IPMU 2022

The 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems

BOOK OF ABSTRACTS

Edited by

S. Boffa, D. Ciucci, I. Couso, J. Medina, D. Ślęzak

July 11-15 2022
University of Milano-Bicocca

UNIVERSITA' DEGLI STUDI DI MILANO BICOCCA
DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE

Title: IPMU 2022 Book of abstracts
THE 19TH INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING AND MANAGEMENT OF UNCERTAINTY IN KNOWLEDGE-BASED SYSTEMS

Editors: Stefania Boffa, Davide Ciucci, Inés Couso, Jesús Medina, Dominik Ślęzak

Publisher: Università degli Studi di Milano-Bicocca - Dip.to di Informatica, Sistemistica e Comunicazione

ISBN 979-12-210-1489-1

Preface

We are very pleased to present you with the Book of Abstract of the 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022). The conference was held during July 11–15, in Milan, Italy. The IPMU conference is organized every two years with the aim of bringing together scientists working on methods for the management of uncertainty and aggregation of information in intelligent systems. Since 1986, the IPMU conference has been providing a forum for the exchange of ideas between theoreticians and practitioners working in these areas and related fields.

Following the IPMU tradition, the Kampé de Fériet Award for outstanding contributions to the field of uncertainty and management of uncertainty was presented. Past winners of this prestigious award are Lotfi A. Zadeh (1992), Ilya Prigogine (1994), Toshiro Terano (1996), Kenneth Arrow (1998), Richard Jeffrey (2000), Arthur Dempster (2002), Janos Aczel (2004), Daniel Kahneman (2006), Enric Trillas (2008), James Bezdek (2010), Michio Sugeno (2012), Vladimir N. Vapnik (2014), Joseph Y. Halpern (2016), Glenn Shafer (2018) and Barbara Tversky (2020). In this 2022 edition, the award was given to Tomaso Poggio for his interdiciplinary work on human and machine intelligence and his fundamental research in computational neuroscience, in particular concerning the computational analysis of vision and learning.

The program included the keynote talk of Tomaso Poggio, as recipient of the Kampé de Feriet Award, and the keynote talks by César Hidalgo (Artificial and Natural Intelligence Toulouse Institute, Toulouse, France), Marianne Huchard (Laboratory of Informatics, Robotics, and Microelectronics, Montpellier, France) and Andrzej Skowron (University of Warsaw, Poland).

To celebrate the 40^{th} anniversary on the seminal paper on Rough Sets by Z. Pawlak, a panel session on rough sets was organized. The panel session witnessed the participation and discussion of renowned researchers on rough sets, including Salvatore Greco, Ernestina Menasalvas and Andrzej Skowron to whom we are grateful for their contribution. The participants shared their memories and experiences related to rough-set-based decision making, applications of rough set approximations, rough set contributions to machine learning, emphasizing the strong points of rough sets and the ways of using them in hybrid solutions.

The IPMU 2022 program consisted of 124 full papers, included in two volumes of the proceedings and 28 abstract collected in the present book.

The organization of the IPMU 2022 conference was possible with the assistance, dedication, and support of many people and institutions. We are particularly thankful to the organizers of special sessions. Such sessions, dedicated to variety of topics and organized by experts, have always been a characteristic feature of IPMU conferences. We gratefully acknowledge the technical co-sponsorship of the IEEE Computational Intelligence Society and the European Society for Fuzzy Logic and Technology (EUSFLAT).

We also acknowledge the support received from the University of Milano-Bicocca and in particular, from the Department of Informatics, Systems and Communications; Our very special and greatest gratitude goes to the authors

who have submitted results of their work and presented them at the conference. Without you this conference would not take place. Thank you!

We hope that these abstracts and the related discussion at the conference provide the readers with multiple ideas leading to numerous research activities, significant publications, and intriguing presentations at future IPMU conferences.

July 6, 2022
Milano

Stefania Boffa
Davide Ciucci
Inés Couso
Jesús Medina
Dominik Ślęzak

Table of Contents

Solution of Fuzzy Fractional Logistic Differential Equations with Generalized Hukuhara Differentiability Concept

Zahra Alijani

Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, Czech Republic
zahraalijani2@gmail.com

The ordinary logistic differential equation of the logistic system

$$\begin{aligned} &u'(t) = \mu u(t)[1 - u(t)],\ t \in [0, T], \\ &u(0) = u_0 \end{aligned} \tag{1}$$

where T is a real number, u_0 is a given initial value and $u : [0, T] \to \mathbb{R}$ is an unknown function that appears in the study of discretization methods for differential equations. Furthermore, the application of the theory of difference equations is increasing rapidly in various fields such as numerical analysis, control theory, financial mathematics, computer science, data science, and economy [3, 6–8]. The fractional logistic differential equation (FLDE) has recently drawn the attention of many scientists. Nieto in [5] has obtained the exact solution of Caputo-Fabrizio FLDE and solved the logistic differential equation of fractional order and non-singular kernel.

One approach to considering uncertainty in a dynamical system to predict the behavior of imprecise real-world phenomena is the fuzzification of the corresponding crisp difference equations [1, 2]. Independently of similar particular formulations of the equation in any model, we expect that the solution accurately reflects the real behavior of the system. Therefore, getting different results through fuzzification of the unique crisp equation may seem unnatural. However, we can consider this fact as an advantage of fuzzy mathematics due to the existence of several choices which can be examined [4].

If u_0 or μ are fuzzy numbers, then the state solution of (1), $u(t)$ is not expected to be a deterministic function. Therefore, the corresponding system is called a fuzzy logistic differential equation. In this study, we solve a fuzzy logistic differential equation, described by

$$\begin{aligned} {}_0^C D_t^{\beta} u(t) &= \mu u(t)[1 - u(t)], \quad t \in [0, T], \\ u(0) &= u_0 \end{aligned} \tag{2}$$

where ${}_0^C D_t^{\beta} u(t)$, stands for the differential derivative of Caputo and $0 < \beta \leq 1$ is the order of the derivative. We consider μ to be a fixed deterministic parameter, and the initial value is a fuzzy function. We study the fuzzy logistic differential equation using the fuzzy Caputo-Fabrizio operator under the generalized Hukuhara differentiability concept. Then, we solve the fuzzy fractional logistic differential equation by giving a power series. Numerical experiments are presented to show the effectiveness of the power series method. Numerical experiments are presented to show the effectiveness of the power series method.

Acknowledgment

The work is supported by ERDF / ESF by the project 'Center for the development of artificial intelligence methods for the automotive industry of the region' No. CZ.02.1.01 / 0.0 / 17-049 / 0008414.

References

1. Z. Alijani, D. Baleanu,B. Shiri, G. C. Wu, Spline collocation methods for systems of fuzzy fractional differential equations, Chaos, Solitons, and Fractals, 131(2019), 1-12.
2. B. Bede, Mathematics of Fuzzy Sets and Fuzzy Logic, Volume 295 of Studies in Fuzziness and Soft Computing, Springer, 2013.
3. F. Brauer, C. Castillo-Chavez, Z. Feng, Mathematical Models in Epidemiology, Springer-Verlag, New York, 2019.
4. A Khastan, Z Alijani, On the new solutions to the fuzzy difference equation $x_{n+1} = A + \frac{B}{x_n}$, Fuzzy Sets and Systems 358, 64-83.
5. J. Nieto, solution of a fractional logistic ordinary differential equation, Applied Mathematics Letters, 123 (2022), 107568.
6. V. V. Tarasova, V.E. Tarasov, Logistic map with memory from economic model, Chaos Solitons Fractals, 95 (2017), 84–91.
7. H.R. Thieme, Mathematics in Population Biology, in: Princeton Series in Theoretical and Computational Biology.
8. C. A. Valentim Jr., J.A. Oliveira, S.A. Rabi, N.A. David, Can fractional calculus help improve tumor growth models, Journal of Computational and Applied Mathematics, 379 (2020), 112964.

The Bandler–Kohout Product of Fuzzy Relations with Undefined Membership Degrees

Libor Běhounek and Martina Daňková

Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, CE IT4Innovations, 30. dubna 22, 701 03 Ostrava, Czech Republic
{Libor.Behounek|Martina.Dankova}@osu.cz

In this contribution, we introduce a generalization of the Bandler–Kohout fuzzy relational product, also known as the BK-product or inf-R composition [1], to admit undefined membership degrees in both the input and output fuzzy relations, and present its basic properties. The generalization provides a method of automated computing (e.g., in database querying) with an error value representing an undefined input degree and propagating to the output fuzzy relation. In the input fuzzy relations, undefined degrees may arise due to an assignment error, non-applicable features, division by zero, and similar failures; the apparatus is, however, not intended for imputation of unknown missing degrees.

We employ a representation of fuzzy sets and fuzzy relations with undefined membership degrees given in [2] (we call such fuzzy sets or relations *partial*). There, undefined degrees are encoded by an extra value $*$ added to the algebra of membership degrees L, and the algebraic operations on L are extended to $L_* = L \cup \{*\}$. The apparatus of partial fuzzy relations has already been used in several applications, e.g., [4,5].

The following definition of the BK-product for partial fuzzy relations uses a specific combination of the extended operations to ensure that the domain of definition of the BK-product is the relational composition of the domains of the input fuzzy relations:

$$(R \lhd S)xz = \bigwedge_{y \in U}{}_{\mathrm{S}} (Rxy \rightarrow_{\mathrm{B}} Syz),$$

where U is a universe of discourse and $\bigwedge_{\mathrm{S}}, \rightarrow_{\mathrm{B}}$ are the following extensions of the lattice infimum $\bigwedge$ and a fuzzy implication $\rightarrow$ on L, respectively, to L_*:

- *Bochvar fuzzy implication:* $\alpha \rightarrow_{\mathrm{B}} \beta = \begin{cases} \alpha \rightarrow \beta & \text{if } \alpha, \beta \neq *, \\ * & \text{otherwise.} \end{cases}$
- *Sobociński infimum:* $\bigwedge_{i \in I}{}_{\mathrm{S}} \alpha_i = \begin{cases} \bigwedge_{\substack{i \in I \\ \alpha_i \neq *}} \alpha_i & \text{if } \alpha_i \neq * \text{ for some } i \in I, \\ * & \text{otherwise.} \end{cases}$

Consider the following order on L_*: $\alpha \leq_{\mathrm{sub}} \beta$ iff $\alpha \leq \beta$ in L or $\alpha = *$; and define $A \subseteq_{\mathrm{sub}} B$ iff $Ax \leq_{\mathrm{sub}} Bx$ for all $x \in U$. Then, based on the properties of the extended operations, we can prove several basic properties of the partial inf-R composition, including, e.g., the antitony of $\lhd$ in the first argument, which takes the following form:

$$\text{if} \quad R_1 \subseteq_{\text{sub}} R_2, \quad \text{then} \quad R_2 \lhd S \subseteq_{\text{sub}} R_1 \lhd S.$$

Following the line of our previous work on sup-T composition of partial fuzzy relations [3], we can automatically transfer the properties of $\lhd$ to several related operations, such as the BK-image, BK-preimage, and graded inclusion. The transfer of the properties to the (lower-dimensional) related notions is ensured by a technical trick introduced in [3] that represents lower-dimensional input arguments (i.e., membership degrees from L_* and partial fuzzy sets) by suitable partial fuzzy relations. The properties of $\lhd$ then apply as well to the related notions; e.g., the BK-preimage of a partial fuzzy set A under a partial fuzzy relation R is defined as

$$(R \overset{\leftarrow}{\lhd} A)x = \bigwedge_{y \in U}{}_{\text{S}} (Rxy \to_{\text{B}} Ay),$$

and from the antitony of $\lhd$, it follows that

$$\text{if} \quad R_1 \subseteq_{\text{sub}} R_2, \quad \text{then} \quad R_2 \overset{\leftarrow}{\lhd} A \subseteq_{\text{sub}} R_1 \overset{\leftarrow}{\lhd} A.$$

The properties of the related notions thus obtained as simple corollaries of the basic properties of the sup-T and inf-R compositions cover a large part of the theory of partial fuzzy relations. Additionally, they provide a simple calculus useful for solving problems in fields dealing with partial fuzzy relations such as approximate reasoning, knowledge representation, or fuzzy relational equations.

Acknowledgement

This research was supported by project No. 20-07851S of the Czech Science Foundation.

References

1. Bandler, W., Kohout, L.J.: Fuzzy relational products and fuzzy implication operators. In: International Workshop of Fuzzy Reasoning Theory and Applications. Queen Mary College, University of London, London (1978)
2. Běhounek, L., Daňková, M.: Variable-domain fuzzy sets—Part II: Apparatus. Fuzzy Sets and Systems **380**, 19–43 (2020)
3. Běhounek, L., Daňková, M.: Sup-T compositions of partial fuzzy relations. In: Mesiar, R., et al. (eds.) Joint Proceedings of the 19th World Congress of IFSA, the 12th Conference of EUSFLAT, and the 11th AGOP. pp. 24–31. Atlantis Press (2021)
4. Daňková, M.: Merging partial fuzzy rule-bases. In: Merelo, J.J., et al. (eds.) Proceedings of the 12th International Joint Conference on Computational Intelligence (IJCCI 2020). pp. 243–251. SciTePress (2020)
5. Štěpnička, M., Cao, N., Běhounek, L., Burda, M., Dolný, A.: Missing values and dragonfly operations in fuzzy relational compositions. International Journal of Approximate Reasoning **113**, 149–170 (2019)

Soft Computing to Aid Biologists Recognize Material Components in Food Products

Gloria Bordogna, [1] Luca Braglia[2], Luca Frigerio[1], Floriana Gavazzi[2], Silvia Gianì[2], Laura Morello[2], Anna Rampini[1], and Diego Breviario[2]

[1] CNR - IREA, Milano, Italy
[2] CNR – IBBA, Milano, Italy
name.surname@cnr.it

An original decision support Web service is described, designed to aid biologists in identifying the animal and plant species of materials present in food products, based on the automatic analysis of their DNA barcodes. The service allows to speed up a rather burden and meticulous task of biologists as well as standardizing the recognition so as to increase their confidence on the results which aim to authenticate the quality of food products. The PCR-based DNA barcoding method, named tubulin-based polymorphism (TBP) [1] can assign specific genomic finger printings to any plant or/and animal species and crop varieties. This fingerprint, or DNA barcode, does not depend on DNA sequencing as the classical DNA barcoding strategy, and is also effective in the rapid recognition of the different species present in a mixture. Each specific DNA barcode consists in a profile of characteristics amplified DNA fragments, corresponding to discrete fluorescence peaks after separation by capillary electrophoresis. Nevertheless, the TBP analysis is subjective to several variabilities depending on the conditions of the experiments and preparation of the sample that may affect the output so that some fuzziness in peak resolution (number of peaks, position and height) may appear. This may compromise effective species and varieties authentication. In animals this may hinder some race recognition. These are the reasons why an automatic and flexible approach was needed both to be able to cope with such variabilities and to speed up the recognition. The automatic recognition algorithm has been defined based on a soft computing approach that flexibly compares the DNA barcode extracted by applying the TBP method to the food sample with the DNA barcodes of M representative elements of animal and plant species or varieties, previously selected and stored in a database by the aid of the Web service functionalities. M is chosen as a multiple of the number of species or varieties so that each species or variety can have an adequate number of representative elements in the database to provide, as far as possible, a coverage of all inter-species or intra-variety variations. The recognition algorithm is conceived as a time series multinomial regression. Although deep learning methods, specifically RNNs and LSTMs, offer a way of processing time series more consistent than traditional models, nevertheless, we could not apply them due to their need of large amounts of training data, which in our case were limited to a maximum of few samples for each species and variety. Our solution was defined based on a kNN algorithm which is particularly effective when classes are characterized by variabilities, like in our case study. For our task we used a modified instance based representation of the DNA barcodes which were cleansed from background noise, compressed and normalized to make them comparable by means of functions flexibly tuned by using a graphic user interface. As far as the distance a number of options were tested including Dynamic Time Warping (DTW) which is considered the most effective to compare time series with different length and shifted along the time dimension as it occurs in our case. Nevertheless, DTW is costly in terms of time, its time complexity is O(x*y), where x and y are the lengths of the two DNA barcodes. Thus, for our specific task, we defined a more efficient parametric fuzzy similarity measure (PFS) based on a fuzzy inclusion, with a lower time complexity O(x*log(x) + y*log(y)), which demonstrated an equal or greatest effectiveness then DTW in the evaluation. A fuzzy inclusion was chosen because a food product can be mixture of animal and plant components, whose approximate number C is generally known: thus, the DNA barcode of a food sample can include approximately C different DNA barcodes of species or varieties. Besides that, the PFS function can be flexibly tuned by specifying N, the desired maximum number of DNA barcodes' peaks to consider, and the value p of the p_norm distance used to compare pairs of peaks. The N smallest p-norm distance values between any pair of peaks are combined to compute the fuzzy inclusion of any single DNA barcode of a representative element in the database within that of the unknown sample. The C species or varieties to which the Kst most similar representatives belong are recognized as compounds of the food product. The Web service (https://lab.s3qr.it/), implemented in node.js with a noSQL database MongoDB, presents several functionalities that the biologist can select to normalize a DNA barcode; to create a new database of DNA barcodes representative of a set of species, and finally to classify DNA barcodes of unknown samples. The algorithm was evaluated using accuracy metrics on two distinct datasets: a dataset of 380 meat samples classified into 12 different animal species, which reported an average accuracy of 0.86, and a dataset of 240 hazelnuts samples, belonging to 13 different varieties of hazelnuts, which reported a lower average accuracy equal to 0.75, mainly due to the fact that some varieties had only one representative in the database. These evaluation of the tool encouraged its engineering for possible use in an operative context of smart food products quality assessment.

Acknowledgements

The work was partially funded by the Spatials3 project (https://hubspatials3.it/)

References

1. Morello L.; Braglia L.; Gavazzi F.; Gianì S.; Breviario D.: Tubulin-Based DNA Barcode: Principle and Applications to Complex Food Matrices. Genes (Basel), 10(3), doi: 10.3390/genes10030229, (2019).

Fuzzy Similarity of Property Graphs

Arnaud Castelltort, Anne Laurent, and Maman Sani Aboubacar

LIRMM, Univ Montpellier, CNRS, Montpellier, France
{firstname.lastname}@lirmm.fr
https://www.lirmm.fr

1 Preliminary Notions

Property graphs have emerged during the last decade and allow to represent concepts and the relationship between concepts.

In our work, we study how to compare two property graphs and we investigate the main future works to use fuzzy similarities in such a context.

A property graph PG is defined as [1]:

- a set of nodes provided with labels and $(key, values)$ properties, and
- relationships between the nodes that are directed and that are provided with a type and $(key, values)$ properties.

It is important to highlight that there exists a difference between labels, types and properties. Figure 1 illustrates the concept of a labeled property graph.

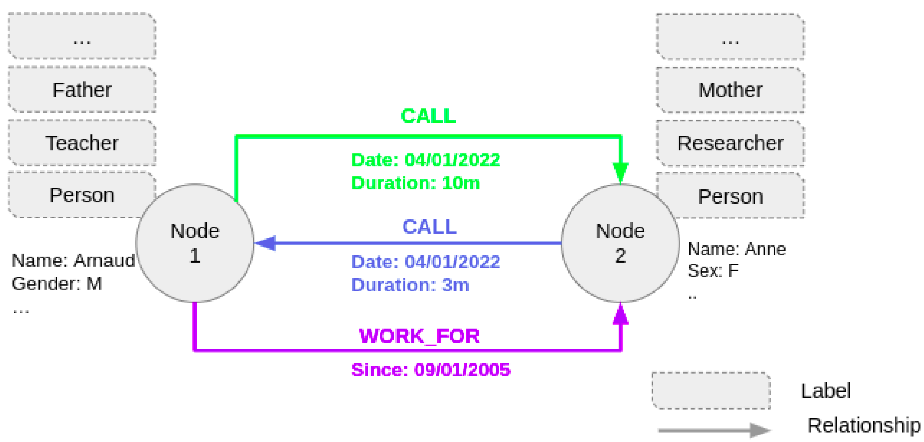

Fig. 1. A labeled property Graph

2 Towards Fuzzy Comparison of Property Graphs

Computing graph similarity is a common operation. Many approaches have been proposed in the literature for similarity searching, from structured, non structured and also graphs and knowledge graphs [4, 5, 7].

However, the proposed approaches do not take into account the graph properties. For this purpose, we propose a similarity measure taking into account the structure of the graph and also the information related to the entities of the graph, i.e the nodes and their properties as well as the relations and their properties.

For computing the similarity $SIM(PG1, PG2)$ between two property graphs $PG1$ with n_i nodes and $PG2$ with n_j nodes, we aggregate diverse measures:

- **Structural comparison** dedicated to the keys of the properties of nodes $SIM_{\mathcal{K}}$ and relationships $SIM_{\mathcal{R}}$;
- **Semantic comparison** dedicated to values of properties of nodes $SIM_{\mathcal{V}}$, that can also be enriched by using an external ontology $SIM_{\mathcal{O}}$.

In our current work, we investigate how fuzzy similarities can be considered in such a context. Fuzzy similarities have been studied for many years [2, 3], including fuzzy similarities for graphs [6]. Our main research topics are thus concentrated on the main specificities of property graphs compared to graphs:

- fuzzy similarity of (key : $values$) properties to consider the similarity of concepts (e.g. sex and gender in Fig. 1) by using background knowledge;
- choice of aggregation operator, should it be a t-norm, t-conorm or OWA [8];
- consideration of fuzzy property graphs, where nodes and relationships are provided with fuzzy properties that must be taken into account. It is for instance the case when relationships are provided with a degree, or when values are fuzzy (e.g., age is "young" in Fig. 1).

References

1. Angles, R.: The property graph database model. In: AMW (2018)
2. Bouchon-Meunier, B., Coletti, G., Lesot, M.J., Rifqi, M.: Towards a conscious choice of a fuzzy similarity measure: A qualitative point of view. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) Computational Intelligence for Knowledge-Based Systems Design. pp. 1–10. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
3. Bouchon-Meunier, B., Rifqi, M., Bothorel, S.: Towards general measures of comparison of objects. Fuzzy Sets and Systems **84**(2), 143–153 (1996)
4. Kriegel, H.P., Schönauer, S.: Similarity search in structured data. In: Int. Conf. on Data Warehousing and Knowledge Discovery. pp. 309–319. Springer (2003)
5. Liang, Y., Zhao, P.: Similarity search in graph databases: A multi-layered indexing approach. In: Int. Conf. on Data Engineering (ICDE). pp. 783–794. IEEE (2017)
6. Perchant, A., Bloch, I.: Fuzzy morphisms between graphs. Fuzzy Sets and Systems **128**(2), 149–168 (2002)
7. Wang, J., Shen, H.T., Song, J., Ji, J.: Hashing for similarity search: A survey. arXiv preprint arXiv:1408.2927 (2014)
8. Yager, R.R.: Families of owa operators. Fuzzy Sets and Systems **59**(2), 125–148 (1993)

From Process Modelling to Planning Simulation Under Uncertainty in Crisis Rescue Management

Cendrella Chahine[1,2], Thierry Vidal[1], Mohamad El Falou[2], and François Pérès[1]

[1] Ecole nationale d'Ingénieurs de Tarbes, 65000 Tarbes, France
[2] Université de Technologie et de Sciences Appliquées Libano-Française, Lebanon

1 Link Between Design Level and Simulation Level

Based on the dramatic experiences that have faced France and Lebanon such as the Beirut explosion in 2020 and Paris attack in 2015, the response phase challenges are related to information, communication, resources and coordination. The operational management of crisis situations requires the mobilization and rapid coordination of the various emergency services. Unfortunately, this interdepartmental coordination is a very delicate exercise due to the diversity of actors intervening on the ground and the heterogeneity of the different organizations. In addition, actors can work in a specific and aggressive environment, so certain factors can affect or restrict their work. Our work is divided into two levels: the design level and the simulation level.

1.1 Design Level

It is essential to analyze the emergency plan to improve the response. The existing emergency plans are available in text form defining the actors involved and their roles. In this level, we use BPMN standard [4] in our functional modeling which represents the relationship between the different activities as interactions between them. This relationship can be external and then represents the relationship between the emergency services (resources). Resource allocation is also present where each process represents the activities to be performed by a specific service. Functional modeling allows us to represent each service separately. Based on previous experiences, unforeseen events can disrupt the actors tasks on site. In this context, we propose a new FMECA [1] approach by adding levels and durations to causes and effects. This dysfunctional modelling allows us to identify the sources of uncertainty that can occur on each activity of the BPMN process. Each activity is represented as a box in SADT [6] that transforms input elements into output elements by taking some limits into account. In addition, it represents which service performs it. In this context, we represented three levels of causes and three levels of consequences in an event tree by adding to each event a duration. The functional and dysfunctional modelling allowed us to see the indicators of each activity in order to analyse its vulnerability.

1.2 Simulation Level

Our terrorist attack scenario is represented in a MAS [5] and simulated in different architectures that we propose in order to compare the behaviours of agents in each of them. This level demonstrates the level of autonomy that can be given to an operational agent. In planning and scheduling [2, 3], we develop algorithms to allow each agent to choose the appropriate standard plan and insert it into its local plan. The approach chosen for each agent to generate and execute its plan is based on the dysfunctional analysis of the design level.

References

1. Deng, Y., Song, L., Zhou, J., Xu, N., Ni, G., Wang, L.: Analysis of Failures and Influence Factors of Critical Infrastructures: A Case of Metro. Advances in Civil Engineering, **2020**, 1-13 (2020).
2. Bidot, J., Vidal, T., Laborie, P., Beck, J.: A theoretic and practical framework for scheduling in a stochastic environment. Journal of Scheduling, **12**(3), 315-344 (2008).
3. Casanova, G., Pralet, C., Lesire, C. and Vidal, T.: Solving dynamic controllability problem of multi-agent plans with uncertainty using mixed integer linear programming. In 'ECAI 2016', The Hague, Netherlands, pp. 930–938. URL: https://hal.archives-ouvertes.fr/hal-02487165
4. Kluza, K., Wiśniewski, P., Jobczyk, K., Ligeza, A., Suchenia, A.: Comparison of selected modeling notations for process, decision and system modeling. In: Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 1095-1098 (2017).
5. El Falou, M., Bouzid, M., Mouaddib, A-I., Vidal, T.: A distributed multi-agent planning approach for automated web services composition. Web Intelligence and Agent Systems Int'l Journal. vol. **10**(4), p.423-445 (2012).
6. Ahmed, F., Robinson, S. and Tako, A.A.: Using the structred analysis and design technique in simulation conceptual modeling. In: Proceedings of the Winter Simulation Conference, pp. 1038–1049 (2014).

A bipolar approach to solve general optimization problems subject to fuzzy relation equations

M. Eugenia Cornejo, David Lobo, and Jesús Medina

Department of Mathematics. University of Cádiz. Spain
{mariaeugenia.cornejo,david.lobo,jesus.medina}@uca.es

In this work, we present a method to solve a general optimization problem constrained by a fuzzy relation equation, under the requirement of having a mixed monotonic objective function, that is, the objective function is order-preserving in some arguments and order-reversing in the remaining arguments. In particular, the developed procedure can be applied to any linear optimization problem subject to a fuzzy relation equation. The underlying idea of the method consists of defining an equivalent optimization problem subject to a bipolar fuzzy relation equation with an order-preserving objective function. As a result, the solutions of the original optimization problem can be computed from the extremal solutions of the bipolar fuzzy relation equation constraint. To this aim, we can make use of some previous works which provide the solution set of a bipolar fuzzy relation equation [1–4].

One of the highlights of the presented approach relies on the fact that the solutions of the optimization problem are analytically described, unlike other existing studies which only provide approximations of such solutions [5–7].

References

1. 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*, 42(17):5779–5793, 2019.
2. 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*, 354:520–532, 2019.
3. 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*, 410:1–18, 2021.
4. 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*, 2022.
5. S.-C. Fang and G. Li. Solving fuzzy relation equations with a linear objective function. *Fuzzy Sets and Systems*, 103(1):107–113, 1999.
6. R. Hassanzadeh, E. Khorram, I. Mahdavi, and N. Mahdavi-Amiri. A genetic algorithm for optimization problems with fuzzy relation constraints using max-product composition. *Applied Soft Computing*, 11(1):551–560, 2011.

Cornejo et al.

7. Y.-K. Wu and S.-M. Guu. Minimizing a linear function under a fuzzy max–min relational equation constraint. *Fuzzy Sets and Systems*, 150(1):147–162, 2005.

Metaheuristics as solutions producers

Hanane El Raoui[1], Marcelino Cabrera[2], and David A. Pelta[3]

[1] Dept. of Management Science, Strathclyde Business School, Glasgow, Scotland G4 0QU, United Kingdom
[2] Dept. of Languages and Information Systems, Universidad de Granada, Spain
[3] Dept. of Comp. Science and A.I., Universidad de Granada, Spain
dpelta@ugr.es

Numerous decision criteria and objectives cannot be clearly incorporated in the computer model formulation of most 'real world' decision-making and/or optimization issues; alternatively, if included, they may raise the complexity of such models to the point that they cannot be solved correctly. Final decisions are frequently made based not just on simulated objectives, but also on the aims, biases, and preferences of the decision makers [1].

Many mathematical optimization algorithms are committed to determining sets of non-inferior solutions to multi-objective formulations or, in the best case, single optimum solutions to single-objective issues [2]. In any scenario, it's important to remember that we're trying to find the best answer for the model.

Considering the existence of unmodeled (or difficult to model) goals and parameters, non-conventional solving methodologies are required to not only search the decision space for optimum solutions, but also to explore the decision area for high-quality alternatives. Such solutions may be sub-optimal from the model's standpoint, but they are helpful from other perspectives.

Hence, metaheuristics' role as solution generators becomes important in two ways: first, because several runs (or a single run in the case of population-based techniques) allow for the generation of a set of potentially good solutions, and second, if a reference solution is available, a new optimization problem can be set up to generate solutions with similar quality but maximally different structure.

The goal of this contribution is to explore how metaheuristics may be used to come up with solutions in a simple problem-solving framework. We look at the problem of tourist trip planning that entails two key elements: the selection of points of interest that are appropriate for the tourist's likes, and the design of itineraries to visit these points (all or a subset of them).

To solve the tourist trip design problem with time-dependent scores (TTDP-TDS), we start with a collection of N POIs. Each node $i \in N$ has an associated interest or score, as well as numerous recommendation factors that weight the interest based on the time of visit. However, because there is a maximum allowed time, not all nodes in the set can be visited. This problem is solved using a crossover-less evolutionary algorithm (EA). A GRASP-like procedure is used to generate members of the initial population. Setting randomly an initial POI, a list of potential POIs is created, and a random element from the list is picked and added to the solution. The procedure is repeated if the time limit is not exceeded. Each generation, K parents are picked by roulette wheel selection from the present population and merged into an intermediate population. A child solution is produced from each parent using one of the numerous mutation operators available. The new solution is integrated into the intermediate population if it is feasible and better than the parent. Finally, the intermediate population is sorted, with the best elements chosen to replace the original population. We will apply the Jaccard's similarity coefficient to measure how similar two solutions *A*, *B* (seen as two sets of POIs) are to identify the relationship between interest and solution diversity. The solutions are highly efficient in terms of decreasing travel time while enhancing visit time. The results clearly emphasize the importance of having a variety of alternatives to select from rather than focusing all the efforts on finding the best one.

As further work, we will explore the Modeling to Generate Alternative Approach (MGA) to produce a set of alternative solutions to our problem. This approach was successfully used in [3,4] to generate a quit diverse set of solutions for the perishable food distribution problem. while also connecting the analysis of the set of solutions with multicriteria decision-making approaches.

Acknowledgement

Authors acknowledge support from 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).

References

1. 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. Raoui, Hanane El, 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. 255-273.

Vectorial extensions of Fuzzy Integrals on sequential information fusion

Mikel Ferrero-Jaurrieta[1], Zdenko Takáč[2], Lubomíra Horánská[2], Graçaliz Dimuro[3], Javier Fernández[1], and Humberto Bustince[1]

[1] Departament of Statistics, Computer Science and Mathematics, Campus Arrosadía, s/n, 31006, Spain.
{mikel.ferrero, fcojavier.fernandez, bustince}@unavarra.es
[2] Institute of Information Engineering, Automation and Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Bratislava, Slovakia.
{zdenko.takac, lubomira.horanska}@stuba.sk
[3] Centro de Ciencias Computacionais, Universidade Federal do Rio Grande, Rio Grande, 96077540, Brazil.
gracalizdimuro@furg.br

Fuzzy integrals are a well-known type of aggregation functions since they are able to consider and model the possible interaction between data. Examples of these are Choquet [1] and Sugeno integrals [2] and its generalizations [3, 4], which have had important advances in fuzzy rule based classification systems, brain-computer interfaces, multi-criteria decision making, image processing, convolutional neural networks or community detection.

Until now different approaches of the discrete fuzzy integrals have been introduced for fusing different information structures, such as Choquet integral for intervals, with respect to admissible orders [5] and with respect to admissible permutations. However, since the Fuzzy integrals are defined for punctual or interval-valued entries, they can not be used as the aggregation operator admitting n-dimensional vector inputs, as required, for example, for aggregating vectors in recurrent neural networks models.

To do so, in this work we present the vectorial extensions of Fuzzy integrals: n-dimensional extensions of the discrete Choquet-like and Sugeno integrals, such that the inputs are n-dimensional vectors, and recovering an n-dimensional vector as output. For both cases, the main definition and the different construction methods are presented and after the main properties of both vectorial extensions are studied.

Finally, as an application of the introduced theory, we use it in different recurrent neural networks units-based models [6]. These recurrent neural networks are modified in their respective information fusion process in order to take into account the possible coalition between data, so the introduced vectorial extensions of fuzzy integrals are used for it.

To check out the performance of this modification, we have tested it in order to solve text classification/sentiment analysis related problems. Results are compared with standard fusion method and also with an order statistic, like the maximum function. We show, validated by statistical analysis, that in most cases vectorial extension fuzzy integral-based models outperform standard ones.

Acknowledgments

The authors gratefully acknowledge the financial support of the Grant PID2019-108392GB-I00 funded by MCIN/AEI/10.13039/501100011033 as well as CNPq (Proc. 301618/2019-4), FAPERGS (Proc. 19/2551-0001660) and Grant VEGA 1/0267/21 and to Tracasa Instrumental and the Immigration Policy and Justice Department of the Government of Navarre.

References

1. Choquet, G.: *Theory of capacities*. In: Annales de l'Institut Fourier, vol. 5, pp. 131–295 (1954)
2. Sugeno, M.: *Theory of fuzzy integrals and its applications*, PhD thesis (1974).
3. Bustince, H., Mesiar, R., Fernández, J., Galar. M., Paternáin, D., Altahi, A., Dimuro, G.P., Bedregal, B. and Takác, Z.: *d-Choquet integrals: Choquet integrals based on dissimilarities*. In: Fuzzy Sets and Systems, vol. 414, pp. 1–27 (2021).
4. Dimuro, G.P., Lucca, G., Bedregal, B., Mesiar, R., Sanz, J., Lin, C.-T., Bustince, H.: *Generalized CF_1F_2-integrals: From Choquet-like aggregation to ordered directionally monotone functions*. In: Fuzzy Sets and Systems, vol. 378, pp. 44–67 (2020).
5. Bustince, H., Galar, M., Bedregal, B., Kolesárová, A., Mesiar, R.: *A New Approach to Interval-Valued Choquet Integrals and the Problem of Ordering in Interval-Valued Fuzzy Set Applications*. In: IEEE Transactions on Fuzzy Systems, vol. 21, 6, pp. 1150-1162 (2013).
6. Graves, A.: *Supervised Sequence Labelling with Recurrent Neural Networks*. In: Studies in Computational Intelligence, Springer, Berlin, Heidelberg (2012).

Parzen-Rosenblatt Window Density Estimation for Restor a Markov Chains Corrupted by NonGaussian Noise

Ali Hamache

Ecole Militaire Polytechnique
PO Box 17, Bordj El Bahri, Algiers 16111
Algeria
ali_hamache@hotmail.fr

The use of the hidden Markov chain model for image classification has known a considerable success. The interest of Markovian modeling of images is the consideration of the contextual information. The idea behind this abstract is to elaborate a new model for the estimation of the parameters of a Markov chain based on Parzen-Rosenblatt window density estimation. The Hidden Markov Chain has a great reputation in multiple fields especially in the field of signal and image processing. They are used like regularisation tool that allows making decisions from a set of data sample while bypassing the problems of complexity of the prohibitive calculation. Such success is due essentially to the existence of multiple techniques and methods associated to, which enable to make different estimation's operations in reasonable time, even if the data sample is enough huge.

Markov's models are well adapted to these problems in reasons of existence sampling algorithms and iterative estimation. In particular, the hidden Markov chains offer quick estimation methods and segmentation using forward-backward recurrence[2], possible thanks to Markovianity of the posterior rule.

Indeed, the posterior distribution of the hidden process can be seen as the fusion of information a priori and observation.
Image classification is among the most important problems in imagery. The development of HMC allowed bypassing rigorously the problems related to prohibitive calculation time due to the important number of pixels in an image. The advantage of those methods is the ability to make a local decision from the all available information in the image. The success of such methods is due to their aptitude to produce good results sometimes exceed human eyes even the presence of noise in the image.

In this abstract, a new classification multi-modal approach which highlights missing, uncertain and imprecise data through the evidence theory and takes into account spatial information through the HMC is shown. From related works we can take a note that an interesting link has been established between HMC and Parzen-Rosenblatt window density by estimating parameter [1],[3],[4].

So,the purpose is to ameliorate the classification process. This approach is base on the data modeling by the Hidden Markov Chains, they allow to take into account the spatial information and offer an estimation methods and quick segmentation tools using forward-backward recurrences. The approach consists on four steps. The first correspond to initialisation the HMC parameters (λ_0), the second step consist on new parameters estimation using an stochastic evidential version(noise density estimation with evidence theory). In the third step, we compute the *posteriori* marginal probability. The last step allow to make decision by MPM technique for the determination of the decision(labels). It should be noted that we need to transform the bi-dimensional set of pixels into a mono-dimensional set. So, to vectorize the images we need a tool to make it, such as *Hilbert-Peano* transformation [5].

The priliminary findings prove that this approch can be considered as a novel classification method applicable to images multispectral analysis based on hidden Markov chains and Parzen-Rosenblatt window, in order to exploit data for better decision-making.

References

1. Hamache, A., Boudaren, M.E.Y., Boukersoul, H., Debicha, I., Sadouk, H., Zibani, R., Habbouchi, A., Merouani, O.: Uncertainty-Aware Parzen-Rosenblatt Classifier for Multiattribute Data.International Conference on Belief Functions,103–111 (2018)
2. Juang,B.H., Rabiner, L.R.: Automatic speech recognition–a brief history of the technology development.
3. Bowman, A.W., Azzalini, A.: Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations, vol. 18. OUP Oxford (1997)
4. Wand, M.P., Jones, M.C.: Kernel smoothing. Crc Press (1994)
5. Giordana, N., Pieczynski, W.: Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation, IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 5, 465-–475 (1997)

Introducing Approximation-based Model Diagnostics into KnowledgePit – A Platform for Organizing Data Mining Challenges

Andrzej Janusz[1,2], Andżelika Zalewska[2], and Dominik Ślęzak[1,2]

[1] Institute of Informatics, University of Warsaw, Warsaw, Poland
[2] QED Software, Warsaw, Poland
andrzej.janusz@qed.pl

Data mining challenges are getting continually growing attention among the data science community. Challenge platforms, such as Kaggle[3], attract thousands of machine learning community members to participate in challenges aiming at solving real-world data science problems. Such challenges not only address specific machine learning tasks but often facilitate more general and innovative applications of data mining techniques and machine learning algorithms. On the one hand, they are appealing to researchers and students for whom competitive challenges can be a source of new interesting data sets and research topics. Moreover, they can also be an attractive addition to academic courses for students who are interested in practical applications of their knowledge. On the other hand, setting up a public machine learning challenge related to a given real-world task is a form of outsourcing this task to the community [6]. It can be beneficial to the organizers and sponsors who define the challenge since it is an inexpensive way to solve the problem that they investigate [5].

KnowledgePit is a Poland-based data mining challenge WWW platform[4], maintained and developed by QED Software[5]. Originally, it was a non-commercial project focusing on collaboration between researchers and international conferences, as well as the promotion of data science in the academic community. In other words, it was created as a hub for students, scientists, and machine learning experts, in which they could compete in solving real-life data science problems. Recently, however, together with companies such as Security On-Demand [2] or Tibco [1], we organized a number of successful challenges for our industrial partners. Consequently, in order to provide even more value to future challenge sponsors, we are currently extending KnowledgePit's functionalities to utilize the XAI tools for the purpose of advanced post-challenge research, including the analysis of the reasons of mistakes made by submitted solutions. We also believe that such functionalities may be useful from an education perspective, whereby KnowledgePit can be considered as a platform for assessing and improving the data science competencies at universities, as well as in companies.

The diagnostics of mistakes made by the challenge participants, as well as investigation of similarities between their submitted models [3], is possible thanks

[3] Kaggle, https://www.kaggle.com
[4] KnowledgePit, https://knowledgepit.ai/
[5] QED Software, https://qed.pl/

to the outcomes of one of our R&D projects, called BrightBox. Therein, a diagnosed model is approximated by rough-set-based surrogate models, taking a form of ensembles of decision reducts [4]. Then, the BrightBox algorithms analyze particular data cases by means of their neighborhoods – other data cases that are classified similarly by the generated reducts (i.e.: objects which support the same decision rules induced by reducts). For example, if a mistake observed for a given object often repeats also within its neighborhood, then we can conclude that the diagnosed model was not trained correctly for such objects from the very beginning. As another example, if the neighborhood is almost empty (i.e.: the corresponding diagnosed object seems to be classified by completely different rules than objects observed so far), then we can conclude that this is a relatively new situation for which the model has not been prepared.

Acknowledgements This research was co-funded by the Polish National Centre for Research and Development in frame of project MAZOWSZE/0198/19.

References

1. Janusz, A., Hao, G., Kałuża, D., Li, T., Wojciechowski, R., Ślęzak, D.: Predicting Escalations in Customer Support: Analysis of Data Mining Challenge Results. In: Proceedings of the 2020 IEEE International Conference on Big Data, IEEE Big Data 2020, Atlanta, GA, USA, December 10-13, 2020. pp. 5519–5526 (2020). https://doi.org/10.1109/BigData50022.2020.9378024
2. Janusz, A., Kałuża, D., Chądzyńska-Krasowska, A., Konarski, B., Holland, J., Ślęzak, D.: IEEE BigData 2019 Cup: Suspicious Network Event Recognition. In: Proceedings of the 2019 IEEE International Conference on Big Data, IEEE Big Data 2019, Los Angeles, CA, USA, December 9-12, 2019. pp. 5881–5887 (2019). https://doi.org/10.1109/BigData47090.2019.9005668
3. Janusz, A., Przyborowski, M., Biczyk, P., Ślęzak, D.: Network Device Workload Prediction: A Data Mining Challenge at Knowledge Pit. In: Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020, Sofia, Bulgaria, September 6-9, 2020. pp. 77–80 (2020). https://doi.org/10.15439/2020F159
4. Ślęzak, D., Grzegorowski, M., Janusz, A., Kozielski, M., Nguyen, S.H., Sikora, M., Stawicki, S., Wróbel, L.: A Framework for Learning and Embedding Multi-sensor Forecasting Models into a Decision Support System: A Case Study of Methane Concentration in Coal Mines. Information Sciences **451-452**, 112–133 (2018). https://doi.org/10.1016/j.ins.2018.04.026
5. Tauchert, C., Buxmann, P., Lambinus, J.: Crowdsourcing Data Science: A Qualitative Analysis of Organizations' Usage of Kaggle Competitions. In: Proceedings of the 53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7-10, 2020. pp. 1–10 (2020). https://doi.org/10.24251/HICSS.2020.029
6. Zimmermann, J.L.: Data Competitions: Crowdsourcing with Data Science Platforms. In: The Machine Age of Customer Insight, pp. 183–197. Emerald Publishing Limited (2021). https://doi.org/10.1108/978-1-83909-694-520211017

DT-EVCLUS: Decision Tree-based Evidential Clustering

Lianmeng Jiao, Haoyu Yang, and Quan Pan

School of Automation, Northwestern Polytechnical University
Xi'an, 710072, People's Republic of China
{jiaolianmeng,quanpan}@nwpu.edu.cn
2020202437@mail.nwpu.edu.cn

In the field of clustering, one of the remaining open problems concerns the description and quantification of cluster-membership uncertainty. A large number of clustering algorithms are proposed by incorporating fuzzy sets, possibility theory, and rough sets to characterize the uncertainty of cluster membership. Recently, evidential clustering [1, 2] has been developed as a promising clustering framework, which uses Dempster-Shafer (DS) belief function theory as an uncertainty model. The defined evidential partition generalizes those hard, fuzzy, possibilistic and rough partition. A survey on evidential clustering algorithms can be found in [3]. Although these evidential clustering algorithms can well model the cluster-membership uncertainty, the resulting cluster assignments are difficult to understand in terms of human cognition, which limits their applications in security, privacy or ethic related fields.

In this study, the unsupervised decision tree model is introduced into the evidential clustering framework to improve the interpretability of the evidential partition. A **D**ecision **T**ree-based **EV**idential **CLUS**tering (DT-EVCLUS) algorithm is proposed to build an evidential clustering decision tree, which uses a path from root to leaf node to achieve the interpretability of each cluster. The proposed algorithm is composed of three main modules, i.e., evidential partition, cutting points evaluation, and cluster merging, in which the first two modules are carried out repeatedly to build an initial unsupervised decision tree and the final module is designed to combine similar leaf nodes to form the final compact clustering model if the number of clusters are available. The proposed algorithm can work in the situation whether the number of clusters is known or not. It can optimize an internal clustering validation index called silhouette metric, which considers both intra-cluster distance and inter-cluster distance to automatically determine the suitable number of clusters. Both synthetic and real data sets are used to evaluate the performance of the proposed algorithm, and the experimental results demonstrate the good performance of the proposal compared with some representative decision tree-based clustering algorithms and evidential clustering algorithms.

References

1. Masson, M.H., Denoeux, T.: ECM: an evidential version of the fuzzy c-means algorithm. Pattern. Recog. **41**(4), 99–110 (2016)

2. Jiao, L., Denoeux, T., Liu, Z.G.,: EGMM: an evidential version of the Gaussian mixture model for clustering. arXiv preprint arXiv: 2010.01333 (2022)
3. Denoeux, T., Kanjanatarakul, O.: Evidential clustering: A review. In: Proceedings of the 5th International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making, pp. 24–35. Springer, Cham (2016)

Uncertainty Measures for Active Learning over Imbalanced Data

Daniel Kałuża[1,2], Andrzej Janusz[1,2], and Dominik Ślęzak[1,2]

[1] Institute of Informatics, University of Warsaw, Warsaw, Poland
[2] QED Software, Warsaw, Poland
daniel.kaluza@qed.pl

Machine learning has achieved spectacular results in various fields, ranging from biology, e.g. AlphaFold2 model for protein third-order structure prediction [3], through computer vision, e.g. skin cancer cells classification [1], to NLP sentiment analysis [6]. However, models of this quality would not be possible without an access to labels for training cases, and the work of human annotators needed to acquire them. Active learning focuses on reducing human labor by interactively asking annotators for labels of samples that are expected to bring the biggest improvement to the generalization performance of the model, and therefore, obtain a better quality of the model with a limited labeling budget [2]. Active learning is also the core idea behind *Label in the Loop* (LITL) – the research project conducted by QED Software with the aim of assisting and assessing subject matter experts involved in the annotation processes.

One aspect of active learning is selecting samples (objects, cases, data items) to be presented to annotators. The sample selection (or sampling, in other words) can be perceived as a decision making process. It can be driven by a number of heuristic laws referring to characteristics of both, samples and experts. One of such laws states that the most informative samples are those for which the current model (being continuously trained) is most uncertain. In the literature, the uncertainty of a model's output is usually expressed by means of Shannon entropy or other model-agnostic measures which reach the maximum value for the discrete uniform distribution of decision class probabilities [5]. This approach is designed to work well for balanced machine learning tasks [4]. However, models created for imbalanced problems are often evaluated using prior-sensitive metrics, such as balanced accuracy. They are optimized to return the prior (rather than uniform) distribution as a prediction in the case of the highest uncertainty. This causes a kind of incompatibility which needs to be addressed.

We propose to consider prior-sensitive uncertainty measures which utilize the knowledge about the decision class distribution of available samples. We show that a simple rescaling trick can be used to shift the maximum of typical measures to any given prior. Moreover, we introduce a new family of uncertainty measures that have some appealing properties. On the one hand, the considered measures are prior-sensitive, and on the other hand, they prioritize selecting samples near the decision border but not exactly from the border as the uncertainty of such samples often includes a significant aleatoric (and thus irreducible) factor. We argue that such measures are better-suited for active learning tasks when the prediction model is being optimized for a prior-sensitive quality metric, and we demonstrate experimental results on several data sets confirming our claims. Finally, we show how to take advantage of our research outcomes in the technology developed within the aforementioned LITL project.

Acknowledgements This research was co-funded by Smart Growth Operational Programme 2014-2020, financed by European Regional Development Fund, in frame of project POIR.01.01.01-00-0213/19, operated by National Centre for Research and Development in Poland.

References

1. Skin Cancer Detection: Applying a Deep Learning based Model driven Architecture in the Cloud for Classifying Dermal Cell Images. Informatics in Medicine Unlocked **18**, 100282 (2020). https://doi.org/https://doi.org/10.1016/j.imu.2019.100282
2. Cassidy, R., Charles, E.S., Slotta, J.D.: Editorial: Active Learning: Theoretical Perspectives, Empirical Studies, and Design Profiles. Frontiers in ICT **6** (2019). https://doi.org/10.3389/fict.2019.00003
3. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zieliński, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D.: Highly Accurate Protein Structure Prediction with AlphaFold. Nature **596**(7873), 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2
4. Settles, B.: Active Learning Literature Survey. Computer Sciences Technical Report 1648, University of Wisconsin-Madison, USA (2009), http://axon.cs.byu.edu/∼martinez/classes/778/Papers/settles.activelearning.pdf
5. Wang, H., Chang, X., Shi, L., Yang, Y., Shen, Y.D.: Uncertainty Sampling for Action Recognition via Maximizing Expected Average Precision. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, IJCAI 2018, Stockholm, Sweden. pp. 964–970. ijcai.org (2018). https://doi.org/10.24963/ijcai.2018/134
6. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., Le, Q.V.: XLNet: Generalized Autoregressive Pretraining for Language Understanding. In: Proceedings of the 33rd Conference on Neural Information Processing Systems, NeurIPS 2019, Vancouver, Canada. Curran Associates Inc. (2019), https://proceedings.neurips.cc/paper/2019/file/dc6a7e655d7e5840e66733e9ee67cc69-Paper.pdf

What If There Are Too Many Outliers

Olga Kosheleva and Vladik Kreinovich

University of Texas at El Paso, El Paso, TX 79968, USA
{olgak,vladik}@utep.edu

Measurements in the ideal case, when there are no outliers. Many measuring instruments measure several quantities at the same time. For example, a meteostation measures temperature, humidity, wind speed, etc.

Let d denote the number of measured quantities, and let $x = (x_1, \ldots, x_d)$ denote the actual values of the measured quantities. Measurements are never absolutely accurate, so the measurement result $x_j = (x_{j1}, \ldots, x_{jd})$ is, in general, different from the actual values x: there is a non-zero difference $x_j - x$, which is known as the *measurement error*. Usually, we have some information about the possible values of the measurement error. For example, we may know the upper bound Δ_i on the absolute value of each component $x_{ji} - x_i$ of the measurement error. In this case, $x_j - x \in [-\Delta_1, \Delta_1] \times \ldots \times [-\Delta_d, \Delta_d]$. In other cases – of probabilistic uncertainty – we know that the expression $\sum\limits_{i=1}^{d} (x_{ji} - x_i)^2 \cdot \sigma_i^{-2}$ does not exceed some constant $\chi^2_{d,\alpha}$. In all such cases, we know that $x_j - x \in U$ for some convex set U – parallelepiped in the Δ_i-case, ellipsoid in the probabilistic case, etc.

If we measure the same (unknown) quantity n times and all n readings $x_1, \ldots, x_n$ of the measuring instrument are measurement results, then from the condition that, for each j, we have $x_j - x \in U$, hence $x \in x_j - U = \{x_j - u : u \in U\}$, we conclude that the set of possible values of x is the intersection $\bigcap\limits_{j=1}^{n} (x_j - U)$ of these n sets.

What if there are outliers. Measuring instruments sometimes malfunction. As a result, in practice, for some of the readings x_j, the difference $x_j - x$ is outside the set U. Such readings can be viewed as particular case of *outliers* – in addition to cases when some actual values are very different from most others. For example, suppose that we several measure temperature with accuracy 1 degree, so that $U = [-1, 1]$, and we get the values 24.5, 25.1, 24.9, and 35. These cannot be all results of measuring the same quantity, since no value can be 1-close to both 24.5 and 35. If we know that the measuring instrument is reasonable reliable and that it can malfunction at most one out of four times, then clearly 35 is an outlier, and the actual value of the temperature belongs to the intersection of the sets $x_j - U$ corresponding to the remaining three reading, i.e., to the set

$$[24.5 - 1, 24.5 + 1] \cap [25.1 - 1, 25.1 + 1] \cap [24.9 - 1, 24.9 + 1] = [24.1, 25.5].$$

But what if we have many outliers? When we have fewer outliers than correct measurements, we can usually detect which readings are outliers and thus, find a reasonable estimate of the actual value x. However, in many cutting-edge measurements, we have fewer correct measurements than outliers, i.e., the proportion ε of correct measurements is smaller than $1/2$. In this case, we do not know which readings are outliers and which are not, so we may have several different possible estimates of x. For example, if in the above 4-readings case, all we know is that at least one fourth of readings are correct, then the actual temperature can be either close to 25 or close to 35 (35 if the only correct measurement is the last one).

In such situations, methods of robust statistics – that work well when there are few outliers – do not work, so all we can conclude is that there are several possible values x; in principle, we can simply say that x is U-close to one of the n readings $x_1, \ldots, x_n$. But this number of readings can be huge, and processing such a huge set can be very time-consuming. It it is therefore desirable to come up with a set of possible values of x which has much fewer than n elements. This is exactly what we do in this paper.

Definition. *Let $X = \{x_1, \ldots, x_n\}$ be a set of points in a d-dimensional space, and let $k < n$ be an integer. By the set X's (n, k)-*compression*, we mean a finite set S with the following property:*

- *for every pair (U, x), for which $U \in \mathbb{R}^d$ is a convex set containing 0, $x \in \mathbb{R}^d$, and $x_j - x \in U$ for at least k different indices j,*
- *there exists en element $s \in S$ for which $s - x \in U$.*

Proposition. *For every d, and for every $\delta > 0$, there exists a constant $c_{d,\delta}$ such that for all ε, if we take $k = \varepsilon \cdot n$, then for every set of n points, there exists an (n, k)-compression with no more that $c_{d,\delta} \cdot \varepsilon^{-(d-0.5+\delta)}$ elements.*

Proof is based on results from [1, 2].

Discussion. Good news is that the above upper bound on the number of elements in a compression does not depend on n at all. We can have billions of measurement results, but we can always compress this information into a finite set whose size remains the same no matter what n we choose.

Acknowledgments. This work was supported by the AT&T Fellowship in Information Technology, by the US National Science Foundation grants 1623190, HRD-1834620, and HRD-2034030, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The authors are thankful to the anonymous referees for valuable suggestions.

References

1. Bárány, I.: Combinatorial Convexity. American Mathematical Society, Providence, Rhode Island (2021)
2. Rubin, N.: Stronger bounds for weak ε-nets in higher dimensions. Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing STOC 2021, Rome, Italy, June 21–25, 2021, pp. 989–1002 (2021)

An approach to Fuzzy Multi-Objective Shortest Path Problems

Irene Mariñas-Collado[1], Susana Montes[1], and Agustina Bouchet[1]

Department of Statistics and Operation Research and Mathematics Didactics, University of Oviedo, 33007 Oviedo, Spain; marinasirene@uniovi.es (I.M.C), montes@uniovi.es (S.M.) bouchetagustina@uniovi.es (A.B.)

One of the most studied classical optimisation problems is the *Shortest Path Problem* (SPP). It is well-known that the SPP consists in finding a path between two vertices (or nodes) in a network so that the sum of its edges' weights is minimised. The history of SPP, together with classical algorithms and applications are reviewed in [7]. As the number of graphs and data used grows with each new application or desired functionality, the development of new algorithms and speedup strategies to handle SP challenges remains an active research topic. Nonetheless, Djikstra's shortest path algorithm [1], which uses dynamic programming, continues to be the current standard for SPP for the case of positive weights.

The value of the path is normally measured in terms of a single attribute (cost, duration, time, risk, etc.) defined in each edge of the input graph. However, in many cases, a single attribute is insufficient to define the preference between routes. As a result, *Multi-Objective Shortest Path* (MOSP) problems arise, in which various attributes are defined on the edges and thus on the paths. This scenario leads to solution sets that can be exponentially sized relative to the input size of the problem. A good review of MOSP models and adaptation of standard algorithms is given by [8]. Numerous approaches involve the transformation of the multi-objective problem into a single criterion optimisation by aggregating or combining the different criteria. But the resulting solution to the single objective optimization problem is usually subjective to the parameter settings chosen by the user [6]. One of the most popular methods for solving MOSP problems is the construction of approximate Pareto sets. In general terms, a set of Pareto paths includes all incomparable 'best' paths [4].

Moreover, in traditional SPPs, there is exact information about the parameters of the problem such as time, costs, and risk. However, real-world environments require dealing with uncertainty and so fuzzy notions can be used. In Fuzzy SPP, the classic costs between nodes are replaced by fuzzy numbers. Despite their potential, fuzzy shortest path problems have received less attention. The fuzzy shortest path problem was first analysed by [2].

The proposed approach is based on existing techniques that can be combined. This involves: 1) Map the fuzzy costs into crisp numbers; 2) Find an approximate Pareto set of optimal solutions; 3) Use ranking methods to establish preferences between the optimal solutions.

There are different approaches to solving the SPP with fuzzy numbers, but one commonly used approach is to use different existing methods of defuzzification, that is, to convert the fuzzy numbers into crisp numbers with which to solve the MOSP problem. This is, for now, the simpler faster solution to combine the fuzzy SPP with the MOSP problem.

To solve MOSP problems, [3] proposed a generalisation of Dijkstra's algorithm with two criteria. This algorithm was generalised to any number of criteria by [5]. Pareto sets are not always large in practice, especially if the criteria are correlated. However, when the number of criteria grows and some are negatively correlated, Pareto set sizes can be unpractical. For this reason, it is important to specify some sort of order amongst the paths contained in the Pareto set. It is proposed to do so by using ranking methods.

Ranking methods are usually studied in the field of social choice theory in which several voters express their preferences over a set of alternatives in the form of rankings. In this scenario, each criterion can be considered a voter, and all the paths in the set can be ordered according to each criterion. The problem is then how to aggregate these rankings in order to rank all the optimal paths. The aim of this approach is to facilitate an ordered list of the best possible paths, according to all criteria, and where the different criteria can be given more or less weight in the voting process, based on different preferences.

Acknowledgements

This research has been partially supported by Spanish MINECO projects PGC2018-098623-B-I00 (I.M-C. and S.M.) and TIN2017-87600-P (A.B.).

References

1. Dijkstra, E.W., et al.: A note on two problems in connexion with graphs. Numerische mathematik **1**(1), 269–271 (1959)
2. Dubois, D.J.: Fuzzy sets and systems: theory and applications, vol. 144. Academic press (1980)
3. Fandel, G., Gal, T.: Multiple Criteria Decision Making Theory and Application: Proceedings of the Third Conference Hagen/Königswinter, West Germany, August 20–24, 1979, vol. 177. Springer Science & Business Media (2012)
4. Hanusse, N., Ilcinkas, D., Lentz, A.: Framing algorithms for approximate multicriteria shortest paths. In: 20th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2020). vol. 85, pp. 11–1 (2020)
5. Martins, E.Q.V.: On a multicriteria shortest path problem. European Journal of Operational Research **16**(2), 236–245 (1984)
6. Odu, G., Charles-Owaba, O.: Review of multi-criteria optimization methods–theory and applications. IOSR Journal of Engineering (IOSRJEN) **3**(10), 1–14 (2013)
7. Schrijver, A.: On the history of the shortest path problem. Documenta Mathematica **17**(1), 155–167 (2012)
8. Tarapata, Z.M.: Selected multicriteria shortest path problems: An analysis of complexity, models and adaptation of standard algorithms. International Journal of Applied Mathematics & Computer Science **17**(2), 269–287 (2007)

Extension of the Two-Sided Normalized probability possibility transformation to the multivariate case

Gilles Mauris, Laurent Foulloy
[1] LISTIC, Univ. Savoie Mont Blanc, F-74000 Annecy
BP 80 439, 74 944 Annecy Le Vieux, France
gilles.mauris@univ-smb.fr,laurent.foulloy@univ-smb.fr

Multivariate data analysis has long been an outstanding issue where the probability theory plays a major role. Recently, different attempts have been made in the literature in terms of possibility theory in order to represent multivariate data. The multivariate possibility distribution can either be defined directly [1], by combining marginal possibility distributions with a t-norm [2][3][4], or by transforming a multivariate probability into a possibility one [5][6][7][8]. In this paper we follow the latter way.

Until now the transformation of multivariate probability distribution is based on the extension of the univariate optimal distribution, that consist in identifying the cuts of a probability distribution to alpha-cuts of the equivalent possibility distributions [9][10]. The optimal transformation is widely used in application fields where the probability distribution is symmetric [11]. The asymmetric case is seldom considered and in fact the optimal transformation does not preserve the asymmetry [12]. Thus previously, we proposed a new univariate probability possibility transformation called two-sided normalized (TSN) transformation [13][14]. In this work we propose to extend this new transformation to the multivariate case.

The most natural way to extend the univariate TSN gives the following expression:

$$\forall x \in R^n, \pi^{TSN}(\vec{x}^n) = \min\left(\frac{F^n(\vec{x}^n)}{F^n(\vec{m}^n)}, \frac{1-F^n(\vec{x}^n)}{1-F^n(\vec{m}^n)}\right) \qquad (1)$$

with F^n the n-joint cumulative distribution and $\vec{x}^n$ the variable vector of n-dimension and $\vec{m}^n$ the vector of the n-dimensional mode. It can be proven that this transformation satisfies the consistency principle of Dubois et al. [15].

Next the following questions are of interest: does this transformation preserve the shape and the asymmetry of the original probability distribution? How does the transformation behave versus the specificity concept? How to define coverage regions in the n-dimensional case? What are the pros and cons of the optimal and two sided normalized transformations?

References

1. Tanaka, H., Ishibuchi, H. : Evidence theory of exponential possibility distributions. International Journal of Approximate Reasoning, Vol. 8, No 2, 123-140 (1993)
2. Vejnarova, J. : On interpretation of t-product extensions of possibility distributions. Fuzzy Sets and Systems, pp.3-17 (2013)
3. Baudrit, C., Couso, I., Dubois, D., : Joint propagation of probability and possibility in risk analysis: towards a formal framework. International Journal of Approximate Reasoning, Vol. 45, No 1, 82-105 (2007)
4. Coroianu, L., Fuller, R., : Nguyen type theorem for extension principle based on a joint possibility distribution. International Journal of Approximate Reasoning, Vol. 95, 22-35 (2018)
5. Ferrero, A., Prioli, M., Salicone, S., Vantaggi, B., : 2d probability possibility transformations. Synergies of soft computing and statistics for intelligent data analysis. Springer, pp. 63-72 (2013)
6. Lesniewska, C., Mauris, G., Atto, A., Mercier, G., : On elliptical possibility distributions. Trans. on Fuzzy Systems,Vol. 8, No 8, pp.1631-39 (2020)
7. Akrout, H., Crouzet, J.F., Strauss, O., : Multidimensional possibility/probability domination for extending maxitive kernel based signal processing. Fuzzy Sets and Systems, Vol 367, pp. 1-18 (2019)
8. Hose, D., Mäck, M., Hanss, M., : On probability-possibility consistency in high-dimensional propagation problems. 3rd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, Crete, Greece, pp. 1-10 (2019)
9. Mauris, G., Lasserre, V., Foulloy, L. : A fuzzy approach for the expression of uncertainty in measurement. Int. Journal of Measurement, Vol 29, No 3, pp.165-177 (2001)
10. Dubois, D., Foulloy, L., Mauris, G., Prade, H. : Probability-possibility transformations, triangular fuzzy sets and probabilistic inequalities. Reliable Computing, Vol 10, No 4, pp. 273-297 (2004)
11. Mauris, G., : A review of relationships between possibility and probability representations of uncertainty in measurement. IEEE Trans. on Instrumentation and Measurement,Vol 63, No 3, pp. 622-632 (2013)
12. Mauris, G., : Representating and approximating symmetric and asymmetric probability coverage intervals by possibility distributions. IEEE Trans. on Instrumentation and Measurement,Vol 58, No 1, pp. 876-888 (2009)
13. Mauris, G., : A possibilitic view of binomial estimation. Proceedings of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2014), Montpellier, France, pp. 396-405 (2014)
14. Foulloy, L., Mauris, G., : A new transformation of continuous unimodal asymmetric probability distributions into possibility distributions. Fuzzy Sets and Systems, Vol. 424, pp. 37-62 (2021)
15. Dubois, D., Prade, H., Sandri, S., : On possibility/probability transformations. In fuzzy logic, Springer, pp. 103-112 (1993)

Towards fuzzy logic programming with generalized quantifiers

Jesús Medina and José Antonio Torné-Zambrano

Department of Mathematics. University of Cádiz. Spain
{jesus.medina, joseantonio.torne}@uca.es

Fixed-point semantics is one of the most useful mechanisms for obtaining information from databases modeled by logic programs [6–9, 11–13]. In the monotonic case, it is characterized through the least fixed-point of the immediate consequence operator $T_{\mathbb{P}}$. In many fuzzy frameworks, the generalization of this operator is given through the supremum (of a set of values). Hence, the presence of some errors in the data (noise), can disrupt the final result. This also happens in multi-adjoint logic programming [8, 9], which is a flexible fuzzy framework embedding the monotonic and residuated logic programing, and fuzzy logic programming frameworks [4, 14], among others.

The main goal of this paper is to provide a further generalization of the $T_{\mathbb{P}}$ given in multi-adjoint logic programming [8] focused on weakening the existential feature of the supremum operator and reduce the impact of imprecise data in computations. Taking into account that standard immediate consequence operator implicitly employs the existential quantifier, it can be generalized using other generalized quantifiers, determined by their associated fuzzy measures [2, 5, 10]. This use is justified due to the fact that generalized quantifiers provide a huge range of different intermediate quantifiers between the existential and the universal one and consequently it is possible to consider weaker quantifiers than the existential one. These operators have already been used in diverse frameworks [1, 3] and in this paper we will use them for introducing the notion of quantified immediate consequence operator $T_{\mathbb{P}}^{Q_\mu}$.

This work will study two interesting examples of $T_{\mathbb{P}}^{Q_\mu}$. In addition, this new operator verifies appealing properties such as the monotonicity, which is fundamental for giving a fixed-point semantics in this framework and the continuity, which is a really important feature in order to ensure that the least fixed point (least model) is reached in a countable number of iterations.

References

1. N. Cao, M. Holčapek, and M. Štěpnička. Extensions of fuzzy relational compositions based on generalized quantifiers. *Fuzzy Sets and Systems*, 339:73 – 98, 2018. Theme: Algebra and Set Theory.
2. N. Cao, M. Stepnicka, and M. Burda. Fuzzy quantifiers and compositions of partial fuzzy relations employing dragonfly algebras. In *2019 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2019, New Orleans, LA, USA, June 23-26, 2019*, pages 1–6. IEEE, 2019.

3. M. Cornejo, J. Díaz-Moreno, J. López-Rodríguez, and J. Medina. Quantified concept-forming operators. *CEUR Workshop Proceedings*, 2668:273–280, 2020.
4. C. V. Damásio and L. M. Pereira. Monotonic and residuated logic programs. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'01*, pages 748–759. Lecture Notes in Artificial Intelligence, 2143, 2001.
5. A. Dvořák and M. Holčapek. L-fuzzy quantifiers of type ⟨1⟩ determined by fuzzy measures. *Fuzzy Sets and Systems*, 160(23):3425 – 3452, 2009. Theme: Computer Science.
6. P. Julián-Iranzo and C. Rubio-Manzano. A sound and complete semantics for a similarity-based logic programming language. *Fuzzy Sets and Systems*, 317:1 – 26, 2017. Theme: Logic and Computer Science.
7. J. Lloyd. *Foundations of Logic Programming*. Springer Verlag, 1987.
8. J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Multi-adjoint logic programming with continuous semantics. In *Logic Programming and Non-Monotonic Reasoning, LPNMR'01*, pages 351–364. Lecture Notes in Artificial Intelligence 2173, 2001.
9. J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Similarity-based unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146:43–62, 2004.
10. M. Štěpnička and M. Holčapek. Fuzzy relational compositions based on generalized quantifiers. In A. Laurent, O. Strauss, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 224–233, Cham, 2014. Springer International Publishing.
11. U. Straccia, M. Ojeda-Aciego, and C. V. Damásio. On fixed-points of multivalued functions on complete lattices and their application to generalized logic programs. *SIAM Journal on Computing*, 38(5):1881–1911, 2009.
12. A. Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.*, 5(2):285–309, 1955.
13. van Emden M.H. Quantitative deduction and its fixpoint theory. *The Journal of Logic Programming*, 1:37–53, 1986.
14. P. Vojtáš. Fuzzy logic programming. *Fuzzy sets and systems*, 124(3):361–370, 2001.

A study about the cardinality of some connectives defined on finite chains

Marc Munar[1,2], Sebastia Massanet[1,2], and Daniel Ruiz-Aguilera[1,2]

[1] Soft Computing, Image Processing and Aggregation (SCOPIA) research group, University of the Balearic Islands, 07122 Palma, Balearic Islands, Spain
{marc.munar,s.massanet,daniel.ruiz}@uib.es
[2] Health Research Institute of the Balearic Islands (IdISBa), 07010 Palma, Spain

Since the introduction of connectives defined on finite chains, the research community has considered them a useful and versatile tool for numerous practical cases, mainly for aggregation of qualitative information and decision making. The study of these operators can be reduced to the study of operators defined on the finite chain $L_n = \{0, 1, ..., n\}$ since every finite chain with $n + 1$ elements is isomorphic to L_n. Because of its nature, since both its domain and its image are finite discrete sets, the number of operators that can be constructed is also finite. Therefore, the cardinality of some families has aroused the interest of researchers; mainly, some authors have dealt with the cardinality of discrete t-norms because they are a family of discrete aggregation functions in which the order of aggregation is invariant with respect to the order in which it is done, thanks to the commutative and associative properties. Despite all efforts, approaching the study from different perspectives such as finite semigroup theory [6, 9], no closed formula for its cardinality has been found yet and only few experimentally-computed values are known [5]. This contrasts with other families of discrete operators, whose cardinality has been tackled too. For instance, for the subfamily of discrete t-norms that are smooth, an expression is known (see [8]): 2^{n-1}. Also, for discrete copulas, introduced and characterized in [7]: a bijection is established between these operators and the group of permutations of n elements, determining that there are $n!$ discrete copulas on L_n. We also find in [4] that the cardinality of discrete idempotent uninorms defined on L_n is 2^n.

In this study we have tried to shed some light on this problem. First, we study a well-known family of unary connectives such as discrete negations, defined as decreasing operators $N : L_n \to L_n$ with some boundary conditions. By applying elementary combinatorial arguments, it has been possible to determine a closed expression for the number of discrete negations defined on L_n. Secondly, the cardinality of some binary operators has been studied. In this case, the problem turns out to be much more complicated. To solve it, the concept of plane partitions and some already known expressions about their cardinality have been considered [1, 10] to derive new ones for the cardinality of some discrete connectives. Specifically, closed expressions have been obtained for:

- The number of aggregation functions, conjunctions, disjunctions, Sheffer strokes (see [2] for more information) and implications defined on L_n.

– The number of commutative aggregation functions, commutative conjunctions, commutative disjunctions and commutative Sheffer strokes defined on L_n.

Finally, we have also tackled the study of the cardinality of some families of discrete implications, obtaining a closed formula for the number of discrete implications which satisfy the following additional properties (see [3] for more information about them):

– The Contrapositive Symmetry (**CP**) with respect to the discrete classical negation $N_C(x) = n - x$, for all $x \in L_n$.
– The Neutrality Principle (**NP**).
– The Consequent Boundary (**CB**).
– The Ordering Principle (**OP**).
– The Identity Principle (**IP**).

Acknowledgments

This paper is part of the R&D&I project PID2020-113870GB-I00 funded by MCIN/AEI/10.13039/501100011033/.

References

1. Andrews, G.E.: Plane partitions (II): The equivalence of the Bender-Knuth and MacMahon conjectures. Pacific Journal of Mathematics **72**(2), 283–291 (1977)
2. Baczyński, M., Berruezo, P., Helbin, P., Massanet, S., Niemyska, W., Ruiz-Aguilera, D.: On the Sheffer stroke operation in fuzzy logic. Fuzzy Sets and Systems **432**, 110–128 (2021)
3. Baczynski, M., Jayaram, B., Massanet, S., Torrens, J.: Fuzzy Implications: Past, Present, and Future. In: Kacprzyk, J., Pedrycz, W. (eds.) Springer Handbook of Computational Intelligence, pp. 183–202. Springer Berlin Heidelberg (2015)
4. De Baets, B., Fodor, J., Ruiz-Aguilera, D., Torrens, J.: Idempotent uninorms on finite ordinal scales. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **17**, 1–14 (2009)
5. De Baets, B., Mesiar, R.: Triangular norms on product lattices. Fuzzy Sets and Systems **104**, 61–75 (1999)
6. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Trends in Logic, Springer Netherlands (2000)
7. Mayor, G., Suñer, J., Torrens, J.: Copula-like operations on finite settings. IEEE Transactions on Fuzzy Systems **13**, 468–477 (2005)
8. Mayor, G., Torrens, J.: On a class of operators for expert systems. International Journal of Intelligent Systems **8**, 771–778 (1993)
9. Petrík, M., Vetterlein, T.: Algorithm for Generating Finite Totally Ordered Monoids. In: Information Processing and Management of Uncertainty in Knowledge-Based Systems - 16th International Conference, IPMU 2016. Communications in Computer and Information Science, vol. 611, pp. 532–543. Springer (2016)
10. Stanley, R.P., Fomin, S.: Enumerative Combinatorics, Cambridge Studies in Advanced Mathematics, vol. 2. Cambridge University Press (1999)

Multi-adjoint attribute implications to digital forensic

M.Eugenia Corjeno, Jesús Medina, and Francisco José Ocaña

Department of Mathematics. University of Cádiz. Spain
{mariaeugenia.cornejo, jesus.medina, franciscojose.ocana}@uca.es

Formal concept analysis is a mathematical tool based on lattice theory to obtain significant information from relational databases. Specifically, formal concept analysis allow us to compute concepts, which describe the relation between a set of attributes and a set of objects, and establish a hierarchy among them, obtaining the algebraic structure known as concept lattice. Usually, databases contain a large amount of data and therefore, it is not useful consider the entire context associated with the data and carry out the corresponding procedures to form the concept lattice. Attribute implications is an alternative to obtain the underlying information in the context, providing rules which models the dataset.

Attribute implications have been widely studied in the classical case [2, 9] and in the fuzzy one [1, 4–6, 10]. However, the powerful advantages provided by the multi-adjoint concept lattice framework [8] for the treatment of attributes implications have not been fully exploited. This work complements the study presented in [3, 7], related to attribute implications in multi-adjoint concept lattices, focusing on the study of the validity of attribute implications in the multi-adjoint frame and its application to digital forensic.

References

1. R. Belohlavek, 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, 2016.2014.
2. P. Cordero, M. Enciso, A. Mora, and M. Ojeda-Aciego. Computing leftminimal 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.
3. 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.
4. C. V. Glodeanu. Knowledge discovery in data sets with graded attributes.*International Journal of General Systems*, 45(2):232–249, 2016.
5. 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.
6. T. Kuhr and V. Vychodil. Fuzzy logic programming reduced to reasoning with attribute implications. *Fuzzy Sets and Systems*, 262:1 – 20, 2015.

7. 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.
8. 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.
9. 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.
10. V. Vychodil. Computing sets of graded attribute implications with witnessed non-redundancy. *Information Sciences*, 351:90 – 100, 2016.

Fuzzy gradual rules based approach for assessing car driver's affective states through physiology

Joseph Onderi ORERO

School of Computing & Engineering Sciences (SCES),
Strathmore University, Kenya.
jorero@strathmore.edu

Affective computing has become a major research interest in the Human Computer Interaction (HCI) community. Hence, there is a need to develop methodologies for assessing user's emotional experiences while interacting with these computer applications. In this context, physiology-based emotionally intelligent paradigms provide an opportunity to enhance human computer interactions by continuously evoking and adapting to the user experiences in real-time [10]. Research in this area has demonstrated the enormous prospects in developing systems equipped with the ability to assess user emotional states using various aggregation of physiological signal absolute value such as mean, minimum or maximum, power spectrum density ... and classical machine learning such as K-Nearest Neighbor, Linear Discriminant Analysis , Artificial Neural Networks, Decision Trees... [12,13,6,3,9]. Nevertheless, there is a need to develop more adequate models to represent the mapping of physiological patterns to users' affective states for real-life emotionally intelligent applications.

To begin with, physiological signals tend to vary from participant to participant and even within the same participant physiological signals vary from time to time. The current methods tend to rely on some forms of normalization using some baseline to tackle this variability. However, the correlation between the baseline and the various emotions also vary from person to person and at different occasions for the same person. Thus, these modelling approaches can not lead to a generalized mapping of affective states to physiological signals irrespective of the person and time. In this study, a way of extracting features that are independent of person and time of expression of the emotion, we consider a model based on gradual rules of the form: *the more or less of A, the more or less of B* [8,4,5,11]. Specifically, we consider the physiological signals variation with time during a particular affective state, such as: *Heart Rate increases with time during Joy more than 60% of the time or Heart Rate increases with time during Disgust less than 40% of the time.* After we extract the support for each gradual item set, we define fuzzy rules to characterize the various emotions.

Secondly, we consider a fuzzy set theory based model [14]. Emotions are conceptual quantities with indeterminate fuzzy boundaries [2]. Therefore, it is necessary to express in fuzzy terms the mapping of affective markers from physiological data. In the context of continuously assessing emotions from physiological signals, change from one emotional state to the next is gradual rather than abrupt. Besides, the physiological data from sensors is itself imperfect, such that it is difficult to express the results in crisp terms [1]. Therefore, it is more natural to formulate a fuzzy set theory based model to represent these continuous transitions, uncertainties and imperfections. In fuzzy approach changes from one rule to another is gradual with fuzzy values $[0, 1]$ instead of crisp values $\{0, 1\}$ in classical machine learning approaches. Therefore, once we extract the covariation of the physiological signals and time, we use the support of the gradual item-sets as input for discretization into fuzzy rules [7]. We validate our model in experimental study to define gradual rules that characterize various physiological features in relation to a car driver's affective states data.

References

1. Bouchon-Meunier, B.: Aggregation and Fusion of Imperfect Information. Physica-Verlag, Spring-Verlag Company (1998)
2. Calvo, R.A., D'Mello, S.: Affect detection: an interdisciplinary review of models, methods, and their applications. IEEE Transactions on Affective Computing **1**, 18 – 37 (2010)
3. Chanel, G., Kierkels, J., Soleymani, M., Pun, T.: Short-term emotion assessment in a recall paradigm. International Journal of Human-Computer Studies **67**, 607–627 (2009)
4. Di-Jorio, L., Laurent, A., Teisseire, M.: Mining frequent gradual itemsets from large databases. In: Advances in Intelligent Data Analysis VIII, 8th International Symposium on Intelligent Data Analysis, IDA 2009, Lyon, France, August 31 - September 2, 2009.
5. Fiot, C., Masseglia, F., Laurent, A., Teisseire, M.: Evolution patterns and gradual trends
6. Kim, J., André, E.: Emotion recognition based on physiological changes in music listening. IEEE Transactions on Pattern Analysis And Machine Intelligence **30**(12), 2067–2083 (2008)
7. Kononenko, I., Kukar, M.: Chapter 7 - data preprocessing. In: Machine Learning and Data Mining
8. Laurent, A., Lesot, M.J., Rifqi, M.: Graank: Exploiting rank correlations for extracting gradual itemsets. In: Eighth International Conference on Flexible Query Answering Systems (FQAS 2009), Roskilde, Denmark (2009)
9. Levillain, F., Orero, J.O., Rifqi, M., Bouchon-Meunier, B.: Characterizing player's experience from physiological signals using fuzzy decision trees. In: Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games, CIG 2010, Copenhagen, Denmark, 18-21 August, 2010
10. Novak, D., Mihelj, M., Munih, M.: A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interacting with Computers **24**, 154–172 (2012)
11. Owuor, D.O., Laurent, A., Orero, J.O.: Mining fuzzy temporal gradual emerging patterns
12. Picard, R., Vyzas, E., Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions Pattern Analysis and Machine Intelligence **23**, 1175–1191 (2001)
13. Rani, P., Sarkar, N., Adams, J.: Anxiety-based affective communication for implicit human machine interaction. Advanced Engineering Informatics **21**, 323–334 (2007)
14. Zadeh, L.: Fuzzy sets. Information Control **8**, 338–358 (1965)

Impact of changes in volumes of external training loads collected with wearable GPS sensors on the occurrence of injuries in football

Tomasz Piłka[1,2], Bartłomiej Grzelak[1,2], Aleksandra Sadurska[1], Tomasz Górecki[1], and Krzysztof Dyczkowski[1]

[1] Adam Mickiewicz University, ul. H.Wieniawskiego 1,61-712 Poznań, Poland
{tomasz.pilka,tomasz.gorecki,chris}@amu.edu.pl, alesad3@st.amu.edu.pl
[2] KKS Lech Poznań, ul. Bulgarska 17, 60-320 Poznań
bartlomiej.grzelak@lechpoznan.pl

This study aimed to check whether the use of rule-based methods, machine learning methods, and fuzzy rule-based methods can be used to analyze and recognize the relationship between the volume and intensity of player training in microcycles to identify injury risk factors in elite male footballers. Based on existing studies, we must be aware that high loads increase injury risk, but those substantial modifications in the training load from microcycle to microcycle are also paid attention to. Recent literature has discussed the need to not only quantify the load in absolute terms but also measure the quantity of change [1, 2, 4].

1 Data collection

The data were collected during one round in the professional football league, among 25 players with an average age of 24 ± 5.26 years. The data were collected using Catapult wearable global positioning trackers [3], both during training and match activities. The data for each microcycle is split into two subsets. In the first one, the activities from the entire training week to the match day are aggregated, in the second – the results from the match day. The data set used in this research contains information about 364 events (20 injuries). For each event, we use variables concerning speed, distance covered, and specific football movements.

2 Methods used

Rule-based method – Expert
A rule-based system consists of if-then rules, a bunch of facts, and an interpreter controlling the application of the rules, given the facts. These *if-then rule* statements are used to formulate the conditional statements that comprise the complete knowledge base.

Fuzzy rule-based methods
In this approach, a fuzzy rule-based decision system was developed. Two approaches have been used to build the decision system. In the first one, the rules prepared by experts were used, in the second the rules were inducted on the basis of available data. In the final stage, the two rule-based systems were combined and optimized. The implementation was done using Simful Python library for fuzzy logic.

Machine learning methods
As a machine learning baseline, we used the gradient boosting method ([5]). It is a machine learning technique, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. It builds the model like other boosting methods do, but it generalizes them by allowing optimization of an arbitrary differentiable loss function. In this approach, we used the algorithm XGBoost.

3 Preliminary results

To check the quality of the XGBoost algorithm we applied 5×10 cross-validation method. Additionally, due to unbalanced classes, we used SMOTE method of oversampling. XGBoost method achieved 100% precision, 90.1% recall, and 90.7% accuracy. F1 score is equal to 94.8%. We correctly predicted all 20 injuries, however, we predicted 34 injuries that were no injuries. We obtained comparable results for the rule-based method.

4 Summary and further work

Our work can be expanded in many directions. We want to add variables with the internal load that refers to the biological stressors that are imposed on the athlete during training or competition [4]. The next step is building a system that will analyze the deviations of the training load and propose their quantitative correction for two weeks in the training process.

References

1. Gabbett, T.J.: The training—injury prevention paradox: should athletes be training smarter and harder? British Journal of Sports Medicine 50(5), 273–280 (2016)
2. Rossi, A., et al.: Effective injury forecasting in soccer with gps training data and machine learning. PLOS ONE 13(7), 1–15 (2018)
3. Catapult, provider of wearable technology for the management of athletes, https://www.catapultsports.com/
4. Bourdon, P.C., et al.: Monitoring athlete training loads: Consensus statement. International Journal of Sports Physiology and Performance 12(s2), 161–170 (2017)
5. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of statistics 29(5), 1189–1232 (2001)

Penalty-based pooling functions for feature reduction on Convolutional Neural Networks

Iosu Rodriguez-Martinez, Pablo Aitor Lizarraga-Guerra[1], Tiago Asmus[1,2], Graçaliz Dimuro[1,2], Francisco Herrera[3,4], and Humberto Bustince[1]

[1] Universidad Publica de Navarra, Pamplona, Spain
{iosu.rodriguez, pablo.lizarraga, bustince}@unavarra.es

[2] Universidade Federal do Rio Grande, Rio Grande, Brazil
{tiagoasmus, gracalizdimuro}@furg.br

[3] Universidad de Granada, Granada, Spain
herrera@decsai.ugr.es

[4] King Abdulaziz University, Jeddah, Saudi Arabia

Penalty-based aggregation functions select the best out of several possible aggregations for a set of values, with respect to the deviation from each aggregated output and the original inputs [1, 2]. They have seen significant application to image downsampling tasks, where the pixel values of small disjoint regions of an input image are fused into single representatives, which try to preserve the highest possible amount of information [3, 4].

Convolutional Neural Networks (CNNs), a family of neural network models capable of automatically extracting the most descriptive visual features of an image in order to solve tasks such as classification [5] or semantic segmentation [6], perform a similar downsampling process on their so called "pooling layers". After extracting "feature images" through the convolution of a series of filters over an input image, disjoint windows of values are aggregated, reducing the dimensionality of the output features. Although this data fusion process has been traditionally performed using simple functions such as the arithmetic mean or the maximum, some alternatives have been proposed over the years: stochastic pooling introduces some regularization into the model by randomly choosing one of the input values according to a multinomial distribution [7]; Mixed Pooling combines the output of both maximum and mean pooling through a convex combination where coefficients are learnt in different ways [8]; the previous strategy has been generalized by us with success to additional increasing functions, in the form of CombPool layers [9].

On this work, we study the suitability of penalty-based aggregation functions for replacing the pooling mechanism of a CNN, allowing for the reduction of different regions of the feature image by different, more suitable aggregations. Further, we introduce a new strategy in which "soft" penalty-based aggregation functions can be used in order to produce coefficients for each term of a convex combination of aggregation functions, relative to the penalty of each of those aggregations.

Acknowledgements Grant PID2019-108392GB-I00/ funded by MCIN/AEI/10.13039/501100011033. T. Asmus and G. Dimuro are partially supported by the projects CNPq (301618/2019-4), FAPERGS (19/2551-0001279-9). F. Herrera is supported by the Andalusian Excellence project P18-FR4961. I. Rodriguez-Martinez is partially supported by the projects PC095-096 FUSIPROD, Tracasa Instrumental (iTRACASA) and Gobierno de Navarra - Departamento de Universidad, Innovación y Transformación Digital.

References

1. Calvo, T., Beliakov, G. (2010). Aggregation functions based on penalties. Fuzzy sets and Systems, 161(10), 1420-1436.
2. Bustince, H., Beliakov, G., Dimuro, G. P., Bedregal, B., Mesiar, R. (2017). On the definition of penalty functions in data aggregation. Fuzzy Sets and Systems, 323, 1-18.
3. Beliakov, G., Bustince, H., Paternain, D. (2011). Image reduction using means on discrete product lattices. IEEE Transactions on Image Processing, 21(3), 1070-1083.
4. Paternain, D., Jurio, A., Beliakov, G. (2012, June). Color image reduction by minimizing penalty functions. In 2012 IEEE International Conference on Fuzzy Systems (pp. 1-7). IEEE.
5. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).
6. Ronneberger, O., Fischer, P., Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
7. Zeiler, M. D., Fergus, R. (2013). Stochastic pooling for regularization of deep convolutional neural networks: 1st International Conference on Learning Representations, ICLR 2013. Paper presented at 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, United States.
8. Lee, C. Y., Gallagher, P. W., Tu, Z. (2016, May). Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. In Artificial intelligence and statistics (pp. 464-472). PMLR.
9. Rodriguez-Martinez, I., Lafuente, J., H. N. Santiago, R., Pereira Dimuro, G., Herrera, F., Bustince, H. Replacing pooling functions in Convolutional Neural Networks by linear combinations of increasing functions. Accepted In Neural Networks. Elsevier.

Fuzzy differential equations: on some connections between solutions from different perspectives

Rosana Rodríguez-López[1,2]

[1] CITMAga, 15782 Santiago de Compostela, España
[2] Departamento de Estatística, Análise Matemática e Optimización, Universidade de Santiago de Compostela, 15782, Santiago de Compostela, España
rosana.rodriguez.lopez@usc.es

After the introduction of the concept of fuzzy set [22], fuzzy differential equations have been proposed as a possible framework to express the features of processes whose factors or data are subject to uncertainty (see, e.g., [10]). However, for these models, we might find some inconveniences in the calculation of the explicit solutions. In the literature, several approaches have been considered, and the relation between the corresponding solutions has also been investigated.

Some perspectives include derivatives for fuzzy functions [8, 19], Zadeh's Extension Principle (e.g., [18]), differential inclusions [9]... With respect to the derivatives for fuzzy-valued functions, we mention H-, GH-, or gH-differentiability, among others (see, for instance, [1–4, 21]). The particular features of the solutions from this perspective have been discussed [3, 5, 11, 12, 20], and, recently, some metric-based derivative concepts have been considered [15, 16].

The relation between the solutions from different approaches has been the main objective of the works [6, 7, 13]. This contribution aims to explore some of the connections between the solutions under strongly generalized derivative and differential inclusions. The main ideas are given for interval [17] and fuzzy cases. In [14], we explain how the change of sign in the coefficient and the type of derivative determine the coincidence or discrepancy between these perspectives.

The solutions under both approaches do not correspond through the introduction of switching points only in the derivative for just a unique equation, thus we combine different types of derivatives and problems depending on the sign of the coefficient. As explained in [14], this allows to establish completely the relationship between the expression of the solution via differential inclusions and the solution to a switching-combination of different related problems and types of derivatives altogether. Analogous ideas are given for the fuzzy case [14].

Acknowledgements
Research partially supported by grants PID2020-113275GB-I00 (AEI/FEDER, UE), and ED431C 2019/02 (GRC Xunta de Galicia).

References

1. Bede, B., Gal, S. G.: Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations. Fuzzy Sets Syst **151**, 581–599 (2005).

2. Bede, B., Gal, S. G.: Solution of fuzzy differential equations based on generalized differentiability. Commun. Math. Anal. **9**, 22–41 (2009).
3. Bede, B., Rudas, I. J., Bencsik, A. L.: First order linear fuzzy differential equations under generalized differentiability. Inf. Sci. **177(7)**, 1648–1662 (2007).
4. Bede, B., Stefanini, L.: Generalized differentiability of fuzzy-valued functions. Fuzzy Sets Syst **230**, 119–141 (2013).
5. Chalco-Cano, Y., Khastan, A., Rodríguez-López, R.: Normalized expression for solutions to linear fuzzy differential equations under combination of differences. In: Alonso, J. M., Bustince, H., Reformat, M. (eds.) Proceedings 2015 IFSA-EUSFLAT, pp. 1382–1388. Atlantis Press, Amsterdam (2015).
6. Chalco-Cano, Y., Román-Flores, H.: Comparation between some approaches to solve fuzzy differential equations. Fuzzy Sets Syst **160**, 1517–1527 (2009).
7. Chalco-Cano, Y., Román-Flores, H.: Some remarks on fuzzy differential equations via differential inclusions. Fuzzy Sets Syst **230**, 3–20 (2013).
8. Dubois, D., Prade, H.: Towards fuzzy differential calculus: part 3, differentiation. Fuzzy Sets Syst **8**, 225–233 (1982).
9. Hüllermeier, E.: An approach to modelling and simulation of uncertain dynamical systems. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **5**, 117–137 (1997).
10. Kaleva, O.: A note on fuzzy differential equations. Nonlinear Anal. **64** 895–900 (2006).
11. Khastan, A., Nieto, J. J., Rodríguez-López, R.: Variation of constant formula for first order fuzzy differential equations. Fuzzy Sets Syst **177(1)** 20–33 (2011).
12. Khastan, A., Rodríguez-López, R.: On the solutions to first order linear fuzzy differential equations. Fuzzy Sets Syst **295(C)** 114–135 (2016).
13. Khastan, A., Rodríguez-López, R.: On linear fuzzy differential equations by differential inclusions' approach. Fuzzy Sets Syst **387**, 49–67 (2020).
14. Khastan, A., Rodríguez-López, R.: Relation between strongly generalized solutions and solutions via differential inclusions: a study for linear fuzzy differential equations (2022).
15. Khastan, A., Rodríguez-López, R., Shahidi, M.: New differentiability concepts for set-valued functions and applications to set differential equations. Inf. Sci. **575**, 355–378 (2021).
16. Khastan, A., Rodríguez-López, R., Shahidi, M.: New metric-based derivatives for fuzzy functions and some of their properties. Fuzzy Sets Syst (2021).
17. Markov, S. M.: Existence and uniqueness of solutions of the interval differential equation $X' = F(t, X)$. Comptes Rendus de l 'Academie Bulg. des Sci. **31(12)** 1519–1522 (1978).
18. Mizukoshi, M. T., Barros, L. C., Chalco-Cano, Y., Román-Flores, H., Bassanezi, R. C.: Fuzzy differential equations and the extension principle. Inf. Sci. **177** 3627–3635 (2007).
19. Puri, M. L., Ralescu, D. A.: Differentials of fuzzy functions. J. Math. Anal. Appl. **91**, 552–558 (1983).
20. Rodríguez-López, R.: On the existence of solutions to periodic boundary value problems for fuzzy linear differential equations. Fuzzy Sets Syst **219**, 1–26 (2013).
21. Stefanini, L., Bede, B.: Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Anal. **71**, 1311–1328 (2009).
22. Zadeh, L.: Fuzzy Sets. Inf. Control. **8**, 338–353 (1965).

Lexicographic Method for Fuzzy Multi-objective Linear Programming

Laura Romero[1], Boris Pérez-Cañedo[1], and José Luis Verdegay[2]

[1] University of Cienfuegos, Cienfuegos 55100, Cuba
lromero@ucf.edu.cu, bpcanedo@gmail.com
[2] University of Granada, Granada, Spain
verdegay@decsai.ugr.es

Mathematical optimization is a valuable tool for decision-making. However, decision-making is not a straightforward task. It is in fact pervaded by uncertainties from the real-world that arise in the analysis, modeling and solution stages of any problem. This is particularly true in the area of automated decision systems (ADS), where some form of expert-like human behavior is attributed to these systems to carry out complex tasks involving uncertain information, several data types, multiple conflicting objectives, and so forth.

Fuzzy optimization addresses optimization problems with uncertain and subjectively defined parameters; it has become a fundamental technique for operations research, artificial intelligence and ADS. Here, we describe a lexicographic method for solving fuzzy multi-objective linear programming (FMOLP) problems, in which the parameters and decision variables can take on fuzzy numbers. We address the problem of finding Pareto optimal fuzzy solutions to the FMOLP problem

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

where $\tilde{x}_j$ (for $j \in J$) are the decision variables, and, for simplicity, only triangular fuzzy numbers, denoted as usual by the triplet $\tilde{a} = (a, b, c)$, are considered. To solve FMOLP problem (1), we use the fuzzy ϵ-constraint method introduced in [1] and obtain the scalarized (mono-objective) problem

$$
\begin{aligned}
& \min\ \tilde{z} \\
& \text{s.t. } \tilde{z} \oplus \sum_{k\in K\backslash\{q\}} \lambda_k \times \tilde{s}_k^+ = \tilde{z}_q\left(\tilde{x}\right) \oplus \sum_{k\in K\backslash\{q\}} \lambda_k \times \tilde{s}_k^- \oplus \left(-M,0,M\right), \\
& \quad \tilde{z}_k\left(\tilde{x}\right) \oplus \tilde{s}_k^+ = \tilde{\epsilon}_k \oplus \tilde{s}_k^- \quad \text{for } k \in K\backslash\{q\}, \\
& \quad \tilde{s}_k^- \preceq \tilde{s}_k^+, \tilde{s}_k^- \geq 0, \tilde{s}_k^+ \geq 0 \quad \text{for } k \in K\backslash\{q\}, \\
& \quad \sum_{j\in J} \tilde{a}_{ij} \otimes \tilde{x}_j \preceq \tilde{b}_i \quad \text{for } i \in I, \\
& \quad \tilde{z}, \tilde{s}_k^-, \tilde{s}_k^+, \tilde{x}_j \in \mathcal{T}(\mathbb{R}) \quad \text{for } j \in J \text{ and } k \in K\backslash\{q\}.
\end{aligned}
\tag{2}
$$

A widespread approach to solving fuzzy mono-objective linear programming problems, such as (2), is to transform them into conventional (non-fuzzy) ones via linear ranking functions. However, using linear ranking functions to solve problem (2) does not guarantee that the obtained solutions are Pareto optimal solutions to problem (1) (see [1]). Alternatively, the following transformation of the mono-objective case introduced in [2] uses multiple indices ($f_t(\cdot)$, $t = 1, 2, \ldots$) lexicographically to compare fuzzy numbers and ensures Pareto optimality [1].

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

We discuss this approach in detail and present some examples to illustrate it.

Acknowledgment

Research supported by projects PID2020-112754GB-I0, MCIN/AEI and B-TIC-640-UGR20 FEDER/Junta de Andalucía.

References

1. Pérez-Cañedo, B., Verdegay, J.L., Miranda Pérez, R.: An epsilon-constraint method for fully fuzzy multiobjective linear programming. Int. J. Intell. Syst. **35**(4), 600–624 (2020). https://doi.org/10.1002/int.22219
2. Pérez-Cañedo, B., Concepción-Morales, E.R.: 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. **123**, 256–269 (2019). https://doi.org/10.1016/j.eswa.2019.01.041

Normal cones corresponding to credal sets of probability intervals

Damjan Škulj

University of Ljubljana, Faculty of Social Sciences, Kardeljeva pl. 5, SI-1000 Ljubljana, Slovenia damjan.skulj@fdv.uni-lj.si

Imprecise probabilistic models [1] are used in situations that cannot be adequately described with precise single probability distributions. The probabilistic information is therefore represented in terms of multiple models compatible with given partial information. The sets of probability distributions representing the compatible models are usually referred to as *credal sets*, and according to most models proposed in literature, they are convex, which means they are closed for convex linear combinations.

In finite probability spaces, credal sets are often induced through finite sets of linear constraints, consequently assuming the form of convex polyhedra, i.e., convex sets with a finite number of extreme points. In the imprecise settings, precise expectations of random variables are replaced by intervals of values corresponding to expectations with respect to individual models within credal sets. The minimal and maximal values, also called lower and upper expectations, are of primary interest, and they are known to be assumed in the extreme points of the credal sets. Identification of the extreme points (see also [4, 5, 7]) is therefore one of the key steps in the corresponding numerical analysis.

We propose a method of characterization of the extreme points based on identification of the corresponding *normal cones*. In the case of credal sets, normal cones are sets of random variables whose expectations are minimized in common extreme points of the credal sets. Special cases of normal cones include sets of so-called *comonotonic* random variables that are known to characterize extreme points corresponding to credal sets of 2-monotone *lower probabilities*, also called *convex capacities* (see e.g., [3]).

Our analysis additionally focuses to normal cones of credal sets corresponding to probability intervals [2]. Let $\mathcal{X}$ denote a finite set with elements $x_1, \ldots, x_n$, and let l and u be given real valued functions on $\mathcal{X}$ such that $0 \le l(x_i) \le u(x_i) \le 1$ for every $i \in \{1, \ldots, n\}$. Let

$$\mathcal{M} = \{p | \; l(x_i) \le p(x_i) \le u(x_i) \; \forall i \in \{1, \ldots, n\}, \sum_{i=1}^{n} p(x_i) = 1\}$$

be the set of all probability mass functions lying between l and u. Then $\mathcal{M}$ is the credal set of the *probability interval* (l, u). Moreover, the probability interval is *coherent* if $\min_{p \in \mathcal{M}} p(x_i) = l(x_i)$ and $\max_{p \in \mathcal{M}} p(x_i) = u(x_i)$ for every $i \in \{1, \ldots, n\}$.

D. Škulj

We show that a minimal normal cone with respect to set inclusion corresponding to an extreme point in $\mathcal{M}$ is of the form

$$N(x, A, B) = \{f: \mathcal{X} \to \mathbb{R} \mid f(a) \leq f(x) \leq f(b)\ \forall a \in A,\ \forall b \in B\}$$

where $x \in \mathcal{X}$ and $A, B \subseteq \mathcal{X}$ such that $A \cup B = \mathcal{X} \backslash \{x\}$, and A and B are non-empty and disjoint. Cones of the above form are explicitly related to the corresponding extreme points. Moreover, an adjacency structure in the set of normal cones is explored, where two cones are considered adjacent if they correspond to adjacent extreme points. To demonstrate usefulness of this new description for the polyhedral structure of the credal sets, some known results are derived and improved in a novel way. Thus we estimate of the maximal number of extreme points in the credal sets and prove that probability intervals belong to the class of 2-monotone lower probabilities.

A full paper version of our results can be found in [6].

Acknowledgment

The author acknowledges the financial support from the Slovenian Research Agency (research core funding No. P5-0168).

References

1. Augustin, T., Coolen, F.P., de Cooman, G., Troffaes, M.C.: Introduction to imprecise probabilities. John Wiley & Sons (2014)
2. Campos, L.D., Huete, J., Moral, S.: Probability intervals: a tool for uncertain reasoning. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **2**(2) (1994)
3. Denneberg, D.: Non-additive measure and integral, vol. 27. Springer Science & Business Media (1994)
4. Miranda, E., Couso, I., Gil, P.: Extreme points of credal sets generated by 2-alternating capacities. International Journal of Approximate Reasoning **33**(1), 95–115 (2003)
5. Miranda, E., Destercke, S.: Extreme points of the credal sets generated by comparative probabilities. Journal of Mathematical Psychology **64—65**, 44 – 57 (2015). https://doi.org/http://dx.doi.org/10.1016/j.jmp.2014.11.004
6. Škulj, D.: Normal cones corresponding to credal sets of lower probabilities. arXiv preprint arXiv:2201.10161 (2022)
7. Wallner, A.: Extreme points of coherent probabilities in finite spaces. International Journal of Approximate Reasoning **44**(3), 339 – 357 (2007). https://doi.org/http://dx.doi.org/10.1016/j.ijar.2006.07.017

On a graded approach to concept lattices in fuzzy rough set theory

Alexander Šostak[a,b], Māris Krastiņš[b], Ingrīda Uļjane[a,b]

[1] Institute of Mathematics and CS University of Latvia, Riga, LV-1459, Latvia
[2] Department of Mathematics, University of Latvia, Riga, LV-1004, Latvia
aleksandrs.sostaks@lumii.lv, aleksandrs.sostaks@lu.lv, mk18032@lu.lv, ingrida.uljane@lu.lv

Formal concept analysis was initiated by R. Wille and B. Ganter [7], [3]. The principal subject of the study and at the same time the main tool of research in formal concept analysis are concept lattices. The basics of the fuzzy concept analysis and the corresponding theory of fuzzy concept lattices were developed by R. Bělohlávek [1]. In turn, I. Düntch and G. Gediga [2] introduced the alternative notion of a concept lattice, now known as a property-oriented concept lattice. Afterwards, Y.Y. Yao [8] defined an object-oriented concept lattice, which is a certain dual to the property-oriented concept lattice. In [4], H.-L. Lai and D. Zhang developed a unifying fuzzy version of Düntch-Gediga and Yao approaches to concept lattices under the name *concept lattices in fuzzy rough set theory.* In addition, this paper contains a deep comparative analysis between the "classic" (that is Wille-Ganter-Bělohlávek) fuzzy concept lattices and concept lattices in fuzzy rough set theory. At present, all directions of (fuzzy) concept analysis and corresponding theories of (fuzzy) concept lattices represent a well developed area of mathematics. Numerous applications of concept analysis and concept lattices in practice can also be found in the literature, in particular, they have important applications in medicine, biology, sociology, data mining, etc. However, to the best of our knowledge, most of full-fledged application are limited to crisp versions of concept analysis. Unfortunately, we could find only fragmentary applications of fuzzy concept lattices for the study of practical problems. This remark concerns both the "classic" *fuzzy* concept analysis and the concept analysis in *fuzzy* rough set theory. The reason for this situation, as far as we understand, is that the exact matching of certain *fuzzy* subsets of objects and corresponding *fuzzy* subsets of their properties (that is the basis for getting the corresponding concept), in the case of fuzzy substances is almost impossible. (In the talk we plan to substantiate this statement; see also our comments in this concern in [6].) To overcome this problem we replace the notion of a fuzzy concept by a more general notion of a graded fuzzy concept and on its base develop a more flexible theory of graded fuzzy concept lattices in rough set theory. In the sequel we sketch the principal steps in the process of realization of the graded approach for the case of object-oriented version of fuzzy concept lattices.

As in most versions of (fuzzy) concept analysis, we start with a fuzzy context, that is a quadruple (X, Y, R, L) where X and Y are sets, $L = (\mathrm{L}, \wedge, \vee, *, \mapsto)$ is a quantale [5] (or a residuated lattice) and $R : X \times Y \to L$ is a fuzzy relation. The sets X and Y are interpreted as universes of objects and properties (at-

tributes) respectively and the value $R(x,y) \in L$ is interpreted as the degree in which object x has the property y. We call the elements $(A,B) \in L^X \times L^Y$ *fuzzy preconcepts* in the fuzzy context (X,Y,R,L) and introduce the lattice structure on the family $\mathbb{P} = L^X \times L^Y$ by setting $(A,B) \barwedge (A',B') = (A \wedge A', B \vee B')$ and $(A,B) \veebar (A',B') = (A \vee A', B \wedge B')$. We define derivation operators $R^{\square}, R^{\lozenge} : L^X \to L^Y$ and $R^{\blacksquare}, R^{\blacklozenge} : L^Y \to L^X$ on the basis of lower and upper images and preimages induced by fuzzy relation R:
$R^{\square}(A)(y) = \bigwedge_{x\in X}(R(x,y) \mapsto A(x))$, $R^{\lozenge}(A)(y) = \bigvee_{x\in X}(R(x,y) * A(x))$, $R^{\blacksquare}(B)(x) = \bigwedge_{y\in Y}(R(x,y) \mapsto B(y))$, $R^{\blacklozenge}(B)(x) = \bigvee_{y\in Y}(R(x,y) * B(y))$.

As different from the standard request in fuzzy concept analysis of the exact correspondence between a fuzzy set of objects and the relevant fuzzy set of attributes and vice versa, we suggest to estimate the *degree* of this correspondence. We define this degree as the measure of inclusion $\hookrightarrow$ on powersets L^X and L^Y induced by the residuum of the quantale L: $A \hookrightarrow A' = \bigwedge_{x\in X}(A(x) \mapsto A'(x))$ for $A, A' \in L^X$ and $B \hookrightarrow B' = \bigwedge_{y\in Y}(B(y) \mapsto B'(y))$ for $B, B' \in L^Y$. In the result we get operators $\mathrm{D}^1_{\mathrm{ob}}, \mathrm{D}^2_{\mathrm{ob}}, \mathrm{D}^3_{\mathrm{ob}}, \mathrm{D}^4_{\mathrm{ob}} : L^X \times L^Y \to L$ defined respectively by
$\mathrm{D}^1_{\mathrm{ob}}(A,B) = A^{\square} \hookrightarrow B$, $\mathrm{D}^2_{\mathrm{ob}}(A,B) = B \hookrightarrow A^{\square}$, $\mathrm{D}^3_{\mathrm{ob}}(A,B) = A \hookrightarrow B^{\blacklozenge}$, $\mathrm{D}^4_{\mathrm{ob}}(A,B) = B^{\blacklozenge} \hookrightarrow A$. Now, the graded fuzzy object oriented fuzzy lattice of the fuzzy context (X,Y,R,L) is defined as the tuple

$$\left((\mathbb{P}, X, Y, R, L), \barwedge, \veebar, \mathrm{D}^1_{\mathrm{ob}}, \mathrm{D}^2_{\mathrm{ob}}, \mathrm{D}^3_{\mathrm{ob}}, \mathrm{D}^4_{\mathrm{ob}}\right).$$

In the talk we plan to expose some properties of graded fuzzy object-oriented concept lattices, illustrate them by examples and discuss possible application of graded fuzzy concept lattices in the research of some problems in practice.

References

1. Bělohlávek, R: Concept lattices and order in fuzzy logic. Annals of Pure and Applied Logic 128 (2004), 277–298.
2. Düntch, I., Gediga, G.: Approximation operators in qualitative data analysis, in: Theory and Application of Relational Structures as Knowledge Instruments, de Swart, H., Orlowska, E., Schmidt, G. and Roubens, M. (Eds.), Springer, Heidelberg, 216-233, 2003.
3. Ganter B., Wille R.: Formal Concept Analysis: Mathematical Foundations. Springer Verlag, Berlin, 1999.
4. Lai, H-L., Zhang, D.: Concept lattices of fuzzy context: Formal concept analysis vs. rough set theory. Intern. J. of Approximate Reasoning 50, (2009), 695–707.
5. Rosenthal, K.I.: Quantales and Their Applications. Pitman Research Notes in Mathematics, vol. 234. Longman Scientific & Technical, 1990.
6. Šostak, A., Uļjane, I., Krastiņš, M.: Gradations of fuzzy preconcept lattices, AXIOMS (2021) 10(1), 41.
7. Wille R.: Concept lattices and conceptual knowledge systems. Computers and Math. with Applications **23** (1992), 493–515.
8. Yao, Y.Y.: Concept lattices in rough set theory. In. Proc. 23rd International Meeting of the North American Fuzzy Information Processing Society Annual Meeting. NAFIPS '04, (2004) pp. 796–801.

Time-dependent author-level metrics

Andrea Stupňanová

Slovak University of Technology in Bratislava, Faculty of Civil Engineering
Radlinského 11, 810 05 Bratislava, Slovak Republic
andrea.stupnanova@stuba.sk

Most author-level metrics (citation indices) do not distinguish the time of publication of the author's paper, and they are based on a simple characterization of each author by a decreasing n-tuple $\mathbf{a} = (a_1, \dots, a_n)$, $a_1 \geq \cdots \geq a_n$, where n is the number of published papers of author A (in the considered database), and a_i denotes the number of citations received by the i-th most cited paper. This applies, among others, to the next indices: h-index of Hirch [4], g-index of Egghe [2] and MAXPROD of Kosmulski [5].

All three recalled citation indices can be seen as particular integrals. In general, each n-tuple $\mathbf{a}$ can be embedded into an infinite sequence $(b_1, b_2, \dots, b_n, b_{n+1}, \dots)$ such that $b_i = a_i$ for $i = 1, \dots, n$, and $b_i = 0$ if $i > n$. Denote the set of all such infinite vectors as $\mathcal{B}$, i.e., $\mathbf{b} \in \mathcal{B}$ if and only if $\mathbf{b}$ is decreasing sequence of integers converging to 0. Evidently, each $\mathbf{b} \in \mathcal{B}$ can be seen as a function $\mathbf{b} : \mathbb{N} \to \mathbb{N}_0$, where $\mathbb{N}$ is the set of all positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Theorem 1. *Let μ be a counting measure on $\mathbb{N}$, i.e., $\mu(E) = \text{card}(E), E \subseteq \mathbb{N}$ and let the author A be characterized by $\frac{a}{n}$ vector $\mathbf{a}$ generating $\mathbf{b} \in \mathcal{B}$. Then:*

(i) the Hirsch index $H(\mathbf{a}) = \mathbf{Su}_\mu(\mathbf{b}) = \bigvee_{i=1} (i \wedge a_i)$;

(ii) the Egghe index $G(\mathbf{a}) = \mathbf{Su}_\mu(\lfloor\sqrt{\mathbf{b}}\rfloor) = \bigvee_{i=1}^{n} (i \wedge \lfloor\sqrt{a_i}\rfloor)$, *where* $\lfloor\cdot\rfloor$ *is the floor of a real number and* $\lfloor\sqrt{\mathbf{b}}\rfloor = (\lfloor\sqrt{b_1}\rfloor, \lfloor\sqrt{b_2}\rfloor, \dots)$;

(iii) the Kosmulski index $\text{MAXPROD}(\mathbf{a}) = \mathbf{Sh}_\mu(\mathbf{b}) = \bigvee_{i=1}^{n} (i \cdot a_i)$,

where **Su** *and* **Sh** *is Sugeno and Shilkret integal, respectively.*

For more details we recommend [1, 8].

In this contribution, we focus on some failure of this indices, namely a non-balanced approach to scholars working for years in their field and to young scholars. In short, they do not consider the time factor, and its performance cannot be diminished with the running time. In such a way, a scholar with no paper neither citations in last decades will preserve his/her (possibly high) h (g, k)-index. Note that there were several attempts to introduce time-dependent (generalization of) h-index. For example, Mannella and Rossi [7] have proposed to consider A_A constant (Academic Age) for each considered scholar and then they have proposed and discussed the index $h/\sqrt{A_A}$. Egghe [3], based on the Lotka's law [6] describing the frequency distribution of scientific productivity, has proposed and discussed the dynamic h-index.

We will follow another approach, surprisingly not yet exploited (up to our best knowledge) in describing time-dependent author's performance. We will apply our approach not only to the standard h-index, but also to some other citation indices. In some database (WoS, SCOPUS, Google Scholar), each paper of a considered author A is characterized not only by the number of its citations (enabling to characterize the global performance of the author A by a vector $\mathbf{a}$, or, equivalently, by a vector $\mathbf{b} \in \mathcal{B}$ derived from $\mathbf{a}$), but also by the proper citation information of that paper, including, among others, the date of publication. Then the author A is characterized by a couple $(\mathbf{a}, \mathbf{y})$ of n-ary vectors, $\mathbf{a} \in \mathbb{N}_0^n$ (and $a_1 \geq \cdots \geq a_n$), and $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{N}^n$, where y_i is the year age of the i-th paper possessing a_i citations. Then the related $\mathbf{c}$ from $[0, \infty]^{\mathbb{N}}$, $\mathbf{c} = (c_1, c_2, \ldots)$ is given as the i-th largest value of the sample $\left(\frac{a_1}{y_1}, \ldots, \frac{a_n}{y_n}\right)$ denoted by c_i, $i = 1, \ldots, n$, and $c_i = 0$ if $i > n$. Then dynamic time-dependent author-level metrics (citation indices) h_T, g_T and k_T are given by $h_T = \mathbf{Su}_\mu(\mathbf{c}), g_T = \mathbf{Su}_\mu(\sqrt{\mathbf{c}})$ and $k_T = \mathbf{Sh}_\mu(\mathbf{c})$. We introduce and exemplify also new time-dependent indices h^T, g^T, k^T and h_T^T, g_T^T, k_T^T.

Acknowledgments. The support of the grants APVV-17-0066 and VEGA 1/0006/19 is kindly announced.

References

1. Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E., Herrera, F.: h-Index: A review focused in its variants, computation and standardization for different scientific fields. Journal of Informetrics **3**(4), 273–289 (2009).
2. Egghe, L.: Theory and practise of the g-index. Scientometrics **69**, 131—152 (2006).
3. Egghe, L.: Dynamic h-index: The hirsch index in function of time. Journal of the American Society for Information Science and Technology **58**(3), 452–454 (2007).
4. Hirsch, J.E.: An index to quantify an individual's scientific research output. Proc. Nat. Acad. Sci. **102**(4), 16569–16572 (2005).
5. Kosmulski, M.: MAXPROD - A new index for assessment of the scientific output of an individual, and a comparison with the h index. Cybermetrics **11**(1), paper 5 (2007).
6. Lotka, A.J.: The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences **16**(12), 317—324 (1926).
7. Mannella R., Rossi P.: On the time dependence of the h-index. Journal of Informetrics **7**(1), 176–182 (2013).
8. Torra, V., Narukawa, Y.: The h-index and the number of citations: Two fuzzy integrals. IEEE Transactions on Fuzzy Systems **16**(3), 795–797 (2008)

A comparative study of fuzzy sets for the assessment of NEA deflection techniques

M. M. Sánchez-Lozano and Fernández-Martínez

University Centre of Defence at the Spanish Air Force Academy,
MDE-UPCT, 30720 San Javier, Spain
{juanmi.sanchez,manuel.fernandez-martinez}@cud.upct.es

Being encouraged by the introduction of the spherical fuzzy sets (SFS in the sequel) [1], we wonder to what extent a ranking of alternatives would vary if either triangular fuzzy sets (TFSs) are considered to deal with a MCDM problem or the most recent extensions of fuzzy sets (such as SFSs) are used with that aim. In this regard, this short communication addresses that question by comparing both the SFS and the TFS versions of the TOPSIS approach to rank a set of alternatives regarding a matter of planetary defense.

Specifically, we illustrate that comparison between rankings by means of a set of alternatives consisting of five deflection technologies for Near Earth Objects (NEAs hereafter), that are asteroids whose orbital trajectories may potentially cross the Earth's orbit. Since they are able to closely approach the Earth's orbit, small perturbations regarding their orbits may place them on a collision path to the Earth's. As such, to perform an exhaustive tracking of these asteroids is highly recommended by the scientific community.

In this paper we assess a set consisting of four NEA deflection technologies, namely, Kinetic Impactor (A_1), Ion-Beam Deflection (A_2), Enhanced Gravity Tractor (A_3), and Laser Ablation (A_4) with respect to the criteria build time (C_1), duration of the active deflection (C2), asteroid rotation (C_3), asteroid composition (C_4), asteroid structure (C_5), asteroid shape (C_6), Technological Readiness Level (C_7), and mission risk (C_8).

In [2] and [3], two assessments regarding such deflection technologies were carried out with respect to the aforementioned set of criteria by means of both the TFN and SFS versions of the AHP approach, respectively. Next, we highlight the main results that were provided by such a comparative analysis involving these fuzzy approaches. First, in [2], the following order of preference for our set of criteria was obtained through the TFS version of the AHP: $C_1 > C_2 > C_7 > C_4 > C_5 > C_8 > C_6 > C_3$. On the other hand, in [3], it was provided the next order of preference for that set of criteria after applying the SFS version of the AHP: $C_1 > C_2 > C_8 > C_4 > C_7 > C_5 > C_3 > C_6$.

Thus, the criteria C_7 (Level of maturity) and C_4 (Asteroid composition) did interchange their relative level of importance from one fuzzy version of the AHP to the other. Specifically, C_4 appeared as being more important than C_7 when applying the SFS AHP. It is also worth mentioning that the attribute C_8 (mission risk) was assigned a greater level of importance by the SFS AHP than by the TFS AHP.

Next we compare the rankings of the alternatives that were provided by both the SFS and the TFS versions of the TOPSIS procedure. The TFS TOPSIS provided the following order of preference for the four deflection technologies: $A_1 > A_2 > A_3 > A_4$, whereas the SFS TOPSIS threw the next ranking for such alternatives: $A_1 > A_2 > A_4 > A_3$. It is worth pointing out that the alternative A_1 (Kinetic Impactor) appears consolidated as the best choice for active NEA deflection purposes according to both the SFS and the TFS TOPSIS procedures. However, we observed that the positions of the alternatives A_3 (Enhanced Gravity Tractor) and A_4 (Laser Ablation) were interchanged when applying the SFS TOPSIS to assess the alternatives instead of the TFS version of that approach. This could be due to the fact that the SFS TOPSIS assigned a greater importance to the criterion C_4 (Asteroid composition) to the detriment of C_7 (Technological Readiness Level).

Acknowledgements

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

References

1. Gündoğdu, F.K., Kahraman, C.: Spherical fuzzy sets and spherical fuzzy TOPSIS method. J Intell Fuzzy Syst. 36, 1, 337–352 (2019). https://doi.org/10.3233/jifs-181401.
2. Sánchez-Lozano, J.M. et al.: Evaluation of NEA deflection techniques. A fuzzy Multi-Criteria Decision Making analysis for planetary defense. Acta Astronaut. 176, 383–397 (2020). https://doi.org/10.1016/j.actaastro.2020.06.043.
3. Fernández-Martínez, M., Sánchez-Lozano, J.M.: Assessment of Near-Earth Asteroid Deflection Techniques via Spherical Fuzzy Sets. Adv Astron. 2021, 1–12 (2021). https://doi.org/10.1155/2021/6678056.

Application of the Sugeno Integral in Fuzzy Rule-Based Classification

Jonata Wieczynski[1], Giancarlo Lucca[2], Eduardo Borges[2], Mikel Ferrero-Jaurrieta[1], Tiago Asmus[1,2], Graçaliz Dimuro[1,2], and Humberto Bustince[1]

[1] Departamento de Estadística, Informática y Matemáticas, Universidad Pública de Navarra, Pamplona, Spain
{jonata.wieczynski,mikel.ferrero,bustince}@unavarra.es
[2] Centro de Ciêcias Computacionais, Universidade Federal de Rio Grande, Rio Grande, Brazil
{giancarlo.lucca,eduardoborges,tiagoasmus,gracalizdimuro}@furg.br

Fuzzy Rule-Based Classification Systems (FRBCS's) [1] is a technique applied to deal with classification problems, which has been applied to diverse problems, e.g., big data, image segmentation, health and others. The Fuzzy Reasoning Method (FRM) used by FRBCSs is composed by four steps. One of them is the aggregation one, where the information of the system's fired rules is aggregated per class. For this step, the FRM normally uses an aggregation function, and, by doing so, the system will have a different final performance whenever we change the used function. The work proposed by Barrenechea et al. [1] introduced a new FRM that accounts the usage of all given information by the fired fuzzy rules when classifying a new instance. To do so, they have considered the Choquet integral [2]. Moreover, they introduced a fuzzy measure that is adapted for each class of the problem. Lucca et al. [3] introduced the concept of pre-aggregation functions, using as example a generalization of the Choquet integral replacing the product operation by different t-norms. This generalization, called C_T-integral, was applied in the FRM to cope with classification problems and elevated the system quality. After that, also considering the Choquet integral as basis, different generalizations where provided and applied. We highlight the CC-integral [4], C_F-integral [5], C_{F1F2}-integrals [5] and gC_{F1F2}-integrals [5]. Additionally, those generalizations were also applied in multi-criteria decision making problems [6] and image processing [5, 7]. On the other hand, the Sugeno integral [8] is another fuzzy integral which has been applied to diverse problems in the literature. More recently it was applied to a Motor-Imagery Brain-Computer Interface [9], where it obtained good results when compared to the standard Choquet integral. Having in consideration that the Choquet integral was used as base to different generalizations and the good results that the Sugeno integral achieved in recent applications, in this work we analyze if the usage of the Sugeno integral as aggregation base in the FRM is able to produce a system with competitive results. To do so, we apply and analyze this new base function in the FRM of a state-of-art classifier and provided an analysis over 33 distinct datasets from the literature.

In the experimental results we have compared our approach against classical FRMs using the maximum and the Choquet integral and the ones composed by the generalizations of the Choquet integral. The results demonstrated that the Sugeno integral is able to provide superior results in many different datasets and also that this method is statistically equivalent to the compared ones.

Acknowledgements

Supported by Navarra de Servicios y Tecnologías, S.A. (NASERTIC), CNPq (301618/2019-4, 305805/2021-5), FAPERGS (19/2551-0001660-3), the Spanish Ministry of Science and Technology (TIN2016-77356-P, PID2019-108392GB I00 (MCIN/AEI/10.13039/501100011033)).

References

1. Barrenechea, E., Bustince, H., Fernandez, J., Paternain, D., Sanz, J.A.: Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems. Axioms **2**(2), 208–223 (2013)
2. Choquet, G.: Theory of Capacities. In: Annales de l'Institut Fourier. (1953–1954)
3. Lucca, G., Sanz, J., Pereira Dimuro, G., Bedregal, B., Mesiar, R., Kolesárová, A., Bustince Sola, H.: Pre-aggregation functions: construction and an application. IEEE Transactions on Fuzzy Systems **24**(2), 260–272 (April 2016)
4. Lucca, G., Sanz, J., Pereira Dimuro, G., Bedregal, B., Asiain, M., Elkano, M., Bustince Sola, H.: CC-integrals: Choquet-like Copula-based aggregation functions and its application in fuzzy rule-based classification systems. Knowledge-Based Systems, 32 – 43 (2017)
5. Dimuro, G.P., Fernández, J., Bedregal, B., Mesiar, R., Sanz, J.A., Lucca, G., Bustince, H.: The state-of-art of the generalizations of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions. Information Fusion **57**, 27 – 43 (2020)
6. Wieczynski., J., Lucca., G., Borges., E., Dimuro., G., Lourenzutti., R., Bustince, H.: CC-separation measure applied in business group decision making. In: Proceedings of the 23rd International Conference on Enterprise Information Systems - Volume 1: ICEIS. pp. 452–462. SciTePress (2021)
7. Bardozzo, F., De La Osa, B., Horanská, L., Fumanal-Idocin, J., delli Priscoli, M., Troiano, L., Tagliaferri, R., Fernandez, J., Bustince, H.: Sugeno integral generalization applied to improve adaptive image binarization. Information Fusion **68**, 37–45 (2021)
8. Sugeno., M.: Theory of Fuzzy Integrals and its Applications. (1974)
9. Ko, L., Lu, Y., Bustince, H., Chang, Y., Chang, Y., Fernadez, J., Wang, Y., Sanz, J.A., Pereira Dimuro, G., Lin, C.: Multimodal fuzzy fusion for enhancing the motor-imagery-based brain computer interface. IEEE Computational Intelligence Magazine **14**(1), 96–106 (2019)

Author Index