de ejemplos del conjunto de pruebas que no activan ninguna regla, utilizando el algoritmo de inferencia basado en la recursión. Seleccionamos un amplio rango de variación de los parámetros (k, t, d), con el fin de encontrar las combinaciones mas eficientes en cuanto a la precisión y el tiempo de aprendizaje en los métodos desarrollados.

La implementación de los métodos genera una mejora en la precisión de las bases de datos, obteniendo métricas del 5 % en promedio sobre todas las bases de datos analizadas con el método de $OChi(k)$, 5.7 % con el método $UChi(t)$ y de 7 % con el método $DChi(d)$. También podemos destacar que a medida que el porcentaje de ejemplos no cubiertos es mayor en el algoritmo original de Chi, la capacidad de mejorar la precisión de los métodos aumenta, obteniendo casos en los que una base de datos puede mejorar su precisión en un 18 % con respecto al Chi original.

Otro punto importante para destacar es que al aumentar el número de reglas de forma considerable con respecto al algoritmo de Chi, el tiempo de inferencia se mantiene constante debido a la aplicación del modelo de inferencia eficiente [3]. Este modelo obtiene buenos resultados cuando el número de reglas es elevado, lo que hace que sea apropiado para la aplicación de estos métodos que expanden el conjunto de reglas.

REFERENCIAS

[1] L.-X. Wang and J. Mendel, "Generating fuzzy rules by learning from examples," IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, no. 6, pp. 1414–1427, 1992.

[2] Z. Chi, H. Yan, and T. Pham, Fuzzy algorithms: with applications to image processing and pattern recognition. World Scientific, 1996, vol. 10.

[3] L. Jara, R. Ariza-Valderrama, J. Fernández-Olivares, A. González, and R. Pérez, "Efficient inference models for classification problems with a high number of fuzzy rules," Applied Soft Computing, vol. 115, p. 108164, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494621010231

[4] K. Bache and M. Lichman, "Uci machine learning repository," 2013.

Representación en 2-Tuplas para el aumento artificial del conocimiento en la evaluación del discurso de odio

Andrés Montoro
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
Andres.Montoro@alu.uclm.es

José A. Olivas
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
JoseAngel.Olivas@uclm.es

Enrique Herrera-Viedma
Ciencias de la Computación e IA
Universidad de Granada
Granada, España
viedma@ugr.es

Adán Nieto
Derecho Público y de la Empresa
Universidad de Castilla-La Mancha
Ciudad Real, España
Adan.Nieto@uclm.es

Francisco P. Romero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
FranciscoP.Romero@uclm.es

Jesús Serrano-Guerrero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
Jesus.Serrano@uclm.es

***Resumen*—El discurso de odio se ha convertido en un mal endémico en nuestra sociedad y las redes sociales no han hecho más que universalizar este fenómeno. La detección y evaluación de este tipo de comportamientos se ha convertido en una prioridad para gobiernos y empresas de medios sociales. Presentamos una propuesta para evaluar el discurso de odio según su intensidad apoyados en un método para el aumento artificial de la base de reglas basado en la representación en 2-Tuplas con el objetivo de minimizar la pérdida de conocimiento. La evaluación del modelo se afronta como trabajo futuro al igual que la mejora de la base de conocimiento y la optimización de la generación automática de reglas.**

***Palabras clave*—discurso de odio, taxonomía, 2-Tuplas, aumento de reglas.**

I. Introducción

Las redes sociales se han convertido en un nuevo espacio de oportunidad delictiva y han dado lugar a una mayor difusión y alcance de los contenidos de odio. El discurso y los delitos de odio envenenan las sociedades, reforzando las tensiones entre los distintos grupos sociales, perturbando la paz y el orden público, lo que pone en peligro la convivencia pacífica. La detección y evaluación de este tipo de comportamientos es extremadamente compleja, e incluso los expertos juristas no se ponen de acuerdo en cómo medir la intensidad o gravedad de este tipo de comportamientos potencialmente delictivos. Debido a la ambigüedad interpretativa, la imprecisión y la incertidumbre se hace necesario el uso de técnicas que sean tolerables a estas condiciones propias del dominio. Se propone un modelo computacional para la evaluación del discurso de odio según su intensidad apoyado en un método de aumento de datos basado en la representación en 2-Tuplas para paliar el problema de las limitaciones del conocimiento experto.

I-A. Antecedentes

Algunos trabajos sobre la detección de odio están basados en el contenido de los mensaje como en [1]. También es lógico asumir que para la detección del odio una herramienta eficaz es el análisis de sentimientos [2]. En los último años, la tendencia va hacia modelos basados en el aprendizaje profundo como en [3]. En todos estos trabajos predomina una característica común, se centran en detectar el discurso de odio empleando técnicas de aprendizaje automático y no tratan de entender el fenómeno de odio.

II. Taxonomía del odio

La evaluación del discurso de violento y de odio se torna un dominio adecuado para emplear la Computación con Palabras propuesta por Zadeh en [4] como motor de razonamiento, para acercarse más a la forma en la que razonaríamos los humanos. Para afrontar el problema es esencial el conocimiento experto humano y bibliográfico para comprender las particularidades del discurso de odio, para ello se ha contado con el apoyo pleno del Instituto de Derecho Penal Europeo e Internacional de la Universidad de Castilla-La Mancha. También se han analizado las legislaciones más importantes en materia anti-odio. Como resultado, ese conocimiento se ha plasmado en forma de taxonomía. En la Figura 1 se puede ver una representación parcial de la taxonomía.

A partir de la taxonomía se ha diseñado una base de conocimiento que permite razonar sobre los conceptos que modela haciendo uso de lógica borrosa. La base de conocimiento no es suficientemente completa debido a la escasez de conocimiento experto y a que la base de reglas se ha construido en base a eventos frecuentes extraídos a partir conocimiento adquirido.

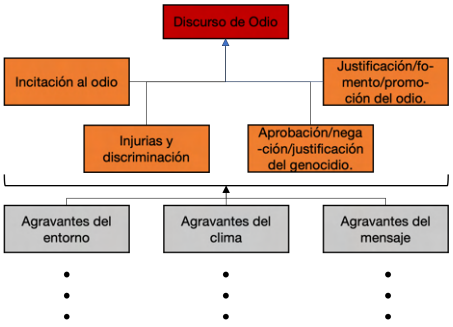


Figura 1. Núcleo y representación parcial de los agravantes de la taxonomía para la comunicación violenta y de odio.

III. Aumento del conocimiento basado en la representación en 2-Tuplas

Se propone una primera aproximación para la modificación de la base de reglas empleando el aumento de éstas de forma artificial a partir de la base de reglas existente, lo que entraña un riesgo para la confianza y fiabilidad del sistema ya que el conocimiento representado es causal. La representación en 2-Tuplas propuesta en [5] puede ser usada como una posible aproximación. Se basa en un modelo de computación simbólica y para paliar la pérdida de conocimiento producido al combinar etiquetas lingüísticas emplean lo que denominan traslación simbólica, que consiste en un valor $alpha$ en el intervalo [-0.5,0.5) que permite mostrar la información perdida en un rango perteneciente a la etiqueta lingüística. En la traslación simbólica se encuentra la clave, ya que cualquier etiqueta lingüística que se mueva en ese intervalo una vez normalizada la función de pertenencia entre 0 y $|S|-1$ siendo $|S|$ la cardinalidad del conjunto borroso, seguirá teniendo la misma etiqueta lingüística con distintos valores de $alpha$. Esto permite lograr un conjunto de reglas más amplio alterando mínimamente el conocimiento. Las fases para conseguir el aumento artificial de reglas basado en la representación en 2-Tuplas de los conjuntos borrosos son las siguientes:

1. Normalización de los conjuntos borrosos en el intervalo $[0, g]$ donde $g = |S| - 1$.
2. Seleccionar la granularidad deseada para el desplazamiento de la etiqueta lingüística representada como número borroso.
3. Establecer el tamaño de la ventana deslizante, que junto con la granularidad no podrá superar un desplazamiento que sobrepase el límite [-0.5,0.5).
4. Recorrer cada una de las reglas de la base de conocimiento aplicando el aumento de datos desplazando cada una de las etiquetas a izquierda y derecha según la granularidad y el tamaño de la ventana. Como el objetivo es evitar la pérdida de conocimiento, el aumento artificial solo se aplica al antecedente, el consecuente permanece inmutable. En la Tabla I se muestra un ejemplo sintético.

Como apunte adicional, en la práctica no es adecuado generar todas las combinaciones posibles porque se perdería la borrosidad del sistema, sobre todo si se considera una granularidad muy pequeña y se cubre todo el espacio en el intervalo [-0.5,0.5). En la Figura 2 se muestra un ejemplo.

Tabla I
Ejemplo del aumento artificial de una regla borrosa representada mediante 2-Tuplas con una granularidad de 0.1 y el tamaño de ventana [-0.3,0.3].

Reglas	Antecedente	Consecuente
Regla original	(Alto, 0) y (Medio, 0)	(Medio, 0)
Regla aumentada 1	(Alto, -0.1) y (Medio, 0)	(Medio, 0)
Regla aumentada 2	(Alto, 0.1) y (Medio, 0.1)	(Medio, 0)
Regla aumentada 3	(Alto, -0.2) y (Medio, 0.1)	(Medio, 0)
Regla aumentada n	(Alto, -0.3) y (Medio, 0.3)	(Medio, 0)

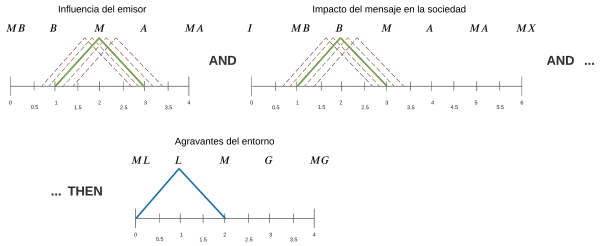


Figura 2. Ejemplo del proceso de aumento artificial de una regla borrosa.

IV. Conclusión

Las redes sociales han universalizado el discurso de odio lo que ha llevado a una lucha activa por parte de los gobiernos y las empresas de Internet contra este. La gran mayoría de las investigaciones y las herramientas existentes para combatir el odio se centra en la detección y no estudian la potencialidad del mismo. En este caso, se ha descrito un modelo para la evaluación de la intensidad del discurso de odio en redes sociales y el aumento artificial de la base de conocimiento empleando la representación 2-Tuplas con el fin de corregir la falta de conocimiento experto. Algunos aspectos a considerar es el incremento del número de expertos para la mejora del conocimiento, la evaluación exhaustiva de éste y por último, la optimización de la generación artificial de reglas para evitar la explosión combinatoria y la consiguiente redundancia.

Referencias

[1] Z. Zhang, y L. Luo, Hate Speech Detection: A Solved Problem? The Challenging Case of Long Tail on Twitter, Semantic Web, vol. 10, pp. 925–945, 2018.

[2] P. Burnap, O. F. Rana, N. Avis, M. Williams, W. Housley, A. Edwards, J. Morgan, y L. Sloan, Detecting tension in online communities with computational Twitter analysis, Technological Forecasting and Social Change, vol. 95, pp 96–108, 2015.

[3] J. Chen, S. Yan, y K. C. Wong, Verbal aggression detection on Twitter comments: convolutional neural network for short-text sentiment analysis, Neural Computing and Applications, vol. 32, pp. 10809–10818, 2020.

[4] L. A. Zadeh, Fuzzy Logic = Computing with Words, IEEE Transactions on Fuzzy Systems, vol. 4, pp. 103–111, 1996.

[5] F. Herrera, y L. Martínez,A 2-Tuple Fuzzy Linguistic Representation Model for Computing with Words, IEEE Transactions on Fuzzy Systems, vol. 8, pp. 746–752, 2000.

An Optimal Best-Worst Prioritization Method under a 2-tuple Linguistic Environment in Decision Making

1st Álvaro Labella
Computer Science Department
University of Jaén
Jaén, Spain
alabella@ujaen.es

2nd Bapi Dutta
The Logistics Institute-Asia Pacific
National University of Singapore
Singapore, Singapore
bapi.iitr@gmail.com

3rd Luis Martínez
Computer Science Department
University of Jaén
Jaén, Spain
martin@ujaen.es

***Abstract*—This is a summary of our article published in Computers & Industrial Engineering [1] to be considered as part of the ESTYLF 2022 key works.**

***Index Terms*—Best-Worst method, 2-tuple linguistic model, Multi-criteria group decision-making**

I. Introduction

Multi-criteria decision-making (MCDM) refers to the evaluation of several alternatives $\{x_1, \ldots, x_m\}$ according to different and usually conflicting criteria $\{C_1, \ldots, C_n\}$ by an expert. The experts' preferences elicitation is key in the resolution of any MCDM problem and pairwise comparison method [2] is one of the most widely used techniques to carry out this step, which consists of comparing pairs of alternatives. However, the greater the number of comparisons, the higher the likelihood that experts provide inconsistent opinions and consequently unreliable results. To overcome this limitation, Rezaei proposed the Best-Worst Method (BWM) [3], in which the expert compares only the best and worst element of the problem with the remaining elements in order to obtain their importance/weights. The BWM proposed by Rezaei considers pairwise comparison assessments represented by crisp numbers, but real-world MCDM problems present vague and uncertain information that cannot be modeled by discrete values. Under these circumstances, linguistic information has been used successfully to model uncertainty in MCDM. In spite of there are many approaches to manage linguistic information, the 2-tuple linguistic representation model [4] stands out because of its simplicity, interpretability, and capability to carry out precise linguistic computations following a Computing with Words (CW) scheme i.e., obtain linguistic results from linguistic premises. For this reason, this contribution consists of a new extension of the BWM for linguistic information based on the 2-tuple linguistic model in which experts provide their preferences by using linguistic pairwise comparisons and whose results are represented by 2-tuple linguistic values modelled in the same expression domain that experts have used to provide their preferences.

II. Preliminaries

The main concepts related to the proposal are briefly explained below. Firstly, the BWM is a MCDM method that derives the importance of the criteria by means of the following steps:

- **Step 1**: Determine a set of decision criteria $\{C_1, \ldots, C_n\}$.
- **Step 2**: Select the best criterion C_B and the worst criterion C_W.
- **Step 3**: Make pairwise comparison among C_B and the rest of the criteria, by obtaining the Best to Others (BO) vector, $BO = \{a_{B1}, a_{B2}, \ldots a_{Bn}\}$.
- **Step 4**: Make pairwise comparison among C_W and the rest of the criteria, by obtaining the Worst to Others (WO) vector, $WO = \{a_{1W}, a_{2W}, \ldots a_{nW}\}$.
- **Step 5**: Compute the weights of criteria by optimization models. Initially, Rezaei proposed the following one:

$$\min\ \xi \quad \text{s.t.} \begin{cases} |\frac{w_B}{w_j} - a_{Bj}| \leq \xi \\ |\frac{w_j}{w_W} - a_{jW}| \leq \xi \\ \sum_{j=1}^{n} w_j = 1, \\ w_j \geq 0, \text{ for all } j = 1, 2, ..., n. \end{cases} \tag{1}$$

On the other hand, the 2-tuple linguistic model represents the information by means of a 2-tuple, (s_i, α), where s_i is a linguistic label belonging to a predefined linguistic term set $S = \{s_1, s_2 \ldots s_g\}$ and $\alpha \in [-0.5, 0.5)$ a numerical value that represents the translation of the fuzzy membership function that represents the closest term s_i. The value of α is defined as:

$$\alpha = \begin{cases} [-0.5, 0.5) & if\ s_i \in \{s_1, s_2, \ldots, s_{g-1}\} \\ [0, 0.5) & if\ s_i = s_0 \\ [-0.5, 0] & if\ s_i = s_g \end{cases}$$

Remark 1: Note that the transformation of a linguistic term $s_i \in S$ into a 2-tuple linguistic value in $\overline{S}$ is:

$$s_i \in S \rightarrow (s_i, 0) \in \overline{S}$$

The 2-tuple linguistic model was extended in order to interpret values from the scale $[0, 1]$ rather than the conventional scale $[0, g]$. In such cases, any value of $\beta \in [0, 1]$ can be mapped to a 2-tuple linguistic value [5]:

$$\Delta_1(\beta) = (s_i, \alpha) \text{ where } \begin{cases} s_i & i = round(\beta \cdot g) \\ \alpha = \beta - \frac{i}{g} & \alpha \in [-\frac{1}{2g}, \frac{1}{2g}) \end{cases} \tag{2}$$

On the other hand, any 2-tuple linguistic value could be translated into an equivalent numerical value $\beta \in [0, 1]$ [5]:

$$\Delta_1^{-1}(s_i, \alpha) = \beta = \frac{i}{g} + \alpha. \tag{3}$$

III. Linguistic Best-Worst method based on 2-tuple model

In our proposal the expert uses a fuzzy linguistic scale to provide the BWM comparisons $S^{BWM} = \{ l_1 =$ Equally Important (EI), $l_2 =$ Weekly Important (WI), $l_3 =$ Fairly Important (FI), $l_4 =$ Very Important (VI), $l_5 =$ Absolutely Important (AI) $\}$. Therefore, from the linguistic term set, S^{BWM}, the expert first compares the best object x_B and the worst object x_W with the rest of the objects and expresses his/her assessments via 2-tuple linguistic values. The expert's preference information can be summarized in a Linguistic Best-Worst Pairwise Comparison Matrix (LBWM):

$$LBWM = \begin{array}{c} \\ c_B \\ c_W \end{array} \begin{array}{c} \begin{array}{ccc} c_1 & \cdots & c_n \end{array} \\ \begin{pmatrix} (d_{B1}, \alpha_{B1}) & \cdots & (d_{Bn}, \alpha_{Bn}) \\ (d_{1W}, \alpha_{1W}) & \cdots & (d_{nW}, \alpha_{nW}) \end{pmatrix} \end{array}$$

where (d_{Bj}, α_{Bj}) is the linguistic preference of the expert against the comparison of the x_B with x_j and (d_{jW}, α_{jW}) denotes the 2-tuple linguistic preference for the comparison of x_W with x_j.

After the expert has completed the preference elicitation tasks by providing LBWM, we attempt to find the priority weights optimally. To do so, we propose the following optimization model:

$$(\mathbf{M-1}) \quad \min \ \xi \quad \text{s.t.} \ \begin{cases} \sum_{j=1}^{n} w_j = 1, \\ |\frac{w_B}{w_j} - \Delta^{-1}((d_{Bj}, \alpha_{Bj}))| \leq \xi, \\ |\frac{w_j}{w_W} - \Delta^{-1}((d_{jW}, \alpha_{jW}))| \leq \xi, \\ w_j \geq 0, \text{ for all } j = 1, 2, ..., n. \end{cases} \tag{4}$$

Solving the optimization model **(M-1)**, we can obtain the optimal numeric priority weights $w^* = (w_1^*, ..., w_n^*)$. To represent the weights linguistically, we introduce a linguistic term set, $S_1 = \{s_0 =$ Very Unimportant (VU), $s_1 =$ Unimportant (U), $s_2^1 =$ Fair (F), $s_3^1 =$ Important (I), $s_4^1 =$ Very Important (VI)$\}$ and maps it to $[0, 1]$ via the Eq, 3. By using Eq. 2 the linguistic optimal priority weights $(l_{w_1^*}, l_{w_2^*}, ..., l_{w_n^*})$ are derived as follows:

$$l_{w^*} = \Delta_1(w*) = (\Delta_1(w_1^*), \Delta_1(w_2^*), ..., \Delta_1(w_n*)) \tag{5}$$

IV. Illustrative example

Let us suppose a travel agency wants to acquire a set of cars for their business. After initial screening they have shortlisted four cars, say $X = \{x_1, x_2, x_3, x_4\}$, as a possible options for the acquisition, which are evaluated over a set of four criteria $C = \{C_1 : Style, C_2 : Power, C_3 : Price, C_4 : Color\}$.

To prioritize the criteria according to the linguistic BWM, the expert is required to identify the best and worst criteria from C considering the company's key needs. Based on such needs, the expert found that the best criterion is C_2 and C_4 is the worst one. Once the best and worst criteria have been identified, the expert makes pairwise comparisons among C_B and C_W with the rest of the criteria. The comparisons are collected in a LBWM:

$$LBWM = \begin{array}{c} \\ c_1 \\ c_4 \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \\ \begin{pmatrix} (EI, 0) & (WI, 0) & (FI, 0) & (AI, 0) \\ (AI, 0) & (FI, 0) & (WI, 0) & (EI, 0) \end{pmatrix} \end{array}$$

Now, we employ the model **(M-1)** to derive the priority weights from L and obtain 2-tuple linguistic priority weights $l_w^* = ((U, 0.01), (U, -0.08), (F, -0.02), (VU, 0.09))$. The 2-tuple linguistic weight is easily understandable to the expert as a linguistic description is much easier to interpret than a numeric one.

V. Conclusions

This paper has proposed an extension on the BWM in a linguistic 2-tuple environment which allows obtaining accuracy and understandable results represented by linguistic 2-tuples values.

As future works, the proposal may be extended to deal with large-scale group decision-making problems and face challenges related to them, such as the polarization of opinions.

Acknowledgements

This work is partially supported by the Spanish Ministry of Economy and Competitiveness through the Spanish National Project PGC2018-099402-B-I00, ERDF and by the Junta de Andalucía, Andalusian Plan for Research, Development, and Innovation (POSTDOC 21-00461).

References

[1] Á. Labella, B. Dutta, and L. Martínez, "An optimal best-worst prioritization method under a 2-tuple linguistic environment in decision making," *Computers & Industrial Engineering*, vol. 155, p. 107141, 2021.

[2] L. L. Thurstone, "A law of comparative judgment." *Psychological review*, vol. 34, no. 4, p. 273, 1927.

[3] J. Rezaei, "Best-worst multi-criteria decision-making method," *Omega*, vol. 53, pp. 49–57, jun 2015.

[4] L. Martínez, R. M. Rodríguez, and F. Herrera, *The 2-tuple Linguistic Model*. Springer International Publishing, 2015.

[5] W. Tai and C. Chen, "A new evaluation model for intellectual capital based on computing with linguistic variable," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3483–3488, 2009.

On the identification of tricky profiles of rankings

Noelia Rico, Camino R. Vela, Irene Díaz
Department of Computer Science, University of Oviedo, Spain
Emails: {noeliarico,crvela,sirene}@uniovi.es

***Abstract*—The runtime it takes to solve the ranking aggregation problem using the Kemeny method grows factorially with the number of alternatives to rank, making it unsuitable for large numbers. Moreover, experimental results have shown that this time can also vary greatly even for profiles of rankings that have the same number of alternatives and voters. In this case, we say that there are some profiles that are more *difficult* to solve, meaning that they require more runtime to find a consensus.**

The aim of the work is to train machine learning models to predict whether a profile is difficult to solve. To do so, first of all, it is necessary to provide some indices to identify different aspects of the profile. Once this is done, these indices can be used to build a dataset, being applied to different profiles of rankings and providing the data obtained as input of machine learning models. These models are trained to forecast how difficult it is to reach the Kemeny ranking in terms of runtime. By using explainable machine larning algorithms, once the models have been trained, the aspects of the profile of rankings that have the most influence in the runtime can be extracted.

***Index Terms*—Ranking aggregation, preferences, uncertainty, indices**

I. Aim of the work

This communication presents a summary of [1], which is a work that emerges from the observation that profiles of rankings with the same number of alternatives and voters sometimes have very different runtime. This is not surprise, as exact algorithms use domain information to discard rankings as possible solution. For this reason, the amount of rankings that can be discarded, which varies for each profile and cannot be known beforehand, impact on the runtime required to find the Kemeny ranking.

Although some authors have previously pointed out that the number of alternatives is not the only factor that determine the runtime required by different algorithms [2], it is not trivial to determine the factors that affect this runtime. The work starts from the premise that there should be some measurable characteristics of the profile of rankings that reflect how difficult is to solve the profile. It seems intuitive that some aspects as, for example, the level of disagreement between the voters in the profile of rankings, should have some impact in this difficulty. Moreover, in order to do fair comparisons between algorithms, the identification of these aspects is a crucial point, and there should be indices to compare different characteristics of the profiles used for evaluating algorithms to find the Kemeny ranking.

In the paper, indices to this aim are reviewed and proposed in order to characterize aspects of the profile of rankings. Then, a dataset is built applying these indices to profiles of rankings for which the runtime is known. This dataset is later feed to supervised machine learning models that, using the indices as input, aim to classify the profiles of rankings in different categories attending to their runtime. The construction of these models allow also to check which of the characteristics are more relevant to predict whether a given a profile is more difficult to solve.

II. Indices for characterizing the profile

We introduce the indices used in [1] in a previous work [3], in which their individual behavior in relation to the runtime is also reviewed. Some of the indices are based in the *outranking matrix* $\mathbf{O}$, which is a square matrix with dimensions equal to the number of alternatives that summarizes in each cell the number of times that the alternative in the row is preferred over the alternative in the column in the profile of rankings. The different indices are listed below, grouped by the characteristics they attend they can be separated in the following categories:

- μ_1, based on Kendall distance [4], computing the averaged Kendall distance between every pair of rankings in the profile.
- Based on the range the different positions an alternative takes in a profile, considering minimum μ_2, maximum μ_3 and average μ_4.
- μ_5, bound to the optimal cost using $\mathbf{O}$.
- μ_6, the number of dirty triplets (i.e. intransitive cycles of length three).
- Based on the margin of the voters preferring one alternative over another one, being a margin an element of the matrix $\mathbf{O}$: the most frequent margin μ_7, margin diversity μ_8 or the frequency of the minimum margin μ_9.
- μ_{10} distance to the Borda ranking [5] to the profile.
- Based on the row sum of $\mathbf{O}$, the number of a priori preferred alternatives μ_{11} and the standard deviation of the vector containing all the row sums μ_{12}.
- μ_{13} to μ_{16}, based on the distribution of preferred alternatives, that are studied in terms of the vector $\vec{\beta} = (\beta_1, \ldots, \beta_n)$, with $\beta_i = \sum_{j=1}^{n} b_{ij}$ and $b_{ij} = 1$ if $o_{ij} > o_{ji}$ and 0 otherwise. More concretely, standard deviation μ_{13}, mean μ_{14}, median μ_{15}, and difference between the last two μ_{16}.

Moreover, the range of all the indices is studied so the values can be normalized for profiles of rankings with different number of voters and alternatives.

III. Machine learning to predict difficulty

A dataset is built including a total of 8400 profiles of rankings with 8, 9 and 10 alternatives and a wide different number of voters. Each instance corresponds to a profile of rankings and it is described by all the indices detailed above and also their runtime, which is computed using the exact search algorithm proposed in [6].

Using this dataset, the aim is to train a supervised machine learning model to predict the runtime required to find the Kemeny ranking for a profile. In the work, models are trained for different variations of this problem. The models must take as input data the values of the indices obtained for the profiles. As some of the indexes take values that depend on the number of alternatives, all the bounds of the indices have been studied and then used to normalize the input data. Also, the runtime variable has been discretized in a different way to each of the problems considered, thus presenting classification instead of a regression problems. Different models are built to solve the following problems:

Problem 1: **Identification of profiles with an outlier runtime** i.e. profiles with highly anomaly time runtime. Trained to two scenarios:

a) Profiles to predict have the same number of alternatives than the one used to train.

b) Profiles to predict have a different number of alternatives than the one used to test.

Problem 2: **Identification of fast profiles**, in profiles of rankings with larger number of alternatives than the ones used to train the model.

For Problem 1, the runtime has been discretized labelling, for each number of alternatives, those with an runtime above $IQ \times Q3 + 1.5$ as *outlier*. Binary classification models are then trained in order to determine whether the profile is an outlier. In the first version (Problem 1.a), actually four different subproblems can be found: on the one hand, fixing the number of alternatives (for example, only profiles of 8 alternatives are used for training and for testing) for the three possible number of alternatives (8, 9 and 10); on the other hand, profiles of 8, 9 and 10 alternatives are considered both in the training process and test evaluation at the same time. In the second version of the first problem (1.b), profiles with 8 and 9 alternatives have been used for the training process and profiles with 10 alternatives have been used for the evaluation.

We consider important this differentiation, as the second way provides a model that can be used as a tool to estimate runtimes for larger profiles of rankings, and consequently to decide in advance whether the problem is addressable or not.

In Problem 2, the runtime is again discretized but in this case labelling the profiles with runtime in the first quartile for each number of alternatives as *fast*. Models are then trained using the same setting than in Problem 1.b, this is, the models are trained with profiles of rankings that have a number of alternatives lower than the number of alternatives used for testing. Again, by doing so, this model can be used to determine in advance whether the computation time required by the algorithm to find the Kemeny ranking is assumable.

Due to how the label class is discretized, all the models present unbalanced data of the target class, so they have been trained using appropriate resampling techniques for dealing with this. The selected machine learning algorithms used to learn the models are decision trees and random forest, because of their explainability.

IV. Results and conclusion

The best results for all the problems are obtained using the random forest algorithm for training the models, obtaining values for the evaluation metrics Sensitivy and the AUC above 0.8. Random forest algorithm makes possible to obtain the importance of each index, which is computed in terms of the impurity of the dataset that remains after being split using the index.

For Problem 1.a, μ_{11} is found the most important index. This is expected as it has a high impact on the number of solutions that can be discarded from the exploration. Also, μ_{12} and μ_{13} are elected as important variables to predict the outcome, which are indices related to the distribution of the votes, which highlights the initial hypothesis that profiles of rankings with the same numbers of voters can behave very different depending on how the voters distributed they votes. The average Kendall distance μ_1 and the distance of the Borda ranking are also listed in the top 5 alternatives. Same alternatives remain the most important for subproblems when models are trained fixing the number of alternatives. For Problem 1.b, the top five features are a permutation of the ones obtained in 1.a, thus reinforcing their importance.

For Problem 2, the 5 most important indices to predict the output using this model are μ_{11}, μ_{14}, μ_8, μ_{10} and μ_{12}. The indexes μ_8 and μ_{12} which are the number of different margins and the distance to the distribution the ranking would have without cycles in the preferences, which resembles the impact of the distribution of the votes in the runtime.

To the best of our knowledge the paper presents for the first time a machine learning-based approach to predict the difficulty of finding the Kemeny ranking from a given profile of rankings. Using this models we also provide a mechanism for the parametrization of the runtime for new algorithms based on the characteristics of the profile.

References

[1] N. Rico, C. R. Vela, and I. Díaz, "Runtime bounds prediction for the kemeny problem," *Journal of Ambient Intelligence and Humanized Computing*, May 2022. [Online]. Available: https://doi.org/10.1007/s12652-022-03881-2

[2] A. Ali and M. Meila, "Experiments with Kemeny ranking: What works when?" *Mathematical Social Sciences*, vol. 64, no. 1, pp. 28 – 40, 2012.

[3] N. Rico, C. R. Vela, and I. Díaz, "An analysis of the indexes measuring the agreement of a profile of rankings," 2021, proceedings of XIX CAEPIA.

[4] M. Kendall, "A new measure of rank correlation," *Biometrika*, 1938.

[5] J. C. Borda, *Mémoire sur les Élections au Scrutin*. Paris: Histoire de l'Académie Royale des Sciences, 1781.

[6] I. Azzini and G. Munda, "A new approach for identifying the Kemeny median ranking," *European Journal of Operational Research*, vol. 281, pp. 388–401, 2020.

Optimizando redes neuronales recurrentes mediante negaciones difusas

Mikel Ferrero-Jaurrieta
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Campus Arrosadía, s/n, 31006 Pamplona, España
mikel.ferrero@unavarra.es

Alfonso Induráin
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Campus Arrosadía, s/n, 31006 Pamplona, España
alfonso.indurain@unavarra.es

Anderson Cruz
Instituto Metrópole Digital
Universidade Federal Do Rio Grande Do Norte
Av. Cap. Mor Gouveia, 3000 - Lagoa Nova, Natal, Brazil
anderson@imd.ufrn.br

Helida Santos
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Campus Arrosadía, s/n, 31006 Pamplona, España
helida.salles@unavarra.es

Javier Fernández
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Campus Arrosadía, s/n, 31006 Pamplona, España
fcojavier.fernandez@unavarra.es

Humberto Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Campus Arrosadía, s/n, 31006 Pamplona, España
bustince@unavarra.es

***Index Terms*—Negación difusa, Red Neuronal Recurrente, LSTM, GRU.**

Las Redes Neuronales Recurrentes (RNN) [1], [2] introducidas en la década de 1980 con el objetivo de modelar de datos secuenciales (como las series temporales o el procesamiento de texto) tenían un gran problema en el proceso de entrenamiento, ya que el valor del gradiente tiende gradualmente a 0. Este problema, conocido como problema del gradiente desvaneciente [1] dio pie a la introducción de un nuevo tipo de redes neuronales recurrentes. Con el objetivo de resolver este problema, en 1997, S. Hochreiter y J. Schmidhuber introdujeron las memorias a corto y largo plazo (*Long Short-Term Memory*, LSTM). Éstas se basan en un mecanismo de puertas, que modelan la información que se retiene, la que se olvida y la que sale, así como las memorias mencionadas anteriormente. Las redes LSTM han tenido varias modificaciones, como LSTM con conexiones de mirilla (*peephole*) [3], CIFG [4] o JANET [5]. En 2014 Cho et. al mejoraron este sistema [6] simplificando la arquitectura de la unidad basada en menos puertas y utilizando sólo una memoria en lugar de dos memorias a corto y largo plazo.

Para simplificar el sistema de puertas, la Unidad Recurrente con puertas (*Gated Recurrent Unit*, GRU) acopla las compuertas, utilizando sólo una de ellas para modelar la actualización de la información de la célula. Esto significa que, en lugar de utilizar las puertas $\mathbf{p}$ y $\mathbf{q}$ de forma independiente, se acoplan, utilizando la siguiente relación $\mathbf{q} = \mathbf{1} - \mathbf{p}$ [5] y en consecuencia se reduce el número de parámetros a aprender. De forma similar, en [4] la red CIFG (*Coupled Input and Forget Gate*) también acopla las puertas de una LSTM de la misma manera.

En este trabajo, consideramos el proceso de actualización de la unidad GRU como un conjunto difuso [7], así como el proceso de olvido de información en las redes LSTM-CIFG. En este sentido, para un elemento concreto, un valor de pertenencia cercano a 0 significa que el elemento no se va a actualizar y un valor cercano a 1 significa que se va a actualizar casi por completo. Por tanto, la operación $1 - x$ puede entenderse como una negación o el complemento del conjunto difuso en cuestión. De esta forma, generalizamos la expresión $1 - x$ de las ecuaciones de la red neuronal recurrente utilizando negaciones difusas [8]–[10], y generando el conjunto difuso complementario a partir de estas negaciones. Se consideran diferentes clases y expresiones de negaciones difusas [11] y se generalizan otras [12]. En ellas se utilizan tanto expresiones fijas como parámetros que son aprendidos por la propia Unidad Recurrente. Para demostrar la eficacia de la modificación que planteamos, lo analizamos en conjuntos de datos de clasificación de texto, y mostramos que nuestro enfoque utilizando diferentes expresiones mejora el rendimiento de las arquitecturas modificadas frente al uso de negación estándar.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Agencia Estatal de Investigación (Gobierno de España) bajo el proyecto PID2019-108392GB-I00 (AEI/10.13039/ 501100011033) y por Tracasa

Instrumental y el Departamento de Política de Inmigración y Justicia del Gobierno de Navarra.

REFERENCIAS

[1] D. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

[2] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, ser. Studies in computational intelligence. Berlin: Springer, 2012. [Online]. Available: https://cds.cern.ch/record/1503877

[3] F. Gers and J. Schmidhuber, "Recurrent nets that time and count," in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, vol. 3, 2000, pp. 189–194 vol.3.

[4] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2017.

[5] J. van der Westhuizen and J. Lasenby, "The unreasonable effectiveness of the forget gate," *CoRR*, vol. abs/1804.04849, 2018. [Online]. Available: http://arxiv.org/abs/1804.04849

[6] K. Cho, B. V. Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," 2014.

[7] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[8] H. Bustince, P. Burillo, and F. Soria, "Automorphisms, negations and implication operators," *Fuzzy Sets and Systems*, vol. 134, no. 2, pp. 209–229, 2003.

[9] H. Zapata, H. Bustince, L. D. Miguel, and C. Guerra, "Some properties of implications via aggregation functions and overlap functions," *International Journal of Computational Intelligence Systems*, vol. 7, pp. 993–1001, 2014.

[10] E. Trillas, "Sobre funciones de negación en la teoría de conjuntos difusos." *Stochastica*, vol. 3, no. 1, pp. 47–60, 1979. [Online]. Available: http://eudml.org/doc/38807

[11] R. R. Yager, "On the measure of fuzziness and negation. ii. lattices," *Inf. Control.*, vol. 44, pp. 236–260, 1980.

[12] "Intuitionistic fuzzy generators application to intuitionistic fuzzy complementation," *Fuzzy Sets and Systems*, vol. 114, no. 3, pp. 485–504, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011498002796

Comparación de Reglas difusas. Aplicado a la toma de decisión en el Modus Ponens Generalizado

1st Asier Urio-Larrea
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
asier.urio@unavarra.es

2nd Graçaliz Pereira Dimuro
Centro de Ciência Computacionais
Universidade Federal do Rio Grande
Rio Grande, Brazil
gracalizdimuro@furg.br, gracaliz.pereira@unavarra.es

3rd Humberto Bustince
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
bustince@unavarra.es

4th Javier Fernandez
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
fcojavier.fernandez@unavarra.es

5th Jose Antonio Sanz
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
joseantonio.sanz@unavarra.es

6th Laura De Miguel
Departamento de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
laura.demiguel@unavarra.es

***Index Terms*—Reglas difusas, Modus Ponens Generalizado, Comparación Reglas**

La introducción por parte de Zadeh, en 1965, de la teoría de conjuntos difusos (FS) [1] permite representar el conocimiento humano, que es de naturaleza imprecisa. Este tipo de conjuntos permite modelar la incertidumbre propia de muchos conceptos semánticos. La modelización del razonamiento humano representado mediante conjuntos difusos necesita de un sistema de inferencia, y uno de esos sistemas es el Modus Ponens Generalizado (GMP) [2]. Este sistema utiliza reglas *si-entonces*, en las cuales los antecedentes son conjuntos difusos.

El caso más básico del GMP se representa de la siguiente manera:

Regla: si u es A entonces v es B
Hecho: u es A'

Conclusión: v es B'

Donde U y V son los universos del discurso, $A, A' \in FS(U)$ y $B, B' \in FS(V)$. De esta manera, conocida una regla y un hecho, se trata de determinar el consecuente B' de ese hecho.

Se han planteado gran variedad métodos de inferencia para resolver el GMP ([2], [3], [4], [5], [6]), el problema es que, además de seleccionar el método a utilizar, todos ellos dependen de diversos parámetros (t-normas,...) que es necesario definir y que afectan al resultado de la inferencia ([7], [8]).

Para poder determinar el mejor método del GMP para un problema dado el objetivo del presente trabajo es la elaboración de un sistema de comparación de reglas, para ello se proponen las propiedades que debería cumplir:

- La comparación de dos reglas deberá ser simétrica.
- Si los antecedentes y consecuentes son iguales en ambas reglas, el resultado de la comparación deberá ser 1.
- Si los soportes de cada par de antecedentes (o consecuentes) son disjuntos, el resultado de la comparación deberá ser 0.

En base a estos axiomas se presenta un método de construcción para el sistema de comparación de reglas mediante la utilización de funciones de overlap [9], indices de overlap [6] y funciones de similitud [10], [11].

El segundo objetivo es la aplicación de la comparación de reglas a la selección de los parámetros más adecuados para la resolución del GMP, mediante un sistema de toma de decisiones. Para ello, se infiere el consecuente por cada uno de los métodos y se construyen nuevas reglas con el antecedente del hecho y los consecuentes calculados. Estas reglas se compara con las reglas utilizadas de cada uno de los métodos y se selecciona aquella que tenga un valor de comparación más alto.

AGRADECIMIENTOS

La investigación ha sido financiada por el proyecto PID2019-108392GB-I00 (AEI/10.13039/501100011033). Asier Urio-Larrea es beneficiario de una ayuda predoctoral de la convocatoria Santander-UPNA para el curso 2021/22. Graçaliz Pereira esta financiada del proyecto CNPq (301618/2019-4) y FAPERGS (19/2551-0001279-9)

REFERENCES

[1] L. Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) 338–353.

[2] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (1) (1973) 28–44. doi:10.1109/TSMC.1973.5408575.

[3] W. Pedrycz, Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data, Fuzzy Sets and Systems 16 (2) (1985) 163–175.

[4] W. Bandler, L. J. Kohout, Semantics of implication operators and fuzzy relational products, International Journal Man-Mach. Stud. 12 (1979) 89–116.

[5] L. Kóczy, K. Hirota, Approximate reasoning by linear rule interpolation and general approximation, International Journal of Approximate Reasoning 9 (3) (1993) 197–225.

[6] S. Garcia-Jimenez, H. Bustince, E. Hüllermeier, R. Mesiar, N. R. Pal, A. Pradera, Overlap indices: Construction of and application to interpolative fuzzy systems, IEEE Transactions on Fuzzy Systems 23 (4) (2015) 1259–1273. doi:10.1109/TFUZZ.2014.2349535.

[7] A. Konguetsof, N. Mylonas, B. Papadopoulos, Fuzzy reasoning in the investigation of seismic behavior, Mathematical Methods In The Applied Sciences 43 (13, SI) (2020) 7747–7757. doi:10.1002/mma.6184.

[8] P. Pagouropoulos, C. D. Tzimopoulos, B. K. Papadopoulos, A method for the detection of the most suitable fuzzy implication for data applications, Evolving Systems 11 (3) (2020) 467–477. doi:https://doi.org/10.1007/s12530-018-9233-0.

[9] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna, Overlap functions, Nonlinear Analysis: Theory, Methods & Applications 72 (3-4) (2010) 1488–1499.

[10] 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) (2008) 525 – 536. doi:https://doi.org/10.1016/j.patrec.2007.11.007.

[11] X. Liu, Entropy, distance measure and similarity measure of fuzzy sets and their relations, Fuzzy Sets and Systems 52 (2) (1992) 305–318.

Funciones de pooling basadas en Penalty aplicadas a la reducción de características en Redes Neuronales Convolucionales

1° Iosu Rodriguez-Martinez
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
iosu.rodriguez@unavarra.es

2° Pablo Lizarraga-Guerra
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
pablo.lizarraga@unavarra.es

3° Tiago Asmus
institute of Mathematics, Statistics and Physics
Universidade Federal do Rio Grande
Rio Grande, Brazil
tiagoasmus@furg.br

4° Graçaliz Dimuro
center for Computational Sciences
Universidade Federal do Rio Grande
Rio Grande, Brazil
gracalizdimuro@furg.br

5° Angela Bernardini
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
angela.bernardini@unavarra.es

6° Francisco Herrera
dept. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
herrera@decsai.ugr.es

7° Humberto Bustince
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
bustince@unavarra.es

Las funciones de agregación basadas en Penalty escogen la mejor agregación de un conjunto de valores, con respecto a la desviación de cada resultado agregado y las entradas de la función [1], [2]. Han sido muy utilizadas en tareas de reducción de imagen, donde los valores de intensidad de los píxeles de regiones disjuntas de la imagen se fusionan generando representantes individuales que tratan de preservar la máxima cantidad de información posible [3], [4].

Las Redes Neuronales Convolucionales (CNNs por sus siglas en inglés), una familia de modelos de red neuronal capaces de extraer las características visuales más importantes de una imagen de forma automática, llevan a cabo un proceso similar en las capas de "pooling". Tras extraer "ìmágenes de características" mediante la convolución de un conjunto de filtros sobre una imagen de entrada, estas capas agregan los valores de ventanas disjuntas, reduciendo la dimensionalidad de las características obtenidas. Pese a que este proceso de fusión de datos se ha realizado tradicionalmente mediante funciones simples como la media aritmética o el máximo, se han propuesto varias alternativas: el pooling estocástico (*Stochastic pooling*) introduce cierta componente de regularización al modelo resultante mediante la selección aleatoria de uno de los valores de la ventana a agregar, en base a una distribución multinomial [5]; El Pooling Mezcla (*Mixed Pooling*) combina las salidas de los pooling máximo y media mediante una combinación convexa cuyos coeficientes pueden ser aprendidos de diferentes formas [6]; En trabajos previos, hemos generalizado la estrategia anterior con éxito a funciones crecientes adicionales, en la forma de capas CombPool [7].

El objetivo de la última propuesta era el de incorporar otras funciones conocidas de la teoría de agregaciones al campo de la fusión de características, tales como la integral de Sugeno o algunas generalizaciones de la misma, adaptadas para operar sobre valores reales. El proceso de pooling se lleva a cabo mediante una operación de reducción de imagen en la que se agregan los valores de ventanas disjuntas $\mathbf{x} \in \mathbb{R}^k$ de las imágenes de características extraídas por la CNN mediante una función $F : \mathbb{R}^k \to \mathbb{R}$. En las capas CombPool, se fijan n funciones crecientes $A_1, \ldots, A_n : \mathbb{R}^k \to \mathbb{R}$ y se utilizan expresiones del tipo $F = \sum_{i=1}^{n} \alpha_i A_i(\mathbf{x})$, con $\alpha_1, \ldots, \alpha_n \in [0, 1]$, como función de pooling. La contribución de cada función A_i se "aprende" optimizando los coeficientes $alpha$ mediante

el descenso del gradiente, del mismo modo que se ajustan el resto de parámetros del modelo.

En este trabajo, estudiamos otras estrategias que permitan combinar distintas funciones de agregación, haciendo uso del concepto de función Penalty. De la misma forma que la mejor reducción para los píxeles de regiones distintas de una imagen pueden venir dados por funciones de agregación distintas, distintas regiones de las imágenes de características extraídas por una CNN pueden presentar información que sea mejor resumida en base a distintas funciones de agregación.

Además, también introducimos una estrategia que aprovecha la información de las funciones penalty para ponderar la contribución de las distintas funciones empleadas en una capa CombPool. De este modo, los coeficientes $\alpha_1, \ldots, \alpha_n$ no se ajustan en base a un proceso puramente supervisado, sino que incorporamos la información devuelta por una función penalty para dar mayor peso a los coeficientes asociados a la agregación que devuelva una menor distancia.

En concreto, dada una función penalty $P : [a, b]^{n+1} \to \mathbb{R}^+$, calculamos el vector de penalties $\mathbf{p} \in (\mathbb{R}^+)^n$, donde $p_i = P(\mathbf{x}, A_i(\mathbf{x}))$ indica el valor asignado por la función P a la reducción ofrecida por la función de agregación A_i para el vector de valores $\mathbf{x}$. El coeficiente $\alpha_i = \lambda_i(1 - \sigma(\mathbf{p})_i)$, donde $\sigma : \mathbb{R}^n \to [0, 1]^n$ es la función "softmax" dada por $\sigma(\mathbf{p})_i = \frac{e^{p_i}}{\sum_{j=1}^{n} e^{p_j}}$ y $\lambda_1, \ldots, \lambda_n$ son parámetros del modelo, es inversamente proporcional al valor devuelto por la función penalty para la reducción dada por A_i.

De este modo, obtenemos una función de agregación basada en penaltis "suaves" que, en lugar de asignar importancia total a la función que minimiza P, tiene en cuenta todas las funciones empleadas. Como puede verse, una función de agregación basada en una función penalty clásica, es una versión estricta del caso anterior en la que $\alpha_i = 1$ si $P(\mathbf{x}, A_i(\mathbf{x})) \leq P(\mathbf{x}, A_j(\mathbf{x}))$ para todo $j \neq i$ y $\alpha_i = 0$ en caso contrario.

Pruebas realizadas utilizando ambas estrategias de pooling ofrecen resultados prometedores en datasets estándar, superando el rendimiento de las capas CombPool base y demostrando que la información aportada por las funciones penalty permite reducciones de información más inteligentes.

AGRADECIMIENTOS

Beca PID2019-108392GB-I00/ financiada por MCIN/AEI/10.13039/501100011033. T. Asmus y G. Dimuro están parcialmente apoyados por los proyectos CNPq (301618/2019-4), FAPERGS (19/2551-0001279-9). F. Herrera está apoyado por el proyecto de Excelencia Andaluz P18-FR4961. I. Rodriguez-Martinez está apoyado parcialmente por el proyecto PC095-096 FUSIPROD, Tracasa Instrumental (iTRACASA) y Gobierno de Navarra - Departamento de Universidad, Innovación y Transformación Digital.

REFERENCIAS

[1] T. Calvo and G. Beliakov, "Aggregation functions based on penalties," *Fuzzy sets and Systems*, vol. 161, no. 10, pp. 1420–1436, 2010.

[2] H. Bustince, G. Beliakov, G. P. Dimuro, B. Bedregal, and R. Mesiar, "On the definition of penalty functions in data aggregation," *Fuzzy Sets and Systems*, vol. 323, pp. 1–18, 2017.

[3] G. Beliakov, H. Bustince, and D. Paternain, "Image reduction using means on discrete product lattices," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1070–1083, 2011.

[4] D. Paternain, A. Jurio, and G. Beliakov, "Color image reduction by minimizing penalty functions," in *2012 IEEE International Conference on Fuzzy Systems*. IEEE, 2012, pp. 1–7.

[5] M. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks: 1st international conference on learning representations, iclr 2013," Jan. 2013, 1st International Conference on Learning Representations, ICLR 2013 ; Conference date: 02-05-2013 Through 04-05-2013.

[6] C.-Y. Lee, P. Gallagher, and Z. Tu, "Generalizing pooling functions in cnns: Mixed, gated, and tree," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 863–875, 2018.

[7] I. Rodriguez-Martinez, J. Lafuente, R. H. Santiago, G. P. Dimuro, F. Herrera, and H. Bustince, "Replacing pooling functions in convolutional neural networks by linear combinations of increasing functions," *Neural Networks*, 2022.

A fuzzy logic-based approach to reason with inconsistent probabilistic theories

Tommaso Flaminio
IIIA – CSIC
08193 Bellaterra, Spain
tommaso@iiia.csic.es

Lluis Godo
IIIA – CSIC
08193 Bellaterra, Spain
godo@iiia.csic.es

Sara Ugolini
IIIA – CSIC
08193 Bellaterra, Spain
sara@iiia.csic.es

Francesc Esteva
IIIA – CSIC
08193 Bellaterra, Spain
esteva@iiia.csic.es

***Abstract*—In this paper we consider the probability logic over Rational Pavelka logic (RPL), denoted FP(RPL), and we explore two possible approaches to reason from inconsistent FP(RPL) theories in a non-trivial way. The first one amounts to replace the logic RPL, that is explosive, by its paraconsistent degree-preserving companion RPL$^{\leq}$. The second one consists of suitably weakening the formulas in an inconsistent theory T, depending on the degree of inconsistency of T.**

***Index Terms*—Probabilistic reasoning; Łukasiewicz fuzzy logic; Paraconsistent reasoning models; Inconsistency measures**

I. Introduction

Reasoning about probability can be properly handled in a fuzzy logical setting by expanding the language of Łukasiewicz fuzzy logic with a unary modality P and interpreting, for every classical formula φ, the modal formula $P\varphi$ as "φ is probable". Clearly, $P\varphi$ is a fuzzy proposition, whose truth-degree can be taken as the probability of φ. More precisely, the fuzzy modal logic FP(Ł), as firstly introduced in [4] and improved in [3], extends the language of Łukasiewicz logic Ł by the unary modal operator P that applies only to classical propositions and uses the ground logic Ł to express the basic properties of a probability function (in particular the finite additivity). Very recently, in [1] the authors have studied in depth the relationship of this fuzzy logic-based approach to more traditional probability logics after Halpern et al. see e.g. [5]. In this paper we will rather consider the probability logic FP(RPL) over Rational Pavelka logic (RPL), the expansion of Łukasiewicz fuzzy logic with rational truth-constants.

In this paper we explore two possible approaches to reason from inconsistent FP(RPL) theories in a non-trivial way. The first one amounts to replace the external logic RPL, that is explosive, by its paraconsistent companion RPL$^{\leq}$. The second one amounts to suitably weaken formulas of an inconsistent theory T depending on the degree of inconsistency of T.

II. The probability logics FP(Ł) and FP(RPL)

Based on a first formalisation in [4], the probability logic FP(Ł) (FP for Fuzzy Probability) was introduced in [3] and defined as a sort of modal extension with a unary operator P over the well-known Łukasiewicz fuzzy logic Ł (see e.g. [3] for details on both Ł and FP(Ł) logics). FP(Ł) allows for reasoning about the probability of classical propositions.

Let L denote the language of classical propositional logic (CPL) built from a countable set V of propositional variables using the classical binary connectives $\wedge$ and $\neg$ Then, the language of FP(Ł) is defined as follows. *Formulas* of FP(Ł) are of two types:

(1) Non-modal: they are the classical logic formulas of L and will be denoted by lower case Greek letters $\varphi, \psi, \ldots$.

(2) Modal: they are built from basic modal formulas of the form $P\varphi$, where $\varphi \in \mathsf{L}$ using the connectives of Ł $(\to_L, \neg)$, and denoted by upper case Greek letters $\Phi, \Psi, \ldots$.

This is a two-layer language, neither nested modalities nor formulas combining non-modal and modal subformulas are not allowed.

Axioms and rules of FP(Ł) are as follows:

(CPL) Axioms and rules of CPL for non-modal formulas;
(Ł) Axioms and rules of Ł for modal formulas;
(P) Axioms and rules for the modality P:
- (P1) $P(\varphi \to \psi) \to_L (P(\varphi) \to_L P(\psi))$
- (P2) $P(\neg\varphi) \leftrightarrow_L \neg P(\varphi)$
- (P3) $P(\varphi \vee \psi) \leftrightarrow_L [P(\varphi) \oplus (P(\psi) \ominus P(\varphi \wedge \psi))]$[1]
- (Nec) if $\vdash_{CPL} \varphi$, derive $P(\varphi)$.

Models of FP(Ł) are probability Kripke structures $K = \langle W, e, \mu\rangle$, where: W is a non-empty set of possible worlds; $e : V \times W \to \{0,1\}$ provides for each world a *Boolean* (two-valued) evaluation of the proposition variables, that is, $e(p,w) \in \{0,1\}$ for each propositional variable $p \in Var$ and each world $w \in W$; and $\mu : 2^W \to [0,1]$ is a finitely additive probability measure on a Boolean algebra of subsets of W such that for each p, the set $\{w \mid e(p,w) = 1\}$ is measurable (cf. [3] 8.4.1). A truth-evaluation e is extended to non-modal formulas in the classical way, to elementary modal formulas as follows:

$$e(P\varphi, w) = \mu(\{w' \in W \mid e(\varphi, w') = 1\}),$$

and to compound modal formulas by using the truth-functions of Ł logic. Actually $e(P\varphi, w)$ does not depend on w and we will write $e(\Phi)$.[2] We will also denote by e_μ the truth-evaluation on modal formulas determined by the model $\langle \Omega, e, \mu\rangle$, where Ω is the set of classical models for L.

[1] Recall that $\Phi \oplus \Psi := \neg\Phi \to_L \Psi$ and $\Phi \ominus \Psi := \neg(\Phi \to_L \Psi)$.

[2] Recall $e(\Phi \to_L \Psi) = \min(1 - e(\Phi) + e(\Psi), 0)$, $e(\neg\Phi) = 1 - e(\Phi)$, $e(\Phi \oplus \Psi) = \min(e(\Phi) + e(\Psi), 1)$ and $e(\Phi \ominus \Psi) = \max(e(\Phi) - e(\Psi), 0)$.

Soundness and completeness of the logic FP(Ł) w.r.t. to the class of probability Kripke models reads as follows: if $T \cup \{\Phi\}$ is a finite set of modal FP(Ł)-formulas, then T proves Φ in FP(Ł), written $T \vdash_{FP} \Phi$, iff for any probability μ on Ω, $e_\mu(\Phi) = 1$ whenever $e_\mu(\Psi) = 1$ for all $\Psi \in T$.

If one wants to formalise reasoning with numeric probabilistic expressions, then one has to replace in FP(Ł) the (external) logic Ł by its expansion with rational truth-constants, the so-called Rational Pavelka logic (RPL for short). So we add to the language of Ł a rational truth constant $\overline{r}$ for every rational $r \in [0, 1]$. As Hájek shows [3], Rational Pavelka logic can be then axiomatized by adding to the axioms Łukasiewicz the following bookkepping axioms:

$$\text{(BK)} \quad \overline{r} \to \overline{s} \equiv \overline{\min(1, 1 - r + s)}$$

for any rational numbers $r, s \in [0, 1]$. The resulting probability logic, FP(RPL), inherits the soundness and completeness results from FP(Ł), where now FP(RPL)-evaluations e_μ are further equired to correctly interpret the truth-constants, that is, $e(\overline{r}) = r$ for every rational $r \in [0, 1]$.

III. Dealing with inconsistent FP(RPL) theories

As the probability logic FP(RPL) is grounded on the RPL logic, the latter being explosive, the logic FP(RPL) is explosive as well. This means that, for any formula Φ in the language of FP(RPL), $\{\Phi, \neg\Phi\} \vdash_{FP} \bot$, and thus $\{\Phi, \neg\Phi\} \vdash_{FP} \Psi$ for any Ψ. Our contribution consists in presenting two approaches to escape the explosion principle in FP(RPL) and to handle inconsistent probabilistic theories in a non-trivial way, briefly introduced below.

A. A paraconsistent probability logic

The first approach consists in replacing RPL by its "degree preserving companion", denoted by RPL$^{\leq}$. Conforming to the usual way of defining deductions in degree-preserving logics, given two modal formulas we define Φ and Ψ, $\Phi \vdash_{FP^{\leq}} \Psi$ iff for every probabilistic Kripke model $\mathcal{M} = (W, e, \mu)$ of FP$^{\leq}(RPL)$, $\|\Phi\|_{\mathcal{M}} \leq \|\Psi\|_{\mathcal{M}}$. This generalises to the more general case in which $T = \{\Phi_1, \ldots, \Phi_n\}$ is any finite set of modal formulas by defining $T \vdash_{FP^{\leq}} \Psi$ iff for all probabilistic Kripke model $\mathcal{M}$,

$$\|\Phi_1 \wedge \ldots \wedge \Phi_n\|_{\mathcal{M}} \leq \|\Psi\|_{\mathcal{M}}.$$

Let us notice that the logic FP$^{\leq}(RPL)$ is not explosive, and hence *paraconsistent*. Indeed, for each classical formula φ that is neither a classical theorem nor a contradiction, $P(\varphi), \neg P(\varphi) \nvdash_{FP^{\leq}} \bot$ because, semantically, one can find a probability μ that assigns $\mu(\varphi) = 1/2$ and this gives

$$\min\{\mu(\varphi), \mu(\neg\varphi)\} = 1/2 > 0.$$

B. An inconsistency-tolerant probabilistic logic

Recall that, from a semantical point of view, the logic $FP(\text{Ł})$ is defined as follows: for any set of FP(Ł)-formulas $T \cup \{\Phi\}$, $T \models_{FP} \Phi$ if, for every probability μ on Boolean formulas, if μ is a model of T then $e_\mu(\Phi) = 1$, where by μ being a model of T we mean that $e_\mu(\Psi) = 1$ for every $\Psi \in T$. We will denote by $\|T\|$ the set probability measures on formulas that are models of T.

Of course, the above definition trivializes in the case T is inconsistent, i.e. when $\|T\| = \emptyset$. But in FP(Ł) one can take advantage of its fuzzy setting and consider the notion of (in)consistency as being fuzzy as well. Indeed, even if T has no models, a situation where, for every probability μ there is always a formula Φ in T such that $e_\mu(\Phi) = 0$, is qualitatively different from a situation where there is a probability μ such that $e_\mu(\Phi) \geq \alpha$ for all $\Phi \in T$, for some value α close to 1. In the former case T is clearly inconsistent, while in the latter case one could say that T is close to being consistent.

This observation justifies to define, for each threshold α, the set of *α-generalised models* of T as follows:

$$\|T\|_\alpha = \{\text{probability } \mu \mid \text{for all } \Psi \in T, e_\mu(\Psi) \geq \alpha\}.$$

Note that the set $\|T\|_1$ coincides with the set of usual models of T. Moreover $\|T\|_\alpha$ is a convex set of probabilities.

Definition 3.1: Let T be a theory of $FP(\text{Ł})$. The consistency degree of T is defined as $Con(T) = \sup\{\beta \in [0, 1] \mid \|T\|_\beta \neq \emptyset\}$. Dually, the inconsistency degree of T is defined as $Incon(T) = 1 - Con(T) = \inf\{1 - \beta \in [0, 1] \mid \|T\|_\beta \neq \emptyset\}$.[3]

The idea we explore in this paper is to use α-generalised models instead of usual models to define a context-dependent inconsistent-tolerant notion of probabilistic entailment.

Definition 3.2: Let T be a theory such that $Con(T) = \alpha > 0$. We define: $T \mid\!\!\approx^* \Phi$ if $e_\mu(\Phi) = 1$ for all probabilities $\mu \in \|T\|_\alpha$.

Note that if $Con(T) > 0$, then $T \not\mid\!\!\approx^* \bot$, hence $\mid\!\!\approx^*$ does not trivialize even if T is inconsistent ($Con(T) < 1$). As an example, if $T = \{P\varphi \leftrightarrow_L \overline{0.4}, P\varphi \leftrightarrow_L \overline{0.3}\}$, that is inconsistent, then $Con(T) = 0.95$ and $T \mid\!\!\approx^* \overline{0.35} \leftrightarrow_L P\varphi$.

The following are some interesting properties of the consequence relation $\mid\!\!\approx^*$: clearly, $\mid\!\!\approx^*$ is not monotonic, while $\mid\!\!\approx^*$ is idempotent, that is, if $S \mid\!\!\approx^* \varphi$ and $T \mid\!\!\approx^* \psi$ for all $\psi \in S$, then $T \mid\!\!\approx^* \varphi$.

Acknowledgments: Esteva, Flaminio and Godo acknowledge support by the PID2019-111544GB-C21 Spanish project ISINC funded by MCIN/AEI/10.13039/501100011033. Ugolini acknowledges the Marie Sklodowska-Curie grant agreement No 890616 (H2020-MSCA-IF-2019).

References

[1] P. Baldi, P. Cintula, C. Noguera. Classical and Fuzzy Two-Layered Modal Logics for Uncertainty: Translations and Proof-Theory. *Int. J. Comput. Intell. Syst.* 13(1): 988-1001, 2020.

[2] G. De Bona, M. Finger, N. Potyka, M. Thimm. Inconsistency Measurement in Probabilistic Logic. In J. Grant, M. V. Martinez (Eds.), *Measuring Inconsistency in Information*, Vol. 73 of Studies in Logic, College Publications, 2018, pp. 235-269.

[3] P. Hájek. *Metamathematics of fuzzy logic,* Kluwer 1998.

[4] P. Hájek, L. Godo, F. Esteva, Probability and Fuzzy Logic. In *Proc. of Uncertainty in Artificial Intelligence* UAI'95, pp. 237-244, 1995.

[5] J. Halpern. *Reasoning about Uncertainty*. MIT Press, 2003.

[3] $Incon(\cdot)$ can be seen as a particular case of distance-based inconsistency mesure, see e.g. [2].

Una metodología para aprender sistemas de clasificación basados en reglas intervalo-valoradas difusas

1st José Antonio Sanz y Humberto Bustince
Departamento of Estadística, Informática y Matemáticas e Instituto de Smart Cities
Universidad Publica de Navarra
Pamplona, España
{joseantonio.sanz,bustince}@unavarra.es

***Abstract*—Este trabajo presenta la propuesta publicada en [1] en la que se definió una metodología para realizar el aprendizaje de un sistema de clasificación basado en reglas intervalo-valoradas difusas. La novedad reside en que los conjuntos difusos intervalo-valorados se utilizan tanto en el proceso de inferencia como en el de aprendizaje, lo que permite explotarlos en todas las fases. Además, dichos conjuntos se utilizarán solamente en caso de que sean necesarios. Es decir, se podrán aprender reglas cuyos antecedentes estén compuestos por: 1) únicamente conjuntos difusos numéricos; 2) únicamente conjuntos difusos intervalo-valorados o 3) una mezcla de ambos tipos de conjuntos. Para ello, se desarrolló una metodología de envoltorio (*wrapper*) compuesta por dos algoritmos evolutivos. El primero de ellos se utiliza para aprender la mejor partición difusa para cada atributo de entrada, a partir de la que se aprenden las reglas, y el segundo se utiliza para optimizar las reglas generadas en aras de incrementar la capacidad predictiva del clasificador.**

***Index Terms*—Problemas de clasificación, sistemas difusos evolutivos, sistemas de clasificación basados en reglas difusas, conjuntos difusos intervalo-valorados**

I. Introduction

Los Conjuntos Difusos Intervalo-Valorados (CDIVs) [2] son una extensión de los Conjuntos Difusos Numéricos (CDNs) [3]. Su uso puede ser beneficioso al abordar problemas difíciles, es decir, cuando la definición de las funciones de pertenencia sea complicada y los CDIVs pueden ofrecer una ventaja al representar la incertidumbre/ignorancia propia de esta etapa [4]. En este sentido, la amplitud de los intervalos permite representar la falta de conocimiento del experto a la hora de definir los CDNs [5]. Debido a las ventajas que presentan los CDIVs, han sido utilizados para afrontar problemas de aprendizaje automático como regresión [6], clustering [7] y clasificación [8].

Principalmente existen dos grandes formas de generar sistemas intervalo-valorados difusos [9]:

1) Generar un sistema basado en reglas difusas inicial al que posteriormente se le añaden los CDIVs, cuya forma se optimiza para tratar de mejorar el rendimiento del sistema.
2) Desarrollar un método que aprenda directamente un sistema intervalo-valorado difuso.

La mayoría de las propuestas para realizar el aprendizaje de Sistemas de Clasificación Basados en Reglas Intervalo-Valoradas Difusas (SCBRIVDs) [10] están basadas en la primera opción. Sin embargo, en esta metodología las reglas se aprenden utilizando CDNs y por tanto, puede que la potencia de los CDIVs no se explote completamente.

El aprendizaje de un SCBRIVDs, sin utilizar un sistema difuso como base, es un reto puesto que su éxito depende directamente de la calidad de las particiones intervalo-valoradas que se utilicen durante el proceso de aprendizaje. Por ello, una pregunta clave es cómo definir las particiones intervalo-valoradas difusas para cada atributo del problema. Además, al generar estos sistemas, otra duda habitual es qué tipo de conjunto difuso, CDNs o CDIVs, es el más apropiado para modelar cada etiqueta lingüística. Es decir, si merece la pena utilizar los CDIVs o no. Por ello, estas dos cuestiones se pueden abordar como un problema de aprendizaje en el que se debe construir la mejor partición difusas para cada variable y en la que se debe determinar la mejor forma de modelar la función de pertenencia para cada etiqueta.

En este trabajo se destaca la metodología publicada en [1] mediante la que se abordaron los problemas mencionados anteriormente. En concreto, la principal contribución de la propuesta realizada en [1] fue la definición de una metodología de envoltorio (*wrapper*) cuyo objetivo principal es realizar el aprendizaje de un SCBRIVD, sin utilizar un sistema difuso inicial, que sea preciso. Para ello, los principales componentes desarrollados fueron:

1) Un método de aprendizaje de reglas intervalo-valoradas difusas. En concreto, se propuso el uso de reglas de asociación basadas en CDIVs puesto que las reglas de asociación difusas se utilizan en uno de los clasificadores difusos más precisos de la actualidad como es FARC-HD [11]. Este método de aprendizaje podría ser diferente ya que la metodología de envoltorio permite cambiarlo

Este trabajo está financiado parcialmente por el Ministerio Español de Ciencia y Tecnología (proyecto PID2019-108392GB-I00 / AEI / 10.13039/501100011033) y por la Universidad Publica de Navarra mediante el proyecto PJUPNA1926.

fácilmente.

2) La propia metodología de envoltorio para lo que se utiliza un proceso evolutivo en el que cada cromosoma representa la unión de todas las particiones difusas del problema a abordar. El proceso evolutivo está compuesto por dos sub-componentes:
   - Un Algoritmo Genético (AG) que se utiliza para aprender la forma de cada CDIV (su posición lateral, la amplitud del soporte de tanto el extremo inferior como el superior así como la amplitud del núcleo del extremo superior). El componente clave del AG es la función de evaluación en la que se aprende el SCBRIVD, utilizando el método de aprendizaje mencionado anteriormente, utilizando las particiones difusas representadas por el cromosoma. El valor de *fitness* es el ratio de ejemplos correctamente clasificador por el SCBRIVD creado.
   - Un algoritmo de Evolución Diferencial (ED) que se utiliza para optimizar el SCBRIVD generado en el punto anterior. En concreto, se ajusta la forma del extremo superior (soporte y núcleo) y se realiza un proceso de selección de reglas. Cabe destacar que este proceso no se realiza para todos los SCBRIVD (cromosomas del GA) sino que se aplica sobre los que su calidad inicial se estima que permita mejorar el mejor rendimiento obtenido hasta ese momento. El valor de *fitness* del cromosoma del GA se actualiza tras aplicar el ED ya que el rendimiento del SCBRIVD habrá mejorado.

Al acabar el proceso de aprendizaje se obtendrá un SCBRIVDs compuesto por reglas interpretables cuyos antecedentes pueden combinar CDNs y CDIVs, ya que el proceso evolutivo permite determinar qué tipo de conjunto difuso es el más apropiado para cada etiqueta. Es decir, los CDIVs se utilizarán cuando sean realmente necesarios y la forma de todas las funciones de pertenencia estará optimizada para abordar el problema de clasificación, lo que implica obtener un clasificador difuso preciso.

La calidad del nuevo método, llamado IVFARC, se evaluó utilizando 29 datasets y se realizó un estudio estadístico adecuado. En concreto, en el estudio experimental se determinó la mejor configuración del nuevo método y se realizó una comparación con respecto a la versión difusa numérica de IVFARC para comprobar que el uso de los CDIVs es beneficioso. Además, para avalar la calidad de la nueva propuesta, se comparó su rendimiento con el de varios clasificadores difusos del estado del arte como FARC-HD [11], FARC-HD utilizando una generalización de la integral Choquet integral [12], FURIA [13] e IVTURS [14]. Finalmente, se utilizaron métricas de complejidad para determinar cuándo el uso de IVFARC es más adecuado. En dicho análisis, se mostró que cuando la dificultad de los problemas crece la utilidad del uso de CDIVs es mejor Este hecho está de acuerdo con la problemática de la definición de las funciones de pertenencia, lo que corrobora la hipótesis del uso de los CDIVs.

REFERENCES

[1] J. A. Sanz and H. Bustince, "A wrapper methodology to learn interval-valued fuzzy rule-based classification systems," *Applied Soft Computing*, vol. 104, p. 107249, 2021.

[2] R. Sambuc, "Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne," Ph.D. dissertation, University of Marseille, 1975.

[3] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.

[4] J. M. Mendel, "Advances in type-2 fuzzy sets and systems," *Information Sciences*, vol. 177, no. 1, pp. 84–110, 2007.

[5] J. Sanz, A. Fernández, H. Bustince, and F. Herrera, "A genetic tuning to improve the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets: Degree of ignorance and lateral position," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 751–766, 2011.

[6] O. Castillo and P. Melin, "A review on interval type-2 fuzzy logic applications in intelligent control," *Information Sciences*, vol. 279, pp. 615 – 631, 2014.

[7] J. Xu, G. Feng, T. Zhao, X. Sun, and M. Zhu, "Remote sensing image classification based on semi-supervised adaptive interval type-2 fuzzy c-means algorithm," *Computers and Geosciences*, vol. 131, pp. 132–143, 2019.

[8] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, and H. Hagras, "A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 4, pp. 973–990, 2015.

[9] J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 808–821, 2006.

[10] J. Sanz, A. Fernández, H. Bustince, and F. Herrera, "Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning," *Information Sciences*, vol. 180, no. 19, pp. 3674–3685, 2010.

[11] J. Alcala-Fdez, R. Alcala, and F. Herrera, "A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 5, pp. 857–872, 2011.

[12] G. Lucca, G. Dimuro, J. Fernandez, H. Bustince, B. Bedregal, and J. Sanz, "Improving the performance of fuzzy rule-based classification systems based on a nonaveraging generalization of CC-integrals named c_{F1F2}-integrals," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 1, pp. 124–134, 2019.

[13] J. Hühn and E. Hüllermeier, "FURIA: an algorithm for unordered fuzzy rule induction," *Data Mining and Knowledge Discovery*, vol. 19, no. 3, pp. 293–319, 2009.

[14] J. Sanz, A. Fernández, H. Bustince, and F. Herrera, "IVTURS: A linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 3, pp. 399–411, 2013.

Relacionando implicaciones de atributos y reglas de decisión

Fernando Chacón-Gómez
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
fernando.chacon@uca.es

M. Eugenia Cornejo
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
mariaeugenia.cornejo@uca.es

Jesús Medina
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
jesus.medina@uca.es

***Resumen*—El análisis de conceptos formales (FCA, del inglés Formal Concept Analysis) y la teoría de conjuntos rugosos (RST, del inglés Rough Set Theory) son herramientas matemáticas de gran utilidad para el manejo y extracción de información contenida en bases de datos. En este trabajo se presenta un estudio comparativo preliminar entre las reglas de decisión de RST y las implicaciones de atributos de FCA.**

***Palabras clave*—Análisis de conceptos formales, teoría de conjuntos rugosos, implicación de atributos, regla de decisión.**

I. Introducción

Ganter y Wille introdujeron FCA como una teoría matemática para la extracción de conocimiento de conjuntos de datos. En esta teoría, los conjuntos de datos son interpretados como contextos, los cuales están formados por una colección de objetos, una colección de atributos y una relación binaria entre ellos. Un tema de investigación interesante en FCA es el estudio de las implicaciones de atributos [9]–[11]. Estas implicaciones son muy útiles para determinar la importancia de los atributos en el contexto.

Por otro lado, Pawlak introdujo la teoría de conjuntos rugosos (RST) [14] para analizar conjuntos de datos conteniendo información completa e imprecisa. En este marco de trabajo, los conjuntos de datos relacionales se interpretan como una tabla de decisión, compuesta de objetos, atributos y aplicaciones definidas para cada atributo caracterizando los objetos de la tabla. En RST, las tablas de decisión pueden interpretarse en términos lógicos gracias a la noción de regla de decisión [15]. De hecho, esta interpretación lógica ayuda significativamente a la extracción de conclusiones a partir de la tabla.

Ambas teorías, FCA y RST, han sido relacionadas en distintos trabajos centrados en diversos temas [1], [3], [4], [12]. Este trabajo complementa la contribución presentada en [5] comparando ambas teorías. El objetivo principal consiste en estudiar la relación entre las reglas de decisión y las implicaciones de atributos, así como su extensión al caso difuso.

Parcialmente financiado por los fondos FEDER 2014-2020 en colaboración con la Agencia Estatal de Investigación (AEI) a través del proyecto con referencia PID2019-108991GB-I00 y el proyecto financiado por la Junta de Andalucía y los fondos FEDER con referencia FEDER-UCA18-108612, y por la European Cooperation in Science & Technology (COST) Action CA17124.

II. Nociones preliminares

Esta sección recuerda algunas nociones preliminares sobre FCA [10], [11] y RST [15]. La extensión de estos conceptos al paradigma multiadjunto puede encontrarse en [8], [13].

Análisis de conceptos formales

Los conjuntos de datos se representan en FCA como contextos formales, que son tuplas $(A, B, \mathcal{R})$ donde A y B son conjuntos no vacíos de atributos y objetos, respectivamente, y $\mathcal{R}$ es una relación entre ellos. Dado un contexto, los operadores de derivación $^{\uparrow}\colon 2^B \to 2^A$ y $^{\downarrow}\colon 2^A \to 2^B$ permiten caracterizar los subconjuntos de atributos a partir de un subconjunto de objetos y viceversa. Es decir, dado $X \subseteq B$, $X^{\uparrow}$ contiene a los atributos que están relacionados con todos los objetos de X y, dado $Y \subseteq A$, $Y^{\downarrow}$ contiene a los objetos que están relacionados con todos los atributos de Y.

Nuestro estudio se centra en las implicaciones de atributos, que son expresiones $C \Rightarrow D$ donde $C, D \subseteq A$, C es el antecedente y D el consecuente. Una caracterización útil de las implicaciones de atributos válidas es aquella que considera los operadores de derivación $(^{\uparrow},^{\downarrow})$ como sigue: $C \Rightarrow D$ es una implicación de atributos válida en $(A, B, \mathcal{R})$ sii $D \subseteq C^{\downarrow\uparrow}$.

Teoría de conjuntos rugosos

Los conjuntos de datos se representan en RST como tablas de decisión, que son tuplas $(U, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$ tales que U y $\mathcal{A}$ son conjuntos no vacíos de atributos y objetos, respectivamente, $\mathcal{A}_d = \mathcal{A} \cup \{d\}$ con $d \notin \mathcal{A}$, $\mathcal{V}_{\mathcal{A}_d} = \{V_a \mid a \in \mathcal{A}_d\}$, donde V_a es el conjunto de valores asociado con el atributo $a \in \mathcal{A}_d$ sobre U, y $\overline{\mathcal{A}_d} = \{\bar{a} \mid a \in \mathcal{A}_d, \bar{a}\colon U \to V_a\}$.

La estructura de las tablas de decisión donde se destaca un atributo permite una transformación directa de las mismas en un conjunto de reglas de decisión. Por tanto, las reglas de decisión permiten describir las tablas de decisión en términos lógicos, proporcionando otra perspectiva de las mismas. Estas reglas son expresiones $\Phi \to \Psi$ donde Φ se construye a partir de la conjunción de pares atributo valor de $\mathcal{A}$ y Ψ es un par atributo valor del atributo d. La probabilidad condicionada de que ocurra el consecuente dado el antecedente permite describir las reglas de decisión. Cuando esta probabilidad es 1 decimos que la regla de decisión es verdadera.

III. Implicaciones de atributos y reglas de decisión

Para relacionar FCA y RST, consideramos tablas de decisión booleanas, es decir, aquellas donde $V_a = \{0, 1\}$ para todo $a \in \mathcal{A}_d$. Esto permite obtener una tabla de decisión booleana a partir de un contexto dado y viceversa. Siguiendo el proceso dado en [3], dado un contexto $(A, B, \mathcal{R})$, es posible construir la tabla de decisión booleana $(B, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$, definiendo el conjunto $\overline{\mathcal{A}_d}$ usando las aplicaciones $\bar{a}\colon B \to \{0, 1\}$ definidas, para cada $a \in \mathcal{A}_d$ y$b \in B$, como $\bar{a}(b) = 1$ si $a\mathcal{R}b$ y $\bar{a}(b) = 0$ en otro caso. Recíprocamente, dada una tabla de decisión booleana $(U, \mathcal{A}_d, \{0, 1\}, \overline{\mathcal{A}_d})$ el contexto $(\mathcal{A}_d, U, R)$ se obtiene definiendo la relación R como $a\mathcal{R}b$, si $\bar{a}(b) = 1$, para todo $a \in \mathcal{A}_d$ y $b \in U$. Teniendo en cuenta este mecanismo, se obtiene la siguiente relación:

$\Phi \to \Psi$ es una regla de decisión verdadera si y solo si $C \Rightarrow D$ es una implicación de atributos válida en $(A, B, \mathcal{R})$,

donde Φ es la conjunción de pares atributo valor de todos los atributos de C, Ψ es el par atributo valor del atributo d, y $\Phi \to \Psi$ es una 1-regla de decisión, es decir todos los atributos presentes en la regla toman el valor 1, y la satisface al menos un objeto $x \in B$.

Por otro lado, también es posible obtener la misma equivalencia sobre la base de las hipótesis de $C^{\downarrow} \neq \varnothing$, puesto que tiene el mismo significado que el antecedente de una regla de decisión cuyos atributos toman el valor 1 y la satisface al menos un objeto de la tabla de decisión. El entorno difuso proporcionará un marco de trabajo más flexible. Sin embargo, la interpretación semántica de ambas nociones es muy diferente, y se requieren distintas adaptaciones.

En las siguientes nociones, consideramos un contexto difuso $(A, B, \mathcal{R})$, una implicación residuada $\to$ definida sobre un retículo completo $(L, \preceq)$ [13], y la tabla de decisión asociada $S = (B, \mathcal{A}_d, \mathcal{V}_{\mathcal{A}_d}, \overline{\mathcal{A}_d})$, donde $\mathcal{A}_d = A$, $V_a = L$, para todo $a \in \mathcal{A}_d$, y $\overline{\mathcal{A}_d}$ es dado por $\mathcal{R}$.

Definición 1. Sea $Y \subseteq \mathcal{A}$. Dado $\Phi \in For(Y)$, con $\Phi = (a, v)$, la aplicación $\|\Phi\|^u_S\colon B \to [0, 1]$ se define como $\|\Phi\|^u_S(x)$ igual a 1 si $v \leq \bar{a}(x)$ y 0, en otro caso, para todo $x \in U$. El resto de fórmulas se definen inductivamente.

Teniendo en cuenta el punto de vista implicativo considerado, por ejemplo, para calcular la validez de un resumen lingüístico implicativo [6], surgen las siguientes nociones.

Definición 2. Sea $\Phi \to \Psi$ una regla de decisión S. Llamamos:

- u-*soporte* de $\Phi \to \Psi$ al valor:

$$supp^u_S(\Phi, \Psi) = \sum_{x \in U} \|\Phi\|^u_S(x) \to \|\Psi\|^u_S(x)$$

- u-*fuerza* de $\Phi \to \Psi$ al valor:

$$\sigma^u_S(\Phi, \Psi) = \frac{supp^u_S(\Phi, \Psi)}{|B|}$$

La noción de u-fuerza de una regla de decisión (difusa) generaliza la noción clásica de implicación de atributos válida.

Teorema 3. *Dado $L = \{0, 1\}$, se verifica que $C \Rightarrow D$ es válida en el contexto $(A, B, \mathcal{R})$ si y solo si la u-fuerza de la regla de decisión $\Phi \to \Psi$ es 1, donde Φ es la conjunción de pares atributo valor de los atributos de C, Ψ es el par atributo valor del atributo d, y $\Phi \to \Psi$ es una 1-regla de decisión.*

Estas relaciones y adaptaciones se continuarán estudiando en profundidad en futuras extensiones de este trabajo.

IV. Conclusiones y trabajo futuro

Este trabajo preliminar continúa analizando la relación entre RST y FCA [3], [4], en este caso, comparando las nociones de regla de decisión en RST e implicación de atributos en FCA. Además de los resultados en el caso clásico, se ha introducido una nueva interpretación de una regla de decisión en el caso difuso, presentando nuevas nociones de soporte y fuerza y comparándolas con la definición clásica de validez de una implicación de atributos. Como trabajo futuro, estas nociones se extenderán para capturar la noción de validez de una implicación de atributos en el marco difuso [2], [7]. Además, se tendrán en consideración otras nociones relacionadas con las implicaciones entre atributos, tales como reglas de asociación, algoritmo de reglas de decisión, eficiencia, etc.

Referencias

[1] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Reducing concept lattices by means of a weaker notion of congruence. *Fuzzy Sets and Systems*, 418:153–169, 2021. Algebra.

[2] R. Belohlávek and V. Vychodil. Attribute dependencies for data with grades I,. *International Journal of General Systems*, 45(7-8):864–888, 2016.

[3] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis. *Fuzzy Sets and Systems*, 2019.

[4] M. J. Benìtez-Caballero, J. Medina, and E. Ramìrez-Poussa. Attribute reduction in rough set theory and formal concept analysis. *Lecture Notes in Computer Science*, 10314:513–525, 2017.

[5] F. Chacón-Gómez, M. E. Cornejo, and J. Medina. Relating decision rules and attribute implications. In *The 16th International Conference on Concept Lattices and Their Applications (CLA 2022)*, 2022. In press.

[6] M. E. Cornejo and J. Medina. Implicative linguistic summaries. *Studies in Computational Intelligence*, 955:21–28, 2022.

[7] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Computing the validity of attribute implications in multi-adjoint concept lattices. In *International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE 2016*, volume II, pages 414–423, 2016.

[8] C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.

[9] D. Dubois, J. Medina, H. Prade, and E. Ramírez-Poussa. Disjunctive attribute dependencies in formal concept analysis under the epistemic view of formal contexts. *Information Sciences*, 2021.

[10] B. Ganter, G. Stumme, and R. Wille, editors. *Formal Concept Analysis: Foundations and Applications*. Lecture Notes in Computer Science. Springer Verlag, 2005.

[11] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.

[12] J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.

[13] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets and Systems*, 160(2):130–144, 2009.

[14] Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.

[15] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992.

On lattice orders on hesitant fuzzy sets

1st Pascual Jara
Departamento de Álgebra
Universidad de Granada
Granada, España
pjara@ugr.es

2nd Luis Merino
Departamento de Álgebra
Universidad de Granada
Granada, España
lmerino@ugr.es

3nd Gabriel Navarro
Departamento de CCIA
Universidad de Granada
Granada, España
gnavarro@ugr.es

4rd Evangelina Santos
Departamento de Álgebra
Universidad de Granada
Granada, España
esantos@ugr.es

***Abstract*—In this communication we show a non-trivial partial order on the class of non-empty subsets in $[0,1]$ that yields a lattice structure compatible with the standard lattice structure on $[0,1]$. As a consequence, the class of all set-valued (or hesitant) fuzzy sets is endowed with a lattice structure whose restriction to (type-1) fuzzy sets yields the classical Zadeh's operations.**

***Index Terms*—Fuzzy sets, interval-valued fuzzy sets, set-valued fuzzy sets, hesitant fuzzy sets, order, lattice.**

I. Introduction and preliminaires

As it was pointed out by Zadeh in [9], the estimation of the membership degrees of fuzzy sets (FS) is a common obstacle in practice. In order to ease this issue, a significant number of extensions have been defined. A prominent example is given by Grattan-Guinness in [2] introducing the notion of set-valued fuzzy sets (SVFSs). Then a natural question is how these membership degrees of the same kind can be compared, that is to say, how they can be endowed with a lattice structure compatible with Zadeh's operations on FSs, or on interval-valued fuzzy sets (IVFSs). Since Grattan-Guineness operations do not satisfy this propierty, Torra in [7] intents to mend this gap by endowing the so-called hesitant fuzzy sets (HFSs) with new operations. HFSs have attracted the attention of a great number of researchers because their applications to real world problems, mainly on information fusion and decision-making [4]. Unfortunately, these do not endow HFSs with a lattice structure, so the question of extending min-max operations to SVFSs remained open. In this communication, a summary of the unpublished work [3], we deal with this problem, stated in [1], and show a lattice structure on SVFSs (or HFSs) extending the Zadeh's operations on FSs and IVFSs.

We recall that a poset is a set P endowed with a relation $\leq$ satisfying reflexivity, antisymmetry and transitivity. The poset P is said to be *bounded* if there exist elements $0, 1 \in P$ such that $0 \leq p$ and $p \leq 1$ for all $p \in P$. A lattice is a set L with two operations $\wedge$ and $\vee$, called infimum and supremum, respectively, verifying associativity, commutativity, idempotency and absorption law. The relationship between both notions is given by the relations $a \leq b$ if and only if $a \vee b = b$ if and only if $a \wedge b = a$. Note then that every lattice is a poset with the former order. Nevertheless, not every poset defines a lattice.

If L is a bounded lattice with operations $\wedge_L$ and $\vee_L$, the class of all L-fuzzy subsets, over a certain universe X, becomes also a bounded lattice with point-wise operations $\sqcap$ and $\sqcup$ defined as $(f \sqcap g)(x) = f(x) \wedge_L g(x)$ and $(f \sqcup g)(x) = f(x) \vee_L g(x)$ for all $x \in X$, where f and g are L-fuzzy sets. For instance, whenever $L = [0,1]$ with the usual min-max lattice structure, we obtain the well-known operations on FSs defined by Zadeh [8]. Similarly, Grattan-Guinness' operations on SVFSs are point-wise extensions of the union and intersection in $\mathcal{P}([0,1])^*$, the class of non-empty subsets in $[0,1]$.

II. Some partial results

We recall that Torra's operations on HFSs are point-wise derived from the operations on $\mathcal{P}([0,1])$, the powerset of $[0,1]$, given by

$$A \cup_H B = \{t \in A \cup B \mid t \geq \mathrm{Inf}(A) \vee \mathrm{Inf}(B)\} \text{ and}$$

$$A \cap_H B = \{t \in A \cup B \mid t \leq \mathrm{Sup}(A) \wedge \mathrm{Sup}(B)\}$$

for any subsets $A, B \subseteq [0,1]$. Unfortunately, it is well-known, see [1] or [6], that they do not provide a lattice structure on $\mathcal{P}([0,1])$. In [6], a partial solution is given. The operation $\cap_H$ is changed by the following one: given $A, B \subseteq [0,1]$,

$$A \cap_C B = \begin{cases} A \cap ([\underline{A}, \underline{B}] \cup B) & \text{if } \underline{A} \leq \underline{B}, \\ B \cap ([\underline{B}, \underline{A}] \cup A) & \text{if } \underline{B} \leq \underline{A}. \end{cases} \quad (1)$$

where $\underline{A}$ and $\underline{B}$ denotes the infimum of A and B, respectively. Nevertheless, $\cup_H$ and $\cap_S$ do not provide a lattice structure on the whole class $\mathcal{P}([0,1])$, but only on the subclass formed by all the closed subsets, yielding the so-called closed-valued fuzzy sets.

Another attempt can be found in [3]. Given a set $A \in \mathcal{P}([0,1])$, we denote by A^- and A^+ the infimum and supremum of A, respectively. We denote by $\mathcal{P}^+([0,1])$ (respectively $\mathcal{P}^-([0,1])$) the class of all right-closed (left-closed) subsets in $[0,1]$, i.e., those for what $A^+ \in A$ ($A^- \in A$). Note that $\mathcal{P}^+([0,1])$ and $\mathcal{P}^-([0,1])$ contain any finite subset in $[0,1]$,

The authors are listed in alphabetical order. All the authors contributed equally to this work. P. Jara, L. Merino and E. Santos are with IMAG and the Department of Algebra, University of Granada. G. Navarro is with CITIC and the Department of Computer Science and Artificial Intelligence, University of Granada.

This work is supported by grant A-FQM-394-UGR20 from Programa Operativo FEDER 2014-2020 and Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía (Spain) and by the "María de Maeztu" Excellence Unit IMAG, reference CEX2020-001105-M, funded by MCIN/AEI/10.13039/501100011033/.

any closed interval in $[0,1]$ and any finite union of them. For simplicity, given two subsets $X, Y \in \mathcal{P}([0,1])$ we say that $X < Y$ if, for all $x \in X$ and all $y \in Y$, $x < y$. Let us define the relations $\leq^r$ on $\mathcal{P}^+([0,1])$ and $\leq^l$ on $\mathcal{P}^-([0,1])$ as follows. Given two subsets $A, B \in \mathcal{P}^+([0,1])$, $A \leq^r B$ if and only if $A^+ \leq B^+$ and $A < B \setminus A$. Similarly, given two subsets $A, B \in \mathcal{P}^-([0,1])$, $A \leq^l B$ if and only if $A^- \leq B^-$ and $A \setminus B < B$.

The relations $\leq^r$ and $\leq^l$ are orders on $\mathcal{P}^+([0,1])$ and $\mathcal{P}^-([0,1])$ respectively, whose minimum and maximum are $\{0\}$ and $\{1\}$. Additionally,

Proposition 1: The class $\mathcal{P}^+([0,1])$ is endowed with a structure of bounded lattice for the order $\leq^r$ given by the operations

$$A \wedge_r B = \{x \in A \cup B \mid x \leq A^+ \wedge B^+\} = A \cap_H B, \text{ and}$$

$$A \vee_r B = (A \cap B) \cup \{x \in A \cup B \mid x > A^+ \wedge B^+\}.$$

Proposition 2: The class $\mathcal{P}^-([0,1])$ is endowed with a structure of bounded lattice for the order $\leq^l$ given by

$$A \wedge_l B = (A \cap B) \cup \{x \in A \cup B \mid x < A^- \vee B^-\}, \text{ and}$$

$$A \vee_l B = \{x \in A \cup B \mid x \geq A^- \vee B^-\} = A \cup_H B.$$

In addition, both orders, $\leq^r$ and $\leq^l$, restricted to the class $\mathbb{I}$ of all closed intervals in $[0,1]$ equals the usual order on $\mathbb{I}$.

III. A LATTICE STRUCTURE ON HESITANT FUZZY SETS

Let us now give an answer to the problem stated in [1, Remark 2]. Denote by $\mathcal{P}^*([0,1])$ the class of all non-empty subsets in $[0,1]$. We define the relation $\leq^0$ on $\mathcal{P}^*([0,1])$ as follows: for any $A, B \in \mathcal{P}^*([0,1])$, $A \leq^0 B$ if and only if $A < B \setminus A$ and,$A \setminus B < B$.

Proposition 3: [3] $(\mathcal{P}^*([0,1]), \leq^0)$ is a partially ordered set.

Given $A, B \in \mathcal{P}^*([0,1])$, we shall denote by $T_{A,B}$ the subset $T_{A,B} = \{t \in A \cup B \setminus A \cap B$ such that $t > A^- \vee B^-\}$, and by $t_{A,B}$ the infimum of the subset $T_{A,B}$. We consider the subsets $X_{A,B} = \{x \in A \cup B \mid x < t_{A,B}\}$ and $\overline{X}_{A,B} = \{x \in A \cup B \mid x \leq t_{A,B}\}$.

The order $\leq^0$ provides a lattice structure on $\mathcal{P}^*([0,1])$. The rules for computing the infimum $\wedge_0$ goes as described here

Theorem 4: [3] Let $A, B \in \mathcal{P}^*([0,1])$ such that $A^- \leq B^-$.

A. If $A^- = B^- = t_{A,B}$, then $A \wedge_0 B = \{B^-\}$.

B. If $A^- < B^-$ and $B^- \in B \setminus A$, then

$$A \wedge_0 B = \{a \in A \mid x < B^-\}$$

F. If $A^- < B^-$, $B^- \notin B \setminus A$ and $B^- = t_{A,B} \in A \setminus B$, then $A \wedge_0 B = \overline{X}_{A,B}$.

D. Otherwise, $A \wedge_0 B = \begin{cases} \overline{X}_{A,B}, & \text{if } t_{A,B} \in A \cap B, \\ X_{A,B}, & \text{if } t_{A,B} \notin A \cap B. \end{cases}$

Note that we are assuming $A^- \leq B^-$. But, otherwise, simply exchange A by B and vice versa. The supremum $\vee_0$ can be computed dually, simply denotes $S_{A,B} = \{s \in A \cup B \setminus A \cap B$ such that $s < A^+ \wedge B^+\}$, and by $s_{A,B}$ its supremum. Let also $Y_{A,B} = \{y \in A \cup B \mid s_{A,B} < y\}$ and $\overline{Y}_{A,B} = \{y \in A \cup B \mid s_{A,B} \leq y\}$.

Theorem 5: [3] Let $A, B \in \mathcal{P}^*([0,1])$ such that $A^+ \leq B^+$.

A*. If $A^+ = B^+ = s_{A,B}$, then $A \vee_0 B = \{A^+\}$.

B*. If $A^+ < B^+$ and $A^+ \in A \setminus B$, then

$$A \vee_0 B = \{b \in B \mid A^+ < b\}.$$

C*. If $A^+ < B^+$, $A^+ \notin A \setminus B$ and $A^+ = s_{A,B} \in B \setminus A$, then $A \vee_0 B = \overline{Y}_{A,B}$.

D*. Otherwise, $A \vee_0 B = \begin{cases} \overline{Y}_{A,B}, & \text{if } s_{A,B} \in A \cap B, \\ Y_{A,B}, & \text{if } s_{A,B} \notin A \cap B. \end{cases}$

Note that the order $\leq^0$ extends the usual order on $[0,1]$ and on $\mathbb{I}$, the class of closed intervals in $[0,1]$, to the class $\mathcal{P}^*([0,1])$. Concretely, the former is defined as $[a,b] \wedge [c,d] = [\min(a,c), \min(b,d)]$ and $[a,b] \vee [c,d] = [\max(a,c), \max(b,d)]$ for each pair of intervals $[a,b]$ and $[c,d]$. Hence

Corollary 6: The lattices $([0,1], \min, \max)$, $(\mathbb{I}, \wedge, \vee)$ and $(\mathcal{P}^*([0,1]), \wedge_0, \vee_0)$ forms a chain of lattices.

Given a universe X, we may then obtain the point-wise lattice structure $(\mathrm{SVFS}(X), \sqcap_0, \sqcup_0)$ compatible with those of $(\mathrm{FS}(X), \min, \max)$ and $(\mathrm{IVFS}(X), \sqcap, \sqcup)$.

Theorem 7: [3] The restrictions of the lattice $(\mathrm{SVFS}(X), \sqcap_0, \sqcup_0)$ to the subclasses $\mathrm{FS}(X)$ and $\mathrm{IVFS}(X)$ yield the lattices $(\mathrm{FS}(X), \min, \max)$ and $(\mathrm{IVFS}(X), \sqcap, \sqcup)$, respectively.

IV. CONCLUSION

The necessity of dealing with uncertainty in real world problems has originated several extensions of Zadeh's fuzzy sets. Nevertheless, the use of these extensions simultaneously requires of certain compatibility between the operations associated to them. We have solved the problem of the seeing, as lattices, FSs and IVFSs inside SVFSs, or equivalently, HFSs, as stated in [1]. This allows also to answer a challenge given in [5] for HFSs (Challenge 5.3), and gives the possibility of treating properly some of the others challenges which we left as future works.

REFERENCES

[1] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, and B. De Baets, "A historical account of types of fuzzy sets and their relationships", IEEE Transactions on Fuzzy Systems 24 (1) (2016), 179–194.

[2] I. Grattan-Guinness, "Fuzzy membership mapped onto interval and many-valued quantities", Zeitschrift für mathematische Logik und Grundladen der Mathematik 22 (1976), 149–160.

[3] P. Jara, L. Merino, G. Navarro and E. Santos, "A lattice structure on hesitant fuzzy sets", submitted, 2022.

[4] R.M. Rodríguez, L. Martínez, V. Torra, Z.S. Xu, F. Herrera "Hesitant fuzzy sets: State of the art and future directions" Int. J. Intell. Syst., 29 (6) (2014), pp. 495-524.

[5] R.M. Rodríguez, B. Bedregal, H. Bustince, Y.C. Dong, B. Farhadinia, C. Kahraman, L. Martínez, V. Torra, Y.J. Xu, Z.S. Xu, F. Herrera, A position and perspective analysis of hesitant fuzzy sets on information fusion in decision making. Towards high quality progress, Information Fusion 29 (2016) 89–97.

[6] F.J. Lobillo, L. Merino, G. Navarro, E. Santos, "Embeddings between lattices of fuzzy sets: An application of closed-valued fuzzy sets", Fuzzy Sets and Systems 352 (2018), 56–72.

[7] V. Torra, "Hesitant fuzzy sets", International Journal of Intelligent Systems 25 (2010), 529–539.

[8] L. A. Zadeh, "Fuzzy sets", Inf. Control 3 (1965), 338–353.

[9] L. A. Zadeh, "Quantitative fuzzy semantics", Information Sciences 3 (1971), 159–176.

Uso de Lógica Difusa para Construir un Sistema Recomendador Médico

Pablo Cordero, Manuel Enciso, Domingo López-Rodríguez, Ángel Mora
Universidad de Málaga
Andalucía Tech
Málaga, Spain
{pcordero,enciso,dominlopez,amora}@uma.es

***Resumen*—En este trabajo, se propone un motor automatizado basado en la Lógica de Simplificación difusa para realizar sugerencias a los usuarios. Los sistemas de recomendación conversacional han demostrado ser un buen enfoque en telemedicina, construyendo un diálogo entre el usuario y el recomendador basado en las preferencias del usuario proporcionadas en cada paso de la conversación. Aquí proponemos un sistema de recomendación conversacional para el diagnóstico médico utilizando la lógica difusa.**

***Index Terms*—Recomendación, diagnóstico, lógica difusa, simplificación**

I. Lógica de Simplificación para Razonamiento Automático

Este trabajo es un resumen del aparecido en [1], y se basa en la variante difusa del Análisis Formal de Conceptos (FCA) [2], que formaliza el conjunto de datos como una relación difusa/graduada entre objetos y atributos. Una de las dos formas de representar el conocimiento son las reglas, también llamadas implicaciones, que pueden obtenerse automáticamente del conjunto de datos. En la versión difusa, las implicaciones de los atributos son fórmulas $A \Rightarrow B$ donde A y B son conjuntos difusos sobre un conjunto de atributos M y, de manera informal, una implicación como $\{a, {}^{0.5}/b\} \Rightarrow \{{}^{0.9}/c\}$ significa que cada objeto que tiene el atributo a con grado 1 (es decir, que posee completamente a), y el atributo b con grado 0.5, entonces tiene el atributo c con grado al menos 0.9.

Una base de implicaciones asociada a un conjunto de datos es un conjunto de implicaciones que permite derivar todas las implicaciones que se satisfacen en el conjunto de datos utilizando reglas de inferencia lógica. Más concretamente, en nuestro trabajo, el conocimiento contenido en el conjunto de datos se formaliza como una base de implicaciones con grados sobre las que razonaremos utilizando la *Lógica de simplificación de atributos difusos* (FASL) [3]. Veremos que el método automatizado basado en esta lógica nos permite llegar a una recomendación.

A continuación, presentamos brevemente FASL. Consideremos un retículo completo residuado $\langle L, \vee, \wedge, \otimes, \to, 0, 1\rangle$ donde $L = \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$ es una discretización del intervalo unidad, $\otimes$ es una t-norma continua a la izquierda (como las de Łukasiewicz o Gödel), y $\to$ es su implicación residuada. Con estas condiciones, un conjunto de datos (un contexto formal difuso en la terminología FCA) es una tupla $\mathbb{K} = \langle G, M, I\rangle$ donde G y M son conjuntos de objetos y atributos respectivamente, y $I \in L^{G\times M}$ es la relación de incidencia que es una relación difusa/gradual entre objetos y atributos. Dado un conjunto de datos $\mathbb{K}$ y dos subconjuntos difusos de atributos $A, B \in L^M$, decimos que $A \Rightarrow B$ es (totalmente) verdadera en $\mathbb{K}$, se denota como $\mathbb{K} \models A \Rightarrow B$, cuando la siguiente propiedad se mantiene para todo $x \in G$: $\bigwedge_{y\in M}(A(y) \to I(x, y)) \leqslant \bigwedge_{y\in M}(B(y) \to I(x, y))$.

El sistema axiomático de FASL se define como sigue: para todo $A, B, C, D \in L^M$ y $c \in L$,

- [Ax] Inferir $A \cup B \Rightarrow A$ (*Axioma*)
- [Mul] De $A \Rightarrow B$, inferir $c\otimes A \Rightarrow c\otimes B$ (*Multiplicación*)
- [Sim] De $A \Rightarrow B$ y $C \Rightarrow D$, inferir $A \cup (C \smallsetminus B) \Rightarrow D$ (*Simplificación*)

donde las operaciones en conjuntos difusos se entienden definidas elemento a elemento. La corrección y completitud de esta lógica están aseguradas si asumimos que M es finito. Además, las reglas de inferencia en FASL proporcionan equivalencias que permiten simplificar los conjuntos de implicaciones: para cualquier $A, B, C, D \in L^M$,

- (**DeEq**) $\{A \Rightarrow B\} \equiv \{A \Rightarrow B \smallsetminus A\}$;
- (**UnEq**) $\{A \Rightarrow B, A \Rightarrow C\} \equiv \{A \Rightarrow B \cup C\}$;
- (**SiEq**) Si $A \subseteq C$ entonces $\{A \Rightarrow B, C \Rightarrow D\} \equiv \{A \Rightarrow B, A \cup (C \smallsetminus B) \Rightarrow D \smallsetminus B\}$.

Las siguientes nociones son fundamentales para los resultados de este trabajo. Dada una base de implicaciones Σ y un conjunto difuso de atributos $A \in L^M$, el *cierre de* A (*con respecto a* Σ), denotado por A^+, se define como el mayor conjunto difuso en M tal que $A \Rightarrow A^+$ se puede derivar de Σ usando las reglas de inferencia. A se llama Σ*-cerrado* si $A^+ = A$.

Teorema 1: Si Σ es una base de implicaciones y $A, B \in L^M$, entonces $A \Rightarrow B$ es válida si y sólo si $B \subseteq A^+$.

En [3], basándonos en estos resultados, se propuso un nuevo método de razonamiento automático para implicaciones con atributos difusos, útil para resolver los problemas clásicos de cálculo de cierres y de derivación sintáctica. El método utiliza las equivalencias anteriores y sustituye las fórmulas (implica-

Supported by Grants TIN2017-89023-P, UMA2018-FEDERJA-001 and PGC2018-095869-B-I00 of the Junta de Andalucia, and European Social Fund.

ciones) por otras equivalentes pero más simples, superando los inconvenientes de otras métodos. Como se ha demostrado en la evaluación experimental, los métodos son factibles desde el punto de vista informático en la misma medida que los métodos clásicos.

II. Sistema de Recomendación Conversacional

A continuación, a modo de ilustración, describimos nuestro sistema de recomendación en el ámbito del diagnóstico médico. En primer lugar, tomaremos un conjunto de datos (contexto formal difuso) en el que los pacientes son los objetos y los atributos pueden ser los síntomas (elementos introducidos en el diálogo) o las enfermedades (los elementos a recomendar). La información original puede ser extremadamente personalizada, ya que utilizamos un enfoque multivaluado y se puede asignar un grado a cada síntoma para cada paciente.

En resumen, nuestro método funciona de la siguiente manera: en cada paso de la conversación, el usuario interactúa con el sistema proporcionando nuevos síntomas y el algoritmo aplica de forma iterativa el operador de cierre difuso para enriquecer el conjunto de síntomas hasta incluir una columna de enfermedad en el conjunto cerrado de atributos, terminando con éxito la conversación proporcionando un diagnóstico como recomendación.

La aplicación del operador de cierre proporciona un conjunto limitado de síntomas, fuertemente relacionados con la conversación en su etapa actual, y también estrecha el espacio de búsqueda guiando así los siguientes pasos de la conversación.

El proceso de conversación se describe con los siguientes pasos, partiendo de la base de implicaciones Σ:

1. El sistema le pide al usuario un síntoma y su grado asociado: $(d_x \mid x)$ donde $x \in M$ y $d_x \in L$.
2. Se calcula el cierre $(d_x \mid x)^+$ con respecto a Σ y el conjunto reducido de implicaciones, que se vuelve a almacenar como Σ.
3. Si el cierre contiene un atributo que identifica una enfermedad, entonces se ha producido un diagnóstico. El sistema detiene el proceso y proporciona la enfermedad como recomendación.
4. En caso contrario, el recomendador pide una retroalimentación al usuario. Los síntomas incluidos en el conjunto $(d_x \mid x)^+ - \{(d_x \mid x)\}$ se muestran al usuario, dándole la oportunidad de aumentar sus grados. Si el usuario desea aumentar alguna de ellas, se inicia un nuevo ciclo del diálogo, pasando al Paso 2. Esta retroalimentación constituye una depuración de la elicitación proporcionada por el usuario.
5. Si el usuario rehúsa dar una respuesta, estando de acuerdo con la información proporcionada, entonces hay que introducir nuevos síntomas para continuar con la conversación, pasando al Paso 1.

III. Caso de Uso: Sistema Diagnóstico

En los últimos años, ha aparecido un número creciente de iniciativas para compartir, organizar y estudiar ciertas patologías cerebrales prevalentes. Hemos recopilado un conjunto de datos de fuentes públicas [4] y curadas con 105 pacientes (objetos en terminología FCA) y 55 atributos como siguen: 53 atributos relacionados con los signos o síntomas, y 2 atributos relacionados con el diagnóstico. Los atributos de los síntomas son multivaluados: para un determinado atributo (síntoma), los grados disponibles van desde *ausente* hasta *extremo*, pasando por *mínimo*, *leve*, *moderado*, *grave* y *severo*. Por lo tanto, todos los atributos pueden considerarse difusos y graduados.

Tabla I
Comparación del sistema recomendador propuesto con otros sistemas usuales y con métodos de aprendizaje automático.

	Prec.	Sens.	Espec.	Val. Pred. Pos.
ALS	0.360	0.333	0.380	0.290
IBCF (Cosine)	0.555	0.475	0.615	0.483
IBCF (Pearson)	0.770	0.466	1.000	1.000
LIBMF	0.491	0.901	0.181	0.455
SVD	0.376	0.515	0.271	0.349
SVDF	0.431	1.000	0.000	0.431
UBCF (Cosine)	0.608	0.967	0.335	0.524
UBCF (Pearson)	0.525	0.783	0.330	0.470
C5.0	0.674	0.636	1.000	1.000
PART	0.883	0.847	0.950	0.970
JRip	0.752	0.814	0.688	0.731
Random Forest	0.953	0.924	1.000	1.000
xgboost	0.818	0.963	0.713	0.706
k-nn	0.589	0.603	0.544	0.815
Proposal	0.982	0.996	0.948	0.955

A partir del dataset, usando el conocido algoritmo Next-Closure para atributos difusos [2], se construye la base canónica de implicaciones del dataset. En este experimento, la base consta de 20663 implicaciones, de las cuales proporcionan información útil para nuestros propósitos 15700 (aquellas con soporte positivo).

Se generó un conjunto de 1000 nuevas observaciones siguiendo la misma distribución estadística de los datos originales. En los experimentos, se ha comparado nuestra propuesta junto a otras alternativas conocidas, como sistemas recomendadores basados en filtrado colaborativo, o métodos de aprendizaje automático tales como random forest o reglas de asociación. Las medidas de calidad que usamos son la precisión, sensibilidad, especificidad y valor predictivo positivo de cada método, destacando que nuestra propuesta es capaz de competir y mejorar los resultados de los esquemas que se pueden considerar el estado del arte actual en esta temática.

Referencias

[1] P. Cordero, M. Enciso, D. López, and A. Mora, "A conversational recommender system for diagnosis using fuzzy rules," *Expert Systems with Applications*, vol. 154, p. 113449, 2020.

[2] R. Belohlavek, "Algorithms for fuzzy concept lattices," ser. Proc. Fourth Int. Conf. on Recent Advances in Soft Computing. Nottingham, United Kingdom, 2002, pp. 200–205.

[3] R. Belohlavek, P. Cordero, M. Enciso, Á. Mora, and V. Vychodil, "Automated prover for attribute dependencies in data with grades," *International Journal of Approximate Reasoning*, vol. 70, pp. 51–67, 2016.

[4] C. Aine, H. J. Bockholt, J. R. Bustillo, J. M. Cañive, A. Caprihan, C. Gasparovic, F. M. Hanlon, J. M. Houck, R. E. Jung, J. Lauriello *et al.*, "Multimodal neuroimaging in schizophrenia: description and dissemination," *Neuroinformatics*, vol. 15, no. 4, pp. 343–364, 2017.

Community detection in social networks and fuzzy measures: analyzing additional information*

1st Inmaculada Gutiérrez
Estadística y Ciencia de los Datos
Facultad de Estudios Estadísticos (UCM)
Madrid, Spain
inmaguti@ucm.es

1st Daniel Gómez
Estadística y Ciencia de los Datos
Facultad de Estudios Estadísticos (UCM)
Instituto de evaluación sanitaria (UCM)
Madrid, Spain
dagomez@estad.ucm.es

1st Javier Castro
Estadística y Ciencia de los Datos
Facultad de Estudios Estadísticos (UCM)
Instituto de evaluación sanitaria (UCM)
Madrid, Spain
jcastroc@estad.ucm.es

1st Rosa Espínola
Estadística y Ciencia de los Datos
Facultad de Estudios Estadísticos (UCM)
Instituto de evaluación sanitaria (UCM)
Madrid, Spain
rosaev@estad.ucm.es

***Abstract*—Community detection problems are one of the most important topics in social network analysis. Most of the algorithms and techniques to find communities in a network, model and represent is as something crisp. However, there exist many real situations in which fuzzy uncertainty appears in a natural way when the network is modeled. In this work, we present a methodology, based on a modification of the well-known Louvain algorithm, that allows us to deal with fuzzy information in networks. The goal is to obtain groups of individuals as realistic as possible, not only considering their crisp connections, but also the synergies and affinities among them.**

***Index Terms*—Extended fuzzy graph, Networks, Community detection, Fuzzy measures**

I. Introduction

In this work we present a review of the paper entitled *A new community detection algorithm based on fuzzy measures'*, presented in the conference INFUS-2019 and published in *Advances in Intelligent Systems and Computing series* [1]. In that preliminary work we set the basis of a novel approach to community detection in social networks [2] considering additional information modeled by fuzzy measures [3].

Problems about Social Network Analysis (SNA) include a set of methodological approaches based on a theoretical frame whose main goal is to analyze the relationships between the involved agents, assuming that they may represent individual people, countries, users in an online site, organizations or just ordinary things. Pondering the interactions between the employees of a large company, your personal connections beyond one degree of separation, the synergies within various animal communities, or the liaison between websites, is essentially an intuitive depiction of what a Social Network is.

Most of these problems are modeled with networks or graphs. A network is a set of nodes or vertices which represent the individuals, and a set of edges or links that represent the liaisons, relations, connections or synergies between them.

This work focuses on the study of the group organization of the individuals of a network. A peculiarity when modeling real situations is that the individuals are usually organized into modular structures. In this type of partitions, whose modules are commonly known as communities, the elements of the same group are highly connected to each other (there are many links between them), whereas hardly any connections between the individuals of different groups can be found.

Classical algorithms to solve the community detection problem in graphs are based on the structure defined by the edges [2], [4]. Nevertheless, it is fair to acknowledge the need to add some extra information besides the connections already displayed by the graph (edges) to increase the quality of the results. We agree that, the more information is analyzed, the more realistic is the solution obtained. Then, assuming that the extra knowledge available is about the relations between the players, regardless of the links that connect them, we work with fuzzy measures [3] (monotonic set functions such as possibility, plausability or capacity measures) to model it. Specifically, we worked in a modification of the well-known Louvain algorithm [5], one of the most popular methods in community detection, to incorporate some additional information. Our proposal is based on the existence of an extended fuzzy graph [6]. That novel idea has led to the development of several approaches, as those about bipolar behaviour [7], [8], on Sugeno measures [9], [10] or polarization measures [11].

II. Some important concepts

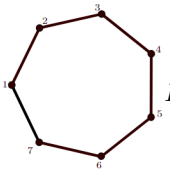
Our work is based on the existence of a crisp graph, $G = (V, E)$, where $V = \{1, \ldots, n\}$ is the set of nodes and $E = \{\{i, j\} : i, j \in V\}$ is the set of edges, and a fuzzy measure [3], a monotonic set function with $\mu(\emptyset) = 0$ and $\mu(V)$=1) $\mu : 2^V \to [0, 1]$ defined over V. Both together characterize the

Supported by the Government of Spain, Gran Plan Nacional de I+D+i PR108/20-28 and PGC2018096509-B-I00.

extended fuzzy graph $\widetilde{G} = (V, E, \mu)$ [1], [6], a representation tool which generalizes fuzzy graphs [12].

On the basis of $\widetilde{G}$, and assuming that μ provides some additional information about the synergies between the individuals of G, besides their connections in E, and useful to find realistic groups of elements, we propose a modification of the Louvain algorithm [5] to incorporate the information of μ to the grouping process. As the Louvain one, our proposal is a multi-phase, iterative heuristic for modularity optimization method [2]. The main difference is that the calculation of modularity, Q, not only considers the structure of G, but also the information of μ. To do so, we calculate the weighted graph associated to μ by means of the Shapley value [13] as $F^{\mu}_{ij} = \phi\left(Sh_i(\mu) - Sh^j_i(\mu), Sh_j(\mu) - Sh^i_j(\mu)\right)_{i,j\in V}$. Being ϕ an aggregation function and $Sh_i(\mu)$ and $Sh^j_i(\mu)$ the Shapley value of i when it is in a coalition with all the individuals of V or $V\backslash\{j\}$ respectively, this graph represents, for each pair of elements, how the absence of an individual of the pair affects the other in the coalition. Then, we combine the information of G with the information of μ, and we calculate the maximum of modularity taking into account both variability sources.

Specifically we work with affinity fuzzy measures which are 2-additives, so the calculation of the Shapley value is immediate. Hence we can affirm that our *Duo Louvain* algorithm, has equal complexity as the Louvain algorithm in these cases. Let us show a simple example of this. We consider the graph $G = (V, E)$ represented by the wheel below with 7 nodes, and the matrix F obtained by an affinity fuzzy measure μ; both together characterize the extended fuzzy graph $\widetilde{G} = (V, E, \mu)$ in which the first four nodes have a good affinity with each other. Any community detection algorithm, specifically those based on modularity optimization, establish a community with four consecutive random nodes, and another with three. That type solution does not fix the affinity modeled by μ and showed with F^{μ}. See how our proposal gathers this affinity



$$F^{\mu} = \begin{pmatrix} 0 & \mathbf{0.083} & \mathbf{0.083} & \mathbf{0.083} & 0 & 0 & 0 \\ \mathbf{0.083} & 0 & \mathbf{0.083} & \mathbf{0.083} & 0 & 0 & 0 \\ \mathbf{0.083} & \mathbf{0.083} & 0 & \mathbf{0.083} & 0 & 0 & 0 \\ \mathbf{0.083} & \mathbf{0.083} & \mathbf{0.083} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Fig. 1: Extended fuzzy graph $\widetilde{G} = (V, E, \mu)$.

TABLE I: Results of Louvain and Duo Louvain algorithms.

	Duo Louvain Algorithm				Louvain Algorithm			
α	Q_A	Q_F	$\alpha Q_A + (1-\alpha)Q_F$	Clusters	Q_A	Q_F	$\alpha Q_A + (1-\alpha)Q_F$	Clusters
1	0,225	-0,025	0,225	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$	0,225	-0,025	0,225	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$
0,9	0,204	0,444	0,228	$\{\{1,2,3,4\},\{5,6,7\}\}$	0,225	-0,025	0,200	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$
0,7	0,204	0,444	0,276	$\{\{1,2,3,4\},\{5,6,7\}\}$	0,225	-0,025	0,150	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$
0,5	0,204	0,444	0,324	$\{\{1,2,3,4\},\{5,6,7\}\}$	0,225	-0,025	0,100	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$
0,3	0,204	0,444	0,372	$\{\{1,2,3,4\},\{5,6,7\}\}$	0,225	-0,025	0,050	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$
0,1	0,204	0,444	0,420	$\{\{1,2,3,4\},\{5,6,7\}\}$	0,225	-0,025	0,000	$\{\{1,2,7\},\{3,4\},\{5,6\}\}$

III. Evaluation of the model

Although the proposal in [1] in was at a very preliminary stage, the excellent reception of the new methodology encouraged us to focus our efforts on evaluating and applying it. Then, in a subsequent paper [6], we showed some computational results based on benchmarking and NMI calculation which allow us to affirm the goodness of the *Duo Louvain* algorithm. We also demonstrate some desirable properties of the affinity fuzzy measures, including the consideration of the interaction index [14]. That evaluation process has been adapted to other scenarios, such as the bipolar case [8] or about the consideration of fuzzy Sugeno λ-measures [9], [10].

IV. Conclusions

In this work we review a work [1] about a modification of Louvain algorithm with which deal with community detection problem in networks in which there exists additional information independent of its structure.

Given a graph, $G = (N, E)$, we assume there exists some additional information about affinity among nodes, which is not inherent in the graph structure, defined by a fuzzy measure, μ. We work in the incorporation this additional information to Louvain algorithm in order to obtain communities consistent with all the available information, not only with the structure of the graph. For the particular case in which this fuzzy measure can be defined as an affinity matrix, the calculation has the same complexity than Louvain algorithm.

That methodology has been adapted to several scenarios [7], [9], evaluated [6], [8], [10] and applied in real cases [11].

References

[1] I. Gutiérrez, D. Gómez, J. Castro, and R. Espínola, "A new community detection algorithm based on fuzzy measures," in *Advances in Intelligent Systems and Computing series*, vol. 1029, 2020, pp. 133–140.

[2] M. Girvan and M. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 2, pp. 7821–7826, 2002.

[3] M. Sugeno, "Fuzzy measures and fuzzy integrals: A survey," *Fuzzy Automata and Decision Processes*, vol. 78, pp. 89–102, 1977.

[4] V. Traag, L. Waltman, and N. van Eck, "From Louvain to Leiden: guaranteeing well-connected communities," *Scientific Reports*, vol. 9, no. 5233, 2019.

[5] V. Blondel, J. Guillaume, R. Lambiotte, and E. Lefevre, "Fast unfolding of communities in large networks," *Journal Of Statistical Mechanics-Theory And Experiment*, vol. 10, no. P10008, 2008.

[6] I. Gutiérrez, D. Gómez, J. Castro, and R. Espínola, "Fuzzy measures: A solution to deal with community detection problems for networks with additional information," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 5, pp. 6217–6230, 2020.

[7] ——, "A new community detection problem based on bipolar fuzzy measures," *Studies in Computational Intelligence*, vol. 955, 2022.

[8] ——, "Multiple bipolar fuzzy measures: an application to community detection problems for networks with additional information," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1636–1649, 2020.

[9] ——, "Fuzzy sugeno λ-measures and theirs applications to community detection problems." Glasgow, UK: IEEE International Conference on Fuzzy Systems, 2020, pp. 1–6.

[10] ——, "Community detection in networks with multiple fuzzy sugeno $\lambda-$measures." FSTA, 2022.

[11] I. Gutiérrez, J. Guevara, D. Gómez, J. Castro, and R. Espínola, "Community detection problem based on polarization measures: An application to twitter: The covid-19 case in spain," *Mathematics*, vol. 9, no. 4, 2021.

[12] A. Rosenfeld, "Fuzzy graphs," *Fuzzy Sets and their Applications*, pp. 77–95, 1975.

[13] L. Shapley, "A value for n-person games," *Contribute. Theory Games*, vol. 2, no. 28, pp. 307–317, 1953.

[14] M. Grabisch, "k-order additive discrete fuzzy measures and their representation," *Fuzzy Sets Systems*, vol. 92, no. 2, pp. 167–189, 1997.

Un enfoque basado en prototipos deformables borrosos para la modelización de la evolución de la probabilidad de victoria en deportes de equipo

Francisco P. Romero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
FranciscoP.Romero@uclm.es

Agustín Mora-Acosta
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
Agustin.Mora1@alu.uclm.es

Eusebio Angulo
Departamento de Matemáticas
Universidad de Castilla-La Mancha
Ciudad Real, España
Eusebio.Angulo@uclm.es

Jesus Serrano-Guerrero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
Jesus.Serrano@uclm.es

Jose A. Olivas
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
JoseAngel.Olivas@uclm.es

Julio A. Lopez-Gómez
Tecnologías y Sistemas de Información
Universidad de Castilla La Mancha
Almadén, España
JulioAlberto.Lopez@uclm.es

***Resumen*—Dentro de la analítica deportiva cobran importancia aquellos métodos de evaluación del juego que se actualizan con el tiempo, ya sea evaluando los estados del juego, los eventos del juego o ambos. Este trabajo presenta un novedoso mecanismo para modelizar la evolución de la probabilidad de victoria utilizando prototipos deformables borrosos.**

***Index Terms*—Prototipos Deformables Borrosos, Probabilidad de victoria, Analítica del Deporte**

I. Introducción

Los modelos de probabilidad de victoria durante el juego (*in-game*) proporcionan la probabilidad de que un determinado equipo gane un partido en función de un estado específico del mismo (por ejemplo, el marcador, el tiempo restante, etc.). Estos modelos se han hecho cada vez más populares en diversos deportes en la última década. Existen diferentes métodos principalmente basados en técnicas de regresión logística y *boosting* que permiten realizar estimaciones de probabilidad de victoria antes del partido o durante el partido [1] [2]. El objetivo de este trabajo es la definición de un modelo que a partir de la evolución de las estadísticas y la probabilidad de victoria *in-game* de un equipo hasta un momento dado de un partido, sea capaz de realizar una predicción de la proyección de esta probabilidad de victoria *in-game* para el futuro cercano. Para ello se han utilizado datos procedentes de partidos de la NBA con información de lo sucedido jugada a jugada.

II. Modelo basado en Prototipos Deformables Borrosos

En primer lugar se propone hacer uso del método de Montecarlo para comprender cómo de probable es que un equipo gane el partido dada la situación actual de éste. Para ello se realizan un número de simulaciones de cómo va a evolucionar su probabilidad de victoria a partir de un determinado instante, tomando como referencia las características de cada equipo en ese momento. Se necesita, por lo tanto, una estimación de la probabilidad de que ocurra un evento u otro en la siguiente jugada en la que dicho equipo tiene la posesión. Con este fin se utilizan, en un primer momento, las estadísticas de cada equipo en ese instante (o reforzando este dato utilizando los históricos en la competición).

Uno de los problemas a la hora de simular las jugadas es establecer el tiempo de partido que va a consumir cada jugada, factor esencial para determinar el momento en el que la simulación debe de terminar. Con este fin se utiliza el tiempo promedio en segundos que han durado las jugadas realizadas por el equipo en posesión durante dicho partido antes del instante a partir del que estamos simulando. Cuando estas simulaciones son efectuadas a partir de la mitad del partido dan resultados muy coherentes, obteniendo simulaciones con escenarios finales muy factibles en cuanto a puntuación.

En la Figura 1 podemos ver la evolución de la probabilidad de victoria *in-game* para las 1500 simulaciones realizadas para un partido a partir de un instante, evoluciones que han sido ordenadas de mayor a menor según esta probabilidad de victoria en el último instante de la simulación, representando en color rojo el primer cuartil, en amarillo la mediana y en azul el tercer cuartil.

Con el fin de obtener escenarios más realistas de la evolución del partido se propone utilizar Prototipos Deformables Borrosos [3] para evaluar el rendimiento de cada equipo en cada instante y modificar acorde a este rendimiento las probabilidades obtenidas a partir de sus estadísticas. Para esto, se utilizan tres valores claves: el indicador ofensivo, el indicador defensivo y el indicador que representa la capacidad de respuesta del equipo. Mediante el escalado y la agregación de estos tres indicadores se obtiene un nuevo indicador de-

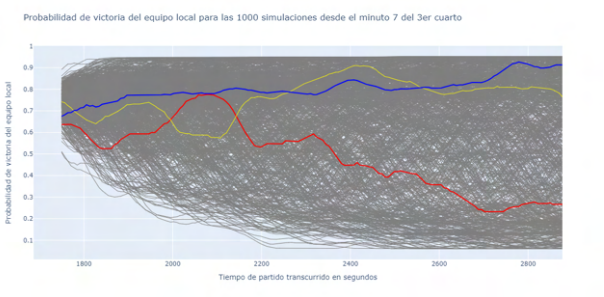


Figura 1. Evolución de la probabilidad de victoria in-game del equipo local en 1000 simulaciones OKC-DEN 12/04/2017

nominado *valor victoria*, el cual representará el rendimiento del equipo en cada instante del partido. Cada situación o escenario del partido tiene una representación prototípica P_j que la define. Para establecer la representación formal de estos prototipos borrosos el eje X viene representado por el *valor victoria* a partir del cual se definen una serie de funciones de pertenencia μ que permiten representar como números borrosos las diferentes prototipos de situación (Figura 2). Además se cuenta con una representación paramétrica basada en los indicadores que definen cada situación (Tabla I).

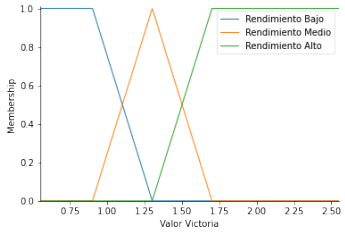


Figura 2. Representación formal de los prototipos de situación del partido.

Clase / Indicador	Defensivo	Ofensivo	Capacidad Respuesta
Clase 0	0.366	0.411	0.256
Clase 1	0.446	0.488	0.345
Clase 2	0.545	0.586	0.472

Tabla I
REPRESENTACIÓN PARAMÉTRICA DE LOS PROTOTIPOS DEFORMABLES BORROSOS.

En la Figura 3 se puede observar la evolución del grado de pertenencia de cada instante del encuentro con los indicadores del equipo local a cada uno de los conjuntos borrosos definidos pudiendo identificarse el cambio de tendencia que ocurrió durante el partido. De esta forma, las estadísticas del equipo para un instante dado se ven modificadas dinámicamente durante la simulación. Para ello, se hace uso de los grados de pertenencia a los conjuntos borrosos de los prototipos del *valor victoria* obtenido a partir de los propios indicadores de dicho equipo en ese momento. En la Figura 4 podemos visualizar la evolución de la probabilidad de victoria *in-game* para el método básico (superior) y para el método que utiliza los prototipos deformables borrosos (inferior).

Figura 3. Grados de pertenencia prototipos Kings-Lakers 10/10/2016

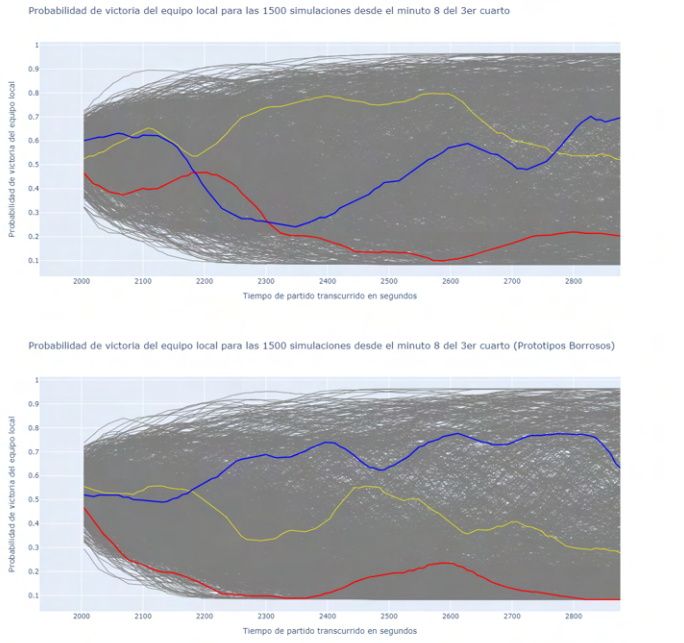


Figura 4. Comparación de la evolución de la probabilidad de victoria del equipo local en las simulaciones realizadas por los dos métodos definidos.

III. CONCLUSIONES

La utilización de Prototipos Deformables Borrosos permite obtener una evolución de la probabilidad de victoria más adaptable a las diferentes circunstancias que se puedan dar en un partido. Otras aproximaciones, como la utilización únicamente del método de Montecarlo o modelos de predicción en series temporales como LSTM ofrecen unos resultados que se adaptan menos al devenir del partido.

REFERENCIAS

[1] P. Robberechts, J. Van Haaren, and J. Davis, "Who will win it? an in-game win probability model for football," *arXiv preprint arXiv:1906.05029*, 2019.

[2] S. Hill, "In-game win probability models for canadian football," *Journal of Business Analytics*, pp. 1–15, 2021.

[3] J. A. Olivas, "Some reflections on the use of interval fuzzy sets for dealing with fuzzy deformable prototypes," in *Claudio Moraga: A passion for multi-valued logic and soft computing*, pp. 55–61, Springer, 2017.

Indicadores sociales en la perspectiva fuzzy: un ensayo sobre los efectos de la pandemia de COVID-19 en la infancia en el nordeste brasileño

1st Antonia Emanuela Oliveira de Lima
Departamento de Estudos Interdisciplinares
Universidade Federal de Ceará
Fortaleza, Brasil
emanuela.lima@ufc.br

2nd Rui Paiva
Instituto Federal de Educação, Ciência e Tecnologia do Ceará
Fortaleza, Brasil
rui.brasileiro@ifce.edu.br

3th Regivan Santiago, 4rd Benjamín Bedregal
Departamento de Informática e Matemática Aplicada
Universidade Federal do Rio Grande do Norte
Natal, Brasil
{regivan, bedregal}@dimap.ufrn.br

5th Humberto Bustince
Departamento de Matematica, Estadistica e Informatica
Universidad Publica de Navarra
Pamplona, España
bustince@unavarra.es

***Resumen*—El presente trabajo surge de la inquietud derivada de un proyecto posdoctoral que tiene como objetivo analizar los efectos de la pandemia del COVID-19 en el contexto de la niñez a través de indicadores sociales. A partir de esto se verificó que algunos indicadores a estudiar tienen características de imprecisión y ambigüedad debido a sus aspectos multidimensionales, lo que lleva a pensar que la teoria fuzzy sea esencial para establecer nuevos indicadores sociales que consigan capturar mejor esas imprecisiones y ambiguedades. Desde esta perspectiva, este trabajo tiene como objetivo discutir cómo la técnica fuzzy puede ser utilizada para definir nuevos indicadores, cuando el fenómeno de estudio es de naturaleza fuzzy. Se propone, a partir de este trabajo, ejemplificar el uso de la teoría fuzzy como metodología en el estudio y determinación de estos indicadores.**

***Palabras clave*—indicadores sociales, teoría fuzzy, modelado de la incertidumbre, políticas públicas, pandemia y la niñez.**

I. Contextualización y justificación

En los últimos años la teoría fuzzy se ha mostrado muy adecuada para analizar y describir algunos fenómenos también en campos de las actividades humanas como economía, gestión de recursos humanos y ciencias sociales [1]–[4]. Este trabajo se basa en el modelado de incertidumbre para medir algunos indicadores sociales que tiene como objetivo analizar los efectos de la pandemia del COVID-19 en el contexto de la infancia, inicialmente en el nordeste brasileño, a través de estos indicadores fuzzy. Los indicadores pueden ser conceptualizados como un conjunto de datos que buscan la descripción de una realidad, una medida de una situación, un registro en un determinado contexto y espacio. Cabe mencionar que estos datos están relacionados con las actividades del sector público para la elaboración de políticas públicas. Según Januzzi [5]: "en el campo aplicado de las políticas públicas, los indicadores sociales son medidas que permiten operacionalizar un concepto abstracto o una demanda de interés programático". Por lo tanto, se pensó en recopilar datos referentes a la efectos de la pandemia de COVID-19 en el contexto de la niñez en el nordeste brasileño. También es de destacar que los indicadores sociales son clasificadas por Januzzy [6] como objetivos y subjetivos. Los indicadores objetivos se refieren a hechos concretos o entidades empíricas de la realidad social, construidos a partir de las estadísticas públicas disponibles (...). También existen indicadores subjetivos, que por otra parte corresponden a medidas construidas a partir de la valoración de individuos o especialistas en relación a distintos aspectos de la realidad" [6]. La variable es cualitativa ordinal, por lo que no es posible calcular media, desviación estándar, varianza. El investigador debe analizar la distribución de frecuencias de cada categoría y/o analizar cómo este indicador se relaciona con otras variables utilizando técnicas adecuadas a su nivel de medición.

Este ensayo propone, a través de la teoría de conjuntos fuzzy, un tratamiento matemático de la subjetividad relacionado con algunos indicadores sociales encontrados gracias al trabajo posdoctoral en curso, los cuales fueron: (1) aumento de la violencia doméstica y (2) retraso en el aprendizaje.

En efecto, el fenómeno de la violencia doméstica se caracteriza muchas veces por dicotomizar la población en la que vive un determinado acto de violencia o cualquier tipo de violencia y quién no. Este enfoque adolece de una limitación importante, ya que la violencia no es sólo un hecho atribuido a un individuo en términos de presencia o ausencia; más bien, es un acontecimiento que se manifiesta en diferentes matices y grados. El enfoque fuzzy modela la violencia como algo que se experimenta en grados, en lugar de que simplemente ocurra o no ocurra para los individuos de la población. En la

misma dirección, retraso en el aprendizaje se entiende como la distancia entre lo que un estudiante sabe y lo que debería saber en su año escolar actual. Su nivel de conocimiento, de lo que ha aprendido, para su grado actual. Como en el caso de la violencia doméstica, existen varios factores que pueden interferir negativa o positivamente en el proceso de aprendizaje infantil. Entre ellos se destacan la desmotivación y baja autoestima del estudiante, aspectos ambientales, económicos, sociales, afectivos, psicológicos, emocionales y familiares. En este contexto, es razonable considerar variables cualitativas medidas en escala ordinal, valores intermedios entre 0 y 1 donde 0 indica la ausencia absoluta del factor, 1 la presencia absoluta y valores intermediarios índican la presencia parcial de este factor, de tal forma que cuanto mayor es el valor, mas fuerte es este factor.

II. Metodología: modelado fuzzy de indicadores

Puede decirse que estos indicadores se constituyen, entonces, como un fenómeno multidimensional, lo que les confiere cierta imprecisión. A partir de esa imprecisión, el presente estudio pretende medir estos dos indicadores a través de la teoría de los conjuntos fuzzy y datos del Censo Demográfico (IBGE), Ministerio de Educación (MEC) y Instituto de Investigaciones Económicas Aplicadas (IPEA). Sobre el cálculo del índice fuzzy que mide el aumento de la violencia doméstica o el retraso en el aprendizaje, con base en consideraciones similares a [4] y [7], tenemos dos situaciones:

(i) las variables tienen una relación positiva con el indicador, es decir, cuando aumenta el valor de la variable, aumenta el indicador (ej.: cuando aumenta el porcentaje de alumnos con baja autoestima, aumenta el retraso en el aprendizaje). La siguiente fórmula se utiliza para definir la función de asociación, el índice inicial para cada factor considerado: $x_{ij} = \frac{N_j - \text{mín}_j}{\text{máx}_j - \text{mín}_j}$, donde: x_{ij} es el valor del índice fuzzy para el factor j calculado para el municipio i; N_j es el valor observado de la serie del factor j para el municipio i; mín_j es el valor mínimo de la serie del factor j; máx_j es el valor máximo de la serie del factor j.

(ii) las variables con una relación negativa con el indicador, es decir, cuando aumenta el valor de la variable, el indicador se reduce (por ejemplo, La presencia de los padres durante el período de cuarentena es un factor que contribuye a la reducción de la violencia doméstica contra los niños, causada por terceros). Se utiliza la siguiente fórmula: $x_{ij} = \frac{\text{máx}_j - N_j}{\text{máx}_j - \text{mín}_j}$.

El establecimiento de estos límites máximos y mínimos depende de la variable bajo análisis. Después de definir la función de asociación, se encuentra la media aritmética de las observaciones de cada variable, formando un *indicador fuzzy elemental*. Así, el índice fuzzy obtenido es relativo, ya que el índice de un municipio en relación a un factor dado depende de los valores de otros municipios en relación al mismo factor.

A partir del índice relativo, se realizó una agregación considerando los pesos de los factores calculados, ponderados por el tamaño de la población en el municipio, con base en la siguiente ecuación $w_j = \ln\left[\frac{n}{n - \sum_{i=1}^{n} x_{ij}}\right] \geq 0$, donde: w_j es el peso del factor j; x_{ij} es el valor del índice fuzzy del factor j calculado para el municipio i; n es el número total de municipios en cada región.

La agregación del índice fuzzy de los indicadores multidimensionales (el aumento de la violencia doméstica o el retraso en el aprendizaje) en dimensiones y en el índice global, según el peso de cada variable, se realizó de la siguiente manera: $\mu_i = \frac{\sum_{j=1}^{n} x_{ij} w_j}{\sum_{j=1}^{n} w_j}$, donde: μ_i es el índice fuzzy multidimensional agregado del municipio i; x_{ij} es el valor del índice fuzzy del factor j calculado para el municipio i y w_j es el peso del factor j.

La metodología propuesta permitió comparar ítems dicotómicos, que evalúan determinados indicadores sociales. El valor fuzzy que mide el aumento de la violencia doméstica o el retraso en el aprendizaje puede interpretarse como una expresión del grado de vulnerabilidad de la unidad analizada (individuo, familia, hogar, unidad geográfica) a la situación de cada indicador. Vulnerabilidad no en el sentido de probabilidad, sino en el sentido de proximidad a la situación de violencia doméstica o retraso en el aprendizaje. Así, un individuo con un valor fuzzy de retraso en el aprendizaje de 0,9 está más cerca de un retraso en el aprendizaje que otro con un valor de 0,4. De la misma forma, un individuo con un valor fuzzy de violencia doméstica de 0,8 está más cerca de ser víctima de violencia doméstica que otro con un valor de 0,3; y es sólo en este sentido que el riesgo de que sea indudablemente víctima de violencia doméstica es mayor que el del segundo individuo.

Con este análisis es posible describir la realidad encontrada y contribuir al diseño, implementación y/o evaluación de políticas públicas dirigidas a la niñez durante la pandemia del COVID-19.

Reconocimiento

Humberto Bustince fue financiado por el proyecto PID2019-108392GB-I00 (AEI/10.13039/501100011033) de la Agencia española de Investigación.

Referencias

[1] M. L. Abdullah, W. S. Wong Abdullah, and A. O. Md. Tap, "Fuzzy sets in the social sciences: An overview of related researches," *Jurnal Teknologi*, vol. 41, no. 1, pp. 43–54, Feb. 2012. [Online]. Available: https://journals.utm.my/jurnalteknologi/article/view/726

[2] F. Bettio, E. Ticci, and G. Betti, "A fuzzy approach to measuring violence against women and its severity," Tech. Rep., 2018.

[3] M. Ciani, F. Gagliardi, S. Riccarelli, and G. Betti, "Fuzzy measures of multidimensional poverty in the mediterranean area: A focus on financial dimension," *Sustainability*, vol. 11, no. 1, 2019. [Online]. Available: https://www.mdpi.com/2071-1050/11/1/143

[4] J. L. Ottonelli, Janaina e Mariano, "Pobreza multidimensional nos municípios da Região do Nordeste," *Revista de Administracão Pública*, vol. 48, pp. 1253–1279, 10 2014.

[5] P. de Martino Jannuzzi, "Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no brasil," in *RSP 2014*, 2014.

[6] ——, "Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais*," 2002.

[7] E. C. Martinetti, "A multidimensional assessment of well-being based on sen's functioning approach," *Rivista Internazionale di Scienze Sociali*, vol. 108, no. 2, pp. 207–239, 2000. [Online]. Available: http://www.jstor.org/stable/41634742

FUZZ-EQ: ecualización de datos para potenciar el poder discriminativo de los clasificadores difusos

Mikel Uriz and Mikel Elkano
Neuraptic AI
Calle de Juan de Mariana 15
25045, Madrid
{elkano,uriz}@neuraptic.ai

Humberto Bustince and Mikel Galar
Institute of Smart Cities (ISC)
Universidad Pública de Navarra (UPNA)
Campus de Arrosadia, 31006 Pamplona
{bustince,mikel.galar}@unavarra.es

***Abstract*—Este keywork es un resumen del artículo publicado en la revista Applied Soft Computing [1] titulado "FUZZ-EQ: A data equalizer for boosting the discrimination power of fuzzy classifiers".**

***Index Terms*—particionamiento difuso, preprocesamiento, sistemas de clasificación basados en reglas difusas, árboles de decisión difusos, transformada integral de probabilidad, cuantil.**

I. Resumen

La lógica difusa ha proporcionado una serie de algoritmos de aprendizaje automático o machine learning con la capacidad de tratar con la incertidumbre, generando además fronteras difusas que permiten mejorar la precisión de los clasificadores. Además, los clasificadores difusos trabajan con términos y etiquetas lingüísticas, lo cual permite generar modelos compuestos por reglas de tipo IF-THEN que son fácilmente interpretables por el ser humano. Estos modelos permiten explicar el razonamiento seguido por el clasificador en sus predicciones y describir la inferencia realizada entre los datos de entrada y las etiquetas de salida. En este trabajo nos hemos centrado en dos modelos de clasificadores difusos ampliamente estudiados en la literatura, los sistemas de clasificación basados en reglas difusas (FRBCSs por sus siglas en inglés) y los árboles de decisión difusos (FDTs).

Seleccionar una estrategia adecuada para definir los términos lingüísticos es un aspecto fundamental en el diseño de un clasificador difuso. Cada término está definido por una función de pertenencia que determina el grado de pertenencia de un valor de entrada al conjunto difuso asociado con el término. En muchas aplicaciones, está función es definida por un experto en la materia.

Sin embargo, existen situaciones donde no podemos involucrar expertos y el sistema debe definir automáticamente estas funciones de pertenencia. Cuando nos enfrentamos a un problema de clasificación, el poder de discriminación del mismo está limitado mayormente por la capacidad de los términos lingüísticos para representar los datos de manera adecuada. Por ejemplo, si consideramos los términos "Baja", "Media" y "Alta" para representar la variable temperatura en el intervalo $[-10°, 40°]C$ y el 99% de los datos observados varían entre $13°C$ y $26°C$, entonces no tendrá sentido que realicemos una partición uniforme del dominio entre las tres etiquetas. En este contexto concreto, deberíamos considerar $13°C$ como temperatura baja y $26°C$ como temperatura alta.

Para evaluar la capacidad de representación de los datos por parte de los términos lingüísticos, Pedrycz [2] formalizó el concepto de *data meaningfulness* de un conjunto difuso. El autor establece que los términos lingüísticos deberían definirse únicamente cuando existan suficientes muestras para justificar su existencia. Teniendo en cuenta esto, la definición de las funciones de pertenencia deberían depender totalmente de la distribución empírica de los datos. Para ajustar la posición y forma de los conjuntos difusos, en la literatura se han propuesto diferentes métodos de particionamiento difusos (FPMs). En el campo concreto de la clasificación, algunos FPMs se centran en optimizar la partición difusa tratando de maximizar la precisión del modelo. Otros métodos priorizan la interpretabilidad sobre la precisión y generan particiones uniformes con parámetros predefinidos. Incluso la combinación de ambas aproximaciones ha sido también utilizada. Además, existen también diferencias sobre cuándo se aplica el particionamiento. Algunos crean la partición antes del proceso de aprendizaje, mientras que otros integran el particionamiento en el propio proceso.

En este trabajo afrontamos el problema descrito desde una perspectiva diferente. En lugar de ajustar los conjuntos difusos a los datos de entrenamiento, nuestra propuesta trata de facilitar el trabajo de los algoritmos de particionamiento ajustando previamente los datos a una hipotética partición uniforme. El método que proponemos, llamado FUZZ-EQ, consiste en transformar la distribución original de los datos en una distribución uniforme mediante la aplicación del teorema de la transformada integral de probabilidad [3], [4]. En concreto, FUZZ-EQ aplica una ecualización del histograma de cada variable de forma que las funciones de pertenencia asociadas a una partición uniforme trivial se ajusten perfectamente a los datos. Nuestra hipótesis es que esta transformación permite a los FPMs obtener una mayor granularidad en regiones con mucha densidad de datos, potenciando así la capacidad discriminativa del modelo. En la Fig. 1 ilustramos los pasos principales que realiza FUZZ-EQ.

Las principales contribuciones de nuestra propuesta son:

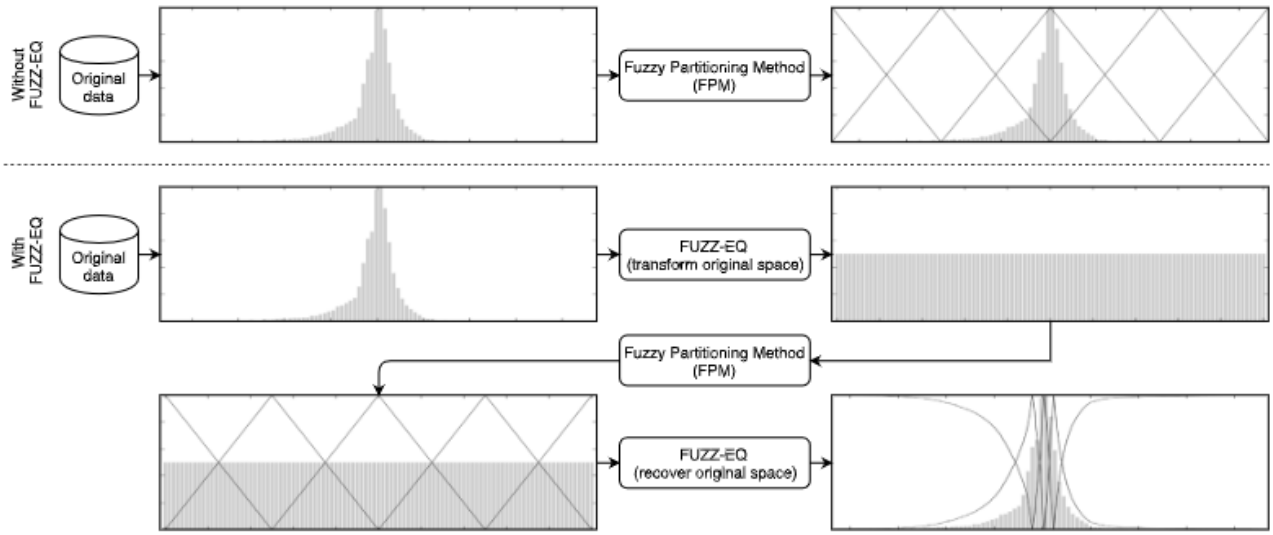


Fig. 1. Principales pasos de FUZZ-EQ.

- Versatilidad y flexibilidad. FUZZ-EQ se puede aplicar con cualquier clasificador, independientemente del FPM aplicado.
- Sin necesidad de optimización. El método permite que un particionamiento difuso uniforme obtenga los mejores resultados sin necesidad de optimizar ninguna otra métrica.
- Interpretabilidad. Cualquier dato en el espacio transformado puede ser proyectado al espacio original, recuperando así la interpretabilidad de los modelos difusos.

Para evaluar la efectividad de nuestra propuesta, hemos realizado un extenso estudio experimental basado en 41 conjuntos de datos de los repositorios UCI y KEEL. Hemos trabajado con 9 clasificadores difusos. Los resultados obtenidos demuestran que FUZZ-EQ es capaz de potenciar el rendimiento de la mayoría de los clasificadores. En aquellos en los que no se observa mejora, el rendimiento se ha mantenido igual. Además, hemos comprobado que el método es escalable a problemas Big Data utilizando el algoritmo CHI-BD como clasificador.

ACKNOWLEDGMENT

Mikel Uriz and Mikel Elkano han sido apoyado por CDTI y el Ministerio de Ciencia e Innovación bajo el proyecto Neotec 2021 (SNEO-20211147). Este trabajo también ha sido apoyado por el Ministerior de Ciencia e Innovación bajo el proyecto PID2019-108392GB-I00 (AEI/10.13039/501100011033).

REFERENCES

[1] M. Uriz, M. Elkano, H. Bustince, M. Galar, "FUZZ-EQ: A data equalizer for boosting the discrimination power of fuzzy classifiers," Applied Soft Computing, vol. 93, no. 106399, 2020.
[2] W. Pedrycz, "Fuzzy equalization in the construction of fuzzy sets", Fuzzy Sets and Systems, vol. 119 pp. 329–335, 2001.
[3] J.E. Angus,"The probability integral transform and related results", SIAM Review, vol. 36, pp. 652–654, 1994.
[4] C.P. Quesenberry, "Probability integral transformations", John Wiley & Sons, Inc., 2004.

Difusión y modelado basado en 2-tuplas lingüísticas de las opiniones de los consumidores en mercados competitivos

Jesús Giráldez-Cru*, Manuel Chica*†, Oscar Cordón* and Francisco Herrera*
* *Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)*
Dpto. de Ciencias de la Computación e Inteligencia Artificial (DECSAI), Univ. de Granada (UGR)
{jgiraldez,mchica,ocordon,herrera}@decsai.ugr.es
†*School of Electrical Engineering and Computing, The University of Newcastle (Australia)*

***Resumen*—Modelar el comportamiento del consumidor es una de las cuestiones fundamentales en marketing. En este trabajo se presenta un modelado del consumidor basado en 2-tuplas lingüísticas, una representación realista de información cualitativa, incluyendo la representación de las percepciones de los consumidores, los procesos de difusión de opiniones entre ellos y las heurísticas de toma de decisión. Todo ello se implementa en un modelo basado en agentes, donde se pueden realizar simulaciones masivas para extraer el comportamiento complejo del mercado.**

***Index Terms*—comportamiento del consumidor, marketing, heurísticas de compra, reconocimiento de marca, toma de decisiones lingüística, 2-tuplas, modelos basados en agentes.**

I. Introducción

Una de las cuestiones fundamentales en marketing es modelar adecuadamente el comportamiento de los consumidores de un mercado [1]. Los enfoques tradicionales, basados en reglas *top-down* a partir de variables globales del mercado, presentan una problemática a la hora de representar los comportamientos heterogéneos de cada consumidor. Alternativamente, los enfoques *bottom-up* permiten modelar comportamientos simples y heterogéneos, de los que se puede derivar el comportamiento complejo del mercado como resultado de su agregación. Los modelos basados en agentes (ABM) [2] son un marco adecuado para este tipo de enfoques.

En este trabajo presentamos un ABM donde se modela el comportamiento del consumidor usando 2-tuplas lingüísticas [3], el cual está basado en las contribuciones de varios artículos previos del grupo de investigación: modelado de percepciones del consumidor [4], difusión de opiniones [5], y heurísticas de toma de decisión [6]. Las percepciones del consumidor representan las opiniones que éste tiene sobre las diferentes marcas o productos en el mercado (en una escala positiva-neutra-negativa), así como el reconocimiento de marca (el conjunto de marcas conocidas para cada consumidor). La difusión de opiniones representa el proceso de comunicación entre consumidores, en los cuales se comparten las opiniones propias (las cuales evolucionan en el tiempo) dentro de una red social y que puede influir a otros consumidores, afectando igualmente a las dinámicas en el reconocimiento de marca (marcas que pasan a ser conocidas o marcas que dejan de serlo). Finalmente, la heurística de toma de decisión representa el proceso de compra de los consumidores y debe contemplar los distintos comportamientos existentes en cada mercado. En las siguientes secciones se resumen las principales contribuciones de estas tres componentes.

II. Percepciones del consumidor

El primer paso en el modelado del consumidor es la representación de sus percepciones de una forma realista. Estas percepciones representan las opiniones del consumidor acerca de los distintos productos en el mercado analizado. Al tratarse de una información cualitativa, en nuestro modelo usamos 2-tuplas lingüísticas para representar estas percepciones [4].

Definición 1 (2-tupla lingüística [3]): Una 2-tupla lingüística es un par $\langle s_k, \alpha \rangle$, en el que $s_k \in S$ es una etiqueta lingüística y $\alpha \in [-0{,}5, 0{,}5)$ es un desplazamiento simbólico que especifica la traslación de la función de pertenencia con respecto a la etiqueta lingüística s_k más cercana.

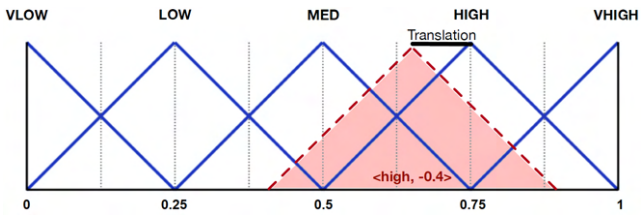


Figura 1. Función de pertenencia triangular para variables lingüísticas del conjunto $S = \{vlow, low, med, high, vhigh\}$, y de la 2-tupla $\langle high, -0{,}4 \rangle$.

En la Figura 1 se representa la función de pertenencia triangular de un conjunto de variables lingüísticas y de la 2-tupla $\langle high, -0{,}4 \rangle$.

En nuestro ABM, el mercado analizado se limita a un conjunto de marcas $B = \{b_1, \ldots, b_n\}$, cada una de las cuales tiene un conjunto $D = \{d_1, \ldots, d_m\}$ de m atributos o *drivers* que guían la decisión de compra de los distintos consumidores y sobre los que éstos tienen percepciones. Estas percepciones se representan mediante una matriz P^x para cada agente x, donde cada elemento $p^x_{i,j} \in P^x$ representa la percepción del agente x sobre el *driver* $d_j \in D$ de la marca $b_i \in B$, y cuyo

valor es una 2-tupla lingüística. Adicionalmente, cada agente asigna un peso $w_j^x \in [0,1]$ (con $1 \leq j \leq m$) a cada *driver*, representando la importancia del *driver* j en la percepción del consumidor concreto. Finalmente, el reconocimiento de marca se representa mediante una variable $a_i^x \in \{0,1\}$ (con $1 \leq i \leq n$), es decir, si el consumidor x conoce o no conoce a la marca i, lo que lógicamente condiciona su decisión de compra y la existencia o no de percepciones sobre esa marca.

III. Difusión de opiniones

La difusión de opiniones representa el proceso en el que los consumidores se comunican y posiblemente cambian sus percepciones debido a estos intercambios de información.

Para ello, nuestro ABM [5] define tres sub-procesos de comunicación basados en el modelo RTR [7]. El primero es un proceso de *radiación*, en el cuál se seleccionan las opiniones que cada agente comunica. El segundo es un proceso de *transmisión*, donde se seleccionan los agentes con los que cada uno se comunica. Y finalmente en el proceso de *recepción* se seleccionan las opiniones que cada agente actualiza.

La comunicación entre los agentes se realiza sobre una red social, sobre la cuál se puede definir también la influencia entre los agentes. Tras este proceso de comunicación, las percepciones de los consumidores se actualizan mediante una regla de fusión de opiniones. Nótese que tanto los procesos de comunicación como la regla de fusión de opiniones usan 2-tuplas lingüísticas.

Adicionalmente, el sistema incorpora un mecanismo de (des)activación de reconocimiento de marca para modelar las dinámicas que cada consumidor tiene en cuanto a marcas que pasan a ser conocidas o marcas que dejan de serlo.

IV. Heurísticas de toma de decisión

La tercera y última componente de nuestro ABM es la heurística de toma de decisión [6]. Esta heurística representa el proceso de compra que realiza cada consumidor, que es la información más importante del mercado desde el punto de vista de marketing.

Aunque cualquier heurística de toma de decisiones se basa en las propias percepciones del consumidor, es bien sabido que cada consumidor utiliza diferentes estrategias según la naturaleza de la compra a la que se enfrenta. Por ejemplo, ciertas decisiones requieren procesos altamente deliberativos con el objetivo de maximizar la utilidad de la decisión (como la compra de un coche), mientras que otras compras se llevan a cabo con estrategias de decisión rápidas y suficientemente satisfactorias (como la compra de un cartón de leche).

Para modelar esta heterogeneidad, nuestro ABM define cuatro heurísticas diferentes, que difieren en cuán involucrado está el consumidor en la decisión. En concreto, las diferencias se encuentran en el grado de exploración (parcial o exhaustivo) tanto de los *drivers* como de las marcas. Estas heurísticas se han diseñado siguiendo estudios sobre comportamiento económico y, dado a que usan las propias percepciones del consumidor, se han diseñado para manejar 2-tuplas lingüísticas para la toma de decisión.

Aunque cada mercado se caracteriza mayormente por una heurística de toma de decisión determinada, no todos los consumidores siguen exactamente esta heurística para decidir sus compras. Para modelar este fenómeno, nuestro ABM integra las cuatro heurísticas en un único proceso, de forma que cada agente puede elegir cualquiera de las cuatro heurísticas con cierta probabilidad. Resultados experimentales muestran que esta estrategia suele ser la más precisa para representar el comportamiento complejo del mercado de la forma más adecuada, dada la heterogeneidad de los consumidores reales.

V. Conclusiones

El ABM propuesto en este trabajo permite modelar el comportamiento individual de cada consumidor de una forma realista, usando 2-tuplas lingüísticas [3], y a partir del cual se puede extraer el comportamiento complejo de todo el mercado. Nuestro modelo incluye la representación de las percepciones del consumidor y el reconocimiento de marca [4], los procesos de difusión de opiniones [5], y las heurísticas de toma de decisión para la compra [6]. Este sistema se ha evaluado experimentalmente usando datos de varios mercados reales y los resultados experimentales muestran que esta representación realista permite mejorar el rendimiento del mismo (en comparación con los datos históricos de ventas en estos mercados). Como trabajo futuro, se propone: (i) personalizar la semántica individual de las percepciones [8]; (ii) analizar los efectos de la comunicación de masas (publicidad o campañas de marketing) en los procesos de difusión de opiniones, y (iii) aprender los comportamientos de toma de decisión del consumidor de manera individual.

Agradecimientos

Trabajo financiado por MICINN, JdA, UGR, AEI, y ERDF (A-TIC-284-UGR18, P18-FR-4262, PY18-4475, IJC2019-040489-I). Agradecemos a ZIO Analytics por los datos aportados.

Referencias

[1] W. Rand and R. T. Rust, "Agent-based modeling in marketing: Guidelines for rigor," *International Journal of Research in Marketing*, vol. 28, no. 3, pp. 181–193, 2011.

[2] J. M. Epstein, *Generative social science: Studies in agent-based computational modeling*. Princeton University Press, 2006.

[3] F. Herrera and L. Martínez-López, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, 2000.

[4] J. Giráldez-Cru, M. Chica, O. Cordón, and F. Herrera, "Modeling agent-based consumers decision-making with 2-tuple fuzzy linguistic perceptions," *International Journal of Intelligent Systems*, vol. 35, pp. 283–299, 2020.

[5] J. Giráldez-Cru, M. Chica, and O. Cordón, "A framework of opinion dynamics using fuzzy linguistic 2-tuples," *Knowledge Based Systems*, vol. 233, p. 107559, 2021.

[6] J. Giráldez-Cru, M. Chica, and O. Cordón, "An integrative decision making mechanism for consumers' brand selection using 2-tuple fuzzy linguistic perceptions and decision heuristics," *International Journal of Fuzzy Systems.* Submitted, 2022.

[7] W. Vermeer, O. Koppius, and P. Vervest, "The Radiation-Transmission-Reception (RTR) model of propagation: Implications for the effectiveness of network interventions," *PLOS ONE*, vol. 13, pp. 1–21, 2018.

[8] H. Liang, C. Li, Y. Dong, and F. Herrera, "Linguistic opinions dynamics based on personalized individual semantics," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 9, pp. 2453–2466, 2021.

Emparejamientos óptimos en contextos L-fuzzy

Cristina Alcalde
Dept. Matemática Aplicada
Universidad del País Vasco-UPV/EHU
Plaza de Europa, 1. 20018 Donostia-San Sebastian
c.alcalde@ehu.es

Ana Burusco
Dept. Estadística, Matemática e Informática
Instituto de Smart Cities
Universidad Pública de Navarra-UPNA
Campus de Arrosadia. 31006 Pamplona
burusco@unavarra.es

***Resumen*—En contextos L-fuzzy (L, X, Y, R) en los que el número de objetos X coincide con el número de atributos Y, se plantea el problema de emparejar cada uno de los objetos con un atributo distinto de manera que los pares resultantes optimicen el grado de emparejamiento que se defina a partir de la relación existente entre los objetos y los atributos. Un ejemplo de este tipo de situaciones puede ser el de emparejar las personas de dos grupos buscando la máxima afinidad. O también asignar taxistas que circulan por la ciudad a varios clientes que los necesitan, minimizando en este caso el tiempo total de espera.**

***Index Terms*—Contexto L-fuzzy, concepto L-fuzzy, retículo de conceptos L-fuzzy**

I. Introducción

Un contexto L-fuzzy es una tupla (L, X, Y, R), donde L es un retículo completo, X e Y son conjuntos de objetos y atributos, y $R \in L^{X\times Y}$ es una relación L-fuzzy entre los objetos y los atributos. Utilizaremos los conceptos L-fuzzy para extraer información de un contexto L-fuzzy ([1], [2]).

Los contextos L-fuzzy son una extensión de los contextos formales de Wille ([6]).

Para obtener la información que se extrae de esos contextos L-fuzzy, definimos en artículos anteriores [1]–[4] los operadores derivación denotados por 1 y 2.

$$A_1(y) = \inf_{x\in X}\{I(A(x), R(x, y))\},\ \forall A \in L^X \tag{1}$$

$$B_2(x) = \inf_{y\in Y}\{I(B(y), R(x, y))\}, \forall B \in L^Y \tag{2}$$

con I un operador de implicación definido en $(L, \leq)$.

La información almacenada en el contexto se extrae mediante conceptos L-fuzzy, los cuales son pares $(A, A_1) \in L^X \times L^Y$ donde $A \in fix(\varphi)$, el conjunto de puntos fijos del operador φ definido como $\varphi(A) = (A_1)_2 = A_{12}$. Estos pares, cuya primera y segunda componentes se llaman extensión e intensión fuzzy, representan grupos de objetos que comparten grupos de atributos.

El conjunto $\mathcal{L} = \{(A, A_1)/A \in fix(\varphi)\}$ con la relación de orden $\preceq$ definida como: $\forall (A, A_1), (C, C_1) \in \mathcal{L}, (A, A_1) \preceq (C, C_1)$ if $A \leq C$(or$C_1 \leq A_1$) es un retículo completo al que llamaremos retículo de conceptos L-fuzzy ([1], [2]).

Además, para cada $A \in L^X$, (o $B \in L^Y$) el concepto L-fuzzy asociado puede determinarse aplicando dos veces los operadores derivación. En el caso de usar una implicación residuada, el concepto L-fuzzy es (A_{12}, A_1) (or (B_2, B_{21})).

Aunque los operadores derivación pueden definirse utilizando cualquier operador de implicación, de cara a simplificar los cálculos, utilizaremos en este trabajo sólamente operadores residuados definidos en $L = [0, 1]$.

II. Grado de emparejamiento entre un objeto y un atributo

Si representamos cada objeto $x_i \in X, i \in \{1, \ldots, n\}$, por el conjunto L-fuzzy $\mathbf{x_i}$ tal que $\mathbf{x_i}(x_i) = 1$ y $\mathbf{x_i}(x) = 0$, para todo $x \neq x_i$, podemos definir el siguiente concepto L-fuzzy:

Definición 1: Para cada $x_i \in X, i \in \{1, \ldots, n\}$, llamaremos *concepto L-fuzzy asociado al objeto* x_i al par $\mathcal{C}_{\mathbf{x_i}} = ((\mathbf{x_i})_\varphi, ((\mathbf{x_i})_\varphi)_1)$, donde $(\mathbf{x_i})_\varphi$ representa el punto fijo del operador φ obtenido a partir de conjunto $\mathbf{x_i}$. Cuando utilizamos una implicación residuada esta expresión se puede escribir como $\mathcal{C}_{\mathbf{x_i}} = ((\mathbf{x_i})_{12}, (\mathbf{x_i})_1)$.

Analogamente, $\mathcal{C}_{\mathbf{y_j}} = ((\mathbf{y_j})_2, (\mathbf{y_j})_{21}), j \in \{1, \ldots, n\}$ será el concepto L-fuzzy asociado al atributo y_j, donde $\mathbf{y_j}$ es el conjunto L-fuzzy que representa a dicho atributo.

En el resto del trabajo, denotaremos por $\mathcal{C}_{\mathbf{x_i}} = (\underline{\mathcal{C}}_{\mathbf{x_i}}, \overline{\mathcal{C}}_{\mathbf{x_i}})$ y $\mathcal{C}_{\mathbf{y_j}} = (\underline{\mathcal{C}}_{\mathbf{y_j}}, \overline{\mathcal{C}}_{\mathbf{y_j}})$. Estos conceptos L-fuzzy son los más próximos a los conjuntos de partida representados por $\mathbf{x_i}$ o $\mathbf{y_j}$ (estudio de un objeto o un atributo), y nos dan una idea general sobre el comportamiento de los objetos y los atributos en el reticulo de conceptos L-fuzzy con bajo coste computacional.

Utilizaremos estos conceptos L-fuzzy para definir el grado de emparejamiento de la siguiente forma:

Definición 2: Llamaremos $GE(x_i, y_j)$ al grado de emparejamiento del objeto x_i y del atributo y_j.

$$GE(x_i, y_j) = \overline{\mathcal{C}}_{\mathbf{x_i}}(y_j) = \underline{\mathcal{C}}_{\mathbf{y_j}}(x_i) \tag{3}$$

Proposición 1: Si la implicación que se utiliza en la definión de operadores derivación para el cálculo de los conceptos L-fuzzy es una implicación residuada, entonces para cada par (x_i, y_j), se cumple que:

$$GE(x_i, y_j) = \overline{\mathcal{C}}_{\mathbf{x_i}}(y_j) = \underline{\mathcal{C}}_{\mathbf{y_j}}(x_i) = R(x_i, y_j) \tag{4}$$

III. Emparejamientos óptimos

A partir de la definición de grado de emparejamiento de un par objeto-atributo $GE(x_i, y_j)$, podemos definir también el grado de emparejamiento que se pueden obtener asociando un atributo distinto a cada uno de los n objetos del contexto.

Definición 3: Sea $n = |X| = |Y|$ y sean σ una permutacioón de $\{1, 2, \ldots, n\}$. Podemos construir el conjunto de pares $\mathcal{P}_\sigma =$

$\{(x_1, y_{\sigma(y_1)}), (x_2, y_{\sigma(y_2)}), \ldots, (x_n, y_{\sigma(y_n)})\}$ y definir el grado de emparejamiento de los elementos de $\mathcal{P}_\sigma$ como:

$$GE(\mathcal{P}_\sigma) = \sum_{1 \leq i \leq n} GE(x_i, y_{\sigma(y_i)}) \tag{5}$$

Este grado de emparejamiento tomará distintos valores en función de la permutación σ elegida y nosotros estaremos interesados de tomar el Grado de emparejamiento máximo (GEM) o el Grado de emparejamiento mínimo (GEm) dependiendo de la naturaleza del problema.

$$GEM = \max_{k \in \{1,2,\ldots n!\}} GE(\mathcal{P}_{\sigma_k}) \tag{6}$$

$$GEm = \min_{l \in \{1,2,\ldots n!\}} GE(\mathcal{P}_{\sigma_l}) \tag{7}$$

Resolver este problema consiste en construir el árbol de estados y aplicar técnicas de ramificación y acotamiento ([5]) que permitan no tener que desarrollar todos los nodos en aras de un bajo coste computacional.

El sistema para generar nodos en el árbol de estados que se sigue es el siguiente: Se irán generando los hijos de un nodo de manera que este nodo seguirá siendo el nodo actual hasta que muera. A continuación, en cada nodo se calcula una cota de los valores de las soluciones situadas por debajo. Si esta cota demuestra que tales soluciones son eventualmente peores que una solución ya encontrada, se abandona la exploración de esa parte del árbol.

El cálculo de la cota se combina con un recorrido en anchura o en profundidad que sirve para podar ciertas ramas del árbol. Además, esta cota sirve también para escoger la rama más prometedora.

A continuación veremos un caso práctico: Los valores que asignamos a cada una de las ramas del árbol de estados son en este caso los grados de pertenencia de los conceptos asociados a los objetos del contexto: $\mathcal{C}_{\mathbf{x_i}} = (\underline{\mathcal{C}}_{\mathbf{x_i}}, \overline{\mathcal{C}}_{\mathbf{x_i}}), i \in \{1, \ldots, n\}$.

Estos valores serán también necesarios para calcular las cotas tal y como vamos a ver.

IV. Caso práctico

Queremos enviar taxistas a 4 clientes que los necesitan minimizando el tiempo total de espera. Si X es el conjunto de clientes e Y el conjunto de taxistas, podemos representar la situación mediante un contexto L-fuzzy (L, X, Y, R) donde hemos escalado los valores a [0,1]. La relación R que mide la distancia entre cada taxista y cada cliente es la siguiente:

Cuadro I
DISTANCIA CLIENTES -TAXISTAS

R	y_1	y_2	y_3	y_4
x_1	0.7	1	0.4	0.3
x_2	0.3	0.8	0.4	0.7
x_3	0.7	1	0.4	0.3
x_4	0.2	0.2	0.2	0.2

Queremos asignar un taxi a cada cliente minimizando el tiempo total de espera. Es decir, buscamos una permutación σ tal que $GE(\mathcal{P}_\sigma) = \min_{k \in \{1,2,\ldots n!\}} GE(\mathcal{P}_{\sigma_k})$

En el árbol de estados que hemos construido, hemos numerado los nodos siguiendo el orden en el que se han generado.

Cada nodo (excepto la raíz) se corresponde con una asignación de un taxista y_j a un cliente x_i y puede tener asociada una cota inferior del tiempo total de atención que se requerirá por ese camino. Dicha cota aparece debajo del nodo. Esta cota se calcula sumando a los tiempos reales de los taxis asignados, la suma de los tiempos mínimos de los taxis no asignados todavía. Por ejemplo, la cota del nodo 6 será 0.9, que es la suma de 0.3+0.3 (taxis asignados) y 0.1+0.2 (valor mínimo de los taxis no asignados).

Desarrollaremos aquellos nodos más prometedores: por ejemplo, el nodo 5 se ha desarrollado antes que los nodos 2, 3 y 4. Y utilizaremos la cota inferior para podar algunas ramas: como ya sabemos que los nodos 3, 4, 7, 8, 13, 14 y 15 no pueden tener valores más pequeños que 0.9 (valor real obtenido en el nodo 11), no es necesario desarrollarlos.

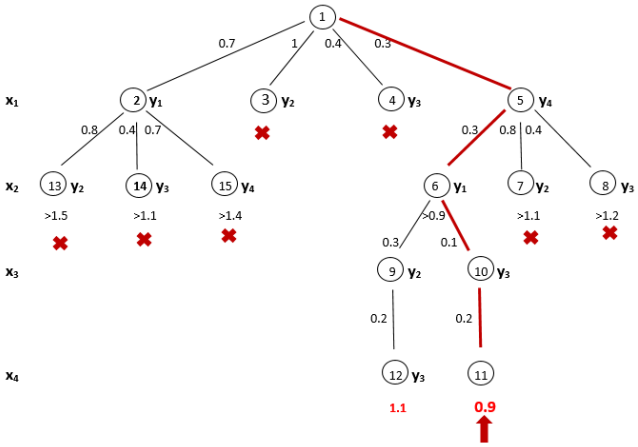


Figura 1. Árbol de estados

Por tanto, el emparejamiento que nos da un tiempo de espera mínimo $GEm = 0{,}9$ es:

$$\mathcal{P}_\sigma = \{(x_1, y_4), (x_2, y_1), (x_3, y_3), (x_4, y_2)\} \tag{8}$$

siendo la permutación $\sigma = (4, 1, 3, 2)$.

Referencias

[1] A. Burusco, R. Fuentes-González, "The Study of the L-fuzzy Concept Lattice", Mathware and Soft Computing 1 (3), pp. 209–218, 1994.

[2] A. Burusco, R. Fuentes-González, "Construction of the L-fuzzy Concept Lattice", Fuzzy Sets and Systems 97 (1), pp. 109–114, 1998.

[3] A. Burusco, R. Fuentes-González, "Concept lattices defined from implication operators", Fuzzy Sets and Systems 114 (1), pp. 431–436, 2000.

[4] A. Burusco, R. Fuentes-González, "The study of the interval-valued contexts", Fuzzy Sets and Systems 121, pp. 69–82, 2001.

[5] E. Horowitz , S. Sahni, S. Rajasekaran, Computer Algorithms, 3ª edición, Computer Science Press, 1998.

[6] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: Rival I. (Ed.), Ordered Sets, Reidel, Dordrecht-Boston, pp. 445-470, 1982.

On Inclusion and Order Relation of Interval-valued Fuzzy Sets

Emilio Torres-Manzanera
Dept. of Statistics
University of Oviedo
Oviedo, Spain
torres@uniovi.es

Agustina Bouchet
Dept. of Statistics
University of Oviedo
Oviedo, Spain
bouchetagustina@uniovi.es

Susana Díaz-Vázquez
Dept. of Statistics
University of Oviedo
Oviedo, Spain
diazsusana@uniovi.es

Susana Montes
Dept. of Satistics
University of Oviedo
Oviedo, Spain
montes@uniovi.es

***Abstract*—In the hazardous marketplace, where well stated membership degrees are not available, interval-valued fuzzy sets (IVFSs) seem to be an appropriate tool. In this contribution we focus on the way to measure how different two IVFSs are and in particular, how to formalize the so-called axiom 5: the more different the IVFSs are, the bigger the dissimilarity between them must be. In connection with this, we present the concept of infimum and supremum of a set of IVFSs and we propose an alternative axiom 5.**

***Index Terms*—interval-valued fuzzy set, order, inclusion, supremum, infimum**

I. Introduction

Investment in stock markets is getting more and more popularity. The daily evolution of the stock prices (Fig. 1) reminds one of the analysis of the dissimilarity among interval-valued fuzzy sets (IVFSs). The lack of information about the price may be assimilated to the imprecise grade of membership of an IVFS. And a measure that quantifies differences in the comparison of yields of different stocks can be enlighten using dissimilarities among IVFSs.

However, the way to measure how different two interval valued fuzzy sets are has been a topic under debate since some time ago [2], [3], [7], [8]. The most discussed point probably being the preservation of the order of the differences: the dissimilarity measure should provide smaller outcomes for closer IVFSs. The formalization of "closer IVFSs" is maybe the most difficult point in the attempt of providing a measure of difference between IVFSs. Some proposals reasonable in the context of fuzzy sets become unfeasible for IVFSs as we discussed in [3].

In this contribution, we take up this point. The definition of inclusion proposed in [3] to formalize closeness is maybe too restrictive. We explore the use of supremum and infimum of a triplet of IVFSs in order to avoid defining inclusion. Yet, we need ordering relations among intervals to define supremum and infimum and we have started the study of their behaviour.

This research was funded by Spanish MINECO projects PGC2018-098623-B-I00 (Emilio Torres-Manzanera, Susana Díaz-Vázquez, and Susana Montes) and TIN-2017-87600-P (Agustina Bouchet).

Fig. 1. In the area of financial investment, dissimilarity between price variation set us off reminiscing interval-valued fuzzy sets. In the image, daily price variation of stock prices of Apple, Bank of America and Exxon Mobil companies.

II. Preliminaries

A. Interval-valued fuzzy sets

In this section, we reproduce some basic notions and properties as presented in [3] in order to understand the main content of this contribution.

An interval-valued fuzzy set (IVFS for short) in X is a mapping $A : X \to L([0,1])$ such that $A(x) = [\underline{A}(x), \overline{A}(x)]$, where $L([0,1])$ denotes the family of closed intervals included in the unit interval $[0,1]$. It is therefore easy to check that an interval-valued fuzzy set A is characterized by two mappings, $\underline{A}$ and $\overline{A}$, from X into $[0,1]$ such that $\underline{A}(x) \leq \overline{A}(x), \forall x \in X$. These functions provide the lower and upper bounds, respectively, of the associated intervals. Observe that if $\underline{A}(x) = \overline{A}(x), \forall x \in X$, then A is a classical fuzzy set. The abbreviation $IVFS(X)$ stands for the set of all the interval-valued fuzzy sets in X.

For IVFSs, we consider the epistemic interpretation. Thus, we assume that there is one actual, real-valued membership degree of an element inside the membership interval of possible membership degrees.

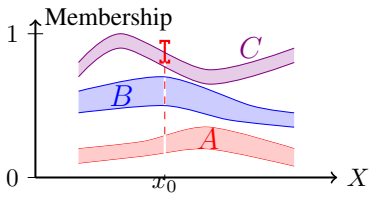


Fig. 2. Membership degrees for A, B and C. A and B are not comparable due to a single element: $A(x) \preceq_o B(x), \forall x \in X - \{x_0\}$ and $B(x_0) \prec_o A(x_0)$; i.e., A is contained in B for all the elements in the universe except for one.

B. Sorting intervals out

The inclusion for IVFSs is directly connected to the way one sorts intervals out. In [6], we can find a summary of the main interval orderings.

Interval dominance [4], the lattice order [5], the lexicographical order of types 1 and 2 [1] or the Xu and Yager order [9] are maybe the most known, but we have used others in [3]. Interval dominance is the strongest of the previous orderings, in the sense that it implies all the other ones.

Notice that we use "ordering" and not "order" because interval dominance and other classical ways of sorting intervals out (not recalled here) are not orders in the mathematical sense. Interval dominance is not reflexive, for example. So we will only use "order" for reflexive, antisymmetric and transitive relations.

Axiom 5 in the definition of dissimilarity provided in [3] is

$$A \subseteq_o B \subseteq_o C, \;\Rightarrow\; \begin{cases} D(A,B) \preceq_o D(A,C) \\ D(B,C) \preceq_o D(A,C)\,. \end{cases} \qquad (1)$$

for $A, B, C \in IVFS(X)$ and $\subseteq_o$ defined as $A \subseteq_o B$ if $A(x) \preceq_o B(x), \forall x \in X$ where $\preceq_o$ is an ordering in $L([0,1])$.

For any ordering $\preceq_o$ in $L([0,1])$, the associated $\subseteq_o$ is a partial order in $IVFS(X)$.

III. The fifth axiom

With the detailed previous concepts we can settle our problem: The 5th axiom (see Eq. 1) is too weak in the sense that very few triplets of IVFSs actually verify the departing properties: only if the three IVFSs considered are ordered by inclusion, the dissimilarity between the extremes of the chain has to be greater than the dissimilarity between the intermediate IVFS and one of the other two. So the axiom becomes an actual restriction only in very specific and particular contexts. Fig. 1 shows an example of a practical situation where these hypothesis (IVFSs ordered by inclusion) do not hold and Fig. 2 shows a case where three IVFSs are not ordered just because of one point in the domain.

So we focus on the formalization of the fifth axiom to cover a wider range of situations. A first approach is to identify the extremes of the chain of IVFSs as the infimum and supremum and study their properties after having given a formal definition for these concepts.

Definition 1. *M is the supremum of an arbitrary collection of IVFSs $\{A_i\}_i \subseteq IVFS(X)$ if and only if $\forall H \in IVFS(X)$, such that $A_i \subseteq_o H\, \forall i$, , then $A_i \subseteq_o M \subseteq_o H$.*

m is the infimum of $\{A_i\}_i \subseteq IVFS(X)$ if and only if $\forall H \in IVFS(X)$, such that if $\forall i$, $H \subseteq A_i$, then $H \subseteq_o m \subseteq_o A_i$.

The first step has been to obtain their explicit expressions for the most common orderings in those cases that there exists. Due to space constraints we only reproduce here the result obtained for interval dominance.

Proposition 2. *Consider $\{A_i\}_i$ a collection of IVFSs of X and the interval dominance [4] as the ordering used ($a \preceq_{ID} b$ if $\overline{a} \leq \underline{b}$) to compare them. Then the supremum M and the infimum m of a set of IVFSs $\{A_i\}_{i\in I}$ are defined by*

$$\underline{M}(x) = \overline{M}(x) = \sup_i\{\overline{A}_i(x)\}, \quad \forall x \in X.$$
$$\underline{m}(x) = \overline{m}(x) = \inf_i\{\underline{A}_i(x)\}, \quad \forall x \in X.$$

Once introduced the concepts of infimum and supremum in Definition 1, we are set to introduce an alternative to axiom 5 (Eq. 1): $\forall A, B, C \in IVFS(X)$,

$$D(A,B) \preceq_o D(\min\{A,B,C\}, \max\{A,B,C\})$$

IV. Conclusion

With the aim of providing a solid Axiom 5 in the definition of dissimilarity between IVFSs, we have started the study of the definitions of supremum and infimum of a set of IVFSs. A first step has been to identify the explicit expressions for the most important orderings between intervals. Another alternative to soften the present formalization of the fifth axiom is to explore other definitions of inclusion between IVFSs. It is our proposal to relax the strong condition of inclusion we have used in previous contributions.

References

[1] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar. Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems*, 220:69–77, 2013.

[2] H. Bustince, C. Marco-Detchart, J. Fernandez, C. Wagner, J.M. Garibaldi, and Z. Takáč. Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders. *Fuzzy Sets and Systems*, 390:23–47, 2020. Similarity, Orders, Metrics.

[3] S. Díaz-Vázquez, E. Torres-Manzanera, I. Díaz, and S. Montes. On the search for a measure to compare interval-valued fuzzy sets. *Mathematics*, 9:3157, 12 2021.

[4] P. Fishburn. Interval ordenings. *Wiley, New York*, 1987.

[5] J. A. Goguen. L-fuzzy sets. *Journal of Mathematical Analysis and Applications*, 18(1):145–174, 1967.

[6] P. Huidobro, P. Alonso, V. Janiš, and S. Montes. Convexity and level sets for interval-valued fuzzy sets. *Fuzzy Optimization and Decision Making*, pages 1–28, 2022.

[7] Z. Takáč, H. Bustince, J.M. Pintor, C. Marco-Detchart, and I. Couso. Width-based interval-valued distances and fuzzy entropies. *IEEE Access*, 7:14044–14057, 2019.

[8] E. Torres-Manzanera, P. Kral, V. Janis, and S. Montes. Uncertainty-aware dissimilarity measures for interval-valued fuzzy sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 28(05):757–768, 2020.

[9] Z. Xu and R. R. Yager. Some geometric aggregation operators based on intuitionistic fuzzy sets. *International Journal of General Systems*, 35(4):417–433, 2006.

Understanding Likes and Dislikes with an Explainable Fuzzy System

Jose Maria Alonso-Moral
Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS)
Universidade de Santiago de Compostela
Santiago de Compostela, SPAIN
https://orcid.org/0000-0003-3673-421X
josemaria.alonso.moral@usc.es

***Abstract*—This paper shows how to design an Explainable Fuzzy System which is ready not only to make accurate classification of songs as likes and dislikes, but also how to explain in natural language the rationale behind such classification regarding numerical features (e.g., acoustics, duration or danceability). First, we have modeled these features, which are vague and imprecise by nature, with global semantics supported by strong fuzzy partitions automatically extracted from data. Then, we have automatically extracted Mamdani-type IF-THEN rules from data. These rules are fired with the usual Mamdani-type min-max inference mechanism and they act as a linguistic layer on top of a fuzzy decision tree which is endowed with global semantics. Thanks to a careful design, the generated fuzzy rule-based classifier is interpretable and it can be deemed as a white box easy to interpret even by non-technical users. Moreover, given a new song, the related fuzzy classification comes along with a textual narrative explanation; what makes the classifier even more accessible and trustworthy for the general public. In practice, automated explanations are just the verbalization of pieces of information associated to an automatic interpretation of the underlying fuzzy inference process. This way of doing is feasible and effective because of the white-box nature of the explainable fuzzy classifier previously designed. Reported results are encouraging because the classifier is characterized by a good balance between performance and interpretability in terms of 10-fold cross-validation. Furthermore, an illustrative example shows how to pave the way from machine learning to humans with the aim of understanding the rationale behind human assessment (in terms of likes and dislikes) associated to well-known songs.**

***Index Terms*—Explainable Artificial Intelligence, Interpretable Fuzzy Modeling, Fuzzy Rule-based Classifiers**

I. Introduction and Related Work

Artificial Intelligence (AI) is becoming pervasive to our daily life, and Explainable AI (or just XAI for brevity) represents an endeavor to endow AI-based systems with explanation capabilities [1]. As a side effect, explainability is expected to contribute to make a responsible and trustworthy AI which ultimately can be even more useful and appreciated in modern society [2]. It is worth noting that XAI can take profit from the ability of fuzzy sets and systems to deal properly with imprecise and vague information [3]. Indeed, Explainable Fuzzy Systems [4], which are carefully designed on top of Interpretable Fuzzy Systems [5], can provide technical and non-technical expert users but also lay end-users with actionable factual and counterfactual human-friendly explanations [6], [7].

The main contribution in this work is the illustration of how to build and use explainable fuzzy systems for generating factual and counterfactual explanations in natural language in a practical use case (Section II). The paper ends with final remarks and pointing out future research work in Section III.

II. Illustrative Use Case

In this section, we pay attention to factual and counterfactual explanations associated with an explainable fuzzy classifier which is automatically learned from the SONGS dataset[1]. This dataset was first used in a Kaggle[2] competition and it contains 2017 data instances. The classification task consists in identifying one out of two classes (Like vs Dislike) in terms of 13 features (Acoustics, Danceability, Duration, Energy, Instrumentalness, Key, Liveness, Loudness, Mode, Speechiness, Tempo, TimeSignature, and Valence). The classifier is built by applying the modeling methodology thoroughly described in [4] and implemented in the open source software GUAJE[3]. GUAJE generates factual and counterfactual explanations in natural language. Notice that, here, for the sake of brevity, Table I summarizes the main results. For comparison purpose, we considered the algorithms Random Forest (RF), J48 and FURIA. More precisely, we used the implementation of these algorithms in the ExpliClas software [8] where they are enriched with linguistic factual explanations.

TABLE I
Interpretability-accuracy trade-off (10-fold cross-validation)

Algorithm	***Accuracy (%)***	***NR***	***TRL***
RF	77.14	-	-
J48	70.85	139	1185
FURIA	72.83	16	65
GUAJE	73.95	125	536

Let us discuss briefly the reported results. On the one hand, RF is a well-known accurate but black-box classifier which is taken as baseline for accuracy in many publications. As expected, it turns up as the most accurate classifier in terms

[1]https://gitlab.citius.usc.es/jose.alonso/xai/-/blob/master/SONGS.txt
[2]https://www.kaggle.com/geomack/spotifyclassification
[3]https://demos.citius.usc.es/guajeOnline/

of the percentage of correctly classified instances (see column *Accuracy (%)* in the table). In addition, we have reported an estimation of the complexity of each classifier in terms of the number of rules (NR) and the total rule length (TRL), which are the two most popular metrics for measuring complexity in the context of fuzzy systems. In this case, RF is made as an ensemble of 100 trees, so in the simplest approach the complexity of the classifier would be computed as the addition of values reported for all the trees. However, unfortunately RF acts as a black box and the ExpliClas software is not able to reporting any value of complexity for the underlying trees. That is the reason why the values of NR and TRL are missing in the table for RF.

On the other hand, J48 is a well-known interpretable classifier which implements a Quinlan's decision tree. Such a tree can be scrutinized from root to leaves with the aim to understand how each single classification is done. Accordingly, J48 is deemed as a white-box classifier and commonly taken as baseline for interpretability in many publications. As expected, it is the least accurate classifier in the table. Also, interestingly, it seems to be the most complex classifier after RF, at least regarding the usual complexity metrics (i.e., it has the highest values for NR and TRL). Since we are addressing a hard classification problem, one single tree struggles to cover properly the whole input space and seems not to be enough (remind that RF combined 100 trees to get acceptable results).

In addition, we have considered two gray-box fuzzy rule-based classifiers. FURIA and GUAJE extract IF-THEN rules from data but FURIA gives priority to accuracy while GUAJE gives priority to interpretability. However, in this illustrative use case, we have forced GUAJE to look for the most accurate system at the cost of penalizing interpretability. Accordingly, GUAJE manages to achieve higher accuracy than FURIA, but also higher values of NR and TRL. At this point, it is also important to remark that GUAJE imposes the use of strong fuzzy partitions with global semantics and this means the generated rules are naturally endowed with linguistic interpretability. On the contrary, FURIA finds out a smaller set of rules with local semantics and rule stretching. Therefore, FURIA lacks linguistic interpretability. Fortunately, ExpliClas is able to set a linguistic layer on top of FURIA rules (as well as it does for J48) and provides users with linguistic explanations. However, such explanations are sometimes misleading and even deceiving because of their low fidelity to the original rules which need to be first approximated linguistically. In other words, in the case that fuzzy sets in the premises of FURIA rules were poorly interpretable, then the linguistic approximation provided by ExpliClas may be too rough and fail to pass from the local semantics (in each single rule) to the global semantics (expected to be cointensive with the user's mental model).

Last but not least, we would like to show how such linguistic explanations look like. For example, we have selected the song "Get Lucky" in the SONGS dataset which is tagged as Like. Indeed, when looking for this song in YouTube we find that it has more than 4M likes. As expected, all algorithms under consideration agree in their classification but the related explanations are slightly different: *The song is classified as Like because:* (**J48**) *danceability, duration and valence are high, instrumentalness, speechness and tempo are low, and energy is medium*; (**FURIA**) *valence and duration are high, and instrumentalness is medium. We have a medium confidence in this classification result because rule firing degree is medium*; (**GUAJE**) *in accordance with rule 64, the song likes in case that energy is not very low and loudness is low. It would dislike if energy were smaller. We have a high confidence in this classification result because rule firing degree is high*.

III. Conclusions and Future Work

We have shown how explainable fuzzy systems can assist experts but also lazy end-users in understanding the rationale behind tagging songs with subjective assessments (i.e., likes and dislikes) in terms of objective information (i.e., numerical features). This kind of systems may be used to increase trust in AI-based recommendation systems but also to assist evaluators in the task of quality assessment for example in the food industry. It is worth noting that this is an ongoing project and as future work, we will compare the reported results with other popular explainers (e.g., LIME or SHAP). In addition, we will address the challenge of running human studies to assess the goodness and effectiveness of different types of explanation.

Acknowledgment

Jose M. Alonso is a *Ramón y Cajal* Researcher (RYC-2016-19802). This work is funded by the Spanish Ministry for Science, Innovation and Universities (grant PID2021-123152OB-C21) and by the Galician Ministry of Culture, Education, Professional Training and University (grants ED431F2018/02, ED431C2018/29, ED431G/08, ED431G2019/04, ED431C2022/19). All grants were co-funded by the European Regional Development Fund (ERDF/FEDER program).

References

[1] A. Barreda et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82–115, 2020.

[2] V. Dignum, *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*. Springer Nature Switzerland AG, 2019.

[3] B. Bouchon-Meunier, A. Laurent, and M.-J. Lesot, *XAI: A Natural Application Domain for Fuzzy Set Theory*. Springer International Publishing, 2022, pp. 23–49.

[4] J. M. Alonso, C. Castiello, L. Magdalena, and C. Mencar, *Explainable Fuzzy Systems. Paving the Way from Interpretable Fuzzy Systems to Explainable AI Systems*. Springer International Publishing, 2021.

[5] J. M. Alonso, C. Castiello, and C. Mencar, *Interpretability of Fuzzy Systems: Current Research Trends and Prospects*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 219–237.

[6] G. Fernandez et al., "Factual and counterfactual explanations in fuzzy classification trees," *IEEE Transactions on Fuzzy Systems*, 2022.

[7] I. Stepin, A. Catala, M. Pereira-Fariña, and J. M. Alonso, "Factual and counterfactual explanation of fuzzy information granules," in *Interpretable Artificial Intelligence: A Perspective of Granular Computing*. Springer International Publishing, 2021, pp. 153–185.

[8] J. M. Alonso and A. Bugarín, "ExpliClas: Automatic generation of explanations in natural language for Weka classifiers," in *IEEE International Conferences on Fuzzy Systems*, 2019.

Fuzzy Agnostic Explainer for Black Box Classifiers

Guillermo Fernandez, Juan A. Aledo, Jose A. Gámez, Jose M. Puerta
Intelligent Systems and Data Mining Lab (SIMD)
University of Castilla-La Mancha
Albacete, Spain
{Guillermo.Fernandez,JuanAngel.Aledo,Jose.Gamez,Jose.Puerta}@uclm.es

***Abstract*—In this work, we tackle the problem of explainability by introducing a novel approach to the well known technique of factual and counterfactual explanations in order to explain why an instance has been classified a certain way by a black box classifier. For that purpose, we follow the methodology proposed by Guidotti et al. in [1] and change the crisp decision tree by a fuzzy decision tree as the interpretable local model that will be trained with a set of generated instances in the neighborhood of the one to be explained.**

***Index Terms*—Explainable Artificial Intelligence, Fuzzy Logic, Metaheuristics, Machine Learning, Rule Based Systems**

I. Introduction

The great increment on computational power in the last decade, and the vast amount of data available, have led to machine learning algorithms being used in numerous fields. More precisely, black box algorithms which benefit greatly from all these data, such as neural networks, ensembles, etc, are showing higher and higher accuracy and can solve problems that a few years back did not have a clear way to be tackled.

The problem with these models lies in their inability to be explained. Their great complexity means that it is not feasible to explain why a certain decision is taken. Because machine learning is expanding to diverse fields, there is a need for the final users to be able to understand the outcome and decision process of a model, particularly in critical applications such as the medical field. It is from this problem where the field of *eXplainable Artificial Intelligence* (XAI) emerges [2].

One well accepted approach to this problem is to build an *explainable* model that can mimic the behavior of the black box in a small region of the space, i.e. the neighborhood of the instance to be explained. If these *explainable* models can generate an explanation regardless of the black box model they are explained, they are called *agnostic*. Some well established proposals are LIME [3] and Anchor [4]. In [1] the authors proposed LORE, an agnostic explanation method which uses a decision tree trained on a neighborhood around the instance to generate factual and counterfactual explanations. In this paper, we present FLORE, an adaptation of LORE which uses a fuzzy decision tree as the interpretable model from which to extract the explanations.

II. Problem Definition

Let us consider a supervised classification problem, consisting of assigning a class c_i belonging to a predetermined set $C = \{c_1, ..., c_m\}$ to an instance $x = (x_1, ..., x_n)$ of values defined over n *input variables* $X_1, X_2, ..., X_n$.

Let us assume that, associated with each input variable X_i, there is a fuzzy (linguistic) variable $F_i = \{v_{i,1} \dots, v_{i,k_i}\}$ defined through a Ruspini partition of k_i ordered fuzzy sets. We will use v_{i,z_i} to denote indistinctly both the fuzzy set and its corresponding associated linguistic label.

Given a value $\delta \in dom(X_i)$, let

$$\mu_i(\delta) = (\mu_{i,1}(\delta), \dots, \mu_{i,k_i}(\delta))$$

be the vector of membership degrees of δ to the k_i fuzzy sets of F_i. In other words, $\mu_{i,z_i}(\delta)$ is the membership degree of the value δ in F_i for the set v_{i,z_i}. Note that because of the type of fuzzy partition considered, $\sum_{z_i=1}^{k_i} \mu_{i,z_i}(\delta) = 1$.

Let B be a black box model already trained so that $B(x) = c \in C$, this is, B can classify the instance x in the class value c from the set of possible class values for the problem.

Our objective is to create an explainer $E(x, c, B)$ that is able to return an explanation $e = \langle R_f, R_{cf} \rangle$, where R_f is a factual and R_{cf} is a counterfactual.

III. Methodology

In this paper, we propose an algorithm that is able to generate an explanation from a black box and an instance, which will be formed by a factual and a counterfactual.

A. Neighborhood

The first step is to generate a neighborhood around the instance. Our algorithm uses the same genetic algorithm in [1] to generate this neighborhood N_x.

B. Fuzzy Decision Tree

The first step is to obtain a new set of fuzzy variables F' defined over N_x. Note that the existing fuzzy variables are not valid because given the local nature of the neighborhood it is likely that all the neighbors in N_x are contained in one of two consecutive fuzzy sets. They will be generated automatically via fuzzy partitioning based on fuzzy entropy, as described in [5]. Then, a fuzzy decision tree is trained using these fuzzy variables and the neighborhood as training data. Particularly, we select the Fuzzy Multiway Decision Tree variant using the aggregated vote inference process.

C. Factual

A factual explanation is the one that help us to explain the classification of an instance into a determined category. In a crisp setting, a factual explanation is usually the single rule fired in the decision tree, which is the only path that leads to the classification.

In fuzzy logic, some proposals considering fuzzy classifiers such as [6], only use the best rule to explain. However, in [7], a study is presented where in a fuzzy setting a factual can be formed by more than a single rule. Given that the inference process takes into account more than a single rule to classify an instance, it can be naive to only take into account a single rule, even if it is the best one of the process of classification. The concept of robustness is introduced, where a factual explanation is considered to be robust when it is composed of enough rules so that, according to the inference mechanism, there do not exist enough fired rules with different consequents to change the outcome.

Our algorithm is compatible with all the factual generation methods explained in that paper, so that, depending on the context or the dataset, one factual explanation or other can be chosen to better suit the necessities.

D. Counterfactual

A counterfactual is defined as the minimum number of changes that have to be made to an instance so that its classification changes. For a crisp setting, this is usually addressed as the minimum number of changes that must be made to the factual rule that explains the instance. However, as explained before, a factual in a fuzzy setting can be formed by multiple rules.

This is why in [7] a counterfactual is defined as a set of changes that modify the classification of the instance. In order to search for the counterfactual, all the rules with different classifications are explored. An order is chosen and the rules are tested iteratively to verify that the changes proposed actually change the class value. As the change proposed is a fuzzy set, it needs to be defuzzified for the change to be applied to the instance.

Again, our algorithm is compatible with both counterfactual search methods explained in that paper, so that both can be applied depending on the context.

E. Explanation mapping

After these steps, the explanation e is defined over the fuzzy variables F'. However, these variables have no linguistic meaning, because a fuzzy set $v'_{i,z_i} \in F'_i$ can mean *high* in the locality but *low* in the global range of the variable. That is why a process of mapping is needed.

This mapping takes a fuzzy set $v'_{i,z'_i} \in F'_i$ returns a fuzzy set $v_{i,z_i} \in F_i$. The universe of F'_i will always be a subset of the universe of F_i. To calculate which set to return, we compute the degree of overlapping between the set v'_{i,z'_i} and all the sets contained in F_i, and return the one with the biggest degree of overlapping. That mapping is then applied to the factual explanation

IV. Experiments

One of the challenges of this work is to find the proper metrics with which to evaluate the explanation. Most of the proposed metrics in the literature are made for crisp classifiers, and adapting them to a fuzzy setting is not trivial.

We will use the metrics proposed in [1]:

- *fidelity*: Measures how similar are the predictions of B and the local explainer in the neighborhood.
- *l-fidelity*: Measures how similar are the predictions of B and the local explainer in the neighborhood for the instances covered by the factual explanation.
- *hit*: Measures if the prediction of B and the local explainer for the instance match.
- *coverage*: Measures how many instances of the dataset are covered by the factual, i.e. how general is the factual.
- *precision*: Measures how many of the covered instances of the dataset match the prediction of the factual.

Most of these metrics use some way of coverage. For a crisp setting it is straight forward, but not for a fuzzy setting. Therefore, it must be carefully chosen when an instance is considered to be covered by a factual explanation [1].

V. Conclusions

This work proposes a novel approach to agnostic explainers by introducing fuzzy logic in the factual and counterfactual generation process. This provides an interesting alternative to the existing methods in the state of the art.

Given that this is an early work in the research line, there is room to expand this work. One such expansions would be to expand on generating natural language explanations from these factuals and counterfactuals so that they can be presented to a final user, taking advantage of the closeness of fuzzy logic to human language.

References

[1] R. Guidotti, A. Monreale, F. Giannotti, D. Pedreschi, S. Ruggieri, and F. Turini, "Factual and counterfactual explanations for black box decision making," *IEEE Intelligent Systems*, vol. 34, no. 6, pp. 14–23, 2019.

[2] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins *et al.*, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information fusion*, vol. 58, pp. 82–115, 2020.

[3] M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[4] ——, "Anchors: High-precision model-agnostic explanations," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.

[5] A. Segatori, F. Marcelloni, and W. Pedrycz, "On distributed fuzzy decision trees for big data," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 1, pp. 174–192, 2017.

[6] I. Stepin, J. M. Alonso, A. Catala, and M. Pereira-Fariña, "Generation and evaluation of factual and counterfactual explanations for decision trees and fuzzy rule-based classifiers," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2020, pp. 1–8.

[7] G. Fernandez, J. A. Aledo, J. A. Gamez, and J. M. Puerta, "Factual and counterfactual explanations in fuzzy classification trees," *IEEE Transactions on Fuzzy Systems*, 2022.

[1] Currently the experiments for FLORE are running, and they will be ready to be published for the date of the conference

IFC-BD: An Interpretable Fuzzy Classifier for Boosting Explainable Artificial Intelligence in Big Data*

Fatemeh Aghaeipoor
School of Computer Science
Institute for Research in Fundamental Sciences (IPM)
Tehran, Iran
f.aghaei@ipm.ir, 0000-0002-4733-0432

Alberto Fernandez
Andalusian Research Institute in Data Science and Computational Intelligence (DASCI)
University of Granada
Granada, Spain
alberto@decsai.ugr.es, 0000-0002-6480-8434

***Abstract*—This is a reduced version of our article published in IEEE Transactions on Fuzzy Systems (IEEE-TFS), presented to ESTYLF 21/22 KeyWorks. In said manuscript, we addressed the classification task in the Big Data context. To do so, we proposed IFC-BD, an Interpretable Fuzzy Classifier for Big Data, aiming at boosting the horizons of explainability by learning a compact yet accurate fuzzy model. IFC-BD was developed in a cell-based distributed framework through the three working stages of initial rule learning, rule generalization, and heuristic rule selection. The findings of the experiments revealed that the proposed algorithm was able to improve the explainability of fuzzy rule-based classifiers as well as their predictive performance.**

***Index Terms*—Explainable Artificial Intelligence, Fuzzy Rule-Based Classification Systems, Big Data, Interpretability, Apache Spark Framework, Scalability**

I. Introduction

Developing and using Machine Learning solutions in the context of Big Data analytics have become a growing demand nowadays. However, due to the higher complexity of this working scenario, and the intrinsic constrains of the models that are being used, it might be difficult to understand the system behavior. This issue contradicts the objectives of the new resurgence of Artificial Intelligence known as eXplainable Artificial Intelligence (XAI) [1].

The objectives of the XAI trend demand the use of more explainable, interpretable, and transparent systems. The ultimate goal is to improve the trust for Machine Learning models in general areas of application. In this regard, the use of rule-based systems in general, and Fuzzy Rule-based Classification Systems (FRBCSs) in particular is advisable [2].

Specifically, FRBCSs provide two key issues for promoting XAI: (1) the use of linguistic fuzzy labels containing a semantic knowledge inspired by the human language; and (2) simple linguistic rules with short antecedents that are known to be manageable for the human user.

This work was partially supported by Andalusian frontier regional project A-TIC-434- UGR20 and by the Spanish Ministry of Science and Technology under project PID2020-119478GB-I00 including European Regional Development Funds.

As commented in the beginning of this work, models developed under the umbrella of Big Data, must face some intrinsic challenges: they are mostly computationally intensive, and even if they can come up with scalable solutions, they may be complex in terms of their Data Base (DB) or Rule Base (RB). In the former case, the conceptual difficulties for the user are related to a high number of variables, as well as high number of granularities. In the latter case, an excessive number of redundant rules or those covering non-dense areas or outliers are not useful.

In our study [3], we proposed an Interpretable Fuzzy Classifier for Big Data (IFC-BD), which aimed at learning a compact yet accurate FRBCS containing less number of short rules. Its design it provided a confident and reliable RB based on a lower granularity, also preserving the original semantics of the linguistic fuzzy labels. The full implementation was carried out in Scala, using the data structures and primitives for the Spark framework.

The remainder of this keywork contribution is structured as follows. Section II explains the methodology used in order to fulfill the proposed goals. Then, Section III presents the conclusions reached.

II. Methodology

The IFC-BD approach was developed following the XAI guidelines to obtain a transparent and interpretable system. To do so, the main features of are listed below, which comprises

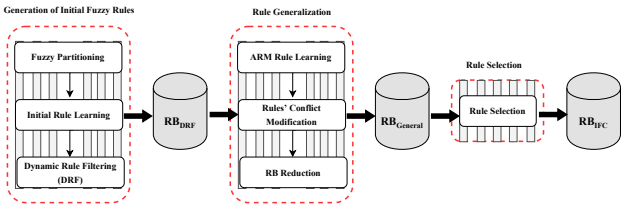


Fig. 1. The general schema of IFC-BD: an illustration of the RB evolution.

TABLE I
RESULTS OF BOTH PERSPECTIVES

Algorithm	CFM-BD^3				CFM-BD^5				Chi-BD-DRF				IFC-BD			
Dataset	#Rules	$\overline{\text{RL}}$	$\overline{\text{TRL}}$	ACC	#Rules	$\overline{\text{RL}}$	$\overline{\text{TRL}}$	ACC	#Rules	$\overline{\text{RL}}$	$\overline{\text{TRL}}$	ACC	#Rules	$\overline{\text{RL}}$	$\overline{\text{TRL}}$	ACC
susy	200.4	2.4	1444	**75.3**	330.4	2.4	3971	73.3	1002.6	18	54140	66.0	**55.6**	2.9	**490**	69.6
BNG-h	**144.2**	2.4	**1045**	84.0	234	2.4	2860	83.9	3230.8	13	126001	**86.4**	157.8	3.0	1398	86.3
BNG-Au	144	2.5	1089	82.0	239.4	2.5	3005	84.2	910.8	14	38254	84.5	**54.6**	2.8	**459**	**84.8**
cov1	443.4	2.5	3361	73.4	738.6	2.6	9477	73.6	1408.4	54	228161	**77.4**	**206.4**	3.0	**1850**	73.9
cov2	436	2.6	3402	72.7	740.8	2.6	9622	74.5	1393	54	225666	**75.6**	**248**	3.0	**2228**	70.3
cov1-vs-2	464.2	2.6	3646	75.0	740	2.6	9632	75.8	1216	54	196992	**76.1**	**188.6**	3.0	**1689**	75.7
higgs	**220.2**	2.7	**1780**	60.8	425.6	2.6	5556	**61.7**	77649.7	28	6522572	57.4	222.6	3.0	1975	56.3
skin	20.4	2.1	127	91.9	47.8	2.2	530	**94.4**	**5**	3	45	89.9	**5**	1.6	**24**	91.9
hyp	102	2.5	774	66.4	172	2.5	2170	66.1	10483.2	10	314496	**67.5**	**81.2**	3.0	**725**	67.3
Avg	241.64	2.5	1852	75.7	407.62	2.5	5203	**76.4**	10811.05	28	856259	75.6	**135.53**	2.8	**1204**	75.1

1) An initial baseline fuzzy classification model was computed in an scalable way, considering all possible rules from the input space. Then, a compact fuzzy rule set was obtained from high density areas by means of a dynamic rule filtering scheme. The former followed the original the Chi-BD-DRF design [4], i.e. rules were generated through three steps of Fuzzy Partitioning, Initial Rule Learning, and Dynamic Rule Filtering (using the support of the rules).
2) General and interpretable rules were extracted from the previous set by means of an enhancement process. To do so, three issues were considered: (1) general rules were derived from the available specific rules through an association rule mining learning scheme; (2) the obtained general rules were aggregated and the possible rules' conflicts are modified using a novel estimated fuzzy rule weight, partially based on the Penalized Certainty Factor; (3) some of the low-confident rules were discarded for the sake of further interpretability improvement. The output was a condense set of shorter antecedent rules, easily manageable by the human-user.
3) Finally, the most influential rules were selected from a heuristic rule selection mechanism. Specifically, a novel metric was considered, namely rule was $P\%$ well-performing if it could classify $P\%$ of its covering examples. Rules were then examined in a class-wise way and those satisfying a minimum user-defined performance threshold were selected to be in the final RB.

For the experimental analysis, we considered both the accuracy and interpretability perspective. In the former case, the standard accuracy rate was selected. In the latter, three measures of #Rules (all available rules), $\overline{\text{RL}}$ (average rule length), and $\overline{\text{TRL}}$ (total length of all the rules) were used. In the case of comparing methods, Chi-BD-DRF [4] and CFM-BD [5] were employed in the experiments. On the one hand, Chi-BD-DRF was the baseline contribution of the study. On the other hand, CFM-BD was the current state-of-the-art fuzzy classifier proposed for Big Data problems. All results are shown in Table I.

Regarding interpretability, given the results in the related metrics per dataset, as well as the overall averages indicated in the last row, it was argued that IFC-BD most likely fulfills the terms of interpretability better than the other models. In the case of predictive performance, values are close to each other in the majority of the datasets. Nevertheless, we have to take this into account Chi-BD-DRF applies many numbers of long rules to yield the best ACC ranking and CFM-BD^5 requires a more complex system composed of more fuzzy sets to achieve this classification performance.

III. CONCLUDING REMARKS

This keywork paper has revisited the published work made in [3], in which IFC-BD, an interpretable fuzzy classification model for Big Data Analytics, was proposed. Its design was focused to be scalable and to obtain a compact rule set of linguistic fuzzy rules, while maintaining a high predictive performance. To achieve this, a complete and efficient learning workflow was developed under the Spark framework. The final RB was obtained by different stages, namely generation, filtering, generalisation, and heuristic rule selection. The effectiveness of IFC-BD was evaluated using 9 different classification datasets versus the baseline and state-of-the-art solutions, showing a very robust behavior in terms of classification performance and complexity measures.

REFERENCES

[1] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inform. Fusion*, vol. 58, pp. 82 – 115, 2020.
[2] A. Fernandez, F. Herrera, O. Cordon, M. J. del Jesus, and F. Marcelloni, "Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?" *IEEE Comput. Intell. Mag.*, vol. 14, no. 1, pp. 69–81, 2019.
[3] F. Aghaeipoor, M. M. Javidi, and A. Fernández, "IFC-BD: An interpretable fuzzy classifier for boosting explainable artificial intelligence in big data," *IEEE Trans. on Fuzzy Syst.*, vol. 30, no. 3, pp. 830 – 840, 2022.
[4] F. Aghaeipoor, M. M. Javidi, I. Triguero, and A. Fernández, "Chi-BD-DRF: Design of scalable fuzzy classifiers for big data via a dynamic rule filtering approach," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2020, pp. 1–7.
[5] M. Elkano, J. A. Sanz, E. Barrenechea, H. Bustince, and M. Galar, "CFM-BD: A distributed rule induction algorithm for building compact fuzzy models in big data classification problems," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 1, pp. 163–177, 2020.

Revisión de modelos de lógica difusa en el ámbito de XAI aplicados a la detención de fraude en tarjetas de crédito

Ángela Escobar
Xauen Cyber Science S.L.
23071 Jaén, España
angela.escobar@xauencybersecurity.com

Pedro González, María José Del Jesus
Departamento de Informática,
Universidad de Jaén
23071 Jaén, España
{pglez, mjjesus}@ujaen.es

Santiago Moral
DCNC Sciences,
Universidad Rey Juan Carlos
28028 Madrid, España
santiago.moral.rubio@urjc.es

***Abstract*—Este artículo analiza el estado del arte en la aplicación de modelos de lógica difusa a la detención de fraude bancario y la forma de tratar la explicabilidad en los mismos, esbozando importantes líneas de investigación abiertas.**

***Index Terms*—lógica difusa, inteligencia artificial explicable, detención de fraude**

I. Introducción

En los últimos años, el uso de tarjetas bancarias, las compras por internet y la banca móvil han experimentado un gran crecimiento, en parte como consecuencia de la transformación digital que se está llevando a cabo en las actividades diarias, tanto en las empresas como en las instituciones que dan servicio directo al ciudadano. Esta transformación digital se ha visto acelerada a raíz de la pandemia provocada por la COVID-19.

Algunos de los problemas que se han agravado por esta transformación digital están vinculados a ataques de *phishing*, *malware*, *carding* o robo de identidad que permiten a los ciberdelicuentes utilizar nuestra información para conseguir grandes beneficios sin un gran esfuerzo. Uno de los principales problemas que provocan estos ataques es el fraude bancario. En este estudio nos centraremos en revisar los modelos de detención de fraude en tarjetas bancarias, fenómeno conocido como *carding*, en los que se utilizan modelos de lógica difusa y estudiaremos la aproximación de éstos hacia modelos de Inteligencia Artificial Explicable o *Explainable Artificial Intelligence* (XAI). XAI hace referencia a la explicabilidad de los métodos utilizados en inteligencia artificial y de los resultados obtenidos por los mismos [1], [2]

En este artículo revisaremos qué se entiende por fraude en tarjetas de crédito. En la sección III revisaremos propuestas en las que se aplica lógica difusa y estudiaremos tanto cómo afrontan la detención de fraude como la explicabilidad del modelo. Por último, describiremos conclusiones y líneas de trabajo abiertas.

II. Fraude en tarjetas de crédito

El fraude en tarjetas bancarias, se produce cuando los datos de la tarjeta se utilizan para efectuar compras o retirar efectivo sin autorización de su propietario [3]. Teniendo en cuenta lo indicado en [4] y [5] podemos categorizar el tipo de fraude según su forma en dos tipos: Tarjeta no presente (CNP, *Card Not Present*), se produce cuando el defraudador conoce los datos de la tarjeta y puede realizar transacciones sin necesidad de tener la tarjeta física en su poder y Tarjeta presente (CP, *Card Present*) se produce cuando se realizan transacciones fraudulentas utilizando una tarjeta física, mediante su robo o falsificación. Revisando datos específicos del fraude bancario, destaca el estudio publicado por el Banco Central Europeo (BCE) [6] que abarca hasta el año 2019, en el que se observa en general una disminución en el fraude bancario a partir de tarjetas de crédito aunque con un incremento en CNP.

III. Algoritmos de lógica difusa aplicados a fraude

Se han realizado diferentes estudios en los que se aplica lógica difusa para la detención de fraude en tarjetas de crédito. La mayoría de las propuestas se centran en la obtención de modelos que maximicen la precisión en la predicción. En la revisión bibliográfica no se han encontrado propuestas que aborden el diseño de modelos de predicción del fraude en tarjeta centrando el interés en la perspectiva XAI, es decir en la extracción de información inteligible y de interés para los expertos sobre el proceso de decisión, que incremente la transparencia y justifique la decisión aportada.

En [7] se determina si una transacción es fraudulenta con un esquema en tres fases: primero se verifican los datos de la tarjeta, después se aplica el algoritmo difuso C-means para determinar si la transacción sigue los patrones de uso normales del usuario según su actividad anterior, y por último, si la transacción se considera sospechosa, se aplica un mecanismo de aprendizaje basado en redes neuronales con retropropagación (FFNNBP) para determinar si se trata de una actividad fraudulenta o de una desviación ocasional de un usuario auténtico. La propuesta se aplica a un *dataset* creado utilizando el simulador descrito en [8] para generar transacciones sintéticas que representan el comportamiento usuarios legítimos y fraudulentos. Su rendimiento se evalúa usando la

métrica ROC que relaciona el ratio de verdaderos positivos frente al de falsos positivos. En este estudio, los autores no tienen en cuenta las características para que el modelo sea explicable, pero la categorización de las transacciones obtenida a partir del comportamiento de los usuarios podría facilitar su comprensión por un experto.

Los mismos autores desarrollan en [9] una propuesta parecida formada por un sistema basado en reglas difusas tipo Mandani (que clasifica las transacciones como genuinas o fraudulentas) y una red neuronal entrenada con un método de retropropagación de regulación bayesiana (que procesa las transacciones clasificadas como fraudulentas para dar una predicción final). Se analizan los resultados, de nuevo sobre datos simulados y desde un punto de vista predictivo. La capacidad explicativa de la propuesta, avalada por un sistema difuso inicial formado por solo 9 reglas tipo Mandani sencillas, se decrementa con el uso posterior de una red neuronal.

En [5] se presenta el algoritmo IDFTC4.5 para generar a a partir de datos un árbol de decisión con el algoritmo C4.5 pero basado en la extensión de la lógica difusa propuesta por Atanassov, la lógica difusa intuicionista que considera valores de pertenencia, no pertenencia e indeterminación. El sistema funciona de forma que cuando se produce una transacción, ésta se une al *dataset* y se calcula su grado de pertenencia, no pertenencia y el grado de indeterminación en base al modelo matemático propuesto. Si el grado de indeterminación no supera el umbral establecido, esta transacción se envía al administrador para su evaluación. En el caso de superarlo, formará parte del conjunto de datos que utilizará el algoritmo para la construcción del árbol, con una etapa de poda para simplificar el modelo. La propuesta se aplica a un conjunto de datos de un banco de Singapur [10]. En el modelo obtenido se analizan diferentes medidas asociadas a precisión (sensibilidad, ratio de falsos positivos y negativos, precisión, *accuracy* y *f-measure*), obteniéndose mejores resultados que los alcanzados por *Random Forest* o máquinas de soporte vectorial (SVM). No se realiza un estudio ni se aportan datos sobre la interpretabilidad del modelo (medida a través de la complejidad del árbol obtenido o del conjunto de reglas derivado), aunque cabe esperar que el nivel de explicabilidad sea superior al de *Random Forest* o SVM.

En [11] se propone un sistema que fusiona la información que aportan un sistema experto basado en reglas difusas con un modelo de comportamiento Fogg. Se analiza el modelo obtenido a partir de un conjunto de datos privados de un proveedor de servicios de pago. De nuevo, solo se estudia desde el punto de vista predictivo, no se considera la explicabilidad del modelo dado que no se aporta información descriptiva sobre los de modelos individuales considerados en la fusión ni sobre la forma de realizar esta. Aunque ninguno de los modelos analizados en este estudio trata de forma específica la explicabilidad de los modelos, combinan lógica difusa con otras técnicas de Inteligencia Artificial y son un punto de partida a considerar en el desarrollo de propuestas XAI para la detección de fraude en tarjetas, dada la capacidad descriptiva de la lógica difusa.

IV. Conclusiones

Los sistemas de detención de fraude (FDS) se han convertido en una herramienta fundamental en la operativa bancaria. Estos sistemas deben clasificar en tiempo real las transacciones que se producen y ofrecer una explicación que responda de forma óptima al usuario final. En esta revisión se ha comprobado que son pocos los modelos de detención de fraude en los que se aplica lógica difusa y ninguno de ellos permite obtener modelos explicables que cumplan con los requerimientos XAI.

Hay múltiples líneas de trabajo abiertas en el área de la detección de fraude para obtener modelos precisos e interpretables. Entre ellas destacan el establecimiento de medidas de evaluación de los conceptos XAI [2], las propuestas de métodos de obtención de modelos interpretables o propuestas de métodos de generación de líneas de explicabilidad, como *Shapley Additive exPlanations* (SHAP) y *Local Interpretable Model-Agnostic Explanations* (LIME), en el dominio de la detección de fraudes con tarjetas de crédito aplicando lógica difusa [12].

Reconocimientos

Este trabajo ha sido parcialmente apoyado por el Ministerio de Ciencia e Innovación de España en el marco del proyecto PID2019-107793 GB-I00 financiado por AEI/ 10.13039/501100011033.

Referencias

[1] Adadi A, Berrada M (2018). Peeking inside the black box: A survey on explainable artificial intelligence (XAI), IEEE Access, vol 6(52), pp. 138-169.

[2] Escobar A, González P & Del Jesus MJ (2021). Revisión y análisis de conceptos en Inteligencia Artificial Explicable. Una aproximación a la unificación de la terminología. XIX Conferencia de la Asociación Española para la Inteligencia Artificial, pp. 41-46.

[3] N26 (2021). Fraudes con tarjeta de débito: todo lo que necesitas saber. Online. https://n26.com/es-es/blog/fraude-tarjeta-de-debito

[4] Jain Y, Tiwari N, Dubey S & Jain S (2019). A comparative analysis of various credit card fraud detection techniques. International Journal of Recent Technology and Engineering, 7(5), pp. 402–407.

[5] Askari SMS & Hussain MA (2020). IFDTC4.5: Intuitionistic fuzzy logic based decision tree for E-transactional fraud detection. Journal of Information Security and Applications, 52

[6] Banco Central Europeo (2021). Seventh report on card fraud. Online. https://www.ecb.europa.eu/pub/cardfraud/html/ecb.cardfraudreport202110~cac4c418e8.en.html

[7] Behera TK & Panigrahi S (2015). Credit Card Fraud Detection: A Hybrid Approach Using Fuzzy Clustering & Neural Network. Proceedings - 2015 2nd IEEE International Conference on Advances in Computing and Communication Engineering, ICACCE 2015, pp. 494–499

[8] Panigrahi S, Kundu A, Sural S & Majumdar A (2009). "Credit card fraud detection a fusion approach using Dempster–Shafer theory and bayesian learning", Information Fusion, vol 10(4), pp. 354–363

[9] Behera TK & Panigrahi S (2017). "Credit Card Fraud Detection Using a Neuro-Fuzzy Expert System". In: Behera, H., Mohapatra, D. (eds) Computational Intelligence in Data Mining. Advances in Intelligent Systems and Computing, vol 556

[10] Quah JT, Sriganesh M (2008). Real-time credit card fraud detection using computational intelligence. Expert Syst Appl, vol 35(4), pp. 1721–32.

[11] Haratinik MR, Akrami M, Khadivi S & Shajari M (2012). "FUZZGY: A hybrid model for credit card fraud detection" 6th International Symposium on Telecommunications (IST), pp. 1088-1093

[12] Psychoula I, Gutmann A, Mainali P, Lee SH, Dunphy P & Petitcolas F (2011). "Explainable Machine Learning for Fraud Detection", Computer, vol 54 (10), pp. 49-59

Sistemas difusos en el borde con autoajuste en línea

Javier Martín Moreno
Dept. de Tecnologías de la Información
Universidad de Huelva
Huelva, España
javier.martin@dti.uhu.es

Samuel López Domínguez
Dept. de Tecnologías de la Información
Universidad de Huelva
Huelva, España
samuel.lopez@dti.uhu.es

Antonio Peregrín Rubio
Instituto Andaluz Interuniversitario de Investigación en Data Science and Computational Intelligence (DaSCI)
Centro de Estudios Avanzados en Física, Matemáticas y Computación
Universidad de Huelva
Huelva, España
peregrin@dti.uhu.es

***Resumen*—** Pequeños dispositivos nos acompañan día a día en infinidad de aplicaciones, utilizando un hardware optimizado, frecuentemente conectados a Internet, y en muchos casos empleando algoritmos de Inteligencia Artificial, con capacidad incluso de aprendizaje o ajuste. En este trabajo se discuten preliminarmente las necesidades de los Sistemas Basados en Reglas Difusas con capacidad de autoajuste en dispositivos hardware con recursos limitados, útiles en aplicaciones actuales.

***Palabras Clave*—** *Sistemas Basados en Reglas Difusas, TinyML, IoT, Edge Computing, Online Learning*

I. Introducción

La computación en el borde (*Edge Computing, (EC)*) [1], es un paradigma que tiene como objetivo procesar los datos en el propio dispositivo, en lugar de enviarlos a la nube para hacerlo mediante grandes clústeres con capacidades de cómputo escalable. De esta forma, se consiguen varias ventajas como:

1) Privacidad de los datos: Al procesarse in situ, los datos no tienen que salir del dispositivo.

2) Tiempo real: La latencia es muy baja o nula, pues sin usar comunicaciones, la respuesta es en tiempo real.

3) Seguridad: Al estar aislados, se previenen mejor los posibles ataques maliciosos más allá del propio dispositivo.

4) Eficiencia energética: Se eliminan los consumos innecesarios, pues se procesa sólo lo indispensable.

Entorno al EC, surge el *TinyML*, un nuevo paradigma que propone integrar los algoritmos clásicos de Aprendizaje Automático (*Machine Learning* (ML)), para dotar de inteligencia a los dispositivos basados en microcontrolador con fuentes de alimentación limitadas [2].

Dentro de esta disciplina, al igual que se adaptaron las técnicas de ML a entornos escalables, hoy en día se están adaptando a los pequeños dispositivos. La opción más popular y accesible es el uso de diversos *frameworks* como *TensorFlow Lite* de Google o *Embedded Learning Library* de Microsoft, entre otros, fundamentalmente orientados a Deep Learning. Esto nos permite desplegar estos modelos de ML en dispositivos como Arduino Nano 33 o Raspberry Pi Pico de forma rápida. Por otro lado, para obtener un rendimiento óptimo y optimizar el consumo de los recursos, podemos implementar nuestros propios modelos de ML, adaptándolos a las restricciones de estos dispositivos, tales como una memoria RAM muy limitada, del orden de KB; almacenamiento limitado, del orden de sólo unos pocos MB; frecuencias de reloj bajas, en el orden de MHz; y muy eficientes en cuanto a consumo energético, pues se fabrican para alimentarse en muchas situaciones con pequeñas baterías.

Así, más allá del uso de *frameworks*, podremos implementar técnicas de ML clásicas para obtener un modelo base que posteriormente podremos adaptar para inferir en un dispositivo con recursos limitados.

Siguiendo esta filosofía, en este trabajo dirigimos la vista al uso de técnicas de Inteligencia Computacional, y en concreto al uso de Sistemas Basados en Reglas Difusas (SBRDs). Estos modelos representan el conocimiento obtenido de los datos de una manera más cercana al conocimiento humano, mediante reglas y términos lingüísticos, además de una alta capacidad para manejar incertidumbre. Por ello, estos dispositivos usados hoy en día en aplicaciones de *IoT* y *EC* pueden estar potenciados por el aprendizaje y ajuste de SBRDs [3].

Este trabajo se organiza así: en la sección II se describe el estudio de los modelos, en la sección III se comentan las aplicaciones, y finalmente en la IV exponer las conclusiones.

II. Descripción del modelo

Actualmente podemos distinguir dos filosofías a la hora de usar modelos de ML en dispositivos de recursos limitados:

1) Por un lado, la filosofía de *TinyML* [4] genérica*,* donde el modelo se aprende o ajusta fuera del dispositivo, para posteriormente adaptarlo a los requerimientos del mismo, de modo que el dispositivo sólo se utiliza para inferir.

2) Por otro lado, la filosofía *TinyOL,* o *TinyML con aprendizaje online* [5]*,* que propone el aprendizaje del modelo, o al menos el ajuste de alguna parte del mismo, en el propio dispositivo al tiempo que realiza su función.

Si bien el primero de ellos es el más frecuente, la filosofía *TinyOL* ofrece una ventaja de interés tal como la capacidad de adaptarse a un entorno que cambie con el tiempo. Sin embargo, para que esto pueda llevarse a cabo, el dispositivo debe poder evaluarse, es decir, cuantificar de la forma más objetiva posible, la precisión con la que está actuando sobre el sistema, lo cual en algunas aplicaciones puede no ser posible.

En este trabajo se propone un mecanismo simple de ajuste en línea para SBRDs (para regresión o clasificación), es decir, un modelo *TinyOL* Difuso sencillo para aplicaciones reales, cuyo funcionamiento consiste en alternar a lo largo del tiempo, períodos de uso de su Base de Conocimiento (BC) actual, con procesos de actualización de esta, basados en los resultados de su autoevaluación llevados a cabo en el propio dispositivo.

La citada BC inicial puede provenir bien de datos previos, si están disponibles, obteniendo las reglas mediante técnicas clásicas de cubrimiento o *clustering*, o reglas de experto.

En concreto, el mecanismo funciona de la siguiente forma:

- El flujo de datos de entrada que recibe el dispositivo se emplea tanto para ofrecer una salida inferida por el SBRD, como para almacenarse localmente, con objeto de crear de un pequeño conjunto de datos local en el propio dispositivo.
- Transcurrido un período de tiempo de trabajo real del SBRD, cuya duración dependerá de un valor apropiado a la aplicación en cuestión, se lleva a cabo un proceso de ajuste de la BC utilizando los datos de ese período de tiempo.

La actualización de una BC para mejorar o mantener su precisión ante posibles cambios del entorno utiliza un ajuste de las etiquetas lingüísticas de sus variables como estrategia ampliamente conocida y válida para este fin. En la Figura 1 se ilustra el mecanismo propuesto; la implementación del mismo debe realizarse de modo que requiera pocos recursos y esfuerzo computacionales. En ese sentido, se deben emplear estrategias orientadas a reducir esta carga tales como:

- Diseñar el SBRD lo más compacto posible, es decir, reducir el número de variables a las estrictamente necesarias, y utilizar un número etiquetas en cada variable muy reducido, de modo que el número de parámetros a ajustar sea pequeño. Estos dos factores anteriores también favorecen un reducido número de reglas, lo cual es necesario para una mejor eficiencia en la fase de evaluación.
- Utilizar ajuste lateral de las etiquetas lingüísticas, con un parámetro (posición de la etiqueta en el universo de discurso de la variable), o a lo sumo dos (posición y anchura) por etiqueta, asociado a una metaheurística de búsqueda eficiente. Ésta, se puede simplificar asimismo empleando valores discretos en lugar de valores continuos.

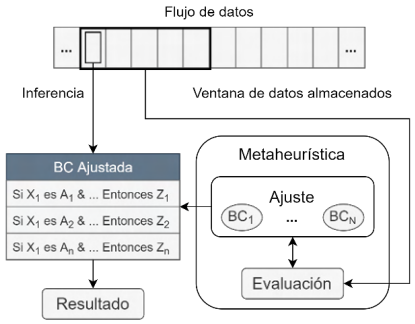


Fig. 1. Diagrama del modelo difuso en el borde con autoajuste online

Dependiendo de la aplicación, el proceso de evaluación puede tener diferentes alternativas:

- Sobre el entorno real: consiste en evaluar las BCs alternativas candidatas sobre la propia aplicación real. El inconveniente es que se trata de un proceso que no siempre es posible llevar a cabo por lo comprometido de la aplicación (es decir, no admita someter a prueba alternativas peores), o la lentitud.
- Sobre una simulación exacta: determinadas aplicaciones, una vez se dispone de los datos recientes, es posible calcular la precisión resultante de utilizar otras alternativas diferentes. Son casos poco frecuentes, pero la evaluación es precisa, por lo que es la opción preferible de ser viable.
- Sobre un modelo subrogado: cuando no es admisible evaluar el modelo sobre el entorno real y la simulación exacta no es una opción, pero es posible disponer de un modelo de la aplicación. Su ventaja es la rapidez, y su inconveniente es requerir un modelo actualizado que puede no ser siempre fiel a la evolución de la aplicación real.

III. Aplicaciones del modelo

Un modelo difuso autoadaptativo, apto para dispositivos *EC* e *IoT*, resulta de interés para gran variedad de aplicaciones. En este trabajo proponemos como ejemplo de aplicación, el desarrollo de un sistema de gestión de fuentes de energía eléctrica para instalaciones de energía renovable en viviendas dotadas de sistemas fotovoltaicos (con paneles solares y almacenamiento: baterías) no aislados, es decir, con conexión a la red eléctrica. Es un modelo de interés debido a que no recurre a grandes baterías (elemento más costoso de estas instalaciones) como lo haría un sistema autónomo, sino a una menor y más económica, para reducir la factura energética.

El objetivo de nuestro modelo, implementado en un dispositivo *EC*, es tomar la decisión de consumir energía renovable o energía convencional de la red eléctrica, de manera que se optimice el coste para usuario de la vivienda. Las decisiones se tomarán en base a factores relevantes para nuestro objetivo, siendo las entradas: el consumo de energía esperado en la vivienda, la energía renovable disponible almacenada, la irradiación solar esperada y el coste de la energía convencional. Por ejemplo, si disponemos de suficiente energía renovable, el consumo esperado es bajo y el coste de la energía convencional es elevado, preferiremos usar energía renovable. Los detalles de la aplicación y sus resultados experimentales se mostrarán en el congreso por limitaciones de espacio y disponibilidad actual.

IV. Conclusiones

Gracias a los avances tecnológicos es posible dotar de inteligencia a los dispositivos con recursos limitados, haciendo que sean capaces de aprender y autoadaptarse a su entorno. Los sistemas difusos son un elemento clásico histórico dentro de estos dispositivos limitados (por ejemplo, en los sistemas de control), gracias tanto a su explicabilidad inherente y un buen balance entre ella y su precisión, así como por sus relativamente restringidas necesidades computacionales. En este trabajo se estudia preliminarmente una forma sencilla de potenciarlos adaptándolos a la filosofía de *TinyOL,* utilizando mecanismos de ajuste que puedan llevar integrados en lugar de ser estáticos, ilustrándolo con un posible mecanismo de ajuste en línea, y se apunta su utilidad, empleado en problemas de actualidad.

Referencias

[1] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, A. Ahmed, "Edge computing: A survey" Future Generation Computer Systems, 2019, vol. 97, pp. 219-235.

[2] C. R. Banbury, et al. "Benchmarking TinyML systems: Challenges and direction," arXiv preprint arXiv:2003.04821, 2020.

[3] A. Fernandez, V. Lopez, M. J. del Jesus, F. Herrera, "Revisiting evolutionary fuzzy systems: Taxonomy, applications, new trends and challenges" Knowledge-Based Systems, 2015, vol. 80, pp. 109-121.

[4] R. Sanchez-Iborra, A. F. Skarmeta, "TinyML-enabled frugal smart objects: Challenges and opportunities," IEEE Circuits and Systems Magazine, 2020, vol. 20, no 3, pp. 4-18.

[5] H. Ren, D. Anicicand T. A. Runkler, "TinyOL: TinyML with online-learning on microcontrollers," 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021. pp. 1-8.

Generación de protoformas en una arquitectura IoT para la atención a la pobreza energética

1st David Díaz Jiménez
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0003-1791-4258

2nd Javier Medina Quero
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0002-8577-8772

3rd José M. Serrano
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0001-5046-0724

4th José Cruz Martínez
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0002-0010-6572

5th Juan Carlos Casas Rosa
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0002-4744-8123

6th Macarena Espinilla Estévez
Departamento de Informática
Universidad de Jaén
Jaén, España
0000-0003-1118-7782

***Resumen*—El seguimiento de indicadores de pobreza energética es fundamental para mejorar el conocimiento y dirigir esta problemática a nivel gubernamental. En este sentido, el paradigma de Internet de las Cosas es clave para monitorizar remotamente los hogares. En este resumen presentamos una arquitectura IoT novedosa para la generación de informes en lenguaje natural a través de protoformas lingüísticas para mejorar la comprensión de dichos informes.**

***Puntos clave*—Internet de las Cosas, protoformas lingüísticas difusas, pobreza energética, arquitectura *cloud*, *fog* y *edge*, computación ubicua.**

I. Introducción

La pobreza energética es uno de los problemas más importantes a los que se enfrenta la sociedad y es sufrido por las personas con menos recursos. Un hogar que padezca de pobreza energética es aquel que no puede acceder a los servicios energéticos esenciales. Dicha pobreza es un factor que repercute en múltiples factores, entre ellos la salud, siendo más notable su impacto en la salud de las personas mayores.

Es fundamental el seguimiento de indicadores de pobreza energética para mejorar el conocimiento y dirigir esta problemática para cumplir con el compromiso adquirido en estrategias nacionales y mundiales [1]. El paradigma de Internet de las Cosas (Internet of Things - IoT) ha sido demostrado como una herramienta eficaz para monitorizar de forma remota y en tiempo real indicadores de pobreza energética. Dicho enfoque se sustenta en el despliegue de sensores en los hogares para realizar el seguimiento de parámetros clave como la temperatura, la humedad o la calidad del aire con el fin de detectar casos complejos.

Generalmente, los informes de pobreza energética están representados con datos cuantitativos junto con representaciones gráficas. Sin embargo, se ha evidenciado que el lenguaje natural es más eficaz que tales representaciones en situaciones, como la pobreza energética, donde es necesario examinar una gran cantidad de datos. Las protoformas lingüísticas difusas permiten describir de manera lingüística [2], la información procesada utilizando lenguaje natural, a través de variables, ventanas temporales difusas y cuantificadores.

Este resumen presenta una arquitectura IoT con capas en el borde, en la niebla y en la nube (en inglés edge, fog y cloud) para la generación de protoformas lingüísticas relativas a pobreza energética. Así, la principal ventaja de la arquitectura es proporcionar descripciones lingüísticas de pobreza energética. Para ello, la propuesta permite un primer nivel de computación de la protoforma en la capa niebla, el cual aligera la comunicación con la capa nube, y permite además realizar un segundo nivel de computación de la protoforma en la capa nube de manera personalizada.

Este resumen se estructura del siguiente modo. En primer lugar, se describe la arquitectura IoT en la Sección II y, posteriormente, la computación de las protoformas lingüísticas difusas en la Sección III. Finalmente, el resumen concluye en la Sección IV.

II. Arquitectura IoT

En esta sección, se presenta la arquitectura IoT, detallándose en primer lugar las capas y, en segundo lugar, la comunicación entre las mismas. La Figura 1 muestra una representación de la arquitectura propuesta, la cual está compuesta de las siguientes capas:

- La capa borde está formada por el conjunto de dispositivos que se ubicarán en cada hogar. Este dispositivo consistirá en una caja, que hemos denominado Green Box, la cual integra sensores de temperatura, que se encargan de recoger los datos de los sensores de la vivienda y enviar por el protocolo MQTT (Message Queing Telemetry Transport) a la capa niebla.
- La capa niebla está compuesta por el conjunto de Raspberries Pi o dispositivos móviles (uno por cada vivienda)

Estre trabajo ha sido parcialmente soportado por el Gobierno de España por el proyecto RTI2018-098979-A-I00 MCIN/ AEI/10.13039/501100011033/, FEDER "Una manera de hacer Europa" y la Universidad de Jaén bajo la Acción 1 con referencia EI_TIC1_2021.

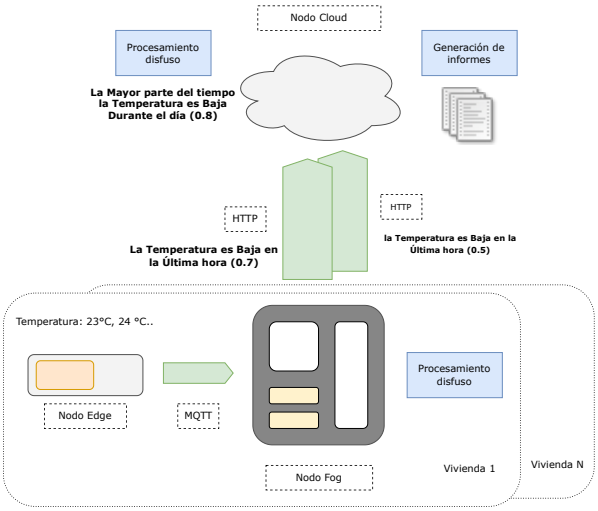


Figura 1. Arquitectura del sistema.

y computarán las protoformas lingüísticas primarias, las cuales enviarán a la capa nube por HTTP.

- La capa nube está compuesta por un servidor que realizará la persistencia de las protoformas primarias y computará protoformas lingüísticas secundarias, las cuales, les serán presentadas a los expertos mediante informes para su posterior análisis de pobreza energética.

Respecto a la comunicación, los datos generados por las Green Boxes son enviados mediante MQTT a cada nodo de la capa niebla de cada vivienda. El flujo de datos está formado por una serie de duplas (vt, t) dónde vt es el valor de temperatura y t la marca temporal. La frecuencia de muestreo de la temperatura puede ser definida con un parámetro fijo, como, por ejemplo, 5 minutos.

Las protoformas lingüísticas primarias que son computadas en el nodo niebla de cada vivienda son enviadas a la capa nube mediante HTTP. Dicho flujo de protoformas primarias está formado por las duplas $Protoform_0 = (Pt, t)$, donde Pt es el valor de la protoforma primaria y t su marca temporal. La frecuencia de envío de la protoforma primaria puede ser definida con un parámetro fijo, como, por ejemplo, 15 minutos.

III. Computación de protoformas primarias y secundarias

En está sección se presenta una propuesta inicial de protoformas lingüísticas difusas en la arquitectura IoT en cuestión de pobreza energética.

La Organización Mundial de la Salud y el Reglamento de Instalaciones Térmicas en los Edificios, recomiendan que las personas mayores tengan una temperatura media en su casa anual de entre 20 y 24 grados, con una humedad relativa media de entre el 40 y el 60 % [3].

Por ello, se establece como relevante incluir un término lingüístico de temperatura con sintaxis $baja$ que es representado por la función trapezoidal con valores $\mu T(vt) = (0, 0, 18, 20)$ expresado en grados Celsius.

Una ventana temporal difusa se describe según la distancia del tiempo actual t_0 a una marca de tiempo dada t como $\Delta t = (|t0 - t|)$.

Como ventanas temporales difusas se definen dos ventanas, una de ellas será utilizada en la protoforma primaria, computada en el nodo niebla, y la otra de ellas será computada en el servidor (capa nube) en la protoforma secundaria. En la protoforma primaria se define el término $H = EnLaUltimaHora$ que es representado por la función trapezoidal con valores $\mu H(\Delta t) = (60, 45, 0, 0)$ expresado en minutos. En la protoforma secundaria se define el término $D = DuranteElDia$ que es representado por la función trapezoidal con valores $\mu D(\Delta t) = (24, 20, 0, 0)$ expresado en minutos.

El cuantificador será utilizado en la protoforma secundaria computada en el servidor, se propone el cuantificador $Q = MayorParteDelTiempo$ que es representado por la función trapezoidal con valores $\mu Q(\Delta t) = (80, 90, 100, 100)$ expresado en porcentaje.

Una vez se han definido los términos lingüísticos y los cuantificadores, se procederá a definir la protoforma primaria, computada en los nodos niebla, y la protoforma secundaria, computada en el servidor.

La protoforma primaria se define como *La Temperatura es Baja en la Última hora*, la cual se calcula cada cierta frecuencia, siendo la propuesta en 15 minutos mediante una ventana deslizante. La protoforma primaria es representada mediante la Ecuación (1)

$$Protoforma_0(vt, t) = (Pt(vt, t), t)$$
$$= V \cup T(vt, t) = \bigcup \mu FT(vt) \cap \mu H(\Delta t) \in [0, 1] \quad (1)$$

La protoforma secundaria se define como *La Mayor parte del tiempo la Temperatura es Baja Durante el día*, la cual se calcula cada cierta frecuencia, siendo la propuesta en 24 horas mediante una ventana deslizante. La protoforma secundaria es representada en la Ecuación (2)

$$Proforoma_2 = QProtoforma_0(Pt, t)$$
$$= Q\ (V \cup T(Pt, t)) = Q\ (\bigcup\ Pt(vt, t) \cap \mu D(\Delta t)) \in [0, 1] \quad (2)$$

Esta protoforma secundaria le permite a los expertos evaluar de manera eficiente las condiciones energéticas de la vivienda.

IV. Conclusiones

En este resumen, hemos planteado una arquitectura IoT para la generación de indicadores de pobreza energética en lenguaje natural mediante protoformas lingüísticas difusas. En futuros trabajos, nos planteamos investigar con expertos energéticos los indicadores de interés clave para la generación de protoformas más relevantes.

Referencias

[1] Ministerio para la transición ecológica y el reto demográfico, "Estrategia Nacional contra la Pobreza Energética", 2022, unpublished.

[2] J. Kacprzyk and S. Zadrożny, "Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools", in Information Sciences, vol. 173, no. 4, pp. 281-304, June 2005.

[3] World Health Organization, et al, "Indoor environment: health aspects of air quality, thermal environment, light and noise", World Health Organization, 1990, unpublished.

UJAmI Location: A Fuzzy Indoor Location System for the Elderly

1º Antonio-Pedro Albín-Rodríguez
Consejería de Educación y Deporte
Junta de Andalucía
Jaén, España
ORCID: 0000-0003-3447-2260

2º Yolanda María De-La-Fuente-Robles
Departamento de Trabajo social
Universidad de Jaén
Jaén, España
ORCID: 0000-0002-2643-0100

3º José-Luis López-Ruiz
Departamento de Informática
Universidad de Jaén
Jaén, España
ORCID: 0000-0003-2583-8638

4ª Ángeles Verdejo-Espinosa
Departamento de Ingeniería Eléctrica
Universidad de Jaén
Jaén, España
ORCID: 0000-0002-7998-553X

5ª Macarena Espinilla Estévez
Departamento de Informática
Universidad de Jaén
Jaén, España
ORCID: 0000-0003-1118-7782

***Abstract*—En este trabajo se resume la propuesta titulada "UJAmI Location: A Fuzzy Indoor Location System for the Elderly" que presenta una metodología de localización difusa en interiores basada en el uso de términos lingüísticos difusos y ventanas temporales difusas para gestionar las fluctuaciones de las balizas BLE y la señal recibida por un dispositivo vestible o móvil. La metodología propuesta es validada a través de un caso de estudio, obteniendo una precisión mayor del 90% y superando en 10 puntos al procesamiento no difuso. Finalmente, es presentado el sistema de localización de interiores denominado UJAmI Location que integra dicha metodología.**

***Index Terms*—BLE balizas, Lógica difusa, RSSI, Sistemas de posicionamiento en interiores**

I. Introducción

El seguimiento de las áreas que visita una persona en su hogar o en una residencia de personas mayores puede arrojar información clave para comprender y monitorizar comportamientos relacionados con la salud como son las rutinas de sueño, la velocidad de la marcha, un comportamiento sedentario, la realización de actividades de la vida diaria o la detección de caídas.

Los sistemas de localización en interiores son una herramienta clave para la monitorización de dichos comportamientos. Se han propuesto múltiples enfoques o modelos que proporcionan metodologías basadas en diferentes tecnologías para localizar personas dentro de espacios cerrados.

En el contexto de localización, es muy común el uso de la banda ultra ancha, Bluetooth de bajo consumo (*Bluetooth Low Energy*, BLE) combinado con un dispositivo que disponga de este tipo de conexión (dispositivos móviles o dispositivos vestibles) o incluso la identificación por radiofrecuencia.

Este trabajo ha sido parcialmente soportado por REMIND project Marie Sklodowska-Curie EU Framework for Research and Innovation Horizon 516 2020, under grant agreement no. 734355, el Gobierno de España por el proyecto RTI2018-098979-A-I00 MCIN/ AEI/10.13039/501100011033/, FEDER "Una manera de hacer Europa" y la Universidad de Jaén bajo la Acción 1 con referencia EI_TIC1_2021.

Entre las tecnologías más utilizadas destaca el uso de transmisores o balizas BLE para la localización en interiores debido a que estos dispositivos son ampliamente utilizados por su excelente rendimiento en términos de batería, pequeño tamaño, peso ligero, alta precisión para la localización y, finalmente, por ser fácilmente desplegables a un bajo coste.

Uno de los principales retos de los sistemas de localización en interiores es lidiar con la incertidumbre inherente a las tecnologías aplicadas en estos sistemas debido a problemas de calibración, pérdida de datos, obstáculos en interiores o limitaciones en el consumo de la batería. Además, existe una brecha importante entre el número de propuestas teóricas en la literatura sobre este tipo de sistemas y las que se desarrollan en sistemas reales para aplicaciones de la vida real.

En este contexto, la propuesta titulada "UJAmI Location: A Fuzzy Indoor Location System for the Elderly" [1] presenta un sistema inteligente de localización en interiores para personas mayores que supera las limitaciones de las actuales soluciones con el uso de dispositivos vestibles con sensores BLE y balizas BLE distribuidos en el ambiente, recogiendo información de proximidad, la cual es procesada de manera inteligente a través de descripciones lingüísticas difusas.

Para ello, este resumen se estructura del siguiente modo: en la Sección II se presenta la metodología de localización difusa en interiores con reseñas sobre la evaluación realizada. La Sección III resume los resultados de evaluación de la metodología. La Sección IV presenta el sistema de localización de interiores que integra dicha metodología. Finalmente, se incluyen las conclusiones en la Sección V.

II. Metodología de localización difusa

La metodología propone, por un lado, un marco teórico y, por otro lado, el procesamiento inteligente para obtener la localización en interiores a través de lógica difusa.

En el marco teórico se define el espacio cerrado y las áreas que se monitorizan, el conjunto de balizas BLE que son ubicadas en cada área, el dispositivo vestible/móvil que lleva cada persona dentro del ambiente y, finalmente, los términos lingüísticos y las ventanas temporales donde se procesará inteligentemente la localización.

Las balizas BLE emiten una señal de fuerza, la cual es leída por el dispositivo que lleva la persona. En concreto, el dispositivo lee el indicador de fuerza de la señal recibida (*Received Signal Strength Indicator*, RSSI) de cada una de las balizas BLE que hay en el ambiente.

El procesamiento inteligente es llevado a cabo a través de operadores de agregación difusos que agregan los flujos de datos de RSSI de cada una de las balizas utilizando ventanas temporales y discriminando por áreas donde se sitúan cada una de las balizas.

III. Evaluación de la metodología

La metodología se ha evaluado a través de un conjunto de datos generados en el apartamento de inteligencia ambiental de la Universidad de Jaén. En concreto, se ha utilizado el conjunto de datos de la Copa UCAmI para obtener patrones de localización. El conjunto de datos fue generado por una persona durante un periodo de 10 días, obteniendo datos de cuatro fuentes heterogéneas, entre ellas, la información de RSSI entre un dispositivo móvil y 15 balizas BLE colocadas en varios objetos del laboratorio inteligente. Para el procesamiento inteligente se ha definido el concepto de proximidad a través de una función difusa con valores de RSSI entre -95 dBm y -85 dBm (Figura 1) y una ventana temporal con valores entre 3 y 5 segundos (Figura 2).

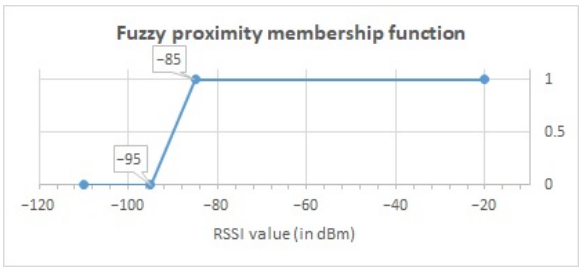


Fig. 1. Representación del término lingüístico proximidad

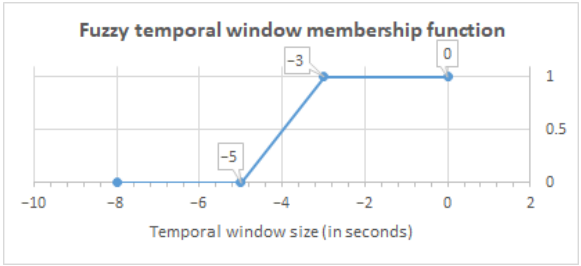


Fig. 2. Representación difusa de la ventana temporal

Para realizar una comparación entre la metodología que utiliza la lógica difusa y la que no la utiliza, se han comparado las fluctuaciones entre áreas, obteniendo la metodología propuesta una precisión del 91,63%, aproximadamente 10 puntos más que la metodología sin utilizar la lógica difusa. Los resultados obtenidos se ilustran en la Tabla I

TABLE I
TABLE TYPE STYLES

Time of Day	Fuzzy Accuracy	Non-Fuzzy Accuracy
Morning	88.13%	75.80%
Afternoon	92.11%	79.30%
Evening	93.36%	92.48%
Full-day	91.63%	81.12%

IV. Sistema inteligente de localización en interiores

La metodología es integrada en un sistema inteligente de localización en interiores, basado en la metodología propuesta, se denominada "UJAmI Location". Dicho sistema ha consistido en una aplicación desarrollada para dispositivos vestibles/móviles con el sistema operativo Android que reconoce balizas BLE dentro de un espacio delimitado y envía la información al servidor, así como un sistema web que procesa la información y gestiona los datos de localización, tanto en tiempo real como a lo largo del tiempo.

El sistema inteligente desarrolla la funcionalidad completa que permite definir todos los elementos necesarios como los datos básicos de identificación (dirección, información de contacto, mapa de localización, etc.), así como las diferentes zonas o áreas en las que se divide el ambiente, la ubicación de las balizas y las personas mayores dentro del espacio con sus dispositivos asignados.

V. Conclusiones

Los datos RSSI generados por dispositivos con BLE a través de balizas pueden ser procesados inteligentemente con computación suave para obtener la localización en interiores. Así, tanto los términos difusos de proximidad con valores de RSSI entre -95 dBm y -85 dBm, como los términos difusos temporales con valores entre 3 y 5 segundos, son adecuados para obtener la localización en interiores.

El procesamiento difuso obtiene una precisión mayor del 90% en el caso de estudio llevado a cabo, superando en 10 puntos al procesamiento no difuso. El sistema inteligente proporciona información útil sobre dónde han estado o están los habitantes en tiempo real, cuánto tiempo llevan allí y las zonas más frecuentadas del espacio.

References

[1] Albín-Rodríguez, A.P., Fuente-Robles, Y.M., López-Ruiz, J.L., Verdejo-Espinosa, M.A. and Espinilla, M. (2021). UJAmI Location: A Fuzzy Indoor Location System for the Elderly. International Journal of Environmental Research and Public Health, 18(16), 8326.

Una nueva aproximación a la descripción lingüística de series temporales y su aplicación a la inquietud en el descanso

1st Carmen Martinez-Cruz
Departamento de Informática
Universidad de Jaén
Jaén, España
cmcruz@ujaen.es

2nd Antonio J. Rueda
Departamento de Informática
Universidad de Jaén
Jaén, España
ajrueda@ujaen.es

3rd Alicia Montoro-Lendínez
Departamento de Informática
Universidad de Jaén
Jaén, España
amlendin@ujaen.es

4th Mihail Popescu
Departament of Health Management and Informatics
University of Missouri
Columbia, USA
popescum@health.missouri.edu

5th James M. Keller
Department of Electrical Engineering and Computer Science
University of Missouri
Columbia, USA
kellerj@missouri.edu

***Abstract*—En los últimos años se ha visto incrementada la necesidad de generación de descripciones y resúmenes en lenguaje natural de series temporales que describen cualquier medida, como la frecuencia cardíaca o la temperatura. Entre otros motivos, se ha evidenciado que las representaciones en el lenguaje natural son más fáciles de interpretar que las representaciones gráficas de datos, en casos que el usuario final carezca de la formación necesaria para la correcta interpretación de estos datos. Este resumen presenta un método de generación de descripciones lingüísticas de series temporales mediante el uso de lógica difusa que fue presentado en [2]. En la propuesta se generan resúmenes de calidad y se atiende a las características más relevantes de las series temporales para un usuario en un contexto concreto. Además, se revisa un caso de estudio real con resultados prometedores para la generación de descripciones lingüísticas de datos sobre la inquietud en el descanso, tomado a partir de los datos generados en la residencia de personas mayores TigerPlace en Columbia, Missouri (USA).**

***Index Terms*—Descripciones lingüísticas difusas, lógica difusa, series temporales.**

I. Introducción

El proceso de generación de descripciones lingüísticas de series temporales (GDLST) consiste en la descripción de datos brutos que incluyen la componente del tiempo en lenguaje natural mediante la construcción de resúmenes personalizados en los que se destaca la información más relevante según el contexto. Este trabajo resume la propuesta presentada en [2] que desarrolla un nuevo método para describir series temporales (ST) con lenguaje natural mediante un análisis de sus características más relevantes usando lógica difusa. Para ello, aborda el estudio de la síntesis lingüística de la evolución del movimiento en la cama (inquietud o desasosiego) de residentes del centro de personas mayores TigerPlace en Columbia, Missouri (USA). Los datos han sido generados por sensores hidráulicos ubicados en los colchones de las camas como parte del proyecto Aging in Place que aborda el reto de ayudar a las personas mayores a envejecer en el entorno escogido por ellas mismas y procesados en el Center for Eldercare and Rehabilitation Technology (ElderTech) de la Universidad de Missouri. Este tipo de resúmenes es de utilidad para dos tipos de usuarios en este contexto: a) residentes y familiares, con textos cortos y nada técnicos, y b) enfermeras, donde se puede usar un lenguaje más técnico. En las secciones II y III se incluye un breve estado del arte y la propuesta presentada, respectivamente. En la Sección IV se exponen las conclusiones a este artículo.

II. Estado del arte

Un proceso de resumen lingüístico de datos mediante sentencias cuantificadas difusas fue incialmente propuesto por Yager [1], y ha sido ampliamente usado en la GDLST [3]. Así, uno de los principales objetivos de GDLST es proporcionar información a personas que con escasos conocimientos sobre un área en concreto o personas más propensas a entender el lenguaje escrito que las representaciones visuales. Un área de aplicación de descripciones lingüísticas será la sanitaria, como el caso que ocupa este artículo [2].

III. Sistema GDLST para la inquietud en el descanso

Las ST analizadas en este artículo para resumir la inquietud en el descanso se han obtenido a partir de cuatro sensores hidráulicos ubicados bajo el colchón de cada uno de los residentes a monitorizar. Los cambios de presión son procesados puesto que son un parámetro que puede indicar la falta de calidad en el sueño o el desasosiego de un residente. El proceso de generación de resúmenes lingüísticos de las series

temporales se lleva a cabo siguiendo la arquitectura general descrita en [3] que resumimos a continuación.

Preprocesamiento de la serie temporal. Se elimina ruido y fluctuaciones de alta frecuencia. Para ello, se realiza el muestreo reducido de recogida de datos, el etiquetado fuera de la cama y un suavizado de los datos para eliminar ruido y enfatizar la tendencia a medio plazo de la inquietud en la serie temporal. Como muestra, en la validación de resultados de esta propuesta los datos brutos de origen se redujeron de 1834 a 508. Tras este procesado, la serie temporal se segmenta utilizando el algoritmo IEPF (Iterative end-point fit), en este caso, con un umbral de 0.25 obteniendo como resultado 14 segmentos.

Descripción de la ST. Las ST y concretamente los segmentos obtenidos, se describen mediante las siguientes características usando lógica difusa: los valores de inquietud en cada punto (que son posteriormente cuantificados), la variabilidad, la tendencia local y los momentos en el tiempo en los que se localiza cada medida. Utilizando estas características, se definen objetos de alto nivel para resumir la ST en este contexto: secciones, picos, conjuntos de picos y espacios sin datos. Todas estas definiciones y las funciones de pertenencia usadas han sido revisadas y aprobadas por un grupo de expertos. Finalmente, se han definido un conjunto de protoformas en función de estos datos.

Generación del resumen lingüistico. Se obtiene un primer resumen lingüístico de cada ST, mediante la instanciación del conjunto de protoformas. En la Tabla I se muestran las protoformas para el modelo de conocimiento de la inquietud en el descanso. A partir de estas instancias y una plantilla sintáctica cercana al estilo lingüístico del perfil del usuario al que va destinado se genera el resumen. Estas plantillas se fijan de acuerdo al modelo de garantía de calidad para el contexto de la inquietud en la cama. Además, también se definen las reglas de selección y transformación de acuerdo a las pautas dadas por el equipo de enfermería de TigerPlace y el equipo de Eldertech para evitar información no relevante o frases inconexas que generen textos no fluidos.

El resultado del procesamiento de las series temporales siguiente el método presentado en [2] es un conjunto de sentencias sencillas y significativas que el equipo de enfermería TigerPlace y de ElderTech, valoraron positivamente. Este resultado sintetiza la información esencial y más relevante de la inquietud en la cama de un residente por la noche, además de ser una herramienta útil para los profesionales. Un ejemplo de resumen obtenido es el siguiente:

Sleeping time was between 22:00 and 06:30. The resident got out of bed 3 times at around 23:45 (00h14m), 00:15 (00h12m) and 03:30 (00h11m). At the beginning of the night, bed restlessness (BR) has a sharply decreasing trend. After that, there is a peak with high BR at around 22:30. Then, BR has an increasing trend with many fluctuations. In the middle of the night, BR has a sharply decreasing trend. After that, BR is very often medium for around 3h00m and steady. At the end of the night, there is a set of peaks with high BR for around 01h30m. In the last part, BR remains steady.

TABLE I
PROTOFORMAS EN EL MODELO DE CONOCIMIENTO DE LA INQUIETUD EN LA CAMA.

N.	Protoformas	Resúmenes
Protoforma por segmento/sección		
1	*At/In the **[part]** of the night.*	**part**: *beginning, end, middle.*
2	***[quantif]** of bed restlessness values in this period are **[value]**(**duration**).*	**quantif:** *almost all, the majority, many.* **value**:*very high, high, medium, low, very low.* **duration**: duration of the event.
3.1	*Bed restlessness is **[local trend]**.*	**local trend**: *sharply increasing, increasing, decreasing, sharply decreasing, steady.*
3.2	*There is a **[feature]** with **[value]** bed restlessness (**[hour]** / **[duration]**).*	**feature**: *set of peaks/peak* **value**: *very high, high, medium, low, very low.* **hour**: hour of the peak. **duration**: duration of the set of peaks.
4	*Volatility is **[var]**.*	**var**: *high, medium, low.*
5	*Getting out of bed at **[hour]** for the **[num]** time (**[duration]**).*	**num**: 1st, 2nd, 3rd, etc. **duration**: duration of the period.
Protoformas globales		
6	*Sleeping time was between **[start hour]** and **[end hour]**.*	**start hour**: starting hour of the TS. **end hour**: end hour of the TS.

IV. CONCLUSIÓN

En este resumen se presenta una visión novedosa que imita la arquitectura general GDLST descrita en [3] con la finalidad de realizar resúmenes de calidad de series temporales para satisfacer el perfil de los receptores del resumen: enfermeros y familiares. Para imitar el comportamiento humano y obtener la estructura básica de la ST se utiliza el algoritmo IEPF. También, se define el modelo de representación del conocimiento con un novedoso mecanismo para describir una ST lingüísticamente procesando las características relevantes en palabras. La propuesta fue exitosa gracias al equipo de ElderTech y las enfermeras del TigerPlace que ayudaron a generar, analizar y dar conocimiento experto al modelado del sistema.

Por último, en relación al trabajo futuro una de las metas es mejorar el modelo de representación lingüística con el objetivo de generar mejores descripciones lingüísticas

AGRADECIMIENTOS

Trabajo financiado por los proyectos EI_TIC1_2021 de la Acción 1 de la Universidad de Jaén y B-TIC-744-UGR20 por FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades.

.

REFERENCES

[1] R. R. Yager, "On linguistic summaries of data", Proc. Knowl. Discov. Databases, pp. 347-366, 1991.
[2] C. Martinez-Cruz, A. J. Rueda, M. Popescu and J. M. Keller, "New Linguistic Description Approach for Time Series and Its Application to Bed Restlessness Monitoring for Eldercare," in IEEE Transactions on Fuzzy Systems, vol. 30, no. 4, pp. 1048-1059, April 2022.
[3] N. Marín and D. Sánchez, "On generating linguistic descriptions of time series", Fuzzy Sets Syst., vol. 285, pp. 6, 2016.

Minería incremental de datos para extracción de reglas de asociación difusas en datos ambientales

1st Ignacio Blanco
Dep. CC. de la Computación e I. A.
Univ. de Granada
Granada, España
iblanco@ugr.es

2nd Luisa M. González-González
Dep. CC. de la Computación
Univ. Central "Marta Abreu" de las Villas
Santa Clara, Cuba
luisagon@uclv.edu.cu

3rd Carmen Martínez-Cruz
Dep. Informática
Univ. de Jaén
Jaén, España
cmcruz@ujaen.es

4th Alain Pérez-Alonso
Dep. Electrónica e Informática
Univ. Técnica Federico Santa María
Concepción, Chile
alain.perez@usm.cl

5th José M. Serrano
Dep. Informática
Univ. de Jaén
Jaén, España
jschica@ujaen.es

***Resumen*—En problemas del mundo real, las bases de datos no son estáticas, y la información se actualiza continuamente. Un ejemplo son los datos procedentes de entornos inteligentes. Precisamente, el análisis de dicha información mediante técnicas tradicionales de minería de datos puede verse dificultado por esta continua variación. En este trabajo se resume la propuesta de dos algoritmos de minería de datos incremental, siguiendo un enfoque inmediato y diferido, respectivamente, para la extracción de reglas de asociación difusas.**

***Palabras clave*—minería incremental de datos, reglas de asociación difusas, sensores, entornos inteligentes**

I. Introducción

La minería de datos continúa vigente como un conjunto de técnicas útiles para el análisis de grandes volúmenes de información. En particular, la extracción de reglas de asociación, en sus múltiples vertientes, sigue siendo objeto de numerosos estudios. En un campo como el de la inteligencia ambiental, donde es habitual trabajar con flujos continuos de datos, procedentes de distintas fuentes, tales como sensores o contenidos en la nube, resulta especialmente interesante contar con herramientas que permitan hallar relaciones o patrones entre las diferentes medidas consideradas.

Sin embargo, la naturaleza continua de dichos flujos de información dificulta la aplicación directa de técnicas tradicionales de minería de datos. Cambios en la información pueden provocar que las reglas obtenidas queden obsoletas rápidamente. En general, si los datos de partida se alteran, es necesario volver a generar e interpretar los resultados desde el principio.

Para afrontar este problema, se han propuesto diferentes soluciones a lo largo de los últimos años, dentro del área conocida como minería de datos incremental. En el presente trabajo, se resumen las aportaciones realizadas en [4], donde se proponen dos algoritmos de minería de datos incremental, mediante los que mantener la validez de un conjunto de reglas de asociación previamente extraido.

Nuestra propuesta no solo se limita a los dos algoritmos. Además, se propone una estructura de almacenamiento de las características de dichas reglas para optimizar el proceso de mantener su validez a lo largo del tiempo.

Este trabajo ha sido parcialmente financiado por el proyecto B-TIC-744-UGR20 *ADIM: Accesibilidad de Datos para Investigación Médica*. Convocatoria FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades.

II. Reglas de asociación

Dado un conjunto de transacciones T, una regla de asociación [1] se puede ver como una implicación de la forma $X \Rightarrow Y$, donde $X, Y \subset It$, siendo It=$\{it_1, it_2, \ldots, it_m\}$ un conjunto de ítems tal que $It \subseteq T$, donde se cumple que $X \neq \emptyset$, $Y \neq \emptyset$ y $X \cap Y = \emptyset$. A X e Y se les denomina, respectivamente, antecedente y consecuente de la regla.

Aunque a lo largo de los años se han propuesto diferentes medidas de la relevancia y el interés de una regla de asociación, las medidas más extendidas son su soporte, definido como la proporción de transacciones sobre el total de T que contienen a $X \cup Y$, y su confianza, obtenida como la fracción de transacciones que contienen a X Y sobre las transacciones que contienen a X. En general, todas las medidas de interés de una regla de asociación comparten una serie de características comunes, relacionadas con el antecedente o consecuente de la regla, excepciones a la misma, etc. Dichas propiedades permiten combinarse para obtener otras métricas en función de ellas.

Tradicionalmente, las reglas de asociación se aplican sobre conjuntos de transacciones. Posteriores estudios han extendido la definición para operar sobre datos cuantitativos o con una alta granularidad. En [2], se presenta una propuesta para extraer reglas de asociación difusas (FAR), donde se consideran transacciones difusas como subconjuntos difusos de ítems, y que generaliza la definición clásica de regla de asociación.

III. Mantenimiento del conocimiento extraído

En problemas del mundo real, los datos no suelen ser estáticos, y varían como resultado de operar sobre la base de datos. En el ámbito de las bases de datos activas, dichas operaciones se representan como eventos, primitivos (PSE, *primitive structural event*) o compuestos (como una combinación de varios PSE). Es habitual considerar tres tipos de PSE: inserción, borrado y actualización de una tupla. Analizando los PSE que afectan a una base de datos, es posible registrar los diferentes estados por los que va pasando la misma.

En [4], se detalla cómo las FAR obtenidas se almacenan como un elemento más de la base de datos. Cada regla cuenta con una serie de características (FMP, *fuzzy measure part*) que, combinadas, nos permiten calcular métricas de interés. Un ejemplo es la confianza de una regla, obtenida como la fracción entre la proporción de casos donde aparecen antecedente y consecuente, y la proporción de casos en los que aparece el antecedente de la regla. Estas características nos pemiten obtener diferentes métricas de forma eficiente. Por ejemplo, en [3] se demuestra que con solo 5 características distintas es posible obtener hasta 20 métricas de interés para una regla de asociación.

De acuerdo con lo anterior, en [4] se proponen dos algoritmos para el mantenimiento incremental de FAR en una base de datos. Ambos tienen la ventaja de que permiten mantener la validez de las FAR sin tener que volver a generarlas de cero, y sin necesidad de volver a revisar la base de datos al completo.

El primer algoritmo sigue un enfoque inmediato. Cada vez que se registra un PSE, el sistema considera si es preciso actualizar las FMP para cada regla almacenada en la base de datos. Si el número de eventos es muy alto (es decir, si la información se modifica con mucha frecuencia), el rendimiento del algoritmo puede verse perjudicado. Es por ello que el segundo algoritmo propone un enfoque diferido, donde se registran igualmente las ocurrencias de distintos PSE, pero en una estructura secundaria, y la actualización de las FMP de las reglas se lleva a cabo, en un solo paso, en una etapa posterior. Por ejemplo, si se inserta una nueva fila y más adelante se elimina, la operación en conjunto no tendría efecto sobre las reglas. El enfoque diferido permite obviar aquellos PSE que no afectan a la base de reglas. En el peor de los casos, habrá que considerar un número de operaciones igual al caso del enfoque inmediato. La figura 1 muestra un esquema del funcionamiento de ambos enfoques.

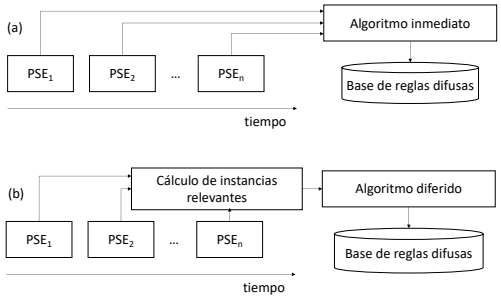


Figura 1. Mantenimiento incremental inmediato (a), y mantenimiento incremental diferido (b).

IV. Experimentos y conclusiones

En [4], se compara el rendimiento de los algoritmos propuestos con el de algoritmos tradicionales como fuzzy Apriori y fuzzy FP-growth, disponibles en KEEL [5]. Los experimentos se ejecutaron a partir de 3 datasets del UCI Machine Learning Repository: *diabetes*, *color texture* y *color moment*. Se extrajo un total de 7 FAR y se midió el tiempo transcurrido tras realizar 5000 operaciones sobre los datos, con el correspondiente mantenimiento de las reglas. En la figura 2 puede comprobarse cómo el tiempo de ejecución de nuestras propuestas es notablemente inferior al de los algoritmos tradicionales. Estos resultados ponen de manifiesto las ventajas de nuestra propuesta de minería incremental en problemas del mundo real, en particular, en su aplicación al análisis de bases de datos procedentes de sensores, aspecto que será estudiado en profundidad en un próximo trabajo.

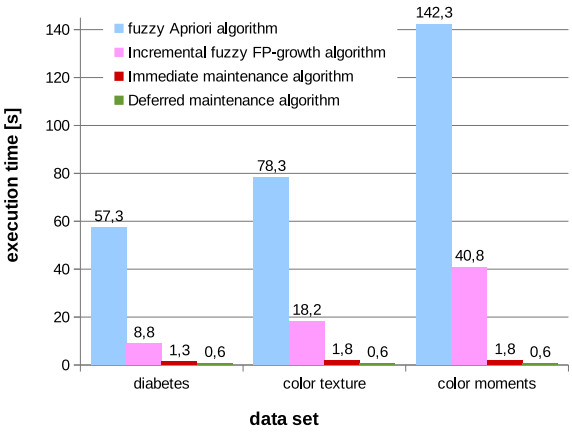


Figura 2. Comparativa entre algoritmos para el mantenimiento de FAR en diferentes datasets.

Referencias

[1] R. Agrawal, T. Imieliński, and A. Swami. "Mining association rules between sets of items in large databases". SIGMOD Rec. 22, 2, 207—216, (1993). https://doi.org/10.1145/170036.170072

[2] M. Delgado, N. Marin, D. Sanchez and M.A. Vila, "Fuzzy association rules: general model and applications", in IEEE Transactions on Fuzzy Systems, vol. 11, 2, 214–225, (2003). https://doi.org/10.1109/TFUZZ.2003.809896

[3] P. Lenca, P. Meyer, B. Vaillant, S. Lallich, "On selecting interestingness measures for association rules: User oriented description and multiple criteria decision aid". Eur. J. Operational Research, vol 184, 2, 610–626, (2008). https://doi.org/10.1016/j.ejor.2006.10.059.

[4] A. Pérez-Alonso, I.J. Blanco, J.M. Serrano et al. "Incremental maintenance of discovered fuzzy association rules". Fuzzy Optim Decis Making 20, 429-–449 (2021). https://doi.org/10.1007/s10700-021-09350-3

[5] I. Triguero, S. González, J.M. Moyano, S. García, J. Alcalá-Fdez, J. Luengo, A. Fernández, M.J. del Jesus, L. Sánchez, F. Herrera. "KEEL 3.0: An open source software for multi-stage analysis in data mining". Int. Journal of Computational Intelligence Systems 10, 1238–1249 (2017). https://doi.org/10.2991/ijcis.10.1.82

A comparative analysis of similarity measures in noisy images

Agustina Bouchet
Dept. of Statistics and O.R.
University of Oviedo
Spain
bouchetagustina@uniovi.es

Susana Díaz-Vázquez
Dept. of Statistics and O.R.
University of Oviedo
Spain
diazsusana@uniovi.es

María Fernández
Dept. of Statistics and O.R.
University of Oviedo
Spain
UO263618@uniovi.es

Susana Montes
Dept. of Statistics and O.R.
University of Oviedo
Spain
montes@uniovi.es

***Abstract*—The analysis of the influence of the noise in an image by means of the use of similarity measures between interval-valued fuzzy sets is proposed in this communication. Thus, some families of similarity measures are compared in accordance to their behaviour for the detection of the noise. The different sensitivity of these measures is clarified by means of the study of twenty five images with ten different degrees of noise, by repeating the analysis one hundred times to generate new instance of the random noise each time.**

***Index Terms*—Similarity measure, interval-valued fuzzy set, noisy image**

I. Introduction

A similarity measure is a function that determines the similarity degree between two objects. This concept has been studied by several authors and it has been applied in different areas, such as image processing, pattern recognition, machine learning, decision making, among other.

In the case of intuitionistic fuzzy sets [1], the conception of similarity measure has been developed over the years [2]. Chen [3], [4], Hong and Kim [5], Fan and Zhangian [6], Dengfeng and Chuntian [7], Mitchell [8] and Liang and Shi [9] have proposed definitions of similarity measures. Although they have some relevant semantic differences, intuitionistic fuzzy sets are proven to be formally equivalent to interval-valued fuzzy sets (IVFS) [10], [11], as proven in [12], [13]. Thus, any similarity for intuitionistic fuzzy sets can be easily translated to a similarity for IVFS. This is important since the use of IVFS has proven to be an optimal tool for modeling images since they allow to characterize the uncertainty present in their structures [14].

Similarity measures between IVFS can be used as a very efficient tool in noisy images. In general, images are immersed in noise due to the way they are acquired and it is difficult to remove from it. Therefore, it is important to be able to quantify the noise present in an image to minimize its effects.

In this work, an analysis of the performance of different similarity measures against images immersed in noise is proposed. This is done comparing the behavior of the different families of similarity measures.

This research has been partially supported by Spanish MINECO projects TIN2017-87600-P (Agustina Bouchet) and PGC2018-098623-B-I00 (Susana Díaz-Vázquez and Susana Montes).

II. Preliminaries

Let X denote the universe of discourse. An IVFS of X is a mapping $A : X \to L([0,1])$ such that $A(x) = [\underline{A}(x), \overline{A}(x)]$, where $L([0,1])$ denotes the family of closed intervals included in the unit interval $[0,1]$. The collection of all the interval-valued fuzzy sets in X will be denoted by $IVFS(X)$.

A function $S : IVFS(X) \times IVFS(X) \to [0,1]$ is a similarity measure if it satisfies the following properties, for any $A, B, C \in IVFS(X)$:

1) $S(A,A) = 1$
2) $S(A,B) = S(B,A)$
3) If $A \subseteq B \subseteq C$ then $S(A,C) \leq S(A,B)$ and $S(A,C) \leq S(B,C)$

III. Similarity measures for IVFS

We consider the universal set as $X = \{x_1, x_2, \ldots, x_n\}$.

Proposition 3.1: For any decreasing map $g : [0,1] \to [0,1]$ with $g(0) = 1$, any map $f : [-1,1] \times [-1,1] \to [0,1]$ with $f(0,0) = 0$, $f(x,y) = f(-x,-y)$ and f increasing in both components in $(0,1]$ and any family of weights $\{w_i\}_{i=1}^n$ with $w_i \geq 0, \forall i$ and $\sum_{i=1}^n w_i = 1$, the function:

$$S_{fg}(A,B) = g\left(\sum_{i=1}^{n} w_i f\left(\overline{A}(x_i) - \overline{B}(x_i), \underline{A}(x_i) - \underline{B}(x_i)\right)\right)$$

is a similarity measure in $IVFS(X)$.

As a particular case of the previous family, we can consider the following cases whenever $g(x) = 1 - \sqrt[p]{x}$, with $p \geq 1$:

- Let $f(x,y) = \left(\frac{|x+y|}{2}\right)^p$, the similarity family obtained is denoted by $S^p_{D\omega}$ and it is the same as the one obtained directly from the similarity between intuitionistic fuzzy sets given by Dengfeng and Chuntian [7]. In the particular case that all the weights are equal, the previous measure is denoted by S^p_D and it is also obtained in [7]. For the case $p = 1$, it is denoted by S_C and it is analogous to the measure introduced previously by Chen [3], [4].
- Let $f(x,y) = \left(\frac{|x|+|y|}{2}\right)^p$, the obtained measure is denoted by S^p_e and it can also be deduced from Liang and Shi [9].
  For the case $p = 1$, we arrive to a measure S_H analogous to the one proposed by Hong and Kim [5].

- The similarity measure defined by $\frac{1}{2}(S_C + S_H)$, which is denoted by S_L coincides with the similarity measure obtained from Fan and Zhangian [6].
- If the map is $f(x,y) = \left(\frac{|x+3y|+|3x+y|}{8}\right)^p$ and the weights are all equal, S_s^p is obtained which can also be deduced from Liang and Shi [9].
- Let all the weights be equal and:
  - $f_1(x,y) = 1/8\ (|x+3y| + |3x+y|)$
  - $f_2(x,y) = 1/2\ |x+y|$
  - $f_3(x,y) = 1/2\ |x-y|$

  Then $f = (\omega_1'\ f_1 + \omega_2'\ f_2 + \omega_3'\ f_3)^p$ with $0 \leq \omega_m' \leq 1$, $\sum_{m=1}^{3} \omega_m' = 1$ defined a similarity measure called S_h^p that coincides with one obtained by Liang and Shi [9].
- Let all the weights be equal and:
  - $f_1(x,y) = |y|^p$, $S_{f_1 g}$ is a similarity measure
  - $f_2(x,y) = |x|^p$, $S_{f_2 g}$ is a similarity measure

  Then, $S_M^p = \frac{1}{2}(S_{f_1 g} + S_{f_2 g})$ is a similarity measure which can be obtained from Mitchell [8].

IV. Images and noise analysis

A grayscale image is defined by a function $I : D_I \subseteq \mathbb{Z}^2 \to \{0, \ldots, 255\}$ where $I(x_{ij})$ represents the gray value of the pixel x_{ij}. In order to consider an image as an IVFS, we define $\underline{A}$ as the normalised value of x_{ij} and $\overline{A}$ is obtained using the Sugeno fuzzy generator [15].

The idea of the experimentation is based on generating noise over the images and, then, apply the similarity measures between the original image and the noisy image, in order to analyze the performance of each similarity measure introduced at the previous section. This will allow us to analyze whether there are more sensitive measures against noise and then, depending on the area in which it is going to be applied, take it into account when the measure has to be chosen.

To analyze the behavior of each similarity measure, the experimentation that is carried out can be summarised below:

- Images: twenty five images of the *Berkeley Segmentation Dataset* are used [16].
- Noise: ten different standard deviation values σ are used to generate the Gaussian noise, from 0.05 to 0.5.
- Parameters of the similarity measures: $p = 3$, $\omega_i = 1/n$ for $S_{D\omega}^p$ and $\omega_m' = 1/3$ for S_h^p.
- Iterations: the analysis was repeated 100 times generating new instances of the random noise each time.

V. Results

Firstly, from the results obtained, it can be seen that not only the similarity values are different in some measures but also the deviation varies as a result of the generated noise. In this way, S_C, S_H and S_L have the greatest deviation. In addition, the atypical points change according to the similarity measures.

Another important characteristic for the similarity measures is their performance. Some similarity measures show similar performance against noise. In this way, three families of similarity measures can be determined. The greatest values are obtained with S_C, S_H and S_L. The lowest values are obtained with S_h^p and S_M^p. Furthermore, the intermediate values are achieved with S_D^p, S_{Dw}^p, S_e^p and S_s^p. Also, the values of all the measures analyzed decrease when the noise level increases as it was expected.

VI. Conclusions

In this communication a general procedure to obtain similarity measures for IVFS is introduced and an analysis between different families of similarities is made, by studying their behaviour against noise images.

According to their performance, we can conclude that there are similarity measures more sensitive against the noise and depending on the application in which it will be used, this should be taking into account.

As a future work, we will analyse in detail the effect of different values of the parameter that defines some of the measure, i.e., the values of p and the weights, and we will make a comparison with alternative methods. We would also like to be able to consider different ways of describing the images by means of IVFS.

References

[1] K. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, pp. 87–96, 1986.

[2] M. Luo and J. Liang, "A novel similarity measure for interval-valued intuitionistic fuzzy sets and its applications," *Symmetry*, vol. 10, no. 10, 2018.

[3] S. Chen, "Measures of similarity between vague sets," *Fuzzy Sets and Systems*, vol. 74, no. 2, pp. 217–223, 1995.

[4] ——, "Similarity measures between vague sets and between elements," *IEEE Trans Syst Man Cybernet*, vol. 27, no. 1, pp. 153—158, 1997.

[5] D. H. Hong and C. Kim, "A note on similarity measures between vague sets and between elements," *Information Sciences*, vol. 115, no. 1, pp. 83–96, 1999.

[6] L. Fan and X. Zhangian, "Similarity measures between vague sets," *J Software*, vol. 12, no. 6, pp. 922—927, 2001.

[7] L. Dengfeng and C. Chuntian, "New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions," *Pattern Recognition Letters*, vol. 23, no. 1, pp. 221–225, 2002.

[8] H. Mitchell, "On the dengfeng–chuntian similarity measure and its application to pattern recognition," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3101–3104, 2003.

[9] Z. Liang and P. Shi, "Similarity measures on intuitionistic fuzzy sets," *Pattern Recognit. Lett.*, vol. 24, pp. 2687–2693, 2003.

[10] I. Grattan-Guinness, "Fuzzy membership mapped onto interval and many- valued quantities," *Zeitschrift für mathematische Logik und Grundladen der Mathematik*, vol. 22, pp. 149–160, 1976.

[11] R. Sambuc, "Fonctions ϕ-floues. application à l'aide au diagnostic en pathologie thyroidienne," Ph.D. dissertation, Université de Marseille, France, 1975.

[12] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 31, pp. 343–349, 1989.

[13] G. Deschrijver and E. Kerre, "On the relationship between some extensions of fuzzy set theory," *Fuzzy Sets and Systems*, vol. 133, pp. 227–235, 2003.

[14] E. Barrenechea, H. Bustince, B. De Baets, and C. Lopez-Molina, "Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 5, pp. 819–830, 2011.

[15] M. Sugeno, "Fuzzy measures and fuzzy integrals—a survey," in *Readings in Fuzzy Sets for Intelligent Systems*, D. Dubois, H. Prade, and R. R. Yager, Eds. Morgan Kaufmann, 1993, pp. 251–257.

[16] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, July 2001, pp. 416–423.

Técnicas de Soft Computing para Sistemas Recomendadores para la Colaboración Científica

Jared D.T. Guerrero-Sosa
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
JaredDavidtadeo.Guerrero@alu.uclm.es

Víctor Hugo Menéndez-Domínguez
Facultad de Matemáticas
Universidad Autónoma de Yucatán
Mérida, México
mdoming@correo.uady.mx

Francisco P. Romero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
FranciscoP.Romero@uclm.es

Jesus Serrano-Guerrero
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
Jesus.Serrano@uclm.es

Jose A. Olivas
Tecnologías y Sistemas de Información
Universidad de Castilla-La Mancha
Ciudad Real, España
JoseAngel.Olivas@uclm.es

María-Enriqueta Castellanos-Bolaños
Facultad de Matemáticas
Universidad Autónoma de Yucatán
Mérida, México
enriqueta.c@correo.uady.mx

***Resumen*—En este trabajo se presentan dos técnicas de lógica borrosa para identificar la afinidad entre dos grupos de investigación con el objetivo de establecer una colaboración científica. A partir del motivo de la cooperación, permiten la evaluación considerando los pesos asignados a los criterios a utilizar.**

***Index Terms*—relaciones borrosas, representación lingüística basada en dos tuplas, colaboración científica**

I. Introducción

Una actividad esencial en la investigación es la colaboración entre miembros de la comunidad científica, ya sea a nivel institucional, nacional e internacional. Los motivos de las colaboraciones son diversos, entre ellos se encuentran la mejora de los indicadores de desempeño para evaluaciones periódicas, la financiación de proyectos, el desarrollo del conocimiento multidisciplinar, así como el incremento de la productividad de una universidad o centro de investigación [1]. Sin embargo, la identificación de autores potenciales para una colaboración puede ser complicada dado que, dependiendo de los motivos de la colaboración se pretende detectar a los mejores candidatos que cumplan con características específicas. Además, la cantidad de investigadores con temáticas afines o de interés puede ser muy amplia. Para solucionar este problema se han propuesto sistemas recomendadores para la colaboración científica. En términos generales, un sistema recomendador consiste en un conjunto de técnicas que pretenden ayudar a un usuario concreto a descubrir elementos que puedan ser de su interés en un contexto a partir de una amplia gama de opciones partiendo de su perfil o preferencias.

Los aspectos relevantes de un posible colaborador se encuentran en los metadatos asociados con su producción científica dado que permiten medir la afinidad entre dos investigadores. Dicha producción se encuentra en las bases de datos bibliográficas, que son grandes gestores de la literatura, siendo algunas multidisciplinares (por ejemplo, Scopus y Web of Science) y otras de propósito específico (por ejemplo, PubMed). Sin embargo se debe considerar que para emitir las recomendaciones existe la imprecisión, la incertidumbre y la vaguedad en los datos, por lo que se propone definir un sistema recomendador para la colaboración científica por medio de técnicas de la lógica borrosa. Este trabajo tiene como propósito presentar dos técnicas basadas en la lógica borrosa para definir recomendaciones de colaboraciones, considerando que cada colaboración tiene un motivo asociado con criterios, los cuales pueden tener mayor o menor importancia dependiendo de la necesidad de la cooperación. El resto del trabajo se organiza como sigue: en la Sección II se presenta un breve estado del arte; en la Sección III, a partir de un ejemplo para la colaboración entre grupos, se muestran dos técnicas basadas en la lógica borrosa, y en la sección IV se presentan las conclusiones.

Este trabajo ha sido desarrollado gracias al apoyo del Consejo Nacional de Ciencia y Tecnologia (CONACYT, México) a través de la beca con número (CVU/Becario): 853088/809847 y del Consejo Quintanarroense de Ciencia y Tecnología (COQCYT, México).

II. Estado del arte

En las propuestas recientes de sistemas recomendadores para la colaboración científica se enfocan en técnicas basadas en redes neuronales [2], algoritmos de agrupamiento [3] y en el contenido [4]. En cuanto a sistemas recomendadores en este ámbito basados en lógica borrosa en [5] se presenta una propuesta para su uso en grupos multidisciplinares por medio del enfoque tradicional basado en contenido pero en conjunto con el modelado lingüístico borroso para representar y manejar la información por medio de etiquetas lingüísticas.

III. Métodos borrosos de recomendación

Hay dos aspectos importantes dentro del ámbito de la investigación que facilitan la recomendación de elementos para la colaboración científica: la afinidad bibliométrica (obtenida con los indicadores que representan la relevancia de la investigación) y la afinidad semántica (considerada a partir

del área de conocimiento y el contenido de sus publicaciones). Dado que los motivos de la colaboración pueden ser diferentes según las circunstancias estos aspectos no siempre tendrán el mismo grado de relevancia. Es por ello que se define un proceso de asignación de pesos a cada uno de los criterios según el motivo de colaboración, donde a mayor relevancia, mayor el peso del criterio. Entre las técnicas de lógica borrosa que permiten el uso de pesos y recomendar elementos están las relaciones borrosas y la representación lingüística basada en dos tuplas. Supongamos un ejemplo: Un grupo de investigación en sistemas distribuidos y paralelos necesita colaborar con otros grupos para la financiación de proyectos, donde uno de los requisitos es la afinidad entre los grupos colaboradores, independientemente de la relevancia en la comunidad científica. Por lo tanto, de los criterios a considerar, la afinidad semántica debe tener un peso mayor que la afinidad bibliométrica.

III-A. Relaciones borrosas

El conjunto de grupos candidatos GC se compone por cinco grupos $GC = \{gc_1, gc_2, gc_3, gc_4, gc_5\}$. La relación borrosa θ obtenida contiene los criterios considerados con valores $[0, 1]$ y cada fila es un grupo candidato gc, la primera columna representa la afinidad bibliométrica y la segunda columna la afinidad semántica.

$$\theta = \begin{pmatrix} 0{,}79 & 0{,}81 \\ 0{,}69 & 0{,}75 \\ 0{,}94 & 0{,}44 \\ 0{,}70 & 0{,}65 \\ 0{,}85 & 0{,}01 \end{pmatrix} \tag{1}$$

Para asignar los valores de los criterios a considerar para definir una colaboración, en un vector borroso V se establecen los pesos $V = [0{,}3\ \ 0{,}9]$ y posteriormente se normaliza el vector $V^N = [0{,}25\ \ 0{,}75]$. Finalmente se componen las relaciones a través de la suma-producto, donde el resultado permite ordenar la prioridad de las posibles colaboraciones siendo el gc_1 el grupo más recomendable.

$$\theta \star V^N = \begin{pmatrix} 0{,}80 \\ 0{,}73 \\ 0{,}56 \\ 0{,}66 \\ 0{,}22 \end{pmatrix} \tag{2}$$

III-B. Representación lingüística basada en dos tuplas

La afinidad entre los grupos de investigación para una colaboración puede representarse por medio de etiquetas lingüísticas las cuales pueden ser operadas por medio del enfoque simbólico a partir de los índices del conjunto de etiquetas, y que para evitar la pérdida de la información se utiliza la representación lingüística basada en dos tuplas, donde cada tupla se compone por un término lingüístico (definido) y un valor numérico α que indica la traslación simbólica.

A partir de los pesos mencionados en V^N del ejemplo anterior y las etiquetas lingüísticas y grados de pertenencia presentados en la Fig. 1 (considerados para la afinidad bibliométrica, semántica y general, que es el resultado que indica qué tan recomendable es colaborar con un grupo) se obtuvieron las dos tuplas para cada uno de los grupos pertenecientes al conjunto GC (Tabla I), siendo el gc_1 el elemento más recomendable.

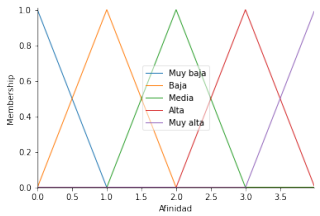


Figura 1. Etiquetas lingüísticas para la afinidad.

Tabla I
RESULTADO REPRESENTACIÓN LINGÜÍSTICA BASADA EN 2-TUPLAS

Grupo	Afinidad bibliométrica	Afinidad semántica	Afinidad general
gc_1	(Alta, 0,16)	(Alta, 0,24)	(Muy alta, -0,13)
gc_2	(Alta, -0,24)	(Alta, 0)	(Muy alta, -0,47)
gc_3	(Muy alta, -0,24)	(Media, -0,24)	(Alta, -0,28)
gc_4	(Alta, -0,2)	(Alta, -0,4)	(Alta, 0,18)
gc_5	(Alta, 0,4)	(Muy baja, 0,04)	(Baja, 0,05)

IV. CONCLUSIONES

En este trabajo se han presentado dos técnicas basadas en la lógica borrosa para un sistema recomendador para la colaboración científica, mismas que se ajustan a la necesidad de recomendar a partir de un motivo específico, por lo que el uso de los pesos dan lugar a recomendar dentro de una necesidad. En lo que concierne al uso de las relaciones borrosas, con los valores obtenidos se identifica de manera ordenada los elementos recomendables para la colaboración. En cuanto al uso de la representación lingüística basada en dos tuplas se identifica con palabras el nivel de afinidad entre los grupos. Como trabajo a futuro se pretende añadir nuevos criterios para la definición de las colaboraciones además de enriquecer los existentes.

REFERENCIAS

[1] L. Aldieri, M. Kotsemir, and C. P. Vinci, "The impact of research collaboration on academic performance: An empirical analysis for some European countries," *Socio-Econ. Planning Sci.*, vol. 62, pp. 13–30, 2018.

[2] X. Chen, T. Tang, J. Ren, I. Lee, H. Chen, and F. Xia, "Heterogeneous Graph Learning for Explainable Recommendation over Academic Networks," in *IEEE/WIC/ACM International Conference on Web Intelligence*, (New York, NY, USA), pp. 29–36, ACM, dec 2021.

[3] M. W. Rodrigues, M. A. J. Song, and L. E. Zárate, "Effectively clustering researchers in scientific collaboration networks: case study on ResearchGate," *Social Network Analysis and Mining*, vol. 11, dec 2021.

[4] T. Koopmann, K. Kobs, K. Herud, and A. Hotho, "CoBERT: Scientific Collaboration Prediction via Sequential Recommendation," in *IEEE International Conference on Data Mining Workshops, ICDMW*, vol. 2021-December, pp. 45–54, 2021.

[5] C. Porcel, A. López-Herrera, and E. Herrera-Viedma, "A Recommender System to Promote Collaborative Research Groups in an Academic Context," in *Computational Intelligence in Decision and Control*, pp. 847–852, World Scientific, aug 2008.

Aprendizaje federado de clasificadores difusos

Samuel López Domínguez
Dept. de Tecnologías de la Información
Universidad de Huelva
Huelva, España
samuel.lopez@dti.uhu.es

Javier Martín Moreno
Dept. de Tecnologías de la Información
Universidad de Huelva
Huelva, España
javier.martin@dti.uhu.es

Antonio Peregrín Rubio
Instituto Andaluz Interuniversitario de Investigación en Data Science and Computational Intelligence (DaSCI)
Centro de Estudios Avanzados en Física, Matemáticas y Computación
Universidad de Huelva
Huelva, España
peregrin@dti.uhu.es

***Resumen*—** Los modelos de Machine Learning actuales requieren de una gran cantidad de datos para ser entrenados. Debido al gran volumen de estos, para su procesamiento, a menudo es necesario el uso de un clúster con un sistema de almacenamiento compartido, lo que implica que los datos sean tratados en un dispositivo diferente a donde se originó, con la pérdida de privacidad que esto conlleva. Como alternativa a estos sistemas y con el objetivo de preservar la privacidad, surge el *Federated Learning* o Aprendizaje Federado. En este trabajo proponemos un método para aprender modelos basados en reglas difusas para clasificación basándonos en este nuevo paradigma. Asimismo, se desarrolló un estudio experimental consiguiendo unos resultados prometedores.

***Palabras clave*—** *Federated Learning*; Clasificación basada en reglas difusas

I. Introducción

En la era de la Inteligencia Artificial (IA) basada en datos en las que nos encontramos, con gran frecuencia, se precisan cantidades ingentes de datos para construir modelos precisos y robustos, lo que implica a menudo el uso de sistemas distribuidos con el objetivo de distribuir la carga de trabajo. Estos sistemas suelen estar basados en un almacenamiento de datos centralizado, por lo que es necesario que en algún momento los datos sean enviados desde el dispositivo donde se generó hacia este almacenamiento. En ocasiones, por este motivo, los sistemas de procesamiento distribuido tradicionales dejan de ser una opción viable cuando nos enfrentamos a un entorno donde la privacidad es un requisito.

Como alternativa a estos sistemas, surge un nuevo paradigma diseñado para preservar la privacidad de los datos, conocido como Aprendizaje Federado (*Federated Learning* (FL)) [1][2][3]. Su diferencia con los sistemas distribuido tradicionales en que no existe transmisión de datos entre dispositivos; se basa en un proceso iterativo conocido como *ronda*, donde un conjunto de nodos clientes descargan un modelo global y lo ajustan utilizando sus datos locales. A continuación, los modelos ajustados localmente, son enviados a un servidor donde se genera un nuevo modelo global aplicando un operador de agregación.

Algunos de los citados operadores de agregación más empleados son:

- FedAvg [4] actualiza el modelo global una vez que recibe las actualizaciones de todos los clientes. El modelo global se actualiza promediando los parámetros de los modelos locales.
- CO-OP [5] actualiza el modelo global cada vez que recibe una actualización de algún cliente. Define una edad asociada a los modelos para no agregar actualizaciones realizadas sobre modelos obsoletos.

El objetivo de este trabajo es presentar un modelo de aprendizaje de clasificadores basado en reglas difusas que pueda hacer frente a las limitaciones del aprendizaje automático distribuido para preservar la privacidad de los datos.

El trabajo está organizado de la siguiente forma: en la sección II se presenta la metodología seguida para la generación de reglas difusas en entornos basados en FL; en la sección III presentamos los resultados obtenidos con el modelo propuesto, para finalizar con una sección IV dedicada a las conclusiones del trabajo.

II. Descripción del Modelo Propuesto

En esta sección se describe el modelo propuesto, el cual denominamos CHI-FL, y que consiste en un algoritmo de aprendizaje distribuido de reglas difusas diseñado con el objetivo de preservar la privacidad de los datos. Para ello, hemos adaptado el algoritmo introducido por Chi *et al.*'s [6] siguiendo los principios del FL.

Nuestra propuesta consiste en un aprendizaje incremental, llevado a cabo a través de las diferentes rondas, donde los conjuntos de datos procesados por cada cliente y ronda son diferentes. Cada ronda federada está compuesta por las siguientes fases:

- La primera fase se encarga de la generación de reglas difusas sin pesos (o factores de confianza, FCs), a partir de los datos disponibles en cada dispositivo para la ronda dada.
- La segunda fase se encarga del cálculo del FC de cada regla, así como la resolución de las posibles reglas contradictorias.

A. Generación de la base de reglas

El procedimiento para generar una base de reglas siguiendo un esquema basado en FL se divide en los siguientes procesos:

1) Generación de reglas en el lado del cliente: cada cliente, utilizando los datos disponibles para la ronda, genera por cada instancia una regla difusa de la siguiente forma: en primer lugar, calcula el grado de pertenencia de la instancia con las distintas etiquetas difusas; a continuación, por cada variable, se selecciona la región difusa para la que obtiene un mayor grado de pertenencia, formando así los antecedentes; finalmente, se usa la clase de la instancia como consecuente de la regla. Finalizado este proceso, se envían las reglas generadas al servidor. En este proceso, cabe señalar que no se calculan los FCs de cada regla.

2) Recepción de las reglas en el lado del servidor: el servidor recibe y almacena las reglas difusas generadas por

los distintos clientes. Una vez que todas son recibidas, el servidor envía una copia a cada uno de los clientes.

B. Cálculo de los FCs y resolución de reglas contradictorias

El objetivo de esta segunda fase es el eliminar las reglas contradictorias (reglas con los mismos antecedentes y distinto consecuente) generadas en la fase anterior. Una práctica común para esta tarea consiste en añadir a cada regla un peso y mantener, de entre las contradictorias, la que tiene el citado peso mayor.

Existen varios métodos para el cálculo del FC. En este trabajo utilizamos el método heurístico conocido como *Penalized Certainty Factor* (PCF) [7]:

$$PCF_{Class_i} = \frac{2*\sum_{x_p \in Class_i} \mu_A(x_p)}{\sum \mu_A(x_p)} - 1 \qquad (1)$$

donde $\mu_A(x_p)$ es el grado de pertenencia del ejemplo x_p con los antecedentes de la regla y $Class_i$ es la clase asociada a dicha regla.

El procedimiento seguido para calcular los FCs siguiendo un esquema basado en FL, se divide en los siguientes procesos:

1) Cálculo de los grados de pertenencia en el lado del cliente: cada cliente, utilizando los datos disponibles para la ronda, calcula por cada regla el sumatorio de los grados de pertencia de la misma para las distintas clases y los envía al servidor.

2) Agregación de los sumatorios en el lado del servidor: el servidor recibe de cada cliente, para cada regla y clase posible, un sumatorio de los grados de pertenencia con dicha clase. A continuación, se agregan estos parámetros de forma individual y se calcula el FC de cada regla aplicando la fórmula descrita en (1). Finalmente se eliminan las reglas contradictorias y se envía una copia a cada uno de los clientes.

En cada ronda, los clientes reciben una nueva base de regla del servidor, la cual fusionan con las anteriores resolviendo las reglas contradictorias.

III. Estudio Experimental

En esta sección se describe el estudio experimental llevado a cabo para comprobar la calidad del modelo propuesto.

A. Configuración

Para la experimentación se han considerado dos problemas de clasificación: *Magic* y *Covtype*. La Tabla I muestra un resumen de las principales características de estos.

TABLA I. Resumen de los datastes empleados

Datasets	#Ejemplos	#Atributos	Clases	#Ejemplos por clase
Magic	19020	10	(g; h)	(12332; 6688)
Covtype_2_vs_1	495141	54	(2; 1)	(283301; 211840)

En el desarrollo de los distintos experimentos se ha hecho uso de validación cruzada de orden 5. Además, para simular un entorno federado, cada conjunto de entrenamiento se dividió en 8 dispositivos y 16 rondas.

En cuanto a los parámetros, hemos utilizado tres términos lingüísticos triangulares para cada variable, el producto como operador de conjunción, el PCF para el cálculo del FC de la regla, y la regla ganadora como método de inferencia.

La métrica empleada para medir la calidad del modelo es el *accuracy*, que se define de la siguiente forma:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

B. Resultados y análisis

En esta sección se muestran los resultados obtenidos con nuestra propuesta comparándolos con los obtenidos con el método original propuesto por Chi *et al*'s.

TABLA II. Accuracy medio obtenido

Datasets	CHI		CHI-FL	
	# Reglas	Accuracy	# Reglas	Accuracy
Magic	337.2	0.7680	337.2	0.7667
Covtype_2_vs_1	7988.4	0.7447	7988.4	0.7454

Como puede observarse en Tabla II, el número de reglas obtenido para ambos métodos es muy similar al del método original. Asimismo, los niveles de precisión son similares, consiguiendo el método original una leve mejor precisión con Magic y al contrario a favor de CHI-LF con *Covtype*.

IV. Conclusiones

En este trabajo hemos presentado un nuevo método para el aprendizaje de reglas difusas en entornos distribuidos el cuál es capaz de preservar la privacidad de los datos. Para ello se ha adaptado el algoritmo propuesto por Chi *et al*'.s, siguiendo un esquema basado en FL, consiguiendo una precisión cercana a la obtenida con la versión secuencial.

Como trabajo futuro se plantea estudiar cómo se comporta este método utilizando más nodos clientes y/o más rondas federadas. Además, se propone estudiar distintos métodos para fusionar las BR en el lado del cliente una vez calculados los FCs.

Referencias

[1] J. Konečný, H. B. McMahan, D. Ramage, & P. Richtárik , "Federated optimization: Distributed machine learning for on-device intelligence", 2016. arXiv preprint arXiv:1610.02527.

[2] H. B. McMahan and Daniel Ramage, "Federated learning: Collaborative machine learning without centralized training data", 2017. https://research.googleblog.com/2017/04/federated-learning-collaborative.html

[3] R. Shokri, & V. Shmatikov, "Privacy-preserving deep learning", in Proceedings of the 22nd ACM SIGSAC conference on computer and communications security 2015, pp. 1310-1321.

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, & B. A. Arcas, (2017, April). "Communication-efficient learning of deep networks from decentralized data", in Artificial intelligence and statistics (pp. 1273-1282). PMLR.

[5] Y. Wang, "Co-op: Cooperative machine learning from mobile devices" (Ph.D thesis), University of Alberta Libraries, 2017.

[6] Z. Chi, H. Yan, & T. Pham, "Fuzzy algorithms: with applications to image processing and pattern recognition", World Scientific (Vol. 10), 1996.

[7] H. Ishibuchi, & T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems", IEEE transactions on fuzzy systems, 13(4), 428-435, 2005.

Construcción de órdenes admisibles y sus aplicaciones en interfaces cerebro-ordenador

Jonata Wieczynski
Dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
jonata.wieczynski@unavarra.es

Anderson Cruz
Institudo Metrópole Digital
Universidade Federal do Rio Grande do Norte
Natal, Brasil
anderson.cruz@ufrn.br

Javier Fernández
Dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
fcojavier.fernandez@unavarra.es

Graçaliz Dimuro
Centro de Ciências Computacionais
Universidade Federal do Rio Grande
Rio Grande, Brasil
gracalizdimuro@furg.br

Humberto Bustince
Dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
bustince@unavarra.es

***Index Terms*—Órdenes admisibles, generación de órdenes totales, interfaces cerebro-ordenador**

Los órdenes en la recta real son bastante intuitivos de interpretar y aplicar. Sin embargo, cuando trabajamos con intervalos, este concepto muchas veces no es tan sencillo de comprender e interpretar. Por ejemplo, un modo usual de definir un orden en el conjunto cerrado de subintervalos en $[0,1]$, que denotamos por $L^I = L(I) = \{X = [\underline{X}, \overline{X}] \mid \underline{I} \leq \underline{X} \leq \overline{X} \leq \overline{I}\}$, donde en este trabajo $I = [\underline{I}, \overline{I}] = [0,1]$, es usando el orden del producto $\leq_P$ en L^I, con $X, Y \in L^I$, $[\underline{X}, \overline{X}] \leq_P [\underline{Y}, \overline{Y}] \iff \underline{X} \leq \underline{Y} \wedge \overline{X} \leq \overline{Y}$.

Este orden es un orden parcial en L^I, esto es, existen elementos $X, Y \in L^I$ que no son comparables. Pero muchas veces necesitamos que todos los elementos sean comparables, y para eso debemos usar órdenes totales. Pensando en eso y en poder usar alguna derivación del orden del producto, Bustince et al. [1] presentaron el concepto de orden admisible, esto es, un orden $\preceq$ en L^I es admisible si: (i) $\preceq$ es un orden total; y (ii) $\preceq$ refina el orden del producto, esto es, para cualquier $X, Y \in L^I$ tenemos que $X \preceq Y$ siempre que $X \leq_P Y$.

En el mismo artículo que introdujeron los órdenes admisibles, los autores presentaron un método de construcción para ellos basados en dos funciones de agregación A y B, donde estas deben tener la propiedad de que dados y dos pares $(x,y),(u,v) \in L^I$ las igualdades $A(x,y) = A(u,v)$ y $B(x,y) = B(u,v)$ se cumplen simultáneamente solo si $(x,y) = (u,v)$.

Teniendo esto en cuenta nosotros extendemos este método de construcción para órdenes totales para cualesquiera dos funciones F, G que tengan la misma propiedad que las agregaciones A y B, esto es, para cualquier $(x_1,x_2),(y_1,y_2) \in L^I$, si $F(x_1,x_2) = F(y_1,y_2)$ y $G(x_1,x_2) = G(y_1,y_2)$ entonces $(x_1,x_2) = (y_1,y_2)$.

Adicionalmente, mostramos un método de construcción para órdenes totales usando una función univariada y además probamos en qué condiciones estos órdenes son admisibles. Por último, proporcionamos algunos ejemplos de órdenes construidos por estos métodos. Como ejemplo de posible aplicación de estos nuevos órdenes consideramos las interfaces cerebro-ordenador (*brain-computer interfaces* - BCI).

Las interfaces cerebro-ordenador se han convertido en un tópico de investigación muy atractivo en los últimos años. Usando desde métodos sencillos como reglas y clasificación binaria hasta redes neuronales complejas, los investigadores intentan comprender mejor el cerebro humano y también optimizar su comunicación con sistemas informáticos.

De entre los métodos de captura de señales del cerebro que no son invasivos - no requieren cirugía - dos son muy usados: la espectroscopía funcional del infrarrojo cercano (*functional near-infrared spectroscopy* - fNIRS) [2] y la electroencefalografía (EEG) [3]. Ambos funcionan con un casco de sensores puesto en la cabeza del individuo. Por un lado, el fNIRS mide los cambios de oxígeno en la hemoglobina (oxihemoglobina y desoxihemoglobina) por dos o más sensores. Y por otro lado, el EEG mide la diferencia de energía potencial entre dos sensores.

En este trabajo, nuestro objetivo es analizar si existen diferencias al usar órdenes admisibles en lugar de parciales para el análisis de datos representados por medio de intervalos en el problema del BCI. Para conseguir este objetivo, utilizaremos una red neuronal adaptada a los intervalos. En cuestión de los datos, utilizaremos datos abiertos y ampliamente difundidos

Este trabajo ha sido financiado en parte por Navarra de Servicios y Tecnologías, S.A. (NASERTIC), en parte por CNPq 301618/2019-4, en parte por FAPERGS 19/2551-0001660-3 y en parte por el Ministerio Español de Ciencia e Innovación PID2019-108392GBI00 (AEI/10.13039/501100011033).

en la literatura como, por ejemplo, los de *BCI Competition IV*[1], donde mediante la red neuronal clasificaremos los datos utilizando múltiples órdenes y estudiando si, tanto la admisibilidad cuanto las órdenes en si afectan al algoritmo.

REFERENCES

[1] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar, "Generation of linear orders for intervals by means of aggregation functions," *Fuzzy Sets and Systems*, vol. 220, pp. 69–77, 2013.

[2] M. Kiani, J. Andreu-Perez, H. Hagras, S. Rigato, and M. L. Filippetti, "Towards understanding human functional brain development with explainable artificial intelligence: Challenges and perspectives," *IEEE Computational Intelligence Magazine*, vol. 17, no. 1, pp. 16–33, 2022.

[3] S. Aggarwal and N. Chugh, "Signal processing techniques for motor imagery brain computer interface: A review," *Array*, vol. 1-2, 2019.

[1] https://www.bbci.de/competition/iv/

Recomendación difusa para la personalización automática de ejercicios de rehabilitación física en pacientes con ictus

Cristian Gmez-Portes
Dept. Information Technologies
University of Castilla-La Mancha
Ciudad Real, Spain
Cristian.Gomez@uclm.es

José J. Castro-Schez
Dept. Information Technologies
University of Castilla-La Mancha
Ciudad Real, Spain
JoseJesus.Castro@uclm.es

Javier Albusac
Dept. Information Technologies
University of Castilla-La Mancha
Ciudad Real, Spain
Javier.Albusac@uclm.es

Carlos Glez-Morcillo
Dept. Information Technologies
University of Castilla-La Mancha
Ciudad Real, Spain
Carlos.Gonzalez@uclm.es

Dorothy N. Monekosso
Dept. Computer Science
Durham University
Durham, United Kingdom
dorothy.monekosso@durham.ac.uk

David Vallejo
Dept. Information Technologies
University of Castilla-La Mancha
Ciudad Real, Spain
David.Vallejo@uclm.es

***Abstract*—El presente artículo es un resumen del trabajo *A Fuzzy Recommendation System for the Automatic Personalization of Physical Rehabilitation Exercises in Stroke Patients*, publicado en la revista Mathematics en Junio de 2021 [Gmez-Portes, C.(2021)]. El ictus se encuentra entre las 10 principales causas de muerte y discapacidad en todo el mundo. Los pacientes que sufren esta enfermedad suelen realizar ejercicios físicos en casa para mejorar su estado. Estos ejercicios son recomendados por los terapeutas en función del nivel de progreso del paciente, y pueden ser supervisados a distancia por ellos si la tecnología es una opción para ambos. Existe el reto de adaptar dinámicamente el plan terapéutico del paciente en tiempo real mientras se rehabilita en casa, ya que su evolución varía a medida que avanza el proceso de rehabilitación. En este contexto, presentamos un sistema difuso que es capaz de adaptar automáticamente el plan de rehabilitación de los pacientes con ictus. El uso de la lógica difusa facilita enormemente el seguimiento y la orientación de los pacientes con ictus. Además, el sistema es capaz de generar automáticamente modificaciones de los ejercicios existentes teniendo en cuenta sus particularidades en cada momento. Se ha realizado un experimento preliminar para mostrar las ventajas de la propuesta, y los resultados sugieren que la aplicación de la lógica difusa puede ayudar a tomar decisiones correctas en función del nivel de progreso del paciente.**

***Index Terms*—rehabilitación remota, sistema de recomendación difuso, ictus, telemedicina**

I. Contexto

En el contexto del diseño y desarrollo de sistemas de rehabilitación remota para pacientes de ictus, nuestro trabajo se ha centrado esencialmente en la creación de un sistema integral de rehabilitación a distancia capaz de evaluar y evaluar y clasificar automáticamente los ejercicios de rehabilitación [Schez-Sobrino, S.(2020)]. Sobre estas bases, también hemos diseñado un lenguaje cuyas frases son procesadas por un software que puede generar automáticamente generar automáticamente *exergames* personalizados que motiven al paciente a realizar ejercicios de rehabilitación. En este resumen, nos centramos en la personalización automatizada e inteligente del proceso de rehabilitación adaptado a cada paciente [Vallejo, D.(2020)]. Así, este trabajo propone un sistema difuso para la recomendación de ejercicios de rehabilitación para pacientes con ictus. Este sistema integra una base de conocimiento experto definida mediante reglas difusas y variables que reflejan aspectos como el rendimiento de un paciente al realizar ejercicios de rehabilitación. El sistema es capaz de adaptar el plan de rehabilitación asignado inicialmente por el terapeuta. Esta adaptación se concreta en recomendaciones de nuevos ejercicios, o ejercicios ya realizados por el paciente, en función de la variación del nivel de del paciente a medida que avanza el proceso de rehabilitación.

Este trabajo ha sido financiado por la Universidad de Castilla-La Mancha, a través de las orgánicas de ayuda a grupos de investigación (2021-GRIN-31017) y del Departamento de Tecnologías y Sistemas de Información (00421372). Por otra parte, este trabajo también es resultado de la estancia de investigación realizada por Cristian Gmez-Portes en la universidad de Durham (Reino Unido), financiada por Becas Santander Investigación para estancias predoctorales en el extranjero.

II. Arquitectura del sistema y módulo de recomendación para fisioterapia

El sistema presentado en este artículo consta de varios componentes que interactúan entre sí. La figura 1 representa la arquitectura general, cuyos componentes se describen brevemente a continuación: i) módulo de conocimiento del dominio, que integra los conocimientos necesarios para que el sistema funcione correctamente; ii) módulo de interfaz, que permite la comunicación entre el paciente/terapeuta y el sistema, e incluye los mecanismos de interacción adecuados para que los pacientes realicen los ejercicios de rehabilitación como si fueran juegos; iii) módulo de seguimiento, que se encarga

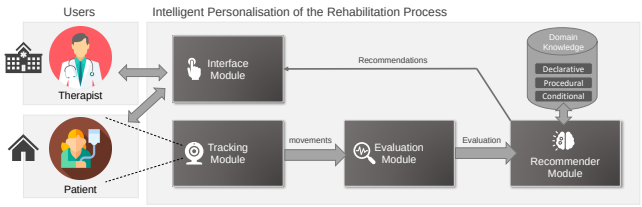


Fig. 1. Visión general de la arquitectura propuesta.

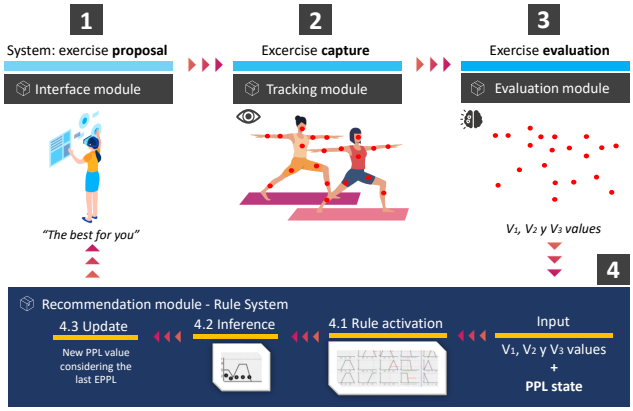


Fig. 2. Vista holística del módulo de recomendación automática.

de reconocer en tiempo real los movimientos de los usuarios para realizar los ejercicios de rehabilitación; iv) módulo de evaluación, que valora el rendimiento de un ejercicio de rehabilitación según el nivel de complejidad que supone para el paciente; y v) módulo de recomendación, que se encarga de modificar el plan de rehabilitación del paciente en función de su estado actual.

Este último módulo hará uso del conocimiento condicional para determinar el ejercicio más apropiado para el paciente en función de su lesión y su nivel actual de progreso en su plan de rehabilitación. La figura 2 muestra una visión holística del módulo de recomendación. Dicho módulo se basa en la definición de una función que determinará una acción en la terapia que se ajustará, en la medida de lo posible, al estado de su recuperación. Por tanto, dicha función modela la relación existente entre el dominio, es decir, el estado del paciente en la recuperación del miembro lesionado, y el codominio, es decir, una acción a realizar en el plan de rehabilitación.

El dominio está determinado por cómo el paciente ha realizado el último ejercicio de su plan, así como por el estado en el que se encuentra respecto a su recuperación. Esto se recoge por medio de un conjunto de variables lingüísticas $V1, V2, V3$, las cuales no tienen en cuenta su evolución, es decir, cómo han progresado. Estas variables representan, respectivamente, i) la diferencia entre el número de pasos realizados por el paciente y el terapeuta, ii) la desviación acumulada al realizar el ejercicio por el paciente, y iii) la diferencia de tiempo entre paciente y terapeuta.

Fig. 3. Representación visual del ejercicio de abducción de hombro izquierdo.

Dado que es importante añadir esta información al dominio para conocer con detalle el estado real del paciente, el sistema incluye una cuarta variable $V4$, con el objetivo de conocer el estado en el que se encuentra un paciente antes de realizar el último ejercicio. Esta variable se denomina *nivel de progreso*. Su valor se tiene en cuenta para elegir los ejercicios del plan de rehabilitación del paciente.

Por otro lado, el codominio debe reflejar el siguiente ejercicio a realizar por los pacientes en función de su estado. En este trabajo de investigación, el plan de rehabilitación inicial establecido por el terapeuta se modifica proponiendo un nuevo ejercicio basado en una dificultad deseable para el paciente. Así, la salida consistirá en un nuevo ejercicio, que se establece en función de la dificultad dada por dos funciones: (1) propone un ejercicio del conjunto o (2) recomienda la repetición del último ejercicio con una ligera modificación.

Es importante destacar que la ejecución del último ejercicio influye en el nivel del progreso del paciente (la figura 3 muestra uno de los ejercicios utilizados en el sistema y gestionados por el módulo de recomendación automática). Dicha información se introduce en la salida del sistema de recomendación mediante una variable lingüística $V5$, denominada *efecto sobre el nivel de progreso* del paciente.

REFERENCES

[Gmez-Portes, C.(2021)] Gmez-Portes, C.; Castro-Schez, J.J.; Albusac, J.; D.N. Monekosso; Vallejo, D.; A Fuzzy Recommendation System for the Automatic Personalization of Physical Rehabilitation Exercises in Stroke Patients. *Mathematics (MDPI)* **2021**, *9(12)*, 1–24.

[Gmez-Portes, C.(2021)] Gmez-Portes, C.; D. Carneros-Prado; Castro-Schez, J.J.; Albusac, J.; D.N. Monekosso; Vallejo, D.; PhyRe Up! A system based on mixed reality and gamification to provide home rehabilitation for stroke patients. *IEEE Access (IEEE)* **2021**, *9*, 139122-139137.

[Gupta, M.M. (1991)] Gupta, M.M.; Qi, J. Theory of T-norms and fuzzy inference methods, *Fuzzy Sets and Systems*, **1991**, *40*, 431–450.

[Schez-Sobrino, S.(2020)] Schez-Sobrino, S.; Vallejo, D.; Monekosso, D.N.; Glez-Morcillo, C.; Remagnino, P. A distributed gamified system based on automatic assessment of physical exercises to promote remote physical rehabilitation, *IEEE Access (IEEE)* **2020**, *8*, 91424–91434.

[Vallejo, D.(2020)] Vallejo, D.; Gmez-Portes, C.; Albusac, J.; Glez-Morcillo, C.; Castro-Schez, J.J. Personalized Exergames Language: A Novel Approach to the Automatic Generation of Personalized Exergames for Stroke Patients. *Applied Sciences (MDPI)* **2020**, *10(20)*, 1–30.

[Zadeh, L.A] Zadeh, L.A. The concept of a linguistic variable and its application to approximate reasoning-I. *Information and Control (Elsevier)* **1975**, *8(3)*, 199–249.

On the definition of some social networks concepts in Granular Fuzzy graphs: The case of Rich Club

1st Clara Simón de Blas
Computer Science and Statistics
Rey Juan Carlos University
Móstoles, Madrid
clara.simon@urjc.es 0000-0002-6963-1295

2nd Daniel Gómez Gonzalez
Faculty of Statistic
Complutense University of Madrid
Madrid, Spain
dagomez@estad.ucm.es or 0000-0001-9548-5781

3rd Regino Criado Herrero
Center for Computational Simulation, Universidad Politécnica de Madrid, Spain.
Data, Complex Networks and Cybersecurity Research Institute, URJC, Spain
regino.criado@urjc.es 0000-0001-7954-6169

***Abstract*—In this work we extend the definition of the set of outstanding nodes in a network, by defining the granular fuzzy rich club. We will focus on networks where relationships between nodes are weighted and directed through a precedence relation in a fuzzy context. Results shows improvements in detecting influential nodes masked in a crisp context.**

***Index Terms*—Granular soft computing, social network analysis**

I. Introduction

The analysis of social networks has become in one of the hottest disciplines of data analysis due to its natural application in areas as diverse as: biology, medicine, transportation. communication, sociology, marketing, mathematics or computer science (see for example [1]–[6] among many others). The well-known term of social network analysis refers to the study of graphs. A graph is usually modeled as a set of nodes with relationships (directed or not, valued or not) between its members. Uncertainty related with the values or existence of the relation between nodes appears in a natural way when modeling real problems.

The inclusion of soft computing in social network problems is clearly a significant improvement to deal with real problems since relationships is more natural to consider soft than something crisp. However, in all these modelizations, it is assumed that it is possible to know for any pair of nodes of a network the existence or not of relationship and its valuation. It is important to mention that for very large networks, this hypothesis is very unrealistic (at least computationally very expensive). With the idea of solving this problem but continuing to incorporate the soft nature of relationships, the concept of Granular Fuzzy Social Network was introduced in [7]. Granular computing is an emerging computing topic that concerns the processing of complex information entities called *information granules* that for this case could avoid the problem of complexity or at least reduce drastically.

Identify applicable funding agency here. If none, delete this.

In this this work we introduce the concept of rich-club coefficient for granular social network that allows us the incorporation of a mechanism for searching power groups in modeled networks in a blurred way with incomplete information represented by granular theory.

This paper is organized as follow: in section 2 we present the fuzzy granular social network. In section 3, we present a novel and natural extension of the concept of rich-club coefficient to the case of granular fuzzy sets. In section 4, we present some short conclusions of this seminal work.

II. Fuzzy Granular social networks

The fuzzy Granular social networks was introduced in [7], [8]. As pointed by the authors, a social network can be viewed from the stand point of nodes and their relationships where global phenomenon ensembles the local behaviors of individuals and their closely related neighborhoods. The relations of a social network can be represented by a graph $G = (V, E)$, where V is the set of all nodes and E represents the relationships (or edges). Let us define by I the unit interval $[0, 1]$.

A fuzzy granular neighborhood defined over the vertex set of a social network G was defined in [7] by means of a function $\phi : V \longrightarrow A(V)$, which assigns every node $c \in V$ to a fuzzy set $\phi_c \in I^V$. If ϕ_c is non empty, we call it the fuzzy neighborhood of the node c, i.e., ϕ_c is the granule defined around the node c. Once this fuzzy neighborhood is introduce the concept of fuzzy granular social network can be described as follow.

Definition 1: [7] Let a family of fuzzy sets associated with each node $c \in V$ be Φ_c. Φ_c represents the neighbourhood sets of node c. A **fuzzy granular social network** is represented by a triple: $S = (C, V, G)$ Where

- V is a finite set of nodes of the network
- $C \subset V$ is a finite set of granule representatives
- G is the finite set of all granules, i.e., $G = \{\bigcup\{\Phi_c,\ c \in C\}\}$.

III. Rich club in granular fuzzy Social network

In this paper, we are interested in a very important and structural property of complex network: the so-called **rich-club** phenomenon. This structural property has been studied for many authors with special relevance in social and computer science. The rich-club phenomenon refers to the tendency of high degree or power-full nodes to be connected with the other high degree nodes. If we will refer to these high degree or powerfull nodes as *rich-nodes*, then we can understand a rich club phenomenon as rich nodes - are much more likely to form tight and well interconnected subgraphs (clubs) than low degree nodes.

In order to define the concept of *rich club*, it is necessary first to introduce the idea of degree in fuzzy granular graphs.

A. The degree in a fuzzy granular social network

Once the concepts of fuzzy granular social network and directed fuzzy granular social network have been defined, inspired by [7] we propose the following concept of granular fuzzy degree as follow:

Definition 2: (Fuzzy Granular Degree of a node). Given a fuzzy granular social network $S = (C, V, G)$. The **fuzzy granular degree** $FGD(c)$ of a node $c \in C$ is defined by the cardinality of the granule centered at the node. Formally

$$FGD(c) = |\phi_c| = \sum_{v \in V} \phi_c(v).$$

Assuming that the granular set $C = V$ is the entire network, in the following figure we show the fuzzy granular degree defined in this section where

$$\phi_c(v) = \frac{1}{d(c,v)+1}$$

being d any distance measure between nodes in crisp graphs.

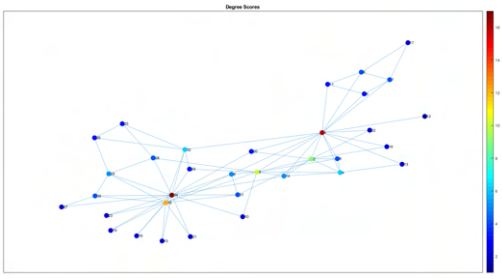


Fig. 1. Degree in the karate club network.

B. The rich coefficient in social network

In order to identify these key nodes, in [9] it was introduced the rich-club coefficient.

Definition 3: Given a graph $G = (V, E)$ and given k an integer, the **rich coefficient** associated with k is defined as

$$\varphi(k) = \frac{|E_k|}{\binom{N_k}{2}}$$

where $G_k = (N_k, E_k)$ is the subgraph generated by those nodes with degree greater or equal to k.

These coefficient can be extended to a fuzzy concept, detecting influential nodes masked in a crisp context.

C. The rich coefficient in fuzzy granular social network

Given the classical definition and once the concept of fuzzy granular degree is introduced a natural extension of the rich coefficient can be defined as follow.

Definition 4: Given a fuzzy granular social network $S = (C, V, G)$, the **granular rich-club** coefficient is defined as

$$\varphi(k)(C, V, G) = \frac{\sum_{i,j \in N_k} \phi_i(j)}{|N_k|(|N_k| - 1)},$$

where $N_k = \{c \in C \ with \ FGD(c) \geq k\}$.

IV. Conclusions

The extension of the crisp definition for rich club, improves in detecting influential nodes masked in a crisp context. It is important to observed that the fuzzy rich club coefficient here introduced for fuzzy granular networks is an extension of the classical coefficient in the case in which the fuzzy granular social network coincides with a standard graph, since the following two statements holds:

- The fuzzy granular degree $FGD(c)$ is the same that classical degree $D(c) = degree(c)$.
- A granular fuzzy social network is a classical network when $\phi_i(j) = a_{ij}$, being $A = (a_{ij})$ the adjacency matrix.

Acknowledgment

This research has been partially supported by the Government of Spain (grant TIN2015-66471-P), the Government of Madrid (grant S2013/ICCE-2845), Complutense University (UCM Research Group 910149 and grant PR26/16-21B-3) and Universidad Politécnica de Madrid.

References

[1] D. Gomez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, and J. Tejada, "Centrality and power in social networks: a game theoretic approach," Mathematical Social Sciences, vol. 46, no. 1, pp. 27–54, 2003.
[2] S. Fortunato, "Community detection in graphs," Physics reports, vol. 486, no. 3-5, pp. 75–174, 2010.
[3] M. O. Jackson, Social and economic networks. Princeton university press, 2010.
[4] S. Wasserman, K. Faust et al., Social network analysis: Methods and applications. Cambridge university press, 1994, vol. 8.
[5] J. Scott, "Social network analysis," Sociology, vol. 22, no. 1, pp. 109–127, 1988.
[6] C. S. de Blas, J. S. Martin, and D. G. Gonzalez, "Combined social networks and data envelopment analysis for ranking," European Journal of Operational Research, vol. 266, no. 3, pp. 990–999, 2018.
[7] S. Kundu and S. K. Pal, "Fgsn: fuzzy granular social networks–model and applications," Information Sciences, vol. 314, pp. 100–117, 2015.
[8] ——, "Fuzzy-rough community in social networks," Pattern Recognition Letters, vol. 67, pp. 145–152, 2015.
[9] S. Zhou and R. J. Mondragón, "The rich-club phenomenon in the internet topology," IEEE Communications Letters, vol. 8, no. 3, pp. 180–182, 2004.

One-sided formal concept analysis from the multi-adjoint algebraic framework

M. José Benítez-Caballero
Departament of Mathematics
Universidad de Cádiz
Puerto Real, Spain
0000-0002-3131-602X

Jesús Medina
Department of Mathematics
Universidad de Cádiz
Puerto Real, Spain
0000-0002-3931-5873

Eloísa Ramírez-Poussa
Department of Mathematics
Universidad de Cádiz
Puerto Real, Spain
0000-0002-2391-5217

Abstract*—This work concerns the research recently published in *M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa "Characterizing one-sided formal concept analysis by multi-adjoint concept lattices", Mathematics 10 (7), 1020.

***Index Terms*—one-sided formal concept analysis, formal concept analysis, adjoint triple, attribute reduction, reduct**

I. Introduction

Formal Concept Analysis (FCA) [5] is one of the most powerful and well-known mathematical tool developed in order to manage information and to reduce the size of databases. There are several approaches in FCA. One of those is the consideration of fuzzy subsets of attributes and fuzzy subsets of objects. For example, multi-adjoint formal concept analysis [7]. This approach generalizes diverse FCA frameworks. In this case, adjoint triples are used in order to define the concept-forming operators [8]–[10]. On the other hand, (generalized) one-sided formal concept analysis, originally studied by Krajči in [6], considers one of the sets, either the set of objects or the set of attributes, as fuzzy and the other one as crisp. In this paper, the relationship between these two theories is studied.

As we previously mentioned, the reduction of the size of a database is a main goal. Therefore, the study of the notion of reduct as well as the analysis of different mechanisms for obtaining reducts are fundamental.

In this paper, the notion of left-sided adjoint triple is presented. This new adjoint triple is essential to describe the (generalized) one-sided formal concept analysis as a particular case of the multi-adjoint formal concept analysis. As a consequence, all the theory developed within multi-adjoint structure can be applied to the one-sided framework. Specifically, the attribute reduction mechanisms and algorithms developed in the multi-adjoint framework [1]. Therefore, some notions and results related to the reduction mechanisms given in the multi-adjoint concept lattice framework will be translated into the the one-sided framework.

This research was partially supported by the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in Project PID2019-108991GB-I00 and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in Project FEDER-UCA18-108612 and by the European Cooperation in Science & Technology (COST) Action CA17124.

The notions and results obtained considering this framework are deeply studied in [4].

II. Comparison with one-sided concept lattices

This section will study the relationship of one-sided concept lattices with the multi-adjoint concept lattices. Specifically, we will show that the one-sided concept lattices are particular cases of the multi-adjoint concept lattices in which only an explicit adjoint triple is considered.

Theorem 1: Let us consider two posets $(P, \leq_P, \top_P, \bot_P)$, and $(Q, \leq_Q, \top_Q, \bot_Q)$. The mappings $\&_l \colon P \times Q \to P$, $\nwarrow_l \colon P \times P \to Q$, and $\swarrow^l \colon P \times Q \to P$ defined as follows

$$z \nwarrow_l x = \begin{cases} \top_Q & \text{if } x \leq_P z \\ \bot_Q & \text{otherwise} \end{cases} \quad (1)$$

$$z \swarrow^l y = \begin{cases} \top_P & \text{if } y = \bot_Q \\ z & \text{if } y \neq \bot_Q \end{cases} \quad (2)$$

$$x \,\&_l\, y = \begin{cases} x & \text{if } y \neq \bot_Q \\ \bot_P & \text{if } y = \bot_Q \end{cases} \quad (3)$$

for all $x, z \in P$, $y \in Q$, form an adjoint triple. This triple $(\&_l, \nwarrow_l, \swarrow^l)$ is called *left-sided adjoint triple*.

The following result shows that the generalized one-sided formal concept is a particular case of the multi-adjoint concept lattice.

Theorem 2: Given an adjoint context $(\mathcal{O}, \mathcal{P}, \mathcal{R})$ and the adjoint frame $(L, \{0,1\}, L, \&_l, \nwarrow_l, \swarrow^l)$, we obtain that $(\mathcal{O}, \mathcal{P}, \mathcal{R})$ is a one-sided formal context and the associated concept-forming operators verify the following equalities:

$$\uparrow X(a) = {\chi_X}^{\uparrow}(a)$$
$$\chi_{\downarrow(f)} = f^{\downarrow}$$

for all $X \subseteq \mathcal{O}$, $a \in \mathcal{P}$, where χ_X and $\chi_{\downarrow(f)}$ are the characteristic mappings of X and $\downarrow(f)$, respectively, and $f \in {L_1}^{\mathcal{P}}$.

Therefore, the (generalized) one-sided concept lattice framework can be seen as a particular case of the multi-adjoint case and, as a consequence, the theory developed in the last one can be applied to the former one. However, due to $\&_l$ not being a t-norm, a similar relationship with a residuated lattice cannot be given.

III. Reducing Left-Sided Concept Lattices

This section will apply the attribute reduction results given in the multi-adjoint framework [11]–[13] to the one-sided framework. Hence, we will reformulate the notions and results needed to compute the reduction of a formal context, when the left-sided adjoint triple is considered. First of all, we will rewrite the notions of consistent set and reduct on the set of objects.

Definition 1: An arbitrary set of objects $B \subseteq \mathcal{O}$ is called an object consistent set of $(\mathcal{O}, \mathcal{P}, \mathcal{R})$ if $\mathcal{M}_l(B, \mathcal{P}, \mathcal{R}_B) \cong_E \mathcal{M}_l(\mathcal{O}, \mathcal{P}, \mathcal{R})$. This is equivalent to say that, for all $\langle X, f\rangle \in \mathcal{M}_l(\mathcal{P}, \mathcal{O}, \mathcal{R})$, there exists a concept $\langle X', f'\rangle \in \mathcal{M}_l(B, \mathcal{P}, \mathcal{R}_B)$ such that $X = X'$.

Moreover, if $\mathcal{M}_l(B \setminus \{x\}, \mathcal{P}, \mathcal{R}_{B\setminus\{x\}}) \ncong_E \mathcal{M}_l(\mathcal{O}, \mathcal{P}, \mathcal{R})$, for all $b \in B$, then B is called an object reduct of $(\mathcal{O}, \mathcal{P}, \mathcal{R})$. The core of $(\mathcal{O}, \mathcal{P}, \mathcal{R})$ is the intersection of all the object reducts of $(\mathcal{O}, \mathcal{P}, \mathcal{R})$.

We are going to introduce the characterization of join-irreducible elements in a left-sided concept lattice framework. Hence, in this case, we are going to use objects to generate the concepts instead of fuzzy attributes.

Proposition 3: The set of $\vee$-irreducible elements of $\mathcal{M}_l$ is:

$$J_F(\mathcal{O}) = \left\{(x^{\uparrow\downarrow}, x^{\uparrow}) \mid x^{\uparrow} \neq \bigwedge\{x_i^{\uparrow} \mid x^{\uparrow} \subsetneq x_i^{\uparrow}\} \text{ and } x^{\uparrow} \neq f_{\perp}\right\}$$

where $\perp$ is the least element in L_2 and $f_{\perp}\colon \mathcal{P} \to L_1$ is the fuzzy subset defined as $f_{\perp}(a) = \perp$, for all $a \in \mathcal{P}$

In the following definition, we introduce a specific notation to the set of objects that generates a concept. This notion is needed in order to reformulate the classification theorems.

Definition 2: A context $(\mathcal{P}, \mathcal{O}, R)$ associated with the concept lattice $(\mathcal{M}_l, \preceq)$ and a concept C of $(\mathcal{M}_l, \preceq)$, the set of objects generating C is defined as the set:

$$\text{Obg}(C) = \{x \in \mathcal{O} \mid \text{such that} \quad \langle x^{\uparrow\downarrow}, x^{\uparrow}\rangle = C\}$$

Now, we will write the classification theorems from the point-of-view of the objects. In this case, the classification is carried out by means of the join-irreducible elements. First of all, the characterization of the objects of the core is presented.

Theorem 4: Given an object $x \in \mathcal{O}$, we have that $x \in C_{\mathcal{O}}$ if and only if there exists a join-irreducible concept C of $(\mathcal{M}_l, \preceq)$ satisfying that $x \in \text{Obg}(C)$ and $\text{card}(\text{Obg}(C)) = 1$.

The following result rewrites the characterization of the relatively necessary objects, in terms of left-sided concept lattices.

Theorem 5: Given an object $x \in \mathcal{O}$, then $x \in K_{\text{g}}$ if and only if $x \notin C_{\text{g}}$ and there exists $C \in J_{\text{F}}(\mathcal{O})$ with $x \in \text{Obg}(C)$ and $\text{card}(\text{Obg}(C)) > 1$, satisfying that $\big(\mathcal{O} \setminus \text{Obg}(C)\big) \cup \{x\}$ is an object consistent set.

Finally, the next proposition shows the characterization of the absolutely unnecessary objects considering the objects generating a concept.

Theorem 6: Given an object $x \in \mathcal{O}$, then $x \in I_{\text{g}}$ if and only if, for any $C \in J_{\text{F}}(\mathcal{O})$, $x \notin \text{Obg}(C)$, or if $x \in \text{Obg}(C)$, then $\big(\mathcal{O} \setminus \text{Obg}(C)\big) \cup \{x\}$ is not an object-consistent set.

As previously mentioned, an object can generate only one concept. Therefore, if $x \in \text{Obg}(C)$ with $C \in J_{\text{F}}(\mathcal{O})$, by Theorem 4 or Theorem 5, the object will be classified as an absolutely or relatively necessary object. Consequently, the sentence "$\big(\mathcal{O} \setminus \text{Obg}(C)\big) \cup \{x\}$ is not an object consistent set" can be erased for the theorem, obtaining the following result.

Corollary 7: Given an object $x \in \mathcal{O}$, then $x \in I_{\text{g}}$ if and only if, for any $C \in J_{\text{F}}(\mathcal{O})$, $x \notin \text{Obg}(C)$.

This characterization of the absolutely unnecessary elements is equivalent to the one presented in classical FCA.

IV. Conclusions and Future Work

In [4], a new adjoint triple has been defined in order to show the connections between the concept-forming operators in one-sided and in multi-adjoint formal concept lattices. Moreover, the study of one-sided formal concept lattices as a particular case of a multi-adjoint formal concept lattices has been developed. Also, the reduction mechanism in one-sided formal concept analysis considering the multi-adjoint philosophy has been presented. After that, The classification theorems over the set of attributes and the set of objects have also been introduced.

As a future work, property-oriented concept lattices and object-oriented concept lattices [13] will be considered in order to produce a deep study of this relationship. Also, the philosophy of one-sided concept lattice will be used in order to reduce an information system given in Rough Set Theory [2], [3].

References

[1] Antoni, L.; Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Attribute classification and reduct computation in multi-adjoint concept lattices. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 1121–1132.

[2] Benítez-Caballero, M.J.; Medina, J.; Ramírez-Poussa, E. FCA attribute reduction in information systems. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*; Medina, J., Ojeda-Aciego, M., Verdegay, J.L., Pelta, D.A., Cabrera, I.P., Bouchon-Meunier, B., Yager, R.R., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 549–561.

[3] Benítez-Caballero, M.J.; Medina, J.; Ramírez-Poussa, E.; Ślęzak, D. Bireducts with tolerance relations. *Inf. Sci.* **2018**, *435*, 26–39.

[4] Benítez-Caballero, M.J., Medina, J., Ramírez-Poussa, E.: Characterizing one-sided formal concept analysis by multi-adjoint concept lattices. Mathematics **10**(7) (2022). DOI: 10.3390/math10071020, https://www.mdpi.com/2227-7390/10/7/1020

[5] Wille, R. Restructuring lattice theory: An approach based on hierarchies of concepts. In *Ordered Sets*; Rival, I., Ed.; Reidel, Dordrecht, Netherlands: 1982; pp. 445–470.

[6] Krajči, S. Cluster based efficient generation of fuzzy concepts. *Neural Netw. World* **2003**, *13*, 521–530.

[7] Medina, J.; Ojeda-Aciego, M.; Ruiz-Calvi no, J. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets Syst.* **2009**, *160*, 130–144.

[8] Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. A comparative study of adjoint triples. *Fuzzy Sets Syst.* **2013**, *211*, 1–14.

[9] Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Algebraic structure and characterization of adjoint triples. *Fuzzy Sets Syst.* **2021**, *425*, 117–139.

[10] Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Implication operators generating pairs of weak negations and their algebraic structure. *Fuzzy Sets Syst.* **2021**, *405*, 18–39.

[11] Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Attribute reduction in multi-adjoint concept lattices. *Inf. Sci.* **2015**, *294*, 41–56.

[12] Cornejo, M.E.; Medina, J.; Ramírez-Poussa, E. Characterizing reducts in multi-adjoint concept lattices. *Inf. Sci.* **2018**, *422*, 364–376.

Reducción de ecuaciones de relaciones difusas mediante retículos de conceptos

David Lobo Palacios
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
david.lobo@uca.es

Víctor López Marchante
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
victor.lopez@uca.es

Jesús Medina Moreno
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
jesus.medina@uca.es

***Resumen*—Este trabajo presenta una primera aproximación a la reducción de ecuaciones en el ámbito de las ecuaciones de relaciones difusas. A raíz de la conexión existente entre las ecuaciones de relaciones difusas y el análisis formal de conceptos, se utilizan procedimientos de reducción de atributos para eliminar subecuaciones redundantes lo que da lugar a una simplificación del proceso de cálculo del conjunto completo de soluciones.**

***Palabras clave*—ecuación de relaciones difusas, retículos de conceptos, reducción de atributos, información redundante.**

I. Introducción

Las ecuaciones de relaciones difusas (FRE, del inglés fuzzy relations equations) introducidas en [1], son una de las principales herramientas utilizadas para distinguir las posibles relaciones existentes en una base de datos. La necesidad de establecer tales relaciones surge en numerosos problemas del mundo real, por lo que las FRE tienen aplicación en campos tan variados como la toma de decisiones, el procesamiento de imágenes, el control difuso y la optimización.

La resolución de las FRE adquiere por tanto un papel relevante, y diferentes autores han trabajado desde la introducción de las FRE hasta nuestros días en caracterizar su resolubilidad y en la obtención de procedimientos que permitan el cálculo de su conjunto de soluciones. En este trabajo, prestaremos especial atención al estudio desarrollado en [2], [3], en el que los autores asocian un retículo de conceptos a cada FRE, que da lugar a una caracterización de su conjunto de soluciones. Además, todo este desarrollo se realiza sobre un ambiente multiadjunto, lo que dota a este enfoque de una gran flexibilidad por la generalidad de las FRE en las que tiene aplicación.

Sin embargo, el procedimiento presentado en [2], [3] tiene un coste computacional elevado cuando el número de ecuaciones o variables es alto, dando lugar a un proceso de resolución cuya complejidad es inabarcable para los sistemas de cómputo habituales. Por este motivo, surge la necesidad de desarrollar estrategias que permitan reducir el proceso de resolución de FRE. En este trabajo, presentamos un primer acercamiento a la eliminación de la información redundante en una FRE, de forma similar a la simplificación que realiza el método de Gauss en la resolución de un sistema de ecuaciones lineales.

Habida cuenta de la relación existente entre FRE y FCA (Formal Concept Analysis), el presente estudio se centra en eliminar las ecuaciones asociadas a aquellos atributos que contienen información redundante en el retículo de conceptos asociado a la FRE. Esto dará lugar a un algoritmo que permite realizar una reducción del número de ecuaciones presentes en una FRE.

Partially supported by the 2014–2020 ERDF Operational Programme in collaboration with the State Research Agency (AEI) in project PID2019-108991GB-I00, and with the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia in project FEDER-UCA18-108612, and by the European Cooperation in Science & Technology (COST) Action CA17124.

II. Preliminares

Una FRE es una igualdad donde uno de sus miembros viene dado por una composición de relaciones difusas. Comencemos definiendo qué se entiende por operador de composición multiadjunto entre relaciones difusas.

Definición 1. Sean U, V, W conjuntos, $(P_1, \preceq_1)$, $(P_2, \preceq_2)$ y $(P_3, \preceq_3)$ conjuntos parcialmente ordenados, $\{(\&_i, \swarrow^i, \nwarrow_i) \mid i \in \{1, \dots, n\}\}$ un conjunto de triples adjuntos y $\sigma\colon V \to \{1, \dots, n\}$ una aplicación. Si P_3 es un semirretículo superior, el operador de sup-composición $\odot\colon P_1^{U\times V} \times P_2^{V\times W} \to P_3^{U\times W}$ se define como

$$R \odot_\sigma S(u, w) = \bigvee \{R(u, v) \,\&_{\sigma(v)}\, S(v, w) \mid v \in V\}$$

A partir de este operador es posible definir dos posibles FRE multiadjuntas.

Definición 2. Una FRE multiadjunta con sup-composición es una igualdad de la forma

$$R \odot_\sigma X = T$$

o de la forma

$$X \odot_\sigma S = T$$

donde, en ambos casos, X es una relación difusa desconocida.

El proceso de resolución de FRE desarrollado en [2] asocia un contexto multiadjunto a cada FRE multiadjunta y, por ende, un retículos de conceptos orientado a propiedades [4].

Definición 3. Sea $R \odot_\sigma X = T$ una FRE multiadjunta, donde $R \in P_1^{U\times V}$, $X \in P_2^{V\times W}$ y $T \in P_3^{U\times W}$. Su contexto

multiadjunto asociado es la tupla (U, V, R, σ) y su retículo de conceptos asociado será denotado por $\mathcal{M}_{\pi N}(U, V, R, \sigma)$.

Varios autores han tratado el problema de eliminar información redundante en una base de datos relacional a partir del estudio de un retículo de conceptos asociado. El método desarrollado en [5]–[7] utiliza conjuntos consistentes y reductos, subconjuntos de atributos que dan lugar a un retículo de conceptos isomorfo al original.

Definición 4. Dado un contexto multiadjunto (A, B, R, σ), diremos que $Y \subseteq A$ es un subconjunto consistente de atributos si

$$\mathcal{M}_{\pi N}(Y, B, R_Y, \sigma_{Y\times B}) \cong \mathcal{M}_{\pi N}(A, B, R, \sigma)$$

Si ademas, para todo $a \in Y$

$$\mathcal{M}_{\pi N}(Y\backslash\{a\}, B, R_{Y\backslash\{a\}}, \sigma_{Y\times B}) \ncong \mathcal{M}_{\pi N}(A, B, R, \sigma)$$

diremos que Y es un reducto.

III. Reducción de ecuaciones

En esta sección se mostrará un ejemplo de cómo aplicar los resultados de reducción de atributos para eliminar subecuaciones de una FRE.

Ejemplo 5. Considérese la FRE

$$R \odot_\sigma X = T$$

donde

$$R = \begin{pmatrix} 0{,}25 & 0{,}5 & 0 & 0{,}5 \\ 0 & 0{,}25 & 0{,}25 & 1 \\ 0{,}25 & 0{,}5 & 0{,}375 & 0 \\ 0{,}25 & 0{,}5 & 0 & 0{,}5 \\ 0{,}25 & 0{,}5 & 0{,}125 & 0{,}5 \end{pmatrix}, \quad T = \begin{pmatrix} 0 \\ 0{,}25 \\ 0{,}125 \\ 0 \\ 0{,}125 \end{pmatrix}$$

y cuyas soluciones son

$$X_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0{,}875 \\ 0 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0{,}75 \\ 0 \end{pmatrix}, \quad X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0{,}625 \\ 0 \end{pmatrix}$$

Utilizando los procedimientos de reducción de atributos adaptados al marco multiadjunto orientado a propiedades mostrados en [6], se obtiene que u_2 y u_3 son atributos absolutamente necesarios, u_1 y u_4 son atributos relativamente necesarios y u_5 es un atributo innecesario. Por ello, $Y_1 = \{u_1, u_2, u_3\}$ e $Y_2 = \{u_2, u_3, u_4\}$ son los únicos dos reductos admitidos por el contexto y contienen toda la información necesaria para la construcción del retículo de conceptos asociado.

La forma natural de reducir la FRE es entonces considerar únicamente las ecuaciones asociadas a los atributos presentes en Y_1 o únicamente las ecuaciones asociadas a los atributos presentes en Y_2.

Por ejemplo, considerando Y_1, la FRE reducida obtenida es

$$R_{Y_1} \odot_\sigma X = T_{Y_1}$$

donde

$$R_{Y_1} = \begin{pmatrix} 0{,}25 & 0{,}5 & 0 & 0{,}5 \\ 0 & 0{,}25 & 0{,}25 & 1 \\ 0{,}25 & 0{,}5 & 0{,}375 & 0 \end{pmatrix}, \quad T_{Y_1} = \begin{pmatrix} 0 \\ 0{,}25 \\ 0{,}125 \end{pmatrix}$$

El cálculo de su conjunto de soluciones lleva a la conclusión de que el conjunto de soluciones es invariante y sigue estando constituido por $X_1\, X_2\, X_3$ y X_4.

Que el conjunto de soluciones permanezca constante a pesar de la eliminación de ciertas ecuaciones no es algo circunstancial, tal y como muestra el siguiente resultado.

Teorema 6. *Sea $R \odot_\sigma X = T$ una FRE multiadjunta resoluble e $Y \subseteq U$ un conjunto consistente de su contexto asociado (U, V, R, σ). Si la FRE $R_Y \odot_\sigma X = T_Y$ se obtiene de eliminar ecuaciones que se correspondan con atributos que no estén presentes en Y, se tiene que $\overline{X}$ es solución de $R_Y \odot_\sigma X = T_Y$ si y solo si es solución de $R \odot_\sigma X = T$.*

IV. Conclusiones

La relación existente entre las FRE y los retículos de conceptos permite aplicar las estrategias de reducción de atributos en el marco de los retículos de conceptos multiadjuntos orientados a propiedades para reducir la FRE original, eliminando la información redundante y manteniendo la totalidad de la información presente en la FRE original. En consecuencia, se produce una simplificación del proceso de cálculo del conjunto de soluciones.

Referencias

[1] E. Sanchez, "Resolution of composite fuzzy relation equations," *Information and Control*, vol. 30, no. 1, pp. 38–48, 1976.

[2] J. C. Díaz-Moreno and J. Medina, "Multi-adjoint relation equations: Definition, properties and solutions using concept lattices," *Information Sciences*, vol. 253, pp. 100–109, 2013.

[3] J. C. Díaz-Moreno and J. Medina, "Using concept lattice theory to obtain the set of solutions of multi-adjoint relation equations," *Information Sciences*, vol. 266, no. 0, pp. 218–225, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025514000127

[4] J. Medina, "Multi-adjoint property-oriented and object-oriented concept lattices," *Information Sciences*, vol. 190, pp. 95–106, 2012.

[5] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa, "Attribute and size reduction mechanisms in multi-adjoint concept lattices," *Journal of Computational and Applied Mathematics*, vol. 318, pp. 388 – 402, 2017, computational and Mathematical Methods in Science and Engineering CMMSE-2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377042716303314

[6] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa, "Attribute reduction in multi-adjoint concept lattices," *Information Sciences*, vol. 294, pp. 41 – 56, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025514009451

[7] J. Medina and E. Ramírez-Poussa, "Towards attribute reduction in object-oriented concept lattices," in *IEEE Intl. Conf. on Fuzzy Systems (FUZZ-IEEE 2017)*, 2017, paper F-0493.

Analizando el impacto de aplicar congruencias locales en contextos formales

Roberto G. Aragón*
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
roberto.aragon@uca.es

Jesús Medina
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
jesus.medina@uca.es

Eloísa Ramírez-Poussa
Departamento de Matemáticas
Universidad de Cádiz
Cádiz, España
eloisa.ramirez@uca.es

Resumen*—Este trabajo hace referencia a la investigación recientemente publicada en *R.G. Aragón, J. Medina, and E. Ramírez-Poussa. Impact of local congruences in variable selection from datasets. Journal of Computational and Applied Mathematics, 404(113416), 2022.

***Palabras clave*—Análisis de conceptos formales, congruencia local, reducción de atributos**

I. Introducción

Una de las líneas de investigación más destacadas en el Ánalisis de Conceptos Formales (FCA, de sus siglas en inglés) es la selección de las variables más representativas de un conjunto de datos, lo que se denomina reducción de atributos [1], [12]–[14], [16]. La eliminación de los datos redundantes no modifica la información contenida en el conjunto de datos, pero sí tiene un gran impacto a la hora de poder extraer la información, pues facilita la obtención eficaz de información fundamental [7], [8], [11], [15].

Existen diversos mecanismos para reducir contextos formales tanto en un marco de trabajo clásico como difuso. Cuando se reduce el número de atributos de un contexto, se induce una relación de equivalencia sobre el retículo de conceptos, cuyas clases de equivalencia no tienen una estructura algebraica cerrada [9], [10]. Recientemente, en [2], [5], se introdujo la noción de congruencia local como una relación de equivalencia definida sobre retículos cuyas clases son subretículos convexos del retículo original. Las congruencias locales pueden ser utilizadas como un mecanismo complementario a los mecanismos de reducción de atributos, para dotar a las clases de equivalencia de una estructura algebraica cerrada.

Cuando la congruencia local se aplica sobre el retículo de conceptos reducido en lugar del retículo original, puede ocurrir que se unan diferentes clases de equivalencia inducidas, lo que puede tener un notable impacto tanto en el retículo de conceptos como en el contexto reducido [3]. En [4], se presentó un estudio sobre las clases de equivalencia inducidas por una reducción de atributos que se alteran tras la aplicación de una congruencia local. La agrupación de conceptos por una congruencia local se puede ver como una eliminación de conceptos del retículo sobre el que se aplica, este hecho nos condujo a estudiar los efectos producidos por estas agrupaciones en el contexto considerado.

* Responsable

Parcialmente financiado por los fondos FEDER 2014-2020 en colaboración con la Agencia Estatal de Envestigación (AEI) a través del proyecto con referencia PID2019-108991GB-I00 y el proyecto financiado por la Junta de Andalucía y los fondos FEDER con referencia FEDER-UCA18-108612, y por la European Cooperation in Science & Technology (COST) Action CA17124.

II. Eliminando conceptos de un retículo de conceptos

En este trabajo se analiza cómo la aplicación de la menor congruencia local, que contiene la relación de equivalencia inducida por una reducción de un contexto (A, B, R) por un subconjunto de atributos $D \subseteq A$, altera el retículo de conceptos asociado al contexto reducido $(D, B, R_{|D\times B})$ y, por lo tanto, modificando también dicho contexto. Este análisis se lleva a cabo en el artículo [6], donde se distinguen tres situaciones diferentes dependiendo del elemento (agrupado por la congruencia local) que se quiere eliminar. Para realizar dicha eliminación se debe distinguir cuándo el elemento a eliminar es supremo irreducible, ínfimo irreducible o un elemento cualquiera que no sea ni supremo ni ínfimo irreducible.

En el FCA la supresión de un concepto supremo irreducible se puede llevar a cabo mediante la eliminación de los objetos del conjunto de objetos que lo generan. Aunque esta eliminación de objetos puede hacer que otros conceptos dependientes del concepto eliminado también desaparezcan. Sin embargo, cuando extraemos únicamente un elemento supremo irreducible de un retículo completo, en general, no se altera la estructura de retículo.

Lema 1. *Dado un retículo completo $(L, \wedge, \vee)$ que satisface la condición de cadena ascendente, y un elemento supremo irreducible $p \in L$, entonces $(L\backslash\{p\}, \wedge_p, \vee_p)$ es un subretículo de $(L, \wedge, \vee)$, donde $\wedge_p, \vee_p$ son la restricción de $\wedge, \vee$ a $L\backslash\{p\}$, y en particular, es un retículo completo.*

De manera dual, se puede obtener un resultado similar si el elemento a eliminar es ínfimo irreducible.

En el caso de los retículos de conceptos, el retículo de conceptos resultante tras la eliminación de los objetos generadores del concepto supremo irreducible no coincide con el conjunto cociente de la menor congruencia que contiene a la relación de equivalencia inducida. Por tanto, en [6], se describe un procedimiento para modificar el contexto sobre cuyo retículo

de conceptos asociado se aplica la congruencia local, para así obtener un nuevo retículo de conceptos isomorfo al retículo completo obtenido tras eliminar un elemento supremo irreducible. Consecuentemente, se está eliminando un concepto supremo irreducible pero preservando el resto de conceptos.

Este procedimiento se presenta en el Algoritmo 1 de [6] como un pseudocódigo. Lo que se obtiene al aplicar este mecanismo es un nuevo contexto cuyo retículo de conceptos es isomorfo al retículo de conceptos original tras eliminarle el elemento supremo irreducible.

Teorema 1. *Dado un contexto (A, B, R) y un concepto supremo irreducible $C_j \in \mathcal{C}(A, B, R)$, se satisface que:*

$$(\mathcal{C}(A, B, R)\backslash\{C_j\}, \leq_j) \cong (\mathcal{C}(A, B^*, R^*), \leq^*),$$

donde (A, B^*, R^*) es el contexto obtenido por el Algoritmo 1, y $\leq_j$ es el orden definido de la restricción de la relación $\leq$ a $\mathcal{C}(A, B, R)\backslash\{C_j\}$.

Dualmente se obtiene un procedimiento para la eliminación de conceptos ínfimo irreducibles (Algoritmo 2). Por último, para el caso donde el concepto a eliminar no es ni supremo ni ínfimo irreducible, es necesario el uso de la compleción de Dedekind-MacNeille para preservar la estructura de retículo completo. Como consecuencia, los dos algoritmos pueden ser aplicados de manera secuencial y, además, los resultados obtenidos de isomorfía de retículos de conceptos pueden ser combinados en un único resultado.

Teorema 2. *Dado un contexto (A, B, R) y un concepto $C \in \mathcal{C}(A, B, R)$, se satisface que:*

$$(DM(\mathcal{C}(A, B, R)\backslash\{C\}), \subseteq) \cong (\mathcal{C}(A^*, B^*, R^*), \leq^*),$$

donde (A^*, B^*, R^*) es el contexto obtenido por el Algoritmo 1 o el Algoritmo 2 o el original en el caso de que C no sea ínfimo ni supremo irreducible, y $\subseteq$ es el orden definido por la compleción de Dedekind-MacNeille.

Por lo tanto, se puede aplicar el Teorema 2 sucesivamente para cada clase de una congruencia local que contenga más de un concepto reducido, pudiéndose así caracterizar el impacto generado en el contexto por la aplicación de una congruencia local sobre el retículo de conceptos asociado.

III. Conclusiones

Todos los resultados mostrados en este trabajo tienen una importancia significativa a la hora de conocer detalladamente las modificaciones realizadas sobre el contexto considerado al aplicar una congruencia local para complementar un mecanismo de reducción de atributos en el FCA. Además, abre una nueva vía de estudio sobre la interpretación semántica de los cambios realizados en el contexto debido a la reducción del retículo de conceptos utilizando congruencias locales, ya que se puede llegar a modificar tanto el conjunto de atributos como el de objetos, eliminando elementos de dichos conjuntos o agregándolos para poder preservar la información contenida en otros conceptos.

Referencias

[1] C. Alcalde and A. Burusco. Study of the relevance of objects and attributes of L-fuzzy contexts using overlap indexes. In J. Medina, M. Ojeda-Aciego, J. L. Verdegay, D. A. Pelta, I. P. Cabrera, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*, pages 537–548, Cham, 2018. Springer International Publishing.

[2] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Reducing concept lattices by means of a weaker notion of congruence. *Fuzzy Sets and Systems*, 2020. https://doi.org/10.1016/j.fss.2020.09.013.

[3] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Impact of local congruences in attribute reduction. In M.-J. Lesot, S. Vieira, M. Z. Reformat, J. P. Carvalho, A. Wilbik, B. Bouchon-Meunier, and R. R. Yager, editors, *Communications in Computer and Information Science*, 1239:748–758, Cham, 2020. Springer International Publishing.

[4] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Identifying non-sublattice equivalence classes induced by an attribute reduction in fca. *Mathematics*, 9(5), 2021.

[5] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. A weakened notion of congruence to reduce concept lattices. In M. E. Cornejo, L. T. Kóczy, J. Medina-Moreno, and J. Moreno-García, editors, *Computational Intelligence and Mathematics for Tackling Complex Problems 2*, pages 139–145. Springer International Publishing, Cham, 2022.

[6] R. G. Aragón, J. Medina, and E. Ramírez-Poussa. Impact of local congruences in variable selection from datasets. *Journal of Computational and Applied Mathematics*, 404(113416), 2022.

[7] R. Bělohlávek, P. Cordero, M. Enciso, A. Mora, and V. Vychodil. Automated prover for attribute dependencies in data with grades. *International Journal of Approximate Reasoning*, 70:51–67, 2016.

[8] R. Bělohlávek and V. Vychodil. Attribute dependencies for data with grades I. *International Journal of General Systems*, 45(7-8):864–888, 2016.

[9] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. A computational procedure for variable selection preserving different initial conditions. *International Journal of Computer Mathematics*, 97(1-2):387–404, 2020.

[10] M. J. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślęzak. Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis. *Fuzzy Sets and Systems*, 391:117–138, 2020.

[11] P. Cordero, M. Enciso, A. Mora, and V. Vychodil. Parameterized simplification logic I: Reasoning with implications and classes of closure operators. *International Journal of General Systems*, 49(7):724–746, 2020.

[12] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294:41–56, 2015.

[13] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute and size reduction mechanisms in multi-adjoint concept lattices. *Journal of Computational and Applied Mathematics*, 318:388–402, 2017.

[14] M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Characterizing reducts in multi-adjoint concept lattices. *Information Sciences*, 422:364 – 376, 2018.

[15] D. Dubois, J. Medina, H. Prade, and E. Ramírez-Poussa. Disjunctive attribute dependencies in formal concept analysis under the epistemic view of formal contexts. *Information Sciences*, 561:31–51, 2021.

[16] J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.

Ecuaciones bipolares de relaciones difusas en retículos residuados con negación involutiva

M. Eugenia Cornejo
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
mariaeugenia.cornejo@uca.es

David Lobo
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
david.lobo@uca.es

Jesús Medina
Departamento de matemáticas
Universidad de Cádiz
Puerto Real, España
jesus.medina@uca.es

Bernard De Baets
KERMIT
Ghent University. Belgium
Gante, Bélgica
bernard.debaets@ugent.be

***Resumen*—En este trabajo se presenta un resumen sobre el estudio publicado recientemente en "M. Eugenia Cornejo, David Lobo, Jesús Medina, Bernard De Baets. Bipolar equations on complete distributive symmetric residuated lattices: The case of a join-irreducible right-hand side, Fuzzy Sets and Systems, 442: 92-108, 2022."**

***Palabras clave*—Ecuación bipolar de relaciones difusas, retículo residuado simétrico, operador de negación, elemento supremo irreducible.**

I. Introducción

La información bipolar está presente de forma natural en el razonamiento humano, desde el análisis de pros y contras en la toma de decisiones hasta el procesamiento de contradicciones. Así, todo sistema inteligente que pretenda imitar el razonamiento humano debería ser capaz de tratar matemáticamente la presencia de información bipolar [8].

Dentro de este contexto de bipolaridad, las ecuaciones bipolares de relaciones difusas (BFRE, del inglés Bipolar Fuzzy Relation Equations) se introdujeron en [9] como una generalización de las ecuaciones de relaciones difusas (FRE, del inglés Fuzzy Relation Equations) [7], [13], donde las incógnitas aparecen junto con su negación lógica simultáneamente. La mayoría, si no todas, las publicaciones actuales sobre BFRE están restringidas al intervalo unidad y son del tipo max-$*$, siendo $*$ una de las tres t-normas básicas: la t-norma mínimo [9]–[11], la t-norma producto [1], [4] y la t-norma Łukasiewicz [12], [14]. Además, los únicos trabajos en los que se asume un operador de negación diferente a la negación estándar son [2], [3], donde se considera la negación residuada de la t-norma producto.

En este trabajo presentamos los resultados publicados en [5], donde se realiza un estudio sobre la resolución de sup-$*$ ecuaciones bipolares sobre un retículo residuado completo distributivo dotado de una negación involutiva. Nótese que este tipo de ecuaciones son la base de las BFRE [6].

En primer lugar, se describe analíticamente el conjunto de soluciones de una sup-$*$ ecuación bipolar, de donde se deduce una condición necesaria para la resolubilidad de dicha ecuación, además de una sencilla caracterización. A continuación, se analiza la estructura algebraica del conjunto de soluciones, proporcionando condiciones para la existencia de una solución máxima (respectivamente mínima) o soluciones maximales (respectivamente minimales). Todos estos resultados generalizan en gran medida aquellos presentes en la literatura actual de BFRE.

Parcialmente financiado por los fondos FEDER 2014-2020 en colaboración con la Agencia Estatal de Investigación (AEI) a través del proyecto con referencia PID2019-108991GB-I00 y el proyecto financiado por la Junta de Andalucía y los fondos FEDER con referencia FEDER-UCA18-108612, y por la European Cooperation in Science & Technology (COST) Action CA17124.

II. Sup-$*$ ecuaciones bipolares

Una sup-$*$ ecuación bipolar es una igualdad en la que uno de los miembros viene dado como el supremo de pares operados con $*$. En particular, las sup-$*$ ecuaciones consideradas en [5] se definen sobre un retículo residuado completo simétrico y distributivo (CDSRL, del inglés Complete Distributive Symmetric Residuated Lattice), cuya definición formal se recuerda a continuación.

Definición 1. Un *retículo residuado completo simétrico y distributivo* es una estructura algebraica $(L, \preceq, *, \to, \neg)$ donde:

(I) $(L, \preceq)$ es un conjunto parcialmente ordenado con elemento máximo $\top$ y elemento mínimo $\bot$.
(II) El supremo y el ínfimo de S existe, para todo $S \subseteq L$.
(III) $(L, *, \top)$ es un monoide conmutativo.
(IV) $(*, \to)$ es un par residuado, es decir, $x * y \preceq z$ si y sólo si $x \preceq y \to z$, para todo $x, y, z \in L$.
(V) $\neg : L \to L$ es un operador unario satisfaciendo las siguientes propiedades:
  a) Si $x \preceq y$, $\neg x \preceq \neg y$.
  b) $\neg\neg x = x$, para todo $x \in L$;
  c) $\neg\top = \bot$ y $\neg\bot = \top$.
(VI) Para todo $x, y, z \in L$, se satisface la siguiente igualdad:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Definición 2. Sean $(L, \preceq, *, \to, \neg)$ un CDSRL y $a_j^+, a_j^-, b \in L$ con $j \in \{1, \ldots, m\}$. Una sup-$*$ ecuación bipolar es una expresión de la forma:

$$\bigvee_{j \in \{1,\ldots,m\}} (a_j^+ * x_j) \vee (a_j^- * \neg x_j) = b \tag{1}$$

donde el valor de las variables $x_1, \ldots, x_m \in L$ se desconoce.

El estudio desarrollado en [5] requiere esencialmente dos condiciones:

(I) El término independiente b es un elemento supremo irreducible del retículo $(L, \preceq)$.

(II) Las componentes del operador $*$ son morfismos del retículo $(L, \preceq)$.

Por un lado, recordemos que un elemento $x \in L$ es *supremo irreducible* si $x = a \vee b$ implica $x = a$ o $x = b$, para todo $a, b \in L$. Por otro lado, un morfismo del retículo $(L, \preceq)$ es una aplicación unaria $f\colon L \to L$ que conserva supremo e ínfimo de subconjuntos no vacíos, es decir, $f(\sup A) = \sup(f(A))$ y $f(\inf A) = \inf(f(A))$, para todo $A \in \mathcal{P}(L) \setminus \{\varnothing\}$.

Para enunciar los principales resultados de [5] haremos uso de los conjuntos

$$S^+ = \{s^{j+} \mid b \preceq a_j^+ \text{ y } a_j^+ \rightsquigarrow b \preceq a_j^+ \to b\}$$

$$S^- = \{s^{j-} \mid b \preceq a_j^- \text{ y } a_j^- \rightsquigarrow b \preceq a_j^- \to b\}$$

donde $s^{j+} = (s_1^{j+}, \ldots, s_m^{j+})$ y $s^{j-} = (s_1^{j-}, \ldots, s_m^{j-})$ vienen dadas por:

$$s_k^{j+} = \begin{cases} a_j^+ \rightsquigarrow b & \text{si} \quad k = j \\ \bot & \text{en otro caso} \end{cases}$$

$$s_k^{j-} = \begin{cases} a_j^- \rightsquigarrow b & \text{si} \quad k = j \\ \bot & \text{en otro caso} \end{cases}$$

Así mismo, consideraremos las tuplas $g^+ = (g_1^+, \ldots, g_m^+)$ y $g^- = (g_1^-, \ldots, g_m^-)$ definidas, para cada $j \in \{1, \ldots, m\}$, como $g_j^+ = a_j^+ \to b$ y $g_j^- = a_j^- \to b$, respectivamente.

Bajo las hipótesis (I) y (II), el siguiente teorema describe el conjunto de soluciones de una sup-$*$ ecuación bipolar.

Teorema 3. *Si la sup-$*$ ecuación bipolar* (1) *es resoluble, su conjunto de soluciones es*

$$D = \Big(\bigcup_{s \in S^+} [s \vee \neg g^-, g^+] \Big) \cup \Big(\bigcup_{s \in S^-} [\neg g^-, g^+ \wedge \neg s] \Big).$$

Del Teorema 3 se escinde la siguiente condición necesaria.

Corolario 4. Si la sup-$*$ ecuación bipolar (1) es resoluble, entonces $\neg g^- \preceq g^+$.

Atendiendo a la forma del conjunto de soluciones mostrada en el Teorema 3, que viene dado como la intersección los conjuntos $\bigcup_{s \in S^+} [s \vee \neg g^-, g^+]$ y $\bigcup_{s \in S^-} [\neg g^-, g^+ \wedge \neg s]$, se obtiene la siguiente caracterización de la resolubilidad de una sup-$*$ ecuación bipolar.

Teorema 5. *La sup-$*$ ecuación bipolar* (1) *es resoluble si y solo si g^+ o $\neg g^-$ es solución.*

Para finalizar, con objeto de entender la estructura algebraica del conjunto de soluciones de una sup-$*$ ecuación bipolar, el siguiente resultado muestra una caracterización para la existencia de una solución máxima (respectivamente mínima) o soluciones maximales (respectivamente minimales).

Teorema 6. *Sea la sup-$*$ ecuación bipolar* (1) *resoluble:*

(1) Si g^+ es solución de (1)*, entonces g^+ es su solución máxima. En otro caso, el conjunto de soluciones maximales de* (1) *es* $\{g^+ \wedge \neg s \mid s \in S^-, \neg g^- \preceq g^+ \wedge \neg s\}$.

(2) Si $\neg g^-$ es solución de (1)*, entonces $\neg g^-$ es su solución mínima. En otro caso, el conjunto de soluciones minimales de* (1) *es* $\{s \vee \neg g^- \mid s \in S^+, s \vee \neg g^- \preceq g^+\}$.

III. Conclusiones

En el trabajo se presentan los principales resultados publicados en [5], en los que se realiza un estudio en profundidad de la resolución de sup-$*$ ecuaciones bipolares sobre un CDSRL. Como en el caso de las sup-$*$ ecuaciones (no bipolares), se pueden obtener resultados duales para inf-$*$ ecuaciones bipolares definidas en un CDSRL.

Referencias

[1] M. E. Cornejo, D. Lobo, and J. Medina, "Bipolar fuzzy relation equations based on the product t-norm," in *Proceedings of the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017.

[2] M. E. Cornejo, D. Lobo, and J. Medina, "Bipolar fuzzy relation equations systems based on the product t-norm," *Mathematical Methods in the Applied Sciences*, vol. 42, no. 17, pp. 5779–5793, 2019.

[3] M. E. Cornejo, D. Lobo, and J. Medina, "On the solvability of bipolar max-product fuzzy relation equations with the product negation," *Journal of Computational and Applied Mathematics*, vol. 354, pp. 520–532, 2019.

[4] M. E. Cornejo, D. Lobo, and J. Medina, "On the solvability of bipolar max-product fuzzy relation equations with the standard negation," *Fuzzy Sets and Systems*, vol. 410, pp. 1–18, 2021.

[5] M. E. Cornejo, D. Lobo, J. Medina and B. De Baets, "Bipolar equations on complete distributive symmetric residuated lattices: The case of a join-irreducible right-hand side," *Fuzzy Sets and Systems*, vol. 442, pp. 92–108, 2022.

[6] B. De Baets, "Analytical solution methods for fuzzy relational equations," in *Fundamentals of Fuzzy Sets*, The Handbooks of Fuzzy Sets Series, chapter 6, vol. 1, pp. 291–340, 2000.

[7] A. Di Nola, S. Sessa, W. Pedrycz and E. Sanchez, "Fuzzy Relation Equations and their Applications to Knowledge Engineering," Kluwer Academic Publishers, 1989.

[8] D. Dubois and H. Prade, "An introduction to bipolar representations of information and preference," *International Journal of Intelligent Systems*, vol. 23, vo. 8, pp. 866–877, 2008.

[9] S. Freson, B. De Baets, and H. De Meyer, "Linear optimization with bipolar max–min constraints," *Information Sciences*, vol. 234, pp. 3–15, 2013.

[10] P. Li and Q. Jin. On the resolution of bipolar max-min equations. *Kybernetika*, 52(4):514–530, 2016.

[11] C.-C. Liu, Y.-Y. Lur, and Y.-K. Wu. Some properties of bipolar max-min fuzzy relational equations. In *Proceedings of the International MultiConference of Engineers and Computer Scientists 2015*, volume 2222 of *Lecture Notes in Engineering and Computer Science*, pages 955–969, 2015.

[12] C.-C. Liu, Y.-Y. Lur, and Y.-K. Wu. Linear optimization of bipolar fuzzy relational equations with max-łukasiewicz composition. *Information Sciences*, 360:149–162, 2016.

[13] E. Sanchez, "Resolution of composite fuzzy relation equations," *Information and Control*, vol. 30, no. 1, pp. 38–48, 1976.

[14] X.-P. Yang, "Resolution of bipolar fuzzy relation equations with max-Łukasiewicz composition," *Fuzzy Sets and Systems*, vol. 397, pp. 41–60, 2020.

Levenberg-Marquardt Damping Factor Fuzzy Updating Strategy: An Ilustratative Example

Lucas C. S. Jardim, Diego C. Knupp, Antônio J. Silva Neto
Instituto Politécnico
Universidade do Estado do Rio de Janeiro
Nova Friburgo, Brazil
{ljardim, diegoknupp, ajsneto}@iprj.uerj.br

Carlos Cruz Corona, David A. Pelta
Dept. de Ciencias de la Computación
Universidad de Granada
Granada, Spain
{carloscruz, dpelta}@decsai.ugr.es

***Abstract*—Fuzzy logic is used to update the damping factor of the classical Levenberg-Marquardt (LM) least-squares minimization method. Canonically, the LM algorithm uses the damping factor to improve convergence, nevertheless, the most common iterative updating technique used to modify this damping parameter during the search only takes in consideration if the objective function have improved or not. The fuzzy updating of such parameter have recently been proposed and takes in consideration the magnitude of improvement of the objective function. In this work a test inverse problem is used to illustrate the behavior of such new approach and the fuzzy system obtained. Results are compared against the classical LM and measured in numbers of iterations until stopping criterion is reach, which indicates a more efficient search for the Fuzzy LM approach.**

***Index Terms*—Levenberg-Marquardt, Inverse Problem, Fuzzy Logic**

I. Introduction

Consider an arbitrary physical problem that can be modeled as an exponential decay. The behavior in time of such physical quantity, namely F, can be written as

$$F(t) = P_1 \exp\left(-P_2\, t\right) \tag{1}$$

where P_1 and P_2 are parameters of the model. If such parameters are known, the physical quantity F is determined for every instant t - this is the so called Direct Problem.

On the other hand, if P_1 and/or P_2 are unknown, but experimental data of F are available for different instants, a inverse problem can be formulated and solved in order to obtain estimates of these unknown parameters.

Considering the maximum likelihood approach, the inverse problem can be implicitly formulated, yielding a global optimization problem, where the squared residuals between predictions of the given model and experimental measurements must be minimized, i.e., the minimum of the objective function

$$S(\mathbf{P}) = \sum_{i=1}^{N_d} \left(Y_i - F(t_i, \mathbf{P})\right)^2 \tag{2}$$

This work has been funded by the projects PID2020-112754GB-I0, MCIN/AEI /10.13039/501100011033 and FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto (BTIC-640-UGR20). Also, this work have been partially funded by CAPES/Brazil (Finance Code 001), (Grant No. 88887.469279/2019-00), CNPq/Brazil, and FAPERJ/Brazil

must be obtained, where N_d is the total number of experimental data, $\mathbf{Y}$ is the vector of such data obtained for different instants t_i.

In this work, the arbitrary physical quantity experimental data is generated by solving the direct problem and adding simulated noise, i.e., a random number drawn from a normal distribution with 0 mean and σ_{exp} standard deviation. These experimental data are generated for exact known values of $\mathbf{P}$, which are, later on, estimated in the solution of the inverse problem.

II. Levenberg-Marquardt with Adaptive Fuzzy Damping Factor

a) Classical Levenberg-Marquardt: The Levenberg-Marquardt (LM) algorithm is used to minimize the objective function 2, and its algorithm is constructed around a iterative step of the form [1], [2]

$$\mathbf{P}_{\text{new}} = \mathbf{P}_{\text{old}} - \left[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}\right]^{-1}\mathbf{J}\mathbf{r} \tag{3}$$

where the current vector of parameters $\mathbf{P}_{\text{old}}$ is updated by adding the incremental modification, generating the new set of parameters $\mathbf{P}_{\text{new}}$, μ is the damping factor, $\mathbf{r}$ is the vector of residuals between experimental and calculated values, $\mathbf{I}$ the identity matrix with order dim($\mathbf{P}$) and $\mathbf{J}$ is the Jacobian matrix written as $\mathbf{J} = \partial F(t_i)/\partial P_j$, with $i = 1, 2, ..., N_d$ and $j = 1, 2$.

The damping factor μ plays an important role in the algorithm convergence. Classically, this parameter is updated multiplying or dividing it by a pre-defined value whenever the objective function value gets higher or lower than the previous iteration, respectively.

b) Fuzzy Levenberg-Marquardt: Recently, Sajedi et. al [3] proposed a modification to the classical LM method by updating the damping factor μ with Fuzzy Logic. The new proposed algorithm uses a fuzzy system with input as the "Percent Error", an relative error obtained for every new iteration that can be written as

$$\text{PE} = 100 \times \frac{S_{\text{new}} - S_{\text{old}}}{S_{\text{old}}} \tag{4}$$

where S_{new} and S_{old} are the new and current objective function values. In other words, if PE < 0, S is improving; if PE ≈ 0, S is converging; if PE > 0, S is getting worse.

Degrees of membership are then associated to the input PE, as presented in Fig. 1, where NLR, NL, NM, NS, Z, PS, PM, PL, and PLR mean larger negative, large negative, medium negative, small negative, zero, small positive, medium positive, large positive, and larger positive, respectively.

The output of the fuzzy is the damping factor μ and its membership function is presented in Fig. 2. In such figure, the terms VS, S, SS, LM, M, UM, SL, L, and VL mean very small, small, slightly small, lesser medium, medium, more medium, slightly large, large, and very large, respectively.

Then, the fuzzy system is constructed with the following rules: If PE is NLR then μ is M. If PE is NL then μ is LM. If PE is NM then μ is SS. If PE is NS then μ is S. If PE is Z then μ is VS. If PE is PS then μ is UM. If PE is PM then μ is SL. If PE is PL then μ is L. If PE is PLR then μ is VL. The defuzzification method chosen is the centroid technique.

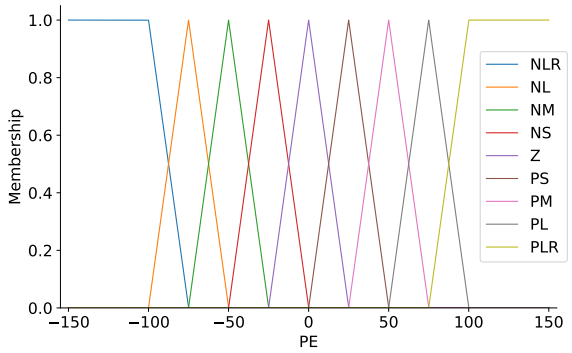


Fig. 1. Memberships of the fuzzy system input PE.

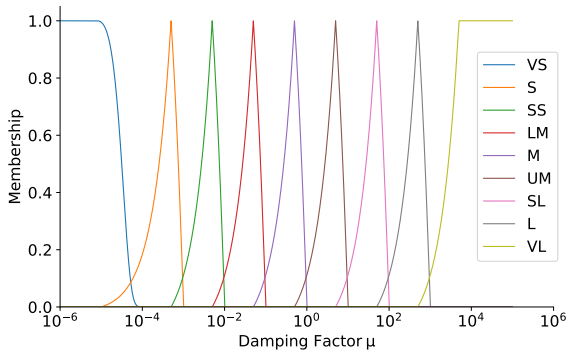


Fig. 2. Memberships of the fuzzy system output μ.

III. Results and Discussion

a) Visualization: To better visualize the Fuzzy System behavior, Fig. 3 shows the complete PE input space for its respective μ output. It is possible to see that μ have a step behavior with range of $10^{-5} \leq \mu \leq 10^{0}$ for $-150 \leq \text{PE} < 0$, $10^{0} < \mu \leq 10^{5}$ for $0 < \text{PE} \leq 0$ and the lowest value possible for PE = 0, which is desirable, meaning that convergence was reached.

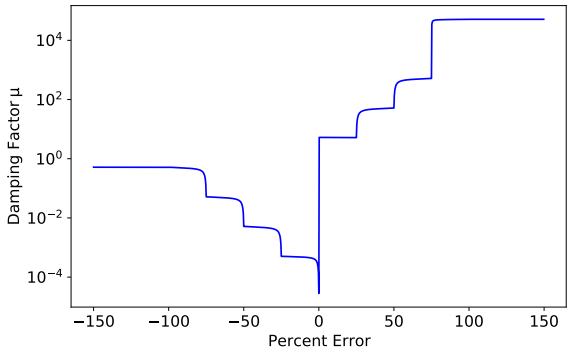


Fig. 3. Complete input vs output of the fuzzy system.

b) Computational Experiments: The "skfuzzy" package from Python 3.9 language is used to perform steps regarding the fuzzy aspects of the routine. The iterative process is stopped when the absolute value of PE $\leq 0.001\%$. Consider the exact values of $P_1 = 10.0$ and $P_2 = 2.5$. The experimental data was generated with $\sigma_{\text{exp}} = 0.1$ and acquired from t = 0.5 to t = 1.55 with steps of 0.05. In Tab. I, the results obtained for 10 executions of both LM and FLM are presented. In each execution, both methods starts at the same point, which is generated randomly around the exact solution. It is possible to see that the FLM was more efficient to converge than the classical LM, with exception of the #3 and #7 executions. Nevertheless, on average, LM needed 9.4 iterations to converge and LMF only 7.1. All the obtained parameters reached a reasonable estimate.

TABLE I
Comparison between Classical (LM) and Fuzzy (FLM) Levenberg-Marquardt methods

	LM			FLM		
Exec.	Iter.	P_1	P_2	Iter.	P_1	P_2
#1	10	10.446723	2.570472	6	10.446722	2.570472
#2	9	10.446767	2.570478	6	10.446722	2.570472
#3	10	10.446722	2.570472	12	10.446720	2.570471
#4	10	10.446722	2.570472	5	10.446706	2.570470
#5	9	10.446683	2.570466	6	10.446711	2.570470
#6	9	10.446744	2.570467	5	10.446656	2.570463
#7	10	10.446724	2.570472	12	10.446720	2.570472
#8	9	10.446760	2.570477	5	10.446724	2.570472
#9	10	10.446722	2.570472	8	10.446720	2.570472
#10	8	10.446747	2.570475	6	10.446710	2.570470
Avg	9.4	10.446665	2.570472	7.1	10.446711	2.570470

References

[1] K. Levenberg, "A method for the solution of certain non-linear problems in least squares." Quarterly of applied mathematics 2.2, 1944, pp. 164–168.

[2] D. W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," Journal of the society for Industrial and Applied Mathematics 11.2, 1963, pp. 431–441.

[3] R. Sajedi, J. Faraji and F. Kowsary, "A new damping strategy of Levenberg-Marquardt algorithm with a fuzzy method for inverse heat transfer problem parameter estimation.," International Communications in Heat and Mass Transfer, vol. 126, 2021, p. 105433.

Toma de Decisiones de Multitud guiada por Análisis de Opiniones

1er Cristina Zuheros
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Universidad de Granada, Granada, Spain
czuheros@ugr.es

2º Eugenio Martínez-Cámara
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Universidad de Granada, Granada, Spain
emcamara@decsai.ugr.es

3er Enrique Herrera-Viedma
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Universidad de Granada, Granada, Spain
viedma@decsai.ugr.es

4º Francisco Herrera
Instituto Andaluz Interuniversitario en Data Science and Computational Intelligence (DaSCI)
Universidad de Granada, Granada, Spain
herrera@decsai.ugr.es

***Resumen*—Las redes sociales son una gran fuente de información que las personas tienden a usar para tomar decisiones aprovechando la sabiduría de la multitud. Afirmamos que la sabiduría de la multitud mejora la calidad de los modelos de toma de decisiones e introducimos el concepto de toma de decisiones de multitud (CDM). Presentamos un resumen del trabajo aceptado [3], que propone el modelo CDM-SA como un modelo CDM capaz de procesar evaluaciones en lenguaje natural para captar la sabiduría de la multitud disponible en redes sociales. El modelo captura los matices de las opiniones de los usuarios a través del análisis de opiniones y trata la falta de información mediante una estrategia de representación escasa. Evaluamos el modelo sobre el conjunto de datos TripR-2020Large, que creamos y publicamos.**

***Index Terms*—Toma de Decisiones de Multitud, Análisis de Opiniones, Representación escasa, Redes Sociales**

I. Introducción

Los modelos de toma de decisiones (DM) evalúan un conjunto de alternativas a través de las evaluaciones de expertos. Coincidimos con la teoría de la sabiduría de la multitud, la cuál afirma que los grupos amplios de personas no expertas son más inteligentes que los grupos reducidos de expertos con amplios conocimientos [1]. Asimismo, manifestamos que las evaluaciones en lenguaje natural mejoran los modelos DM [2].

Este trabajo resume las características del modelo *Crowd Decision Making guided by Sentiment Analysis* (CDM-SA) presentado en [3]. Se trata de un modelo DM a gran escala que permite aprovechar la sabiduría de las multitudes disponible en redes sociales, siguiendo la metodología SA-MpMcDM [2]. El modelo procesa las evaluaciones de una cantidad masiva de usuarios utilizando modelos de análisis de opiniones de aprendizaje profundo que captan la mayoría de los matices de las evaluaciones. Asimismo, el modelo trata la posible falta información siguiendo una estrategia de representación escasa.

II. Preliminares

Esta sección se centra en las bases del modelo CDM-SA. La Sección II-A presenta la metodología SA-MpMcDM y la Sección II-B presenta la teoría de la sabiduría de la multitud.

A. Metodología SA-MpMpDM

La metodología SA-MpMcDM [2] es un modelo DM multi-personal y multi-criterio para la toma de decisiones más inteligente, capaz de procesar evaluaciones sin restricciones mediante un modelo de análisis de opiniones (SAM). Permite hacer evaluaciones mediante lenguaje natural e incluso con valores numéricos. Por tanto, es adecuada para inferir el conocimiento experto a partir de las reseñas de redes sociales.

B. Sabiduría de la multitud

La teoría de la sabiduría de la multitud afirma que los grandes grupos son más inteligentes que pocas personas de élite [1]. Los requisitos necesarios para que estos grupos tomen decisiones sabias es mantener la diversidad de sus miembros, la independencia de sus pensamientos, y la descentralización del grupo. Por consiguiente, las redes sociales son un entorno adecuado para captar la sabiduría de la multitud.

III. Toma de decisiones de multitud guiada por el análisis de opiniones

Definimos la toma de decisiones de multitud (CDM) como un modelo DM basado en la sabiduría de la multitud que integra evaluaciones en lenguaje natural de un gran número de personas [3]. Las redes sociales son entornos adecuados para CDM, ya que conectan a millones de personas que valoran libremente temas a través lenguaje natural. Presentamos el modelo CDM-SA para captar la sabiduría de la multitud disponible en las redes sociales. Su flujo de trabajo tiene cinco etapas como se muestra en la Fig. 1 y describimos a continuación:

1) Rastreo de redes sociales: Se descargan las reseñas de diversas alternativas $X = \{x_1, ..., x_n\}$ disponibles en una red social y se recopilan los usuarios que evalúan al menos una de ellas en un conjunto de expertos $E = \{e_1, ..., e_l\}$. La reseña proporcionada por e_k para x_i es un texto en lenguaje natural, t_i^k, que puede ser complementado con valoraciones numéricas.

Fig. 1. Flujo de trabajo del modelo CDM-SA.

2) Extracción de las opiniones: Se infieren las opiniones de los textos t_i^k mediante múltiples SAMs para obtener un resultado más preciso. Cada opinión está compuesta por: 1) un término de aspecto que indica si la palabra de entrada es una entidad de interés, 2) una categoría que indica el criterio asociado al término de aspecto, y 3) una polaridad que proporciona la distribución de probabilidad que representa un valor de opinión como $(softmax(x)_{negative}, softmax(x)_{neutral}, softmax(x)_{positive})$. Recogemos las categorías identificadas y los criterios que la red social permite evaluar numéricamente en un conjunto de criterios $C = \{c_1, ..., c_m\}$.

3) Representación de los matices de las opiniones: Se agregan las opiniones inferidas y se representan sus matices en matrices. Definimos el concepto de tripleta de opinión para representar cómo de negativa, neutra y positiva es una opinión, lo que permite captar sus matices. Primeramente, transformamos la polaridad de las opiniones en tripletas de opinión, por lo que la polaridad de una opinión se representa como $\{(negative, softmax(x)_{negative}), (neutro, softmax(x)_{neutro}), (positivo, softmax(x)_{positivo})\}$). A continuación, agregamos las opiniones mediante un sistema de votación de SAMs de forma que las opiniones de t_i^k sobre el criterio c_j se combinan en tripletas de opinión representativas teniendo en cuenta la cantidad y diversidad de opiniones. Finalmente, las tripletas de opinión representativas asociadas a e_k se recopilan en una matriz de evaluación textual individual. Asimismo, transformamos las calificaciones numéricas opcionales proporcionadas por e_k en tripletas de opinión y las recopilamos en una matriz de evaluación numérica individual. Fusionamos las evaluaciones textuales y numéricas obteniendo la matriz de preferencia individual IP^k del experto e_k.

4) Agregación colectiva bajo representación escasa: Se fusionan las matrices IP^k en una matriz colectiva. Las matrices tienen falta de información si los expertos no evalúan todas las alternativas o criterios, por lo que diseñamos un modelo de representación escasa para agregarlas que asigna puntuaciones a los expertos considerando su implicación y extremismo.

5) Clasificación de las alternativas: Se proporciona un *ranking* de las alternativas. Agregamos la evaluación colectiva ponderando los criterios mediante la atención de los expertos.

IV. Caso de estudio: elección de un restaurante

El modelo CDM-SA permite resolver situaciones de decisión reales. En este caso de estudio, elegimos un restaurante a través de la evaluación de los usuarios de la red social TripAdvisor. Recopilamos las reseñas de restaurantes de esta plataforma construyendo el conjunto de datos TripR-2020Large, que anotamos a nivel de aspecto y publicamos.[1] Contiene 474 reseñas en inglés de 132 expertos que evalúan 4 restaurantes londinenses X = {*Oxo Tower*, *J. Sheekey*, *The Wolseley*, *The Ivy*}, que son desglosadas en 2.522 oraciones.

Las reseñas son analizadas por 4 SAMs: *end_to_endCNN* [4], *JointModel* [5], *TASBERT* [6], and *DOC-ABSADeepL* [1]. Las categorías identificadas se recogen en un conjunto de criterios C= {*Restaurant*, *Food*, *Drinks*, *Ambience*, *Service*, *Location*}. Captamos los matices de las opiniones en tripletas de opinión y las agregamos mediante el sistema de votación de SAMs. Construimos la matriz IP^k para cada experto e_k recogiendo tanto las opiniones inferidas de los textos como las calificaciones numéricas opcionales proporcionadas en TripAdvisor. Estas matrices se agregan a través del modelo de representación escasa obteniendo la evaluación colectiva. Generamos la evaluación final para cada restaurante ponderando los criterios a través de la atención de los expertos. El *ranking* final es *The Ivy*>*J. Sheekey*>*The Wolseley*>*Oxo Tower*.

[1] https://github.com/ari-dasci/OD-TripR-2020Large.

V. Conclusiones

En nuestro estudio concluimos que: 1) el modelo CDM-SA captura la sabiduría de la multitud disponible en redes sociales, 2) las tripletas de opinión recogen los matices de opinión, 3) el modelo de representación escasa gestiona la falta de información, y 4) la base de datos the TripR-2020Large permite evaluar modelos DM que incorporan análisis de opiniones.

Agradecimientos

Este trabajo está respaldado por los proyectos PRE2018-083884, PID2019-103880RB-I00, PID2020-119478GB-I00, PID2020-116118GA-I00, EQC2018-005-084-P, P18-FR-4961 financiados por MCIN/AEI/10.13039/5011 00011033, "ERDF A way of making Europe", "ESF Investing in your future" y "Proyectos I+D+i Junta de Andalucía 2018".

Referencias

[1] Surowiecki, J. (2005). The wisdom of crowds. Anchor.

[2] Zuheros, C., Martínez-Cámara, E., Herrera-Viedma, E., & Herrera, F. (2021). Sentiment analysis based multi-person multi-criteria decision making methodology using natural language processing and deep learning for smarter decision aid. Case study of restaurant choice using TripAdvisor reviews. Information Fusion, 68, 22-36.

[3] Zuheros, C., Martínez-Cámara, E., Herrera-Viedma, E., & Herrera, F. (2022). Crowd Decision Making: Sparse Representation guided by Sentiment Analysis for leveraging the Wisdom of the Crowd. IEEE Transactions on Systems, Man, and Cybernetics: Systems. In press.

[4] Schmitt, M., Steinheber, S., Schreiber, K., & Roth, B. (2018). Joint Aspect and Polarity Classification for Aspect-based Sentiment Analysis with End-to-End Neural Networks. In Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 1109-1114).

[5] Li, Y., Yang, Z., Yin, C., Pan, X., Cui, L., Huang, Q., & Wei, T. (2020, October). A Joint Model for Aspect-Category Sentiment Analysis with Shared Sentiment Prediction Layer. In China National Conference on Chinese Computational Linguistics (pp. 388-400). Springer, Cham.

[6] Wan, H., Yang, Y., Du, J., Liu, Y., Qi, K., & Pan, J. Z. (2020). Target-aspect-sentiment joint detection for aspect-based sentiment analysis. In Proc. of the AAAI Conf. on Artificial Intelligence (Vol. 34, No. 05, pp. 9122-9129).

Automorphisms on normal and convex fuzzy truth values revisited

1st Susana Cubillo
Dept. Matemática aplicada a las TIC
Universidad Politecnica de Madrid
scubillo@fi.upm.es

2nd Carmen Torres-Blanc
Dept. Matemática aplicada a las TIC
Universidad Politecnica de Madrid
ctorres@fi.upm.es

3rd Luis Magdalena
Dept. Matemática aplicada a las TIC
Universidad Politecnica de Madrid
lmagdalena@fi.upm.es

***Abstract*—This work is a summary of our paper 'Automorphisms on normal and convex fuzzy truth values revisited', published in the journal 'Fuzzy sets and systems' (vol. 431, pp. 143-159, (2022)). That paper focused on the partially ordered set of convex and normal functions from $[0,1]$ to $[0,1]$, characterizing the automorphisms in this set preserving the function 1.**

***Index Terms*—automorphisms, normal and convex functions, linguistic variables**

I. Introduction

Type-2 fuzzy sets (T2FSs) were introduced by Zadeh in 1975 ([11]) as an extension of type-1 fuzzy sets. As is well known, the degree of membership of an element in a T2FS is a fuzzy set on $[0,1]$, or equivalently, a label of the linguistic variable "TRUTH". In this way, a T2FS is determined by a membership function $\mu : X \to \mathbf{M}$, where $\mathbf{M} = [0,1]^{[0,1]} = Map([0,1],[0,1])$ is the set of functions from $[0,1]$ on $[0,1]$ ([2], [7]).

In the last decade, many researchers have studied type-2 fuzzy sets theory and its application to several fields of science (see [3], [4], [5], [6], [10]). In particular, many operations, properties and results obtained in the T1FSs have been extended to T2FSs, throughout Zadeh's Extension Principle. This is the case of the automorphisms in **M** studied by Walker and Walker ([7], [8]).

Our cited paper focused on characterizing order automorphisms in **M**. Furthermore, since the degrees of membership of type-2 fuzzy sets represent linguistic terms of the "TRUTH" variable, our objective was to characterize the automorphisms that provide symmetry between the labels representing "TRUE" and those representing "FALSE".

On the other hand, in general, linguistic labels are often constructed with normal and convex functions in $[0,1]$, so we restricted ourselves to the partially ordered set $(\mathbf{L}, \sqsubseteq)$, where **L** is the subset of normal and convex functions of **M**, and $\sqsubseteq$ the partial order defined in previous papers devoted to T2FSs ([2], [9]). Also, the labels representing "TRUE" are increasing and the labels representing "FALSE" are decreasing, so we were interested in order automorphisms in **L** that transform increasing functions in increasing ones, and decreasing functions in decreasing ones. As the only function in **L** that is both increasing and decreasing is the constant function **1** (**1** (x) = 1 for all $x \in [0,1]$), our goal was to characterize automophisms in $(\mathbf{L}, \sqsubseteq)$ that preserve order and leave the constant function **1** fixed.

J. Harding, C. L. Walker and E. A. Walker, had already presented in some of their works a family of functions in **L** (see [9]), assuming that they are automorphisms. In our paper, first we prove that, in fact, they are automorphisms. Furthermore, the cited authors showed that those functions satisfied the condition of leaving the function **1** fixed. The proof of these properties only requires checking the necessary conditions. But, despite the significant studies carried out by these authors, the problem of determining if there are other automorphisms that fulfill this same property was left open. Then, we closed this question, obtaining that, in fact, these automorphisms are the only ones that verify that property. This is not a trivial task. In fact, it was necessary to establish some intermediate results, which finally made it possible to obtain the appropriate functions to write any automorphism as considered by J. Walker et al. in [9].

II. Previous steps to attain the characterization

In [9], the authors considered the following family of functions $F : \mathbf{L} \to \mathbf{L}$, defined as

$$F(f) = (\beta \circ f^L \circ \alpha) \wedge (\gamma \circ f^R \circ \alpha),$$

with α, β and γ automorphisms in $([0,1], \leq)$, assuming that they are automorphisms in $(\mathbf{L}, \sqsubseteq)$, and showing that they leave the function **1** fixed.

Then, we address the problem of knowing if there are other automorphisms satisfying the same conditions. So, the main result of our paper was to show that the only automorphisms that leave the function **1** fixed are those presented in [9]. It is worth noting that from a mathematical point of view, proving that those functions are automorphisms leaving the function **1** fixed simply requires checking the conditions that should be verified. But in order to prove that any automorphism satisfying that condition belongs to the family presented in [9], it is necessary to define some right functions, which is not a trivial task. In fact, as far as we know, this characterization had not been obtained, even though some researchers, such as J. Harding, C. L. Walker and E. A. Walker, have carried out important work on automorphisms in **L**.

This research has been partially supported by the Government of Spain (grant PGC2018-096509-B-I00) and Universidad Politécnica de Madrid.

This result was achieved following some steps.

- Firstly, we proved that, in fact, if α, β and γ are automorphisms in $([0,1],\leq)$ then the function $F:\mathbf{L}\to\mathbf{L}$, such that $F(f)=(\beta\circ f^L\circ\alpha)\wedge(\gamma\circ f^R\circ\alpha)$ is an automorphism in $(\mathbf{L},\sqsubseteq)$, which leaves the function $\mathbf{1}$ fixed. But F is not an automorphism in $\mathbf{M}$.
- More, we showed with some examples that to be sure that this operator F is well defined, the function α must be the same in both terms.
- Next, we studied some properties of the automorphisms in $(\mathbf{L},\sqsubseteq)$ leaving the function $\mathbf{1}$ fixed. These properties will lead us to construct the required representation of the automorphism, in some steps.
- We prove that if $F:\mathbf{L}\to\mathbf{L}$ is an automorphism, with $F(\mathbf{1})=\mathbf{1}$, the image of a singleton $\overline{a}$ is also a singleton $\overline{b}$. This property will allow us to construct an automorphism $\alpha_F:([0,1]\leq)\to([0,1]\leq)$, defined by $\alpha_F(a)=b$ if $F(\overline{a})=\overline{b}$.
- Then, we obtained the perhaps more significant result of the paper, as it is a key point for the rest of the proofs:

  Let F be an automorphism in $(\mathbf{L},\sqsubseteq)$, such that $F(\mathbf{1})=\mathbf{1}$. For all $a\in[0,1]$, if f and g are decreasing (or both increasing) functions, and $f(a)=g(a)$, then $(F(f))(\alpha_F(a))=(F(g))(\alpha_F(a))$,
- Some intermediate functions were defined:

  Let $f_{a,1}$ be the function in $\mathbf{L}$ given by

$$f_{a,1}(x)=\begin{cases}1, & \text{if } x\neq 1,\\ a, & \text{if } x=1.\end{cases}$$

  and the function in $\mathbf{L}$ given by

$$f^{a,0}(x)=\begin{cases}a, & \text{if } x=0,\\ 1, & \text{if } x\neq 0.\end{cases}$$

  These functions will lead us to define the automorphism β_F .
- Similarly, two functions in $\mathbf{L}$, $f_{a,m}$ and $f^{a,m}$:

$$f_{a,m}(x)=\begin{cases}1, & \text{if } x<m,\\ a, & \text{if } x\geq m.\end{cases}$$

  and the function in $\mathbf{L}$ given by

$$f^{a,m}(x)=\begin{cases}a, & \text{if } x\leq m,\\ 1, & \text{if } x>m.\end{cases}$$

  will allow us to define the automorphism γ_F.

III. The main result

A latest result before attaining the main theorem was presented:

Let F be an automorphism in $(\mathbf{L},\sqsubseteq)$, with $F(\mathbf{1})=\mathbf{1}$. Then, there are three order automorphisms in $([0,1],\leq)$, α_F, β_F and γ_F, such that for all $x\in[0,1]$ it holds:

if f is a decreasing function, it is $(F(f))(x)=\beta_F(f(\alpha^{-1}(x)))$, and

if f is an increasing function, it is $(F(f))(x)=\gamma_F(f(\alpha^{-1}(x)))$.

All the previous results will finally lead to obtaining the desired characterization.

F is an automorphism in $(\mathbf{L},\sqsubseteq)$, such that $F(\mathbf{1})=\mathbf{1}$, if and only if, there exist three order automorphisms in $[0,1]$, α_F, β_F and γ_F, such that

$$(F(f))(x)=\beta_F(f^R(\alpha_F^{-1}(x)))\wedge\gamma_F(f^L(\alpha_F^{-1}(x)))\text{ for all }x\in[0,1].$$

IV. Conclusions

In the paper published in the journal 'Fuzzy Sets and Systems', to which we have referred ([1]), we have attained a characterization of the family of automorphisms presented in the work of Walker et al. [9]. This characterization was achieved through some intermediate steps, generally non-trivial.

However, the existence of automorphisms that do not meet that property remains to be analyzed. This is a study to which the authors will dedicate future research.

References

[1] S. Cubillo, C. Torres-Blanc and L. Magdalena, "Automorphisms on normal and convex fuzzy truth values revisited", Fuzzy sets and systems, vol. 431, pp. 143-159, (2022).

[2] J. Harding, C.L. Walker and E.A. Walker, "The Truth Value Algebra of Type-2 ", Fuzzy Sets, Order Convolutions of Functions on the Unit Interval, Chapman and Hall/CRC (2016).

[3] A.K. Shukla, S.K Banshal, T. Seth, A. Basu, R. John and P.K. Muhur, "A bibliometric overview of the field of type-2 fuzzy sets and systems" [discussion forum], IEEE Computational Intelligence Magazine, vol.15 (1), pp. 89-98 (2020).

[4] C. Torres-Blanc, S. Cubillo and P. Hernández-Varela, "New negations on the membership functions of type-2 fuzzy sets", IEEE Transactions on Fuzzy Systems, vol. 27 (7), pp. 1397-1406 (2019).

[5] C. Torres-Blanc, S. Cubillo, and P. Hernández, "Aggregation operators on type-2 fuzzy sets", Fuzzy Sets and Systems, vol. 324, pp. 74-90 (2017).

[6] C. Torres-Blanc, P. Hernández-Varela and S. Cubillo, "Self-contradiction for type-2 fuzzy sets whose membership degrees are normal and convex functions", Fuzzy Sets and Systems, vol. 352, pp. 73-91 (2018).

[7] C.L. Walker and E.A. Walker, "Automorphisms of the algebra of fuzzy truth values", Int. J of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 14 (6), pp. 711-732 (2006).

[8] C.L. Walker and E.A. Walker,"Automorphisms of the algebra of fuzzy truth values II", Int. J of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 16 (5), pp. 627-643 (2008).

[9] C.L. Walker and E.A. Walker, "Automorphisms of subalgebras of the algebra of fuzzy truth values of type-2 fuzzy sets", Proc. 29^{th} Linz Seminar on Fuzzy Set Theory (LINZ 2008), pp. 118-120 (2008).

[10] D. Wu and J.M. Mendel, "Similarity measures for closed general type-2 fuzzy sets: Overview, comparisons, and a geometric approach", IEEE Transactions on Fuzzy Systems, vol. 27 (3), pp. 515-526 (2019).

[11] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I", Information Sciences, vol. 8 (3), pp. 199-249 (1975).

Operadores de Consecuencias y de Interior Borrosos y Relaciones Borrosas*

J. Elorza
Dpto de Física y Matemática Aplicada, Fac. Ciencias
Universidad de Navarra
Pamplona, España
jelorza@unav.es

J. Recasens
Sec. Matemàtiques, ETSAV
Universitat Politècnica de Catalunya
Sant Cugat del Vallès, España
j.recasens@upc.edu

***Resumen*—Se estudia la relación entre los operadores de consecuencias y de interior borrosos generados por relaciones borrosas R y los generados por los $*$-preórdenes obtenidos a partir de las filas y columnas de R mediante el bien conocido Teorema de Representación para $*$-preórdenes.**

***Palabras Clave*—Operador de consecuencias borroso, operador de interior borroso, relación borrosa, $*$-preorden, Teorema de Representación para $*$-preórdenes**

I. Introducción

Es sabido que a partir de un $*$-preorden R en un universo X (Definición 2) se pueden deducir un operador de consecuencias c_R y un operador de interior i_R [3], [4]:

- $c_R(\mu)(x) = \sup_{y\in X}\{\mu(y) * R(y,x)\}$
- $i_R(\mu)(x) = \inf_{y\in X}\{R(x,y) \to_* \mu(y)\}$.

La relación entre R, c_R e i_R ha sido ampliamente estudiada. Los principales resultados son:

- A partir de un operador c de L^X (L un retículo residuado completo) se puede derivar una relación borrosa R_c y si $c = c_R$ con R un $*$-preorden, entonces $R = R_{c_R}$ y $c_R = c_{R_c}$.
- A partir de un operador i de L^X se puede obtener una relación borrosa R_i y si $i = i_R$ con R un $*$-preorden R, entonces $R = R_{i_R}$ y $i_R = i_{R_i}$.

Se ha prestado menos atención a la relación entre operadores borrosos y relaciones borrosas generales. El estudio más relevante es [1] donde, combinando c_R e i_R, Bodenhofer obtiene un operador de consecuencias $\gamma_R = i_R \circ c_R$ y un operador de interior $\delta_R = c_R \circ i_R$ a partir de una relación borrosa cualquiera R. A partir de γ_R o δ_R no se puede recuperar R en general. La relación entre estos operadores y R se obtiene a partir de un caso especial del Teorema de Representación para $*$-preórdenes dado en [8]: las columnas y filas de una relación borrosa R generan dos $*$-preórdenes R_{col} y R_{fil}. La relación entre estos preórdenes y γ_R y δ_R se estudiará en la sección III. Los resultados más importantes son

- Si γ_R (δ_R) es el operador de consecuencias (de interior) generado per la relación borrosa R, entonces R_{γ_R} (R_{δ_R}) es el $*$-preorden R_{col} ($*$-preorder R_{fil}) generado por las columnas (filas) de R según el Teorema de Representación de $*$-preórdenes.
- $c_{R_{col}} \le \gamma_R$ e $i_{R_{fil}} \ge \delta_R$.

Una versión extendida de este trabajo se ha enviado a la revista Fuzzy Sets and Systems y actualmente se halla en fase de revisión.

II. Preliminares

Esta sección contiene definiciones sobre operadores y relaciones borrosas que se usarán en la siguiente sección y mostramos el Teorema de Representación para $*$-preórdenes.

A lo largo de todo el trabajo, todos los subconjuntos borrosos se supondrán de un universo X y valorados en un retículo residuado completo $\langle L, \wedge, \vee, *, \to_*, 0, 1\rangle$.

Definición 1. *Una relación borrosa en X es una aplicación $R : X \times X \to L$. Dado $x \in X$, el subconjunto borroso R_x ($_xR$) de X definido para todo $y \in X$ por $R_x(y) = R(x,y)$ ($_xR(y) = R(y,x)$) se llama una fila (columna) de R.*

Definición 2. [9] *Una relación borrosa R en un universo X es un $*$-preorden si para todo $x, y, z \in X$*

1. *$R(x,x) = 1$ (Reflexividad)*
2. *$R(x,y) * R(y,z) \le R(x,z)$ ($*$-transitividad).*

Teorema de Representación 3. [8] *Sea R una relación borrosa en X. R es un $*$-preorden de X si, y sólo si, existe una familia $(\mu_i)_{i\in I}$ de subconjuntos borrosos de X tal que para todo $x, y \in X$,*

$$R(x,y) = \inf_{i\in I}\{\mu_i(y) \to_* \mu_i(x)\}.$$

En particular, si se consideran las familias $(R_x)_{x\in X}$ y $(_xR)_{x\in X}$ de filas y columnas de una relación borrosa R se obtiene la siguiente definición.

Definición 4. *Sea R una relación borrosa de X y $(R_x)_{x\in X}$, $(_xR)_{x\in X}$ las familias de filas y columnas de R respectivamente. Entonces los $*$-preórdenes R_{fil} y R_{col} de X se definen para todo $x, y \in X$ como*

$$R_{fil}(x,y) = \inf_{z\in X}\{R_z(y) \to_* R_z(x)\}$$
$$R_{col}(x,y) = \inf_{z\in X}\{_zR(y) \to_* {}_zR(x)\}.$$

III. Operadores de Consecuencias, de Interior y Relaciones Borrosas

En esta sección se verá la relación entre operadores de consecuencias, de interior y relaciones borrosas. A partir de

los dos operadores c_R e i_R se generan otros dos, $\gamma_R = i_R \circ c_R$ y $\delta_R = c_R \circ i_R$, que son de consecuencias y de interior para cualquier relación borrosa R [1]. La relación entre γ_R, c_R y R se estudiará con ayuda del Teorema de Rerpesentación de $*$-preórdenes. Usando la dualidad entre γ_R y δ_R (Proposición 13) también se establecerá la relación entre δ_R, i_R y R.

Definición 5. [4] *Un operador borroso $c : L^X \to L^X$ en X es un operador de consecuencias si para todo par de subconjuntos borrosos μ, ν de X se satisface*

1. $\mu \leq c(\mu)$
2. *Si* $\mu \leq \nu$, *entonces* $c(\mu) \leq c(\nu)$
3. $c(c(\mu)) = c(\mu)$.

Definición 6. *Un operador borroso $i : L^X \to L^X$ en X es un operador de interior borroso si para todo par de subconjuntos borrosos μ, ν de X se satisface*

1. $\mu \geq i(\mu)$
2. *Si* $\mu \leq \nu$, *entonces* $i(\mu) \leq i(\nu)$
3. $i(i(\mu)) = i(\mu)$.

Definición 7. [1], [2], [4], [7] *Sea R una relación borrosa en X. Los operadores borrosos c_R e i_R de X se definen para cada subconjunto borroso μ de X por*

$$c_R(\mu)(x) = \sup_{y \in X} \{\mu(y) * R(y, x)\} \qquad \forall x \in X$$
$$i_R(\mu)(x) = \inf_{y \in X} \{R(x, y) \to_* \mu(y)\} \qquad \forall x \in X.$$

Proposición 8. [4] *Si R es un $*$-preorden de X, entonces c_R e i_R son un operador de consecuencias y de interior respectivamente.*

R puede recuperarse a partir de c_R o i_R, si es un $*$-preorden (Proposición 10).

Definición 9. [5], [7] *Sea p un operador en L^X. Entonces las relaciones borrosas R_p y R^p de X se definen por*

$$R_p(x, y) = p(\varphi_x)(y)$$
$$R^p(x, y) = \inf_{\alpha \in L} \{p(\varphi_x \to_* \underline{\alpha})(y) \to_* \alpha\}$$

para todo $x, y \in X$, donde, dado $\alpha \in L$, el subconjunto borroso $\underline{\alpha}$ de X se define para todo $x \in X$ por $\underline{\alpha}(x) = \alpha$ y φ_x es el singleton $\{x\}$.

Proposición 10. [4] *Sea R un $*$-preorden en X. Entonces $R_{c_R} = R^{i_R} = R$.*

Si la relación R no es un $*$-preorden, entonces ni c_R es un operador de consecuencias borroso ni i_R de interior. Además R_{c_R} y R^{i_R} no coinciden con R. El resto de esta sección está dedicada a clarificar esta situación.

Proposición 11. [1]

- *Si R es una relación en X, entonces $\gamma_R = i_R \circ c_R$ es un operador de consecuencias.*
- *Si R es una relación en X, entonces $\delta_R = c_R \circ i_R$ es un operador de interior.*
- *Si R es un $*$-preorden de X, entonces $\gamma_R = c_R$ y $\delta_R = i_R$.*

Definición 12. *Dada una relación borrosa R de X, la relación inversa R^{-1} de R es la relación borrosa de X definida para todo $x, y \in X$ por $R^{-1}(x, y) = R(y, x)$.*

La siguiente proposición expresa la dualidad entre γ_R y $\delta_{R^{-1}}$

Proposición 13. *Sea R una relación borrosa de X. Entonces γ_R y δ_R son operadores duales en el sentido de que para todo subconjunto borroso μ de X*

$$\gamma_R(\mu) = \inf_{\alpha \in L} \{\delta_{R^{-1}}(\mu \to_* \underline{\alpha}) \to_* \underline{\alpha}\}$$
$$\delta_R(\mu) = \inf_{\alpha \in L} \{\gamma_{R^{-1}}(\mu \to_* \underline{\alpha}) \to_* \underline{\alpha}\}.$$

Proposición 14. *Sea R una relación de X. Entonces $R_{\gamma_R} = R_{col}$ y $R^{\delta_R} = R_{fil}$.*

El siguiente resultado muestra relaciones de desigualdad entre γ_R y $c_{R_{col}}$ y entre i_R y δ_R.

Proposición 15. *Sea R una relación borrosa en X. Entonces $c_{R_{col}} \leq \gamma_R$ e $i_{R_{fil}} \geq \delta_R$.*

IV. Conclusiones

Se ha estudiado la relación entre relaciones borrosas y operadores de consecuencias y de interior borrosos. Se pueden definir dos operadores c_R e i_R a partir de R que, en general, no son de consecuencias o de interior. Sin embargo, sus composiciones $\gamma_R = i_R \circ c_R$ y $\delta_R = c_R \circ i_R$ sí lo son; esto clarifica la situación para $*$-preórdenes donde c_R e i_R son operadores de consecuencias y de interior porque, en este caso, $\gamma_R = c_R$ y $\delta_R = i_R$.

Agradecimientos

Los autores agradecen las ayudas del Ministerio de Economía y Competitividad bajo los proyectos PGC2018-098623-B-I00 y PGC2018-095709-B-C21 (financiado por MCIN/AEI/10.13039/501100011033 y FEDER "Una manera de hacer Europa"), respectivamente.

Referencias

[1] U. Bodenhofer, A unified framework of opening and closure operators with respect to arbitrary relations. Soft Computing 7, 220–227 (2003).
[2] D. Boixader, J. Jacas, J. Recasens, Upper and lower approximation of fuzzy sets, Int. J. of General Systems 29, 555–568 (2000).
[3] N. Carmona, J. Elorza, J. Recasens, J. Bragard, Permutable Fuzzy Consequence and Interior Operators and their Connection with Fuzzy Relations, Information Sciences 310, 36–51 (2015).
[4] J. Elorza, P. Burillo, On the relation between fuzzy preorders and fuzzy consequence operators, Int. J. Uncert., Fuzz. Knowl.-Based Syst. 7, 219–234 (1999).
[5] J. Elorza, R. Fuentes-González, J. Bragard, P. Burillo, On the relation between fuzzy closing morphological operators, fuzzy consequence operators induced by fuzzy preorders and fuzzy closure and co-closure systems, Fuzzy Sets and Systems 218, 73–89 (2013).
[6] J.C. Fodor, M. Roubens, Structure of transitive valued binary relations. Math. Soc. Sci. 30, 71–94 (1995).
[7] N.N. Morsi, M.M. Yakout, Axiomatics for fuzzy rough sets, Fuzzy Sets and Systems 100, 327–342 (1998).
[8] L. Valverde, On the Structure of F-indistinguishability Operators. Fuzzy Sets and Systems 17, 313–328 (1985).
[9] L.A. Zadeh, Similarity relations and fuzzy orderings, Inform. Sci. 3, 177–200 (1971).

Modelización difusa intuicionista para simulación de computación cuántica

1st Anderson Cruz
Instituto Metrópole Digital
Universidade Federal do Rio Grande do Norte (UFRN)
Natal, Brasil
anderson@imd.ufrn.br

2nd Helida Santos
Dept. de Estadística, Informática y Matemáticas, Institute of Smart Cities (ISC)
Universidade Pública de Navarra (UPNA)
Pamplona, España
helida.salles@unavarra.es

3rd Renata Reiser
Lab. Ubiquitous and Parallel Systems
Universidade Federal de Pelotas (UFPEL)
Pelotas, Brasil
reiser@inf.ufpel.edu.br

4th Javier Fernández
Dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra (UPNA)
Pamplona, España
fcojavier.fernandez@unavarra.es

5thHumberto Bustince
Dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra (UPNA)
Pamplona, España
bustince@unavarra.es

***Resumen*—En los últimos años, la computación cuántica ha ganado atención creciente. Los desafíos actuales del área están relacionados con las tecnologías usadas para desarrollar ordenadores y algoritmos cuánticos. En este contexto, la área de investigación de simulación de la computación cuántica es cada vez más relevante. En esta perspectiva, este trabajo presenta una modelización de cómo los conectivos difusos intuicionistas pueden ser interpretados vía estados cuánticos. Además estos conectivos preservan su representación a partir de la aplicación de propiedades cuánticas. Por lo tanto, esos conectivos pueden simular el comportamiento de puertas cuánticas y sus informaciones pueden ser representadas por registros cuánticos.**

***Palabras Clave*— Conectivos difusos intuicionistas, computación cuántica, simulación de circuitos cuánticos.**

El creciente interés por la computación cuántica se justifica por su aplicación en la teoría de la computación (vea [1]–[3]), en particular desde del punto de vista de la optimización, debido a que no hay limite teórico para la velocidad de resolución de tareas computacionales cuánticas (ver [4] para más detalles).

En este sentido, dentro de los algoritmos cuánticos, podemos citar: el algoritmo de Grover el cual se usa para hallar un elemento en una base con n entradas, y requiere $O(\sqrt{n})$ pasos [5], en contraste con la complejidad del mejor algoritmo clásico el cual es $O(n)$. También se encuentra el algoritmo de Shor que transformó el problema de factorización en primos de un entero en un problema de complejidad polinomial [6].

Además, debido al comportamiento de los sistemas cuánticos se espera que ellos sean una herramienta más eficiente para solucionar problemas del mundo cuántico relacionados con las áreas de la física, la química y la biología.

Por otro lado, aún existen grandes desafíos tecnológicos para controlar y operar efectivamente una cantidad de qubits que haga los algoritmos cuánticos útiles para los problemas reales. Estos desafíos incluyen los problemas con respecto al tiempo de vida útil de los qubits así como la fiabilidad de las puertas cuánticas.

Por consiguiente, las limitaciones de implementación y de disponibilidad de los ordenadores cuánticos han dado paso a soluciones de computación clásica que simulen los comportamientos de la computación cuántica. Aunque la computación cuántica sea aplicable en unos años, la simulación de la computación cuántica es importante para: comprender el funcionamiento de esos tipos de ordenadores; realizar la integración entre la computación clásica y la cuántica; y construir estructuras lógicas de lenguajes de programación cuántica. Además, la relevancia de los simuladores es independiente del tipo de tecnología en la que los ordenadores cuánticos serán implementados al final y es determinante para las interpretaciones de las estructuras de decisión y repetición de los lenguajes de programación cuántica de alto nivel que se desarrollen en el futuro.

Según la interpretación de [7], [8], los operadores de la lógica difusa en la computación cuántica pueden ser generalizados usando los conjuntos difusos intuicionistas de Atanassov (CDIA) [9], que tienen la posibilidad de representar tanto el conocimiento como el desconocimiento, mediante las funciones de pertenencia y no pertenencia.

Por tal razón, en este trabajo presentamos una forma de extender los conjuntos difusos para un enfoque intuicionista. A partir de tal extensión, presentamos las descripciones de los operadores de union, intersección, complemento, implicación, coimplicación, diferencia y codiferencia. Así, también podemos simular los circuitos cuánticos y representar la evolución temporal de los estados cuánticos correspondientes.

Para atender la perspectiva de simulación de circuitos cuánticos, otros enfoques son posibles, sin embargo los re-

Este trabajo ha sido financiado por el Ministerio Español de Ciencia e Innovación - PID2019-108392GBI00 (AEI/10.13039/501100011033) y Fundación La Caixa.

sultados de la implementación de este método en *clusters* de ordenadores clásicos (ver [10]) evidencian una buena perspectiva practica para la modelización presentada en este trabajo.

En un futuro planeamos presentar la modelización de otras puertas lógicas y estructuras lógicas de programación de alto nivel, así como estudiar la aplicación de los operadores de (pre)agregación en puertas cuánticas.

REFERENCIAS

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[2] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6, p. 467–488, 1982.

[3] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, "Elementary gates for quantum computation," *Physical Review A*, vol. 52, no. 5, pp. 3457–3467, 1995.

[4] C. H. Bennett, "Time/space trade-offs for reversible computation," *SIAM Journal on Computing*, vol. 18, no. 4, p. 766–776, 1989.

[5] C. Cafaro and S. Mancini, "On Grover's search algorithm from a quantum information geometry viewpoint," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1610–1625, 2012.

[6] C. Bennett and P. Shor, "Quantum information theory," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2724–2742, 1998.

[7] M. Mannucci, "Quantum fuzzy sets: Blending fuzzy set theory and quantum computation," *CoRR*, vol. abs/cs/0604064, 2006. [Online]. Available: http://arxiv.org/abs/cs/0604064

[8] A. B. de Avila, M. Schmalfuss, R. Reiser, and V. Kreinovich, "Fuzzy xor classes from quantum computing," in *Artificial Intelligence and Soft Computing - 14th International Conference, ICAISC 2015, Zakopane, Poland, June 14-28, 2015, Proceedings, Part II*, ser. Lecture Notes in Computer Science, L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, and J. M. Zurada, Eds., vol. 9120. Springer, 2015, pp. 305–317. [Online]. Available: https://doi.org/10.1007/978-3-319-19369-4_28

[9] K. T. Atanassov, *Intuitionistic Fuzzy Sets - Theory and Applications*, ser. Studies in Fuzziness and Soft Computing. Physica-Verlag, 1999, vol. 35. [Online]. Available: https://doi.org/10.1007/978-3-7908-1870-3

[10] A. B. de Avila, "Hybrid-gm: Parallel model for quantum computing targeted to hybrid architectures," p. 96, 2022, Thesis - UFPEL.

Razonamiento por similaridad y razonamiento difuso

Juan Luis Castro
Departamento de Ciencias de la
Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
castro@decsai.ugr.es

Resumen—El aprendizaje automático requiere de un conjunto de ejemplos de cierto tamaño para que los modelos sean entrenados y ajustados adecuadamente. Sin embargo, hay problemas donde no se dispone de un conjunto de ejemplos suficiente. En el caso de que dispongamos de una relación de similaridad entre los posibles casos del problema de forma que los casos similares tengan una alta probabilidad de pertenecer a la misma clase, podemos utilizar esta relación para actualizar nuestra creencia en la clase de cada caso de acuerdo a los casos más cercanos con clase conocida. En este trabajo estudiaremos un modelo que aplica esa idea, llamado razonamiento por similaridad y su relación con el razonamiento difuso. Concretamente:
- Presentaremos el modelo de razonamiento por similaridad
- Demostraremos que el Modus Ponens generalizado puede verse como razonamiento por similaridad
- Demostraremos que los sistemas basados en reglas difusas es un caso particular de razonamiento por similaridad.

Palabras Clave—Razonamiento por simlilaridad, Lógica difusa, Modus Ponens Generalizado, Sistemas basados en lógica difusa.

I. INTRODUCCION

El aprendizaje automático requiere de un conjunto de ejemplos de un buen tamaño para que los modelos sean entrenados y ajustados adecuadamente. Bajo estas condiciones, los algoritmos del estado del arte, principalmente modelos de Deep Learning, están consiguiendo unos excelentes resultados. Sin embargo, hay muchos problemas donde no solo es que no se disponen de esos datasets de un tamaño suficiente, sino que resulta complicado llegar a conseguirlos. Esto puede ser debido a un elevado coste económico, o a un elevado coste en cuanto al tiempo requerido o simplemente a la dificultad para obtener cada nuevo caso. En estas circunstancias estas metodologías de Aprendizaje automático no consiguen un resultado adecuado y por tanto hay que buscar en otras metodologías.

Los sistemas basados en buscar casos similares para obtener la solución al caso actual han sido utilizados con éxito en el ámbito de la Inteligencia Artificial [1]. Desde el clásico k-NN hasta complejos sistemas de razonamiento basados en casos han sido utilizados con éxito para obtener soluciones a problemas complejos del ámbito de la Inteligencia Artificial. En la gran mayoría de estos casos se ha planteado esta metodología como un procedimiento algorítmico o heurístico [5]. Sin embargo, este es un esquema de razonamiento muy utilizado por los seres humanos, que argumentan en muchos casos sus decisiones en base a casos similares. Por ello, nos planteamos en este trabajo analizar desde un punto de vista de la lógica este tipo de razonamiento, incorporándolo como esquema de inferencia y analizando las propiedades del sistema lógico obtenido.

II. RAZONAMIENTO POR SIMILARIDAD

A. *Modelo difuso de razonamiento por similaridad*

En el ámbito de la lógica difusa, podríamos describir el razonamiento por similaridad [4] con el siguiente esquema de inferencia:

"X es_similar_a Y" es S

P(X) es A

P(Y) es m(S,A)

donde m es un modificador que traslade a Y la propiedad P mediante la relación de similaridad S.

Por ejemplo, podríamos utilizar este esquema para realizar la siguiente deducción:

X es_similar_a Y

Tamaño(X) es grande

Tamaño(Y) es mas_o_menos(grande)

B. *Modelo probabilistico del razonamiento por similaridad*

Si utilizamos un esquema probabilístico, podríamos describir el razonamiento por similaridad [3] con el siguiente esquema de inferencia:

Prob(X=Y) = α

Prob(X es P) = β

$$\text{Prob}(Y \text{ es } P) \geq \alpha * \beta$$

Que a su vez podemos generalizar al esquema

$$\text{Prob}(X=Y) \geq \alpha$$
$$\text{Prob}(X \text{ es } P) \geq \beta$$
$$\text{Prob}(Y \text{ es } P) \geq \alpha * \beta$$

III. Modus ponens generalizado es un ejemplo de razonamiento por similaridad

El modus ponens generalizado se define por el siguiente esquema de inferencia:

If X is A, then Y is B

X is A'

Y is B'

donde

$$B'(y) = \sup_x \{\min(A'(x), I(A(x), B(y))\} \quad (1)$$

siendo I una función de implicación.

En el caso particular en que I se una t-norma, que es el caso más utilizado en las aplicaciones de la lógica difusa, la ecuación (1) se transforma en

$$B'(y) = I(\max (A(x) \wedge A'(x)), B(y)). \quad (2)$$

y en estas circunstancias sería equivalente al siguiente esquema de inferencia por similaridad:

Llamemos X_A a la variable valor de verdad de X es A, e $Y_{X \text{ es } A}$ a la propiedad Y cuando X es A

X_A es_similar_a $X_{A'}$ es S

Propiedad_Y(X_A)=B

Propiedad_Y($X_{A'}$)=m(S,B)

Y ahora, tomando como modificador m=I y como similaridad $S(X_A, X_{A'})(x) = \max (A(x) \wedge A'(x))$, obtenemos que

$$m(S,B) = B' \quad (3)$$

y así obtenemos el modus ponens generalizado como un caso particular de razonamiento por similaridad.

IV. Sistemas basados en reglas difusas como un ejemplo de sistema razonando por similaridad

Si aplicamos razonamiento por similaridad para un conjunto de casos, similares obtendremos diferentes valores para las propiedades a deducir para el nuevo caso. Si aplicamos razonamiento probabilístico, la integración mediante una función de agregación [2] por encima del máximo mantendría la semántica de las deducciones como cotas inferiores de la probabilidad de que se verifique la propiedad. De esta forma podemos definir la extensión como la agregación mediante una t-co-norma de los distintos valores. A este proceso lo llamaremos extensión por similaridad de la función de creencia de partida:

$$P^{Sim}{}_{(S,D)}(e,y_i)=s(\{P(e,y_i)\} \cup \{P^{Sim}{}_{S,\{c\}}(e,y_i) / c \text{ es similar a } e\}) \quad (4)$$

donde s es una t-conormo

En particular, si aplicamos razonamiento por similaridad para cada una de las reglas de un sistema basado en reglas obtendremos distintos valores para la variables consecuentes, concretamente uno por cada regla, que si lo integramos mediante la extensión por similaridad a través de una t-conorma, obtendremos exactamente la salida difusa del sistema basado en reglas difusas. De esta forma deducimos que la inferencia de un sistema de reglas difusas no es mas que un ejemplo particular de aplicación de un sistema de razonamiento por similaridad.

References

[1] Castro, J.L., Francisco, M.: Similarity fuzzy semantic networks and inference. an application to analysis of radical discourse in twitter. In: Artificial Intelligence and Soft Computing. Proceedings of ICAISC 2022. Lecture Notes in Computer Science, Springer International Publishing (Jun 2022)

[2] Dujmoví c, J., Torra, V.: Aggregation Functions in Decision Engineering: Ten Nec- essary Properties and Parameter-Directedness. In: Kahraman, C., Cebi, S., Ce- vik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds.) Intelligent and Fuzzy Tech- niques for Emerging Conditions and Digital Transformation. pp. 173–181. Lecture Notes in Networks and Systems, Springer International Publishing, Cham (2022)

[3] Francisco, M., Castro, J.L.: A fuzzy model to enhance user profiles in microblogging sites using deep relations. Fuzzy Sets and Systems 401, 133–149 (Dec 2020).

[4] Klawonn, F., Castro, J.L.: Similarity in fuzzy reasoning. Mathware and Soft Com- puting (1995)

[5] Luo, M., Zhao, R.: Fuzzy reasoning algorithms based on similarity. Journal of In- telligent & Fuzzy Systems 34, 213–219 (Jan 2018).

Nuevos resultados sobre el f-índice de inclusión

Nicolás Madrid
Matemática Aplicada
Universidad de Málaga
Málaga, España
nicolas.madrid@uma.es

Manuel Ojeda-Aciego
Matemática Aplicada
Universidad de Málaga
Málaga, España
aciego@uma.es

***Resumen*—En este documento presentamos alguno de los últimos resultados teóricos obtenidos para el f-índice de inclusión. Estos resultados motivan el uso de dicho índice como una nueva forma de representar la inclusión entre dos conjuntos difusos y como un operador de inferencia lógica. En este resumen recordamos dos: se satisfacen los axiomas de Sinha-Dougherty (convenientemente adaptados al marco teórico del f-índice de inclusión) y, además, corresponde a una elección optimal de una implicación difusa residuada para llevar a cabo la inferencia Modus Ponens.**

***Index Terms*—Fuzzy inclusion, Residuated Implication, Modus Ponens, Zadeh's Computational Rule Inference.**

I. Introducción y preliminares

El concepto de inclusión es, sin lugar a dudas, uno de los conceptos más importantes en la teoría de conjuntos. Por este motivo, su generalización difusa ha captado la atención de multitud de investigadores. Sin embargo, a día de hoy no hay consenso sobre cómo extender este concepto y la mayoría de los investigadores se decanta por dos opciones. Por un lado, tenemos una opción constructiva, introducida originalmente por Zadeh, en la que las medidas de inclusión entre dos conjuntos difusos A y B se definen a través de una implicación difusa $\to$ y de la fórmula:

$$S(A,B) = \bigwedge_{u\in\mathcal{U}} A(u) \to B(u), \tag{1}$$

donde $\mathcal{U}$ denota el universo de discurso sobre el que se definen los conjuntos difusos; y la otra, , axiomática, en la que destacan los trabajos de Sinha-Dougherty [9] y de Kitainik [2]. Los axiomas proporcionados por Kitainik tenían como objetivo caracterizar las medidas de inclusion de Zadeh (1) para implicaciones contrapositivas, aunque recientemente se ha demostrado que eso solo sucede cuando el universo $\mathcal{U}$ es finito [3]. Fuera como fuese, y puesto que los axiomas propuestos por Sinha-Dougherty son más restrictivos que los de Kitainik, recordaremos aquí únicamente los primeros:

Definición 1: La función $S\colon \mathcal{F}(\mathcal{U})\times\mathcal{F}(\mathcal{U}) \to [0,1]$ es una *SD-medida de inclusión* si satisface los siguientes axiomas para todo conjunto difuso A, B y C:

Parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, la Junta de Andalucía, la Universidad de Málaga y los fondos FEDER a través de los proyectos PGC2018-095869-B-I00 (MCIU/AEI/FEDER) y UMA2018-FEDERJA-001 (JA/UMA/FEDER)

(SD1) $S(A,B) = 1$ si y solo si $A(u) \le B(u)$ para todo $u \in \mathcal{U}$.
(SD2) $S(A,B) = 0$ si y solo si existe $u \in \mathcal{U}$ tal que $A(u) = 1$ y $B(u) = 0$.
(SD3) Si $B(u) \le C(u)$ para todo $u \in \mathcal{U}$ entonces $S(A,B) \le S(A,C)$.
(SD4) Si $B(u) \le C(u)$ para todo $u \in \mathcal{U}$ entonces $S(C,A) \le S(B,A)$.
(SD5) Si $T\colon \mathcal{U} \to \mathcal{U}$ es una función bijectiva en el universo, entonces $S(A,B) = S(T(A),T(B))$.
(SD6) $S(A,B) = S(B^c,A^c)$.
(SD7) $S(A\cup B,C) = \min\{S(A,C),S(B,C)\}$.
(SD8) $S(A,B\cap C) = \min\{S(A,B),S(A,C)\}$.

donde las operaciones unión, intersección y complementario se definen como: (union) $(A\cup B)(u) = \max\{A(u),B(u)\}$; (intersección) $(A\cap B)(u) = \min\{A(u),B(u)\}$; y (complementario) $A^c(u) = 1 - A(u)$.

f-índice de inclusión

El f-índice de inclusión se define en tres pasos. En el primero, definimos cuales son los "grados" con los que representaremos o mediremos la inclusión entre conjuntos difusos.

Definición 2: El conjunto de f-índices de inclusión, denotado por Ω, es el conjunto formado por las funciones crecientes $f\colon [0,1]\to[0,1]$ que cumplen $f(x) \le x$ para todo $x \in [0,1]$.

En el segundo de los pasos, definiremos cuándo un par de conjuntos difusos satisface un "grado" de f-inclusión.

Definición 3: Sean A y B dos conjuntos difusos y sea $f \in \Omega$. Decimos que A *está f-incluido en* B (denotado por $A \subseteq_f B$) si y solo si la desigualdad $f(A(u)) \le B(u)$ se cumple para todo $u \in \mathcal{U}$.

Finalmente, hay que tener en cuenta que no representaremos la inclusión con una única f-inclusión, sino con todas ellas. En concreto la idea es representar la inclusion siguiendo la idea de que "cuantas más f-inclusiones se satisfagan entre dos conjuntos difusos, mayor será la inclusión". Puede demostrarse [4] que el conjunto de las f-inclusiones que satisfacen dos conjuntos difusos tiene estructura de subretículo y, de hecho, está caracterizado por el máximo de todos ellas. Por ese motivo, el tercero de los pasos consiste en definir el f-índice de inclusión de la siguiente forma:

Definición 4 (f-índice de inclusión): Sean A y B dos conjuntos difusos, el *f-índice de inclusión* de A en B, denotado por $\text{Inc}(A, B)$, se define como

$$\text{Inc}(A, B) = \text{máx}\{f \in \Omega \mid A \subseteq_f B\}$$

II. Relación con los axiomas de Sinha-Dougherty

En [4] y [6] obtuvimos una serie de resultados que muestran que el f-índice de inclusión está en plena sintonía con los axiomas de Sinha-Dougherty. En primer lugar, tenemos el siguiente resultado, donde puede observarse que el f-índice de inclusión satisface 7 de los 8 axiomas tras una adaptación natural de éstos al marco teórico de los f-índices de inclusión:

Teorema 1 (Véase [4]): Dados conjuntos difusos A, B y C

1. $\text{Inc}(A, B) = id$ si y solo si $A(u) \leq B(u)$ para todo $u \in \mathcal{U}$.
2. $\text{Inc}(A, B) = \bot$ si y solo si existe una sucesión $\{u_i\}_{i\in\mathbb{N}} \subseteq \mathcal{U}$ tal que $A(u_i) = 1$ para todo $i \in \mathbb{N}$ y $\bigwedge_{i\in\mathbb{N}} B(u_i) = 0$.

2*. Si $\mathcal{U}$ es finito, entonces $S(A, B) = 0$ si y solo si existe $u \in \mathcal{U}$ tal que $A(u) = 1$ y $B(u) = 0$.

3. Si $B(u) \leq C(u)$ para todo $u \in \mathcal{U}$ entonces $\text{Inc}(C, A) \leq \text{Inc}(B, A)$.
4. Si $B(u) \leq C(u)$ para todo $u \in \mathcal{U}$ entonces $\text{Inc}(A, B) \leq \text{Inc}(A, C)$;
5. Sea $T\colon \mathcal{U} \to \mathcal{U}$ una transformación sobre $\mathcal{U}$, entonces $\text{Inc}(A, B) = \text{Inc}(T(A), T(B))$.
6. $\text{Inc}(A, B \cap C) = \text{Inc}(A, B) \wedge \text{Inc}(A, C)$.
7. $\text{Inc}(A \cup B, C) = \text{Inc}(A, C) \wedge \text{Inc}(B, C)$.

Se puede apreciar que el axioma (SD2) solo se cumple si el universo de discurso es finito y que el axioma (SD6) relacionado con la contraposición de la medida de inclusión no aparece representado en el teorema anterior. Con respecto a este último axioma, recientemente en [6] se presentó la siguiente medida de inclusión basada en el f-índice de inclusión.

Definición 5: Sean A y B dos conjuntos difusos, la *medida de inclusión* de A en B inducida por el f-índice de inclusión se define como

$$M_{\text{Inc}}(A, B) = 2 \int_0^1 \text{Inc}(A, B)(x)dx.$$

Teorema 2 (Véase [6]): Sean A y B dos conjuntos difusos, entonces $M_{\text{Inc}}(A, B) = M_{\text{Inc}}(B^c, A^c)$.

En otras palabras, aunque la inclusión de A en B y la de B^c en A^c no vienen dados por el mismo f-indice de inclusión, "la cantidad" de inclusión de ambos f-índices sí es la misma. Por lo tanto, puede afirmarse que el f-índice de inclusión sí modela la idea tras el axioma (SD6).

III. El f-índice de inclusión y Modus Ponens

En [8] se ha relacionado el f-índice de inclusión con pares adjuntos de la siguiente forma:

Teorema 3: Sean A y B dos conjuntos difusos definidos sobre un universo finito $\mathcal{U}$. Entonces existe un par adjunto $(*, \to)$ tal que

$$Inc(A; B)(x) = x * \left(\bigwedge_{u\in\mathcal{U}} A(u) \to B(u) \right)$$

Obsérvese que la composición de los operadores de un par adjunto tal y como vienen dados en el resultado anterior es la regla de inferencia composicional de Zadeh . Como consecuencia, el f-índice de inclusión puede verse como un operador difuso de inferencia al estilo del Modus Ponens.

$$\begin{array}{lll} & A \Rightarrow B & \equiv \text{Inc}(A; B) \\ & A(u) & \equiv \ \beta \\ \hline \therefore & B(u) & \geq \text{Inc}(A; B)(\beta) \end{array} \tag{2}$$

De hecho, se puede afirmar que la inferencia realizada por el f-índice de inclusión sería la mayor de entre todas las posibles inferencias que podrían llevarse a cabo a través de pares adjuntos (ver [8]). Resultados iniciales apuntan a la existencia de un Modus Ponens Generalizado con todas las propiedades propuestas por Baldwin y Pilsworth; reinterpretando la simetría en términos de conexiones de Galois.

Conclusiones y detalles dejados en el tintero

Hemos visto en este breve resumen que hay indicios sólidos para afirmar que el f-índice de inclusión es un operador conveniente para representar la inclusión entre conjuntos difusos. Hemos especificado claramente en el resumen que el f-índice de inclusión es altamente coherente con los axiomas de Sinha-Dougherty y que es óptimo para la realización de la inferencia Modus Ponens. Sin embargo hay muchos detalles que no hemos podido incluir aquí sobre el f-índice de inclusión, como por ejemplo: su relación con el f-índice de contradicción [1] y el cuadrado de oposición resultante [5], las estructuras residuadas que pueden definirse sobre el conjunto de f-índice de inclusión o las medidas de entropía y similaridad inducidas [7].

Referencias

[1] H. Bustince, N. Madrid, and M. Ojeda-Aciego. *The Notion of Weak-Contradiction: Definition and Measures*. IEEE Trans. Fuzzy Systems, **23**(4) (2015), 1057–1069.

[2] L. M. Kitainik. *Fuzzy Inclusions and Fuzzy Dichotomous Decision Procedures*.Theory and Decision Library, 4 (1987) Springer Netherlands, Dordrecht, 154–170.

[3] N. Madrid and C. Cornelis. *Kitainik axioms do not characterize the class of inclusion measures based on contrapositive fuzzy implications* Fuzzy Sets and Systems (In press).

[4] N. Madrid and M. Ojeda-Aciego. *A view of f-indexes of inclusion under different axiomatic definitions of fuzzy inclusion*. Lecture Notes in Artificial Intelligence, **10564** (2017), 307–318.

[5] N. Madrid and M. Ojeda-Aciego. *On contradiction and inclusion using functional degrees*. Intl J. of Computational Intelligence Systems, **13**(1) (2020), 464–471.

[6] N. Madrid and M. Ojeda-Aciego. *Functional degrees of inclusion and similarity between L-fuzzy sets*. Fuzzy Sets and Systems, **390** (2020), 1–22.

[7] N. Madrid and M. Ojeda-Aciego. *Measures of inclusion and entropy based on the φ-index of inclusion*. Fuzzy Sets and Systems, **390** 423:29-54, 2021.

[8] N. Madrid and M. Ojeda-Aciego. *The f -index of inclusion as optimal adjoint pair for fuzzy modus ponens*. (Submitted)

[9] D. Sinha and E. R. Dougherty. *Fuzzification of set inclusion: Theory and applications*. Fuzzy Sets and Systems, **55**(1) (1993), 15–42.

Sobre la exactitud de los datos imprecisos. El caso de los problemas de decisión multicriterio

David A. Pelta, María T. Lamata, José L. Verdegay, Carlos Cruz
Grupo de Trabajo en Modelos de Decisión y Optimización
Depto. de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada, España
{dpelta, mtl, verdegay, carloscruz}@ugr.es

***Abstract*—La resolución de un problema donde existe información de naturaleza difusa (imprecisa), requiere primero decidir si dicha naturaleza difusa se va a considerar o no. Si la respuesta es que sí, entonces se debe elegir una manera adecuada para representar dicha información.**

Tomando como ejemplo los problemas de decisión multicriterio, analizamos que ocurre con las salidas (los rankings de las alternativas) si la imprecisión se considera o no. Supuesto que la imprecisión de la que hablamos es de naturaleza difusa, para representarla se utilizan números difusos triangulares, puesto que son los que menos datos requieren para ser definidos.

Se resuelve un conjunto de problemas de decisión multicriterio generados aleatoriamente, considerando o no el carácter difuso de los datos. A continuación, se comparan las salidas correspondientes y se estudian las variaciones en el ranking de la mejor alternativa detectada.

Los resultados muestran que los cambios en la mejor alternativa son menores. Esta situación, unida al hecho de que el ranking "verdadero" es desconocido, plantea el debate de si en este contexto específico tiene sentido la representación explícita del carácter difuso de los datos de entrada.

I. Introduction

Supongamos que se pide resolver un problema de decisión multicriterio (MCDM) donde los datos de entrada son imprecisos, y esta imprecisión es de naturaleza difusa. Usualmente, se modeliza el problema mediante una matriz de decisión $A^{m\times n}$, donde cada fila se asocia con una alternativa $\{A_1, A_2, \ldots, A_m\}$ y cada columna representa un criterio en consideración $\{C_1, C_2, \ldots, C_n\}$. Se denota el valor de una alternativa A_i en el criterio C_j como x_{ij}, valor que supondremos impreciso. Finalmente se asume que el decisor es capaz de reflejar la importancia relativa de los criterios mediante un vector de n valores o pesos $W = \{w_1, w_2, \ldots, w_n\}$.

En esta situación, se plantean dos cuestiones: 1) ¿ debe tenerse en cuenta el carácter difuso de los datos de entrada? y 2) si la respuesta es sí, entonces ¿cómo se elige una representación adecuada para los x_{ij}?.

En general, gran parte de la literatura asume que la respuesta a la primera pregunta es SI y se dedica un amplio esfuerzo al estudio de posibles representaciones de la información difusa.

Este trabajo ha sido financiado por los proyectos PID2020-112754GB-I0, MCIN/AEI /10.13039/501100011033 y FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto (BTIC-640-UGR20).

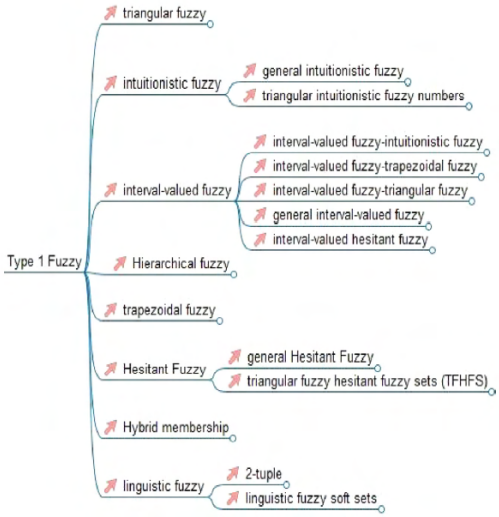


Fig. 1. Posibles representaciones de un número difuso (adaptada de [6]).

Si se supone que el valor de A_i en C_j es *aproximadamente* x_{ij}, entonces los números difusos [3] son herramientas adecuadas para modelizar ese tipo de imprecisión, dando lugar a los MCDM difusos [4]. La Figura 1 (adaptada de [6]) muestra posibles representaciones para un valor x_{ij}.

Un elemento a resaltar es que *para definir un dato difuso (impreciso), necesitamos disponer de un conjunto de valores absolutamente precisos*. Así, un *trapezoidal fuzzy number* requiere más información que un *triangular fuzzy number*, pero el primero requiere menos información que un *interval-valued fuzzy number*. Si se consideran los *type-2 fuzzy sets* la situación es aún peor.

Partiendo de resultados previos presentados en [1], [2], [5], en esta contribución y en el contexto de los problemas de decisión multicriterio, se pone el foco en la siguiente pregunta: *¿cambia sustancialmente la solución del problema si se considera o no el carácter difuso de los datos de entrada?*. En otras palabras, ¿cuáles son las consecuencias de ignorar la imprecisión?

A efectos de evitar confusiones, en esta contribución nos referimos estrictamente a la imprecisión derivada de la naturaleza difusa de los datos.

II. Experimentos

Para responder a las preguntas, se diseñó el siguiente experimento.

Se generaron 5400 problemas (matrices de decisión) al azar con imprecisión, donde cada x_{ij} se representó con un número difuso triangular simétrico. Esta representación se ha elegido puesto que es la que menos información requiere por parte del decisor.

Se consideraron diferentes números de alternativas, criterios y niveles de dispersión (límites izquierdo/derecho del número difuso). Para ver el caso donde se ignora la imprecisión (el carácter difuso de los datos), se derivaron 5400 problemas crisp equivalentes, utilizando solamente el valor central de cada número difuso. Se asignó el mismo peso a todos los criterios.

Cada problema k (con/sin imprecisión) se resolvió utilizando el TFN-TOPSIS (la adaptación para trabajar con números difusos triangulares) y el método TOPSIS "estandar", dando lugar a a dos rankings (r_k^1, r_k^2), respectivamente, que se deben comparar. Cada comparación puede dar lugar a una de las siguientes opciones:

1) *Sin cambio*: en ambos rankings, la mejor alternativa es la misma.
2) *Cambio menor*: en ambos rankings, las dos primeras alternativas son las mismas pero en orden diferente.
3) *Cambio mayor*: si no ocurren ninguna de las condiciones anteriores.

Finalmente, se suman los casos acumulados en cada opción (sobre 5400).

III. Resultados y Conclusiones

En 4788 (88%) casos la mejor alternativa fue la misma para los dos casos (considerando o no la imprecisión). En otras palabras, solo se produjeron cambios en la primer alternativa en 455 (9%) + 157 (3%) casos (números de casos con cambios menores y mayores). En los primeros, las dos primeras alternativas aparecieron intercambiadas, lo que indica que los correspondientes "scores" fueron muy similares. Se puede decir que no hay evidencias suficientes para considerar que una alternativa es mejor que otra.

Desde otro punto de vista, en el 97% de los casos no hubo cambios o fueron menores, lo cual implica que la mejor alternativa se pudo identificar independientemente de la consideración o no de la imprecisión en los datos de entrada. Solo se observaron cambios mayores en 157 casos (un 3%, sobre 5400 casos) . A medida que el número de alternativas se incrementó (de 5 a 50 a 100), el número de casos con cambios mayores se incrementó (del 1% al 3% y al 5%, respectivamente).

A medida que el número de criterios se incrementó (de 5 a 10 a 15), el número de casos con cambios mayores disminuyó (del 3.7% al 2.6% y al 2.4%, respectivamente).

Se observó que al incrementar el nivel de dispersión, también se reducía el número de casos con cambios mayores. Con el máximo nivel de dispersión, los casos con cambios mayores fueron solo del 2.4%.

Por tanto, *para el tipo de problema MCDM considerado y en el escenario de experimentación definido*, los resultados no justifican la modelización explícita del carácter difuso de los datos de entrada.

Aunque consideramos sólo un tipo de imprecisión que puede modelarse con números difusos triangulares simétricos, se debe analizar si el uso de extensiones más sofisticadas de dichos números puede proporcionar algún beneficio.

En el caso más sencillo de un número difuso triangular asimétrico, deben definirse *con precisión* tres valores por cada entrada de la matriz de decisión. Más allá de eso, la definición de un solo dato en la matriz de decisión utilizando una extensión más compleja de los números difusos, requerirá un *mayor conocimiento preciso* por parte del decisor, lo que parece una contradicción: el usuario necesita más información para definir un valor impreciso.

Todo esto plantea la necesidad de discutir si tratar con entradas más complejas, manteniendo la salida en los mismos términos (una permutación en nuestro contexto) es un enfoque razonable.

References

[1] B. Ceballos, M. T. Lamata, and D. A. Pelta. A comparative analysis of multi-criteria decision-making methods. *Progress in Artificial Intelligence*, 5(4):315–322, 2016.

[2] B. Ceballos, M. T. Lamata, and D. A. Pelta. Fuzzy multicriteria decision-making methods: A comparative analysis. *International Journal of Intelligent Systems*, 32(7):663—-753, 2017.

[3] D. Dubois and H. Prade. Fuzzy numbers: An overview. In J. Bezdek, editor, *Analysis of Fuzzy Information*, volume 2, pages 3–39. CRC Press, 1988.

[4] C. Kahraman, S. C. Onar, and B. Oztaysi. Fuzzy multicriteria decision-making: A literature review. *International Journal of Computational Intelligence Systems*, 8(4):637–666, 2015.

[5] D. A. Pelta, M. T. Lamata, J. L. Verdegay, C. Cruz, and A. Salas. Against artificial complexification: Crisp vs. fuzzy information in the TOPSIS method. In *Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and the 11th International Summer School on Aggregation Operators (AGOP)*. Atlantis Press, 2021.

[6] M. M. Salih, B. Zaidan, A. Zaidan, and M. A. Ahmed. Survey on fuzzy topsis state-of-the-art between 2007 and 2017. *Computers & Operations Research*, 104:207–227, 2019.

Some new Trends in Computable Aggregations

Luis Garmendia Salvador
Dep. de Ingenería del Software e Inteligencia Artificial
Universidad Complutense de Madrid
Madrid, Spain
lgarmend@fdi.ucm.es

Luis Magdalena
Dep. de Matemática Aplicada
Universidad Politécnica de Madrid
luis.magdalena@upm.es

Daniel Gómez González
Dep. de Estadística y Ciencia de los Datos
Universidad Complutense de Madrid
Madrid, Spain
dagomez@ucm.es

Francisco Javier Montero de Juan
Departamento de Estadística e Investigación Operativa
Universidad Complutense de Madrid
Madrid, Spain
monty@mat.ucm.es

***Abstract*—The concept of aggregation goes beyond functions on [0, 1] to the power of algorithms on lists of objects. New concepts as recursivity, non deterministic, map reducible or hierarchical can be studied for Computable Aggregations.**

Keywords—Computable aggregation, aggregation function, recursivity, non deterministic, map reducible, hierarchical

I. INTRODUCTION: AGGREGATION FUNCTIONS

Aggregation is a fundamental part of science. The process of aggregating the information is a key tool for most of the knowledge based systems. The concept of aggregation operator is extended to program to implement it, generalising the concept to new trends to explore.

Definition. An aggregation operator [1] is usually defined as a function

$$Ag: [0, 1]^n \rightarrow [0, 1]$$

where it is usually assumed the following properties

- Ag is monotonic, non decreasing.
- $Ag(0, \ldots, 0) = 0$
- $Ag(1, \ldots, 1) = 1$

II. COMPUTABLE AGGREGATION

Definition: Let L < T > be a finite and non-empty list of n elements with type T. A computable aggregation [8] is a program P that transform the list L < T > into an element of T, being T a complete lattice.

III. RECURSIVE COMPUTABLE AGGREGATIONS

Several recursive computable aggregations are studied giving a recursive function ϕ that operates the aggregation of a list of n values with a new value, without the need to look again to the list when one more value is added to recompute the aggregation. This means that an aggregation can be computed by running n times the recursive function.

Definition. A computable aggregation P is recursive if and only if there exists a sequence of binary programs defined only with lists of dimension two $\{\phi_n\}_{n>1}$

verifying:

$$P(L) = \begin{cases} 0 & \text{If } |L| = 0 \\ x_1 & \text{If } L = \{x_1\} \\ \phi_2(x_1, x_2) & \text{If } L = \{x_1, x_2\} \\ \vdots & \vdots \\ \phi_n(P(L \setminus \{x_n\}), x_n) & \text{If } L = \{x_1, \ldots, x_n\}. \end{cases}$$

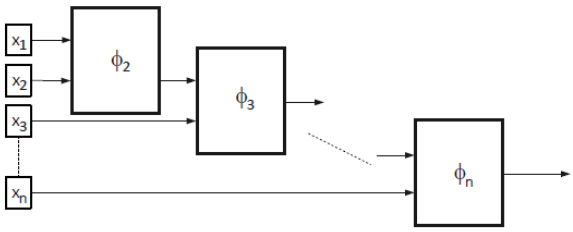


Fig. 1. Recursive (Hierarchical) Computable Aggregation

IV. PARALLER COMPUTABLE AGGREGATIONS

Given a hierarchical structure such as a tree and some computable aggregations for its levels, we can define the concept of tree hierarchical computable aggregation depending on the depth of the tree.

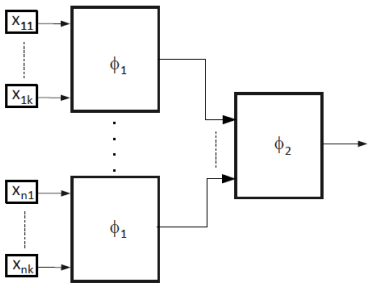


Fig. 2. 3-level tree (Hierarchical) Computable Aggregation

V. PRIORIZED COMPUTABLE AGGREGATIONS

The concept of prioritized aggregation operators is presented by Yager in [10], where he considers the existence of many situations in which a decision process involving the aggregation of several criteria, requires those criteria to be ordered and

grouped to establish a prioritization in the aggregation process.

Here it is important to note that if we have a motorbike with low value in safety, most probably it is not even necessary to evaluate the criteria related with the cost, since the global value of the aggregation should be low.

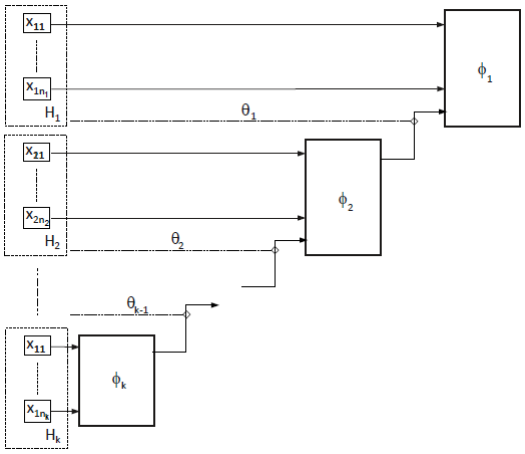


Fig. 3. Prioritized (Hierarchical) Computable Aggregation

VI. NON DETERMINISTIC COMPUTABLE AGGREGATIONS

Definition. A computable aggregation P over the set T is a non-deterministic computable aggregation if and only if the program implementing it is non deterministic.

Definition. (Empirical distribution of a computable aggregation). Given a computable aggregation P, and given a list L1 ∈ L < T >, the distribution of results obtained after m executions of P over L_1 represented by $LP^m(L_1)$

Example. Sample Mean distribution

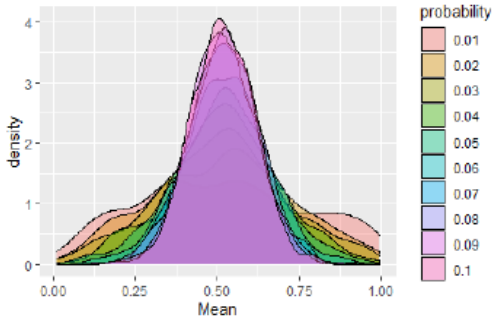


Fig. 4. $LP^{1000}_{Mean,p}(L)$, L= [0.1, 0.4, 0.2, 0.2, 0.9, 1, 0.8, 0.2, 0.7, 0.5].

Definition. Strong ≤-monotonicity of non deterministic computable aggregations).

A non-deterministic computable aggregation P isstrong ≤-monotone if and only if, when L1 ≤ L2 then DPL1m ≤ DPL2m any m ≥ 1.

Definition. (Max-Min preorder). A list L1 in L is Max-Min lower or equal than a list L2 in L, if and only if the highest degree in L1 is lower or equal than the minimum value of L2.

Proposition. The Probability sampled and k-sampled computable aggregations are Strong $\leq_{MaxMin}$ - Monotonic.

Definition. (Theoretical distribution of a computable aggregation). Given a computable aggregation P that aggregates a list L1 ∈ L < T >, let us denote by $P(L_1)$ the theoretical distribution after all possible realizations of P over the fixed list L_1.

Proposition. P is population asymptotic monotone if P is $\leq_{S1to1}$-asymptotic monotone.

Proposition. $P_{Ag,k}$ computable aggregation is population monotone

ACKNOWLEDGMENT

This research has been partially supported by the Government of Spain (grant PGC2018-096509-B-I00 and grant PID2020-112502RB / AEI / 10.13039/501100011033)

REFERENCES

[1] G. Beliakov, A. Pradera, and T. Calvo. Aggregations Functions: A guide for Practitioners. Springer. 2007.

[2] R. Gonzalez-del-Campo, L. Garmendia, and J. Montero, "Complexity of increasing phi-recursive computable aggregations," in XIX Congreso Espanol sobre Tecnologías y Logica Fuzzy (CAEPIA-ESTYLF), pp. 187–191. 2018.

[3] L. Garmendia, D. Gómez, L. Magdalena, J. Montero Analyzing Non-deterministic Computable Aggregations, pp. 551-564. 2020.

[4] L. Garmendia, D. Gómez, L. Magdalena, J. Montero Empirical Monotonicity of Non-deterministic Computable Aggregations. AGOP. 2021.

[5] L. Magdalena, L. Garmendia, D. Gomez, R. G. del Campo, J. T. Rodríguez, and J. Montero, "Types of recursive computable aggregations," FUZZ-IEEE, pp. 1–6. 2019.

[6] L. Magdalena, D. Gómez, L. Garmendia, J. Montero. Population Monotonicity of Non-deterministic Computable Aggregations. FUZZ-IEEE. 2021.

[7] L. Magdalena, D. Gómez, L. Garmendia, J. Montero Analysing Monotonicity in Non-deterministic Computable Aggregations: the probabilistic case. Information Sciences. 2022.

[8] J. Montero, R. G. del Campo, L. Garmendia, D. Gomez, and J. Rodríguez, "Computable aggregations," Information Sciences, vol. 460, pp. 439–449. 2018.

[9] Ronald R. Yager. Families of OWA operators, Fuzzy Sets and Systems, 59: 125-148. 1993.

[10] R. R. Yager, "Prioritized aggregation operators," International Journal of Approximate Reasoning, vol. 48, no. 1, pp. 263–274. 2008.

Monotonía direccional de funciones intervalo-valuadas: Diferencias entre orden parcial y órdenes totales

Mikel Sesma-Sara
Institute of Smart Cities,
Public University of Navarra
Pamplona, Spain
mikel.sesma@unavarra.es

Radko Mesiar
Faculty of Civil Engineering,
Slovak University of Technology
Bratislava, Slovakia
mesiar@math.sk

Humberto Bustince
Institute of Smart Cities
Public University of Navarra
Pamplona, Spain
bustince@unavarra.es

***Abstract*—La monotonía direccional surge como una relajación de la condición de monotonia exigida a las funciones de agregación. En el caso de extender el dominio de este tipo de funciones al caso intervalar, existe la adaptación de la condición de monotonía direccional al ámbito de las fuciones intervalo-valuadas. Sin embargo, esta condición se basa en el orden parcial entre intervalos. En este trabajo, exploramos la posibilidad de definir el concepto de monotonía direccional para funciones intervalo-valuadas con respecto a un orden admisible. Estudiamos qué propiedades se mantienen y cuáles dejan de ser válidas.**

***Index Terms*—Función de agregación, monotonía direccional, función intervalo-valuada, orden admisible**

I. Introducción

El objeto de una función de agregación es representar una tupla de n números mediante un único valor. Este tipo de funciones satisfacen dos condiciones de contorno y una condición de monotonía y han sido estudiadas tanto desde el punto de vista teórico como práctico [1], [2].

Una de las tendencias de investigación en la teoría de las funciones de agregación es extender el tipo de información a agregar. Por ejemplo, para agregar intervalos [3], [4], que presentan el problema de que el orden natural en este contexto no es total, sino parcial.

Otra tendencia en el ámbito de las funciones de agregación es relajar la condición de monotonía. Existen varias formas de monotonía relajada [5]–[8], muchas de las cuales están basadas en el concepto de monotonía direccional [6].

Existen trabajos aunando ambas tendencias. En particular, se ha definido el concepto de monotonía direccional para funciones intervalo-valuadas [9], [10]. Sin embargo, solo se ha hecho teniendo en cuenta el orden parcial entre intervalos.

En este trabajo, estudiamos cómo afecta el hecho de considerar un orden total entre intervalos en la condición de monotonía direccional, concluyendo que ciertas propiedades relevantes no se conservan.

II. Preliminares

Denotamos por $L([0,1])$ el conjunto de intervalos cerrados en $[0,1]$. El orden natural de este conjunto, considerando el orden producto, es parcial. Este orden está definido de la siguiente manera: $[a,b] \leq_L [c,d]$ si $a \leq c$ y $b \leq d$, para cualesquiera $[a,b],\ [c,d] \in L([0,1])$.

En este trabajo, nos centramos en funciones $F : L([0,1])^n \to L([0,1])$, donde el concepto de monotonía viene determinado por el orden parcial. Es decir, todos los argumentos han de ser monótonos con respecto a $\leq_L$.

Para definir el concepto de monotonía direccional para funciones cuyo dominio es $L([0,1])^n$, se considera que las direcciones son vectores pertenecientes al espacio vectorial ordenado $(\mathbb{R}^2)^n$ [9]. Por lo tanto, se define la monotonía direccional de la siguiente manera.

Definición 1: Sea $\mathbf{v} = (a_1, b_1, \ldots, a_n, b_n) \in (\mathbb{R}^2)^n$ tal que $(a_i, b_i) \neq \vec{0}$ para algún $i \in \{1, \ldots, n\}$. Una función $F : L([0,1])^n \to L([0,1])$ es $\mathbf{v}$-creciente si para todo $\mathbf{x} \in L([0,1])^n$ y $c > 0$ tales que $\mathbf{x} + c\mathbf{v} \in L([0,1])^n$, se cumple que $F(\mathbf{x}) \leq_L F(\mathbf{x} + c\mathbf{v})$.

En la Definición 1 se define la monotonía con respecto al orden parcial del conjunto de intervalos. Para adaptar este concepto al caso de órdenes totales, conviene recordar la noción de orden admisible. Un orden admisible [11] es un orden total que extiende el orden parcial entre intervalos. Es decir, una relación de orden $\preceq$ en $L([0,1])$ es admisible $\preceq$ si es una relación de orden total en $L([0,1])$, y para todos $[a,b],\ [c,d] \in L([0,1])$, si $[a,b] \leq_L [c,d]$ entonces $[a,b] \preceq [c,d]$.

Nótese que la elección del intervalo $[0,1]$ proviene del contexto de las funciones de agregación, pero todos los resultados son extensibles a cualquier otro intervalo cerrado.

III. Propiedades de la monotonía direccional con respecto al orden parcial

Entre las propiedades de las funciones intervalo-valuadas que son direccionalmente monótonas con respecto al orden parcial, encontramos las siguientes.

Por un lado, el conjunto de direcciones para la cuales una función es direccionalmente creciente es cerrado con respecto a combinaciones lineales de dichas direcciones.

Teorema 1: Sean $a, b > 0$ y $\vec{v}, \vec{u} \in (\mathbb{R}^2)^n \backslash \{\vec{0}\}$ tales que para todos $\mathbf{x} \in L([0,1])^n$ y $c > 0$ que cumplen $\mathbf{x} + c(a\vec{v} + b\vec{u}) \in L([0,1])^n$, se satisface que o bien $\mathbf{x} + ca\vec{v} \in L([0,1])^n$ o $\mathbf{x} + cb\vec{u} \in L([0,1])^n$. Entonces, si una función $F : L([0,1])^n \to L([0,1])$ es $\mathbf{v}$-creciente y también $\mathbf{u}$-creciente, entonces F es $(a\vec{v} + b\vec{u})$-creciente.

Esto permite que podamos caracterizar la monotonía estándar estudiando la monotonía direccional para funciones $F : L([0,1])^n \to L([0,1])$. Para ello, consideramos $\{\vec{e}_i\}_{i=1}^{2n}$, la base canónica de $(\mathbb{R}^2)^n$.

Teorema 2: Sean $F : L([0,1])^n \to L([0,1])$ y $\{\vec{e}_i\}_{i=1}^{2n}$ la base canónica de $(\mathbb{R}^2)^n$. Así, F es creciente si y solo si F es $\vec{e}_i$-creciente para todo $i \in \{1, \ldots, 2n\}$.

Otra propiedad importante es que podemos construir funciones intervalo-valuadas crecientes con respecto a una dirección cualquiera mediante funciones $f : [0,1]^n \to [0,1]$ direccionalmente monótonas. Para ello, utilizamos las funciones intervalo-valuadas representables [12].

Definición 2: Una función $F : L([0,1])^n \to L([0,1])$ es representable si existen dos funciones $f, g : [0,1]^n \to [0,1]$ que cumplan $f(x_1, \ldots, x_n) \leq g(y_1, \ldots, y_n)$, siempre que $x_i \leq y_i$ para todo $i \in \{1, \ldots, n\}$ y que cumplan

$$F([\underline{x_1}, \overline{x_1}], \ldots, [\underline{x_n}, \overline{x_n}]) = [f(\underline{x_1}, \ldots, \underline{x_n}), g(\overline{x_1}, \ldots, \overline{x_n})].$$

Las direcciones de monotonía de una función representable están totalmente determinadas por las direcciones de crecimiento de sus funciones f y g.

Teorema 3: Sea $F : L([0,1])^n \to L([0,1])$ una función representable con respecto a $f, g : [0,1]^n \to [0,1]$. Sean $\vec{a} = (a_1, \ldots, a_n)$ y $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ tales que $\vec{a}, \vec{b} \neq \vec{0}$. Entonces, F es $((a_1, b_1), \ldots, (a_n, b_n))$-creciente si y solo si f es $\vec{a}$-creciente y g es $\vec{b}$-creciente.

IV. Monotonía direccional con respecto a un orden admisible

En esta sección presentamos el concepto de monotonía direccional para el caso en que tenemos un orden admisible en $L([0,1])$ en lugar del orden parcial.

Definición 3: Sea $\mathbf{v} = (a_1, b_1, \ldots, a_n, b_n) \in (\mathbb{R}^2)^n$ tal que $(a_i, b_i) \neq \vec{0}$ para algún $i \in \{1, \ldots, n\}$ y sea $\preceq$ un orden admisible en $L([0,1])$. Una función $F : L([0,1])^n \to L([0,1])$ es $\mathbf{v}$-creciente con respecto a $\preceq$ si para todo $\mathbf{x} \in L([0,1])^n$ y $c > 0$ tales que $\mathbf{x} + c\mathbf{v} \in L([0,1])^n$, se cumple que $F(\mathbf{x}) \preceq F(\mathbf{x} + c\mathbf{v})$.

El hecho de considerar órdenes admisibles influye en las propiedades que cumplen las funciones que cumplan la condición de la Definición 3.

En el caso de la monotonía con respecto a la dirección correspondiente a la combinación de dos direcciones (Teorema 1) se cumple también en este contexto considerando un orden admisible.

Sin embargo, los Teoremas 2 y 3 no se cumplen para funciones direccionalmente monótonas con respecto a un orden admisible.

En concreto, solo podemos garantizar una de las implicaciones del *si y solo si*.

Teorema 4: Sean $F : L([0,1])^n \to L([0,1])$, $\preceq$ un orden admisible en $L([0,1])$ y $\{\vec{e}_i\}_{i=1}^{2n}$ la base canónica de $(\mathbb{R}^2)^n$. Si F es creciente con respecto a $\preceq$, entonces F es $\vec{e}_i$-creciente con respecto a $\preceq$ para todo $i \in \{1, \ldots, 2n\}$.

Teorema 5: Sean $F : L([0,1])^n \to L([0,1])$ una función representable con respecto a $f, g : [0,1]^n \to [0,1]$ y $\preceq$ un orden admisible en $L([0,1])$. Sean $\vec{a} = (a_1, \ldots, a_n)$ y $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ tales que $\vec{a}, \vec{b} \neq \vec{0}$. Si F es $((a_1, b_1), \ldots, (a_n, b_n))$-creciente, entonces f es $\vec{a}$-creciente y g es $\vec{b}$-creciente.

V. Conclusiones

Hemos introducido el concepto de monotonía direccional para funciones intervalo-valuadas en un contexto de orden total. El estudio de este tipo de monotonía con respecto a un orden admisible pierde varias de las propiedadades más interesantes que se cumplen en el caso del orden parcial.

Agradecimientos

Financiación de MCIN, proyecto PID2019-108392GB-I00/AEI/10.13039/501100011033 y APVV-18-0052 (Slovak Research and Development Agency).

References

[1] M. Grabisch, J. Marichal, R. Mesiar, and E. Pap, *Aggregation functions*. Cambridge University Press, 2009.

[2] D. Paternain, J. Fernandez, H. Bustince, R. Mesiar, and G. Beliakov, "Construction of image reduction operators using averaging aggregation functions," *Fuzzy Sets and Systems*, vol. 261, pp. 87–111, 2015.

[3] G. Beliakov, H. Bustince, S. James, T. Calvo, and J. Fernandez, "Aggregation for Atanassov's intuitionistic and interval valued fuzzy sets: The median operator," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 3, pp. 487–498, 2012.

[4] G. Deschrijver, "Generalized arithmetic operators and their relationship to t-norms in interval-valued fuzzy set theory," *Fuzzy Sets and Systems*, vol. 160, no. 21, pp. 3080–3102, 2009.

[5] T. Wilkin and G. Beliakov, "Weakly monotonic averaging functions," *International Journal of Intelligent Systems*, vol. 30, no. 2, pp. 144–169, 2015.

[6] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar, "Directional monotonicity of fusion functions," *European Journal of Operational Research*, vol. 244, no. 1, pp. 300–308, 2015.

[7] G. Beliakov, T. Calvo, and T. Wilkin, "Three types of monotonicity of averaging functions," *Knowledge-Based Systems*, vol. 72, pp. 114–122, 2014.

[8] M. Sesma-Sara, J. Lafuente, A. Roldán, R. Mesiar, and H. Bustince, "Strengthened ordered directionally monotone functions. Links between the different notions of monotonicity," *Fuzzy Sets and Systems*, vol. 357, pp. 151–172, 2019.

[9] M. Sesma-Sara, L. De Miguel, R. Mesiar, J. Fernandez, and H. Bustince, "Interval-valued pre-aggregation functions: a study of directional monotonicity or interval-valued functions," in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2019, pp. 1–6.

[10] M. Sesma-Sara, R. Mesiar, and H. Bustince, "Weak and directional monotonicity of functions on Riesz spaces to fuse uncertain data," *Fuzzy Sets and Systems*, vol. 386, pp. 145–160, 2020.

[11] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar, "Generation of linear orders for intervals by means of aggregation functions," *Fuzzy Sets and Systems*, vol. 220, pp. 69–77, 2013.

[12] G. Deschrijver and C. Cornelis, "Representability in interval-valued fuzzy set theory," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 15, no. 3, pp. 345–361, 2007.

Pertenencias difusas para modelar el contexto en un cuadro

1er Javier Fumanal-Idocin
Estadística, Informática, Matemáticas
Universidad Pública de Navarra
Pamplona, España
javier.fumanal@unavarra.es

2º Humberto Bustince
Estadística, Informática, Matemáticas
Universidad Pública de Navarra
Pamplona, España
bustince@unavarra.es

3º Óscar Cordón
DeCSAI
Universidad de Granada
Granda, España
ocordon@decsai.ugr.es

***Abstract*—El análisis automático de arte consiste en utilizar diversos métodos de procesamiento para clasificar y categorizar obras de arte [1], [2]. A la hora de trabajar con este tipo de imágenes, debemos tener en cuenta consideraciones diferentes a las del manejo clásico de imágenes, ya que las obras de arte se modifican definitivamente según el autor, la escena que se delinea o su forma estética [3], [4]. Estos datos adicionales mejoran las señales visuales obtenidas de las imágenes y pueden conducir a un mejor rendimiento. Sin embargo, esta información debe modelarse e integrarse junto con las características visuales de la imagen. Esto a menudo se realiza utilizando modelos de aprendizaje profundo, pero son costosos de entrenar [3]–[5]. En este artículo, utilizamos el algoritmo Fuzzy C-Means para crear una estrategia de incrustación basada en membresías difusas para extraer información relevante de los grupos presentes en la información contextual. Ampliamos un sistema de clasificación estado del arte existente utilizando esta estrategia para obtener una nueva versión que presenta resultados similares o mejores sin entrenar modelos de aprendizaje profundo adicionales.**

I. Agradecimientos

El trabajo de Javier Fumanal Idocin y Humberto Bustince ha sido financiado por el proyecto PID2019-108392GB-I00 (AEI/10.13039/ 501100011033).

El trabajo de Oscar Cordón ha sido financiado por el gobierno de España, EXASOCO (PGC2018-101216-B-I00), incluyendo fondos de desarrollo regional europeo (ERDF).

References

[1] Noa Garcia and George Vogiatzis. How to read paintings: semantic art understanding with multi-modal retrieval. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, pages 0–0, 2018.

[2] Giovanna Castellano, Vincenzo Digeno, Giovanni Sansaro, and Gennaro Vessio. Leveraging knowledge graphs and deep learning for automatic art analysis. *Knowledge-Based Systems*, page 108859, 2022.

[3] Noa Garcia, Benjamin Renoust, and Yuta Nakashima. Context-aware embeddings for automatic art analysis. In *Proceedings of the 2019 on International Conference on Multimedia Retrieval*, pages 25–33, 2019.

[4] Cheikh Brahim El Vaigh, Noa Garcia, Benjamin Renoust, Chenhui Chu, Yuta Nakashima, and Hajime Nagahara. Gcnboost: Artwork classification by label propagation through a knowledge graph. *arXiv preprint arXiv:2105.11852*, 2021.

[5] Bingqing Guo and Pengwei Hao. Analysis of artistic styles in oil painting using deep-learning features. In *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pages 1–4. IEEE, 2020.

Identify applicable funding agency here. If none, delete this.

Agregación y definición de un marco algebraico para series temporales difusas: Una aplicación en el ámbito de la oferta y la demanda*

1st Luis Rodriguez-Benitez
Dep. de Tecnologías y Sistemas de Información
Escuela Superior de Informática
Universidad de Castilla-La Mancha
Ciudad Real, España
Luis.Rodriguez@uclm.es

2nd Juan Moreno-Garcia
Dep. de Tecnologías y Sistemas de Información
Escuela de Ingeniería Industrial y Aeroespacial
Universidad de Castilla-La Mancha
Toledo, España
juan.moreno@uclm.es

3rd Ester del Castillo-Herrera
Dep. y Sistemas de Información
Escuela Superior de Informática
Universidad de Castilla-La Mancha
Ciudad Real, España
ester.castillo@uclm.es

4th Jun Liu
Artificial Intelligence Research Centre
School of Computing
University of Ulster
Northern Ireland, UK
j.liu@ulster.ac.uk

5th Luis Jimenez-Linares
Dep. y Sistemas de Información
Escuela Superior de Informática
Universidad de Castilla-La Mancha
Ciudad Real, España
luis.jimenez@uclm.es

***Abstract*—Se propone un método para agregar la información contenida en conjuntos de series temporales (TS) en una Fuzzy Time Series (FTS). Hace uso de la técnica de agregación KDE para obtener la FTS a partir de las TS de entrada. Se obtiene un nuevo FTS que permite obtener información más rica y operar en condiciones de incertidumbre. Finalmente, se introducen también las operaciones necesarias para calcular la función de pertenencia de cualquier TS en el FTS agregado.**

***Index Terms*—Fuzzy Time Series, agregación de series de tiempo, estimación de la densidad del núcleo**

I. Introducción

Una TS es un conjunto de observaciones regulares ordenadas en el tiempo de una característica cuantitativa de un fenómeno medida en puntos sucesivos del tiempo. Song y Chissom [1] sentaron las bases de las FTS que se han utilizado ámpliamente en los últimos años. Los valores FTS son conjuntos difusos [2], [3]. El uso de conjuntos difusos para la modelización de series temporales permite no sólo incorporar la capacidad de los modelos difusos para aproximar funciones, sino también añadir la legibilidad asociada a las variables lingüísticas que las hace más accesibles al análisis de los no expertos en el campo de aplicación. En la propuesta presentada los conjuntos de TS de entrada se resumen en un único FTS basándose en el concepto de agregación utilizando KDE [4], [5]. También se establece un marco para la operación entre ellos utilizando el principio de extensión de Zadeh [2] para construir operaciones entre FTS. Finalmente, se define el cálculo de la función de pertenencia de cualquier TS a la FTS agregada. Como caso de uso, se aplica a series con datos de oferta y demanda de un producto, concretamente en la energía.

A los proyectos TRA2016-76914-C3-2-P y PID2020-112967GB-C32, PID2020-112967GB-C33 financiadas por MCIN/AEI/ 10.13039/501100011033 y por ERDF A way of making Europe.

II. Agregación de TS y operaciones entre FTS

Como exponen Song y Chissom en [1], la principal diferencia entre las TS convencionales y las FTS consiste en que las observaciones en las primeros son números reales mientras que en la de las segundas son conjuntos difusos. Una FTS, denotada por $F(t)$, puede definirse como una colección de conjuntos difusos $f_i(t)$ sobre el universo del discurso $Y(t)$, siendo $Y(t)$ un subconjunto de $\mathbb{R}$, $t = (0, 1, 2, \ldots)$.

A. Agregación de un conjunto de TS en una FTS

El primer objetivo de este trabajo consiste en agregar un conjunto de TS en una FTS. Una colección de TS se representa como $S = \{s_0, \ldots, s_{n-1}\}$, donde cada TS se formaliza como $s_i = \{v_{(i,t_0)}, v_{(i,t_1)}, \ldots, v_{(i,t_{m-1})}\}$ donde t_j representa el tiempo de cada observación. Como ejemplo asociado al caso de uso desarrollado posteriormente en este trabajo, el conjunto S podría representar la colección del consumo eléctrico horario de un edificio a lo largo de 24 horas (varios jueves). El intervalo de tiempo real respecto a una correspondencia con la actividad de consumo, de modo que se considera que el día comienza a las 5 de la mañana al mismo tiempo que la actividad en el edificio. La Figura 1 muestra 19 TS correspondientes a 19 jueves pertenecientes a las semanas laborales del 06/02/2014 al 03/07/2014, donde $t_0 = 5$ a.m. y $t_{23} = 4$ a.m.

Sea el dominio $Y(t) = \{v_{(0,t)}, \ldots, v_{(n-1,t)}\}$ con $t \in (t_0,\ldots,t_{m-1})$ y $DE_t(v)$ una función de densidad que estima la probabilidad de que v pertenezca a $Y(t)$. El conjunto

difuso f_t sobre el dominio $Y(t)$ se define como la función de pertenencia $\mu_t(v)$:

$$\mu_t(v) = \frac{DE_t(v)}{max\{DE_t(v)|v \in Y(t)\}} \quad (1)$$

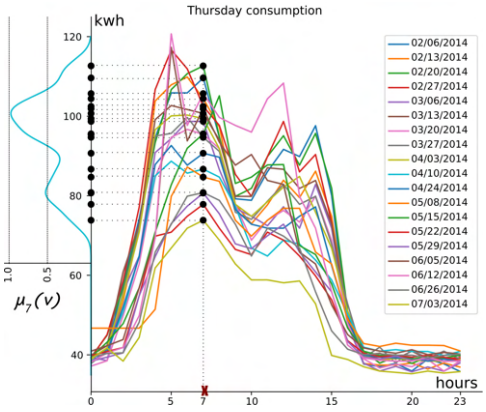


Fig. 1. Conjunto de consumo de los jueves con el conjunto difuso f_7 y su función de pertenencia $\mu_7(v)$

En este trabajo se emplea la técnica KDE [4], [5] con funciones kernel gaussianas para obtener las funciones de densidad DE_t. Por ejemplo, en la Figura 1 se muestra cómo se obtiene el conjunto difuso f_7 dentro de su función de pertenencia $\mu_7(v)$ siendo el dominio $Y(7)$=(109,63, 104,4 , 112. 65, 102.47, 98.69, 100.69, 95.42, 94.68, 99.4, 86.68, 90.68, 84.69, 95.56, 77.84, 80.68, 101.89, 105.68, 80.8 , 73.84) que es el consumo en el instante t_7 para las 19 TS.

Ahora, una vez que se ha obtenido un conjunto difuso f_t para cada instante t, se puede definir la FTS sobre la colección S: $FTS(S) = \{f_{t_0}, f_{t_1}, \ldots, f_{t_{m-1}}\}$.

Esta definición se extiende a un conjunto arbitrario T con $t \subset T$ tal que si $r \in T$, su función de pertenencia f_r es

$$\mu_r(v) = \begin{cases} \frac{DE_r(v)}{max\{DE_r(v)|v \in Y(r)\}} & r \in t \\ \mu_{t_{down}}(v)(1-\alpha) + \alpha\mu_{t_{up}}(v) & t_0 \leq r \leq t_{m-1}, r \notin t \\ 0 & r < t_0 \ or \ r > t_m \end{cases} \quad (2)$$

con $t_{down} = max_j\{j < r | j \in t\}$, $t_{up} = min_i\{i > r | i \in t\}$ y $\alpha = \frac{r - t_{down}}{t_{up} - t_{down}}$.

En resumen, el grado de pertenencia es 0 si r está fuera del rango de valores de t, el valor del conjunto difuso si r está en t y la interpolación lineal entre dos valores de t si r está entre ellos. Como ejemplo, la Figura 2 muestra la serie temporal del consumo de electricidad para 19 jueves laborables, donde el color y la anchura de las líneas verticales indican el grado de pertenencia de los conjuntos difusos f_t.

B. Definición de las operaciones con FTS

Sean A y B dos FTS con funciones de pertenencia $\mu_A(t,x),\ x \in D_x$ y $\mu_B(t,y),\ y \in D_y$ y $T : D_x \times D_y :\rightarrow D_z$ sea una función definida sobre el dominio $D_x \times D_y$ con valores en $D_z = T^{-1}(x,y)$. La extensión de la función T al dominio

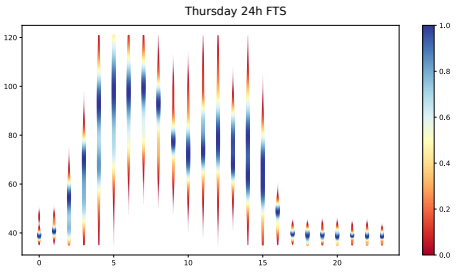


Fig. 2. FTS resultado de la agregación de las TS con el consumo de los 19 jueves

de las FTS se realiza definiendo $\tilde{T}(A,B) = Z$, donde Z es una nueva FTS con función de pertenencia $\mu_Z(t,z),\ z \in D_z$ que sigue el principio de extensión propuesto por Zadeh, tal que:

$$\mu_Z(t,z) = \bigvee_{x \in D_x, y \in D_y} \{\mu_A(t,x) \bigwedge \mu_B(t,y) | T(x,y) = z\} \quad (3)$$

donde $\bigvee$ es una t-conorma y $\bigwedge$ una t-norma.

C. Función de pertenencia de una serie temporal en una FTS

Estamos interesados en conocer si una TS puede ser considerada como la serie prototipo del conjunto de series agregadas. Para ello, es necesario definir una función de pertenencia de una TS a la FTS resultante de la agregación. Sea $s' = \{v_{t_0}, \ldots, v_{t_{m-1}}\}$ una TS definida en el mismo dominio que el conjunto S. Entonces el grado medio de pertenencia de esta serie a la $FTS(S)$ se define como sigue:

$$\mu_{FTS(S)}(s') = \frac{1}{m} \sum_{i=0}^{m-1} \mu_{t_i}(v_{t_i}) \quad (4)$$

III. Conclusiones y trabajos futuros

Podemos concluir que la técnica que se ha desarrollado permite trabajar a diferentes niveles con FTS. En primer lugar, se ha propuesto un método para la agregación de series temporales en un FTS. Posteriormente, se ha definido un marco para realizar operaciones de FTS. También se presenta la definición de una función para calcular la pertenencia de una TS a una FTS. Como trabajo futuro estamos considerando la obtención de descripciones lingüísticas de estas FTS y su combinación mediante operadores lingüísticos, donde la temporalidad debe jugar un papel importante.

References

[1] Qiang Song, Brad S. Chissom, "Fuzzy time series and its models,", Fuzzy Sets and Systems, vol. 54(3), 1993.
[2] L.A. Zadeh, "Fuzzy sets,", Information and Control, vol. 8(3), 338–353, 1965.
[3] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I,", Information Sciences, vol. 8(3), 1975.
[4] Emanuel Parzen, "On Estimation of a Probability Density Function and Mode,", The Annals of Mathematical Statistics, vol. 33(3), 1962.
[5] Matej Kristan, Aleš Leonardis, Danijel Skočaj, "Multivariate online kernel density estimation with Gaussian kernels,", Pattern Recognition, vol. 44(10-11), 2011.

Evaluación de diferentes funciones de agregación en la capa de pooling de una Red Neuronal Convolucional

1° Pablo Ursúa-Medrano
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
pablo.ursua@unavarra.es

2° Iosu Rodriguez-Martinez
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
iosu.rodriguez@unavarra.es

3° Angela Bernardini
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
angela.bernardini@unavarra.es

4° Javier Fernandez
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
fcojavier.fernandez@unavarra.es

5° Humberto Bustince
dept. de Estadística, Informática y Matemáticas
Universidad Pública de Navarra
Pamplona, España
bustince@unavarra.es

***Abstract*—Las redes neuronales convolucionales (CNN) han tenido mucho éxito en tareas de procesamiento de imágenes, ya que son muy útiles para extraer automáticamente características locales de las imágenes [1]. La arquitectura de los modelos CNN está formada por capas que tratan la información de entrada de forma consecutiva, donde cada una de ellas juega un papel específico para realizar este proceso de extracción de características. Estas características se utilizan posteriormente para resolver diferentes problemas, como la clasificación [2], [3] o la segmentación semántica [4].**

En este trabajo nos centraremos en una de las capas que las CNNs implementan en su arquitectura. En concreto, la capa de pooling. Esta capa se centra en la tarea de agregar los datos que la operación de extracción de características, conocida como convolución, ha generado. Esta capa agrega los valores de cada ventana disjunta de datos de la imagen, reduciendo la dimensionalidad de las características que el modelo está extrayendo. Habitualmente estos valores se agregan utilizando funciones como la media aritmética o el máximo, a pesar de algunas pruebas para su sustitución [5], [6].

Nuestro objetivo con este trabajo es explorar la posibilidad de sustituir estas funciones por otras bien conocidas de la teoría de la agregación. En particular, intentaremos utilizar integrales difusas, dada su capacidad para modelar posibles coaliciones entre los datos a agregar [7], fuciones de grouping,otros estadísticos de orden y combinaciones lineales de todas estas funciones [8], entre otras.

AGRADECIMIENTOS

Grant PID2019-108392GB-I00/ funded by MCIN/AEI/10.13039/501100011033. Pablo Ursúa and I. Rodríguez-Martínez are partially supported by the project PC095-096 FUSIPROD. I. Rodríguez-Martínez is also partially supported by Tracasa Instrumental (iTRACASA) and Gobierno de Navarra - Departamento de Universidad, Innovación y Transformación Digital.

REFERENCIAS

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[5] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," *arXiv preprint arXiv:1301.3557*, 2013.

[6] C.-Y. Lee, P. W. Gallagher, and Z. Tu, "Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree," in *Artificial intelligence and statistics*. PMLR, 2016, pp. 464–472.

[7] G. Beliakov, H. B. Sola, and T. C. Sánchez, *A practical guide to averaging functions*. Springer, 2016.

[8] I. Rodríguez-Martínez, J. Lafuente, R. H. Santiago, G. P. Dimuro, F. Herrera, and H. Bustince, "Replacing pooling functions in convolutional neural networks by linear combinations of increasing functions," *Neural Networks*, 2022.

Group Definition Based on Flow in Community Detection. A Review

María Barroso[1,3], Inmaculada Gutiérrez[1], Daniel Gómez[1,2], Javier Castro[1,2] and Rosa Espínola[1,2]
[1]*Faculty of Statistics, Complutense University Puerta de Hierro, s/n, 28040 Madrid, Spain*
[2]*Instituto de Evaluación Sanitaria. Complutense University*
[3]E-mail: mbarro10@ucm.es

***Abstract*—The community detection problem is one of the hottest topic in the field of social network analyses. Although most real-life problems can be modelled using directed networks, the vast majority of methods in the literature are devoted to non-directed networks, many of them based on modularity. In this research, a new definition of group based on flow is developed, using a fuzzy measure in order to add additional information. Adapting the Louvain Algorithm, it is possible to find not only dense communities, but also well communicated structures. To address this problem about community detection with additional information, a new fuzzy measure which models the flow capacity of the network is defined. Then, we propose a novel algorithm which can deal with the interactions among individuals.**

***Index Terms*—Community detection, Directed Network, Flow, Fuzzy Measures, Louvain Algorithm, Flow Capacity Louvain**

I. Introduction

In this review, we summarise our research in [1], in which we introduced a novelty algorithm in the field of community detection problem. The definition of what a community is makes an important issue in those analyses, which turns particularly challenging in directed networks. We consider the classical flow concept [2] when dealing with networks to which some additional information is added.

Considering only density in directed networks do not reflect its underlying structure. Due to the direction of its edges it is necessary to provide other type of connectivity. Three approaches about what may be considered as a community in directed networks are deeply explained in [3]: random walk, where groups are formed by nodes that are more likely to stay inside that outside; co-citations, in which the communities could contain member non-connected since they share other pattern, for instance, nodes with the same followers; and flow, where the amount of information which is possible to move in a network is considered.

In particular, we focus on the flow idea. The key question is how to incorporate it into the community detection. In this sense, we use functions to manage imprecise and vague information: fuzzy measures [4]. Inspired by the methodology introduced in [5], we work with a flow fuzzy measure which represents the communication capacity within a set of nodes in directed networks. Then, on the basis of the extended fuzzy graph, we introduce the flow extended fuzzy graph. Given a crisp directed graph $G = (V, E)$ and a fuzzy measure $\mu : 2^V \longrightarrow [0, 1]$ defined over the set of nodes, the triplet $\widetilde{G} = (V, E, \mu)$ obtained from considering together the graph with the fuzzy measure, is called extended fuzzy graph.

On the other hand, other important issue in community detection is to determine the goodness of a partition. We consider the directed modularity [6], $Q_d(G, P) = \frac{1}{m}\sum_{i,j}\left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]\delta(c_i, c_j)$, adapted from the modularity [7] used in non-directed networks (so Q_d is based on the density concept). It can be seen that Q_d only considers the structural information of the graph. On the basis of the modularity, one of the most used methods to describe pattern structures is the Louvain algorithm [8]. It has become very popular because of its speed and effectiveness, especially with large networks. It is important to note that is not suitable for searching communities in which the idea of the flow is considered.

Then, we provide a modification of Louvain algorithm [5], which will allow us to consider not only the topological information provided by the adjacency matrix of networks, but also the additional information defined by the new fuzzy measure modelling the flow ability of a group in a directed network. Once it is defined this method, we give a simple example to show how it works.

II. The flow capacity measure

The first step purpose is the definition of a new fuzzy measure, $\mu^F(S) = \frac{\sum_{i,j\in S} f_{ij}}{\sum_{i,j\in V} f_{ij}}$, which represents the flow capacity between two nodes of a directed network. Propitiously, because of its properties, we can affirm that μ^F is a 2-additive fuzzy measure [9].

Once we have this fuzzy measure μ^F, which represents the additional information, in order to summarize it and facilitates its understanding, we suggest to build the weighted graph associated with it, G_{μ^F}, using the interaction index proposed by Grabisch [9]. Specifically, the adjacency matrix of the G_{μ^F} is the matrix I^D (see equation (1)) and, hence, the μ^F can be re-formulated as a summation which involves the elements of that matrix, where:

$$I_{ij}^D = \frac{f_{ij}}{\sum_{l,m\in V} f_{lm}} \tag{1}$$

Because of its definition, I^D is a non-negative and 1-normalized matrix.

Finally, we define a new representation model to deal with community detection problems in directed networks: the flow

extended fuzzy graph as $\widetilde{G} = (V, E, \mu^F)$. This triplet is obtained from considering together the crisp directed graph with a fuzzy measure which models the flow of G.

III. Community detection problems based on flow measures in directed networks

In a previous point, a new fuzzy measure that describes the flow ability of a set of nodes in a directed network was defined. Subsequently, we approach the community detection problems using that fuzzy measure in order to add further information. We refresh the idea in [5], and we propose a modification of the directed Louvain algorithm, a multiphase heuristic method based on the directed modularity (Q_d) [6]. On the basis of $\widetilde{G} = (V, E, \mu^F)$, it is defined the Flow Capacity Louvain algorithm.

In this proposal is introduced the directed interaction I^D seen in (1), as a matrix which summarizes the information of μ^F. Besides, it is taken into account the matrix A, being the adjacency matrix of the graph G. Then, we combine the matrices A and I^D by means of a linear combination, obtaining the matrix $M = \alpha A + (1-\alpha)\, I^D$, and being $\alpha \in [0, 1]$ a parameter of importance. Note that any other aggregation may have been considered instead. We reckon that α assigns a weight to the original adjacency matrix, and $(1-\alpha)$ represents the importance of the flow defined by the fuzzy measure μ^F. Note that when $\alpha = 1$ the additional information is not considered and the problem will be the classical community detection problem. Or else, if $\alpha = 0$, only the flow capacity among nodes will be allowed to set the partition in the clustering.

Taking our suggestion into account, we handle the Flow Capacity Louvain algorithm in community detection problems when we try to find clusters with maximum flow in directed network. To do it, the objective function will incorporate the matrix M, searching the partition such that maximizes its modularity. In this sense, the best partition will be the one that has the higher $Q_d(M)$, that is, the one with which the value of $\alpha Q_d(A) + (1-\alpha)\, Q_d(I^D)$ is maximum.

In the end, we integrate to the heuristic Louvain algorithm a more global knowledge of the problem.

IV. A simple example to show how the Flow Capacity Algorithm works

We illustrate the Flow Capacity Algorithm performing in a synthetic example which displays the relevance of our contribution. We consider the unweighted directed graph showed in Fig. 1, which its structure can approach a wheel.

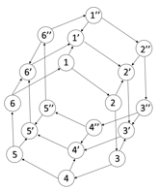


Fig. 1. Directed wheel with 18 nodes.

The graph is obtained with three circles, connecting the first and the third with the middle circle, linking each vertex with its equivalent.

With the aim of getting the communities in the clustering Louvain method, there are several partitions that maximize the modularity, $Q_d(G)$, cutting the wheel in three identical transversal zones. One detected partition is $P_1 = \{\{1, 1', 1'', 2, 2', 2''\}; \{3, 3', 3'', 4, 4', 4''\}; \{5, 5', 5'', 6, 6', 6''\}\}$, which the density is the highest and the directed modularity [6] is maximum. Nevertheless, considering those communities, the members inside the group are not well communicate and, therefore the amount of information moved by the flow is null.

On the other hand, the Flow Capacity Louvain (with $\alpha = 0.5$) will find another partition highly dense but also with communities which could pass information among their members, being $P_2 = \{\{1, 2, 3, 4, 5, 6\}; \{1', 2', 3', 4', 5', 6'\}; \{1'', 2'', 3'', 4'', 5'', 6''\}\}$. To do so, we work with the $Q_d(\widetilde{G})$, where we merge the flow information of the members by means of μ^F.

Despite that both definitions, density and flow, are clearly similar, using an algorithm based only on density does not achieve groups that maximize flow.

V. Conclusions

Our study serves to adapt algorithms which solve the community detection problems in directed networks considering further information to the topological structure. Particularly, we provide the Flow Capacity Louvain Algorithm which considers the flow among nodes defined by the fuzzy measure μ^F, and searches dense and well connected communities.

Acknowledgment

This research has been partially supported by PGC2018096509-B-I00.

References

[1] M. Barroso, I. Gutiérrez, D. Gómez, J. Castro, and R. Espínola, "Group definition based on flow in community detection," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer, 2020, pp. 524–538.

[2] A. V. Goldberg and R. E. Tarjan, "A new approach to the maximum-flow problem," *Journal of the ACM (JACM)*, vol. 35, no. 4, pp. 921–940, 1988.

[3] F. Malliaros and M. Vazirgiannis, "Clustering and community detection in directed networks: A survey," *Physics Reports*, vol. 533, no. 4, pp. 95–142, 2013.

[4] M. Sugeno, "Fuzzy measures and fuzzy integrals—a survey," in *Readings in fuzzy sets for intelligent systems*. Elsevier, 1993, pp. 251–257.

[5] I. Gutiérrez, D. Gómez, J. Castro, and R. Espínola, "A new community detection algorithm based on fuzzy measures," in *Advances in Intelligent Systems and Computing series*, vol. 1029, Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making. INFUS 2019. Springer, Cham, 2020.

[6] A. Arenas, J. Duch, A. Fernández, and S. Gómez, "Size reduction of complex networks preserving modularity," *New Journal of Physics*, vol. 9, no. 6, p. 176, 2007.

[7] M. Newman and M. Girvan, "Finding and evaluating community structure in networks." *Physical Reviews*, vol. 69, p. 026113, 2004.

[8] V. Blondel, J. Guillaume, R. Lambiotte, and E. Lefevre, "Fast unfolding of communities in large networks," *Journal Of Statistical Mechanics-Theory And Experiment*, no. P10008, 2008.

[9] M. Grabisch, "k-order additive discrete fuzzy measures and their representation," *Fuzzy sets and systems*, vol. 92, no. 2, pp. 167–189, 1997.

Nueva dimensión de sentimiento difusa para mejorar el análisis multidimensional en redes sociales

Karel Gutiérrez-Batista
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
karel@decsai.ugr.es

Maria-Amparo Vila
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
vila@decsai.ugr.es

Maria J. Martin-Bautista
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
mbautis@decsai.ugr.es

El análisis de sentimientos, también conocido como minería de opiniones, es una de las tareas más relevantes relacionadas con el procesamiento del lenguaje natural (PLN). Esta tarea permite el análisis de opiniones, sentimientos, emociones y valoraciones que hacen los usuarios sobre un determinado producto o servicio.

Gracias al crecimiento y auge alcanzado por las redes sociales, se han convertido en fuentes de información primarias para el análisis de sentimientos, donde los usuarios tienden a comentar sobre productos y servicios. Es por esto, que muchos de los estudios dedicados al análisis de sentimientos utilizan datos provenientes de redes sociales [1], [2].

El procesamiento y análisis automático de datos textuales de redes sociales constituye un reto de mayor dificultad comparado con datos textuales bien elaborados (correos electrónicos, bibliotecas digitales, sitios web de noticias, etc.), ya que los usuarios expresan ideas, opiniones y sentimientos sobre diferentes temas utilizando un lenguaje informal. Si a esto le sumamos que trabajar con datos de redes sociales implica manejar grandes volúmenes de datos, realizar resúmenes y agregaciones, es de esperar que el uso de tecnologías como Data Warehouses y Online Analytical Processing (OLAP) reciban especial atención en este contexto.

El estudio y desarrollo de herramientas que integren tareas relacionadas con el PLN (análisis de sentimiento, etiquetado de la categoría gramatical, reconocimiento de entidades, detección de tópicos, etc.) y el modelo multidimensional son de gran importancia en el proceso de Inteligencia de Negocios, ya que permiten análisis más exhaustivos para establecer las estrategias adecuadas. Tal es el caso en [3], donde los autores proponen integrar tanto datos textuales como las opiniones extraídas de los documentos, permitiendo a los usuarios realizar análisis más detallados aprovechando las facilidades que ofrece el modelo multidimensional.

Tomando como referencia la problemática planteada anteriormente, en el presente trabajo proponemos crear una dimensión difusa para facilitar el análisis de sentimiento en redes sociales mediante el uso de herramientas de análisis de sentimiento, lógica difusa y las ventajas del modelo multidimensional (más detalles en [4]). La propuesta es completamente automática y no supervisada, lo que nos permite experimentar con datos etiquetados y sin etiquetar. A continuación, listamos las principales contribuciones de este trabajo:

- Crear una Dimensión de Sentimiento Difusa (DSD) a partir de textos de redes sociales que facilite el análisis de sentimiento multidimensional en redes sociales.
- Establecer un proceso automático de agrupación de documentos, teniendo en cuenta los sentimientos expresados en los textos.
- Desarrollar un proceso adaptativo (completamente automático) para seleccionar las *etiquetas lingüísticas* y definir las funciones de pertenencia difusa para cada etiqueta.
- Definir las extensiones de consulta y almacenamiento sobre la DSD.

Para ver en profundidad cada una de estas contribuciones, ver el trabajo original presentado en [4].

Para el análisis de sentimientos se han utilizado las herramientas TextBlob [5] y VADER [6], las cuales entre otras funcionalidades permiten determinar la polaridad del sentimiento de un documento. La polaridad asignada a cada documento permite establecer términos lingüísticos. Estos términos lingüísticos brindan la posibilidad de abordar el problema utilizando la lógica difusa, lo cual es un enfoque factible teniendo en cuenta la incertidumbre inherente del lenguaje natural. Además, permite utilizar etiquetas diferentes a las que han sido empleadas para anotar los documentos (para el caso de un conjunto de datos etiquetado).

Teniendo en cuenta que el principal objetivo del presente trabajo es crear una dimensión de sentimientos difusa que facilite el análisis multidimensional de sentimientos difuso en redes sociales, hemos establecido varios conjuntos de términos lingüísticos difusos (ver trabajo original [4]) que serán los niveles de la jerarquía de nuestra DSD. Se han seleccionado estos conjuntos, ya que no tiene mucho sentido el uso de un mayor número de etiquetas para el análisis de sentimientos. Se debe destacar que no es obligatorio crear la dimensión con una jerarquía que contenga estos niveles, en otras palabras, la jerarquía puede tener un único nivel o dos dependiendo del nivel de granularidad deseado para el análisis. De la misma manera, se pueden establecer tantos conjuntos de términos

lingüísticos como se deseen. Además, hemos decidido que la cantidad de términos lingüísticos en cada conjunto sea impar como se sugiere en la literatura cuando se establecen términos lingüísticos.

Para una mejor comprensión, a continuación resumimos los pasos principales que componen el proceso de asignación de etiquetas. Los pasos son los siguientes:

1) Asignar una polaridad de sentimiento a cada tweet o revisión usando las herramientas TextBlob y VADER,
2) Agrupar los documentos considerando la polaridad de sentimiento asignada a cada documento,
3) Determinar el número óptimo de grupos (utilizando técnicas como el método del codo, el coeficiente de silueta, entre otros),
4) Seleccionar los términos lingüísticos apropiados que mejor se ajusten al conjunto de datos analizado,
5) Crear la función de pertenencia difusa para cada término lingüístico, y
6) Determinar el grado de pertenencia de cada documento en cada término lingüístico.

En las Tablas I y II se muestran los resultados de aplicar los pasos anteriores. Los valores en negritas de la Tabla II indican el término lingüístico con mayor grado de pertenencia.

TABLA I
ASIGNACIÓN DE POLARIDAD A CADA TWEET

tweet	airline sentiment	textblob polarity	vader compound
@VirginAmerica seriously would pay $30 a flight for seats that didn't have this playing. it's really the only bad thing about flying VA	negative	-0.21	-0.59
@VirginAmerica yes, nearly every time I fly VX this "ear worm" won't go away :)	positive	0.47	0.69
@VirginAmerica Really missed a prime opportunity for Men Without Hats parody, there. https://t.co/mWpG7grEZP	neutral	0.2	0.15
@virginamerica Well, I didn't…but NOW I DO!	positive	1	-0.35
@VirginAmerica it was amazing, and arrived an hour early. You're too good to me	positive	0.47	0.77

TABLA II
GRADO DE PERTENENCIA DE LOS TWEETS A CADA TÉRMINO LINGÜÍSTICO

id	airline sentiment	very negative	negative	neutral	somewhat positive	positive	very positive
1	negative	0.0	**0.85**	0.21	0.0	0.0	0.0
2	positive	0.0	0.0	0.0	0.04	**0.89**	0.0
3	neutral	0.0	0.0	0.25	**0.79**	0.02	0.0
4	positive	0.0	0.0	0.02	**0.68**	0.31	0.0
5	positive	0.0	0.0	0.0	0.01	**0.39**	0.12

Para demostrar la viabilidad de la propuesta, se experimentó con datos reales de redes sociales (Twitter) y reseñas de películas (IMDB). En la Tabla III se muestran las características más importantes de los conjuntos de datos. Los resultados muestran la capacidad de la DSD, así como la forma en que la DSD permite dotar a los analistas con una herramienta de consulta robusta (para más detalles ver [4]).

Además del análisis anterior, se realizó un análisis de los resultados de rendimiento del proceso de asignación de términos lingüísticos difusos (FLTAP de las siglas en inglés de Fuzzy Linguistic Terms Assignment Process) en términos de precisión. Debemos resaltar que FLTAP no es un clasificador al uso, es más que eso, ya que puede trabajar sobre un conjunto de datos sin etiquetar, y además permite asignar etiquetas lingüísticas descriptivas para lograr una mejor comprensión de los datos analizados. Este proceso puede verse como un proceso de clasificación si el conjunto de datos utilizado ha sido etiquetado.

Para evaluar el desempeño de FLTAP, lo comparamos con seis algoritmos de clasificación. Tres de ellos son algoritmos de clasificación clásicos (Naive Bayes (NB), Support Vector Machines (SVM) y Decision Trees (DT)) y los otros tres son algoritmos de clasificación basados en aprendizaje profundo, Sequential Neural Networks (SNN), Convolutional Neutral Networks (CNN) y la CNN basada en multicanal (MCNN). Como se puede apreciar en la Tabla IV, los resultados del proceso de asignación de términos lingüísticos difusos son aceptables teniendo en cuenta que el enfoque propuesto no es un algoritmo de clasificación.

TABLA III
DESCRIPCIÓN DE LOS CONJUNTOS DE DATOS

Dataset	Documents	Classes	Vocabulary size
Sentiment140	20000	2	29606
US Airline Sentiment	14640	3	15051
Binary Classification	2000	2	27139
Multi-class Classification	25000	5	15240

TABLA IV
RESULTADOS DE LA EXPERIMENTACIÓN EN TÉRMINOS DE PRECISIÓN

Dataset	NB	SVM	DT	SNN	CNN	MCNN	FLTAP
Sentiment140	**0.75**	073	0.67	0.73	0.73	0.74	0.68
US Airline Sentiment	0.66	0.71	0.65	0.74	0.77	**0.78**	0.7
Binary Classification	0.84	**0.86**	0.68	0.77	0.73	0.81	0.76
Multi-class Classification	0.62	0.62	0.64	**0.71**	0.67	0.68	0.64

AGRADECIMIENTOS

La presente investigación ha sido financiada parcialmente por la Junta de Andalucía y el programa operativo FEDER en el marco del proyecto BigDataMed (P18-RT-2947

REFERENCIAS

[1] K. Howells and A. Ertugan, "Applying fuzzy logic for sentiment analysis of social media network data in marketing," *Procedia Computer Science*, vol. 120, pp. 664 – 670, 2017, 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 22-23 August 2017, Budapest, Hungary.

[2] L. Yang, X. Geng, and H. Liao, "A web sentiment analysis method on fuzzy clustering for mobile social media users," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–13, 2016.

[3] N. U. Rehman, A. Weiler, and M. H. Scholl, "Olaping social media: the case of twitter," in *Advances in Social Networks Analysis and Mining 2013, ASONAM '13, Niagara, ON, Canada - August 25 - 29, 2013*, J. G. Rokne and C. Faloutsos, Eds. ACM, 2013, pp. 1139–1146.

[4] K. Gutiérrez-Batista, M. A. Vila, and M. J. Martín-Bautista, "Building a fuzzy sentiment dimension for multidimensional analysis in social networks," *Appl. Soft Comput.*, vol. 108, p. 107390, 2021.

[5] S. Loria, "Textblob: Simplified text processing." [Online]. Available: https://textblob.readthedocs.io/en/dev/index.html

[6] C. J. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *ICWSM*, 2014.

Estudio de la cardinalidad de algunas familias de conjunciones discretas

Marc Munar*†, Sebastia Massanet*†, Daniel Ruiz-Aguilera*†

* Grupo de investigación en Soft Computing, Procesamiento de Imágenes y Agregación (SCOPIA),
Universitat de les Illes Balears, Palma, Illes Balears
† Instituto de Investigación Sanitaria Illes Balears (IdISBa)
{marc.munar,s.massanet,daniel.ruiz}@uib.es

***Index Terms*—Cadenas finitas, Operadores discretos, Cardinalidad de conjunciones discretas suaves, Matrices de signo alterno**

Al procesar información cualitativa en el razonamiento humano o en la computación con palabras, los conectivos lógicos discretos se han convertido en una herramienta fundamental y valiosa para poder tratarla con precisión. Para analizar las propiedades de esos conectivos de forma genérica, su estudio puede reducirse al estudio de operadores definidos sobre $L_n = \{0, 1, \ldots, n\}$, ya que toda cadena finita formada por $n+1$ elementos es isomorfa a L_n. Con esta formulación, como el dominio y la imagen de un operador discreto son conjuntos finitos, el número de operadores que pueden construirse sobre L_n también es finito. En este contexto, desde la introducción de estos operadores, la comunidad investigadora ya ha centrado sus esfuerzos en intentar encontrar la cardinalidad de algunas clases de operadores discretos. Cronológicamente, en [1] se encuentra el primer estudio sobre la cardinalidad de las t-normas discretas suaves, determinando que pueden construirse 2^{n-1} sobre L_n. En este punto, cabe destacar que la suavidad de un operador discreto es el equivalente al concepto de continuidad para operadores definidos sobre $[0, 1]$; fue introducida por primera vez en [2] bajo la idea que cambios sutiles en los argumentos deben provocar cambios sutiles en el resultado final. Siguiendo con la cardinalidad, en [3] se recogen algunos resultados sobre t-normas discretas en general, haciendo mención en que aún no se conoce una expresión cerrada sobre su cardinalidad. Ahora bien, para intentar facilitar futuras investigaciones se calcularon computacionalmente el número de t-normas discretas para $n \leqslant 13$. Siguiendo este problema, en [4] se introdujeron las cópulas discretas, caracterizándose a partir de las matrices de permutación y concluyendo que hay $n!$ sobre L_n. En [5] se encuentra un estudio similar, caracterizando las cuasi-cópulas discretas en términos de las matrices biestocásticas generalizadas y las cópulas discretas irreductibles en términos de las matrices de signo alterno (por brevedad, ASM, ver [6], [7] para más información).

Siguiendo esta línea de investigación, los autores, en [8], ya han determinado la cardinalidad de las siguientes familias de operadores usando resultados conocidos relacionados con las llamadas particiones del plano (ver [9] para más información):

- El número de funciones de agregación, conjunciones, disyunciones, Sheffer *strokes* e implicaciones definidas en L_n.
- El número de funciones de agregación conmutativas, conjunciones conmutativas, disyunciones conmutativas y Sheffer *strokes* conmutativas definidas en L_n.

Además, también en [8] se ha determinado la cardinalidad para las implicaciones discretas que satisfacen propiedades adicionales, como son: la Simetría Contrapositiva (**CP**) respecto a la negación discreta clásica $N_C(x) = n - x$, para todo $x \in L_n$; el Principio de Neutralidad por la izquierda (**NP**); la Frontera Consecuente (**CB**); el Principio de Ordenación (**OP**); y, finalmente, el Principio de Identidad (**IP**).

En este estudio, usando resultados conocidos sobre las ASM, se han conseguido caracterizar las conjunciones discretas suaves. Primero, se ha construido una biyección entre las conjunciones discretas suaves $\mathcal{C}_n^{\text{SMT}}$ y las ASM, dando lugar al siguiente resultado:

Teorema 1. *Sea $n \geqslant 1$. Un operador binario $C : L_n^2 \to L_n$ es una conjunción suave discreta si, y solo si, existe una matriz de signo alterno $n \times n$ $A = (a_{ij})$ tal que, para todo $x, y \in L_n$:*

$$C(x, y) = \begin{cases} 0, & \text{si } x = 0 \text{ o } y = 0, \\ \sum_{i=1}^{x} \sum_{j=1}^{y} a_{ij}, & \text{en otro caso.} \end{cases} \tag{1}$$

A partir de esta correspondencia biyectiva, pueden extraerse algunas propiedades adicionales, como son:

- Del estudio efectuado en [10] en el que se obtiene una expresión para la cardinalidad de las ASM de tamaño $n \times n$, se puede concluir que esa misma expresión es válida para el número de conjunciones discretas suaves sobre L_n.
- De las propiedades obtenidas en [4], existe una biyección entre el conjunto formado por conjunciones discretas suaves asociativas y el conjunto de las matrices de permutación expresables mediante sumas ordinales de matrices de Łukasiewicz.

Para finalizar, con el objetivo de arrojar luz al problema de la cardinalidad de las t-normas discretas, también se han

Esta publicación es parte del proyecto de I+D+i PID2020-113870GB-I00, financiado por MCIN/AEI/10.13039/501100011033/.

n	$\left\lvert\mathcal{C}_n^{\mathrm{COM}} \cap \mathcal{C}_n^{\mathrm{NE}}\right\rvert$		$\lvert\mathcal{T}_n\rvert$	
1	1	$= 1 \cdot 10^0$	1	$= 1 \cdot 10^0$
2	2	$= 2 \cdot 10^0$	2	$= 2 \cdot 10^0$
3	7	$= 7 \cdot 10^0$	6	$= 6 \cdot 10^0$
4	42	$= 4.2 \cdot 10^1$	22	$= 2.2 \cdot 10^1$
5	429	$\approx 4.3 \cdot 10^2$	94	$= 9.4 \cdot 10^1$
6	7,436	$\approx 7.4 \cdot 10^3$	451	$\approx 4.5 \cdot 10^2$
7	218,348	$\approx 2.2 \cdot 10^5$	2,386	$\approx 2.4 \cdot 10^3$
8	10,850,216	$\approx 1.1 \cdot 10^7$	13,775	$\approx 1.4 \cdot 10^4$
9	911,835,460	$\approx 9.1 \cdot 10^8$	86,417	$\approx 8.6 \cdot 10^4$
10	129,534,272,700	$\approx 1.3 \cdot 10^{11}$	590,489	$\approx 5.9 \cdot 10^5$
11	31,095,744,852,375	$\approx 3.1 \cdot 10^{13}$	4,446,029	$\approx 4.4 \cdot 10^6$
12	12,611,311,859,677,500	$\approx 1.3 \cdot 10^{16}$	37,869,449	$\approx 3.8 \cdot 10^7$
13	8,639,383,518,297,652,500	$\approx 8.6 \cdot 10^{18}$	382,549,464	$\approx 3.8 \cdot 10^8$

TABLA I: Valores de $\left\lvert\mathcal{C}_n^{\mathrm{COM}} \cap \mathcal{C}_n^{\mathrm{NE}}\right\rvert$, con $n \leq 13$, obtenidos usando la Ecuación (2). En la segunda columna, los valores de la cardinalidad de las t-normas discretas $\mathcal{T}_n$ obtenidos computacionalmente en [3].

estudiado las conjunciones discretas conmutativas $\mathcal{C}_n^{\mathrm{COM}}$ y con elemento neutro n, $\mathcal{C}_n^{\mathrm{NE}}$, obteniendo una expresión para las mismas.

Proposición 1. *Para todo $n \geqslant 1$,*

$$\left\lvert\mathcal{C}_n^{\mathrm{COM}} \cap \mathcal{C}_n^{\mathrm{NE}}\right\rvert = \left\lvert\mathcal{C}_n^{\mathrm{SMT}}\right\rvert = \prod_{i=0}^{n-1} \frac{(3i+1)!}{(n+i)!}. \tag{2}$$

De esta forma, se ha determinado la cardinalidad de una familia de operadores a los que solo les resta la propiedad asociativa para ser t-normas discretas, concluyendo que la expresión obtenida es una cota para la cardinalidad de las t-normas discretas. En la tabla I se ilustran algunos valores usando la expresión obtenida, así como una comparativa con los valores experimentales de la cardinalidad de t-normas discretas.

REFERENCES

[1] G. Mayor and J. Torrens, "On a class of operators for expert systems," *International Journal of Intelligent Systems*, vol. 8, no. 7, pp. 771–778, 1993.

[2] L. Godo and C. Sierra, "A new approach to connective generation in the framework of expert systems using fuzzy logic," in *Proceedings of the Eighteenth International Symposium on Multiple-Valued Logic*, 1988, pp. 157–162.

[3] B. De Baets and R. Mesiar, "Triangular norms on product lattices," *Fuzzy Sets and Systems*, vol. 104, no. 1, pp. 61–75, 1999.

[4] G. Mayor, J. Suñer, and J. Torrens, "Copula-like operations on finite settings," *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 4, pp. 468–477, 2005.

[5] I. Aguiló, J. Suñer, and J. Torrens, "Matrix representation of discrete quasi-copulas," *Fuzzy Sets and Systems*, vol. 159, no. 13, pp. 1658–1672, 2008.

[6] D. P. Robbins and H. Rumsey, "Determinants and alternating sign matrices," *Advances in Mathematics*, vol. 62, no. 2, pp. 169–184, 1986.

[7] D. M. Bressoud and J. G. Propp, "How the Alternating Sign Matrix Conjecture Was Solved," *Notices of the American Mathematical Society*, vol. 46, pp. 637–646, 1999.

[8] M. Munar, S. Massanet, and D. Ruiz-Aguilera, "A study about the cardinality of some connectives defined on finite chains," in *The 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022)*, Milan, Italy, 2022, aceptado en congreso.

[9] R. P. Stanley and S. Fomin, *Enumerative Combinatorics*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999, vol. 2.

[10] D. Zeilberger, "Proof of the alternating sign matrix conjecture," *Electronic Journal of Combinatorics*, vol. 3, no. 2, 1996.

Incertidumbre en problemas de optimización robusta en el tiempo: clasificación y oportunidades

Pavel Novoa-Hernández
Modelos de Decisión y Optimización
Universidad de Granada
Granada, España
pavelnovoa@correo.ugr.es

David A. Pelta
Dept. Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada
Granada, España
dpelta@ugr.es

***Abstract*—La optimización robusta en el tiempo es un tema relativamente reciente dentro de la computación evolutiva, para la que no existe hasta ahora una caracterización adecuada de la incertidumbre presente en la función objetivo del problema. Este trabajo propone una clasificación de este aspecto, y organiza las contribuciones actuales de acuerdo a las categorías de la clasificación propuesta. Algunas oportunidades de investigación son propuestas.**

***Index Terms*—Robust optimization over time, Uncertainty, Surrogate model, Forecasting**

I. Motivación

La optimización robusta en el tiempo (en lo adelante ROOT) se define como la tarea de encontrar, para cada instante de tiempo, la solución óptima tanto para el instante de tiempo actual, como para un conjunto de instantes del futuro [1], [2]. Este requerimiento implica el manejo de un importante nivel de incertidumbre, dado que el futuro no siempre puede ser caracterizado con exactitud. De hecho, a falta de información específica sobre el futuro, solo se tendría al pasado (espacios de búsqueda ya vistos por el algoritmo) para intentar pronosticar adecuadamente la calidad de una determinada solución.

Aunque en la actualidad existen contribuciones importantes sobre ROOT, este campo adolece de una caracterización adecuada sobre la incertidumbre presente en los problemas. En tal sentido, el presente trabajo tiene un objetivo doble: por un lado, se propone una clasificación de los problemas ROOT teniendo en cuenta la posibilidad o no, de evaluar exactamente las funciones objetivo; y por otro, organizar las contribuciones existentes de acuerdo a los escenarios de incertidumbre propuestos por nuestra clasificación. De esta forma, esperamos detectar las principales tendencias y oportunidades de investigación futura.

II. Problemas de optimización robusta en el tiempo

Concretamente, un problema ROOT se define como:

$$\max_{x \in \Omega} \mathcal{R}(x, t) \tag{1}$$

donde $\Omega \subseteq \mathbb{R}^D$ es el espacio de búsqueda y $x \in \Omega$ es una solución candidata. $t \in \mathbb{N}$ es un ambiente del problema. La función $\mathcal{R} : \mathbb{R}^D \times \mathbb{N} \rightarrow \mathbb{R}$ define la robustez de una solución dada x en el ambiente t. La tarea consiste entonces en encontrar el conjunto de soluciones $X^* = \{x_t^* \ : \ t = 1, 2, ..., N\}$ tal que x_t^* maximiza $\mathcal{R}$.

Existen varias formas de definir a $\mathcal{R}$. Por ejemplo, se prodía asumir como la agregación de varios valores de T funciones objetivo futuras. T se conoce como *ventana de tiempo*. Cuando en esta definición se emplea un media aritmética simple, recibe el nombre de *Fitness promedio* [2]. Otra alternativa es definir la robustez como la cantidad de ambientes del futuro en los que una determinada solución mantiene un nivel de calidad aceptable (e.g sobre un cierto umbral predefinido). En este caso, se conoce como *Tiempo de supervivencia* [2].

III. Clasificación propuesta

Como se advirtió anteriormente, si se considera la posibilidad de evaluar exactamente o no, las funciones objetivo de cada intervalo de tiempo de la ejecución, entonces tenemos tres escenarios posibles: (1) las funciones objetivo del futuro pueden evaluarse con exactitud, (2) solo las funciones objetivos del pasado pueden evaluarse con exactitud, y (3) no es posible evaluar exactamente las funciones objetivo del pasado o del futuro. Estos escenarios se ilustran en el diagrama de flujo de la Fig. 1, junto con el número de contribuciones que se han publicado hasta ahora por cada uno de estos. Adicionalmente, en la Tabla I se detalla qué contribuciones se han hecho por cada escenario.

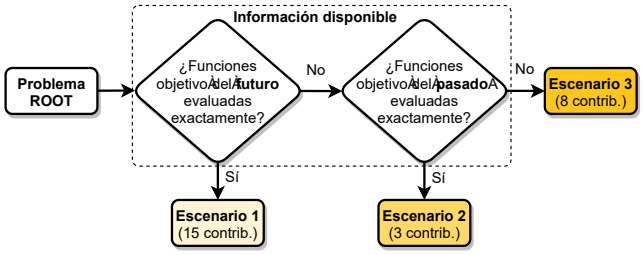


Fig. 1. Posibles escenarios de ROOT según la información disponible en el problema.

Este trabajo ha sido financiado por los proyectos PID2020-112754GBI0, MCIN/AEI /10.13039/501100011033 y FEDER/Junta de Andalucía- Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto (BTIC-640-UGR20).

TABLA I
CONTRIBUCIONES POR ESCENARIO DE PROBLEMA ROOT.

Esc.	Descripción	Contribuciones
1	Las funciones objetivo del futuro pueden evaluarse exactamente.	[3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17]
2	Solo las funciones del pasado pueden evaluarse exactamente.	[18], [19], [20]
3	Ni las funciones objetivo del futuro, ni las del pasado pueden evaluarse exactamente.	[2], [21], [22], [23], [24], [25], [26], [1]

Nótese que de los tres escenarios, el *Escenario 1* es el que ha sido más abordado. Sin embargo, hay que tener en cuenta que este representa la situación ideal de que es posible conocer el futuro con total certidumbre. De manera que aquí, si se adoptase una definición de robustez como el *Fitness promedio*, entonces es posible aplicar algoritmos típicos de optimización dinámica evolutiva [27]. En cambio, los escenarios 2 y 3 requieren de mecanismos adicionales para lidiar con la incertidumbre.

IV. OPORTUNIDADES DE INVESTIGACIÓN FUTURA

El presente trabajo ha abordado un aspecto muy específico de la optimización robusta en el tiempo: la incertidumbre de la funciones objetivo. La clasificación propuesta, y posterior organización de la literatura, permiten identificar las oportunidades de investigación futura siguientes:

- abordar/proponer problemas de los escenarios 2 y 3;
- aplicar enfoques para el manejo de incertidumbre no probabilística, esto es, adoptando enfoques como la lógica difusa o teoría de conjuntos aproximados [28].
- caracterizar de manera más general a los problemas ROOT, esto es, teniendo en cuenta otros factores como la definición de la robustez.

REFERENCES

[1] X. Yu, Y. Jin, K. Tang, and X. Yao, "Robust optimization over time — a new perspective on dynamic optimization problems," in *IEEE Congress on Evolutionary Computation*, July 2010, pp. 1–6.

[2] H. Fu, B. Sendhoff, K. Tang, and X. Yao, "Robust optimization over time: Problem difficulties and benchmark problems," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 5, pp. 731–745, 2015.

[3] L. Adam and X. Yao, "A simple yet effective approach to robust optimization over time," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2019, pp. 680–688.

[4] R. Ankrah, B. Lacroix, J. McCall, A. Hardwick, A. Conway, and G. Owusu, "Racing strategy for the dynamic-customer location-allocation problem," in *2020 IEEE Congress on Evolutionary Computation (CEC)*, 2020, pp. 1–8.

[5] M. Chen, Y. Guo, H. Liu, and C. Wang, "The evolutionary algorithm to find robust pareto-optimal solutions over time," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–18, 2015.

[6] I. Essiet and Y. Sun, "Tracking variable fitness landscape in dynamic multi-objective optimization using adaptive mutation and crossover operators," *IEEE Access*, vol. 8, pp. 188 927–188 937, 2020.

[7] H. Fu, B. Sendhoff, K. Tang, and X. Yao, "Characterizing environmental changes in robust optimization over time," in *2012 IEEE Congress on Evolutionary Computation, CEC 2012*, 2012.

[8] Y.-N. Guo, J. Cheng, S. Luo, D. Gong, and Y. Xue, "Robust dynamic multi-objective vehicle routing optimization method," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 6, pp. 1891–1903, nov 2018.

[9] J. Guzmán-Gaspar and E. Mezura-Montes, "Robust optimization over time with differential evolution using an average time approach," in *2019 IEEE Congress on Evolutionary Computation (CEC)*, 2019, pp. 1548–1555.

[10] J.-Y. Guzmán-Gaspar, E. Mezura-Montes, and S. Domínguez-Isidro, "Differential evolution in robust optimization over time using a survival time approach," *Mathematical and Computational Applications*, vol. 25, no. 4, p. 72, oct 2020.

[11] J. Guzmán-Gaspar and E. Mezura-Montes, "Differential evolution variants in robust optimization over time," in *2019 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, 2019, pp. 164–169.

[12] Y. Huang, Y. Ding, K. Hao, and Y. Jin, "A multi-objective approach to robust optimization over time considering switching cost," *Information Sciences*, vol. 394-395, pp. 183–197, 2017.

[13] Y.-J. Huang, Y.-C. Jin, and K.-R. Hao, "Robust optimization over time for constrained optimization based on swarm intelligence," *Kongzhi yu Juece/Control and Decision*, vol. 35, no. 3, pp. 740–748, 2020.

[14] Y. Liu and H. Liang, "A root approach for stochastic energy management in electric bus transit center with pv and ess," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[15] Y. Liu and H. Liang, "A three-layer stochastic energy management approach for electric bus transit centers with pv and energy storage systems," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1346–1357, March 2021.

[16] D. Yazdani, T. T. Nguyen, J. Branke, and J. Wang, "A new multi-swarm particle swarm optimization for robust optimization over time," in *Applications of Evolutionary Computation*, G. Squillero and K. Sim, Eds. Cham: Springer International Publishing, 2017, pp. 99–109.

[17] D. Yazdani, J. Branke, M. N. Omidvar, T. T. Nguyen, and X. Yao, "Changing or keeping solutions in dynamic optimization problems with switching costs," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ser. GECCO '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1095—1102.

[18] M. Fox, S. Yang, and F. Caraffini, "An experimental study of prediction methods in robust optimization over time," in *2020 IEEE Congress on Evolutionary Computation (CEC)*, July 2020, pp. 1–7.

[19] H. Fu, B. Sendhoff, K. Tang, and X. Yao, "Finding robust solutions to dynamic optimization problems," in *Applications of Evolutionary Computation*, ser. Lecture Notes in Computer Science, A. I. Esparcia-Alcázar, Ed., vol. 7835. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 616–625.

[20] Y. Guo, M. Chen, H. Fu, and Y. Liu, "Find robust solutions over time by two-layer multi-objective optimization method," in *2014 IEEE Congress on Evolutionary Computation (CEC)*, July 2014, pp. 1528–1535.

[21] Y. Huang, Y. Jin, and Y. Ding, "New performance indicators for robust optimization over time," in *2015 IEEE Congress on Evolutionary Computation (CEC)*, May 2015, pp. 1380–1387.

[22] Y. Huang, Y. Jin, and K. Hao, "Decision-making and multi-objectivization for cost sensitive robust optimization over time," *Knowledge-Based Systems*, vol. 199, p. 105857, 2020.

[23] S. Hutterer and M. Affenzeller, "Policy function approximation for optimal power flow control issues," in *1st International Workshop on Simulation for Energy, Sustainable Development and Environment, SESDE 2013, Held at the International Multidisciplinary Modeling and Simulation Multiconference, I3M 2013*, 2013, pp. 66–70.

[24] Y. Jin, K. Tang, X. Yu, B. Sendhoff, and X. Yao, "A framework for finding robust optimal solutions over time," *Memetic Computing*, vol. 5, no. 1, pp. 3–18, 2013.

[25] P. Novoa-Hernández, D. A. Pelta, and C. C. Corona, "Approximation models in robust optimization over time - an experimental study," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, July 2018, pp. 1–6.

[26] D. Yazdani, T. Nguyen, and J. Branke, "Robust optimization over time by learning problem space characteristics," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 1, pp. 143–155, 2018.

[27] P. Novoa-Hernández, C. C. Corona, and D. A. Pelta, "Self-adaptation in dynamic environments - a survey and open issues," *International Journal of Bio-Inspired Computation*, vol. 8, no. 1, pp. 1–13, 2016.

[28] C. Wang, X. Qiang, M. Xu, and T. Wu, "Recent advances in surrogate modeling methods for uncertainty quantification and propagation," *Symmetry*, vol. 14, no. 6, 2022.