8th European Symposium
on Computational Intelligence
and Mathematics

# ESCIM 2016

Sofia, Bulgaria
October 5th-8th, 2016

# PROCEEDINGS

Editors:
László Kóczy, Jesús Medina

Associate Editors:
María Eugenia Cornejo-Piñero, Juan Carlos Díaz-Moreno,
Janusz Kacprzyk, Vassia Atanassova
Eloísa Ramírez-Poussa, María José Benítez-Caballero

Proceedings of ESCIM 2016



Edition: 1st
First published: 2016

ISBN: 978-84-617-5119-8
Published and printed by:
Universidad de Cádiz (Dept. Matemáticas), Spain

# Organization

## General Chairs

| | |
|---|---|
| László T. Kóczy | Széchenyi István University, Hungary |
| Janusz Kacprzyk | Polish Academy of Sciences, Poland |
| Jesús Medina | Universidad de Cádiz, Spain |

## International Program Committee

| | |
|---|---|
| Krassimir Atanassov | Bulgarian Academy of Sciences, Bulgaria |
| Pedro Cabalar | University of A Coruña, Spain |
| Joao Paulo B. Carvalho | University of Lisbon, Portugal |
| Agata Ciabattoni | TU Wien, Austria |
| Davide Ciucci | University of Milano-Bicocca, Italy |
| Bernard De Baets | University of Ghent, Belgium |
| Juan Carlos Díaz-Moreno | University of Cádiz, Spain |
| Chris Cornelis | University of Granada, Spain |
| Christian G. Fermueller | TU Wien, Austria |
| Péter Foldesi | Széchenyi István University, Hungary |
| Lluis Godo | Artificial Intelligence Research Institute, Spain |
| Stanislav Krajči | UPJS Kosice, Slovakia |
| Ondrej Krídlo | UPJS Kosice, Slovakia |
| Piotr Kulczycki | Cracow University of Technology, Poland |
| Manuel Ojeda-Aciego | University of Málaga, Spain |
| David Pearce | Universidad Politécnica de Madrid, Spain |
| Jozef Pócs | Slovak Academy of Sciences. Slovakia |
| Claudiu Pozna | Transilvania University of Braşov, Romania |
| Alex Tormási | Széchenyi István University, Hungary |
| Esko Turunen | Tampere University of Technology, Finland |

## Organizing Committee

| | |
|---|---|
| Vassia Atanassova | Bulgarian Academy of Sciences, Bulgaria |
| María Eugenia Cornejo-Piñero | Universidad de Cádiz, Spain |
| Jesús Medina | Universidad de Cádiz, Spain |
| Eloísa Ramírez-Poussa | Universidad de Cádiz, Spain |

## Sponsoring Institutions

Universidad de Cádiz, Spain
Széchenyi István University, Hungary
Hungarian Fuzzy Association

# Table of Contents

# Preface

This eighth edition of the European Symposium on Computational Intelligence and Mathematics (ESCIM 2016) consolidates this symposium among one of the most important events in Europe related to computational intelligence and mathematics.

Its participants provide their best advances in the area and share new challenges in order to create an atmosphere capable of creating synergies among the authors, and the keynote speakers focus on achieving these goals and new ones, such as the preparation of a European project.

The general scope and particular topics are perfectly in line with the very essence, philosophy and policy of *Horizon 2020*, the new EU Framework Programme for Research and Innovation. As it can be seen from the relevant contributions presented at ESCIM 2016, and then included in the procedings, a rich and powerful set of various tools and techniques of computational intelligence and mathematics has shown its full potential in solving a wide variety of problems of science, technology, business, etc. Therefore, one can say that computational intelligence must be a fundamental tool in order to solve the new problems and challenges which the society faces. This importance concerns virtually all disciplines in engineering, computer science, data sciences, physics, chemistry, material sciences, etc. to just name a few. Possible *Horizon 2020* projects which would concern all these disciplines would have a huge impact and should without doubt be funded.

The European Symposium on Computational Intelligence and Mathematics arises as a merger between the Győr Symposia on Computational Intelligence (successfully organized from 2008 to 2014 in Győr, Hungary), and the International Workshop on Mathematics and Soft Computing, which combines the area of computational intelligence, which has become one of the main research topics at the Széchenyi István University in the last years, and the different developments in mathematics applied to computer science. The location has been changed but preserves the philosophy of the past Győr Symposia enriched from a more mathematical perspective. That is, bringing together scientists and engineers working in the field of computational intelligence and mathematics to solve current challenges in these fundamental areas.

Sofia will hold this edition, ESCIM 2016, from October 5$^{th}$ to 8$^{th}$, 2016, and it is organized by members of the Bulgarian Academy of Sciences and University of Cádiz, Spain. Thank you very much in particular to Ms. Vassia Atanassova for her amiability and wonderful local organization.

Finally, after a regorous review, 20 out of 43 submissions by authors from 10 different countries were accepted by the members of the Program Committee. The best full papers have been published in a special issue of the International Journal of Intelligent

Systems. The extensions of the best papers have been reviewed to be included in an issue of International Journal of Computational Intelligence Systems.

This Symposium proceedings volume contains the 20 contributions presented during ESCIM 2016 and the 5 presentations of the keynote speakers, which have covered the different topics of the symposium:

- Aggregation functions
- Aggregations for extensions of fuzzy sets
- Data mining and knowledge discovery
- Formal concept analysis
- Fuzzy control
- Fuzzy decision analysis, decision making, optimization and design
- Fuzzy databases and information retrieval
- Fuzzy measures and integrals
- Fuzzy sets and fuzzy logic
- General operators in Computer Science
- Interval-valued fuzzy sets
- Knowledge extraction, representation, and modeling
- Learning, adaptive, and evolvable fuzzy systems
- Logic programming
- Mathematical foundations of fuzzy sets and fuzzy systems
- Rough set theory

The plenary speakers play an important role in this symposium. We would like to thank them for their outstanding contributions to research and leadership in their respective fields. There five plenary lectures cover the different areas of the symposium in charge of renowned researchers such as László Kóczy, Manuel Ojeda-Aciego, Krassimir Atanassov, Jesús Medina and António E. Ruano.

This symposium has its recognition due to the great quality of the contributions. Thank you very much to all the participants for their contributions to the symposium program and all the authors for their submitted papers. We are also indebted to our colleagues members of the Program Committee, since the successful organization of this symposium would not have been possible without their work. Finally, we acknowledge the support received from the Department of Mathematics of the University of Cádiz, the Széchenyi István University, Győr, and the Hungarian Fuzzy Association.

October 2016

Conference Chairs
ESCIM 2016

# Fuzzy Signature Sets Are L-fuzzy

László T. Kóczy

Széchenyi István University, Győr, Hungary
and
Budapest University of Technology and Economics
koczy@sze.hu

## Abstract

When Zadeh introduced the concept of fuzzy sets where $\mu(x) : X \to [0, 1]$, soon his that time student Goguen extended the idea to L-fuzzy sets where the unit interval [0, 1] was replaced by arbitrary algebraic lattice in the manner that the membership degrees are defined by $\mu(x) : X \to L$. In the late 1970s, we introduced a practical extension to fuzzy sets: vector valued Fuzzy (VVF) sets $\mu(x) : X \to [0, 1]^n$. The $n$-dimensional unit hypercube may be certainly interpreted as a lattice under the usual partial ordering $\leq$. This concept was necessary for a certain industrial application, classifying microscopic images of steel alloys.

Much later we proposed a further extension of the idea by allowing the vectorial membership degree components being vectorial themselves, the new concept called Fuzzy Signature (FSig). This way the degree of $x$ belonging to a FSig set is expressed by a nested membership degree vector (with arbitrary depth), or illustrated by a rooted tree graph where each leaf has a membership degree. As in the following various applications (medical diagnosis, built construction evaluation, fuzzy communication of robots, warehouse optimization, etc.) it was necessary to manipulate partly different FSig-s at the same time, the internal nodes of the graphs were attached fuzzy aggregations so partial reduction and transformation of the FSig becomes possible, in order to combine FSig-s of partially different, but essentially similar structure. While many applications were completed and they worked all right, the algebraic structure of FSig-s has never been analysed as far.

The present keynote talk is an attempt to define a series of operations, such as lattice meet and join, and two variations of partial ordering among FSig-s belonging to a certain "family". Based on these it is possible to define an algebraic lattice over the set of nested vectors (within a family), and so, it will be proved that Fuzzy Signature Sets are a special case of Goguen's L-fuzzy sets, thus the "new" concept is in fact a possible realisation of an "old" definition and thus it fits in the existing mathematical system of the fuzzy theory.

# Model Reduction Method Supported by Fuzzy Cognitive Map to Promote Circular Economy

Adrienn Buruzs[1], Miklós F. Hatwágner[2], László T. Kóczy[3]

[1] Department of Environmental Engineering, Széchenyi István University, Győr, Hungary
`buruzs@sze.hu`
[2] Department of Information Technology, Széchenyi István University, Győr, Hungary
`miklos.hatwagner@sze.hu`
[3] Department of Information Technology, Széchenyi István University, Győr, Hungary
`koczy@sze.hu`

**Abstract.** The aim of the present paper is to develop an integrated method that may provide assistance to decision makers during system planning, design, operation and evaluation.

In order to support the realization of circular economy it is essential to evaluate local needs and conditions which help to select the most appropriate system components and resource needs. Each of these activities requires careful planning, however, the model of circular economy offers a comprehensive interdisciplinary framework. The aim of this research was to develop and to introduce a practical methodology for evaluation of local and regional opportunities to promote circular economy.

**Keywords:** factors, fuzzy cognitive maps, model reduction, circular economy, sustainability.

## 1 Introduction

The circular economy provides a coherent framework for systems level redesign and as such offers an opportunity to harness innovation and creativity to enable a positive and restorative economy. Our linear take-make-dispose approach is leading to scarcity, volatility, and pricing levels that are unaffordable for our economy's manufacturing base. The circular economy is an economy in which today's goods are tomorrow's resources, forming a virtuous cycle that fosters prosperity in a world of finite resources.

In order to support future models for circular economy, the method of Fuzzy Cognitive Maps (FCM) [1–3] was selected. This method is capable to simulate the operation of a model as long as the input data sets are available that include the factors with significant effects on the system, and also the historical time series of these factors, which together allow the representation of the features of factors describing the operation of the systems .

In the first phase of the research, the significant 33 factors (Table 1) of the waste management systems were determined. The time series of these elements were developed based on the results of a text mining method. Next, model calculations were made on the basis of fuzzy graph structure [4–7].

Table 1**.** The identified factors of the system and their concept IDs (CID)

| Main factor | Sub-factor | CID | Main factor | Sub-factor | CID |
|---|---|---|---|---|---|
| Technology (C1) | Engineering knowledge | C1.1 | Society (C4) | Public opinion | C4.1 |
| | Technological system and its coherence | C1.2 | | Public health | C4.2 |
| | Local geographical and infrastructural conditions | C1.3 | | Political and power factors | C4.3 |
| | Technical requirements in the EU and national policy | C1.4 | | Education | C4.4 |
| | Technical level of equipment | C1.5 | | Culture | C4.5 |
| Environment (C2) | Impact on environmental elements | C2.1 | | Social environment | C4.6 |
| | Waste recovery | C2.2 | | Employment | C4.7 |
| | Geographical factor | C2.3 | Law (C5) | Monitoring and sanctioning | C5.1 |
| | Resource use | C2.4 | | Internal and external legal coherence (domestic law) | C5.2 |
| | Wildlife (social acceptance) | C2.5 | | General waste management regulation in the EU | C5.3 |
| | Environmental feedback | C2.6 | | Policy strategy and method of implementation | C5.4 |
| Economy (C3) | Composition and income level of the population | C3.1 | Institution (C6) | Publicity, transparency (data management) | C6.1 |
| | Changes in public service fees | C3.2 | | Elimination of duplicate authority | C6.2 |
| | Depreciation and resource development | C3.3 | | Fast and flexible administration | C6.3 |
| | Economic interest of operators | C3.4 | | Cooperation among institutions | C6.4 |
| | Financing | C3.5 | | Improvement of professional standards | C6.5 |
| | Structure of industry | C3.6 | | | |

During the simulation of this complex model, the author proposed a new method for model reduction [8], where three different distance definitions were introduced. The essence of this method is to create clusters of factors and using these clusters to develop new reduced models. Thus, reducing the number of factors, the model is more easily understandable and realistic. The authors' main aim with the reduced model is to support strategic decision-making process to ensure long-term sustainability of material and waste management systems.

## 2 Application of model reduction to FCM

On the basis of the gathered data we constructed the connection matrix. FCM consists of nodes and weighted arcs, which are graphically illustrated as a signed weighted graph with feedback. Nodes of the graph stand for the concepts describing behavioural characteristics of the system. Signed weighted arcs represent the causal relationships that exist among concepts and interconnect them. Concepts represent conceptual characteristics of the system and weight represents the cause and effect influence of one concept on another. In general, concepts represent key-factors and characteristics of the modelled system. The relationships between concepts are described using a degree of causality. Experts describe this degree of influence using linguistic variables for every weight; so weight for any interconnection can range from [–1, 1]. The degree of causal relationship between different factors of the FCM can have either positive or negative sign and values of weights express the degree of the causal relationship. Linkages between concepts express the influence one concept on another.

Taking into account the special characteristics of the studied engineering problem, the number of clusters in the reduced FCM model have to be around six (based on the expert consensus [9]), or at least much less than the number of the factors in the detailed model.

The authors have made several attempts to find the right reduced model in order to determine the right clusters. On the one hand, if only some of the factors are merged, to achieve the goal of unification is not possible. On the other hand, if there are too many factors to be merged then the interpretation of the combined factors is difficult or almost impossible. The intention was to reduce the original matrix by a matrix with around 15 clusters.

The above mentioned three methods are different only in the metrics representing the similarity between the factors. This approach may be considered as a strong generalization of the state reduction procedure of sequential circuits or finite state machines [10–12]. The essence of the methods is to create clusters from the factors and apply these clusters as the new factors in the new, simplified model. The base for binding these clusters is the use of a tolerance relation (a reflexive and symmetric, but non-transitive relation) [13]

Research was divided into two main phases. In the first phase of the research, the significant 33 factors (Table 1) of the waste management systems were determined [9]. The time series of these elements were developed based on the results of a text mining method [10]. Next, model calculations were made on the basis of fuzzy graph structure [11–13].

## 3 Results

The authors have chosen matrices, each including 15 clusters. The tables below (Table 2–4) introduce the clusters in the reduced connection matrix, using different metrics.

Table 2**.** The clusters in the matrix produced by Metric A

| | |
|---|---|
| 1 | C3.1+C3.2+C3.3+C3.4+C3.5+C2.5+C1.3+C6.1+C6.2+C6.3+C6.4+C6.5+C4.3+C4.4+C4.5+C4.6 |
| 2 | C3.3+C3.4+C3.6+C2.3+C2.5+C1.3+C6.2+C6.3+C6.4+C4.3+C4.5 |
| 3 | C3.1+C2.1+C2.5+C1.3+C6.1+C6.2+C6.4+C4.3+C4.7 |
| 4 | C3.3+C3.6+C2.2+C6.2+C6.3+C6.4 |
| 5 | C3.3+C3.4+C3.5+C2.3+C2.5+C1.1+C1.3+C6.2+C6.3+C6.4+C6.5+C4.3+C4.5 |
| 6 | C3.1+C3.2+C2.4+C6.2+C6.3+C4.3 |
| 7 | C3.1+C3.2+C3.3+C3.5+C2.5+C2.6+C1.3+C6.1+C6.2+C6.4+C6.5+C4.3 |
| 8 | C3.3+C3.5+C2.3+C5.1+C5.3+C1.1+C1.2+C1.4+C1.5+C6.3+C6.4+C4.3+C4.5 |
| 9 | C3.3+C3.5+C2.5+C5.2+C1.1+C1.3+C6.1+C6.2+C6.4+C6.5+C4.2+C4.3+C4.4+C4.5+C4.6 |

| | |
|---|---|
| 10 | C3.3+C3.5+C2.3+C5.3+C5.4+C1.1+C1.4+C1.5+C6.2+C6.4+C6.5+C4.3+C4.5 |
| 11 | C3.2+C3.3+C3.4+C3.5+C2.5+C1.1+C1.3+C6.1+C6.2+C6.3+C6.4+C6.5+C4.2+C4.3+C4.4+C4.5+C4.6 |
| 12 | C3.1+C3.2+C3.3+C3.4+C3.5+C1.2+C1.5+C6.1+C6.2+C6.3+C6.4+C6.5+C4.3+C4.4+C4.5+C4.6 |
| 13 | C3.2+C3.3+C3.4+C3.5+C1.1+C1.2+C1.4+C1.5+C6.1+C6.2+C6.3+C6.4+C6.5+C4.2+C4.3+C4.4+C4.5+C4.6 |
| 14 | C3.3+C3.5+C5.1+C5.2+C1.1+C1.4+C1.5+C6.4+C4.1+C4.2+C4.3+C4.4+C4.5 |
| 15 | C3.1+C3.3+C3.5+C2.5+C2.6+C1.3+C6.1+C6.2+C6.4+C4.3+C4.7 |

Table 3. The clusters in the matrix produced by Metric B

| | |
|---|---|
| 1 | C3.1+C3.2+C3.3+C3.4+C3.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.4+C6.4+C4.6 |
| 2 | C3.2+C3.3+C3.4+C3.5+C3.6+C5.3+C5.4+C1.1+C1.2+C1.4+C1.5 |
| 3 | C3.1+C3.5+C2.1+C2.3+C2.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.3+C4.4+C4.5+C4.6 |
| 4 | C3.2+C3.3+C3.5+C3.6+C2.2+C2.5+C1.1+C1.2+C1.5 |
| 5 | C3.1+C3.3+C3.5+C2.3+C2.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.3+C1.4+C6.1+C6.3+C6.4+C6.5+C4.4+C4.5+C4.6 |
| 6 | C3.5+C3.6+C2.3+C2.4+C2.5+C5.3+C1.1+C4.4 |
| 7 | C3.1+C3.2+C3.3+C3.5+C2.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.4+C6.4+C4.6 |
| 8 | C3.1+C3.4+C3.5+C2.6+C5.2+C5.3+C5.4+C1.1+C1.4+C6.1+C4.4+C4.5+C4.6 |
| 9 | C3.3+C3.5+C2.2+C2.3+C2.5+C5.1+C1.1+C1.2+C1.5+C6.5+C4.4 |
| 10 | C3.1+C3.3+C3.4+C3.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.4+C6.1+C6.4+C6.5+C4.4+C4.5+C4.6 |
| 11 | C3.1+C3.3+C3.4+C3.5+C5.4+C1.1+C1.2+C1.4+C6.1+C6.2+C6.4+C6.5+C4.3+C4.4+C4.6 |
| 12 | C3.1+C3.3+C2.3+C2.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.3+C1.4+C6.1+C6.3+C4.1+C4.2+C4.4+C4.5+C4.6 |
| 13 | C3.1+C3.3+C2.3+C2.5+C5.2+C5.3+C5.4+C1.1+C1.2+C1.3+C1.4+C6.1+C6.3+C6.4+C6.5+C4.2+C4.4+C4.5+C4.6 |
| 14 | C3.1+C3.2+C3.3+C3.4+C3.5+C5.3+C5.4+C1.1+C1.2+C1.4+C6.4+C4.3+C4.6 |
| 15 | C3.1+C3.3+C2.3+C2.5+C5.2+C5.3+C1.1+C1.2+C1.3+C1.4+C6.1+C6.3+C6.4+C6.5+C4.2+C4.4+C4.5+C4.6+C4.7 |

Table 4. The clusters in the matrix produced by Metric C

| | |
|---|---|
| 1 | C3.1+C3.2+C3.3+C3.4+C3.5+C1.1+C1.4+C6.1+C6.4 |
| 2 | C3.3+C3.4+C3.5+C3.6+C5.3+C5.4+C1.1+C1.2+C1.5 |
| 3 | C3.1+C2.1+C2.5+C2.6+C4.5 |
| 4 | C3.6+C2.2 |
| 5 | C3.3+C3.5+C2.3+C2.5+C5.3+C5.4+C1.1+C1.4+C1.5+C6.4+C4.3+C4.4 |
| 6 | C2.4+C2.5+C2.6 |
| 7 | C3.1+C3.3+C3.5+C2.5+C1.1+C1.3+C1.4+C6.1+C6.3+C6.4+C4.3+C4.4 |
| 8 | C3.1+C3.2+C2.6 |
| 9 | C2.1+C2.3+C2.5+C5.1+C1.1+C1.3+C4.4+C4.5 |
| 10 | C3.3+C3.5+C2.5+C5.2+C5.3+C5.4+C1.1+C1.4+C1.5+C6.4+C4.3+C4.4 |
| 11 | C3.2+C3.3+C3.4+C3.5+C5.4+C1.1+C1.2+C1.4+C6.1+C6.4 |

| | |
|---|---|
| 12 | C3.1+C3.3+C3.4+C3.5+C1.1+C1.4+C6.1+C6.2+C6.3+C6.4+C6.5+C4.3 |
| 13 | C2.5+C5.1+C5.2+C5.3+C5.4+C1.1+C1.4+C1.5+C4.1+C4.2+C4.4+C4.5+C4.6 |
| 14 | C3.1+C3.3+C2.5+C1.1+C1.3+C1.4+C6.1+C6.3+C6.4+C4.3+C4.4+C4.5+C4.6 |
| 15 | C3.1+C2.5+C1.1+C1.3+C1.4+C6.1+C6.3+C6.4+C4.3+C4.4+C4.5+C4.6+C4.7 |

The above tables show that there are overlapping between the clusters. Accordingly, a factor is listed 1 to 15 times in the clusters in the new models. The role of the factors in the clusters is introduced in Table 5.

Table 5. The role (frequency) of the factors in the new models

| | | Metric A | Metric B | Metric C |
|---|---|---|---|---|
| C1.1 | Engineering knowledge | 15 | 11 | 11 |
| C1.2 | Technological system and its coherence | 12 | 2 | 2 |
| C1.3 | Local geographical and infrastructural conditions | 5 | 4 | 4 |
| C1.4 | Technical requirements in the EU and national policy | 11 | 9 | 9 |
| C1.5 | Technical level of equipment | 3 | 4 | 4 |
| C2.1 | Impact on environmental elements | 1 | 2 | 1 |
| C2.2 | Waste recovery | 2 | 1 | 2 |
| C2.3 | Geographical factor | 7 | 2 | 2 |
| C2.4 | Resource use | 1 | 1 | 1 |
| C2.5 | Wildlife (social acceptance) | 9 | 9 | 9 |
| C2.6 | Environmental feedback | 1 | 3 | 3 |
| C3.1 | Composition and income level of the population | 11 | 7 | 7 |
| C3.2 | Changes in public service fees | 5 | 3 | 3 |
| C3.3 | Depreciation and resource development | 12 | 8 | 8 |
| C3.4 | Economic interest of operators | 6 | 4 | 4 |
| C3.5 | Financing | 12 | 7 | 7 |
| C3.6 | Structure of industry | 3 | 2 | 2 |
| C4.1 | Public opinion | 1 | 1 | 1 |
| C4.2 | Public health | 3 | 1 | 1 |
| C4.3 | Political and power factors | 3 | 6 | 6 |
| C4.4 | Education | 10 | 7 | 7 |
| C4.5 | Culture | 7 | 5 | 5 |
| C4.6 | Social environment | 11 | 3 | 3 |
| C4.7 | Employment | 1 | 1 | 1 |
| C5.1 | Monitoring and sanctioning | 1 | 3 | 2 |

| | | | | |
|---|---|---|---|---|
| C5.2 | Internal and external legal coherence (domestic law) | 9 | 2 | 2 |
| C5.3 | General waste management regulation in the EU | 12 | 4 | 4 |
| C5.4 | Policy strategy and method of implementation | 11 | 5 | 5 |
| C6.1 | Publicity, transparency (data management) | 7 | 6 | 6 |
| C6.2 | Elimination of duplicate authority | 1 | 1 | 1 |
| C6.3 | Fast and flexible administration | 4 | 4 | 4 |
| C6.4 | Cooperation among institutions | 8 | 8 | 8 |
| C6.5 | Improvement of professional standards | 6 | 1 | 1 |

Based on international experience, perhaps it is still surprising that the most important element of the system is 'Engineering knowledge'. It is followed by financial, technological and legal factors (C1.2, C3.3, C3.5, C1.4, and C5.3).

The authors concluded that Metric A to B and C shows a match of 75% in terms of the most common elements however B to C and vica versa shows a match of 94%.

The authors also verified that 12 factors out of the most common 16 elements occur often in each cluster (max. 15, min. 4 times). In this sense, the metrics gave a very similar outcome.

As a result of the introduces fuzzy cognitive modelling techniques and the proposed new model reduction method it can be stated that the above listed factors can be of greatest importance on the management of the circular economy as a sustainable waste and material management system.

Since the factors having the most significant effect on a system's sustainability receive the proper emphasis during the design and operation process, the effect of the other factors contributes also to the long-term management of the system.

The model formulated on the basis of the proposed method can be an example of how an environmentally and socially-economically mission can be done in a way to be able to provide a favourable solution from economic, legal and institutional point of view.

## 4 Conclusions

The study presented here has got a limit, which the authors have faced during the research, and which may also occur as a problem for future applications. Six to eight professionals with extensive experience in their fields are needed to support the fuzzy cognitive map methodology. The work of the group of experts needed to be moderated by an environmental specialist, who also interprets the results. So, at this stage of the evaluation the expertise and experience is of great importance. The support of an IT staff member is also required who performs the simulations based on the in-put data and help in producing results.

Therefore, it can be concluded that for carrying out the proposed assessment methodology, expert knowledge is needed to ensure the reliability of the results obtained.

# 5 Future research

The authors' purpose is to continue the investigation to understand the deeper context of the circular economy and try to develop a refined model.

## References


1. Zadeh, L. A.: Fuzzy sets. Information and Control, Volume 8, Issue 3, June 1955, Pages 338–353.
2. Axelrod, R.: Structure of Decision: The Cognitive Map of Political Elites. Princeton University Press, 1975.
3. Kosko, B.: Fuzzy Cognitive Maps. Int. J. of Man–Machine Studies, 24(1): 55–75, 1985.
4. Buruzs A., Hatwágner F. M., Kóczy T. L.: Expert-Based Method of Integrated Waste Management Systems for Developing Fuzzy Cognitive Map. In: Taher, A. A., Quanmin, Z. (ed.): Complex System Modelling and Control through Intelligent Soft Computations. Springer International Publishing Switzerland, 2015, pp. 111-137, Studies in Fuzziness and Soft Computing, 319.
5. Buruzs A., Hatwágner F. M., Földesi P., Kóczy L. T.: Fuzzy Cognitive Maps Applied in Integrated Waste Management Systems. In: IEEE HNICEM-ISCIII 2014. Paper 63. 4 p.
6. Buruzs A., Kóczy L. T., Hatwágner F. M.: Using Fuzzy Cognitive Maps approach to identify integrated waste management system characteristics. In: Baranyi P. (ed.) 5th IEEE International Conference on Cognitive Infocommunications: CogInfoCom 2014, IEEE Hungary Section, 2014. pp. 141-147. ISBN:978-1-4799-7279-1.
7. Buruzs A., Kóczy T. L., Hatwágner F. M.: Studies on the sustainability of integrated waste management systems. In: Kóczy T., Tormási A., Kulczycki, P., Talabă, D., Kacprzyk, J. (ed.): Proceedings of the 6th Győr Symposium and 3rd Hungarian-Polish and 1st Hungarian-Romanian Joint Conference on Computational Intelligence. Győr, Széchenyi István University, pp. 201-204.
8. Hatwágner F. M., Buruzs A., Földesi P., Kóczy T. L.: Strategic Decision Support in Waste Management Systems by State Reduction in FCM Models. Lecture Notes in Computer Science 8836: pp. 447-457., 21st International Conference, ICONIP 2014. Kuching, Malaysia, ISBN 978-3-319-12642-5; 978-3-319-12643-2.
9. Buruzs A., Hatwágner F. M., Pozna, R. C., Kóczy L. T.: Advanced Learning of Fuzzy Cognitive Maps of Waste Management by Bacterial Algorithm. In: Pedrycz, W., Reformat, M. Z. (ed.): Proceedings of the 2013 Joint IFSA World Congress NAFIPS Annual Meeting (IFSA/NAFIPS), Edmonton: IEEE, 2013. pp. 890-895, ISBN:978-1-4799-0347-4
10. Paul, M. C. (1959) Minimizing the Number of States in Incompletely Specified Sequential Switching Functions, in: IRE Transactions on Electronic Computers, Vol. EC-8, Issue 3, pp. 356-367.
11. Das, M., Chakraborty, M. K., Ghoshal, T. K. (1998) Fuzzy tolerance relation, fuzzy tolerance space and basis, in: Fuzzy Sets and Systems, Vol. 97, Issue 3, pages 361-369.
12. Klir, G. J., Folger, T. A. (1987) Fuzzy Sets, Uncertainty and Information, Prentice Hall.
13. Hatwágner F. M., Buruzs A., Földesi P., Kóczy L. T.: A New State Reduction Approach for Fuzzy Cognitive Map With Case Studies for Waste Management Systems. In: Somnuk, P-A., Thien, W. A. (ed.): Computational Intelligence in Information Systems: Proceeding of the 4th INNS Symposia Series on Computational Intelligence in Information Systems (INNS-CIIS 2014). 362 p. Cham: Springer Int. Publ. Switzerland, 2015. pp. 119-127. Advances in Intelligent Systems and Computing; 331. (ISBN:978-3-319-13152-8; 978-3-319-13153-5), 2014.

# Analyzing the Process of Constructing Reducts in Multi-adjoint Concept Lattices ⋆

M. Eugenia Cornejo[1], Jesús Medina[2], and Eloísa Ramírez-Poussa[2]

[1] Department of Statistic and O.R., University of Cádiz, Spain
`mariaeugenia.cornejo@uca.es`
[2] Department of Mathematics, University of Cádiz, Spain
`jesus.medina@uca.es, eloisa.ramirez@uca.es`

**Abstract.** Attributes contained in a database usually provide redundant information, hence an essential part in Formal Concept Analysis is getting procedures to remove the unnecessary attributes. In this work, we present some properties related to the attributes that generate irreducible elements of a multi-adjoint concept lattice. These properties are very useful in order to obtain minimal sets of attributes keeping the knowledge of the database.

**Keywords:** attribute reduction, reduct, multi-adjoint concept lattice.

## 1 Introduction

An important part of Formal Concept Analysis (FCA) [1, 4, 8, 13] is the reduction of the number of attributes, since databases usually contain redundant information. One of the main goals of this mathematical theory is to obtain minimal subsets of attributes preserving the knowledge contained in the initial database, from which we can build a concept lattice isomorphic to the original one. From these minimal subsets of attributes, called reducts, interesting information about data is obtained. For instance, the construction process of the concept lattice is made easier since the difficulty depends on the number of attributes and objects. In addition, we can also obtain information about the attributes implications in our context.

On the other hand, due to its broad appeal new fuzzy extensions of this theory were introduced [2, 3, 7, 12]. In this paper, we consider FCA within the multi-adjoint concept lattice framework, which was introduced in [9, 10]. Specifically, we study how to obtain reducts considering an attribute reduction method based on a characterization of the meet-irreducible elements of a multi-adjoint concept lattice presented in [6]. This mechanism classifies the set of attributes in absolutely necessary, relatively necessary and absolutely unnecessary attributes, as in Rough Set Theory (RST) [11]. Therefore, it can be used to build reducts.

⋆ Partially supported by the Spanish Economy and Competitiveness Ministry (MINECO) project TIN2016-76653-P

In addition, since the meet-irreducible elements of a concept lattice are related to the obtained attribute classification, we will focus on the attributes that generate meet-irreducible elements of a multi-adjoint concept lattice. We will study properties about this kind of attributes and will analyze its influence on the reducts of the context. Finally, we also include an illustrative example.

## 2 Attribute classification in multi-adjoint concept lattices

First and foremost, it is convenient to mention that the notions corresponding to the multi-adjoint concept lattice environment are supposed known by the reader. For that reason, we will only fix the necessary framework to present our study in this section. We will consider a multi-adjoint framework $(L_1, L_2, P, \&_1, \dots, \&_n)$ and a multi-adjoint context $(A, B, R, \sigma)$. The tuple of the multi-adjoint frame is composed by two complete lattices $(L_1, \preceq_1)$ and $(L_2, \preceq_2)$, a poset $(P, \leq)$ and the conjunctors of a family of adjoint triples. With respect to the context, this is formed by sets of attributes and objects $A$ and $B$, respectively, a $P$-fuzzy relation $R\colon A \times B \to P$ and a mapping $\sigma$ which associates any element in $A \times B$ with some particular adjoint triple in the frame.

Hereunder, the attribute classification theorems obtained by using the characterization of the meet-irreducible elements of a multi-adjoint concept lattice will be reminded. The set of meet-irreducible elements will be denoted as $M_F(A)$ [6]. In addition, we need to define the following sets:

$$E_{a_i,x} = \{a_j \in A \setminus \{a_i\} \mid \text{there exists } x' \in L_1, \text{ satisfying } \phi^{\downarrow}_{a_i,x} = \phi^{\downarrow}_{a_j,x'}\}$$

where $a_i \in A$, $x \in L_1$ and $\phi_{a_i,x}$, $\phi_{a_j,x'}$ are the fuzzy attributes defined in [6].

**Theorem 1.** *Given $a_i \in A$, we have that:*

*(1) It is absolutely necessary, $a_i \in C_f$, if and only if there exists $x_i \in L_1$, such that $\langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \in M_F(A)$, satisfying that $\langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \neq \langle \phi^{\downarrow}_{a_j,x_j}, \phi^{\downarrow\uparrow}_{a_j,x_j} \rangle$, for all $x_j \in L_1$ and $a_j \in A$, with $a_j \neq a_i$.*

*(2) It is relatively necessary, $a_i \in K_f$, if and only if $a_i \notin C_f$ and there exists $\langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \in M_F(A)$ satisfying that $E_{a_i,x_i}$ is not empty and $A \setminus E_{a_i,x_i}$ is a consistent set.*

*(3) It is absolutely unnecessary, $a_i \in I_f$, if and only if, for each $x_i \in L_1$, we have that $\langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \notin M_F(A)$, or if $\langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \in M_F(A)$, then $A \setminus E_{a_i,x_i}$ is not a consistent set.*

The previous theorem plays a key role in the computation of the minimal sets of attributes, which are called reducts. An initial survey about the building process of reducts will be carried out in the following section, since they can provide a significant reduction of the computational complexity of the concept lattice. The reader can find more information about all these notions in [5, 6].

## 3 Properties of reducts in multi-adjoint concept lattices

In order to obtain reducts in multi-adjoint concept lattices is fundamental to know whether an attribute should be considered or not. It is easy to see that unnecessary attributes must be removed whereas the absolutely necessary attributes must be included in all reducts. Due to different reducts can be obtained when the set of relatively necessary attributes is nonempty, our main goal will consist in giving the first steps in order to select this kind of attributes.

Taking into account the relationship between the given attribute classification and the meet-irreducible elements of a concept lattice, this section will analyze the attributes that generate the meet-irreducible elements of a multi-adjoint concept lattice. The following definition will be very useful for this survey.

**Definition 1.** *Given a multi-adjoint frame* $(L_1, L_2, P, \&_1, \dots, \&_n)$ *and a context* $(A, B, R, \sigma)$ *with the associated concept lattice* $(\mathcal{M}, \leq)$*. Let* $C$ *be a concept of* $(\mathcal{M}, \leq)$*, we define* the set of attributes generating $C$ *as the set:*

$$\mathit{Atg}(C) = \{a_i \in A \mid \textit{there exists} \quad \phi_{a_i,x} \in \Phi \quad \textit{such that} \quad \langle \phi_{a_i,x}^{\downarrow}, \phi_{a_i,x}^{\downarrow\uparrow} \rangle = C\}$$

From now on, interesting properties corresponding to the attributes of a multi-adjoint context will be introduced. These properties will be helpful in the construction process of reducts.

**Proposition 1.** *If* $C$ *is a meet-irreducible concept of* $(\mathcal{M}, \leq)$*, then* $\mathit{Atg}(C)$ *is a nonempty set.*

A characterization of the singleton sets of attributes generating a concept is shown in the proposition below.

**Proposition 2.** *If* $C$ *is a meet-irreducible concept of* $(\mathcal{M}, \leq)$ *satisfying that* $\mathit{card}(\mathit{Atg}(C)) = 1$*, then* $\mathit{Atg}(C) \subseteq C_f$*.*

In general, the converse of the previous result is not satisfied. However, if the set of absolutely necessary attributes is nonempty, the existence of a meet-irreducible element $C$ satisfying that $\mathrm{card}(\mathrm{Atg}(C)) = 1$ is guaranteed.

**Proposition 3.** *If the attribute* $a \in C_f$ *then there exists* $C \in M_F(A)$ *such that* $a \in \mathit{Atg}(C)$ *and* $\mathit{card}(\mathit{Atg}(C)) = 1$*.*

The corollary below is a straightforward consequence of the previous properties.

**Corollary 1.** *If* $C$ *is a meet-irreducible concept of* $(\mathcal{M}, \leq)$ *and* $\mathit{Atg}(C) \cap K_f \neq \varnothing$ *then* $\mathit{card}(\mathit{Atg}(C)) \geq 2$*.*

The following result states that there does not exist an absolutely necessary attribute belonging to $\mathrm{Atg}(C)$ being $C$ a meet-irreducible concept obtained from a relatively necessary attribute.

**Proposition 4.** *Let* $C$ *be a meet-irreducible concept.* $\mathit{Atg}(C) \cap K_f \neq \varnothing$ *if and only if* $\mathit{Atg}(C) \cap C_f = \varnothing$*.*

The next proposition shows the required condition in order to guarantee that each relatively necessary attribute is generated by only one meet-irreducible element of the concept lattice.

**Proposition 5.** *If $\mathcal{G}_K = \{Atg(C) \mid C \in M_F(A)$ and $Atg(C) \cap K_f \neq \varnothing\}$ is a partition of $K_f$, each attribute in $K_f$ generates only one meet-irreducible element of the concept lattice.*

Finally, a sufficient condition to ensure that all the reducts have the same cardinality is established.

**Theorem 2.** *When the set $\mathcal{G}_K = \{Atg(C) \mid C \in M_F(A)$ and $Atg(C) \cap K_f \neq \varnothing\}$ is a partition of $K_f$, then all the reducts $Y \subseteq A$ have the same cardinality. Specifically, the cardinality is $card(Y) = card(C_f) + card(\mathcal{G}_K)$.*

The computation of reducts for a particular multi-adjoint concept lattice will be shown in the following section, by using an interesting example. Specifically, the example will let us clarify the obtained properties in Proposition 5 and Theorem 2.

## 4 An illustrative example

First of all, we will establish a multi-adjoint frame and context in which we will compute the reducts. Consider the frame $([0,1]_{10}, [0,1]_4, [0,1]_5, \preceq, \&_G^*)$ composed by three complete lattices which are regular partitions of the unit interval in 10, 4 and 5 pieces, respectively, and the discretization of the Gödel conjunctor $\&_G^*$ defined on $[0,1]_{10} \times [0,1]_4$. The set of attributes $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, the set of objects $B = \{b_1, b_2, b_3\}$, the relation $R\colon A \times B \to [0,1]_5$ given by Table 1 and the mapping $\sigma$ which is constantly $\&_G^*$ form the context $(A, B, R, \sigma)$.

**Table 1.** Relation R

| $R$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| $b_1$ | 0.6 | 0.2 | 0.2 | 0 | 1 | 0.6 |
| $b_2$ | 0.8 | 0.4 | 0.6 | 0.6 | 1 | 0.8 |
| $b_3$ | 0.6 | 0.6 | 0.2 | 0 | 0 | 0 |

As we mentioned above, the attribute classification will facilite us the computation of reducts. Therefore, we will need to know the meet-irreducible elements of the concept lattice and the fuzzy-attributes associated with them in order to obtain this classification. Taking into account Theorem 1 and the Hasse diagram of the concept lattice exposed in Figure 1, we have:

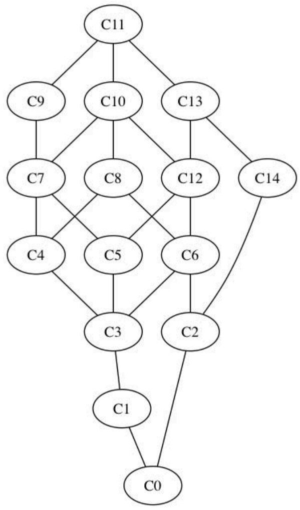
$$C_f = \{a_1, a_2\}$$
$$K_f = \{a_3, a_4, a_5, a_6\}$$

Since $a_1$ and $a_2$ are absolutely necessary attributes, they must be included in all reducts. Hereunder, we will examine the attributes generating each meet-irreducible concept in order to select the relatively necessary attributes that should be included in each reduct.

$$\begin{array}{ll} \text{Atg}(C_1) = \{a_3, a_4\} & \text{Atg}(C_{10}) = \{a_1\} \\ \text{Atg}(C_8) = \{a_1\} & \text{Atg}(C_{13}) = \{a_2\} \\ \text{Atg}(C_9) = \{a_5, a_6\} & \text{Atg}(C_{14}) = \{a_2\} \end{array}$$

**Fig. 1.** Meet-irreducible concepts $M_F(A)$ and Hasse diagram of $(\mathcal{M}, \preceq)$.

| $M_F(A)$ | Fuzzy-attributes generating meet-irreducible concepts |
|---|---|
| $C_1$ | $\phi_{a_3,0.7}, \phi_{a_3,0.8}, \phi_{a_3,0.9}, \phi_{a_3,1}$ $\phi_{a_4,0.7}, \phi_{a_4,0.8}, \phi_{a_4,0.9}, \phi_{a_4,1}$ |
| $C_8$ | $\phi_{a_1,0.9}, \phi_{a_1,1}$ |
| $C_9$ | $\phi_{a_5,0.1}, \phi_{a_5,0.2}, \phi_{a_5,0.3}, \phi_{a_5,0.4}$ $\phi_{a_5,0.5}, \phi_{a_5,0.6}, \phi_{a_5,0.7}, \phi_{a_5,0.8}$ $\phi_{a_5,0.9}, \phi_{a_5,1}$ $\phi_{a_6,0.1}, \phi_{a_6,0.2}, \phi_{a_6,0.3}, \phi_{a_6,0.4}$ $\phi_{a_6,0.5}, \phi_{a_6,0.6}$ |
| $C_{10}$ | $\phi_{a_1,0.7}, \phi_{a_1,0.8}$ |
| $C_{13}$ | $\phi_{a_2,0.3}, \phi_{a_2,0.4}$ |
| $C_{14}$ | $\phi_{a_2,0.5}, \phi_{a_2,0.6}$ |




Clearly, the set $\mathcal{G}_K = \{\text{Atg}(C) \mid C \in M_F(A) \text{ and } \text{Atg}(C) \cap K_f \neq \varnothing\} = \{\text{Atg}(C_1), \text{Atg}(C_9)\}$ is a partition of $K_f$ because $\text{Atg}(C_1)$ and $\text{Atg}(C_9)$ are disjoint subsets of $K_f$. As a consequence, applying Proposition 5, we can ensure that each attribute in $K_f$ generates only one meet-irreducible element of the concept lattice. This fact can be easily seen in the table displayed in Figure 1. In addition, Theorem 2 guarantees that all the reducts have the same cardinality as it is shown below:

$$\begin{array}{ll} Y_1 = \{a_1, a_2, a_3, a_5\} & Y_3 = \{a_1, a_2, a_4, a_5\} \\ Y_2 = \{a_1, a_2, a_3, a_6\} & Y_4 = \{a_1, a_2, a_4, a_6\} \end{array}$$

These reducts give rise to concept lattices isomorphic to the original one.

## 5 Conclusions and future work

We have provided an initial study about the construction process of reducts in multi-adjoint concept lattices, where the attribute classification introduced in [6] has played a fundamental role. We have presented interesting properties and results in order to emphasize the significance of the selection of the relatively necessary attributes for determining the reducts.

As a future work, we will continue analyzing more properties corresponding to reducts with the purpose of generate them in the most suitable way. Another important task will be to develop an algorithm to compute reducts.

## References


1. L. Antoni, S. Krajči, and O. Krídlo. Constraint heterogeneous concept lattices and concept lattices with heterogeneous hedges. *Fuzzy Sets and Systems*, 2015, In Press.
2. R. Bělohlávek. Lattices of fixed points of fuzzy Galois connections. *Mathematical Logic Quartely*, 47(1):111–116, 2001.
3. A. Burusco and R. Fuentes-González. Construction of the $L$-fuzzy concept lattice. *Fuzzy Sets and Systems*, 97(1):109–114, 1998.
4. J. Chen, J. Li, Y. Lin, G. Lin, and Z. Ma. Relations of reduction between covering generalized rough sets and concept lattices. *Information Sciences*, 304:16 – 27, 2015.
5. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. On the classification of fuzzy-attributes in multi-adjoint concept lattices. *Lecture Notes in Computer Science*, 7903:266–277, 2013.
6. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294:41 – 56, 2015.
7. S. Krajči. A generalized concept lattice. *Logic Journal of IGPL*, 13(5):543–550, 2005.
8. J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
9. J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.
10. 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.
11. Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
12. S. Pollandt. *Fuzzy Begriffe*. Springer, Berlin, 1997.
13. L. Wei and J.-J. Qi. Relation between concept lattice reduction and rough set reduction. *Knowledge-Based Systems*, 23(8):934–938, 2010.

# Uncertainty Tolerance and Behavioral Stability Analysis of Fixed Structure Fuzzy Cognitive Maps

Miklós F. Hatwágner[1] and László T. Kóczy[2]

[1] Department of Information Technology, Széchenyi István University, Győr, Hungary
miklos.hatwagner@sze.hu
[2] Department of Information Technology, Széchenyi István University, Győr, Hungary and Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary
koczy@sze.hu, koczy@tmit.bme.hu

**Abstract.** Fuzzy Cognitive Maps (FCMs) are widely applied to describe, model and simulate complex systems. It supports decision making and helps better understand the operation of multicomponent systems. FCMs can be considered as digraphs: the nodes represent the components of the system, while the arcs express the cause-effect relationships among these components. There are two possible main ways of model creation: expert-based and automated methods.

Expert based methods are based on the knowledge, experience and opinions of experts of the investigated field. As a consequence of this approach, the resulting models contain more or less subjective information that may decrease the trustworthiness of the model, even if the final model aggregates the information provided by individual models of several experts.

This phenomenon can be avoided by using automated methods. The drawback of this approach is that these methods require the availability of historical data.

A novel technique is proposed in this paper in order to investigate the effect of slightly modified model parameters (arc weights) on simulation results. The most influencing parameters can be detected this way, and their values can be revised in order to model the studied system in a more accurate way. It proved to be especially useful in practice when the model is created by experts. The paper presents an example as well to demonstrate that the performed modifications can give rise to different kind of model behaviors, different amount of fixed-point attractors.

**Keywords:** fuzzy cognitive maps, model uncertainty, behavioral stability, multiobjective optimization, Bacterial Evolutionary Algorithm.

## 1 Introduction

Fuzzy Cognitive Maps (FCM) are a proper tool for decision makers to describe complex systems, including their components and the direction and strength of relationships among these components. Several papers can be found in the literature about how to create models and use them for simulation purposes and decision support [1]. The novelty of this paper is to provide a method to analyze the uncertainty and behavioral stability of models. This way the effect of small weight changes on model behavior can be studied. There are at least two good reasons why does it worth perform such an analysis.

If the model is provided by experts, viz. human beings, the connection matrix of the model is sometimes not perfect. In practice, it is not easy to define the weight of a connection between two components of the investigated system if the size of that system is huge. Obviously, they did their best, but the number of connections is commensurate with the square of the number of components. For example, if the investigated system has 10 components, the number of connections can be up to 90. It is very hard to see the system as a whole if it has so many details and chose well-balanced weights to represent the real relationships. If these weights are poorly estimated, the simulation of the system will lead to states that may never occur in the real world.

The results of this paper are interesting even if the weights are already well chosen, or defined by a sort of machine learning or optimization algorithm, because if the behavior of the modified models is known, we can see how and where should the model be changed in order to eventuate a better operation, or what effects jeopardize the operation of it.

The next section describes briefly the theoretical background of our investigation, including FCMs and Bacterial Evolutionary Algorithm (BEA) which is used to find the most interesting, slightly modified model versions. It is followed by the description of our suggested method. The next section contains the analysis results of an example model. Finally, the directions of further improvements and conclusions are summarized.

## 2 Theoretical background

### 2.1 A short overview of Fuzzy Cognitive Maps

The first cognitive maps were created by Axelrod [2] to support political decision making. His technique was further improved by Kosko [3, 4], and Fuzzy Cognitive Maps (FCM) was born. FCMs are fuzzy signed directed graphs: the nodes of the graph represent the components of the modeled system. These are also called 'concepts' in FCM theory. They can have any real value in the [0 +1] interval, describing e.g. the state of a valve (fully or partly opened, closed) or a gauge (current fluid level) [5]. (In some special environments, [−1 +1] interval is used [6].) The cause-effect relationships among concepts are expressed by the weight of arcs. These values have to be in the [−1 +1] interval; positive sign means positive effect on the connected concept, and vice versa. The absolute value of the connection weight depends on the strength of that connection. The system can be depicted by a graph or a connection matrix. In most cases the main diagonal contains zeros according to Kosko's original idea, that is, the concepts cannot influence their own state directly. The FCM used to demonstrate the proposed method are described by Fig. 1 and Table 1.

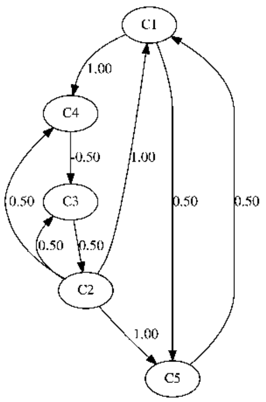



**Fig. 1** Graph representation of the investigated FCM model

Table 1. Connection matrix

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| C1 | 0 | 0 | 0 | 1 | 0.5 |
| C2 | 1 | 0 | 0.5 | 0.5 | 1 |
| C3 | 0 | 0.5 | 0 | 0 | 0 |
| C4 | 0 | 0 | -0.5 | 0 | 0 |
| C5 | 0.5 | 0 | 0 | 0 | 0 |

If the initial states of concepts and the connection matrix are known, a simulation can be started to investigate the concept values at discrete time steps and the behavior of the modeled system. The next state of the model can be calculated by Eq. (1).

$$A_i^{(t+1)} = f\left(\sum_{j=1}^{M} w_{ji}\, A_j^{(t)} + A_i^{(t)}\right), i \neq j \tag{1}$$

$A_i^{(t)}$ represents the value of concept $i$ at time $t$ (also called 'activation' value), $w_{ji}$ is the weight of the directed arc between concepts $j$ and $i$, $M$ is the number of concepts and $f$ is the threshold function.

Note that a version of this equation was first proposed in [7] in a more general form and was called FCM of Type II. The main advantage of this version of FCM is that it takes into account the current value ('state') of a concept during the calculation of the next value of it. In some sense, the concepts 'remember' their current state, and this memory influences their next state. Several systems in real life behave like this. On the contrary, Kosko's original equation (FCM of Type I) defines the next activation value of a concept using only the effects of all other concepts on it. Furthermore, it can be seen that equation (2) and a connection matrix containing only zeros in its main diagonal (see e.g. Table 1) behaves in the same way as Kosko's equation with a connection matrix that has ones in the main diagonal.

Several threshold functions are known in the literature [6]. The sigmoidal one in Eq. (2) will be applied in this paper.

$$f = \frac{1}{1+e^{-\lambda x}} \tag{2}$$

The threshold function guarantees that concept values remain in the allowed range. The λ parameter of the function specifies the steepness of the function and influences the values of concepts, but their order always remains the same. A widely applied value of λ is 5.

The simulation of a FCM that uses Eq. (2) can lead to three different outcomes [6]:

1. In most cases the values of concepts converge quickly to a final, stable value. This vector of concept values often called as ‘fixed point attractors’.
2. Sometimes a series of $n$ state vectors can be observed repeatedly after a specific time step of the simulation. This is called a ‘limit cycle’.
3. The last possible outcome is, when the values of concepts never stabilize, and the model behaves chaotically.

In most cases the fixed point attractors are the desired outcomes, because this way the future states of the system, the consequences of different decisions can be predicted. But sometimes the chaotic or limit cycle behavior proves to be useful, e.g. if the goal is to predict time series data [8].

### 2.2 Bacterial Evolutionary Optimization

Bacterial Evolutionary Algorithm (BEA) is a global optimization algorithm. BEA was originally proposed by Nawa and Furuhashi at the end of 1990s as a new evolutionary algorithm [9, 10]. This algorithm was established as a further development of the already existing Pseudo-Bacterial Genetic Algorithm [11] and the classical Genetic Algorithm [12]. The name of the algorithm indicates that its operations mimics the process of the evolution of bacteria. Individual bacteria represents possible solutions of a problem. BEA keeps a record of all available bacteria, i.e., solutions or model variants, called the bacterium population. BEA creates successive generations of the population by using its two main operators, bacterial mutation and gene transfer, until some kind of termination condition is fulfilled, e.g. a pre-defined number of generations is calculated or the algorithm is unable to find better solutions after evaluating some generations. Finally, the last generation is considered as the result, i.e., a group of model variants with interesting properties. During the simulation process, the bacterial mutation creates new versions of bacteria with random modifications. With other words, this operator is liable for the exploration of the search space. Depending on some parameters governing the spread or deviation of the mutation results, its properties balance between ‘globalness’ and convergence speed. The other operator, namely gene transfer, combines the genetic information of pairs of bacteria. Thus it performs the exploitation of the genetic data.

Bacterial mutation optimizes the bacteria individually. For every bacteria of the population, the operator does the following. First, it creates $K$ copies, the so-called ‘clones’ of the original bacterium. Then it chooses from the genes of the bacterium (these are the non-zero elements of the original connection matrix), and it randomly modifies the value of the selected gene in the clones. The next step is the evaluation of the modified bacteria. If one of the clones proved to be better than the original bacterium, the new value of the gene is copied back to the original bacterium. This process is repeated while all the genes of the original bacteria will not be mutated. At the end of the mutation, all the clones are dropped.

The gene transfer operator combines the genetic data of the bacteria in the hope that it can produce better bacteria. The original version of the gene transfer operator divides the population into two, equally sized parts. The bacteria with better objective function value get into the superior part, the others in the inferior part. Next, the operator repeats T times the following activities. It chooses one bacterium from the superior part and another bacterium from the inferior part. After that, it copies some randomly selected genes of the superior bacterium into the other. Certainly, after such a modification, the objective function has to be evaluated again. Before the next iteration of the gene transfer, the population has to be sorted again considering the objective function values. If the modification of the inferior bacteria was successful enough, it can migrate into the superior part.

## 3 Description of the applied method

The weights of connections in FCM models are represented by real numbers in the [−1 +1] interval, thus theoretically the number of possible weight values of a specific connection is infinite. Obviously, we have to define some levels of weight values in order to decrease the number of possible configurations. For example, if the number of levels is five, the values −1, −0.5, 0, +0.5 and +1 can be applied. According to our experiences, it does not have any significant advantages to increase the number of levels over nine, especially if the original model did not contain more levels.

In order to decrease the computational demand of our investigations and because experts can decide with high reliability whether two components of the system are connected or not, the cells of the connection matrix containing zeros is left untouched.

The behavior of modified models are tested with random generated initial state vectors, called scenarios. The initial values of components (represented by concepts of FCM) also often have some pre-defined levels, e.g. five, according to the used linguistic variables (the fifth possible level, namely −1 is not used in Fig. 1). Despite of these restrictions it is still practically impossible to perform an exhaustive check of all possible initial state vectors on big models. For example, if the number of allowed levels is five and the number of components (concepts) is 10, then the number of initial state vectors is $5^{10}$ = 9 765 625. Every simulation is very time consuming process thus the number of required simulations must be limited. A smaller amount of scenarios, e.g. 1 000 seems enough in most cases and in our specific example case as well to reveal the general behavior of different models, but the exact number obviously depends on the specific model. The constant value of 5 for lambda is used because otherwise the number of possible models would be much greater.

The proposed method explores 'interesting' modified versions of the investigated model. In most cases, the models have only a few fixed-point attractors. In some special cases however several fixed points or limit cycles/chaotic behavior may occur. These cases are rare and therefore more interesting and illuminating than the others. In many cases, the chaotic behavior should be avoided because the system never stabilizes and its state cannot be predicted.

The method uses an implementation of BEA. It runs FCM simulations as objective function to measure the quality of model versions. A randomly modified model version (with respect to the constraints described above) is represented by an individual bacterium. In this paper the optimization executed as an example was stopped after 5 generations and the last generation was considered as result. The population contains 50 bacteria. BEA automatically produces newer and newer modified versions of an original model during mutation and gene transfer and examines their behavior. The initial concept values in scenarios have 5 different levels (0, 0.25, 0.5, 0.75 and 1) while the modified connection matrices have 9 different levels (−1, −0.75, −0.5, …, +1). The method performs 1 000 simulations with the model according to the 1 000 scenarios, identifies and counts the different fixed point attractors and also counts the number of scenarios that lead to chaotic behavior or limit cycles (these two behaviors are not distinguished at the moment).

The applied method compares the bacteria (model variants) in order to decide which one is better than the other. It ranks the models on the grounds of two values: the number of different fixed-point attractors and the sum of chaotic cases and limit cycles. One model is better than the other (with other words, one model dominates the other) if it has at least one property that is better than the other's while all other properties are not worse. For example, if Model 1 has 4 fixed point attractors and never produces chaotic behavior and Model 2 has only 3 different fixed points and also never produces chaotic behavior then Model 1 is better than Model 2.

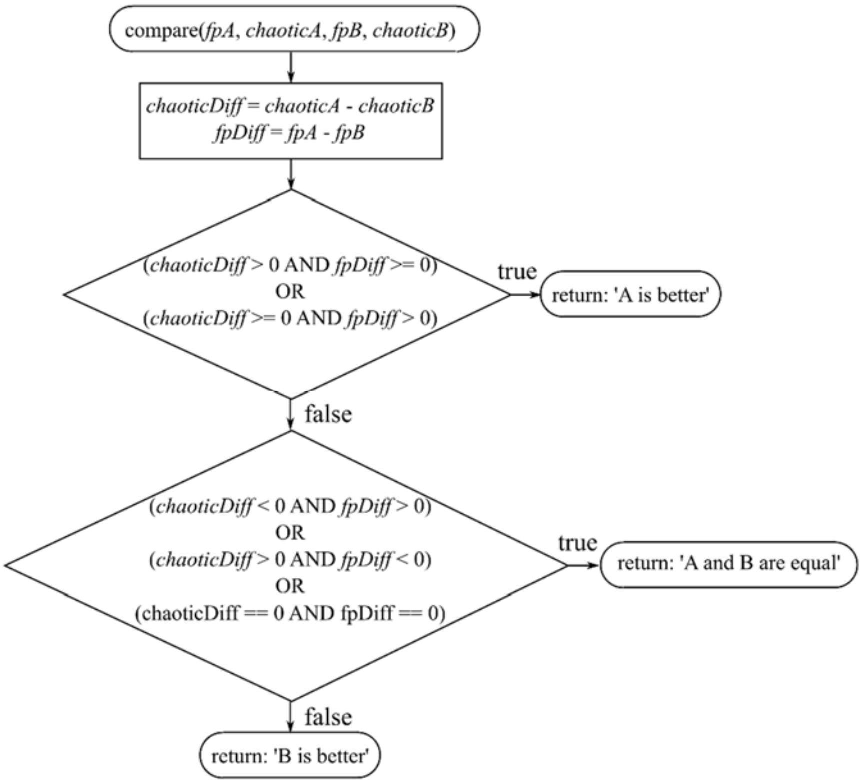



**Fig. 2** Comparison function of two model variants

The ranking of bacteria should be performed by searching for the members of Pareto-fronts of bacteria, where the criteria are the number of fixed point attractors and limit cycles/chaotic cases. All members of the current front are considered ‘equally good’ for the algorithm, and the first front contains the best model variants. In order to find the next front, the bacteria of the current front have to be removed, and the whole process is repeated until all the bacteria are classified into different Pareto-fronts. As this paper documents our first attempt to investigate the uncertainty and behavioral stability of human-made FCM models, only a quick-and-dirty solution is implemented in the program, and it must be changed in the future because it does not necessarily provides the same results in all cases. Currently a Quick sort algorithm sorts the bacteria and it uses the comparison function described by Fig. 2. (*fp* is the number of different fixed point attractors, *chaotic* is the number of scenarios resulted in chaotic behavior. These parameters are given for two models (*A* and *B*) at the same time.)

## 4 Results

Due to the limited extent of the paper only three interesting model variants are shown: model #1, #34 and #48. The first 47 models never behave chaotically, and the number of fixed-points ranges from 3 to 15. The first model variant is the only one that has exactly 3 fixed-point attractors. Both model #34 and #48 have 10 fixed-point attractors, but model #48 has 4 chaotic cases/limit cycles as well, while model #34 has none of these. Concepts that have different final values in fixed-point attractors are emphasized with bold typeface. The percentages in parenthesis show the proportions of different fixed points.

Table 2 Fixed point attractors of Model #1 ($\lambda = 5$)
Chaotic cases or limit cycles were not found.

| FP. ID | C1 | C2 | C3 | **C4** | C5 |
|---|---|---|---|---|---|
| 1 (23.3%) | 0.01 | 0.99 | 1.00 | **0.85** | 1.00 |
| 2 (76.5%) | 0.01 | 0.99 | 1.00 | **0.14** | 1.00 |
| 3 (0.2%) | 0.01 | 0.99 | 1.00 | **0.51** | 1.00 |

Table 3 Connection matrix of Model #1 ($\lambda = 5$)

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| C1 | 0 | 0 | 0 | -0.75 | 0.5 |
| C2 | -0.25 | 0 | 0.25 | -0.5 | 0.25 |
| C3 | 0 | 0 | 0 | 0 | 0 |
| C4 | 0 | 0 | 0.25 | 0 | 0 |
| C5 | -0.75 | 0 | 0 | 0 | 0 |

Table 4 Fixed point attractors of Model #34 ($\lambda = 5$)
Chaotic cases or limit cycles were not found.

| FP. ID | **C1** | **C2** | **C3** | **C4** | **C5** |
|---|---|---|---|---|---|
| 1 (19%) | **0.15** | **0.01** | **0.99** | **0.99** | **0.99** |
| 2 (5.7%) | **0.85** | **0.99** | **0.03** | **0.11** | **0.01** |
| 3 (20.7%) | **0.99** | **0.01** | **0.99** | **1.00** | **0.15** |
| 4 (8.6%) | **0.85** | **0.99** | **0.03** | **0.75** | **0.01** |
| 5 (18.6%) | **0.90** | **0.01** | **0.99** | **1.00** | **0.90** |
| 6 (13.9%) | **0.10** | **0.99** | **0.03** | **0.03** | **0.10** |
| 7 (12.8%) | **0.01** | **0.99** | **0.03** | **0.03** | **0.85** |
| 8 (0.1%) | **0.02** | **0.99** | **0.03** | **0.03** | **0.53** |
| 9 (0.2%) | **0.98** | **0.01** | **0.99** | **1.00** | **0.46** |
| 10 (0.4%) | **0.45** | **0.01** | **0.99** | **1.00** | **0.98** |

Table 5 Connection matrix of Model #34 ($\lambda = 5$)

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| C1 | 0 | 0 | 0 | 0.25 | -0.5 |
| C2 | -0.5 | 0 | -0.75 | -0.75 | -0.5 |
| C3 | 0 | -1 | 0 | 0 | 0 |
| C4 | 0 | 0 | 0 | 0 | 0 |
| C5 | -0.5 | 0 | 0 | 0 | 0 |

Table 6 Fixed point attractors of Model #48 ($\lambda = 5$)
4 chaotic cases/limit cycles were found.

| FP. ID | **C1** | **C2** | **C3** | **C4** | **C5** |
|---|---|---|---|---|---|
| 1 (13.1%) | **0.00** | **0.91** | **0.91** | **0.91** | **0.91** |
| 2 (21.7%) | **0.86** | **0.99** | **0.03** | **0.00** | **0.00** |
| 3 (9.3%) | **0.00** | **0.99** | **0.03** | **0.15** | **0.86** |
| 4 (9.7%) | **0.99** | **0.16** | **0.99** | **0.00** | **0.02** |
| 5 (25.8%) | **0.07** | **0.99** | **0.03** | **0.08** | **0.09** |
| 6 (16.2%) | **0.00** | **0.15** | **1.00** | **0.99** | **0.99** |
| 7 (3.5%) | **0.00** | **0.99** | **0.11** | **0.86** | **0.86** |
| 8 (0.1%) | **0.02** | **0.91** | **0.90** | **0.89** | **0.33** |
| 9 (0.1%) | **0.00** | **0.99** | **0.06** | **0.49** | **0.86** |
| 10 (0.1%) | **0.98** | **0.40** | **0.96** | **0.00** | **0.01** |

Table 7 Connection matrix of Model #48 ($\lambda = 5$)

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| C1 | 0 | 0 | 0 | -1 | -0.75 |
| C2 | -0.5 | 0 | -0.75 | -0.5 | -0.5 |
| C3 | 0 | -0.5 | 0 | 0 | 0 |
| C4 | 0 | 0 | 0.25 | 0 | 0 |
| C5 | -1 | 0 | 0 | 0 | 0 |

# 5 Conclusions and future research

The proposed method is able to generate small modifications on FCM models that lead to very different model behavior. It is especially useful to find relationships that are very sensitive to changes and may cause unexpected simulation results. These connection weights can be further investigated, the importance of these changes can be consulted with experts.

The method could be further improved to increase its usefulness in practice, however. For example, only one lambda value, namely 5 was applied but a different lambda value can also affect the simulation results. In some fields the differentiation of chaotic cases and limit cycles would be important, and other details of the implementation should be improved. The program and sample data can be downloaded at http://rs1.sze.hu/~hatwagnf/escim2016/escim2016.zip

**Acknowledgement.** The authors would like to thank to the National Research, Development and Innovation Office (NKFIH) K105529, K108405 grant for the support of the research.

## References

1. Papageorgiou, E. I. (Ed.). (2013). Fuzzy Cognitive Maps for Applied Sciences and Engineering: From Fundamentals to Extensions and Learning Algorithms (Vol. 54). Springer Science & Business Media.
2. Axelrod, R. (Ed.). (1976). Structure of decision: The cognitive maps of political elites. Princeton University Press.
3. Kosko, B. (1986). Fuzzy cognitive maps. Int. J. of Man-Machine Studies, 24(1), 65-75.
4. Kosko, B. (1988). Hidden patterns in combined and adaptive knowledge networks. International Journal of Approximate Reasoning, 2(4), 377-393.
5. Parsopoulos, K. E., Papageorgiou, E. I., Groumpos, P. P., & Vrahatis, M. N. (2003, December). A first study of fuzzy cognitive maps learning using particle swarm optimization. In Evolutionary Computation, 2003. CEC'03. The 2003 Congress on (Vol. 2, pp. 1440-1447). IEEE.
6. Tsadiras, A. K., (2007). Inference using Binary, Trivalent and Sigmoid Fuzzy Cognitive Maps, Proceedings of the 10th International Conference on Engineering Applications of Neural Networks (EANN 2007) (vol. 284, pp. 327-334)
7. Stylios, C. D., & Groumpos, P. P. (1999, June). Mathematical formulation of fuzzy cognitive maps. In Proc. of the 7th Mediterranean Conf. on Control and Automation (pp. 2251-2261).
8. Lu, W., Pedrycz, W., Liu, X., Yang, J., & Li, P. (2014). The modeling of time series based on fuzzy information granules. Expert Systems with Applications, 41(8), 3799-3808.
9. Nawa, N. E., & Furuhashi, T. (1998, April). A study on the effect of transfer of genes for the bacterial evolutionary algorithm. In Knowledge-Based Intelligent Electronic Systems, 1998. Proceedings KES'98. 1998 Second International Conference on (Vol. 3, pp. 585-590). IEEE.
10. Nawa, N. E., & Furuhashi, T. (1999). Fuzzy system parameters discovery by bacterial evolutionary algorithm. Fuzzy Systems, IEEE Transactions on, 7(5), 608-616.
11. Nawa, N. E., Hashiyama, T., Furuhashi, T., & Uchikawa, Y. (1997, April). A study on fuzzy rules discovery using pseudo-bacterial genetic algorithm with adaptive operator. In Evolutionary Computation, 1997. IEEE International Conference on (pp. 589-593). IEEE.
12. Goldberg, D. E. (1989). Genetic algorithms in search optimization and machine learning (Vol. 412). Reading Menlo Park: Addison-Wesley.

# Towards Multi-adjoint Logic Programming with Negations ⋆

Maria Eugenia Cornejo[1], David Lobo[2], and Jesús Medina[2]

[1] Department of Statistics and O.R., University of Cádiz. Spain
mariaeugenia.cornejo@uca.es
[2] Department of Mathematics, University of Cádiz. Spain
david.lobo@alum.uca.es, jesus.medina@uca.es

**Abstract.** An initial study on multi-adjoint logic programming with negations is presented in this paper. In particular, the existence of stable models for the mentioned logic programs is guaranteed.

**Keywords:** multi-adjoint logic program, negation, stable model.

## 1 Introduction

Multi-adjoint logic programming was presented as a generalization of different non-classical of logic programming theories in [11]. The main feature of this logical framework is related to the use of several implications in the rules of a same logic program, as well as general operators in the bodies of the rules. Recent papers have shown an special interest in this research topic [4, 5, 10, 16].

This manuscript will enrich the semantic structure of multi-adjoint logic programs with an additional operator, that is, a negation operator. Our goal will focus on establishing the first steps to define the semantics for multi-adjoint logic programs with negations by using stable models.

To the best of our knowledge, only sufficient conditions have been stated in order to guarantee the existence and the unicity of stable models in some approaches [3, 6, 13–15]. An important survey on the existence and unicity of stable models for residuated logic programs with negations defined on the unit interval was introduced in [1, 8, 9].

Following the idea shown in [1], we will provide the necessary requirements to ensure the existence of stable models for multi-adjoint programs with negations defined on any convex compact set of an euclidean space.

## 2 Multi-adjoint logic programming

Multi-adjoint logic programming was introduced in [11] generalizing different non-classical logic programming frameworks such that the residuated logic programming [2] and the initial work presented in [17]. The following notion plays a key role in the extension to logic programming to the fuzzy case.

⋆ Partially supported by the Spanish Economy and Competitiveness Ministry (MINECO) project TIN2016-76653-P

**Definition 1.** *Let $(P, \leq)$ be a partially ordered set. We say that $(\&, \leftarrow)$ is an* adjoint pair *with respect to $(P, \leq)$ if the mappings $\&, \leftarrow : P \times P \to P$ satisfy that:*

1. *$\&$ is order-preserving in both arguments.*
2. *$\leftarrow$ is order-preserving in the first argument (the consequent) and order-reversing in the second argument (the antecedent).*
3. *The equivalence $x \leq y \leftarrow z$ if and only if $x \,\&\, z \leq y$ holds, for all $x, y, z \in P$.*

Adjoint pairs are the basic operators to make the computations in fuzzy logic programming. Considering a bounded lattice together with different adjoint pairs, whose conjunctors satisfy the boundary conditions with respect to the top element of the lattice, allows us to define the semantic structure of the multi-adjoint logic programs.

**Definition 2.** *A* multi-adjoint lattice *is a tuple $(L, \preceq, \leftarrow_1, \&_1, \ldots, \leftarrow_n, \&_n)$ verifying the following properties:*

1. *$(L, \preceq)$ is bounded lattice, i.e. it has bottom $(\bot)$ and top $(\top)$ elements;*
2. *$(\&_i, \leftarrow_i)$ is an adjoint pair in $(L, \preceq)$, for all $i \in \{1, \ldots, n\}$;*
3. *$\top \,\&_i\, \vartheta = \vartheta \,\&_i\, \top = \vartheta$, for all $\vartheta \in L$ and for all $i \in \{1, \ldots, n\}$.*

Once we have shown the notions of adjoint pair and multi-adjoint lattice, we can proceed to include the definition of multi-adjoint logic program. Before, it is important to say that we will assume that the reader is familiar with the concepts associated with the syntax and the semantics of propositional languages.

From now on, we will consider a fixed set of propositional symbols denoted as $\Pi$ and a given language denoted as $\mathfrak{F}$. The syntax of a multi-adjoint logic program is based on a set of rules and facts as it is explained below.

**Definition 3.** *Let $(L, \preceq, \leftarrow_1, \&_1, \ldots, \leftarrow_n, \&_n)$ be a multi-adjoint lattice. A* multi-adjoint logic program *is a set of rules of the form $\langle (A \leftarrow_i \mathcal{B}), \vartheta \rangle$ such that:*

1. *The* rule *$(A \leftarrow_i \mathcal{B})$ is a formula of $\mathfrak{F}$;*
2. *The* confidence factor *$\vartheta$ is an element (a truth-value) of $L$;*
3. *The* head *of the rule $A$ is a propositional symbol of $\Pi$.*
4. *The* body *formula $\mathcal{B}$ is a formula of $\mathfrak{F}$ built from propositional symbols $B_1, \ldots, B_n$ ($n \geq 0$) by the use of conjunctors $\&_1, ..., \&_n$ and $\wedge_1, ..., \wedge_k$, disjunctors $\vee_1, ..., \vee_l$, aggregators $@_1, \ldots, @_m$ and elements of $L$.*
5. Facts *are rules with body $\top$.*

In this paper, we are interested in studying a special kind of multi-adjoint logic programs which are called multi-adjoint normal logic programs. The main difference between multi-adjoint normal logic programs and the previous ones is that the multi-adjoint lattice is enriched with a negation operator.

## 3 Syntax and semantics of multi-adjoint normal logic programs

This section will begin with the definition of multi-adjoint normal logic programs. Note that, the considered negation operator $\neg : L \to L$ is a decreasing mapping verifying the equalities $\neg(\bot) = \top$ and $\neg(\top) = \bot$.

**Definition 4.** *Let* $(L, \preceq, \leftarrow_1, \&_1, \dots, \leftarrow_n, \&_n, \neg)$ *be a multi-adjoint lattice with negation. A* multi-adjoint normal logic program $\mathbb{P}$ *is a finite set of weighted rules of the form:*

$$\langle p \leftarrow_i @[p_1, \dots, p_m, \neg p_{m+1}, \dots, \neg p_n]; \vartheta\rangle$$

*where* $i \in \{1, \dots, n\}$*,* $@$ *is an aggregator operator,* $\vartheta$ *is an element of* $L$ *and* $p, p_1, \dots, p_n$ *are propositional symbols such that* $p_j \neq p_k$*, for all* $j, k \in \{1, \dots, n\}$*.*

Henceforth, the set of propositional symbols appearing in a multi-adjoint normal logic program $\mathbb{P}$ will be denoted as $\Pi_{\mathbb{P}}$. In order to define the semantics of multi-adjoint logic programs with negations, we need to include the notion of fuzzy interpretation and model.

**Definition 5.** *A fuzzy* L-interpretation *is a mapping* $I : \Pi_{\mathbb{P}} \to L$ *which assigns a truth value to every propositional symbol appearing in* $\mathbb{P}$*. The set of all interpretations is denoted* $\mathcal{I}_{\mathfrak{L}}$*, where* $\mathfrak{L}$ *is a multi-adjoint algebra in which the multi-adjoint lattice is defined. We say that:*

*(1) A weighted rule* $\langle p \leftarrow_i @[p_1, \dots, p_m, \neg p_{m+1}, \dots, \neg p_n]; \vartheta\rangle$ *is* satisfied *by* $I$ *if and only if* $\vartheta \preceq \hat{I}\,(p \leftarrow_i @[p_1, \dots, p_m, \neg p_{m+1}, \dots, \neg p_n])$*.*
*(2) An interpretation* $I \in \mathcal{I}_{\mathfrak{L}}$ *is a* model *of a multi-adjoint normal logic program* $\mathbb{P}$ *if and only if all weighted rules in* $\mathbb{P}$ *are satisfied by* $I$*.*

In order to make the difference between the syntax and semantics of the operators, we will write $\omega$ to denote an operator symbol in $\Omega$ and $\dot{\omega}$ to denote the interpretation of the previous operator symbol under $\mathfrak{L}$.

The immediate consequence operator is generalized for the multi-adjoint normal logic programs framework as follows.

**Definition 6.** *Let* $\mathbb{P}$ *be a multi-adjoint normal logic program. The* immediate consequence operator *is the mapping* $T^{\mathfrak{L}}_{\mathbb{P}} : \mathcal{I}_{\mathfrak{L}} \to \mathcal{I}_{\mathfrak{L}}$ *defined as*

$$T_{\mathbb{P}}(I)(p) = \sup\{\vartheta \,\dot{\&}_i\, \hat{I}(\mathcal{B}) \mid \langle p \leftarrow_i \mathcal{B}; \vartheta\rangle \in \mathbb{P}\}$$

*where* $p \in \Pi_{\mathbb{P}}$ *and* $\langle p \leftarrow_i \mathcal{B}; \vartheta\rangle$ *denote the rules of* $\mathbb{P}$*.*

When a multi-adjoint logic program (without negation operator) is considered, its immediate consequence operator is monotonic. This fact leads us to characterize the models of a multi-adjoint logic program by using the post fix points of the immediate consequence operator. Knaster-Tarski's fix point theorem together with the mentioned characterization guarantees that the least model of the multi-adjoint program coincides with its least fix point.

However, the immediate consequence operator of a multi-adjoint normal logic program is not necessarily monotonic. Consequently, the existence of the least model cannot be guaranteed. In order to avoid this problem, stable models arose.

## 4 Existence of stable models

The semantics of multi-adjoint normal logic programming is based on the notion of stable model, which was introduced in [7]. The strategy to define this semantics is similar to the employed approach for normal residuated logic programs.

From a multi-adjoint normal logic program $\mathbb{P}$ and a fuzzy $L$-interpretation $I$, we are going to construct a multi-adjoint program without any negation. The mechanism consists in substituting each rule in $\mathbb{P}$ such as

$$\langle p \leftarrow_i @[p_1, \dots, p_m, \neg p_{m+1}, \dots, \neg p_n]; \vartheta\rangle$$

by the rule

$$\langle p \leftarrow_i @_I[p_1, \dots, p_m]; \vartheta\rangle$$

where the operator $\dot{@}_I \colon L^m \to L$ is defined as

$$\dot{@}_I[\vartheta_1, \dots, \vartheta_m] = \dot{@}[\vartheta_1, \dots, \vartheta_m, \dot{\neg}\, I(p_{m+1}), \dots, \dot{\neg}\, I(p_n)]$$

for all $\vartheta_1, \dots, \vartheta_m \in L$. The new obtained program will be called the *reduct* of $\mathbb{P}$ with respect to the interpretation $I$ and it will be denoted as $\mathbb{P}_I$.

In what follows, the notion of stable model of a multi-adjoint normal logic program is introduced.

**Definition 7.** *Given a multi-adjoint normal logic program $\mathbb{P}$ and a fuzzy $L$-interpretation $I$, we say that $I$ is a* stable model *of $\mathbb{P}$ if and only if $I$ is a minimal model of $\mathbb{P}_I$.*

The next result shows one of the most important features of stable models.

**Proposition 1.** *Any stable model of $\mathbb{P}$ is a minimal fix point of $T_{\mathbb{P}}$.*

Since the immediate consequence operator is not necessarily monotonic, the counterpart of proposition above is not satisfied in general.

A detailed survey about the existence of stable models for normal residuated logic programs defined on $[0, 1]$ was carried out in [9]. Later, this study was generalized in [1], where the existence of stable models was proven for normal residuated logic programs defined on any convex compact set of an euclidean space. This section will only focus on the survey of the existence of stable models in the multi-adjoint logic programming framework.

First and foremost, the convexity and the compactness of the set of fuzzy $K$-interpretations of a multi-adjoint normal logic program defined on a lattice with convex compact support is guaranteed.

**Proposition 2.** *Let $\mathbb{P}$ be a multi-adjoint normal logic program defined on a complete lattice $(K, \preceq, \leftarrow_1, \&_1, \dots, \leftarrow_n, \&_n, \neg)$ where $K$ is a convex (compact, respectively) set in an euclidean space $X$. Then the set of fuzzy $K$-interpretations of $\mathbb{P}$ is a convex (compact, respectively) set in the set of mappings defined on $X$.*

Before introducing the existence theorem, we will expose the idea the proof is based on. Given a fuzzy $K$-interpretation $I$, if we demonstrate that the mapping which associates each $I$ with the least fix point of $T_{\mathbb{P}_I}$, that is $R(I) = \text{lfp}(T_{\mathbb{P}_I})$, is continuous then we can apply the Schauder's Fixed Point Theorem [12]. This fact allows us to ensure that there exists $I$ being the least fix point of $T_{\mathbb{P}_I}$. Moreover, since $\mathbb{P}_I$ is a multi-adjoint logic program without any negation, this fix point is actually the minimal model of $\mathbb{P}_I$. Consequently, we obtain that $I$ is a stable model of $\mathbb{P}$.

**Theorem 1.** *Let $(K, \preceq, \leftarrow_1, \&_1, \dots, \leftarrow_n, \&_n, \neg)$ be a multi-adjoint lattice where $K$ is a non-empty convex compact set in an euclidean space, and a finite multi-adjoint normal logic program $\mathbb{P}$ defined on this lattice. If $\&_1, \dots, \&_n$, $\neg$ and the aggregator operators in the body of the rules of $\mathbb{P}$ are continuous operators, then $\mathbb{P}$ has at least a stable model.*

Finally, an illustrative example is shown in order to clarify the existence theorem presented above.

*Example 1.* Let $(X, \oplus, \otimes, \mathbb{R})$ be an euclidean space where $X$ is a space of triangular functions defined as

$$f_n(z) = \begin{cases} 10(z-n)+1 & \text{if} \quad n-0.1 \leq z \leq n \\ 10(n-z)+1 & \text{if} \quad n \leq z < n+0.1 \\ 0 & \text{otherwise} \end{cases}$$

and $\oplus, \otimes$ are mappings from $X$ to $X$ defined as $f_n \oplus f_m = f_{n+m}$ and $k \otimes f_n = f_{k \cdot n}$, respectively, where $n, m, k \in \mathbb{R}$.

We will see that $K = \{f_x \mid x \in [0,1]\}$ together with the ordering relation $f_n \leq f_m$ if and only if $n \leq m$, for all $n, m \in \mathbb{R}$, is a convex and compact set.

With respect to the convexity, we have to prove that $t \otimes f_x \oplus (1-t) \otimes f_y \in K$, for all $f_x, f_y \in K$ and $t \in [0,1]$. This can be easily proven considering that $t \cdot x + (1-t) \cdot y \in [0,1]$, for all $x, y, t \in [0,1]$.

On the other hand, taking into account that $\{f_0, f_1\} \subset K$ and $f_x \leq f_1$, for all $f_x \in K$, we have that $K$ is a bounded and closed set. Then, the compactness is straightforwardly obtained.

Once we have seen that $K$ is a convex compact set in $X$, applying Theorem 1, we can guarantee that every multi-adjoint normal logic program $\mathbb{P}$ defined on the multi-adjoint lattice $(K, \leq, \leftarrow_1, \&_1, \dots, \leftarrow_n, \&_n, \neg)$, where the conjunctors and the negation are continuous operators, has at least a stable model.

## 5 Conclusions and future work

Our contribution has consisted in applying the philosophy of the multi-adjoint paradigm in the syntax and the semantics of normal logic programs. Specifically, we have provided an original procedure to build reducts from multi-adjoint normal logic programs. In addition, we have extended the theorem associated with the existence of stable models given in [1] to the multi-adjoint case.

As a future work, we will develop the mathematical theory corresponding to the study of the unicity of stable models for multi-adjoint normal logic programs in a more general domain, such as the set of subintervals $\mathcal{C}([0,1])$.

## References


1. M. E. Cornejo, D. Lobo, and J. Medina. Stable models in normal residuated logic programs. In J. Kacprzyk, L. Koczy, and J. Medina, editors, *7th European Symposium on Computational Intelligence and Mathematices (ESCIM 2015)*, pages 150–155, 2015.
2. C. V. Damásio and L. M. Pereira. Monotonic and residuated logic programs. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'01*, pages 748–759. Lecture Notes in Artificial Intelligence, 2143, 2001.
3. M. Fitting. The family of stable models. *The Journal of Logic Programming*, pages 17(2–4):197–225, 1993.
4. P. Julian, G. Moreno, and J. Penabad. On fuzzy unfolding: A multi-adjoint approach. *Fuzzy Sets and Systems*, 154(1):16–33, 2005.
5. P. Julian, G. Moreno, and J. Penabad. An improved reductant calculus using fuzzy partial evaluation techniques. *Fuzzy Sets and Systems*, 160:162–181, 2009.
6. Y. Loyer and U. Straccia. Epistemic foundation of stable model semantics. *Journal of Theory and Practice of Logic Programming*, pages 6:355–393, 2006.
7. N. Madrid and M. Ojeda-Aciego. Towards a fuzzy answer set semantics for residuated logic programs. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '08)*, pages 260–264, 2008.
8. N. Madrid and M. Ojeda-Aciego. On coherence and consistence in fuzzy answer set semantics for residuated logic programs. In *Fuzzy Logic and Applications, 8th International Workshop (WILF 2009)*, pages 60–67, 2009.
9. N. Madrid and M. Ojeda-Aciego. On the existence and unicity of stable models in normal residuated logic programs. *International Journal of Computer Mathematics*, 89(3):310–324, 2012.
10. J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.
11. 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.
12. C. Minazzo and K. Rider. Théorèmes du point fixe et applications aux equations différentielles. *Université de Nice-Sophia Antipolisl.*
13. T. Przymusinski. Well-founded semantics coincides with three-valued stable semantics. *Fundamenta Informaticae*, 13:445–463, 1990.
14. U. Straccia. Query answering in normal logic programs under uncertainty. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 8th European Conference, ECSQARU'05*, pages 687–700, 2005.
15. U. Straccia. Query answering under the any-world assumption for normal logic programs. In *Proceedings, 10th International Conference on Principles of Knowledge Representation and Reasoning*, pages 329–339, 2006.
16. H. Strass, S. Muï¿½oz-Hernández, and V. Pablos-Ceruelo. Operational semantics for a fuzzy logic programming system with defaults and constructive answers. In *IFSA-EUSFLAT 2009*, pages 1827–1832, 2009.
17. P. Vojtáš. Fuzzy logic programming. *Fuzzy sets and systems*, 124(3):361–370, 2001.

# Bonds in a Fuzzy Environment

Manuel Ojeda-Aciego

Universidad de Málaga
aciego@uma.es

## Abstract

Formal Concept Analysis (FCA) has become a very active research topic, both theoretical and practical; its wide applicability justifies the need of a deeper knowledge of its underlying mechanisms, and one important way to obtain this extra knowledge turns out to be via generalization.

Several fuzzy variants of generalized FCA have been introduced and developed both from the theoretical and the practical side. Most of the generalizations focus on including extra features (fuzzy, possibilistic, rough, etc.); however, not much have been published on the suitable general version of certain specific notions, such as the bonds between formal contexts.

One of the motivations for introducing the notion of bond was to provide a tool for studying mappings between formal contexts, somehow mimicking the behavior of Galois connections between their corresponding concept lattices. In this talk we will deal with generalizations of the notion of bond in an L-fuzzy setting.

# Comparison of Krill Herd Algorithm and Flower Pollination Algorithm in Clustering Task

Piotr A. Kowalski[1,2], Szymon Łukasik[1,2],
Małgorzata Charytanowicz[2,3], Piotr Kulczycki[1,2]

[1] Faculty of Physics and Applied Computer Science
AGH University of Science and Technology
al. Mickiewicza 30, 30-059 Cracow, Poland
{pkowal,slukasik,mchat,kulczycki}@agh.edu.pl
[2] Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, PL-01-447 Warsaw, Poland,
{pakowal,slukasik,kulczycki}@ibspan.waw.pl
[3] Institute of Mathematics and Computer Science
The John Paul II Catholic University of Lublin
Konstantynów 1 H, 20-708 Lublin, Poland
mchmat@kul.lublin.pl

**Abstract.** The Krill Herd Algorithm (KHA) and Flower Pollination Algorithm (FPA) are modern heuristic techniques that are applicable for the purposes of deriving best solution within several optimization tasks. This paper is a comparison with regard to quality of result, of utilizing these metaheurisctics procedures for the data clustering task when applied to twelve data sets drawn from the UCI Machine Learning Repository. Here, the Celinski-Harabsz Index served to validate the cluster division criteria. Moreover, for broader comparison, the well-known and commonly employed k-means method was applied. Via the Rand Index, the quality of the results were assessed. The notion that KHA and FPA are effectively employed in such work is supported.

**Keywords:** clustering, krill herd algorithm, flower pollination algorithm, biologically inspired algorithm, optimization, metaheuristic.

## 1 Introduction

Exploratory Data Analysis mainly consists of addressing the issues of clustering, classification, data and dimension reduction, as well as outliers detection. The main goal of the clusterization procedure is to break up the confederate data collection into smaller subsets called 'clusters'. This action is unsupervised, being achieved by way of information directly derived from the data set itself. This division algorithm has been successfully applied within a wide variety of situations, including, in particular, those of engineering [6], but also in control tasks [9], economics [7] or even within the environmental sciences [4].

The division of data is carried out in such a way as to minimize or maximize the adapted to quality measure called the 'clustering index'. This measure is most frequently based on simple statistical relationships, such as average distance or variance, etc. The prime intent is that of assigning individual elements of the investigated data set, to particular clusters. Therefore, it can be concluded that the defined task is basically an NP-hard combinatorial optimization problem. Thus, metaheuristic procedures can be naturally applied for optimization purposes.

Optimization problems are encountered when deriving solutions to many engineering issues. The optimization task may be considered as that of choosing the best possible use of limited resources (time, money, etc.), while attempting to bring about certain goals. In achieving this, the optimization problem can be notated in a formal way. Let us introduce the 'cost function' (in some instances, this function can be also called the 'fitness function') $K$:

$$K : A \longmapsto \mathbb{R}, \tag{1}$$

(in some instances, this function can be also called the 'fitness function') where $A \subset \mathbb{R}^n$. The optimization task consists of finding the value $x^* \in A$, such that for every $x \in A$, the following relationship is true:

$$K(x) \geq K(x^*). \tag{2}$$

Although the optimization problem can be easily defined and described, determining its solution is already a very difficult issue. To resolve this problem, certain optimization algorithms are commonly used.

The presented research compares the quality of results achieved through applying these two Nature-derived clustering algorithms, with that gained through utilizing the k-means [11] procedure. The biological inspired group of algorithm is employed so as to find the best position of the cluster centre, and then the particular element of the investigated data set are assigned to a particular group. In the presented approach, the Krill Heard Algorithm [5] and the Flower Pollination Algorithm [12], both being continuous optimisation procedures, are employed. For assigning the particular solution in each iteration of the optimisation task, the Celinski-Harabasz Clustering Index [3] is subsequently used. For global comparative purposes, final results based on all considered clustering procedures are measured by way of the Rand Index [1].

The paper is organised as follow. In next section some general information concerning selected metaheuristic procedures KHA and FPA are being covered. In Section III reader can find a short information about clustering approach with Celinski-Harabasz validity criteria. Comparison results of numerical verification with discussion is contained in Section IV. Finally, conclusions are provided regarding the proposed clustering methods.

## 2 Optimization metaheurisctics

The KHA is a global optimization procedure inspired by the natural behavioural activity of the Antarctic Krill Swarm. This procedure was introduced by Gandomi and Alavi, in the paper [5]. A characteristic feature of krill, and the inspiration of the present algorithm [5, 6], is the ability for the activity of individual krill to be modelled within that of a large herd that is even hundreds of meters in size.

The operation of the algorithm begins via the initialization of data structures, i.e. describing individuals, as well as the whole population. Initializing the data structure representing a single krill, means situating it in a certain place (at the 'solution space') by giving it a set of coordinates. For this purpose, it is recommended to employ random number generation according to a uniform distribution. Like other algorithms inspired by Nature, each individual represents one possible solution of the problem under consideration. After the initialization phase of the algorithm, it continues into a series of iterations. The first step of each is to calculate a fitness function $K$ value for each individual of the population. This is equivalent to the calculation of optimized functions by way of arguments which are the coordinates of the krill's position. Then, for each individual of the population, the vector which indicates its displacement in the solution space is calculated. The movement of the krill individual is described by an equation dependent on three factors:

$$\frac{dX_i}{dt} = N_i + F_i + D_i, \tag{3}$$

where $N_i$ refers to the movement induced by the influence of other krills, $F_i$ is a movement associated with the search for food, and $D_i$ constitutes a random diffusion factor. In this algorithm, evolutionary operators such as mutation and crossover are also utilized.

The FPA is also a global optimization procedure, however, this one is inspired by the nature of the flower pollination process. This algorithm was introduced by Yang, in 2012 [12]. This procedure begins in a similar way to that of KHA. The main difference is in the main loop. Here, some random value is generated, and depending on this value, either global pollination or local pollination is carried out. The first process is inspired by the movement of insects flying over long distances so as to achieve the act of distant pollination. This step of the heuristic algorithm corresponds to the so-called 'exploration of the solution space'. Herein, the Levy flight distribution is employed in the mathematical realisation of this behaviour. The second process is inspired by a different, local pollination process that of 'self-pollination'. Here, the procedure of local search of solution is connected with the exploitation of the solution space that is under consideration. To perform either the global or local search process, a parameter for control of switching is implemented [10].

As a result of applying the KHA or the FPA, the best value of cost function K and the argument for which it was calculated, are achieved.

## 3 Clusterization Procedure

Let us assume that $Y$ is a data set matrix with dimensions $D$ and $M$, respectively. Each data collection element is represented by one column of $Y$. The main goal of the clustering procedure is to divide the data set and assign particular elements of $Y$ to the distinguished clusters $CL_1, \ldots, CL_C$,. In this procedure each cluster is characterised by a point called as the 'centre of cluster', which is calculated as:

$$O_c = \frac{1}{\#CL_c} \sum_{x_i \in CL_c} y_i, \tag{4}$$

where $\#CL_c$ denotes the number of elements assigned to the $c$th cluster. Similarly, the centre of gravity for all the investigated elements $y_1, \ldots, y_M$ can be defined as:

$$O_Y = \frac{1}{M} \sum_{i=1}^{M} y_i. \tag{5}$$

In this approach, the optimisation task in the clusterization procedure is based on the metaheuristic algorithms KHA and FPA. Herein, each element of the optimisation task is encoded as a collection of cluster centroids. Therefore, the product value $D \cdot C$ expresses the dimensionality of a particular optimization task. In this case, the number of the clusters is established at the beginning of the grouping process. Moreover, the assignment of individual elements to particular cluster is made on basis of the rule of the newest centroid point. Thus, for each point of $Y$, the distances to all cluster- centroids is calculated. In addition, the investigated point $y_i$ belongs to the cluster $CL_c$ if the Euclidean distance $dist(y_i, O_c)$ is the smallest. In this work, the Celinski-Harabasz Index $[3, 2]$ is applied as a criterion for assessing the quality of the data set division. This clustering index is implemented within the optimisation metaheuristic algorithm as a cost function. The Celinski-Harabasz criterion has its foundation within the concept of data set variance. This index is defined as:

$$I_{CH} = \frac{V_B}{V_W} \frac{M - C}{C - 1}, \tag{6}$$

where $V_B$ and $V_W$ denote overall between-cluster and within-cluster variance respectively. These are calculated according to the following formulas:

$$V_B = \sum_{c=1}^{C} \#CL_c \|O_c - O_Y\|^2, \tag{7}$$

and

$$V_W = \sum_{c=1}^{C} \sum_{y_i \in CL_c} \|y_i - O_c\|^2, \tag{8}$$

here, $\| \cdot \|$ is the $L^2$ norm (Euclidean distance) between the two vectors.

It should be underlined that high values of Celinski-Harabasz Index results point towards well-defined partitions. Because of the aforementioned properties, the following forms of cost functions are formulated:

$$K_{CH} = \frac{1}{I_{CH}} + \#CL_{\text{empty}}, \tag{9}$$

where $\#CL_{\text{empty}}$ denotes the number of clusters without any assigned element. More information about this index can be found at [3].

## 4 Numerical studies

In order to obtain numerical results regarding quality of the verification tasks, twelve sets of data obtained from the UCI Machine Learning Repository were taken into consideration [8]. In Table 1, results in the form of Rand Index mean value and resulting standard deviation are presented. In the first part of the table, results based on the well-known kmeans clustering procedure are revealed, in subsequent parts, results based on KHA-clustering and FPA-clustering, respectively, are shown. The best achieved results are bolded. All presented investigations were repeated 30 times.

**Table 1.** Achieved results of comparison

| | k-means clustering | | KHA-clustering | | FPA-clustering | |
|---|---|---|---|---|---|---|
| Data set | $\overline{R}$ | $\sigma_R$ | $\overline{R_{KHA}}$ | $\sigma_{R_{KHA}}$ | $\overline{R_{FPA}}$ | $\sigma_{R_{FPA}}$ |
| S1 | 0.9748 | 0.0093 | 0.9782 | 0.0078 | **0.9950** | 0.0018 |
| S2 | 0.9760 | 0.0072 | **0.9839** | 0.0053 | 0.9837 | 0.0037 |
| S3 | 0.9522 | 0.0072 | 0.9548 | 0.0053 | **0.9583** | 0.0026 |
| S4 | 0.9454 | 0.0056 | 0.9484 | 0.0048 | **0.9487** | 0.0023 |
| Iris | 0.8458 | 0.0614 | 0.8872 | 0.0145 | **0.8931** | 0.0000 |
| Ionosphere | 0.5945 | 0.0004 | 0.5573 | 0.0124 | **0.5946** | 0.0000 |
| Seeds | 0.8573 | 0.0572 | 0.8709 | 0.0156 | **0.8839** | 0.0000 |
| Sonar | 0.5116 | 0.0016 | **0.5145** | 0.0078 | 0.5128 | 0.0000 |
| Vehicle | 0.5843 | 0.0359 | 0.6076 | 0.0194 | **0.6101** | 0.0006 |
| WBC | 0.5448 | 0.0040 | **0.5456** | 0.0000 | **0.5456** | 0.0000 |
| Wine | 0.7167 | 0.0135 | 0.7257 | 0.0073 | **0.7299** | 0.0000 |
| Thyroid | **0.5844** | 0.0982 | 0.4535 | 0.0339 | 0.5128 | 0.0000 |

Based on the results reported in Table 1, it can be noted that it is only in the case of the Thyroid data collection that the application of the classic k-means procedure achieved the generation of better results than did the utilization of the KHA-clustering and FPA-clustering algorithms. In other applications, in a comparison between these metaheuristic methods, the FPA-clustering won 8 times, the KHA won 2 times and in the clustering of the WBC data set, the same results were received. A quite interesting observation is that regarding the stability of the obtained results. In the case of the k-means algorithm, the standard deviation of the results is several times higher than that of other methods. Especially worth emphasizing, is the pronounced negligible standard deviation of results based on FPA-clustering for 7 data set cases. It, thus, can be mainly concluded from the aforementioned results, that, unequivocally, both heuristic methods generate better solutions to the problem of clustering than does the classic k-means method.

## 5 Conclusions

This paper was a presentation of a comparison in quality terms, of the two metaheuristic algorithms - KHA and FPA - when applied for optimisation purposes within a data clustering problem. In the numerical verification, twelve data sets, taken from the UCI repository, were used for comparison purposes. The location of the cluster centre then was investigated by way of optimisation procedures. Here, the Celinski-Harabasz Index was first used to determine the quality in the heuristic solution. For comparative purposes, the well-known k-means procedure was subsequently performed. Based upon the obtained results, one can note that in almost all the data set cases, the metaheuristic algorithms demonstrated greater quality of result when compared with that gained by way of the k-means procedure. In particular, the FPA procedure revealed a much greater stability and greater quality.

## References


1. E. Achtert, S. Goldhofer, H. P. Kriegel, E. Schubert, and A. Zimek. Evaluation of clusterings – metrics and visual support. In *2012 IEEE 28th International Conference on Data Engineering*, pages 1285–1288, April 2012.
2. Olatz Arbelaitz, Ibai Gurrutxaga, Javier Muguerza, Jesús M Pérez, and Inigo Perona. An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1):243–256, 2013.
3. Tadeusz Calinski and Jerzy Harabasz. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1):1–27, 1974.
4. Małgorzata Charytanowicz, Jerzy Niewczas, Piotr Kulczycki, Piotr A. Kowalski, Szymon Łukasik, and Sławomir Żak. Complete gradient clustering algorithm for features analysis of x-ray images. In Ewa Pietka and Jacek Kawa, editors, *Information Technologies in Biomedicine*, volume 69 of *Advances in Intelligent and Soft Computing*, pages 15–24. Springer Berlin Heidelberg, 2010.
5. Amir Hossein Gandomi and Amir Hossein Alavi. Krill herd: A new bio-inspired optimization algorithm. *Communications in Nonlinear Science and Numerical Simulation*, 17(12):4831–4845, 2012.
6. Piotr A. Kowalski and Szymon Łukasik. Experimental study of selected parameters of the krill herd algorithm. In *Intelligent Systems'2014*, pages 473–485. Springer Science Business Media, 2015.
7. Piotr Kulczycki, Malgorzata Charytanowicz, Piotr A Kowalski, and Szymon Lukasik. The complete gradient clustering algorithm: properties in practical applications. 2012.
8. M. Lichman. UCI machine learning repository, 2013.
9. S. Łukasik, P.A. Kowalski, M. Charytanowicz, and P. Kulczycki. Fuzzy models synthesis with kernel-density-based clustering algorithm. In *Fuzzy Systems and Knowledge Discovery, 2008. FSKD '08. Fifth International Conference on*, volume 3, pages 449–453, Oct 2008.
10. Szymon Łukasik and Piotr A. Kowalski. Study of flower pollination algorithm for continuous optimization. In *Intelligent Systems'2014*, pages 451–459. Springer Science Business Media, 2015.
11. Lior Rokach and Oded Maimon. Clustering methods. In Oded Maimon and Lior Rokach, editors, *Data Mining and Knowledge Discovery Handbook*, pages 321–352. Springer US, 2005.
12. X.-S. Yang. Flower pollination algorithm for global optimization. *Lecture Notes in Computer Science*, 7445:240–249, 2012.

# On the Lower Limit of Possibilistic Correlation Coefficient for Identical Marginal Possibility Distributions

István Á. Harmati[1] and Robert Fullér[2,3]

[1] Department of Mathematics and Computational Sciences,
Széchenyi István University, Győr, Hungary
harmati@sze.hu
[2] Department of Informatics
Széchenyi István University, Győr, Hungary
[3] Institute of Applied Mathematics
John von Neumann Faculty of Informatics
Óbuda University, Budapest, Hungary
robert.fuller@nik.uni-obuda.hu

**Abstract.** In their recent paper Fullér et al. [1] defined a new measure of interactivity between fuzzy numbers, the so-called $f$-weighted possibilistic correlation coefficient, which can be determined from the joint possibility distribution. They also left two open questions related to the lower limit of the f-weighted possibilistic correlation coefficient of marginal possibility distribution with the same membership function.
In this paper we answer the more general version of the questions, for a large class of fuzzy and quasi fuzzy numbers.

**Keywords:** possibility theory, possibility distribution, correlation coefficient, possibilistic correlation, fuzzy numbers, quasi fuzzy numbers.

## 1 Introduction

In probability theory the expected value of functions of random variables plays a fundamental role in defining the basic characteristic measures of probability distributions. For example, the variance, covariance and correlation of random variables can be computed as the expected value of their appropriately chosen real-valued functions. In probability theory we can use the principle of *expected value* of functions on fuzzy sets to define variance, covariance and correlation of possibility distributions, while in possibility theory we can use the principle of *average value* of appropriately chosen real-valued functions to define mean value, variance, covariance and correlation of possibility distributions.

In [1] the authors proposed a new measure on interactivity between fuzzy numbers. This possibilistic correlation coefficient is based on a set of probabilistic correlation coefficients defined on the $\gamma$-level sets of the joint possibilitiy distribution. Namely, each level set of a (two dimensional) possibility distribution is equipped with a uniform probability distribution, then the correlation coefficient between random variables whose joint distribution is uniform on the $\gamma$-level set is computed in the standard probabilistic way. Finally, integration of these probabilistic correlation coefficients over the set of all membership grades [2, 4] with appropriate weights gives the possibilistic correlation coefficients.

At the end of their paper [1], the authors proposed two open questions:

1. Can one define a joint possibility distribution $C$ with the non-symmetrical but identical marginal distributions $A(x) = B(x) = 1-x$ for all $x \in [0, 1]$ and a weighting function $f$ for which the $f$-weighted index of interactivity (possibilistic correlation coefficient) could go below the value of $-3/5$?
2. What is the lower limit for $f$-weighted possibilistic correlation coefficient $\rho_f(A, B)$ between non-symmetrical marginal possibility distributions with the same membership function?

Harmati [3] answered question 1, and question 2 for the case when $A(x) = B(x) = 1 - x$. Later, Hong [5] gave another proof for the same case.

We note here that these are non-trivial questions. For example, in probability theory if the marginal distributions are Gaussians, then their correlations coefficient can be anything between $-1$ and 1, while if the marginal distributions are exponentials with parameter $\lambda = 1$, then the lower limit of their correlation coefficient is $1 - \pi^2/6$.

In the remaining part of the paper we show that the lower limit of the possibilistic correlation coefficient is $-1$, if the marginal distribution $A$ and $B$ have the same membership functions which are strictly increasing or stictly decreasing. Finally we extend the result to quasi fuzzy numbers, too.

## 2 Preliminaries

**Definition 1.** *A fuzzy number $A$ is a fuzzy set of $\mathbb{R}$ with a normal, fuzzy convex and continuous membership function of bounded support.*

Fuzzy numbers can be considered as possibility distributions.

**Definition 2.** *A fuzzy set $C$ in $\mathbb{R}^2$ is said to be a joint possibility distribution of fuzzy numbers $A, B$, if it satisfies the relationships*

$$\max\{x \mid C(x, y)\} = B(y), \qquad \textit{and} \qquad \max\{y \mid C(x, y)\} = A(x), \tag{1}$$

*for all $x, y \in \mathbb{R}$. Furthermore, $A$ and $B$ are called the marginal possibility distributions of $C$.*

**Definition 3.** *A $\gamma$-level set (or $\gamma$-cut) of a possibility distribution $C$ is a non-fuzzy set denoted by $[C]^\gamma$ and defined by*

$$[C]^\gamma = \begin{cases} \{(x, y) \in \mathbb{R}^2 \mid C(x, y) \geq \gamma\} & \text{if } \gamma > 0 \\ \text{cl}(\text{supp}C) & \text{if } \gamma = 0 \end{cases} \tag{2}$$

*where $\text{cl}(\text{supp}C)$ denotes the closure of the support of $C$.*

Fullér, Mezei and Várlaki introduced a new definition of possibilistic correlation coefficient [1] between marginal distributions of the joint possibility distribution that improves the earlier definition introduced by Carlsson, Fullér and Majlender [2].

**Definition 4 (see [1]).** *Let $f\colon [0,1] \to \mathbb{R}$ a non-negative, monotone increasing function with the normalization property $\int_0^1 f(\gamma)d\gamma = 1$. The $f$-weighted possibilistic correlation coefficient of fuzzy numbers $A$ and $B$ (with respect to their joint distribution $C$) is defined by*

$$\rho_f(A,B) = \int_0^1 \rho(X_\gamma, Y_\gamma) f(\gamma)\mathrm{d}\gamma, \tag{3}$$

*where*

$$\rho(X_\gamma, Y_\gamma) = \frac{\mathrm{cov}(X_\gamma, Y_\gamma)}{\sqrt{\mathrm{var}(X_\gamma)}\sqrt{\mathrm{var}(Y_\gamma)}},$$

*and, where $X_\gamma$ and $Y_\gamma$ are random variables whose joint distribution is uniform on $[C]^\gamma$ for all $\gamma \in [0,1]$, and $\mathrm{cov}(X_\gamma, Y_\gamma)$ denotes their probabilistic covariance.*

In other words, the $f$-weighted possibilistic correlation coefficient is the $f$-weighted average of the probabilistic correlation coefficients $\rho(X_\gamma, Y_\gamma)$ for all $\gamma \in [0,1]$.

## 3 Results for fuzzy numbers

In this section we follow the construction given in [3]. The fuzzy numbers $A$ and $B$ have the same membership function $F(x)$ with the following properties:

1. $F(0) = 1$ and $F(1) = 0$ and
2. $F(x)$ is strictly monotone decreasing in $[0,1]$.

We note here that by scaling and shifting the membership function, we can always let the support of the fuzzy numbers be the $[0,1]$ interval, and the correlation coefficient is invariant under scaling and shifting.

Let have $0 < c \leq 1$ and $k \geq 1$ and consider the following sets:

$$H_1 = \left\{(x,y) \in \mathbb{R}^2 |\ 0 \leq x, y \leq 1,\ y \leq cx^k\right\}\ , \tag{4}$$

$$H_2 = \left\{(x,y) \in \mathbb{R}^2 |\ 0 \leq x, y \leq 1,\ y \geq \sqrt[k]{\frac{x}{c}}\right\} \tag{5}$$

We define the two-dimensional joint possibility distribution as the following:

$$C(x,y) = \begin{cases} F(x)\,, & \text{if } (x,y) \in H_1 \\ F(y)\,, & \text{if } (x,y) \in H_2 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

The possibilistic correlation coefficient is the weighted average of probabilistic correlation coefficients of $X_\gamma$ and $Y_\gamma$, where $X_\gamma$ and $Y_\gamma$ are random variables whose joint distributions is uniform on the $\gamma$-cut of the joint possibility distribution. The $\gamma$-level set is $[C]^\gamma = H_1^\gamma \cup H_2^\gamma$, where (with $\delta = F^{-1}(\gamma)$):

$$H_1^\gamma = \left\{(x,y) \in \mathbb{R}^2 |\ 0 \le x, y \le \delta,\ y \le cx^k\right\}$$
$$H_2^\gamma = \left\{(x,y) \in \mathbb{R}^2 |\ 0 \le x, y \le \delta,\ y \ge \sqrt[k]{\frac{x}{c}}\right\} \tag{7}$$

If we do the similar computations as in [3], we get the correlation coefficient of $X_\gamma$ and $Y_\gamma$:

$$\rho(X_\gamma, Y_\gamma) = \frac{\frac{1}{4}c\delta^{k+1} - \left[\frac{k+1}{2k+4}\delta + c\frac{k+1}{8k+4}\delta^k\right]^2}{\frac{k+1}{2k+6}\delta^2 + c^2\frac{k+1}{18k+6}\delta^{2k} - \left[\frac{k+1}{2k+4}\delta + c\frac{k+1}{8k+4}\delta^k\right]^2} \tag{8}$$

If $k \to \infty$ and $c \to 0$ in the above equation then the limit is $-1$.

**Theorem 1.** *Let $C(x,y)$ be a joint possibility distribution with marginal possibility distributions $A(x) = B(x) = F(x)$ for all $x \in [0,1]$, where $F$ is a strictly monotone decreasing continuous function, $F(0) = 1$ and $F(1) = 0$. Let $X_\gamma$ and $Y_\gamma$ be random variables whose joint probability distribution is uniform on $[C]^\gamma$ for all $\gamma \in [0,1]$. Then for all $\gamma \in [0,1]$*

$$\inf_C \rho(X_\gamma, Y_\gamma) = -1 \tag{9}$$

*Proof.* The statement follows from the limit of Eq. (8).

**Theorem 2.** *Let $C(x,y)$ be a joint possibility distribution with marginal possibility distributions $A(x) = B(x) = F(x)$ for all $x \in [0,1]$, where $F$ is a strictly monotone increasing continuous function, $F(0) = 0$ and $F(1) = 1$. Let $X_\gamma$ and $Y_\gamma$ be random variables whose joint probability distribution is uniform on $[C]^\gamma$ for all $\gamma \in [0,1]$. Then for all $\gamma \in [0,1]$*

$$\inf_C \rho(X_\gamma, Y_\gamma) = -1 \tag{10}$$

*Proof.* This immediately follows if we rotate the previous joint possibility distribution by 180 degrees around the centre of the square $[0,1] \times [0,1]$.

From the previous theorems we get lower limit for the possibilistic correlation coefficient for identical “one-legged” marginal possibility distribution:

**Theorem 3.** *Let $C(x,y)$ be a joint possibility distribution with marginal possibility distributions $A(x) = B(x) = F(x)$ for all $x \in [0,1]$, where*

1. *$F$ is a strictly monotone decreasing continuous function, $F(0) = 1$ and $F(1) = 0$, or*
2. *$F$ is a strictly monotone increasing continuous function, $F(0) = 0$ and $F(1) = 1$.*

*Let $X_\gamma$ and $Y_\gamma$ be random variables whose joint probability distribution is uniform on $[C]^\gamma$ for all $\gamma \in [0,1]$, then*

$$\inf_C \rho_f(A,B) = \inf_C \int_0^1 \rho(X_\gamma, Y_\gamma) f(\gamma)\, \mathrm{d}\gamma = -1 \tag{11}$$

*Proof.* We have seen that we can define a family of joint possibility distribution with the properties above and with the property of $\rho(X_\gamma, Y_\gamma) \to -1$ for all $\gamma \in [0, 1]$. Since the possibilistic correlation coefficient is the $f$-weighted average of the probabilistic correlation coefficients, then the statement follows.

## 4 Result for quasi fuzzy numbers

The quasi fuzzy number is a generalization of the fuzzy number for unbounded support:

**Definition 5.** *A quasi fuzzy number $A$ is a fuzzy set of $\mathbb{R}$ with a normal, fuzzy convex and continuous membership function for which* $\lim_{x\to-\infty} A(x) = 0$ *and* $\lim_{x\to\infty} A(x) = 0$.

**Theorem 4.** *Let $C(x, y)$ be a joint possibility distribution with marginal possibility distributions $A(x) = B(x) = F(x)$ for all $x \in [0, 1]$, where*

1. *$F(x)$ is a strictly monotone decreasing continuous function if $0 \le x$, $F(0) = 1$ and $\lim\limits_{x\to\infty} F(x) = 0$ and $F(x)$ is zero if $x < 0$, or*
2. *$F(x)$ is a strictly monotone increasing continuous function if $x \le 0$, $F(0) = 1$ and $\lim\limits_{x\to-\infty} F(x) = 0$ and $F(x)$ is zero if $x > 0$.*

*Let $X_\gamma$ and $Y_\gamma$ be random variables whose joint probability distribution is uniform on $[C]^\gamma$ for all $\gamma \in [0, 1]$, then*

$$\inf_C \rho_f(A, B) = \inf_C \int_0^1 \rho(X_\gamma, Y_\gamma) f(\gamma)\,\mathrm{d}\gamma = -1 \tag{12}$$

*Proof.* We show a family of joint distributions, for which $\rho_f(A, B) \to -1$. We prove the theorem for the first case, the second behaves similarly. Let us define a joint possibility distribution by its $\gamma$-levels:

$$[C_q]^\gamma = \begin{cases} [\min(A(x), B(y))]^\gamma & \text{if } \gamma \le \gamma_0 \\ [C']^\gamma & \text{if } \gamma > \gamma_0 \end{cases} \tag{13}$$

where $[C']^\gamma$ is the scaled version of the $\gamma$-level set defined in Eq. (7), with the scaling factor of $\delta_0 = F^{-1}(\gamma_0)$, so $[C']^\gamma = {H'}_1^\gamma \cup {H'}_2^\gamma$, where (with $\delta = F^{-1}(\gamma)$):

$$\begin{aligned} {H'}_1^\gamma &= \left\{(x, y) \in \mathbb{R}^2 \mid 0 \le x, y \le \delta,\ y \le cx^k \delta_0^{k-1}\right\} \\ {H'}_2^\gamma &= \left\{(x, y) \in \mathbb{R}^2 \mid 0 \le x, y \le \delta,\ y \ge \sqrt[k]{\frac{x}{c}}\delta_0^{1-1/k}\right\} \end{aligned} \tag{14}$$

The $\min$ operator implies zero correlation [1] on $\gamma$-level sets for $\gamma \le \gamma_0$.

On the other hand, if $\gamma > \gamma_0$, then the $\gamma$ cuts have the same shape as in the previous section, so the correlation coefficient between $X_\gamma$ and $Y_\gamma$ goes to $-1$ if $c \to 0$ and $k \to \infty$. So by an appropriate weighting function we get

$$\inf_C \rho_f(A, B) = \inf_C \int_0^1 \rho(X_\gamma, Y_\gamma) f(\gamma)\,\mathrm{d}\gamma = -1$$

## 5 Summary

We answered the open question published in [1] for fuzzy numbers and quasi fuzzy numbers with "one-legged" membership function, i.e. the lower limit of the possibilistic correlation coefficient is $-1$ (the upper limit is 1, when the joint possibility distribution is concentrated along a line).

The result tells us that in this case from the marginal possibility distributions one cannot infer to the value of possibilitic correlation coefficient, so the shape of the marginal distribution does not mean any restrictions. The result can be used in fuzzy time-series models and fuzzy statistics when the theoretical joint distributions are not known.

The general question is still open, the case when $A$ and $B$ have the same non-symmetrical membership function, but their membership function is not strictly increasing neither decreasing on their whole support, so our future work is the generalization of the result to a larger class of fuzzy and quasi fuzzy numbers.

## References


1. R. Fullér, J. Mezei and P. Várlaki, An improved index of interactivity for fuzzy numbers, *Fuzzy Sets and Systems* 165, 50–60 (2011) doi:10.1016/j.fss.2010.06.001
2. C. Carlsson, R. Fullér and P. Majlender, On possibilistic correlation, *Fuzzy Sets and Systems*, 155, 425-445 (2005), doi: 10.1016/j.fss.2005.04.014
3. I. Á. Harmati, A note on $f$-weighted possibilistic correlation for identical marginal possibility distributions, *Fuzzy Sets and Systems*, 165, 106–110 (2011) doi: 10.1016/j.fss.2010.11.005
4. R. Fullér, P. Majlender, On interactive fuzzy numbers, *Fuzzy Sets and Systems*, 143, 355–369 (2004) doi: 10.1016/S0165-0114(03)00180-5
5. D. H. Hong, J. D. Kim, The Lower Limit for Possibilistic Correlation Coefficient. Applied Mathematical Sciences. Vol. 9, No.121, 6041–6047 (2015)

# Manipulating Positive and Negative Attributes in Implications

P. Cordero, M. Enciso, A. Mora, and J. M. Rodríguez-Jiménez

University of Málaga, Andalucía Tech, Spain
{pcordero,enciso}@uma.es,
{amora,jmrodriguez}@ctima.uma.es

**Abstract.** In several areas such as artificial intelligence, data mining, database theory, formal concept analysis, etc. the implications represent relationships between set of attributes. Usually, they only specify the presence of the attribute (positive) but not their absence (negative). In this work, we propose how to manipulate positive and negative attributes by means of rules, in a logic-style, for inferring new knowledge. A new logic for the treatment of implications with positive and negative attributes is introduced as the first step for the subsequent design of automated methods.

## 1 Introduction

In the area of knowledge discovering, one of the main issues is the mining of patterns among data stored in a data set. Different patterns have been defined, but connection between attribute set is one of the most widespread proposals. This pattern is approached by using the attribute implication but, in the classical case, it reveals only the presence of some attributes (positive attributes) in the data set. In this work, we consider not only the presence but also the absence of attributes, that is, positive and negative attributes.

Considering negative attributes could increase our knowledge and, taking into account this knowledge, some benefits that we can obtain in applications. Thus, the use of positive and negative attributes has been demanded in basket analysis and many efforts have been dedicated to discover this knowledge in a more efficient way. For instance, if a customer usually buys a set of products, but avoids getting other itemset, such information can be analyzed, for marketing purposes, to distribute products in markets or to predict whether or not a new product will be accepted.

More specifically, Formal Concept Analysis (FCA) explores binary data stored in a table with objects as rows and attributes as columns. In these tables the presence of an attribute in an object is marked, whereas the absence of this attribute is not taken into account.

There are only a few, but promising, works for the management of both positive and negative attributes in FCA.

In [3] the use of negative information is handled considering the relation of a learning model described in terms of FCA with a standard model of Machine Learning called version spaces. A version space is the set of all possible functions compatible with training data and in the above paper the authors use positive and negative information to generate a classifier.

In [4], the authors have approached this problem in a easy way by joining the data set, named formal context in this area, with its complementary so that the classical FCA methods and techniques can be used to obtain the concept lattice. Now, the corresponding implications consider positive and negative information because of the use of complementary attributes. However, the main problem is that dataset size has grown, increasing the complexity.

Another approach is to consider the connection among positive, negative and mixed information. In this way, Missaoui et.al. [5] have approached the generation of mixed implications from two given sets of implications only with positive and negative attributes respectively.

We have deepened theoretically in this line opened by Missaoui et al. In [9], we have extended the classical FCA framework with new derivation operators constituting a Galois connection. Moreover, we have proposed some mining algorithms to derive directly mixed implications following this line [7, 8].

In this work, we introduce a set of inference rules to manage mixed implications. As far as we know, to provide a sound and complete logic to manipulate mixed implications is an open problem in the literature.

The structure of the paper is the following: Section 2 shows some preliminaries. The new set of rules to manipulate implications with positive and negative attributes is presented in Section 3. Finally, some conclusions and future works appear in Section 4.

## 2 Preliminaries

In FCA, the input data are stored in a binary table named a *formal context* which formally is defined as the triple $\mathbb{K} = \langle G, M, I\rangle$ where $G$ and $M$ are finite non-empty sets and $I \subseteq G \times M$ is a binary relation.

The elements in $G$ are named objects, the elements in $M$ attributes and $\langle g, m\rangle \in I$ means that the object $g$ has the attribute $m$.

From this triple, two mappings ${}^{\uparrow} : 2^G \to 2^M$ and ${}^{\downarrow} : 2^M \to 2^G$, named derivation operators, are defined as follows: for any $X \subseteq G$ and $Y \subseteq M$,

$$X^{\uparrow} = \{m \in M \mid \langle g, m\rangle \in I \text{ for all } g \in X\} \tag{1}$$

$$Y^{\downarrow} = \{g \in G \mid \langle g, m\rangle \in I \text{ for all } m \in Y\} \tag{2}$$

$X^{\uparrow}$ is the subset of all attributes shared by all the objects in $X$ and $Y^{\downarrow}$ is the subset of all objects that have the attributes in $Y$. The pair $(\uparrow, \downarrow)$ constitutes a Galois connection between $2^G$ and $2^M$ and, therefore, both compositions are closure operators.

A pair of subsets $\langle X, Y\rangle$ with $X \subseteq G$ and $Y \subseteq M$ such that $X^{\uparrow} = Y$ and $Y^{\downarrow} = X$ is named a *formal concept* where $X$ is its *extent* and $Y$ its *intent*. These extents and intents coincide with closed sets wrt the closure operators because $X^{\uparrow\downarrow} = X$ and $Y^{\downarrow\uparrow} = Y$. Thus, the set of all formal concepts is a lattice, named *concept lattice*, with the relation

$$\langle X_1, Y_1\rangle \leq \langle X_2, Y_2\rangle \text{ if and only if } X_1 \subseteq X_2 \text{ (or equivalently, } Y_2 \subseteq Y_1) \tag{3}$$

This concept lattice will be denoted by $\mathfrak{B}(\mathbb{K})$.

The concept lattice can be characterized in terms of *attribute implications* being expressions $A \to B$ where $A, B \subseteq M$. An implication $A \to B$ holds in a context $\mathbb{K}$ if $A^{\downarrow} \subseteq B^{\downarrow}$. That is, any object that has all the attributes in $A$ has also all the attributes in $B$. These implications can be syntactically managed in a logical style [1]. The main aim of this paper is to provide a sound and complete logic to manage implications that also considers negative information.

## 3 MixAtL: a logic for mixed attribute implications

First, we introduce our extension of the notion of attribute implication to consider positive and negative information. This will be the basis for a further development of automated reasoning methods for this kind of implications and they also open the door to study the definition of canonical representations of implicational systems.

Now, we formally introduce the novel logic to manage mixed implications, that is, considering positive and negative attributes. To do this, we present: the language, the semantics and a sound and complete axiomatic system.

### 3.1 The language of $\mathbf{MixAtL}$

**Definition 1.** *Given a finite set of attributes $M$, the language of* $\mathbf{MixAtL}$ *is*

$$\mathcal{L}_M = \{X \to Y \mid X, Y \subseteq M \cup \overline{M}\}.$$

*Formulas in $\mathcal{L}_M$ are named* mixed attribute implications*.*

Hereinafter, the lowercase character $m$ (possibly with subindexes) is reserved for positive attributes. We use $\overline{m}$ to denote the negation of the attribute $m$ and $\overline{M}$ to denote the set $\{\overline{m} \mid m \in M\}$ whose elements will be named negative attributes. Arbitrary elements in $M \cup \overline{M}$ are going to be denoted by the first letters in the alphabet: $a$, $b$, $c$, etc. and $\overline{a}$ denotes the opposite of $a$, i.e. the symbol $a$ could represent a positive or a negative attribute and, if $a = m \in M$ then $\overline{a} = \overline{m}$ and if $a = \overline{m} \in \overline{M}$ then $\overline{a} = m$.

The subsets of $M \cup \overline{M}$ will be denoted by uppercase characters $A$, $B$, $C$, etc. and we will use the following notation: for each $A \subseteq M \cup \overline{M}$,

- $\overline{A}$ is the set of the opposite of attributes $\{\overline{a} \mid a \in A\}$
- $\mathrm{Pos}(A) = \{m \in M \mid m \in A\}$ and $\mathrm{Neg}(A) = \{m \in M \mid \overline{m} \in A\}$
- $\mathrm{Tot}(A) = \mathrm{Pos}(A) \cup \mathrm{Neg}(A)$

Note that $\mathrm{Pos}(A), \mathrm{Neg}(A), \mathrm{Tot}(A) \subseteq M$.

### 3.2 The semantics of $\mathbf{MixAtL}$

Once we have defined the language and we have set the notation to be used, we now introduce the semantics of $\mathbf{MixAtL}$. Classical derivation operators are extended as follows:

**Definition 2.** *Let $\mathbb{K} = \langle G, M, I \rangle$ be a formal context. We define the operators $^{\Uparrow}$ : $2^G \to 2^{M \cup \overline{M}}$ and $^{\Downarrow} : 2^{M \cup \overline{M}} \to 2^G$ as follows: for $X \subseteq G$ and $Y \subseteq M \cup \overline{M}$,*

$$X^{\Uparrow} = \{m \in M \mid \langle g, m \rangle \in I \text{ for all } g \in X\} \cup \{\overline{m} \in \overline{M} \mid \langle g, m \rangle \notin I \text{ for all } g \in X\}$$

$$Y^{\Downarrow} = \{g \in G \mid \langle g, m \rangle \in I \text{ for all } m \in Y\} \cap \{g \in G \mid \langle g, m \rangle \notin I \text{ for all } \overline{m} \in Y\}$$

It is remarkable that the extended derivation operators have similar properties as the classical ones.

**Theorem 1.** *For any formal context $\mathbb{K} = \langle G, M, I \rangle$, the pair $(^{\Uparrow}, ^{\Downarrow})$ is a Galois connection between $(2^G, \subseteq)$ and $(2^{M \cup \overline{M}}, \subseteq)$.*

As a consequence of the above theorem, similarly that occurs in the classical case, both compositions $^{\Uparrow} \circ ^{\Downarrow}$ and $^{\Downarrow} \circ ^{\Uparrow}$ are closure operators and lead to the notion of mixed concept lattice.

**Definition 3.** *Let $\mathbb{K} = \langle G, M, I \rangle$ be a formal context. A* mixed formal concept *in $\mathbb{K}$ is a pair of subsets $\langle X, Y \rangle$ with $X \subseteq G$ and $Y \subseteq M \cup \overline{M}$ such $X^{\Uparrow} = Y$ and $Y^{\Downarrow} = X$.*

Thus, if $\langle X, Y \rangle$ is a mixed concept, then $X^{\Uparrow\Downarrow} = X$ and $Y^{\Downarrow\Uparrow} = Y$. In addition, the set of all mixed formal concepts is a lattice, named *mixed concept lattice*, with the relation

$$\langle X_1, Y_1 \rangle \leq \langle X_2, Y_2 \rangle \text{ if and only if } X_1 \subseteq X_2 \text{ (or equivalently, } Y_2 \subseteq Y_1)$$

This mixed concept lattice will be denoted by $\mathfrak{B}^{\#}(\mathbb{K})$.

Now, we have all needed notions in order to provide a meaning for formulas in the language (implications).

**Definition 4.** *Let $\mathbb{K} = \langle G, M, I \rangle$ be a formal context and $A \to B \in \mathcal{L}_M$. The context $\mathbb{K}$ is a model for $A \to B$, denoted by $\mathbb{K} \models A \to B$, if $A^{\Downarrow} \subseteq B^{\Downarrow}$, or equivalently $B \subseteq A^{\Downarrow\Uparrow}$.*

*Example 1.* Considering the formal context $\mathbb{K} = \langle G, M, I \rangle$ where the set of objects is $G = \{o_1, o_2, o_3, o_4\}$, the set of attributes is $M = \{m_1, m_2, m_3, m_4, m_5\}$ and $I$ is the binary relation depicted in Table 1, we have that $\mathbb{K} \not\models m_2 \to m_4$ and $\mathbb{K} \models m_2 \to \overline{m}_4$ whereas $\mathbb{K} \not\models m_2 \to m_3$ either $\mathbb{K} \not\models m_2 \to \overline{m}_3$.

| I | $m_1$ | $m_2$ | $m_3$ | $m_4$ | $m_5$ |
|---|---|---|---|---|---|
| $o_1$ | | × | × | | × |
| $o_2$ | × | × | | | |
| $o_3$ | | × | × | | × |
| $o_4$ | | | × | × | |

**Table 1.** A formal context

As usual, given a set of mixed attribute implications $\Sigma \subseteq \mathcal{L}_M$ and a formal context $\mathbb{K}$, the expression $\mathbb{K} \models \Sigma$ denotes $\mathbb{K} \models A \to B$ for all $A \to B \in \Sigma$ and $\Sigma \models A \to B$ denotes that any model for $\Sigma$ is also model for $A \to B$.

### 3.3 An axiomatic system for $\mathbf{MixAtL}$

Now, we propose a set of inference rules to reason with mixed rules, that is, implications with positive and negative attributes.

Once we have the implicational system with negative attributes, we can explore all the posible implications with the set of inference rules built by supplementing Armstrong's axioms with the following ones.

The axiomatic system for $\mathbf{MixAtL}$ considers two axiom schemes and four inference rules. They are the following:

[Ref] Reflexivity: If $B \subseteq A$ then $\vdash A \to B$.
[Cont] Contradiction: $\vdash a\overline{a} \to M\overline{M}$.
[Augm] Augmentation: $A \to B \vdash A \cup C \to B \cup C$.
[Trans] Transitivity: $A \to B, B \to C \vdash A \to C$.
[Rft] Reflection: $Aa \to b \vdash A\overline{b} \to \overline{a}$.
[Tr] True: $a \to \overline{a} \vdash \varnothing \to \overline{a}$.

**Theorem 2.** *The* $\mathbf{MixAtL}$ *axiomatic system is sound and complete.*

## 4 Conclusions

We tackle the manipulation of positive and negative attributes in the framework of Formal Concept Analysis. Although it has been partially addressed in some works, in the area of Formal Concept Analysis this issue is still considered to be an open problem and only some works have tried to solve it.

We propose a set of rules constituting a sound and complete axiomatic system to manipulate implications having mixed attributes. It is the first stage for future developments of automated methods to reason with this kind of implications.

For achieving a further generalization of our proposal, it could be interesting to explore the work of D. Ciucci [2] where orthopairs, pairs of disjoint sets, are introduced as a useful tool to manage uncertainty and to collect positive and negative information. Orthopairs are related to recent works in FCA [6] where positive and negative operators are described to manipulate opposite information.

## Acknowledgments

Supported by project regional no. TIN2014-59471-P of the Science and Innovation Ministry of Spain, co-funded by the European Regional Development Fund (ERDF).

## References

1. Armstrong, W.: Dependency structures of data base relationships. Proc. IFIP Congress. North Holland, Amsterdam pp. 580–583 (1974)
2. Ciucci, D.: Orthopairs and granular computing. Granul. Comput. 1, 159 – 170 (2016)

3. Ganter, B., Kuznetsov, S.: Hypotheses and version spaces. Lecture Notes in Computer Science 2746 (2003)
4. Gasmi, G., Yahia, S.B., Nguifo, E.M., Bouker, S.: Extraction of association rules based on literalsets. LNCS 4654, 293–302 (2007)
5. Missaoui, R., Nourine, L., Renaud, Y.: Computing Implications with Negation from a Formal Context. Fundamenta Informaticae 115(4), 357–375 (dec 2012)
6. Qi, J., Qian, T., Wei, L.: The connections between three-way and classical concept lattices. Knowledge-Based Systems 91, 143 – 151 (2016)
7. Rodríguez-Jiménez, J., Cordero, P., Enciso, M., Mora, A.: A generalized framework to consider positive and negative attributes in formal concept analysis. In: CLA. CEUR Workshop Proceedings, vol. 1252, pp. 267–278. CEUR-WS.org (2014)
8. Rodríguez-Jiménez, J., Cordero, P., Enciso, M., Mora, A.: Data mining algorithms to compute mixed concepts with negative attributes: an application to breast cancer data analysis. Mathematical Methods in the Applied Sciences (2016)
9. Rodriguez-Jimenez, J., Cordero, P., Enciso, M., Rudolph, S.: Concept lattices with negative information: a characterization theorem. Information Sciences http://dx.doi.org/10.1016/j.ins.2016.06.015 (2016)

# Attribute Reduction in Fuzzy Formal Concept Analysis from Rough Set Theory *

M. José Benítez[1], Jesús Medina[2], and Dominik Ślęzak[2,3]

[1] Department of Mathematics, University of Cádiz. Spain
{mariajose.benitez, jesus.medina}@uca.es
[2] Institute of Mathematics, University of Warsaw. Poland
slezak@mimuw.edu.pl
[3] Infobright Inc., Canada / Poland
slezak@infobright.com

**Abstract.** Reducing the number of attributes by preventing the occurrence of incompatibilities and eliminating existing noise in the original data is an important goal in different frameworks as Rough Set Theory (RST) and Formal Concept Analysis (FCA). We have recently presented a novel reduction method in RST based on bireducts using similarity relations. This paper applies this method in the FCA framework since an illustrative example.

## 1 Introduction

Rough Set Theory (RST) and Formal Concept Analysis (FCA) are two fundamental mathematical tools for modelling and processing incomplete information in information systems, which can extract pieces of information from databases. One of the main goals in these frameworks is to reduce the unnecessary attributes in the database, preserving the information that can be extracted from the database.

Different types of mechanism have been studied in RST and FCA [5, 6, 9, 11]. One of the most recent ones in RST considers the use of bireducts, which generalizes the classical RST-based notions of reducts [10, 13, 14].

Recently, in [3, 4], we have extended the notions of reducts and bireducts introduced in rough set theory for attribute reduction based on similarity relations defined on attribute values. The characterizations of the new reducts and bireducts were given in terms of the corresponding generalizations of the discernibility function.

This paper shows the possibility of considering such attribute reduction in a Fuzzy Concept Lattice framework, for that, a worked out example is considered. This procedure will formally be detailed in an extended version and the comparison with other similar approaches in FCA, such as the ones given in [1, 2, 7].

* M. José Benítez and Jesús Medina are partially supported by the Spanish Economy and Competitiveness Ministry (MINECO) project TIN2016-76653-P. Dominik Ślęzak is partially supported by the Polish National Science Centre grants DEC-2012/05/B/ ST6/03215 and DEC-2013/09/B/ST6/01568.

## 2 Fuzzy formal concept analysis

Formal Concept Analysis (FCA) is a mathematical theory to extract information from databases, in which two different sets (the set of attributes $A$ and the set of objects $B$) and a relation $R$ between them are considered [8]. The triple $(A, B, R)$ is called *context* and the mappings[1] $^{\uparrow}\colon 2^B \to 2^A$, $^{\downarrow}\colon 2^A \to 2^B$, are defined, for each $X \subseteq B$ and $Y \subseteq A$, as follows:

$$X^{\uparrow} = \{a \in A \mid \text{for all } b \in X, aRb\} = \{a \in A \mid \text{if } b \in X,\ \text{then } aRb\} \tag{1}$$
$$Y^{\downarrow} = \{b \in B \mid \text{for all } a \in Y, aRb\} = \{b \in B \mid \text{if } a \in Y,\ \text{then } aRb\} \tag{2}$$

A *concept* in the context $(A, B, R)$ is defined to be a pair $(X, Y)$, where $X \subseteq B$, $Y \subseteq A$, and which satisfy $X^{\uparrow} = Y$ and $Y^{\downarrow} = X$. The element $X$ of the concept $(X, Y)$ is the *extent* and $Y$ the *intent*.

$\mathcal{B}(A, B, R)$ is the set of concepts in a context $(A, B, R)$, which is a complete lattice [8], with the inclusion order on the left argument or the opposite of the inclusion order on the right argument, that is, for each $(X_1, Y_1), (X_2, Y_2) \in \mathcal{B}(A, B, R)$, we have $(X_1, Y_1) \leq (X_2, Y_2)$ if $X_1 \subseteq X_2$ (or, equivalently, $Y_2 \subseteq Y_1$). The meet $\wedge$ and join $\vee$ operators are defined by:

$$(X_1, Y_1) \wedge (X_2, Y_2) = (X_1 \wedge X_2, (Y_1 \vee Y_2)^{\downarrow\uparrow})$$
$$(X_1, Y_1) \vee (X_2, Y_2) = ((X_1 \vee X_2)^{\uparrow\downarrow}, Y_1 \wedge Y_2)$$

for all $(X_1, Y_1), (X_2, Y_2) \in \mathcal{B}(A, B, R)$.

Attribute reduction theory in formal concept analysis pursues to reduce the set of attributes in order to not alter the number of concepts in the concept lattice, that is, obtaining a new concept lattice isomorphic to the original one. The definitions of finer, consistent set, reduct and core on concept lattices are also recalled.

**Definition 1.** *Given a context $(A, B, R)$, if there exists a set of attributes $Y \subseteq A$ such that $\mathcal{B}(A, B, R) \cong \mathcal{B}(Y, B, R_Y)$, then $Y$ is called a* consistent *set of $(A, B, R)$. Moreover, if $\mathcal{B}(Y \smallsetminus \{y\}, B, R_{Y \smallsetminus \{y\}}) \ncong \mathcal{B}(A, B, R)$, for all $y \in Y$, then $Y$ is called a* reduct *of $(A, B, R)$.*

*The* core *of $(A, B, R)$ is the intersection of all the reducts of $(A, B, R)$.*

## 3 $\delta$-decision bireducts

This section recalls the basic notion of information system, the definitions of $\delta$-information bireduct and bidimensional $\delta$-discernibility function, and the characterization theorem of the $\delta$-decision bireducts.

[1] Ganter and Wille used originally the notation $'$ for this operator, hence they were called derivation operators. We have changed the notation in order to differentiate between the mapping on the set of object and on the set of attributes.

**Definition 2.** *An* information system $(U, \mathcal{A})$ *is a tuple, where* $U = \{x_1, \dots, x_n\}$ *and* $\mathcal{A} = \{a_1, \dots, a_m\}$ *are finite, non-empty sets of objects and attributes, respectively. Each* $a$ *in* $\mathcal{A}$ *corresponds to a mapping* $f_a \colon U \to V_a$*, where* $V_a$ *is the value set of* $a$ *over* $U$*. For every subset* $B$ *of* $\mathcal{A}$*, the* $B$*-indiscernibility relation*[2] $I_B$ *is defined as the equivalence relation*

$$I_B = \{(x_i, x_j) \in U \times U \mid \text{for all } a \in B, f_a(x_i) = f_a(x_j)\}, \tag{3}$$

*where each class can be written as* $[x]_B = \{x_i \mid (x, x_i) \in I_B\}$*.* $I_B$ *produces a partition on* $U$ *denoted as* $U/I_B = \{[x]_B \mid x \in U\}$*.*

**Definition 3.** *The pair* $(B, X)$*, where* $B \subseteq \mathcal{A}$ *and* $X \subseteq U$*, is called* $\delta$-information bireduct *if and only if all pairs* $i, j$ *of* $X$ *are* $\delta$*-discordant by* $B$ *and the following properties hold:*

1. *There is no* $C \subsetneq B$ *such that all pairs* $i, j \in X$ *are* $\delta$*-discordant by* $C$*.*
2. *There is no* $X \subsetneq Y$ *such that all pairs* $i, j \in Y$ *are* $\delta$*-discordant by* $B$*.*

**Definition 4.** *Let* $\mathbb{A} = (U, A)$ *be an information system, the conjunctive normal form* $\tau^{bi}_{\mathcal{A}} = \bigwedge\{i \vee j \bigvee\{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \mid i, j \in U\}$*, where the elements of* $U$ *and* $\mathcal{A}$ *are the propositional symbols of the language, is called the* bidimensional $\delta$-discernibility function*.*

The following theorem characterizes the $\delta$-decision bireducts.

**Theorem 1.** *Given a decision system* $\mathbb{A} = (U, A \cup \{d\})$*, an arbitrary pair* $(B, X)$*,* $B \subseteq \mathcal{A}$*,* $X \subseteq U$*, is a* $\delta$*-decision bireduct if and only if the cube* $\bigwedge_{b \in B} b \wedge \bigwedge_{i \notin X} i$ *is a cube in the RDNF of* $\tau^{bir}_{\mathbb{A}}$*.*

## 4 Application to fuzzy formal concept analysis

This section illustrates with an example, how the mechanism given from Theorem 1 on information systems can be considered in a fuzzy concept lattice framework. Specifically, we have considered an example of the multi-adjoint concept lattice introduced in [12], which study the suitable journal for submitting a paper. In this example, several journals and different parameters appearing in the ISI Journal Citation Report have been taken into account.

The following sets of attributes and objects, together with the relation in Table 1, will be the elements of the context which will be considered.

$$A = \{\text{Impact Factor}(im), \text{Immediacy Index}(ii), \text{Cited Half-Life}(c), \text{Best Position}(b)\}$$
$$B = \{\text{AMC}, \text{CAMWA}, \text{FSS}, \text{IEEE-FS}, \text{IJGS}, \text{IJUFKS}, \text{JIFS}\}$$

where the "best position" means the best quartile of the different categories under which the journal is included, and the journals considered are Applied Mathematics and Computation (AMC), Computer and Mathematics with Applications (CAMWA), Fuzzy

[2] When $B = \{a\}$, i.e., $B$ is a singleton, we will write $I_a$ instead of $I_{\{a\}}$.

Sets and Systems (FSS), IEEE transactions on Fuzzy Systems (IEEE-FS), International Journal of General Systems (IJGS), International Journal of Uncertainty Fuzziness and Knowledge-based Systems (IJUFKS), Journal of Intelligent and Fuzzy Systems (JIFS).

The fuzzy relation between them, $R\colon A \times B \to P$, is the normalization of the information in the JCR to the regular partition of $[0,1]$ into 100 pieces, that is $[0,1]_{100} = \{0.00, 0.01, 0.02, \ldots, 0.99, 1.00\}$.

**Table 1.** Fuzzy relation between the objects and the attributes.

| $R$ | AMC | CAMWA | FSS | IEEE-FS | IJGS | IJUFKS | JIFS |
|---|---|---|---|---|---|---|---|
| Impact Factor | 0.34 | 0.21 | 0.52 | 0.85 | 0.43 | 0.21 | 0.09 |
| Immediacy Index | 0.13 | 0.09 | 0.36 | 0.17 | 0.1 | 0.04 | 0.06 |
| Cited Half-Life | 0.31 | 0.71 | 0.92 | 0.65 | 0.89 | 0.47 | 0.93 |
| Best Position | 0.75 | 0.5 | 1 | 1 | 0.5 | 0.25 | 0.25 |

We are going to define a tolerance relation in function of the distance between two objects, so:

$$T(a_i,a_j) = \begin{cases} 1 & \text{if} \quad d(a_i,a_j) \leqslant 0.1 \\ 0.7 & \text{if} \quad d(a_i,a_j) \leqslant 0.2 \\ 0.4 & \text{if} \quad d(a_i,a_j) \leqslant 0.3 \\ 0.1 & \text{if} \quad d(a_i,a_j) \leqslant 0.4 \\ 0 & \text{otherwise} \end{cases}$$

Let consider a fixed threshold as $\delta = 0.5$. We are building the discernibility matrix as a symmetric one:

$$\begin{pmatrix} \varnothing & & & & & & \\ \{c,b\} & \varnothing & & & & & \\ \{ii,c,b\} & \{im,ii,c,b\} & \varnothing & & & & \\ \{im,c,b\} & \{im,b\} & \{im,c\} & \varnothing & & & \\ \{c,b\} & \{im\} & \{ii,b\} & \{im,c,b\} & \varnothing & & \\ \{b\} & \{c,b\} & \{im,ii,c,b\} & \{im,b\} & \{im,c,b\} & \varnothing & \\ \{c,b\} & \{c,b\} & \{im,ii,b\} & \{im,c,b\} & \{im,b\} & \{c\} & \varnothing \end{pmatrix}$$

If we study the $0.5$-bidimensional discenibility function, we obtain:

$$\begin{aligned}\tau^{bir}_{0.5} = {} & \{1 \vee 2 \vee c \vee b\} \wedge \{1 \vee 3 \vee ii \vee c \vee b\} \wedge \{1 \vee 4 \vee im \vee c \vee b\} \\ & \wedge \{1 \vee 5 \vee c \vee b\} \wedge \{1 \vee 6 \vee b\} \wedge \{1 \vee 7 \vee c \vee b\} \\ & \wedge \{2 \vee 3 \vee im \vee ii \vee c \vee b\} \wedge \{2 \vee 4 \vee im \vee b\} \wedge \{2 \vee 5 \vee im\} \\ & \wedge \{2 \vee 6 \vee c \vee b\} \wedge \{2 \vee 7 \vee c \vee b\} \wedge \{3 \vee 4 \vee im \vee c\} \\ & \wedge \{3 \vee 5 \vee ii \vee b\} \wedge \{3 \vee 6 \vee im \vee ii \vee c \vee b\} \wedge \{3 \vee 7 \vee im \vee ii \vee b\} \\ & \wedge \{4 \vee 5 \vee im \vee c \vee b\} \wedge \{4 \vee 6 \vee im \vee b\} \wedge \{4 \vee 7 \vee im \vee c \vee b\} \\ & \wedge \{5 \vee 6 \vee im \vee c \vee b\} \wedge \{5 \vee 7 \vee im \vee b\} \wedge \{6 \vee 7 \vee c\}\end{aligned}$$

Thus, we obtain 51 $0.5$-information bireducts. Some of them are:

$(B_1, X_1) = (\{\text{Impact Factor}, \text{Immediacy Index}\}, \{1, 3, 4\})$
$(B_2, X_2) = (\{\text{Best Position}\}, \{1, 2, 4, 7\})$
$(B_3, X_3) = (\{\text{Impact Factor}\}, \{4, 5, 6\})$
$(B_4, X_4) = (\{\text{Impact Factor}, \text{Immediacy Index}, \text{Cited Half-Life}\}, \{2, 3, 4, 5, 6, 7\})$
$(B_5, X_5) = (\{\}, \{4, 7\})$

Therefore, the theory developed throughout this paper provides a new reduction method in FCA following the RST philosophy which is similar to the one given in FCA.

There exist several papers which reduce the size of the concept lattice considering similarities [1, 2, 7]. In the future, we will study the relation among these methods and the (bi)reduction proposed here for FCA.

## References


1. R. Bělohlávek. Similarity relations in concept lattices. *Journal of Logic and Computation*, 10(6):823–845, 2000.
2. R. Bělohlávek, J. Dvořák, and J. Outrata. Fast factorization by similarity in formal concept analysis of data with fuzzy attributes. *Journal of Computer and System Sciences*, 73(6):1012–1022, 2007.
3. M. Benítez, J. Medina, and D. Ślęzak. Delta-information reducts and bireducts. In J. M. Alonso, H. Bustince, and M. Reformat, editors, *2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15), Gijón, Spain*, pages 1154–1160. Atlantis Press, 2015.
4. M. Benítez, J. Medina, and D. Ślęzak. Reducing information systems considering similarity relations. In J. Kacprzyk, L. Koczy, and J. Medina, editors, *7th European Symposium on Computational Intelligence and Mathematices (ESCIM 2015)*, pages 257–263, 2015.
5. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. On the use of irreducible elements for reducing multi-adjoint concept lattices. *Knowledge-Based Systems*, 89:192–202, 2015.
6. C. Cornelis, R. Jensen, G. Hurtado, and D. Ślęzak. Attribute selection with fuzzy decision reducts. *Information Sciences*, 180:209–224, 2010.
7. A. Formica. Ontology-based concept similarity in formal concept analysis. *Information Sciences*, 176(18):2624–2641, 2006.

8. B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.
9. A. Janusz and D. Ślęzak. Rough set methods for attribute clustering and selection. *Applied Artificial Intelligence*, 28(3):220–242, 2014.
10. N. Mac Parthalain and R. Jensen. Simultaneous feature and instance selection using fuzzy-rough bireducts. In *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, pages 1–8, July 2013.
11. J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
12. 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.
13. D. Ślęzak and A. Janusz. Ensembles of bireducts: Towards robust classification and simple representation. In T.-h. Kim, H. Adeli, D. Ślęzak, F. Sandnes, X. Song, K.-i. Chung, and K. Arnett, editors, *Future Generation Information Technology*, volume 7105 of *Lecture Notes in Computer Science*, pages 64–77. Springer Berlin Heidelberg, 2011.
14. S. Stawicki and D. Ślęzak. Recent advances in decision bireducts: Complexity, heuristics and streams. *Lecture Notes in Computer Science*, 8171:200–212, 2013.

# On the $n$-ary Generalization of Dual Bonds

Ondrej Krídlo[1] and Manuel Ojeda-Aciego[2]

[1] University of Pavol Jozef Šafárik, Košice, Slovakia*
[2] Universidad de Málaga, Departamento de Matemática Aplicada, Spain**

**Abstract.** We propose the generalization of the notion bond between two formal contexts to the case of $n$ formal contexts. The first properties of the $n$-ary bonds are given, together with a method for building $n$-ary bonds. This construction enables to formalize some inference rules within the research topic of building a sequent calculus for formal contexts.

## 1 Introduction

One can often find problems in which it is necessary to decide whether certain tabular information is a consequence of underlying information given as a set of tables. This enables the possibility of interpreting each table as a formal context and, then, the problem translates to that of 'being a (logical) consequence' from a set of formal contexts.

The notion of logical consequence between formal contexts has been recently introduced in [4] and, to the best of our knowledge, it has been the only work focused on how to introduce logical content within the machinery of Formal Concept Analysis (FCA), apart from its ancient roots anchored in the Port-Royal logic, the so-called logical information systems [2, 6], and the fact that the concluding section of [5] states that dual bonds could be given a proof-theoretical interpretation in terms of consequence relations.

Specifically, in [4], taking into account that the category ChuCors (of contexts and Chu correspondences) is *-autonomous, and hence a model of linear logic, the authors obtained a preliminary notion of logical consequence relation between contexts which, together with the interpretation of the additive (resp. multiplicative) conjunction as the cartesian product (resp. bond) of contexts, enable to prove the correctness of the corresponding rules of the sequent calculus of the conjunctive fragment of linear logic.

The treatment of the disjunctive fragment of linear logic in terms of constructions between contexts suggests the need to introduce a more general version of the notion of (dual) bond, which allows to properly formalize the introduction and the elimination rules for the two different disjunctions of linear logic.

In this work, we propose the generalization of the notion bond between two formal contexts to the case of $n$ formal contexts, and obtain the first properties of the $n$-ary bonds, together with a method for building them.

---

* Partially supported by grants VEGA 1/0073/15 and APVV-15-0091
** Partially supported by Spanish Ministry of Science project TIN2015-70266-C2-1-P, co-funded by the European Regional Development Fund (ERDF)

## 2 Preliminaries

In order to make the manuscript self-contained, the fundamental notions and their main properties are recalled in this section.

### 2.1 Context, concept and concept lattice

**Definition 1.** *An algebra $\langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1\rangle$ is said to be a* ***complete residuated lattice*** *if*

1. *$\langle L, \wedge, \vee, 0, 1\rangle$ is a complete lattice with least element $0$ and greatest element $1$,*
2. *$\langle L, \otimes, 1\rangle$ is a commutative monoid,*
3. *$\otimes$ and $\rightarrow$ are adjoint, i.e. $a \otimes b \leq c$ if and only if $a \leq b \rightarrow c$, for all $a, b, c \in L$, where $\leq$ is the ordering generated by $\wedge$ and $\vee$.*

We now introduce the notions of *$L$-fuzzy context*, *extended derivation operations*, and *$L$-fuzzy concept* provided by Bělohlávek [1]. Notice that we will use the notation $Y^X$ to refer to the set of mappings from $X$ to $Y$.

**Definition 2.** *Let $L$ be a complete residuated lattice, an $L$-**fuzzy context** is a triple $\langle B, A, r\rangle$ consisting of a set of objects $B$, a set of attributes $A$ and an $L$-fuzzy binary relation $r$, i.e. a mapping $r\colon B \times A \longrightarrow L$, which can be alternatively understood as an $L$-fuzzy subset of $B \times A$.*

*Given an $L$-fuzzy context $\langle B, A, r\rangle$, a pair of mappings $\uparrow\colon L^B \longrightarrow L^A$ and $\downarrow\colon L^A \longrightarrow L^B$ can be defined for every $f \in L^B$ and $g \in L^A$ as follows:*

$$\uparrow(f)(a) = \bigwedge_{o \in B} (f(o) \rightarrow r(o, a)) \qquad \downarrow(g)(o) = \bigwedge_{a \in A} \big(g(a) \rightarrow r(o, a)\big) \tag{1}$$

**Definition 3.** *An $L$-**fuzzy concept** of an $L$-context $\mathcal{C} = \langle B, A, r\rangle$ is a pair $\langle f, g\rangle \in L^B \times L^A$ such that $\uparrow(f) = g$ and $\downarrow(g) = f$. The first component $f$ is said to be the* ***extent*** *of the concept, whereas the second component $g$ is the* ***intent*** *of the concept. The set of all $L$-fuzzy concepts associated to a fuzzy context $\langle B, A, r\rangle$ will be denoted as $L$-$\mathrm{FCL}(B, A, r)$. The set of all extents or intents of $\mathcal{C}$ will be denoted by $\mathrm{Ext}(\mathcal{C})$ or $\mathrm{Int}(\mathcal{C})$ respectively.*

### 2.2 Intercontextual structures

Two main constructions have been traditionally considered in order to relate two formal contexts: the bonds and the Chu correspondences.

**Definition 4.** *Let $\mathcal{C}_i = \langle B_i, A_i, r_i\rangle$ for $i \in \{1, 2\}$ be two formal contexts. A* bond *between $\mathcal{C}_1$ and $\mathcal{C}_2$ is any relation $\beta \in L^{B_1 \times A_2}$ such that, when interpreted as a table, its columns are extents of $\mathcal{C}_1$ and its rows are intents of $\mathcal{C}_2$. All bonds between such contexts will be denoted by $L$-$\mathrm{Bonds}(\mathcal{C}_1, \mathcal{C}_2)$.*

*Another equivalent definition of bond between $\mathcal{C}_1$ and $\mathcal{C}_2$ defines it as any relation $\beta \in L^{B_1 \times A_2}$ such that $\mathrm{Ext}(\langle B_1, A_2, \beta\rangle) \subseteq \mathrm{Ext}(\mathcal{C}_1)$ and $\mathrm{Int}(\langle B_1, A_2, \beta\rangle) \subseteq \mathrm{Int}(\mathcal{C}_2)$*

Dual bonds *between $\mathcal{C}_1$ and $\mathcal{C}_2$ are bonds between $\mathcal{C}_1$ and the transposed of $\mathcal{C}_2$. The transposed of a context $\mathcal{C} = \langle B, A, r\rangle$ is defined as a new context $\mathcal{C}^* = \langle A, B, r^t\rangle$ with $r^t(a, b) = r(b, a)$ for all $(b, a) \in B \times A$.*

The notion of Chu correspondence between contexts can be seen as an alternative inter-contextual structure which, instead, links intents of $\mathcal{C}_1$ and extents of $\mathcal{C}_2$.

**Definition 5.** *Consider $\mathcal{C}_1 = \langle B_1, A_1, r_1\rangle$ and $\mathcal{C}_2 = \langle B_2, A_2, r_2\rangle$ two formal contexts. An $L$-*Chu correspondence *between $\mathcal{C}_1$ and $\mathcal{C}_2$ is any pair $\varphi = \langle \varphi_L, \varphi_R\rangle$ of mappings $\varphi_L \colon B_1 \longrightarrow \mathrm{Ext}(\mathcal{C}_2)$ and $\varphi_R \colon A_2 \longrightarrow \mathrm{Int}(\mathcal{C}_1)$ such that for all $(b_1, a_2) \in B_1 \times A_2$ it holds that*

$$\uparrow_2\big(\varphi_L(b_1)\big)(a_2) = \downarrow_1\big(\varphi_R(a_2)\big)(b_1)$$

*All $L$-Chu correspondences between $\mathcal{C}_1$ and $\mathcal{C}_2$ is denoted by* $\mathrm{Chu}(\mathcal{C}_1, \mathcal{C}_2)$.

The notions of bond and Chu correspondence are interchangeable; specifically, we can consider the bond $\beta_\varphi$ associated to a Chu correspondence $\varphi$ from $\mathcal{C}_1$ to $\mathcal{C}_2$ defined for $b_1 \in B_1, a_2 \in A_2$ as follows

$$\beta_\varphi(b_1, a_2) = \uparrow_2\big(\varphi_L(b_1)\big)(a_2) = \downarrow_1\big(\varphi_R(a_2)\big)(b_1)$$

Similarly, we can consider the Chu correspondence $\varphi_\beta$ associated to a bond $\beta$ defined by the following pair of mappings:

$$\varphi_{\beta L}(b_1) = \downarrow_2\big(\beta(b_1)\big) \text{ for all } o_1 \in B_1$$
$$\varphi_{\beta R}(a_2) = \uparrow_1\big(\beta^t(a_2)\big) \text{ for all } a_2 \in A_2$$

The set of all bonds (resp. Chu correspondences) between two formal contexts endowed with the ordering given by set inclusion is a complete lattice. Moreover, both complete lattices are dually isomorphic.

Note that if $\varphi \in \mathrm{Chu}(\mathcal{C}_1, \mathcal{C}_2)$, then we can consider $\varphi^* \in \mathrm{Chu}(C_2^*, C_1^*)$ defined by $\varphi_L^* = \varphi_R$ and $\varphi_R^* = \varphi_L$.

### 2.3 Tensor product

**Definition 6.** *The tensor product $\mathcal{C}_1 \boxtimes \mathcal{C}_2$ of contexts $\mathcal{C}_i = \langle B_i, A_i, R_i\rangle$ for $i \in \{1, 2\}$ is defined as the context $\langle B_1 \times B_2, \mathrm{Chu}(\mathcal{C}_1, \mathcal{C}_2^*), R_\boxtimes\rangle$ where*

$$R_\boxtimes\big((b_1, b_2), \varphi\big) = \downarrow_2\big(\varphi_L(b_1)\big)(b_2).$$

The properties of the tensor product were shown in [3], together with the result that $L$-ChuCors with $\boxtimes$ is symmetric and monoidal. The structure of the formal concepts of a tensor product context was established as an ordered pair formed by a bond and a set of Chu correspondences.

**Lemma 1.** *Let $(\beta, X)$ be a formal concept of the tensor product $\mathcal{C}_1 \boxtimes \mathcal{C}_2$, it holds that $\beta = \bigwedge_{\psi \in X} \beta_\psi$ and $X = \{\psi \in \mathrm{Chu}(\mathcal{C}_1, \mathcal{C}_2^*) \mid \beta \le \beta_\psi\}$.*

Note that, due to the structure of complete lattice, any set of bonds is closed under intersection; hence any extent of the tensor product of any two contexts is a bond between such contexts.

## 3 Introducing $n$-ary bonds

Due to the symmetric and monoidal properties of $\boxtimes$ on $L$-ChuCors we can introduce the notion of $n$-ary tensor product $\boxtimes_{i=1}^{n}\mathcal{C}_i$ of $n$ formal contexts $\mathcal{C}_i$ for $i \in \{1, \ldots, n\}$. Hence, it is possible to consider a notion of $n$-ary bond that we can imagine as any extent of this $n$-ary tensor product. This notion is formally introduced as follows:

**Definition 7.** *Let $\mathcal{C}_i = \langle B_i, A_i, r_i\rangle$ be formal contexts for $i \in \{1, \ldots, n\}$. An $n$-*ary dual bond *between $\{\mathcal{C}_i\}_{i=1}^{n}$ is an $n$-ary relation $\beta : \prod_{i=1}^{n} B_i \longrightarrow L$ such that for all $i \in \{1, 2, \ldots, n\}$ and any $(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n) \in \prod_{j=1, j\neq i}^{n} B_i$ it holds that $\beta(b_1, \ldots, b_{i-1}, \_ , b_{i+1}, \ldots, b_n) \in \mathrm{Ext}(\mathcal{C}_i)$. The set of all $n$-ary dual bonds between contexts $\mathcal{C}_1, \ldots, \mathcal{C}_n$ will be denoted by $nL\text{-Bonds}(\mathcal{C}_1, \ldots, \mathcal{C}_n)$.*

The two following results can be obtained from the definition and the preliminary results introduced above.

**Lemma 2.** *Let $\mathcal{C}_i = \langle B_i, A_i, r_i\rangle$ be formal contexts for $i \in \{1, \ldots, n\}$. Any extent of $n$-ary tensor product $\boxtimes_{i=1}^{n}\mathcal{C}_i$ is an $n$-ary dual bond between such formal contexts.*

**Lemma 3.** *Let $\{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ be a set of $n$ formal contexts and $\beta$ be some $n$-ary dual bond between such contexts. Let $\mathcal{D}_i^{\beta}$ be a new formal context defined as $\langle B_i, \prod_{j=1,j\neq i}^{n} B_j, \mathcal{R}_i\rangle$ where $\mathcal{R}_i(b_i, (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)) = \beta(b_1, \ldots, b_n)$ for any $i \in \{1, 2, \ldots, n\}$. Then $\mathrm{Ext}(\mathcal{D}_i^{\beta}) \subseteq \mathrm{Ext}(\mathcal{C}_i)$.*

## 4 Building $n$-ary bonds

Instead of generating bonds as extents of a tensor product, which is a redundant process, there is another easy way to compute bonds between two formal contexts via their direct product, which is built on the operation of fuzzy disjunction (the direct product of two $L$-fuzzy contexts was studied in details in [3]). This process of generating bonds can be extended to more than two formal contexts as it shown below.

In our framework, the fuzzy disjunction $\oplus : L \times L \longrightarrow L$ will be constructed in terms of fuzzy implication as

$$k \oplus m = \neg k \to m = (k \to 0) \to m$$

**Definition 8.** *The* direct product *of two $L$-contexts $\mathcal{C}_1 = \langle B_1, A_1, r_1\rangle$ and $\mathcal{C}_2 = \langle B_2, A_2, r_2\rangle$ is a new $L$-context $\mathcal{C}_1 \Delta \mathcal{C}_2 = \langle B_1 \times A_2, A_1 \times B_2, \Delta\rangle$, where $\Delta$ is defined as $\Delta((o_1, a_2), (a_1, o_2)) = r_1(o_1, a_1) \oplus r_2(o_2, a_2)$.*

**Theorem 1 (See [3]).** *Let $\mathcal{C}_i = \langle B_i, A_i, r_i\rangle$ be $L$-contexts for $i \in \{1, 2\}$, where $L$ satisfies the double negation law, and let $\beta$ be an $L$-multifunction between $B_1$ and $A_2$. If $\beta \in \mathrm{Ext}(\mathcal{C}_1 \Delta \mathcal{C}_2)$, then $\beta \in L\text{-Bonds}(\mathcal{C}_1, \mathcal{C}_2)$.*

By simply transposing the second context, it is easy to generate dual bonds as extents of $\mathcal{C}_1 \Delta \mathcal{C}_2^{*}$.

If the underlying structure of truth-values satisfies the double negation law ($\neg\neg k = k$ for any $k \in L$), then $\oplus$ is a commutative and associative operation. Hence we can consider the following extension into $n$-ary disjunction of $n$ values of $L$.

**Definition 9.** *Let $k_1, \ldots, k_n \in L$ be arbitrary values. The $n$-ary disjunction of values $\{k_1 \ldots, k_n\}$ is defined as*

$$\bigoplus_{i=1}^{n} k_i = k_1 \oplus (k_2 \oplus (k_3 \oplus (\ldots (k_{n-1} \oplus k_n)))) \,.$$

**Definition 10.** *Let $\mathcal{C}_1, \ldots, \mathcal{C}_n$ be $n$ arbitrary $L$-contexts. The* direct product *of such contexts is defined by*

$$\varDelta_{i=1}^{n} \mathcal{C}_i = \left\langle \prod_{i=1}^{n} B_i, \prod_{i=1}^{n} A_i, r_\Delta \right\rangle$$

*where $r_\Delta((b_1 \ldots, b_n), (a_1, \ldots, a_n)) = \bigoplus_{i=1}^{n} r_i(b_i, a_i)$.*

**Lemma 4.** *Given $n$ arbitrary $L$-contexts $\mathcal{C}_1, \ldots, \mathcal{C}_n$, it holds that*

$$\mathrm{Ext}(\varDelta_{i=1}^{n} \mathcal{C}_i) \subseteq nL\text{-}\mathrm{Bonds}(\mathcal{C}_1, \ldots, \mathcal{C}_n) \,.$$

## 5 Conclusion

We have introduced the notion of $n$-ary dual bond as a technical tool within our wider research line on the development of a sequent calculus for formal contexts. The first properties of this new notion have been presented, together with a characterization which allows a straightforward construction of $n$-ary dual bonds.

As future work, we will use the new notion in order to formalize the context-based version of the introduction and the elimination rules for the two kinds of disjunction of linear logic, underlying the theory of bonds and Chu correspondences.

## References


1. R. Bělohlávek. *Fuzzy Relational Systems: Foundations and Principles*. Kluwer Academic Publishers, 2002.
2. P. Cellier, S. Ferré, M. Ducassé, and T. Charnois. Partial orders and logical concept analysis to explore patterns extracted by data mining. *Lecture Notes in Artificial Intelligence*, 6828:77–90, 2011.
3. O. Krídlo, S. Krajči, and M. Ojeda-Aciego. The category of $L$-Chu correspondences and the structure of $L$-bonds. *Fundamenta Informaticae*, 115(4):297–325, 2012.
4. O. Krídlo and M. Ojeda-Aciego. Towards a sequent calculus for formal contexts. Submitted, 2016.
5. M. Krötzsch, P. Hitzler, and G.-Q. Zhang. Morphisms in context. *Lecture Notes in Computer Science*, 3596:223–237, 2005.
6. O. Ridoux and S. Ferré. Introduction to logical information systems. *Information Processing and Management*, 40:383–419, 2004.

# Toward the Use of the Contraposition Law in Multi-adjoint Lattices ⋆

Nicolás Madrid

Department of Applied Mathematics
University of Málaga, Spain
nicolas.madrid@uma.es

**Abstract.** This paper deals with the issue of combining involutive negations and the contraposition law. Specifically, instead of considering the negation $x \to 0$, we consider arbitrary involutive negations and we define a multi-adjoint lattice where we can apply the contraposition law by incorporating to the original implication a set of new implications.

**Keywords:** multiadjoint lattices, involutive negation, contraposition law, residuated pairs, adjoint triple.

## 1 Introduction

In logic, the basic connectives (as the implication ($\to$), the conjunction ($\wedge$), the disjunction ($\vee$) and the negation ($\neg$)) hold interesting relationships among them via tautologies. For instance, in terms of propositional logic, some of the most famous are:

- $p \wedge q \equiv \neg(\neg p \vee \neg q)$ (the Morgan's law)
- $p \to q \equiv \neg q \to \neg p$ (the contraposition law)
- $p \to q \equiv \neg p \vee q$ (the material implication)
- $\neg(\neg p) \equiv p$ (the double negation law)

Depending on how those connectives are generalized into fuzzy logic, many of such relationships between connectives could be no longer valid [10, 4]. It is worth to take into account that the most approaches of fuzzy logic focus on residuated lattice, which is motivated by the modus ponens inference. So the relationships between connectives are subordinated by the relationships imposed by the adjoint property (the core of residuated lattices).

In this paper we focus on the satisfability of the double negation law and the contraposition law. The former is modelled in fuzzy logic by means of involutive negations [3]. The latter, in the residuated lattice framework, requires the use of a specific negation, the one defined by $x \to 0$. The issue is that in general, the negation $x \to 0$ is not involutive [3]. Some authors motivate the use of fuzzy logics based on Łukasiewicz connectives because it is defined on the theory of residuated lattices and it holds, among others, the double negation and the contraposition law [10]. However, real applications require connectives different from the Łukasiewicz ones. This approach aims at the development

⋆ This work was supported by the Spanish governmental project TIN15-70266-C2-P-1

of a theoretical framework where it is possible to consider arbitrary residuated pairs, arbitrary involutive negations and still, a framework where it is possible to apply "a kind of" contraposition law in the operators considered.

The background of this paper is multi-adjoint lattice theory[9], an approach that has shown to be useful in several fields like Logic Programming [8], Concept Lattice Analysis [5, 6] and Relation Equations [2]. The main difference between multi-adjoint lattices and residuated lattices, is that firstly, the former can contain several residuated pairs; i.e. several conjunctions and implications. And secondly, it allows to use non-commutative conjunctions, which implies the existence of two different residuated implications per each (non-commutative) conjunction. The idea behind this approach is to consider a set of different implications $\to_1, \ldots, \to_n$ related by the contraposition formula; in the sense that for each $i \in \{i, \ldots, n\}$ there exists a $j \in \{i, \ldots, n\}$ such that $x \to_i y \equiv n(x) \to_j n(y)$. Thus, a contraposition rule (understanding rule like a method to substitute a formula by another) could be used by jumping from one implication onto another. This idea contrasts with other studies about negations on multi-adjoint lattices, like [1], where negations are defined from implications by the formula $x \to 0$.

The paper starts by showing some preliminaries in Section 2. Then, in Section 3, the main contribution is presented by defining firstly a multi-adjoint lattice to model the contraposition law and, subsequently, a sequence of results that relate the different implications appearing in the mentioned multi-adjoint lattice. Finally, Section 4 shows the conclusions and future work.

## 2 Preliminaries

As I have stated in the introduction, this approach is framed in the field of multi-adjoint lattices. The core of multi-adjoint lattices is the notion of adjoint triple that generalizes the well-known notion of residuated pair, which is widely used in fuzzy logic.

**Definition 1** **([7]).** *Let $(P_1, \leq_1)$, $(P_2, \leq_2)$, $(P_3, \leq_3)$ be three posets. I say that the mappings $\&\colon P_1 \times P_2 \to P_3$, $\searrow\colon P_2 \times P_3 \to P_1$, and $\nearrow\colon P_1 \times P_3 \to P_2$ form an* adjoint triple *among $P_1, P_2$ and $P_3$ whenever:*

$$\begin{aligned} x \leq_1 y \searrow z \quad & \textit{if and only if} \\ x \,\&\, y \leq_3 z \quad & \textit{if and only if} \\ y \leq_2 x \nearrow z \quad & \end{aligned} \tag{1}$$

*for all $x \in P_1$, $y \in P_2$ and $z \in P_3$.*

The main difference with respect to residuated pairs is that, in the case of adjoint triples, the operator used to model the conjunction ($\&$) is not necessarily commutative. This fact implies that I need two implications, denoted in Definition 1 by $\searrow$ and $\nearrow$, to model the adjoint (also called residuated) property. The following result present some properties of the operator used in adjoint triples.

**Lemma 1.** *If $(\&, \swarrow, \nwarrow)$ is an adjoint triple w.r.t. $P_1, P_2, P_3$, then*

1. *$\&$ is order-preserving on both arguments, i.e. if $x_1, x_2, x \in P_1$, $y_1, y_2, y \in P_2$ and $x_1 \leq_1 x_2$, $y_1 \leq_2 y_2$, then $(x_1 \,\&\, y) \leq_3 (x_2 \,\&\, y)$ and $(x \,\&\, y_1) \leq_3 (x \,\&\, y_2)$.*
2. *$\swarrow$, $\nwarrow$ are order-preserving on the first argument and order-reversing on the second argument, i.e., if $x_1, x_2, x \in P_1$, $y_1, y_2, y \in P_2$, $z_1, z_2, z \in P_3$ and $x_1 \leq_1 x_2$, $y_1 \leq_2 y_2$, $z_1 \leq_3 z_2$, then $(z_1 \swarrow y) \leq_1 (z_2 \swarrow y)$, $(z \swarrow y_2) \leq_1 (z \swarrow y_1)$, $(z_1 \nwarrow x) \leq_2 (z_2 \nwarrow x)$ and $(z \nwarrow x_2) \leq_2 (z \nwarrow x_1)$.*

In this paper I consider a simplified structure of adjoint triples, specifically that where $(P_1, \leq_1), (P_2, \leq_2)$ and $(P_3, \leq_3)$ are considered equal to a lattice $(L, \leq)$. I can define now the notion of multi-adjoint lattice.

**Definition 2 ([7]).** *A multi-adjoint lattice is a tuple*

$$(L, \leq, (\&_1, \swarrow_1, \nwarrow_1), (\&_2, \swarrow_2, \nwarrow_2), \ldots, (\&_k, \swarrow_k, \nwarrow_k))$$

*where $(L, \leq)$ is a complete lattice and $(\&_i, \swarrow_i, \nwarrow_i)$ is an adjoint triple on $(L, \leq)$ for each $i \in \{1, \ldots, k\}$.*

In this paper I consider multi-adjoint latticesI enriched with an involutive negation[1]. Let us recall that an involutive negation on a lattice $(L, \leq)$ is an operator $n\colon L \to L$ such that $n(0) = 1$, $n(1) = 0$ and $n(n(x)) = x$ for all $x \in L$.

## 3 Contraposition rule and multi-adjoint lattices

I start this section by considering just an adjoint triple $(\&, \searrow, \nearrow)$ defined on a lattice $L$ with an involutive negation $n$. As I have introduced in precedent sections, my goal is to formalize the contraposition inference rule in a multi-adjoint environment. Thus, given $x, y \in L$, my purpose is to construct a multi-adjoint framework where I can combine $x \searrow y$,$x \nearrow y$,$n(y) \searrow n(x)$ and $n(y) \nearrow n(x)$. I propose to consider different adjoint triples, specifically two, one associated to the implication $n(y) \searrow n(x)$ and another to the implication $n(y) \nearrow n(x)$.

**Definition 3.** *Let $(\&, \searrow, \nearrow)$ be an adjoint triple defined in a lattice L with an involutive negation $n$. The $\searrow_n$-*adjoint triple *$(\&_n, \searrow_n, \nearrow_n)$ of $(\&, \searrow, \nearrow)$ is given by the following operators:*

- *$x \nearrow_n y = n(x \,\&\, n(y))$ for all $x, y \in L$*
- *$x \,\&_n\, y = n(x \nearrow n(y))$ for all $x, y \in L$.*
- *$x \searrow_n y = n(y) \searrow n(x)$ for all $x, y \in L$*

*Similarly, the $\nearrow^n$-*adjoint triple *$(\&^n, \searrow^n, \nearrow^n)$ of $(\&, \searrow, \nearrow)$ is given by the following operators:*

- *$x \nearrow^n y = n(y) \nearrow n(x)$ for all $x, y \in L$*
- *$x \,\&^n\, y = n(y \searrow n(x))$ for all $x, y \in L$*
- *$x \searrow^n y = n(n(y) \,\&\, x)$ for all $x, y \in L$ .*

[1] Also called strong negation in many approaches of fuzzy logic.

The definition above requires a proof to justify the use of the term 'adjoint triple' in it.

**Lemma 2.** *Let* $(\&, \searrow, \nearrow)$ *be an adjoint triple defined in a lattice* $L$ *with an involutive negation* $n$*. Then* $(\&_n, \searrow_n, \nearrow_n)$ *and* $(\&^n, \searrow^n, \nearrow^n)$ *(as given in Definition 3) are adjoint triples as well.*

Two remarks. Firstly, I have opted to use the subindex and supindex notation to mark the implication where is applied the contraposition rule. Thus, the adjoint triple associated to the supindex ($^n$) denotes those operators associated to the application of the contraposition formula to $\nearrow$. In parallel, the adjoint triple associated to the subindex ($_n$) is the one where the contraposition formula is applied to $\searrow$. Secondly, note that the other implication in the adjoint triple coincides with the one that extends the crisp logic equivalence $x \to y \equiv \neg(x \wedge \neg y)$, which is linked to a common interpretation of the implication in crisp logic, i.e., *the only case where* $x \to y$ *is false is when* $x$ *is true and* $y$ *false*. So, the notation of sup and sub-indexes together with the second remark simplifies the memorization and use of formulae in Definition 3.

The next proposition presents an interesting relation among the sequential composition of the construction of $\nearrow^n$-adjoint triples and $\searrow_n$-adjoint triples. Note that by composing twice such constructions, there are, a priori, four possible new adjoint triples, namely:

- $((\&_n)_n, (\searrow_n)_n, (\nearrow_n)_n)$; i.e. the $\searrow_n$-adjoint triple of $(\&_n, \searrow_n, \nearrow_n)$.
- $((\&_n)^n, (\searrow_n)^n, (\nearrow_n)^n)$; i.e. the $\nearrow^n$-adjoint triple of $(\&_n, \searrow_n, \nearrow_n)$.
- $((\&^n)^n, (\searrow^n)^n, (\nearrow^n)^n)$; i.e. the $\nearrow^n$-adjoint triple of $(\&^n, \searrow^n, \nearrow^n)$.
- $((\&^n)_n, (\searrow^n)_n, (\nearrow^n)_n)$; i.e. the $\searrow_n$-adjoint triple of $(\&^n, \searrow^n, \nearrow^n)$.

The next result shows that the four possibilities coincide, up to commutative, either with $(\&, \searrow, \nearrow)$, $(\&_n, \searrow_n, \nearrow_n)$ or $(\&^n, \searrow^n, \nearrow^n)$.

**Proposition 1.** *Let* $(\&, \searrow, \nearrow)$ *be an adjoint triple defined in a lattice* $L$ *with an involutive negation* $n$*. Then the following equalities hold:*

1. $x(\&_n)_n y = x(\&^n)^n y = x \,\&\, y$.
2. $x(\&^n)_n y = x(\&_n)^n y = y \,\&\, x$.
3. $x(\searrow^n)^n y = x \searrow y$.
4. $x(\nearrow^n)^n y = x \nearrow y$.
5. $x(\searrow_n)_n y = x \searrow y$
6. $x(\nearrow_n)_n y = x \nearrow y$
7. $x(\nearrow^n)_n y = x \searrow_n y$.
8. $x(\searrow^n)_n y = x \nearrow_n y$.
9. $x(\nearrow_n)^n y = x \searrow^n y$
10. $x(\searrow_n)^n y = x \nearrow^n y$

The proposition above shows that given an adjoint triple $(\&, \searrow, \nearrow)$ defined on a lattice $L$ with an involutive negation $n$, the multi-adjoint framework given by:

$$(L, \leq, (\&, \searrow, \nearrow), (\&_n, \searrow_n, \nearrow_n), (\&^n, \searrow^n, \nearrow^n)) \tag{2}$$

is a close framework where it is possible to apply the contraposition rule associated to $n$ an indefinitely (numerable) number of times. For instance, from the equality 8 of Proposition 1, for every $x, y \in L$ we have:

$$x \nearrow_n y = n(y) \searrow^n n(x).$$

So it is possible to apply the contraposition rule to the implication $\searrow^n$ by using the implication $\nearrow_n$. Note that Proposition 1 covers all the possible application of the contraposition rule in the multi-adjoint framework of equation (2).

There is another interesting consequence of the Proposition 1: applying twice the construction of the $\searrow^n$-adjoint triple (resp. the $\nearrow^n$-adjoint triple) yields the original adjoint triple.

**Corollary 1.** *Let $(\&, \searrow, \nearrow)$ be an adjoint triple defined on a lattice $L$ with an involutive negation $n$. Then, the $\searrow^n$-adjoint triple (resp. $\nearrow^n$-adjoint triple) of the $\searrow^n$-adjoint triple (resp. $\nearrow^n$-adjoint triple) of $(\&, \searrow, \nearrow)$ is $(\&, \searrow, \nearrow)$.*

To end the section, I present a brief study restricted to residuated pairs $(\&, \to)$, which is a special case of adjoint triples where the operator $\&$ is commutative. Note that in such a case, both implications of an adjoint triple, $\searrow$ and $\nearrow$, coincides and are represented directly by $\to$. Anyway, we can construct the $\searrow^n$-adjoint triple of a residuated pair $(\&, \to)$ and, obviously, it coincides with its respective $\nearrow^n$-adjoint triple. However, note that in general $\nearrow^n \neq \searrow^n$. Thus, the $\searrow^n$-adjoint triple of a residuated pair differs, in general, from a residuated pair. So, the adjoint triple framework with a non-commutative operator $\&$ is necessary even if the origin of the approach focus on residuated pairs. Given a residuated pair $(\&, \to)$ and an involutive negation $n$, the multi-adjoint framework associated by equation (2) is:

$$(L, \leq, (\&, \to), (\&_n, \searrow_n, \nearrow_n))$$

where, thanks to Proposition 1, the following equalities hold:

- $x \to y = n(y) \searrow_n n(x)$,
- $x \searrow_n y = n(y) \to n(x)$,
- $x \nearrow_n y = n(y) \nearrow_n n(x)$

for all $x, y \in L$. Moreover, we have the following result, which characterises when the multiadjoint framework associated to a residuated pair is a residuated lattice again. Note that it is equivalent to show that

$$\&_n = \& \quad \text{and} \quad \searrow_n = \nearrow_n = \to$$

or equivalently

$$(\&, \to) \equiv (\&_n, \searrow_n, \nearrow_n)$$

i.e., the $\searrow^n$-adjoint triple coincides with the original residuated lattice.

**Proposition 2.** *Let $(\&, \to)$ be a residuated pair on $L$ and let $n$ be an involutive negation on $L$. Then, the multi-adjoint framework associated by equation* (2) *to $(L, \leq (\&, \to))$ is*

$$(L, \leq (\&, \to))$$

*if and only if $n(x) = x \to 0$.*

Note that the proposition above requires that the negation defined by $n(x) = x \to 0$ is involutive.

## 4 Conclusion and future work

In this paper I have presented an approach aimed at providing a theoretical background where the double negation and the contraposition rule can be applied in a broader way. Instead of considering the typical negation $x \to 0$, which rarely holds the double negation law, the approach considers arbitrary involutive negations and multi-adjoint lattices. Thus, given an adjoint triple (or simply a residuated pair), we have defined a multi-adjoint framework with exactly three adjoint triples that is close by the contraposition rule.

As a future work there are two different research lines. On the one hand I find interesting the development of a formal logic theory based on this approach. On the other hand, the extension of the approach for non-involutive negations.

## References


1. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Adjoint negations, more than residuated negations. *Information Science*, 345:355–371, 2016.
2. J. C. Díaz and J. Medina. Multi-adjoint relation equations: Definition, properties and solutions using concept lattices. *Information Science*, 253:100–109, 2013.
3. F. Esteva, L. Godo, P. Hájek, and M. Navara. Residuated fuzzy logics with an involutive negation. *Arch. Math. Log.*, 39(2):103–124, 2000.
4. P. Hajek. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publishers, 1998.
5. J. Konecny, J. Medina, and M. Ojeda-Aciego. Multi-adjoint concept lattices with heterogeneous conjunctors and hedges. *Annals of Mathematics and Artificial Intelligence*, 72(1-2):73–89, 2014.
6. J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Science*, 180(5):712–725, 2010.
7. 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.
8. J. Medina, M. Ojeda-Aciego, and P. Vojtás. A completeness theorem for multi-adjoint logic programming. In *Proceedings of the 10th IEEE International Conference on Fuzzy Systems*, pages 1031–1034, 2001.
9. J. Medina, M. Ojeda-Aciego, and P. Vojtás. Similarity-based unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146(1):43–62, 2004.
10. V. Novák, I. Perfilieva, and J. Mockor. *Mathematical Principles of Fuzzy Logic*, volume 517 of *The Springer International Series in Engineering and Computer Science*. Springer, 1999.

# On Pseudo-fixed Points of the Intuitionistic Fuzzy Quantifiers and Operators

Krassimir Atanassov

Department of Bioinformatics and Mathematical Modelling
Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 105, Sofia-1113, Bulgaria
and Prof. Asen Zlatarov University, Bourgas-8000, Bulgaria
krat@bas.bg

**Abstract.** In this paper, the pseudo-fixed points of the intuitionistic fuzzy quantifiers and operators from modal and level types, are described.

**Keywords:** intuitionistic fuzzy operator, intuitionistic fuzzy quantifier, pseudo-fixed point.

## 1 Introduction

The first research in the area of intuitionistic fuzzy logics started 30 years ago. Sequentially, intuitionistic fuzzy propositional calculus, intuitionistic fuzzy predicate logic, intuitionistic fuzzy modal logic and intuitionistic fuzzy temporal logic were introduced and developed.

In intuitionistic fuzzy predicate logic, firstly, intuitionistic fuzzy quantifiers that are analogous of the standard logic quantifiers were defined, and after this a series of their extensions arised. In intuitionistic fuzzy modal logic, firstly, intuitionistic fuzzy modal operators that are analogues of standard modal logic operators "necessity" and "possibility" were introduced and after this, a series of their extensions and modifications were defned. Level operators, that are intuitionistic fuzzy analogues of the standard fuzzy set operators, were introduced, too.

Here, the pseudo-fixed points of all these operators will be described.

## 2 Main results

In classical logic (e.g., [13, 14, 16]), to each proposition (sentence) we juxtapose its truth value: truth – denoted by 1, or falsity – denoted by 0. In the case of fuzzy logic [15], this truth value is a real number in the interval $[0, 1]$ and it is called "truth degree" or "degree of validity". In the intuitionistic fuzzy case (see [2, 3, 5, 6] we add one more value – "falsity degree" or "degree of non-validity" – which is again in interval $[0, 1]$. Thus, to the proposition $p$, two real numbers, $\mu(p)$ and $\nu(p)$, are assigned with the following constraint:

$$\mu(p), \nu(p) \in [0, 1] \text{ and } \mu(p) + \nu(p) \leq 1. \tag{1}$$

Therefore, if $\mathcal{S}$ is a set of propositions (or more general, formulas), then we can construct function $V : \mathcal{S} \to [0,1] \times [0,1]$ that is defined for every $p \in \mathcal{S}$ by

$$V(p) = \langle \mu(p),\ \nu(p) \rangle.$$

In [11], the pair $\langle \mu(p), \nu(p) \rangle$ that satisfies condition (1) is called "Intuitionistic Fuzzy Pair" (IFP).

Now, we introduce the separate quantifiers and operators from modal and level types.

Let $x$ be a variable, obtaining values in set $E$ and let $P(x)$ be a predicate with a variable $x$. Let

$$V(P(x)) = \langle \mu(P(x)), \nu(P(x)) \rangle.$$

The IF-interpretations of the (intuitionistic fuzzy) quantifiers *for all* ($\forall$) and *there exists* ($\exists$) are introduced in [3, 12, 9] by

$$V(\exists x P(x)) = \langle \sup_{y \in E} \mu(P(y)), \inf_{y \in E} \nu(P(y)) \rangle,$$

$$V(\forall x P(x)) = \langle \inf_{y \in E} \mu(P(y)), \sup_{y \in E} \nu(P(y)) \rangle.$$

Their geometrical interpretations are illustrated in Figures 1 and 2, respectively, where $x_1, ..., x_5$ are possible values of variable $x$ and $V(x_1), ..., V(x_5)$ are their IF-evaluations.

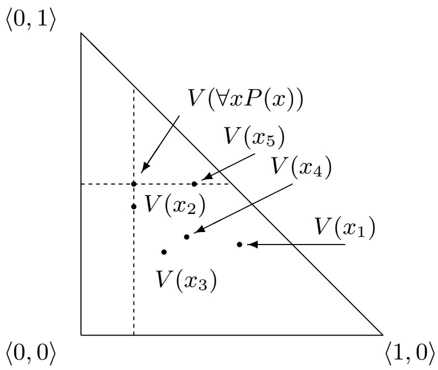



Fig. 1.

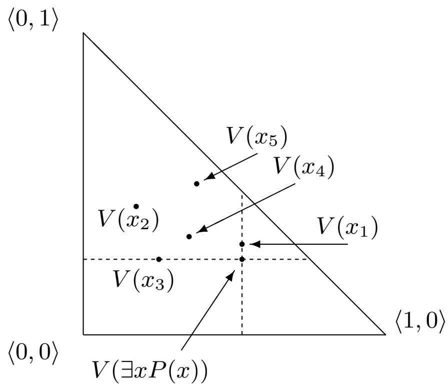



Fig. 2.

In [5], we introduced the following six quantifiers and studied some of their properties.

$$V(\forall_\mu x P(x)) = \{\langle x, \inf_{y \in E} \mu(P(y)), \nu(P(x)) \rangle | x \in E\},$$

$$V(\forall_\nu x P(x)) = \{\langle x, \min(1 - \sup_{y \in E} \nu(P(y)), \mu(P(x))), \sup_{y \in E} \nu(P(y)) \rangle | x \in E\},$$

$$V(\exists_\mu x P(x)) = \{\langle x, \sup_{y\in E} \mu(P(y)), \min(1 - \sup_{y\in E} \mu(P(y)), \nu(P(x)))\rangle | x \in E\},$$

$$V(\exists_\nu x P(x)) = \{\langle x, \mu(P(x)), \inf_{y\in E} \nu(P(y))\rangle | x \in E\},$$

$$V(\forall^*_\nu x P(x))$$
$$= \{\langle x, \min(1 - \sup_{y\in E} \nu(P(y)), \mu(P(x))), \min(\sup_{y\in E} \nu(P(y)), 1 - \mu(P(x)))\rangle | x \in E\},$$

$$V(\exists^*_\mu x P(x))$$
$$= \{\langle x, \min(\sup_{y\in E} \mu(P(y)), 1 - \nu(P(x))), \min(1 - \sup_{y\in E} \mu(P(y)), \nu(P(x)))\rangle | x \in E\}.$$

Let the possible values of the variable $x$ be $a, b, c$ and let their IF-evaluations $V(a)$, $V(b), V(c)$ be shown on Fig. 3. The geometrical interpretations of the new quantifiers are shown in Figs. 4 – 9.

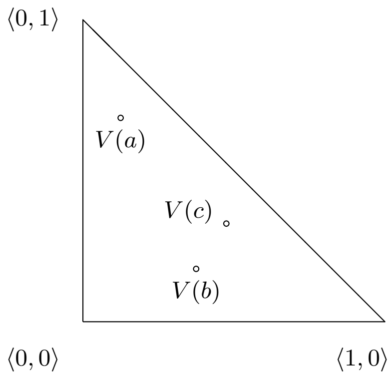



Fig. 3.

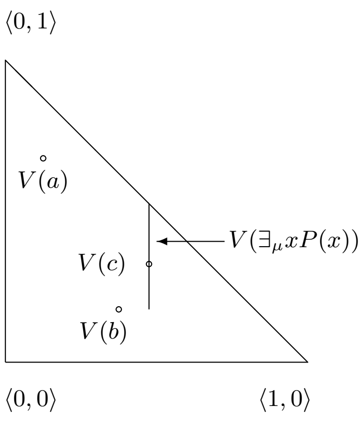



Fig. 4.

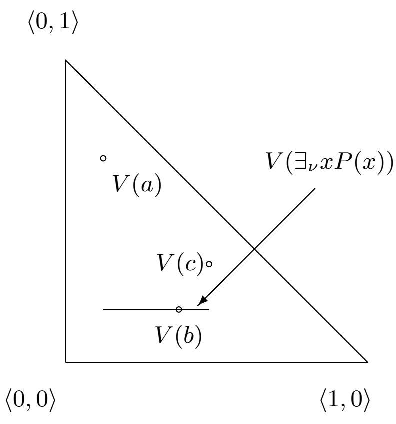



Fig. 5.

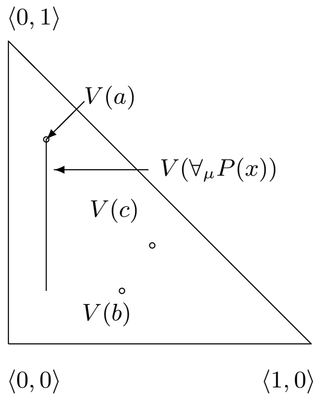



Fig. 6.

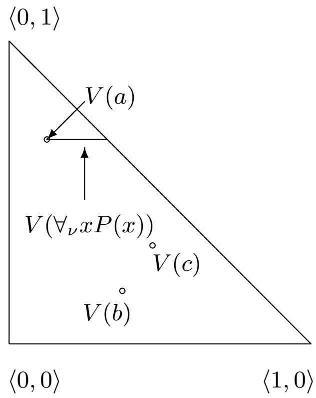



Fig. 7.

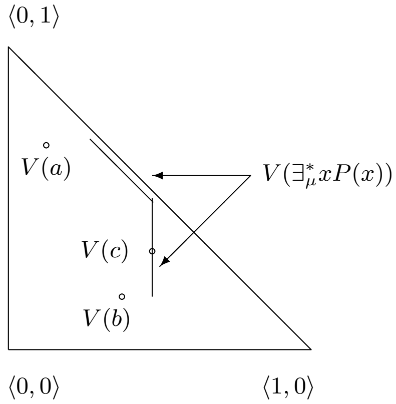



Fig. 8.

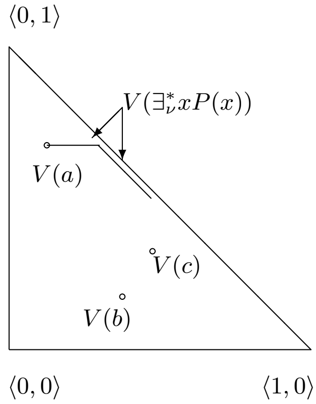



Fig. 9.

For the formula $A$ for which $V(A) = \langle a, b\rangle$, where $a, b \in [0, 1]$ and $a + b \leq 1$, following [1], we define the two modal operators "necessity" and "possibility" by $V(\Box A) = \langle a, 1-a\rangle$, $V(\Diamond A) = \langle 1-b, b\rangle$, respectively.

The geometrical interpretations of both operators are given in Figures 10 and 11, respectively.

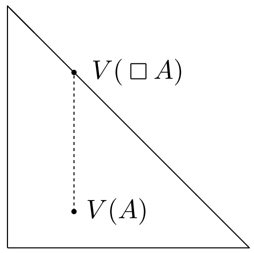



Fig. 10.

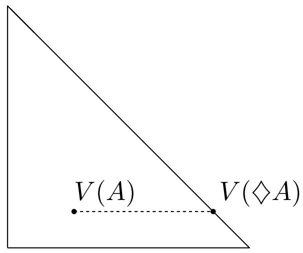



Fig. 11.

Following [7], we discuss another modal operator, without an analogue in modal logic.

For the formula $A$, for which $V(A) = \langle a, b \rangle$, where $a, b \in [0, 1]$ and $a + b \leq 1$, we define the new modal operator "$\bigcirc$" by:

$$V(\bigcirc A) = \langle \frac{a}{a+b}, \frac{b}{a+b} \rangle.$$

The first two modal operators are extended (see, e.g., [3, 4]) to operators $D_\alpha$, $F_{\alpha,\beta}$, $G_{\alpha,\beta}$, $H_{\alpha,\beta}$, $H^*_{\alpha,\beta}$, $J_{\alpha,\beta}$, $J^*_{\alpha,\beta}$ and $X_{\alpha,\beta,\gamma,\delta,\varepsilon,\eta}$, defined by:

$$\begin{aligned}
&V(D_\alpha(A)) = \langle a + \alpha.(1 - a - b), b + (1 - \alpha).(1 - a - b) \rangle,\\
&V(F_{\alpha,\beta}(A)) = \langle a + \alpha.(1 - a - b), b + \beta.(1 - a - b) \rangle, \text{ for } \alpha + \beta \leq 1,\\
&V(G_{\alpha,\beta}(A)) = \langle \alpha.a, \beta.b \rangle,\\
&V(H_{\alpha,\beta}(A)) = \langle \alpha.a, b + \beta.(1 - a - b) \rangle,\\
&V(H^*_{\alpha,\beta}(A)) = \langle \alpha.a, b + \beta.(1 - \alpha.a - b) \rangle,\\
&V(J_{\alpha,\beta}(A)) = \langle a + \alpha.(1 - a - b), \beta.b \rangle,\\
&V(J^*_{\alpha,\beta}(A)) = \langle a + \alpha.(1 - a - \beta.b), \beta.b \rangle,\\
&V(X_{\alpha,\beta,\gamma,\delta,\varepsilon,\eta}(A)) = \langle \alpha.a + \beta.(1 - a - \gamma.b), \delta.b + \varepsilon.(1 - \eta.a - b) \rangle,
\end{aligned}$$

for $\alpha + \varepsilon - \varepsilon.\eta \leq 1$, $\beta + \delta - \beta\gamma \leq 1$, $\beta + \varepsilon \leq 1$.

Their geometrical interpretations are given in Fig. 12–18.

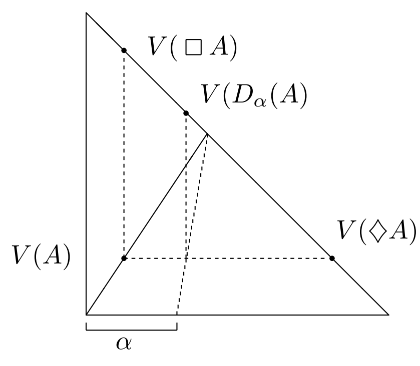



Fig. 12.

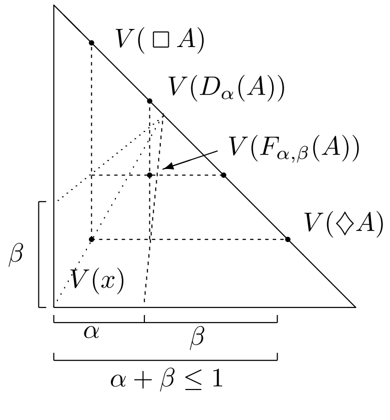



Fig. 13.

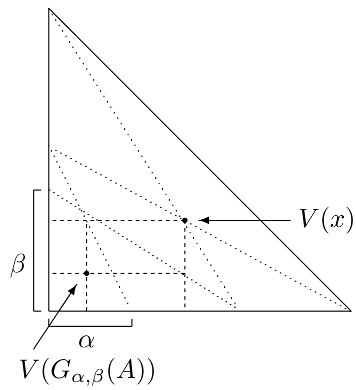



Fig. 14.

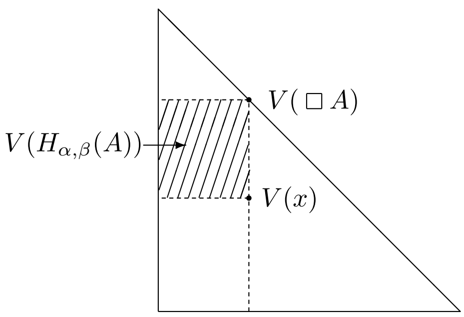



Fig. 15.

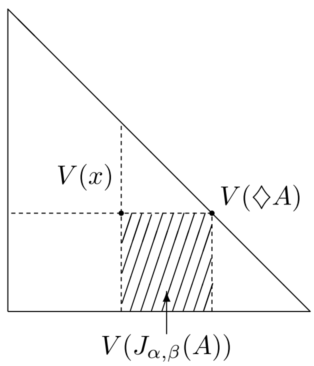



Fig. 16.

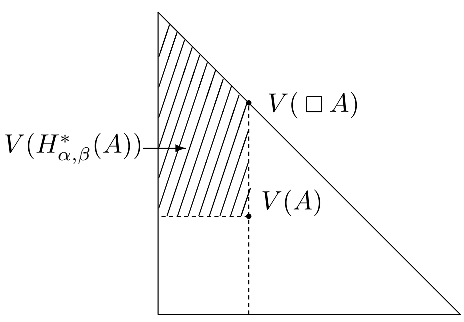



Fig. 17.

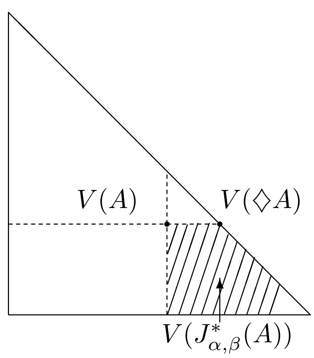



Fig. 18.

Let for operator $Y$ and for IFP $\langle a, b\rangle$: $Y(\langle a, b\rangle) = \langle a, b\rangle$. Then, we call that the IFP is a fixed point for operator $Y$. But, when operator $Y$ is defined over elements of $\mathcal{S}$, i.e., when for formula $A$, $V(Y(A)) = \langle \mu(Y(A)), \nu(Y(A))\rangle$, then we will call that $A$ is a pseudo-fixed point for operator $Y$. In this case, the equality $\langle \mu(Y(A)), \nu(Y(A))\rangle = \langle \mu(A), \nu(A)\rangle$ holds (see [10]).

Now, we formulate and prove the following assertion, related to the extended modal operators.

**Theorem 1.** For every two formulas $A, B$ there exists an operator $Y \in \{F_{\alpha,\beta}, G_{\alpha,\beta}, H_{\alpha,\beta}, H^*_{\alpha,\beta}, J_{\alpha,\beta}, J^*_{\alpha,\beta}\}$ and there exist real numbers $\alpha, \beta \in [0, 1]$ such that $V(A) = V(Y_{\alpha,\beta}(B))$.

*Proof.* Let everywhere $V(A) = \langle a, b\rangle$ and $V(B) = \langle c, d\rangle$, where $a, b, c, d \in [0, 1]$ and $a + b \leq 1$ and $c + d \leq 1$.

The following 9 cases are possible for $a, b, c$ and $d$.

| | | |
|---|---|---|
| Case 1. | $a = c$ and $b = d$ | Then, for $Y$: if $Y = F$, then, $\alpha = \beta = 0$; if $Y = G$, then, $\alpha = \beta = 1$; if $Y = H$ or $Y = H^*$, then, $\alpha = 1$ and $\beta = 0$; if $Y = J$ or $Y = J^*$ then, $\alpha = 0$ and $\beta = 1$. |
| Case 2. | $a > c$ and $b = d$ | Then, $Y = F$ and $\alpha = \frac{a-c}{1-c-d}$ and $\beta = 0$ (we shall note that $1 - c - d > 1 - a - b \geq 0$) or $Y = J$ or $Y = J^*$, and $\alpha$ has the above form and $\beta = 1$. |
| Case 3. | $a < c$ and $b = d$ | Then, $Y = G$ and $\alpha = \frac{a}{c}$ and $\beta = 1$ (we note that $c > a \geq 0$) or $Y = J$ or $Y = J^*$, and $\alpha$ has the above form and $\beta = 0$. |
| Case 4. | $a = c$ and $b > d$ | Then, $Y = F$ and $\alpha = 0$ and $\beta = \frac{b-d}{1-c-d}$ (we note that $1 - c - d > 1 - a - b \geq 0$) or $Y = H$ or $Y = H^*$, and $\alpha = 1$ and $\beta$ has the above form. |
| Case 5. | $a > c$ and $b > d$ | Then, $Y = F$ and $\alpha = \frac{a-c}{1-c-d}$ and $\beta = \frac{b-d}{1-c-d}$ (we note that $1 - c - d > 1 - a - b \geq 0$) |
| Case 6. | $a < c$ and $b > d$. | Then, there are two subcases: |
| 6.1. | | if $b \leq 1 - c$, then, $Y = H$ and $\alpha = \frac{a}{c}$ and $\beta = \frac{b-d}{1-c-d}$ or $Y = H^*$ and $\alpha = \frac{a}{c}$ and $\beta = \frac{b-d}{1-a-d}$ (we note that $1 - a - d > 1 - c - d \geq b - d > 0$ and $c > a \geq 0$) |
| 6.2. | | if $b > 1 - c$, then, $Y = H^*$ and $\alpha = \frac{a}{c}$ and $\beta = \frac{b-d}{1-a-d}$ (we note that $1 - a - d \geq b - d > 0$ and $c > a \geq 0$) |

| | | |
|---|---|---|
| Case 7. | $a = c$ and $b < d$. | Then, $Y = G$ and $\alpha = 1$ and $\beta = \frac{b}{d}$ (we note that $d > b \geq 0$)<br>or $Y = J$ or $Y = H^*$, and $\alpha = 0$ and $\beta$ has the above form. |
| Case 8. | $a > c$ and $b < d$. | Then, there are two subcases: |
| 8.1. | | if $a \leq 1 - d$, then, $Y = J$ and $\alpha = \frac{a-c}{1-c-d}$ and $\beta = \frac{b}{d}$<br>or $Y = J^*$ and $\alpha = \frac{a-c}{1-b-c}$ and $\beta = \frac{b}{d}$ (we note that $1 - c - b > 1 - c - d \geq a - c > 0$ and $d > b \geq 0$) |
| 8.2. | | if $a > 1 - d$, then, $Y = J^*$ and $\alpha = \frac{a-c}{1-b-c}$ and $\beta = \frac{b}{d}$ (we note that $1 - c - b \geq a - c > 0$ and $d > b \geq 0$) |
| Case 9. | $a < c$ and $b < d$. | Then, $Y = G$ and $\alpha = \frac{a}{c}$ and $\beta = \frac{b}{d}$ (we shall note that $c > a \geq 0$ and $d > b \geq 0$). |

The following operators of modal type are similar to the standard modal operators (see Figs. 19 and 20). Let for formula $A$: $V(A) = \langle \mu, \nu \rangle$. Then (see, e.g., [3]),

$$V(\boxplus A) = \left\langle \frac{\mu}{2}, \frac{\nu + 1}{2} \right\rangle,$$

$$V(\boxtimes A) = \left\langle \frac{\mu + 1}{2}, \frac{\nu}{2} \right\rangle.$$

In [4] these operators are extended, as follows:

$$V(\boxplus_\alpha A) = \langle \alpha\mu, \alpha\nu + 1 - \alpha \rangle,$$

$$V(\boxtimes_\alpha A) = \langle \alpha\mu + 1 - \alpha, \alpha\nu \rangle,$$

where $\alpha \in [0, 1]$,

$$V(\boxplus_{\alpha,\beta} A) = \langle \alpha\mu, \alpha\nu + \beta \rangle,$$

$$V(\boxtimes_{\alpha,\beta} A) = \langle \alpha\mu + \beta, \alpha\nu \rangle,$$

where $\alpha, \beta, \alpha + \beta \in [0, 1]$.

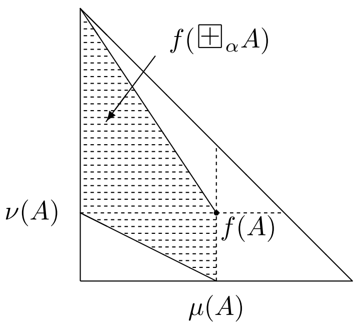



Fig. 19.

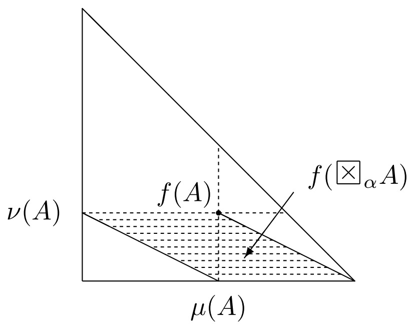



Fig. 20.

Finally, we determine all pseudo-fixed points of all quantifiers and operators, defined above. The proof of this Theorem is similar to the above one.

**Theorem 2.** For all $\alpha, \beta \in [0, 1]$ the pseudo-fixed point(s) of:

(a) $\exists$ are all elements $A \in S$ for which $V(A) = \langle 1, 0\rangle$,

(b) $\forall$ are all elements $A \in S$ for which $V(A) = \langle 0, 1\rangle$,

(c) $\exists_\mu$ are all elements $A \in S$ for which

$$\mu(A) = \sup_{x \in \mathcal{S}} \mu(x)$$

and in the more general case, all elements $A \in S$ for which $V(A) = \langle 1, 0\rangle$,

(d) $\exists_\nu$ are all elements $A \in S$ for which

$$\nu(A) = \inf_{x \in \mathcal{S}} \nu(x)$$

and in the more general case, all elements $A \in S$ for which $V(A) \in [0, 1] \times \{0\}$,

(e) $\forall_\mu$ are all elements $A \in S$ for which

$$\mu(A) = \inf_{x \in \mathcal{S}} \mu(x)$$

and in the more general case, all elements $A \in S$ for which $V(A) \in \{0\} \times [0, 1]$,

(f) $\forall_\nu$ are all elements $A \in S$ for which

$$\nu(A) = \sup_{x \in \mathcal{S}} \nu(x)$$

and in the more general case, all elements $A \in S$ for which $V(A) = \langle 0, 1\rangle$,

(g) $\Box, \Diamond, \bigcirc$ are all elements $A \in S$ for which $\mu(A) + \nu(A) = 1$,

(h) $D_\alpha$ are all elements $A \in S$ for which $\mu(A) + \nu(A) = 1$,

(i) $F_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) + \nu(A) = 1$ and $\alpha + \beta \leq 1$,

(j) $G_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) = \nu(A) = 0$,

(k) $H_{\alpha,\beta}, H^*_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) = 0$ and $\nu(A) = 1$,

(l) $J_{\alpha,\beta}, J^*_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) = 1$ and $\nu(A) = 0$,

(m) $\boxplus, \boxplus_\alpha$ are all elements $A \in S$ for which $\mu(A) = 0$ and $\nu(A) = 1$,

(n) $\boxtimes, \boxtimes_\alpha$ are all elements $A \in S$ for which $\mu(A) = 1$ and $\nu(A) = 0$,

(o) $\boxplus_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) = 0$, $\nu(A) = 1$ and $\alpha + \beta = 1$,

(p) $\boxtimes_{\alpha,\beta}$ are all elements $A \in S$ for which $\mu(A) = 1$, $\nu(A) = 0$ and $\alpha + \beta = 1$,

(q) $P_{\alpha,\beta}$ are all elements $A \in S$ for which $\alpha \leq \mu(A) = 1$ and $0 \leq \nu(A) \leq \beta$,

(r) $Q_{\alpha,\beta}$ are all elements $A \in S$ for which $0 \leq \mu(A) = \alpha$ and $\beta \leq \nu(A) \leq 1$.

## 3 Conclusion

In intuitionistic fuzzy sets theory there are some operators that by the moment do not have analogue in intuitionistic fuzzy logics. In future, we hope that they will be defined for intuitionistic fuzzy logic case and their properties, including these, discussed in the present paper, will be studied.


## Acknowledgements

The author is grateful for the support provided by the project Ref. No. DFNI-I-02-5 funded by the Bulgarian Science Fund.


## References


1. Atanassov K. T. Intuitionistic Fuzzy Sets, 7th Youth Scientific Session, ITCR-BAS, Sofia, 20-23 July 1983. Reprinted: Int. J. Bioautomation, 2015, 19(4), Suppl. 2, S131-S136.
2. Atanassov K. T. Two Variants of Intuitionistic Fuzzy Propositional Calculus, Mathematical Foundations of Artificial Intelligence Seminar, Sofia, 1988, Preprint IM-MFAIS-5-88. Reprinted: Int. J. Bioautomation, 2015, 19(4), Suppl. 2, S146-S155.
3. Atanassov, K. Intuitionistic Fuzzy Sets, Springer, Heidelberg, 1999.
4. Atanassov, K. On Intuitionistic Fuzzy Sets Theory, Springer, Berlin, 2012.
5. Atanassov, K. On Intuitionistic Fuzzy Logics: Results and Problems. In: Modern Approaches in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics, Volume 1: Foundations. (Atanassov, K., M. Baczynski, J. Drewniak, J. Kacprzyk, M. Krawczak, E. Szmidt, M. Wygralak, S. Zadrozny, eds.), SRI-PAS, Warsaw, 2014, 23-49.

6. Atanassov, K. Intuitionistic fuzzy logics as tools for evaluation of Data Mining processes, Knowledge-Based Systems, Vol. 80, 2015, 122-130.
7. Atanassov, K. A new topological operator over intuitionistic fuzzy sets. Notes on Intuitionistic Fuzzy Sets, Vol. 21, 2015, No. 3, 90-92.
8. Atanassov, K. Uniformly expanding intuitionistic fuzzy operator, Notes on Intuitionistic Fuzzy Sets. Vol. 22, 2016, No. 1, 48-52.
9. Atanassov, K., G. Gargov. Elements of intuitionistic fuzzy logic. I. Fuzzy sets and Systems Vol. 95, 1998, No. 1, 39-52.
10. Atanassov, K., N. Nikolov. Remark on the pseudo-fixed points of the intuitionistic fuzzy logic operators. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems Vol. 9, 2001, No. 1, 77-80.
11. Atanassov, K., E. Szmidt, J. Kacprzyk. On intuitionistic fuzzy pairs, Notes on Intuitionistic Fuzzy Sets, Vol. 19, 2013, No. 3, 1-13.
12. Gargov, G., K. Atanassov. Two results in intuitionistic fuzzy logic, Comptes Rendus de l'Academie bulgare des Sciences, Tome 45, 1992, No. 12, 29-31.
13. Mendelson, E. Introduction to Mathematical Logic, Princeton, NJ: D. Van Nostrand, 1964.
14. Shoenfield, J. R. Mathematical Logic (2nd ed.), Natick, MA, A. K. Peters, 2001.
15. Zadeh, L. Fuzzy logics, Computer, 1988, 83-93.
16. http://plato.stanford.edu/entries/generalized-quantifiers/

# On a New Ordering between Intuitionistic Fuzzy Pairs

Peter Vassilev and Todor Stoyanov

Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences, Acad. G. Bonchev Str.
Bl. 105, Sofia-1113, Bulgaria
peter.vassilev@gmail.com

**Abstract.** In this paper we investigate orderings between intuitionistic fuzzy pairs and find some relationships between them. Our purpose is to establish a meaningful automated way to determine an object which best fits a given object, when this specific object is compared to all others and the result of this comparison is in the form of intuitionistic fuzzy pairs. Our idea is to obtain a way of comparison which may be used to identify potentially similar structures, texts, geometric patterns, which can later be processed by other methods to verify or discard this initial hypothesis.

**Keywords:** intuitionistic fuzzy pairs, similarity, orderings.

## 1 Introduction

Let us suppose that we are given an object (e.g. an input string, geometric shape) $I$ and set of objects (of the same nature) $O = \{O_1, \ldots, O_n\}$, and a function $f$ that assigns similarity score to every pair $(I, O_j)$. When this score is a real number, we can easily find the objects that are most similar to $I$, due to the natural ordering of the real numbers which allows comparison of any two of them. However, if the function assigns an ordered pair of real numbers to each couple of objects, which reflect similarity and dissimilarity in some sense, we are faced with a particular problem, since only partial ordering exists and some elements are not comparable. There are two ways in which such problem can be overcome - one is to use algorithms to generate only comparable pairs. The other is to introduce different orderings which whenever possible imply similar behaviour. In a manner of speaking we are looking for something like an analytic continuation of a function, but applied to the ordering. This allows for greater freedom when considering the function (or algorithm) that assigns the IFP as the range of the function is thus extended.

## 2 Preliminaries

Here we remind some basic definitions which will be used further.

**Definition 1 (cf. [2]).** *An intuitionistic fuzzy pair (IFP) is an ordered couple of real non-negative numbers $\langle a, b\rangle$, with the additional constraint:*

$$a + b \leq 1. \tag{1}$$

This concept is very important in practice since many methods implementing intuitionistic fuzzy techniques lead to estimates in the form of IFPs.

One way to measure which of two IFPs is "better" is by using an ordering:

**Definition 2 (cf. [1, 2]).** *Given two IFPs:* $u = \langle u_1, u_2 \rangle$ *and* $v = \langle v_1, v_2 \rangle$*, we say that* $u$ *is less or equal to* $v$*, and we write:*

$$u \le v,$$

*iff*

$$\begin{cases} u_1 \le v_1 \\ u_2 \ge v_2. \end{cases} \tag{2}$$

It is obvious that $\le$ is a partial ordering, because from its definition it is evident that it is transitive, reflexive and antisymmetric but there exist $u$ and $v$, for which conditions (2) are not satisfied.

**Definition 3.** *We will say that a partial ordering* $\preceq_i$ *is stricter than partial ordering* $\preceq_j$*, when* $u \preceq_i v$*, implies* $u \preceq_j v$*, but not vice versa.*

A partial ordering stricter than $\le$ is the following:

**Definition 4 (cf.[3]).** *Given two IFPs* $u = \langle u_1, u_2 \rangle$ *and* $v = \langle v_1, v_2 \rangle$*, we say that* $u$ *is (definiteness-based) less or equal to* $v$*, and we write:*

$$u \preceq_\sigma v,$$

*iff*

$$\begin{cases} (u_1 + u_2)u_1 \le (v_1 + v_2)v_1 \\ (u_1 + u_2)u_2 \ge (v_1 + v_2)v_2. \end{cases} \tag{3}$$

Further we will propose a new ordering which may be viewed as a continuation of the above orderings.

## 3 The new partial ordering

Here we will propose a new partial ordering.

**Definition 5.** *Given two IFPs* $u = \langle u_1, u_2 \rangle$ *and* $v = \langle v_1, v_2 \rangle$*, we say that* $u$ *is (first-component biased) less or equal to* $v$*, and we write:*

$$u \preceq_\mu v$$

*iff*

$$\begin{cases} u_1 \le v_1 \\ u_2(1 - u_1) \ge v_2(1 - v_1). \end{cases} \tag{4}$$

The fact that $u \preceq_\mu v$ is transitive, reflexive and antisymmetric is evident from the definition.

**Theorem 1.** *The partial ordering $\leq$ is stricter than $\preceq_\mu$ .*

*Proof.* First we will show that $u \leq v$, implies $u \preceq_\mu v$.

From (2) we have $u_2 \geq v_2$. From the first (identical for both cases) inequality we have:

$$u_1 \leq v_1 \Leftrightarrow (1 - u_1) \geq (1 - v_1).$$

Thus, multiplying both inequalities we obtain:

$$u_2(1 - u_1) \geq (1 - v_1)v_2.$$

Hence, $u \leq v$, implies $u \preceq_\mu v$.

It remains to show that $u \preceq_\mu v$ does not imply $u \leq v$.

Let us consider $u$ and $v$ such that $u_1 < v_1$ and

$$u_2(1 - u_1) = (1 - v_1)v_2.$$

Then, we must have (assuming $v_2 \neq 0$ and $u_1 \neq 1$):

$$\frac{u_2}{v_2} = \frac{(1 - v_1)}{(1 - u_1)}$$

But $(1 - v_1) < (1 - u_1)$, hence $u_2 < v_2$. Thus, $u \preceq_\mu v$ but $u \not\leq v$. □

As can be seen from the proof, more IFPs are comparable with the new ordering.

We will briefly explain our motivation in its construction. The first component (if the pair is an intuitionistic fuzzy estimate) usually has the meaning of similarity, agreement or compatibility. Hence, in determining which is the closest ("most similar") element it should have a leading role. The second condition in $\preceq_\mu$, reflects the "degree of disagreement" scaled with the maximal theoretically possible "degree of disagreement" for each pair.

Let us suppose that we have a predefined object $I$ and a set of objects $O = \{O_1, ..., O_n\}$. If we have a function $f$ which measures the similarity between the objects and is such that:

$$f(I, O_i) = f(O_i, I) = \langle a_i, b_i \rangle,\ i = 1, \ldots, n.$$

If for some $k$, $\langle a_k, b_k \rangle$, we have:

$$\langle a_i, b_i \rangle \preceq_\mu \langle a_k, b_k \rangle, i = 1, \ldots, n,$$

then we can conclude that $O_k$ is potentially the most similar to $I$.

Note that since we have that if two elements are comparable by $\preceq_\sigma$ and $u \preceq_\sigma v$, implies $u \leq v$, and $u \leq v$ in turn implies $u \preceq_\mu v$, then $u \preceq_\mu v$ preserves the ordering for the elements sorted by the stricter relations.

## 4 Conclusion

In the present work we have proposed a new partial ordering which is consistent with $\preceq_\sigma$ and $\leq$, whenever all are applicable. It has, however, the benefit of being applicable whenever the above mentioned are not.

## Acknowledgment

Work presented here is supported by Program for career development of young scientists, Bulgarian Academy of Sciences, Grant number DFNP-123/2016 "Finding approximate matching of string or parts of biomedical signals with defined models with the help of intuitionistic fuzzy estimates".

## References

1. Atanassov, K. T. *On Intuitionistic Fuzzy Sets Theory.* Springer, Berlin, 2012
2. Atanassov, K., E. Szmidt and J. Kacprzyk, "On Intuitionistic Fuzzy Pairs", *Notes on Intuitionistic Fuzzy Sets*, Vol. 19, No. 3, 2013, 1–13.
3. Vassilev, P. A note on new partial ordering over intuitionistic fuzzy sets. Annual of Informatics Section Union of Scientists in Bulgaria, 8, 2015 (to appear)

# Generalized Net for Coordination and Synchronization of Human and Computer-based Expert Type Decision Support Activities in a Multiagent Setting

Eulalia Szmidt[1], Evdokia Sotirova[2], and Janusz Kacprzyk[1], and Krassimir Atanassov[2,3]

[1] Systems Research Institute – Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
{szmidt,kacprzyk}@ibspan.waw.pl
[2] "Prof. Asen Zlatarov" University
Prof. J. Jakimov Str. 1, Bourgas-8010, Bulgaria
esotirova@btu.bg
[3] Dept. of Bioinformatics and Mathematical Modelling
Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
krat@bas.bg

**Abstract.** The processes of decision making by experts and by decision making tools are described by a Generalized Net (GN). It is shown that the GN-model can organize, coordinate and synchronize the work of the experts and/or decision making tools with aim to obtain the best results of their activity.

**Keywords:** decision making tool, expert system, generalized net, multiagent system

## 1 Introduction

The Generalized Nets (GNs, [2, 4] are tools for modelling of parallel processes, including as partial cases the standard Petri nets and all their modifications and extensions (as Time Petri nets, E-nets, Colour Petri nets, Predicative-Transition Petri nets, Fuzzy Petri nets, etc.). These proofs are published in *"AMSE Press"* journals in 1980s and discussed in [2]. Here, we give an illustration: by Time Petri nets we can describe the functioning of a cinema hall with projections in 10 AM, 12 AM, 2 PM, etc. By E-nets we can describe the functioning of the same cinema hall with projections with duration 1:45, 1:35, etc., hours long. In none of these cases, the respective type of nets can show us what is the interval betweet two of projections. GNs, having the two temporal components of the two Petri net types, can give us the respective information.

The apparatus of the GNs is used for modelling of different processes in the areas of Artificial Intelligence, medicine and biology, economics, industry and many others (see, e.g., [7]).

In Section 2, we give short remarks from GN-theory, in Section 3 – description of a GN that organizes, coordinates and synchronizes the work of experts and/or decision making tools united in a multiagent setting for solving of concrete problem. The GN-model is based on some criteria that determine: the choice of experts and/or decision making tools, the correctness of the evaluations and of the final evaluations. As the process of expert decision making, as well as the processes of functioning of different decision making tools, are represented by GNs in [8]. So, the present GN can be interpreted as a GN over which hierarchical operators are defined (see [2, 4]).

The paper is a continuation and illustration of the ideas from [7].

## 2 Short remarks of the theory of the generalized nets

The GNs are defined in a way that is different in principle from the ways of defining the other types of PNs [2, 4].

The first basic difference between GNs and the ordinary PNs is the "place – transition" relation. Here, the transitions are objects of more complex nature. A transition may contain $m$ input places and $n$ output places, where the integers $m, n \geq 1$.

Formally, every transition is described by a seven-tuple (Figure 1):

$$Z = \langle L', L'', t_1, t_2, r, M, \square \rangle,$$

where:

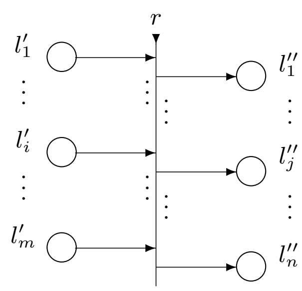



**Fig. 1.** The form of transition

**(a)** $L'$ and $L''$ are finite, non-empty sets of places (the transition's input and output places, respectively); for the transition in Figure 1 these are $L' = \{l'_1, l'_2, \ldots, l'_m\}$ and $L'' = \{l''_1, l''_2, \ldots, l''_n\}$;

**(b)** $t_1$ is the current time-moment of the transition's firing;

**(c)** $t_2$ is the current value of the duration of its active state;

**(d)** $r$ is the transition's *condition* determining which tokens will pass (or *transfer*) from the transition's inputs to its outputs; it has the form of an Index Matrix (IM; see [6]):

$$r = \begin{array}{c|ccccc} & l''_1 & \ldots & l''_j & \ldots & l''_n \\ \hline l'_1 & & & & & \\ \vdots & & & r_{i,j} & & \\ l'_m & & & & & \end{array};$$

$r_{i,j}$ is the predicate that corresponds to the $i$-th input and $j$-th output place $(1 \leq i \leq m, 1 \leq j \leq n)$. When its truth value is "*true*", a token from the $i$-th input place transfers to the $j$-th output place; otherwise, this is not possible;

**(e)** $M$ is an IM of the capacities $m_{i,j}$ of transition's arcs, where $m_{i,j} \geq 0$ is a natural number:

$$M = \begin{array}{c|ccccc} & l''_1 & \ldots & l''_j & \ldots & l''_n \\ \hline l'_1 & & & & & \\ \vdots & & & m_{i,j} & & \\ l'_m & & & & & \end{array};$$

**(f)** $\Box$ is an object of a form similar to a Boolean expression. It contains as variables the symbols that serve as labels for a transition's input places, and $\Box$ is an expression built up from variables and the Boolean connectives $\wedge$ and $\vee$. When the value of a type (calculated as a Boolean expression) is "*true*", the transition can become active, otherwise it cannot.

The ordered four-tuple

$$E = \langle\langle A, \pi_A, \pi_L, c, f, \theta_1, \theta_2\rangle, \langle K, \pi_K, \theta_K\rangle, \langle T, t^o, t^*\rangle, \langle X, \Phi, b\rangle\rangle$$

is called a GN if:

**(a)** $A$ is a set of transitions;

**(b)** $\pi_A$ is a function giving the priorities of the transitions, i.e., $\pi_A : A \to N$, where $N = \{0, 1, 2, \ldots\} \cup \{\infty\}$;

**(c)** $\pi_L$ is a function giving the priorities of the places, i.e., $\pi_L : L \to N$, where $L = pr_1 A \cup pr_2 A$, and $pr_i X$ is the $i$-th projection of the $n$-dimensional set, where $n \in N, n \geq 1$ and $1 \leq k \leq n$ (obviously, $L$ is the set of all GN - places);

**(d)** $c$ is a function giving the capacities of the places, i.e., $c : L \to N$;

**(e)** $f$ is a function that calculates the truth values of the predicates of the transition's conditions (for the GN described here, let the function $f$ have the value "*false*" or "*true*", that is, a value from the set $\{0, 1\}$;

**(f)** $\theta_1$ is a function which indicates the next time-moment when a certain transition $Z$ can be activated, that is, $\theta_1(t) = t'$, where $pr_3 Z = t, t' \in [T, T + t^*]$ and $t \leq t'$. The value of this function is calculated at the moment when the transition ceases to function;

**(g)** $\theta_2$ is a function which gives the duration of the active state of a certain transition $Z$, i. e., $\theta_2(t) = t'$, where $pr_4 Z = t \in [T, T + t^*]$ and $t' \geq 0$. The value of this function is calculated at the moment when the transition starts to function;

**(h)** $K$ is the set of the GN's tokens.

**(i)** $\pi_K$ is a function which gives the priorities of the tokens, that is, $\pi_K : K \to N$;

**(j)** $\theta_K$ is a function which gives the time-moment when a given token can enter the net, that is, $\theta_K(\alpha) = t$, where $\alpha \in K$ and $t \in [T, T + t^*]$;

**(k)** $T$ is the time-moment when the GN starts to function. This moment is determined with respect to a fixed (global) time-scale;

**(l)** $t^o$ is an elementary time-step, related to the fixed (global) time-scale;

**(m)** $t^*$ is the duration of the functioning of the GN;

**(n)** $X$ is the set of all initial characteristics which the tokens can obtain on entering the net;

**(o)** $\Phi$ is the characteristic function that assigns new characteristics to every token when it makes the transfer from an input to an output place of a given transition.

**(p)** $b$ is a function which gives the maximum number of characteristics a given token can obtain, that is, $b : K \rightarrow N$.

A given GN may not have some of the above components. In these cases, any redundant component will be omitted. The GNs of this kind form a special class of GNs called *"reduced GNs"*.

## 3 Generalized net model of a decision making process

Let us have the experts $E_1, ..., E_e$ who participate in procedures for decision making. Let the $i$-th expert have score $\langle \rho_i, \sigma_i \rangle$, where $\rho_i$ is its degree of correctness (precision, ...) and $\sigma_i$ – its degree of incorrectness (imprecision, ...) and they correspond to the number of the expertises in which the expert $E_i$ has given correct or incorrect estimation, respectively. If the expert $E_i$ had participated in $N_i$ expertises in the past and if his/her evaluation in the $(N_i + 1)$-st expertise is correct, then his/her score will be changed to the form

$$\langle \frac{N_i \rho_i + 1}{N_i + 1}, \frac{N_i \sigma_i}{N_i + 1} \rangle,$$

if his/her evaluation in the $(N_i + 1)$-st expertise is incorrect, then his/her score will be changed to the form

$$\langle \frac{N_i \rho_i}{N_i + 1}, \frac{N_i \sigma_i + 1}{N_i + 1} \rangle,$$

while if he/she had not diven any evaluation in the $(N_i + 1)$-st expertise, then his/her score will be changed to the form

$$\langle \frac{N_i \rho_i}{N_i + 1}, \frac{N_i \sigma_i}{N_i + 1} \rangle.$$

In our GN-model, each expert will be represented by a token ($E_i$ – by $\varepsilon_i$). Initially, all tokens $\varepsilon_1, ..., \varepsilon_e$ stay in place $l_5$. When a problem for solving arises, the model will determine which experts must participate in the process, if they are available. In the present model, we assume that if an expert participates in another procedure, the present problem will wait for him/her to finish his/her previous activity and the work over the new problem will start when the necessary expert is already available. In next research, we will discuss the possibility for replacing of one expert with another. Here, we assume that the expert always gives correct evaluations, i.e., each his/her evaluation $\langle \alpha_{i,k}, \beta_{i,k} \rangle$ for the problem $\pi_k$, satisfies the inequality $\alpha_{i,k} + \beta_{i,k} \leq 1$, where $\alpha_{i,k}$ and $\beta_{i,k}$ are the positive and negative degrees of the evaluation of expert $E_i$.

In future research, we will discuss the case, when we will correct the expert estimations, when they are wrong, using the procedures from [3, 5].

Let us have different decision making tools $T_{1,1}, ..., T_{1,t_1}, ..., T_{n,1}, ..., T_{n,t_n}$, where $T_{1,1}, ..., T_{1,t_1}$ can be, e.g., $t_1$ different data bases, $T_{2,1}, ..., T_{2,t_2}$ – $t_2$ different expert systems, etc. Among these decision making tools there will be different types of neural networks, genetic algorithms, ant colony optimization procedures, and others.

In the GN-model, each tool will be represented by a token ($T_{i,j}$ – by $\tau_{i,j}$, where $1 \leq j \leq t_i$). Let us assume that each tool gives (in the present model – always correct) evaluation $\langle \gamma_{i,j,k}, \delta_{i,j,k} \rangle$, for the problem $\pi_k$, where $\gamma_{i,j,k} + \delta_{i,j,k} \leq 1$.

These tokens stay in places $m_1, ..., m_n$.

There are four tools that we call “solvers” (criteria) that solve concrete situations in the global process - $\kappa_1$, $\kappa_2$, $\kappa_3$ and $\kappa_4$. The first of them, using necessary criteria, determines which decision making tools are necessary for solving problem $\pi$. It stays in place $l_7$. The second solver checks whether the joint result of work of the decision making tools that are necessary for solving the current problem, is correct. It stays in place $l_{13}$. The third solver checks whether the final results of work of the experts or of the decision making tools, are correct. It stays in place $l_{15}$. The fourth solver is activated when the third solver determines that the final result is not correct. The fourth solver determines who - other experts or decision making tools shall repeat the decision making procedure. It stays in place $l_{19}$.

There are $a$ aggregation algorithms $A_1, ..., A_a$ that aggregate expert's and decision making tool's evaluations for the current problem. They are represented by tokens $\alpha_1, ..., \alpha_a$ that stay in place $l_{15}$.

In separate time-moments, tokens $\pi_1, \pi_2, ...$ enter the GN. They represent the separate problems that must be solved by experts or decision making tools. For simplicity, the current $k$-th token $\pi_k$ will be denoted by $\pi$. Each of these tokens enters place $l_1$ with initial characteristic “problem; list of suitable experts or decision making tools who can solve the problem”.

The GN-model (see Fig. 2) contains 7 transitions, $19 + n$ places and the above described types of tokens.

The GN-transitions are the following.

$$Z_1 = \langle \{l_1, l_{17}\}, \{l_2, l_3\},$$

$$\begin{array}{c|cc} & l_2 & l_3 \\ \hline l_1 & W_{1,2} & W_{1,3} \\ l_{17} & W_{17,2} & W_{17,3} \end{array} \rangle,$$

where:

- $W_{1,2} = W_{17,2} =$“the decision making process uses experts”;
- $W_{1,3} = W_{17,3} =$“the decision making process uses decision making tools”.

Token $\pi$ from place $l_1$ enters place $l_2$ with initial characteristic “pre-formulation of the problem in a form suitable for experts” and it enters place $l_3$ with initial characteristic “pre-formulation of the problem in a form suitable for decision making tools”, while token $\pi$ from place $l_{17}$ enters places $l_2$ and $l_3$ without any new characteristic, if it will use other experts or other decision making tools, and it will obtain as a next characteristic

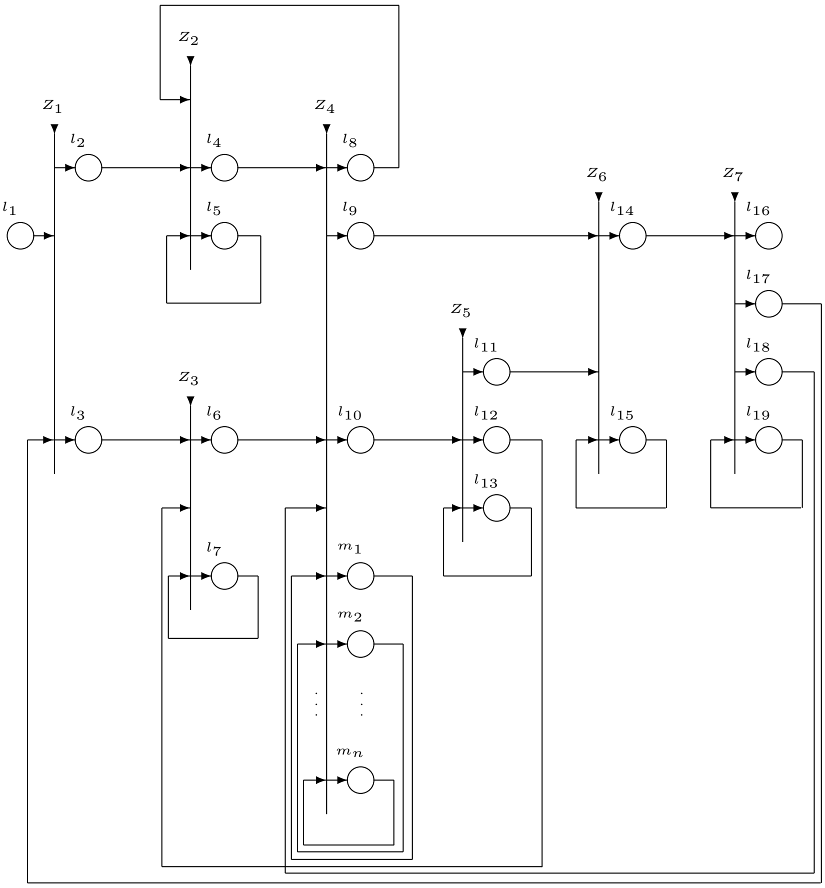



Fig. 2.

| | |
|---|---|
| "pre-formulation of the problem in suitable for experts form;, list of the suitable experts", | if on the previous cycles it used decision making tools |
| "pre-formulation of the problem in suitable for decision making programs form", | if on the previous cycles it used experts |

$$Z_2 = \langle \{l_2, l_5, l_8\}, \{l_4, l_5\},$$

$$\begin{array}{c|cc} & l_4 & l_5 \\ \hline l_2 & W_{2,4} & false \\ l_5 & W_{5,4} & true \\ l_8 & false & true \end{array} \rangle,$$

where:

- $W_{2,4}$ ="the experts from the last token characteristic are available in place $l_5$",
- $W_{5,4}$ ="there is a token in place $l_2$".

Let tokens $\varepsilon_{q_1},...,\varepsilon_{q_s}$ represent the experts who are necessary for the solution of problem $\pi$.

Token $\pi$ from place $l_2$ enters place $l_4$ where it unites with tokens $\varepsilon_{q_1},...,\varepsilon_{q_s}$ arriving from place $l_5$, and the new token obtains the characteristic "list of scores of experts $\varepsilon_{q_1},...,\varepsilon_{q_s}$".

Tokens from place $l_8$ enter place $l_5$ without new characteristics.

$$Z_3 = \langle \{l_3, l_7, l_{12}\}, \{l_6, l_7\},
\begin{array}{c|cc}
 & l_6 & l_7 \\
\hline
l_3 & W_{3,6} & false \\
l_7 & W_{7,6} & true \\
l_{12} & false & true
\end{array} \rangle,$$

where:

- $W_{3,6}$ ="the decision making tools from the last token characteristic are available in place $l_7$",
- $W_{7,6}$ ="there is a token in place $l_3$".

Token $\kappa_1$, that has a higher priority than the $\pi$-tokens, first makes a loop within place $l_7$ and this action represents the process of determining of the decision making tools that are necessary for the solution of problem $\pi$.

The $\pi$-tokens from places $l_3$ or $l_{17}$ enter place $l_6$ with the characteristic "list of decision making tools $\tau_{i_1,j_1},...,\tau_{i_r,j_r}$ that are necessary for solving of problem $\pi$".

$$Z_4 = \langle \{l_4, l_6, l_{18}, m_1, ..., m_n\}, \{l_8, l_9, l_{10}, m_1, ..., m_n\},
\begin{array}{c|cccccc}
 & l_8 & l_9 & l_{10} & m_1 & ... & m_n \\
\hline
l_4 & true & true & false & false & \dots & false \\
l_6 & false & false & true & false & \dots & false \\
l_{18} & false & false & false & W_{18,1} & \dots & W_{18,m} \\
m_1 & false & false & false & true & \dots & false \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
m_n & false & false & false & false & \dots & true
\end{array} \rangle,$$

where for each $i$: $(1 \leq i \leq m)$, $W_{18,i}$ ="a decision making tool from the current type had participated in the decision making process related to current token $\pi$".

Token $\pi$ from place $l_4$ splits to $s+1$ tokens – tokens $\varepsilon_{q_1},...,\varepsilon_{q_s}$ and the original token $\pi$. Tokens $\varepsilon_{q_1},...,\varepsilon_{q_s}$ (that have united with token $\pi$ on the previous time-step) enter place $l_8$ without new characteristics. In a future research of the authors, we will study the number of participations of the separate decision making tools in the processes of problems solving. So, the $\varepsilon$-tokens will obtain as a characteristic in place $l_8$

the number of current participation. Token $\pi = \pi_k$ enters place $l_9$ with a characteristic "$\{\langle\alpha_{q_1,k}, \beta_{q_1,k}\rangle, ..., \langle\alpha_{q_s,k}, \beta_{q_s,k}\rangle$; the concrete aggregation procedure for results of work of the experts for solving problem $\pi$".

The reason behind the processes of uniting and splitting of the $\varepsilon$-tokens is that human experts need longer time for their process of decision making, compared to decision making tools. Thus, the tokens that represent decision making tools will not unite or split with token $\pi$. So, token $\pi = \pi_k$ from place $l_6$ enters place $l_{10}$ with a characteristic "$\{\langle\gamma_{i_1,j_1,k}, \delta_{i_1,j_1,k}\rangle, ..., \langle\gamma_{i_r,j_r,k}, \delta_{i_r,j_r,k}\rangle\}$", where the evaluations $\langle\gamma_{i_1,j_1,k}, \delta_{i_1,j_1,k}\rangle$, ..., $\langle\gamma_{i_r,j_r,k}, \delta_{i_r,j_r,k}\rangle$ are given by decision making tools $\tau_{i_1,j_1}, ..., \tau_{i_r,j_r}$, as determined by criterion $\kappa_1$.

In the last time-moment of the GN-working over current token $\pi$, token $\pi'$ from place $l_{18}$ splits to $r$ equal tokens that enter those of places, which contain $\tau$-tokens (i.e., decision making tools) that participated in solving the problem, represented by $\pi$. There, the tokens, generated by the token $\pi'$, unite with the concrete $\tau$-tokens, adding as a new characteristic of the $\tau$-tokens the current number of the procedure, in which the respective token participated.

$$Z_5 = \langle\{l_{10}, l_{13}\}, \{l_{11}, l_{12}, l_{13}\},\ \begin{array}{c|ccc} & l_{11} & l_{12} & l_{13} \\ \hline l_{10} & W_{10,11} & W_{10,12} & false \\ l_{13} & false & false & true \end{array}\rangle,$$

where:

- $W_{10,11} =$"the current characteristic of the token $\kappa_2$ is positive",
- $W_{10,12} = \neg W_{10,11}$, where $\neg P$ is the negation of predicate $P$.

Token $\kappa_2$, that has higher priority than $\pi$-tokens, first makes a loop within place $l_{13}$ and this action represents the process of checking whether the joint results of work of the decision making tools that are necessary for solving problem $\pi$, are correct. As we mentioned above, in the present research, we assume that each of these tools work correctly, but there is no guarantee that a combination of all results continues to be correct. So, token $\kappa_2$ obtains as a current characteristic "evaluation of the joint results of work of the decision making tools for solving problem $\pi$".

When the current characteristic of the token $\kappa_2$ is positive, the $\pi$-token from place $l_{10}$ enters place $l_{11}$ with the characteristic "the concrete aggregation procedure for results of work of the decision making tools for solving problem $\pi$".

In the opposite case, it enters place $l_{12}$ without a new characteristic.

$$Z_6 = \langle\{l_9, l_{11}, l_{15}\}, \{l_{14}, l_{15}\},\ \begin{array}{c|cc} & l_{14} & l_{15} \\ \hline l_9 & true & false \\ l_{11} & true & false \\ l_{15} & false & true \end{array}\rangle.$$

Token $\kappa_3$, that has higher priority than $\pi$-tokens, first makes a loop within place $l_{15}$ and this action represents the process of aggregation of the experts' or decision making tools's estimations for the current problem $\pi$.

The $\pi$-tokens from places $l_9$ or $l_{11}$ enter place $l_{14}$ with the characteristic "final (aggregated) evaluation (solution) of problem $\pi$".

$$Z_7 = \langle \{l_{14}, l_{19}\}, \{l_{16}, l_{17}, l_{18}, l_{19}\},$$
$$\begin{array}{c|cccc} & l_{16} & l_{17} & l_{18} & l_{19} \\ \hline l_{14} & W_{14,16} & W_{14,17} & W_{14,18} & false \\ l_{19} & false & false & false & true \end{array} \rangle,$$

where:

- $W_{14,16} = W_{14,18} =$"the current characteristic of the token $\kappa_4$ is positive",
- $W_{14,17} = \neg W_{14,16}$.

Token $\kappa_4$, that has higher priority than $\pi$-tokens, first makes a loop within place $l_{19}$ and this action represents the process of check whether the final result for the current problem $\pi$ is correct or not. It obtains the characteristic "final evaluation (solution) of current problem $\pi$".

When the current characteristic of the token $\kappa_4$ is positive, the $\pi$-token from place $l_{14}$ splits to two tokes $\pi$ and $\pi'$ that enter places $l_{16}$ and $l_{18}$ with the characteristic "final (correct) evaluation (solution) of problem $\pi$".

When the current characteristic of the token $\kappa_4$ is negative, the $\pi$-token from place $l_{14}$ enters place $l_{17}$ to repeat the process of decision making. It obtains the characteristic "list of new experts or decision making tools that can solve the problem".

## 4 Conclusion

The so constructed GN-model can be used for simulation of different situations that can arise in the process of decision making in a multiagent setting containing human and computer based expert type decision making tools.

## References


1. Alexieva, J., E. Choy, E. Koycheva, Review and bibloigraphy on generalized nets theory and applications, In A Survey of Generalized Nets (E. Choy, M. Krawczak, A. Shannon and E. Szmidt, Eds.), Raffles KvB Monograph No. 10, 2007, 207–301.
2. Atanassov, K., Generalized Nets, World Scientific. Singapore, London, 1991.
3. Atanassov, K., Intuitionistic Fuzzy Sets, Springer, Heidelberg, 1999.
4. Atanassov, K., On Generalized Nets Theory, Prof. M. Drinov Acad. Publ. House, Sofia, 2007.
5. Atanassov, K., On Intuitionistic Fuzzy Sets Theory, Springer, Berlin, 2012.
6. Atanassov, K., Index Matrices: Towards an Augmented Matrix Calculus, Springer, Cham, 2014.
7. Atanassov, K., Generalized Nets as a Tool for the Modelling of Data Mining Processes. In :- Innovative Issues in Intelligent Systems (V. Sgurev, R. Yager, J. Kacprzyk, V. Jotsov, Eds.), Springer, Cham, 2016, 161-215.
8. Atanassov, K., G. Gluhchev, S. Hadjitodorov, J. Kacprzyk, A. Shannon, E. Szmidt, V. Vassilev. Generalized Nets Decision Making and Pattern Recognition. Warsaw School of Information Technology, Warszawa, 2006.

# Enriching Prolog Systems by Means of Interval-valued Fuzzy Sets

Clemente Rubio-Manzano[1] and Martin Pereira-Fariña[2]

[1] Department of Information Systems, University of the Bío-Bío, Chile
clrubio@ubiobio.cl
[2] Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), Universidade de Santiago de Compostela, Spain
martin.pereira@usc.es

**Abstract.** In this paper we analyze the benefits of incorporating interval-valued fuzzy sets into Prolog systems. A syntax, semantics and implementation for this extension is presented and formalized. The main application of our approach is to show that fuzzy logic programming frameworks can work together with lexical resources and ontologies in order to improve their capabilities for knowledge representation and reasoning.

**Keywords.** interval-valued fuzzy sets, approximate reasoning, lexical knowledge resources, fuzzy logic programming, fuzzy Prolog.

## 1 Introduction and motivation

Nowadays, lexical knowledge resources as well as ontologies of concepts are widely employed for modelling domain independent knowledge [8, 9] and by automated reasoners [1]. This makes possible to add, more or less easily, general knowledge automatically and independently of the designer into an approximate reasoning system.

Inside the frames of fuzzy logic programming [13, 12, 11, 6], we argue that lexical reasoning also can be useful in order to make easier knowledge representation tasks. Usually, this type of knowledge is expressed linguistically, which involves vagueness and uncertainty and, for this reason, it is important to select the most appropriate tool for modelling it. Fuzzy set theory (FS) is a good candidate, but it shows some particular limitations for our aim: i) there is an additional uncertainty level because words mean different things to different people; ii) it is difficult to accept that all individuals agree on the same membership function associated to words; and, iii) there not exists an unique and standard semantic measure, hence different metrics will provide different results.

In particular, in the field of fuzzy logic programming and fuzzy Prolog systems, little attention has been paid to the impact of a high degree of uncertainty and vagueness of lexical knowledge in the definition their knowledge bases and inference processes.

*Example 1. Suppose that we extract from Internet two people's opinions about a particular football player. The first one says "Ronaldo is a normal player" and the second one says "Ronaldo is a bad player". If we consider the label for qualifying the highest quality (e.g., "excellent") as a basic element, this lexical knowledge could be modelled by using two annotated facts as "football_player(Ronaldo,excellent):-0.8."*

*and "football_player(Ronaldo,excellent):-0.6.", respectively. In this case, we use "football_player(ronaldo,excelent):-0.6." given the infimum is usually employed. However, as it can be observed, we lost information using this approach.*

Another problem appears when semantic measures are applied (in the sense of WordNet Similarity [10]), since they are available different valid alternative models and each of them have their advantages and disadvantages. For that reason, we propose an alternative frame for modelling them in order to minimize loss of information.

*Example 2. Suppose we have the fact "loves(a,b)" and we extract the closeness between "loves" and "desires" by using two different semantics measures obtaining* $0.8$ *and* $0.6$*. Therefore, in order to represent this semantic knowledge we could employ two facts either "desires(a,b):-0.8" or "desires(a,b):-0.6".*

In order to address these limitations we propose to enrich fuzzy Prolog systems with interval-valued fuzzy sets (IVFSs), since they allow us to model better than FS the uncertainty associated to lexical knowledge. Several advantages are pointed out for dealing with environments with high uncertainty or imprecision by using IVFSs [7]. Other authors, from different areas, have shown that IVFSs can generate better results than those obtained using FSs [2]. For instance, the facts of the Example 1 can be combined into a single IVFS, obtaining a single fact "football_player(ronaldo,excelent):-[0.6,0.8]" or "desires(a,b):-[0.6,0.8]" in the Example 2.

## 2 Simple interval-valued fuzzy Prolog: Syntax, semantics and implementation

The design of a programming language consists of three important steps. The first one is the definition of the syntax. The second one consists in performing a formal study of its semantics. Finally, an implementation for the system must be provided. We are going to follow the guidelines established in [3] [3] (for the first and second steps) and the guidelines detailed in [4] (for the third step) since it is an efficient and standard way for implementing a Prolog system.

### 2.1 Approximate deductive reasoning

When we consider a collection of imprecise premises and a possible imprecise conclusion inferred by them in a Prolog program, we are applying a process of approximate deductive reasoning. Thus, assuming a propositional frame, an imprecise premise is a fact qualified by a degree of truth; e.g. "John is tall with [0.2,0.5]" or, if we assume a set-based frame, it denotes a degree of membership; i.e., "the membership of John to the set of tall people is [0.2,0.5]". It is worth noting that this degree is modelled by means IVFSs (where $[\underline{\alpha}, \overline{\alpha}] \in L([0,1])$ and $\underline{\alpha} \leq \overline{\alpha}$; $\underline{\alpha}$ is the lower bound and $\overline{\alpha}$ the upper one). The conclusion inferred from an imprecise premise must be also qualified by the same type of degree; e.g. "John is a good player with [0.2,0.5]". Therefore, approximate deductive reasoning is based on multi-valued modus ponens [5]: if we have (Q,$[\underline{\alpha}, \overline{\alpha}]$) and ($A \leftarrow Q$,$[\underline{\beta}, \overline{\beta}]$) then we can deduce $(A, T([\underline{\alpha}, \overline{\alpha}], [\underline{\beta}, \overline{\beta}]))$ with $T$ a t-norm defined

[3] We assume familiarity with the theory and practice of logic programming

on the lattice $L([0,1])$. In a Prolog context, only a set of clauses are employed in a process of inference when a query is launched, this fact is very important because it should be taken into account in order to define the declarative semantics. Given a set of interval-valued fuzzy clauses employed in a refutation, we will take its least approximation degree ($\lambda$) by using a particular t-norm. This concept plays an important role in our framework.

*Example 3.* Suppose the following interval-valued fuzzy program $\Pi = \{C_1$:p(a) [0.5,0.8]; $C_2$: q(a) [0.5,0.6]; $C_3$:s(a) [0.1,0.1]; $C_4$: r(X):- p(X), q(X). $\}$. If the query "?.-r(X)" is launched then the answer X=a with [0.5,0.6] is obtained by using a t-norm $T$, in this case the minimum. Note that the third fact "s(a) [0.1,0.1]" is not employed in the process of inference, hence the least approximation degree for the set of clauses $C_1, C_2$ and $C_4$ by using the t-norm minimum is $\lambda = T([0.5, 0.6], [0.5, 0.8], [1, 1]) = [0.5, 0.6]$ and the least approximation degree for the clause $C_3$ is $\lambda = [0.1, 0.1]$.

We denote this set of clauses as $\Pi_\lambda^T$ (it is also an interval-valued fuzzy program) and it provides us with a useful meaning about interval-valued fuzzy Prolog programs.

### 2.2 Syntax

An Interval-valued fuzzy program is basically formed by a classical logic program and a set of IVFSs, which are used for annotating the facts by means of an interval-valued fuzzy approximate degree: $p(t_1,\ldots,t_n)[\underline{\alpha},\overline{\alpha}]$. An interval-valued fuzzy fact is a Horn clause of the form $A[\underline{\alpha},\overline{\alpha}]$ and an interval-valued fuzzy program is a finite set of interval-valued fuzzy facts and a set of interval-valued fuzzy clauses $A \leftarrow B_1,\ldots,B_n[\underline{\beta},\overline{\beta}]$, where $A$ is called the head, and $B_1,\ldots,B_n$ denote a conjunction which is called the body (variables in a clause are assumed to be universally quantified).

### 2.3 Declarative semantics

We are going to provide the semantics of the syntax presented in the previous section. In our framework, truth is a matter of interval-valued degree $[\underline{\alpha},\overline{\alpha}]$ with $0 \leq \underline{\alpha} \leq \overline{\alpha} \leq 1$. An interval-valued fuzzy Interpretation $\mathcal{I}$ is a pair $\langle \mathcal{D}, \mathcal{J} \rangle$ where $\mathcal{D}$ is the domain of the interpretation and $\mathcal{J}$ is a mapping which assigns meaning to the symbols of $\mathcal{L}$: specifically n-ary relation symbols are interpreted as mappings $\mathcal{D}^n \longrightarrow L([0,1])$. In order to evaluate open formulas, we have to introduce the notion of variable assignment. A variable assignment, $\vartheta$, w.r.t. an interpretation $\mathcal{I} = \langle \mathcal{D}, \mathcal{J} \rangle$, is a mapping $\vartheta : \mathcal{V} \longrightarrow \mathcal{D}$, from the set of variables $\mathcal{V}$ of $\mathcal{L}$ to the elements of the interpretation domain $\mathcal{D}$. This notion can be extended to the set of terms of $\mathcal{L}$ by structural induction as usual. The following definition formalizes the notion of *valuation of a formula* in our framework.

**Definition 1.** *Given an interval-valued fuzzy interpretation $\mathcal{I}$ and a variable assignment $\vartheta$ in $\mathcal{I}$, the valuation of a formula w.r.t. $\mathcal{I}$ and $\vartheta$ is:*

1. $\mathcal{I}(p(t_1,\ldots,t_n))[\vartheta] = \bar{p}(t_1\vartheta,\ldots,t_n\vartheta)$, *where* $\mathcal{J}(p) = \bar{p}$; *ii)* $\mathcal{I}(A_1,\ldots,A_n))[\vartheta] = T\{\mathcal{I}(A_1)[\vartheta],\ldots,\mathcal{I}(A_n)[\vartheta]\}$;
2. $\mathcal{I}(A \leftarrow \mathcal{Q})[\vartheta] = 1$ *if* $I(A) >= I(Q)$; $\mathcal{I}(A \leftarrow \mathcal{Q})[\vartheta] = \mathcal{I}(\mathcal{A})[\vartheta]$ *if* $I(A) < I(Q)$;
3. $\mathcal{I}((\forall x)\mathcal{C})[\vartheta] = \inf\{\mathcal{I}(\mathcal{C})[\vartheta'] \mid \vartheta' \text{ } x\text{–equivalent to } \vartheta\}$ *where $p$ is a predicate symbol, $A$ and $A_i$ atomic formulas and $\mathcal{Q}$ any body, $\mathcal{C}$ any clause, $T$ is any left-continuous t-norm defined on $L([0,1])$. An assignment $\vartheta'$ is $x$–equivalent to $\vartheta$ when $z\vartheta' = z\vartheta$ for all variables $z \neq x$ in $\mathcal{V}$.*

We perform the generalization of the notion of Herbrand interpretation to the interval-valued fuzzy case.

**Definition 2.** *Given, a first order language $\mathcal{L}$, an interval-valued fuzzy Herbrand interpretation for $\mathcal{L}$ is a mapping $\mathcal{I} : \mathcal{B}_{\mathcal{L}} \longrightarrow L([0,1])$.*

Next, we introduce the notion of Interval-valued Fuzzy Herbrand Model and the logical consequence in our framework.

**Definition 3.** *An Interval-valued fuzzy Herbrand Interpretation is a model of an interval-valued fuzzy clause $\mathcal{C}[\underline{\alpha}, \overline{\alpha}]$ if and only if $\mathcal{I}(C) \geq [\underline{\alpha}, \overline{\alpha}]$ and it is a model of an interval-valued fuzzy program $\Pi^T_\lambda$ if and only if $\mathcal{I}(C)$ is a model for each clause $\mathcal{C}[\underline{\alpha}, \overline{\alpha}] \in \Pi^T_\lambda$.*

**Theorem 1.** *Let $\Pi^T_\lambda$ be an Interval-valued fuzzy program and suppose $\Pi^T_\lambda$ has a model. Then $\Pi^T_\lambda$ has a Herbrand model.*

**Definition 4.** *Let $\Pi^T_\lambda$ be an interval-valued fuzzy program. Let $\mathcal{A}$ be an interval-valued fuzzy clause of $\Pi^T_\lambda$. Then $\mathcal{A}$ is a logical consequence of $\Pi^T_\lambda$ if and only if for each interval-valued fuzzy interpretation $I$, if $I$ is a model for $\Pi^T_\lambda$ then $\mathcal{I}$ is a model for $\mathcal{A}$.*

**Proposition 1.** *$\mathcal{A}$ is a logical consequence of an interval-valued fuzzy program $\Pi^T_\lambda$ if and only if for every interval-valued fuzzy Herbrand interpretation $\mathcal{I}$ for $\Pi^T_\lambda$, if $\mathcal{I}$ is a model for $\Pi^T_\lambda$, it is an interval-valued fuzzy Herbrand model for $A$*

Interval-valued fuzzy interpretations have an important property which allow us to characterize the semantics of an interval-valued fuzzy program $\Pi^T_\lambda$. We can employ $T$ in order to establish the least model of an interval-valued fuzzy program $\Pi^T_\lambda$.

**Proposition 2.** *Let $\Pi^T_\lambda$ an interval-valued fuzzy program and $T$ be a t-norm defined on $L([0,1])$. Let $\mathcal{M}_1, ..., \mathcal{M}_n$ be a non-empty set of model for $\Pi^T_\lambda$. Then $T(\mathcal{M}_1, ..., \mathcal{M}_n) \geq \lambda$ is a model for $\Pi^T_\lambda$.*

**Definition 5.** *Let $\Pi^T_\lambda$ be an interval-valued fuzzy program and a t-norm $T$ defined on $L([0,1])$. The least model for $\Pi^T_\lambda$ employing $T$ is defined as follows: $\mathcal{M}^T_\Pi = T\{\mathcal{I}(A) \geq \lambda \mid A \in B_L\}$*

Note that, $\mathcal{M}^T_\Pi$ is the intersection of all models for $\Pi^T_\lambda$ by employing a particular t-norm $T$. This distinction is necessary because in the implementation phase a t-norm must be indicated before executing a program (t-norm minimum is usually employed by default).

**Theorem 2.** *Let $\Pi^T_\lambda$ an interval-valued fuzzy program. Let $\mathcal{M}^T_\Pi$ be the least model of $\Pi^T_\lambda$ employing $T$. Let $\mathcal{A} \in \mathcal{B}_{\mathcal{L}}$ a ground atom of the interval-valued fuzzy Herbrand base. $\mathcal{M}^T_\Pi(\mathcal{A}) \geq \lambda$ if and only if $\mathcal{A}$ is logical consequence of $\Pi^T_\lambda$.*

Note that, the operational semantics will compute the highest approximation degree associated to all logical consequences obtained and it depends on the query launched in the construction of the different programs $\Pi^T_\lambda$. For example, for the program of Example 3 which was created by the query "r(X)" we obtain that $(p(a), [0.25, 0.48])$ is a logical consequence by using the t-norm product. Additionally, we also obtain that $(p(a)[0.5, 0.8])$ is a logical consequence because the program $\Pi^T_\lambda = \{$ p(a) [0.5,0.8] $\}$ is created by the query "p(X)". We take $(p(a)[0.5, 0.8])$, that is, we are interested in those logical consequences associated to queries.

### 2.4 Operational semantics

We begin by providing definitions of an interval-valued SLD-derivation and an interval-valued fuzzy SLD-refutation that will be used later for showing the soundness and the completeness of the system.

**Definition 6.** *Let $\mathcal{G}$ be $\leftarrow A_1, \ldots, A_m, \ldots, A_k$ and $C$ be either $A[\underline{\alpha}, \overline{\alpha}]$ or $A \leftarrow B_1, \ldots, B_q$. Then $G'$ is derived from G and C using mgu $\theta$ if the following conditions hold (G' is the interval-fuzzy resolvent of G and C): i) $A_m$ is an atom called the selected atom in G; ii) $\theta$ is a mgu of $A_m$ and A; iii) $G'$ is the interval-valued fuzzy goal $\leftarrow (A_1, \ldots, B_1, \ldots, B_q, , \ldots, A_k)\theta$ with $[\underline{\alpha}_{G'}, \overline{\alpha}_{G'}] = T([\underline{\alpha}_C, \overline{\alpha}_C], [\underline{\alpha}_G, \overline{\alpha}_G])$*

**Definition 7.** *An interval-valued fuzzy SLD-derivation of $\Pi \cup G$ is a successful interval-valued SLD-derivation of $\Pi \cup G$ which has the empty clause as the last goal in the derivation. If $G_n$ is the empty clause, we say that the derivation has length $n$. The empty clause is derived from $\leftarrow (A_1, \ldots, A_m, \ldots, A_k)[\underline{\alpha}_G, \overline{\alpha}_G]$ and $A(t_1, \ldots, t_q)[\underline{\alpha}_A, \overline{\alpha}_A] \leftarrow$ with $[\underline{\alpha}_{G_n}, \overline{\alpha}_{G_n}] = T([\underline{\alpha}_A, \overline{\alpha}_A], [\underline{\alpha}_G, \overline{\alpha}_G])$*

**Definition 8.** *Let $\Pi$ be an interval-valued fuzzy program and G be an interval-valued fuzzy goal. An interval-valued fuzzy computed answer $\langle \theta, [\underline{\beta}, \overline{\beta}] \rangle$ for $\Pi \cup G$ is the substitution obtained by restricting the composition $\theta_1, \ldots, \theta_n$ to the variables of G, where $\theta_1, \ldots, \theta_n$ is the sequence of mgu's employed in the finite interval-valued fuzzy SLD-derivation of $\Pi \cup G$ with an interval-valued approximation degree $[\underline{\beta}, \overline{\beta}]$*

**Definition 9.** *Let $\Pi$ be an interval-valued fuzzy program, G be an interval-valued fuzzy goal $\leftarrow (A_1, \ldots, A_k)$ and $\langle \theta, [\underline{\beta}, \overline{\beta}] \rangle$ be an answer for $\Pi \cup G$. We say that $\langle \theta, [\underline{\beta}, \overline{\beta}] \rangle$ is an interval-valued fuzzy correct answer if $\forall (A_1, \ldots, A_k)\theta$ is a logical consequence of $\Pi$.*

## 3 Implementation

In this section, we briefly explain the incorporation of interval-valued fuzzy sets into the Bousi˜Prolog system (a beta version of this incorporation can be founded at the URL: `http://www.face.ubiobio.cl/~clrubio/bousiTools/`). We present the structure and main features of its abstract machine. It is an extension of the SWAM (Similarity Warren Abstract Machine) which was created for the execution of Bousi Prolog programs. We have appropriately modified some machine instructions and SWAM structures to carry out the interval-valued fuzzy resolution. It is important to note that, to the best of our knowledge, this is the first WAM implementation that supports interval-valued fuzzy resolution.

*Example 4.* Suppose the interval-valued fuzzy Prolog program P={coordinate(a)[0.2,0.4]; fast(a)[0.9,1.0]; tall(a)[0.8,0.9]; good_player(X):-tall(X), fast(X), coordinate(X)}. The SWAM generated for this program is:

```
00:good_player:trust_me [1.0,1.0]   11:coordinate:trust_me   [0.2,0.4]
01:            allocate             12:           get_constant a A0
02:            get_variable Y0 A0   13:           proceed
03:            put_value Y0 A0      14:           fast: trust_me [0.9,1.0]
04:            call coordinate (11) 15:           get_constant a A0
05:            put_value Y0 A0      16:           proceed
06:            call fast   (14)     17:tall:      trust_me   [0.8,0.9]
07:            put_value Y0 A0      18:           get_constant a A0
08:            call tall   (17)     19:           proceed
09:            deallocate
10:            proceed
```

The SWAM allows to obtain the answer: "X=a with [0.2,0.4]".

## 4 Conclusions and future work

We have formally defined and efficiently implemented a simple interval-valued fuzzy programming language using interval-valued fuzzy sets for modelling the uncertainty and imprecision of the knowledge associated to lexical resources. As future work we propose to extend our language and to provide results of soundness and completeness. Additionally, we want to develop a fully integrated framework in which interval-valued fuzzy sets and interval-valued fuzzy relations can be combined in a same framework.

## References


1. P. Leon-Arauz, J. Gomez-Romero, F. Bobillo. A Fuzzy Ontology Extension of WordNet and EuroWordnet for Specialized Knowledge. Proceedings of the TKE 2012, pp. 139-154, 2012.
2. H. Bustince. Interval-valued Fuzzy Sets in Soft Computing International Journal of Computational Intelligence Systems, Vol.3, No. 2 pp. 215-222 (2010).
3. J.W. Lloyd. Foundations of Logic Programming. Springer-Verlag, Berlin, 1987. Second edition.
4. P. Julian-Iranzo and C.Rubio-Manzano. A Similarity-Based WAM for Bousi Prolog. IWANN (2009).
5. P. Hajek. Metamathematics of fuzzy logic (Vol. 4). Springer Science and Business Media (1998).
6. L. E. Van Hung et al. Fuzzy linguistic logic programming and its applications. Theory and Practice of Logic Programming, 9(03), 309-341 (2009)
7. I. B. Turksen, Four methods of approximate reasoning with interval-valued fuzzy sets. International journal of approximate reasoning, 3(2), 121-142 (1989)
8. G.A. Miller WordNet: A Lexical Database for English Communications of the ACM, 38, (1995) 39–41.
9. H. Liu and P. Singh. Commonsense Reasoning in and over Natural Language. Proc. of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, (2004), 293–306.
10. T. Pedersen et al. WordNet:: Similarity: measuring the relatedness of concepts. In Demonstration papers at HLT-NAACL 2004 (pp. 38-41). Association for Computational Linguistics.
11. U. Straccia. Managing uncertainty and vagueness in description logics, logic programs and description logic programs. Springer (2008).
12. J. Medina and M. Ojeda-Aciego. Multi-adjoint logic programming. IPMU 2004.
13. P. Julian-Iranzo, et al. A fuzzy logic programming environment for managing similarity and truth degrees. arXiv preprint arXiv:1501.02034.

# Multi-adjoint Frameworks, towards a More Flexible Formal World

Jesús Medina

Department of Mathematics, University of Cádiz, Spain
jesus.medina@uca.es

## Abstract

Multi-adjoint logic programming [10–12], multi-adjoint fuzzy rough sets [1], multi-adjoint concept lattices [6–9], multi-adjoint fuzzy relation equations [2–5], etc. These are different frameworks in which the common factor is the multi-adjoint philosophy. This is based on the consideration of an general algebraic structure, called multi-adjoint lattices or algebras, in which the adjoint triples are the underline operators, and the possibility of considering different adjoint triples at the same time.

These operators are defined on general posets or lattices, depend on the specific considered framework, and they do not need to be commutative and/or associative. These general properties together with the consideration of several adjoint triples provides an extra level of flexibility in the framework in which this structure is considered.

This work will present the multi-adjoint algebras, introduce diverse examples, analyze the main features and properties in the main frameworks in which they have been considered and show the first demo of a software which involves different of these frameworks.

## References

1. C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.
2. J. C. Díaz and J. Medina. Multi-adjoint relation equations: Definition, properties and solutions using concept lattices. *Information Sciences*, 253:100–109, 2013.
3. J. C. Díaz and J. Medina. Solving systems of fuzzy relation equations by fuzzy property-oriented concepts. *Information Sciences*, 222:405–412, 2013.
4. J. C. Díaz-Moreno and J. Medina. Using concept lattice theory to obtain the set of solutions of multi-adjoint relation equations. *Information Sciences*, 266(0):218–225, 2014.
5. J. C. Díaz-Moreno, J. Medina, and E. Turunen. Minimal solutions of general fuzzy relation equations on linear carriers. an algebraic characterization. *Fuzzy Sets and Systems*, pages –, 2016.
6. J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.
7. J. Medina and M. Ojeda-Aciego. On multi-adjoint concept lattices based on heterogeneous conjunctors. *Fuzzy Sets and Systems*, 208:95–110, 2012.
8. J. Medina, M. Ojeda-Aciego, J. Pócs, and E. Ramírez-Poussa. On the Dedekind-Macneille completion and formal concept analysis based on multilattices. *Fuzzy Sets and Systems*, 303:1–20, 2016.

9. J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. Formal concept analysis via multi-adjoint concept lattices. *Fuzzy Sets and Systems*, 160(2):130–144, 2009.
10. J. Medina, M. Ojeda-Aciego, A. Valverde, and P. Vojtáš. Towards biresiduated multi-adjoint logic programming. *Lecture Notes in Artificial Intelligence*, 3040:608–617, 2004.
11. 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.
12. J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Similarity-based unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146:43–62, 2004.

# Characterizing the Join Irreducible Elements of Multi-adjoined Object-oriented Fuzzy Concept Lattices *

Jesús Medina and Eloísa Ramírez-Poussa

Department of Mathematics, University of Cádiz, Spain
{jesus.medina@uca.es, eloisa.ramirez}@uca.es

**Abstract.** This paper introduces a first relation between multi-adjoint formal and object-oriented concept lattices, which is focused on a characterization of the irreducible elements of a multi-adjoint object-oriented concept lattice by the one given in the multi-adjoint concept lattice framework.

**Keywords:** formal concept analysis, multi-adjoint object-oriented concept lattice, irreducible element.

## 1 Introduction

Formal concept analysis (FCA) and rough set theory are two interesting tools for data analysis. FCA is used to extract information from databases transforming the information contained in a database into a mathematical structure called concept lattice. These databases contain a set of attributes $A$ and a set of objects $B$ related between them by means of a relation $R \subseteq A \times B$, from which we obtain pieces of information called concepts. Establishing a hierarchy among these concepts, we obtain the algebraic structure of a concept lattice from which we can develop the data analysis. This mathematical theory has become a very attractive research topic from both theoretical [15] and applied perspectives [1, 5–8, 14], due to the large number of possible applications. A new general approach to FCA, using the philosophy of the multi-adjoint framework was presented in [11, 12].

Rough set theory was introduced by Pawlak [13], as an approach for dealing with the problem of how to understand and handle imperfect knowledge or information in uncertain conditions. Because of its many practical applications, many researches have contributed to develop this theory. Hence, after its introduction, in order to consider two different sets (the set of objects and the set of attributes) two extensions of this theory were introduced, *property-oriented concept lattice* [2] and *object-oriented concept lattices* [16]. In [9] a generalisation of the classical property and object-oriented concept lattices considering the multi-adjoint paradigm was introduced.

On the other hand, databases usually contain a large number of attributes. Hence, attribute reduction plays an important role on the reduction of the computational complexity to build the concept lattices, in all these frameworks. In [10] the relation among

* Partially supported by the Spanish Economy and Competitiveness Ministry (MINECO) project TIN2016-76653-P

attribute reduction in formal, object-oriented and property-oriented concept lattices in the classical case was presented. In this paper, in the fuzzy case, a characterization of the join irreducible ($\vee$-irreducible) elements of a multi-adjoint object-oriented concept lattice, which is an essential part in order to provide an attribute reduction mechanism, is introduced and how it can be proved from the corresponding characterization of the meet irreducible ($\wedge$-irreducible) elements of a multi-adjoint concept lattice [4].

## 2 Preliminary notions

In this section, we briefly introduce some preliminary necessary notions and results to understand this work. Due to lack of space, we assume that the reader is familiar with the main concepts concerning the three frameworks considered in this paper. More information about these notions can be found in [4, 9].

When we work FCA in a multi-adjoint environment, we need to fix a multi-adjoint frame $(L_1, L_2, P, \&_1, \dots, \&_n)$ where $(\&_i, \swarrow^i, \nwarrow_i)$ is an adjoint triple with respect to $L_1$, $L_2$ and $P$, for all $i \in \{1, \dots, n\}$. In addition, in the multi-adjoint object-oriented concept lattice environment, the considered multi-adjoint object-oriented frame is $(L_1, L_2, P, \&_1, \dots, \&_n)$ where the family of adjoint triples $(\&_i, \swarrow^i, \nwarrow_i)$ is defined on the carriers $L_1, P$ and $L_2$. In a similar way, for the property-oriented concept lattice environment, the multi-adjoint property-oriented frame $(L_1, L_2, P, \&_1, \dots, \&_n)$ is composed by a family of adjoint triples $(\&_i, \swarrow^i, \nwarrow_i)$ with respect to $P, L_2, L_1$. In all these frameworks, $(L_1, \leq_1)$, $(L_2, \leq_2)$ and $(P, \leq)$ are two complete lattices and a poset, respectively.

In addition, for each environment we need to fix a context $(A, B, R, \sigma)$ such that $A$ and $B$ are sets of attributes and objects, respectively, $R$ is a $P$-fuzzy relation $R\colon A \times B \to P$ and $\sigma$ is a mapping which associates any element in $A \times B$ with some particular adjoint triple of the corresponding framework.

We will write $L_2^B$ and $L_1^A$ to represent the set of mappings $g\colon B \to L_2$, $f\colon A \to L_1$, respectively. Taking into account these mappings we define the concept-forming operators for the different frameworks as follows:

– In the multi-adjoint framework[1]

$$g^{\uparrow_\sigma}(a) = \inf\{R(a,b) \swarrow^{\sigma(a,b)} g(b) \mid b \in B\} \tag{1}$$
$$f^{\downarrow^\sigma}(b) = \inf\{R(a,b) \nwarrow_{\sigma(a,b)} f(a) \mid a \in A\} \tag{2}$$

– In the multi-adjoint object-oriented framework

$$g^{\uparrow_N}(a) = \inf\{g(b) \swarrow^b R(a,b) \mid b \in B\} \tag{3}$$
$$f^{\downarrow^\pi}(b) = \sup\{f(a) \,\&_b\, R(a,b) \mid a \in A\} \tag{4}$$

– In the multi-adjoint property-oriented framework

$$g^{\uparrow_\pi}(a) = \sup\{R(a,b) \,\&_b\, g(b) \mid b \in B\} \tag{5}$$
$$f^{\downarrow^N}(b) = \inf\{f(a) \nwarrow_b R(a,b) \mid a \in A\} \tag{6}$$

[1] In this paper, in order to simplify the notation, we will write $^\uparrow$ and $^\downarrow$ instead of $^{\uparrow_\sigma}$ and $^{\downarrow^\sigma}$, respectively.

It is worth highlighting that the pair $(^{\uparrow}, ^{\downarrow})$ is a antitone Galois connection, while the pairs $(^{\uparrow_N}, ^{\downarrow^{\pi}})$ and $(^{\uparrow_\pi}, ^{\downarrow^N})$ are isotone Galois connections. The multi-adjoint concept lattice will be denotes as $(\mathcal{M}, \preceq)$, the multi-adjoint object-oriented concept lattice as $(M_{N\pi}, \preceq)$ and the multi-adjoint property-oriented concept lattice as $(M_{\pi N}, \preceq)$.

In this work, we will give a characterization of the irreducible element of $(M_{N\pi}, \preceq)$ following the idea presented in [4]. In order to recall the characterization of the $\wedge$-irreducible elements of a multi-adjoint concept lattice $(\mathcal{M}, \preceq)$, we introduce the following definition and result [3, 4].

**Definition 1.** *For each $a \in A$, the fuzzy subsets of attributes $\phi_{a,x} \in L_1^A$ defined, for all $x \in L_1$, as*

$$\phi_{a,x}(a') = \begin{cases} x & \text{if } a' = a \\ \perp_1 & \text{if } a' \neq a \end{cases}$$

*will be called* fuzzy-attributes*, where $\perp_1$ is the minimum element in $L_1$. The set of all fuzzy-attributes will be denoted as $\Phi = \{\phi_{a,x} \mid a \in A, x \in L_1\}$.*

**Theorem 1 ([4]).** *The set of $\wedge$-irreducible elements of $\mathcal{M}$, $M_F(A)$, is formed by the pairs $\langle \phi_{a,x}^{\downarrow}, \phi_{a,x}^{\downarrow\uparrow} \rangle$ in $\mathcal{M}$, with $a \in A$ and $x \in L_1$, such that*

$$\phi_{a,x}^{\downarrow} \neq \bigwedge \{\phi_{a_i,x_i}^{\downarrow} \mid \phi_{a_i,x_i} \in \Phi, \phi_{a,x}^{\downarrow} \prec_2 \phi_{a_i,x_i}^{\downarrow}\}$$

*and $\phi_{a,x}^{\downarrow} \neq g_{\top_2}$, where $\top_2$ is the maximum element in $L_2$ and $g_{\top_2}\colon B \to L_2$ is the fuzzy subset defined as $g_{\top_2}(b) = \top_2$, for all $b \in B$.*

The following result shows the effect of exchanging the order on the posets by the opposite one. This result has been very useful to prove the characterization of the irreducible elements of $(M_{N\pi}, \preceq)$.

**Lemma 1 ([9]).** *Given the posets $(P_1, \leq_1)$, $(P_2, \leq_2)$, $(P_3, \leq_3)$ and an adjoint triple $(\&, \swarrow, \nwarrow)$ with respect them, we obtain that:*

1. *$(\&^{op}, \nwarrow, \swarrow)$ is an adjoint triple with respect to $P_2$, $P_1$, $P_3$.*
2. *$(\swarrow, \&, \nwarrow_{op})$ is an adjoint triple with respect to $P_3^\partial$, $P_2$, $P_1^\partial$.*
3. *$(\nwarrow, \&^{op}, \swarrow^{op})$ is an adjoint triple with respect to $P_3^\partial$, $P_1$, $P_2^\partial$.*
4. *$(\nwarrow_{op}, \swarrow^{op}, \&^{op})$ is an adjoint triple with respect to $P_1$, $P_3^\partial$, $P_2^\partial$.*

Considering a multi-adjoint object-oriented frame $(L_1, L_2, P, \&_1, \dots, \&_n)$ and a context $(A, B, R, \sigma)$. If we consider $(L_1, \preceq_1)$, $(L_2, \preceq_2^{op})$ and $(P, \leq^{op})$ instead of $(L_1, \preceq_1$ $)$, $(L_2, \preceq_2)$ and $(P, \leq)$, the pairs $(^{\uparrow_N}, ^{\downarrow^{\pi}})$ and $(^{\uparrow}, ^{\downarrow})$, defined in the multi-adjoint object-oriented frame $(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \dots, \&_n)$ and in the multi-adjoint frame $(L_1, L_2, P, \preceq_1$ $, \preceq_2^{op}, \leq^{op}, \nwarrow_{1,op}, \dots, \nwarrow_{n,op})$ respectively, coincide.

**Lemma 2.** *If we consider the posets $(L_1, \preceq_1)$, $(L_2, \preceq_2^{op})$ and $(P, \leq^{op})$ instead of $(L_1, \preceq_1$ $)$, $(L_2, \preceq_2)$ and $(P, \leq)$, then we obtain that the pair $(^{\uparrow_N}, ^{\downarrow^{\pi}})$ is equal to the pair $(^{\uparrow}, ^{\downarrow})$ given in Equations (1) and (2).*

## 3 Relating multi-adjoint formal and object-oriented concept lattices

In this section, considering the previous results and comments, a characterization of the set of $\vee$-irreducible elements of a multi-adjoint object-oriented concept lattice is introduced. Beside this interesting property, in a similar way, the results given in [4], can be translate to the object-oriented framework and, analogously, to the property-oriented concept lattice as well. Due to the lack of space they will be present in future extensions of this work.

From now on, a multi-adjoint object-oriented frame $(L_1, L_2, P, \&_1, \dots, \&_n)$ and a context $(A, B, R, \sigma)$ will be fixed.

From Theorem 1, the characterization of the set of $\vee$-irreducible elements of a multi-adjoint object-oriented concept lattice is given below.

**Proposition 1.** *Let $(A, B.R, \sigma)$ be a formal context and a multi-adjoint object-oriented frame $(L_1, L_2, P, \&_1, \dots, \&_n)$. The set of $\vee$-irreducible elements of $\mathcal{M}_{N\pi}$, $J_o(A, B, R, \sigma)$, is:*

$$\Big\{\langle \phi_{a,x}^{\downarrow^\pi}, \phi_{a,x}^{\downarrow^\pi\uparrow_N}\rangle \mid \phi_{a,x}^{\downarrow^\pi} \neq \bigvee\{\phi_{a_i,x_i}^{\downarrow^\pi} \mid \phi_{a_i,x_i} \in \varPhi, \phi_{a_i,x_i}^{\downarrow^\pi} \prec_2 \phi_{a,x}^{\downarrow^\pi}\} \text{ and } \phi_{a,x}^{\downarrow^\pi} \neq g_\perp\Big\}$$

*where $\perp$ is the minimum element in $L_2$ and $g_\perp \colon B \to L_2$ is the fuzzy subset defined as $g_\perp(b) = \perp$, for all $b \in B$.*

The demonstration of the previous result can be developed in a similar way to the one given in [4]. Whereas, this proof can be straightforwardly obtained from the fact that if $(\&_i, \swarrow^i, \nwarrow_i)$ are adjoint triples defined on the carriers $L_1, P$ and $L_2$ in the multi-adjoint object-oriented frame, then $(\nwarrow_{i,op}, \swarrow^{i,op}, \&_i^{op})$ are adjoint triples with respect to $(L_1, \preceq_1)$, $(L_2, \preceq_2^\partial)$ and $(P, \leq^\partial)$ in the multi-adjoint frame, by Lemma 1. Therefore, we are considering the dual ordering on the lattice $L_2$ and, as a consequence, the infima turn into suprema.

*Example 1.* Let $(L, \preceq, L, \preceq, L, \preceq, \&_G^*)$ be a multi-adjoint object-oriented frame, where $L = [0,1]_2$ is the regular partitions of $[0,1]$ in 2 pieces, and $\&_G^*$ is the discretization of the Gödel conjunctor defined on $L \times L$. We consider a context $(A, B, R, \sigma)$, where $A = \{a_1, a_2, a_3\}$, $B = \{b_1, b_2\}$, $R\colon A \times B \to L$ is given by the table shown in the left side of Figure 1 and $\sigma$ is constantly $\&_G^*$.

The multi-adjoint object-oriented concept lattice $(M_{N\pi}, \preceq)$, associated with the considered framework and context is composed by the following 6 concepts:

$$\begin{aligned}
C_0 &= \langle\{0/b_1, 0/b_2\}, \{0/a_1, 0/a_2, 0/a_3\}\rangle\\
C_1 &= \langle\{0/b_1, 0.5/b_2\}, \{0/a_1, 1.0/a_2, 0/a_3\}\rangle\\
C_2 &= \langle\{0.5/b_1, 0/b_2\}, \{0.5/a_1, 0/a_2, 1.0/a_3\}\rangle\\
C_3 &= \langle\{0.5/b_1, 0.5/b_2\}, \{0.5/a_1, 1.0/a_2, 1.0/a_3\}\rangle\\
C_4 &= \langle\{1.0/b_1, 0.5/b_2\}, \{1.0/a_1, 1.0/a_2, 1.0/a_3\}\rangle\\
C_5 &= \langle\{1.0/b_1, 0/b_2\}, \{1.0/a_1, 0/a_2, 1.0/a_3\}\rangle
\end{aligned}$$

In this case, analyzing the fuzzy-attributes, we can see that $\langle \phi_{a,0.0}^{\downarrow^\pi}, \phi_{a,0.0}^{\downarrow^\pi\uparrow_N}\rangle = C_0$, for all $a \in A$, and $\langle \phi_{a_1,0.5}^{\downarrow^\pi}, \phi_{a_1,0.5}^{\downarrow^\pi\uparrow_N}\rangle$ is also $C_0$. From the rest of the fuzzy-attributes we

| $R$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $b_1$ | 1 | 0 | 0.5 |
| $b_2$ | 0 | 0.5 | 0 |

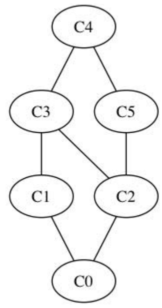



**Fig. 1.** Relation R (left side) and Hasse diagram of $(M_{N\pi}, \preceq)$ (right side) of Example 1.

obtain the following concepts:

$$\begin{aligned} \langle \phi^{\downarrow^\pi}_{a_1,1.0}, \phi^{\downarrow^\pi\uparrow_N}_{a_1,1.0}\rangle &= C_5 \\ \langle \phi^{\downarrow^\pi}_{a_2,0.5}, \phi^{\downarrow^\pi\uparrow_N}_{a_2,0.5}\rangle = \langle \phi^{\downarrow^\pi}_{a_2,1.0}, \phi^{\downarrow^\pi\uparrow_N}_{a_2,1.0}\rangle &= C_1 \\ \langle \phi^{\downarrow^\pi}_{a_3,0.5}, \phi^{\downarrow^\pi\uparrow_N}_{a_3,0.5}\rangle = \langle \phi^{\downarrow^\pi}_{a_3,1.0}, \phi^{\downarrow^\pi\uparrow_N}_{a_3,1.0}\rangle &= C_2 \end{aligned}$$

Therefore, the concepts $C_0, C_1, C_2$ and $C_5$ are related to the fuzzy-attributes. From the Hasse diagram we can observe that the set of $\vee$-irreducible elements of multi-adjoint object-oriented concept lattice $(M_{N\pi}, \preceq)$ is formed by the concepts:

$$J_o(A, B, R, \sigma) = \{C_1, C_2, C_5\}$$

In addition, we can assert that all these concepts satisfy the conditions of the Proposition 1. For example, if we consider the concept $C_5 = \langle \phi^{\downarrow^\pi}_{a_1,1.0}, \phi^{\downarrow^\pi\uparrow_N}_{a_1,1.0}\rangle$, we can easily see that the extension of this element can not be expressed as supremum of elements: $\phi^{\downarrow^\pi}_{a_j,x_j} \in \Phi$, such that $\phi^{\downarrow^\pi}_{a_j,x_j} \prec_2 \phi^{\downarrow^\pi}_{a_1,1.0}$; in addition $\phi^{\downarrow^\pi}_{a_1,1.0} \neq g_\perp$. Consequently, the concept $C_5$ is a $\vee$-irreducible element. In the same situation we find the concepts $C_1$ and $C_2$.

In a similar way, a similar result can be proven for the multi-adjoint property-oriented concept lattice framework. It will be presented in future extensions of this work.

## 4 Conclusions and further works

In this paper, we have considered the relationship between formal concept analysis and rough set theory, in order to present a characterization of the irreducible elements in multi-adjoint object-oriented concept lattices. As it was shown in [4], this characterization plays an important role in the study of attribute reduction in multi-adjoint object-oriented concept lattices.

As future work, we will continue studying more properties in order to obtain algorithms which will allow us to reduce the set of attributes in multi-adjoint object-oriented and property-oriented concept lattices.

## References


1. G. Arévalo, S. Ducasse, S. Gordillo, and O. Nierstraszn. Generating a catalog of unanticipated schemas in class hierarchies using formal concept analysis. *Information and Software Technology*, 52(11):1167–1187, 2010.
2. Y. Chen and Y. Yao. A multiview approach for intelligent data analysis based on data operators. *Information Sciences*, 178(1):1–20, 2008.
3. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. On the classification of fuzzy-attributes in multi-adjoint concept lattices. *Lecture Notes in Computer Science*, 7903:266–277, 2013.
4. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294:41 – 56, 2015.
5. A. Formica. Concept similarity in fuzzy formal concept analysis for semantic web. *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, 18(2):153–167, 2010.
6. B. Galitsky and J. D. L. Rosa. Concept-based learning of human behavior for customer relationship management. *Information Sciences*, 181(10):2016–2035, 2011.
7. B. Ganter, G. Stumme, and R. Wille, editors. *Formal Concept Analysis: Foundations and Applications*. Lecture Notes in Computer Science. Springer Verlag, 2005.
8. C. Kumar and S. Srinivas. Mining associations in health care data using formal concept analysis and singular value decomposition. *Journal of Biological Systems*, 18(4):787–807, 2010.
9. J. Medina. Multi-adjoint property-oriented and object-oriented concept lattices. *Information Sciences*, 190:95–106, 2012.
10. J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
11. J. Medina and M. Ojeda-Aciego. Multi-adjoint t-concept lattices. *Information Sciences*, 180(5):712–725, 2010.
12. 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.
13. Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
14. J. Poelmans, P. Elzinga, S. Viaene, and G. Deden. Formal concept analysis in knowledge discovery: A survey. *Lecture Notes in Computer Science*, 6208:139–153, 2010.
15. K.-S. Qu and Y.-H. Zhai. Generating complete set of implications for formal contexts. *Knowledge-Based Systems*, 21:429–433, 2008.
16. Y. Y. Yao. A comparative study of formal concept analysis and rough set theory in data analysis. *Lecture Notes in Artificial Intelligence*, 3066:59–68, 2004.

# Minimal Solutions of Finite Fuzzy Relation Equations on Linear Carriers by Cubes ⋆

Jesús Medina and Juan Carlos Díaz-Moreno

Department of Mathematics, University of Cádiz, Spain
{jesus.medina,juancarlos.diaz}@uca.es

## Abstract

Fuzzy relation equations, introduced by E. Sanchez in the seventies [11], is an important tool for managing and modeling uncertain or imprecise datasets, which has useful applied to, e.g. approximate reasoning, time series forecast, decision making, fuzzy control, etc. Different results and properties have been studied from its introduction [1, 3, 5].

One of the most important task in this area is the computation of the minimal solutions of a solvable fuzzy relation equation [2, 4, 10, 12, 13, 15–17]. However, restrictive frameworks have been considered in order to solve this important challenge.

The study of characterizations and mechanisms in order to compute the solution set and the minimal solutions, when they exist, is one of our main goals in the last years. See, for example, the recent papers [7, 6, 9].

This paper is a continuation of our last contribution [9] focuses on presenting a new procedure in order to obtain the minimal solutions of this fuzzy relation equation.

The considered general algebraic structure is given by a complete linear lattice $L, \leq$ in which a residuated operator $\odot\colon L \times L \to L$ is defined satisfying that it is order preserving in both arguments and there exists another operator $\to\colon L \times L \to L$, satisfying the following adjoint property with the conjunctor $\odot$

$$x \odot y \preceq z \quad \text{if and only if} \quad y \preceq x \to z \tag{1}$$

for each $x, y, z \in L$. This property is equivalent to say that $\odot$ preserves supremums in the second argument; $x \odot \bigvee\{y \mid y \in Y\} = \bigvee\{x \odot y \mid y \in Y\}$, for all $Y \subseteq L$.

From this operator and the sets $U$, $V$ and $W$, the following composition of the matrices $R\colon U \times V \to L$ and $X\colon V \times W \to L$ is introduced

$$(R \circ X)\langle u, w\rangle = \bigvee\{R\langle u, v\rangle \odot X\langle v, w\rangle \mid v \in V\}.$$

which leads to define the general *fuzzy relation equation*:

$$R \circ X = T, \tag{2}$$

where $R\colon U \times V \to L$, $T\colon U \times W \to L$ are given finite $L$-fuzzy relations and $X\colon V \times W \to L$ is unknown;

⋆ Partially supported by the Spanish Economy and Competitiveness Ministry (MINECO) project TIN2016-76653-P

This fuzzy relation equation has a solution if and only if

$$(R \Rightarrow T)\langle v, w\rangle = \bigwedge\{R\langle u, v\rangle \to T\langle u, w\rangle \mid u \in U\}$$

is a solution and, in that case, it is the greatest solution, see [8, 11, 14]. Moreover, the minimal solutions, when they exist, were studied in [9].

This paper introduces a further step presenting a new mechanism based on the notions of disjunctive and conjunctive normal form of a formula in the sense of classical logic and it is an alternative procedure to the one obtained in [9].

The procedure is illustrated in the following example. More details will be given during the conference and in future extensions.

*Example 1.* In this example we will consider the standard MV–algebra, that is, $L = [0, 1]$ is the unit interval, $\odot\colon L \times L \to L$ is the Łukasiewicz operator defined by $x \odot y = \max\{0, x + y - 1\}$ and $\to\colon L \times L \to L$ its residuated implication, defined by $y \to z = \min\{1, 1 - y + z\}$, for all $x, y, z \in [0, 1]$.

Let $U = \{u_1, u_2, u_3\}$, $V = \{v_1, v_2, v_3\}$, $W = \{w_1, w_2, w_3\}$ and the fuzzy relation equations defined from the following tables

| $R$ | $v_1$ | $v_2$ | $v_3$ |
|---|---|---|---|
| $u_1$ | 0.9 | 0.5 | 0.9 |
| $u_2$ | 0.2 | 0.9 | 0.7 |
| $u_3$ | 0.8 | 0.6 | 0.9 |

and

| $T$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $u_1$ | 0.8 | 0.4 | 0.7 |
| $u_2$ | 0.6 | 0.7 | 0.3 |
| $u_3$ | 0.8 | 0.4 | 0.6 |

Once we have checked that $R \circ (R \Rightarrow T) = T$, we can ensure that the equation $R \circ X = T$ is solvable and its greatest solution is $R \Rightarrow T$, which is expanded in the following table

| $R \Rightarrow T$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0.9 | 0.5 | 0.8 |
| $v_2$ | 0.7 | 0.8 | 0.4 |
| $v_3$ | 0.9 | 0.5 | 0.6 |

For the first position $((u_1, w_1))$ we have that

$$\begin{aligned}
R\langle u_1, v_1\rangle \odot (R \Rightarrow T)\langle v_1, w_1\rangle &= 0.8\\
R\langle u_1, v_2\rangle \odot (R \Rightarrow T)\langle v_2, w_1\rangle &= 0.2\\
R\langle u_1, v_3\rangle \odot (R \Rightarrow T)\langle v_3, w_1\rangle &= 0.8
\end{aligned}$$

Therefore, $V_{u_1 w_1} = \{v_1, v_3\}$ and we obtain the set $Z_{u_1 w_1} = \{(v_1, 0.9), (v_3, 0.9)\}$.

Following a similar procedure in positions $((u_2, w_1))$ and $((u_3, w_1))$, we obtain $Z_{u_2 w_1} = \{(v_2, 0.7), (v_3, 0.9)\}$ and $Z_{u_3 w_1} = \{(v_3, 0.9)\}$.

From these sets we consider the CNF formula:

$$\begin{aligned}
f = {}& ((v_1, 0.9) \vee (v_3, 0.9)) \wedge ((v_2, 0.7) \vee (v_3, 0.9)) \wedge\\
& (v_3, 0.9)
\end{aligned}$$

the reduced DNF of the formula $f$ is computed:

$$\begin{aligned}
f = {}& ((v_1, 0.9) \wedge (v_2, 0.7) \wedge (v_3, 0.9)) \vee ((v_1, 0.9) \wedge\\
& (v_3, 0.9)) \vee ((v_2, 0.7) \wedge (v_3, 0.9)) \vee (v_3, 0.9)
\end{aligned}$$

The following fuzzy sets arise from the obtained reduced DNF
$X^1_{w_1} = \begin{pmatrix} 0.9 \\ 0.7 \\ 0.9 \end{pmatrix}, X^2_{w_1} = \begin{pmatrix} 0.9 \\ 0 \\ 0.9 \end{pmatrix}, X^3_{w_1} = \begin{pmatrix} 0 \\ 0.7 \\ 0.9 \end{pmatrix}, X^4_{w_1} = \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix},$
and we can ensure that

$$X_{w_1} = \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix}$$

is the only minimal element of $\{X^1_{w_1}, X^2_{w_1}, X^3_{w_1}, X^4_{w_1}\}$.

This procedure is repeated to the other two columns. From the second column of $R \Rightarrow T$, we obtain the fuzzy sets

$$X^1_{w_2} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0 \end{pmatrix}, X^2_{w_2} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.5 \end{pmatrix}, X^3_{w_2} = \begin{pmatrix} 0 \\ 0.8 \\ 0.5 \end{pmatrix}$$

Hence, the only minimal elements of $\{X^1_{w_2}, X^2_{w_2}, X^3_{w_2}\}$ are $\begin{pmatrix} 0 \\ 0.8 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.8 \\ 0 \end{pmatrix}$.

According to the third column of $R \Rightarrow T$, we obtain the solutions

$$X^1_{w_3} = \begin{pmatrix} 0.8 \\ 0.4 \\ 0 \end{pmatrix}, X^2_{w_3} = \begin{pmatrix} 0.8 \\ 0 \\ 0.6 \end{pmatrix}$$

which are both the unique minimal elements of $\{X^1_{w_3}, X^2_{w_3}\}$.

Thus, we finally obtain four minimal solutions of Equation (2), which are detailed below.

| $X_1$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0 | 0.8 |
| $v_2$ | 0 | 0.8 | 0.4 |
| $v_3$ | 0.9 | 0.5 | 0 |

| $X_2$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0 | 0.8 |
| $v_2$ | 0 | 0.8 | 0 |
| $v_3$ | 0.9 | 0.5 | 0.6 |

| $X_3$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0.5 | 0.8 |
| $v_2$ | 0 | 0.8 | 0.4 |
| $v_3$ | 0.9 | 0 | 0 |

| $X_4$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0.5 | 0.8 |
| $v_2$ | 0 | 0.8 | 0 |
| $v_3$ | 0.9 | 0 | 0.6 |

## Acknowledgements

Jesús Medina and Juan Carlos Díaz were partially supported by the Spanish Science Ministry project TIN2012-39353-C04-04.

## References


1. W. Bandler and L. Kohout. Semantics of implication operators and fuzzy relational products. *Int. J. Man-Machine Studies*, 12:89–116, 1980.
2. E. Bartl. Minimal solutions of generalized fuzzy relational equations: Probabilistic algorithm based on greedy approach. *Fuzzy Sets and Systems*, 260(0):25 – 42, 2015.
3. R. Bělohlávek. *Fuzzy Relational Systems: Foundations and Principles*. Kluwer Academic Publishers, 2002.
4. L. Chen and P. Wang. Fuzzy relation equations (ii): The branch-point-solutions and the categorized minimal solutions. *Soft Computing–A Fusion of Foundations, Methodologies and Applications*, 11:33–40, 2007.
5. B. De Baets. Analytical solution methods for fuzzy relation equations. In D. Dubois and H. Prade, editors, *The Handbooks of Fuzzy Sets Series*, volume 1, pages 291–340. Kluwer, Dordrecht, 1999.
6. J. C. Díaz and J. Medina. Multi-adjoint relation equations: Definition, properties and solutions using concept lattices. *Information Sciences*, 253:100–109, 2013.
7. J. C. Díaz-Moreno and J. Medina. Using concept lattice theory to obtain the set of solutions of multi-adjoint relation equations. *Information Sciences*, 266(0):218–225, 2014.
8. J. C. Díaz-Moreno, J. Medina, and M. Ojeda-Aciego. On basic conditions to generate multi-adjoint concept lattices via galois connections. *International Journal of General Systems*, 43(2):149–161, 2014.
9. J. C. Díaz-Moreno, J. Medina, and E. Turunen. Minimal solutions of general fuzzy relation equations on linear carriers. an algebraic characterization. *Fuzzy Sets and Systems*, pages –, 2016.
10. J.-L. Lin, Y.-K. Wu, and S.-M. Guu. On fuzzy relational equations and the covering problem. *Information Sciences*, 181(14):2951–2963, 2011.
11. E. Sanchez. Resolution of composite fuzzy relation equations. *Information and Control*, 30(1):38–48, 1976.
12. B.-S. Shieh. Solution to the covering problem. *Information Sciences*, 222(0):626–633, 2013.
13. E. Shivanian. An algorithm for finding solutions of fuzzy relation equations with max-lukasiewicz composition. *Mathware and Soft Computing*, 17:15–26, 2010.
14. E. Turunen. On generalized fuzzy relation equations: necessary and sufficient conditions for the existence of solutions. *Acta Universitatis Carolinae. Mathematica et Physica*, 028(1):33–37, 1987.
15. Q.-Q. Xiong and X.-P. Wang. Solution sets of fuzzy relational equations on complete Brouwerian lattices. *Information Sciences*, 177(21):4757–4767, 2007.
16. Q.-Q. Xiong and X.-P. Wang. Fuzzy relational equations on complete Brouwerian lattices. *Information Sciences*, 193(0):141–152, 2012.
17. C.-T. Yeh. On the minimal solutions of max-min fuzzy relational equations. *Fuzzy Sets and Systems*, 159(1):23–39, 2008.

# The Existence of Generalized Inverses of Fuzzy Matrices

Miroslav Ćirić and Jelena Ignjatović

University of Niš, Faculty of Sciences and Mathematics
Višegradska 33, 18000 Niš, Serbia
{miroslav.ciric, jelena.ignjatovic}@pmf.edu.rs

**Abstract.** In this paper we show that all fuzzy matrices with entries in a complete residuated lattice possess the greatest generalized inverses of certain types, and we determine criteria for the existence of the greatest generalized inverses of other types. Moreover, we provide an iterative method for computing these greatest generalized inverses, which terminates in a finite number of steps, for example, for all fuzzy matrices with entries in a complete Heyting algebra.

## 1 Introduction and preliminaries

It is well-known that all systems composed of Moore-Penrose equations are solvable for matrices over a field of complex numbers. This implies the existence of all types of generalized inverses defined by these systems, such as the $g$-inverse, outer inverse, reflexive $g$-inverse, last-squares $g$-inverse, minimum-norm $g$-inverse, and Moore-Penrose inverse. Although the group inverse does not necessarily exist, the Drazin inverse always exists. However, the situation is completely different when the generalized inverses are considered in the context of semigroups, the most general context in which they are studied. None of these types of generalized inverses does not necessarily exist in a semigroup, or an involutive semigroup.

The aim of this paper is to show that fuzzy matrices, with entries in an arbitrary complete residuated lattice, are somewhere between. It is easy to see that fuzzy matrices always possess certain types of generalized inverses, such as generalized inverses defined by the equation (2), or those defined by some of the equations (3), (4) and (5) given below. For example, the zero matrix is always such a generalized inverse. However, we will show that fuzzy matrices also have other inverses of these types, and in particular, we show that they possess the greatest such inverses. The equation (1) behaves differently than others, and those types of generalized inverses whose definitions include this equation do not necessarily exist. Here we determine criteria for the existence of these types of generalized inverse, including the criteria for the existence of all previously listed important types of generalized inverses. In addition, we provide methods for computing the greatest inverses of these types. The method is iterative and does not necessarily terminate in a finite number of steps for every fuzzy matrix, but it terminates, for example, for all fuzzy matrices with entries in a complete Heyting algebra.

Notice that various types of generalized inverses of fuzzy matrices, mainly those with entries in Heyting algebras, the Gödel structure and Boolean matrices, have been studied in [1–7, 10], but here we use an original approach, different than the approaches that were used in the mentioned papers.

Throughout this paper, $\mathbb{N}$ will denote the set of all natural numbers, and for any $n \in \mathbb{N}$ we write $[1, n] = \{k \in \mathbb{N} \mid 1 \leqslant k \leqslant n\}$.

A *residuated lattice* is an algebra $\mathbb{L} = (L, \wedge, \vee, \otimes, \to, 0, 1)$ such that

(L1) $(L, \wedge, \vee, 0, 1)$ is a lattice with the least element $0$ and the greatest element $1$,
(L2) $(L, \otimes, 1)$ is a commutative monoid with the unit $1$,
(L3) $\otimes$ and $\to$ satisfy the *residuation property*: for all $x, y, z \in L$,

$$x \otimes y \leqslant z \;\Leftrightarrow\; x \leqslant y \to z.$$

If, additionally, $(L, \wedge, \vee, 0, 1)$ is a complete lattice, then $\mathbb{L}$ is called a *complete residuated lattice*. A (complete) residuated lattice in which the operations $\otimes$ and $\wedge$ coincide is called a (*complete*) *Heyting algebra*.

Other important special cases of complete residuated lattices, defined on the real unit interval $[0, 1]$ with $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$, are the *Łukasiewicz structure* ($x \otimes y = \max(x + y - 1, 0)$, $x \to y = \min(1 - x + y, 1)$), and the *Gödel structure* ($x \otimes y = \min(x, y)$, $x \to y = 1$ if $x \leqslant y$ and $= y$ otherwise).

For a complete residuated lattice $L$ and $n \in \mathbb{N}$, the set of all $n \times n$ matrices with entries in $\mathbb{L}$ will be denoted by $\mathbb{L}^{n\times n}$. Such matrices will be called *fuzzy matrices*. For a fuzzy matrix $A \in \mathbb{L}^{n\times n}$ and $i, j \in [1, n]$, the $(i, j)$-entry of $A$ will be denoted by $A(i, j)$. Fuzzy matrices will be ordered coordinatewise: for $A, B \in \mathbb{L}^{n\times n}$, $A \leqslant B$ if and only if $A(i, j) \leqslant B(i, j)$, for all $i, j \in [1, n]$. Endowed with this ordering, $\mathbb{L}^{n\times n}$ forms a complete lattice in which the meet $\bigwedge_{i\in I} A_i$ and the join $\bigvee_{i\in I} A_i$ of a family $\{A_i\}_{i\in I}$ of fuzzy matrices are defined by

$$\big(\bigwedge_{i\in I} A_i\big)(i, j) = \bigwedge_{i\in I} A_i(i, j), \qquad \big(\bigvee_{i\in I} A_i\big)(i, j) = \bigvee_{i\in I} A_i(i, j).$$

for all $i, j \in [1, n]$. For $A \in \mathbb{L}^{n\times n}$, the set $(A] = \{X \in \mathbb{L}^{n\times n} \mid X \leqslant A\}$ will be called the *down set* determined by $A$. The product of two fuzzy matrices $A, B \in \mathbb{L}^{n\times n}$ is a fuzzy matrix $AB \in \mathbb{L}^{n\times n}$ defined by

$$AB\,(i, j) = \bigvee_{k=1}^{n} A(i, k) \otimes B(k, j),$$

for all $i, j \in [1, n]$. For fuzzy matrices $A, B \in \mathbb{L}^{n\times n}$, the *right residual* of $B$ by $A$, denoted by $A\backslash B$, and the *left residual* of $B$ by $A$, denoted by $B/A$, are fuzzy matrices in $\mathbb{L}^{n\times n}$ defined by

$$(A\backslash B)(i, j) = \bigwedge_{k=1}^{n} A(k, i) \to B(k, j), \quad (B/A)(i, j) = \bigwedge_{k=1}^{n} A(j, k) \to B(i, k),$$

for all $i, j \in [1, n]$. It is easy to check that the following *residuation property* holds:

$$AB \leqslant C \;\;\Leftrightarrow\;\; A \leqslant C/B \;\;\Leftrightarrow\;\; B \leqslant A\backslash C,$$

for arbitrary fuzzy matrices $A, B, C \in \mathbb{L}^{n\times n}$.

The *transpose* of a fuzzy matrix $A \in \mathbb{L}^{n\times n}$ is denoted by $A^*$.

## 2 The main results

Let us consider the equations

(1) $AXA = A$,
(2) $XAX = X$,
(3) $(AX)^* = AX$,
(4) $(XA)^* = XA$,
(5) $AX = XA$,

where $A \in \mathbb{L}^{n\times n}$ is a given fuzzy matrix and $X$ is an unknown fuzzy matrix taking values in $\mathbb{L}^{n\times n}$. For any $\gamma \subseteq \{1, 2, 3, 4, 5\}$, the system consisting of the equations $(i)$, for $i \in \gamma$, is denoted by ($\gamma$), and solutions to ($\gamma$) are called $\gamma$-*inverses* of $A$. The set of all $\gamma$-inverses of $A$ will be denoted by $A\gamma$.

Commonly, a $\{1\}$-inverse is called a *g-inverse* (abbreviation for "generalized inverse") or an *inner inverse*, a $\{2\}$-inverse is called an *outer inverse*, a $\{1, 2\}$-inverse is called a *reflexive g-inverse* or a *Thierrin-Vagner inverse*, a $\{1, 3\}$-inverse is known as a *last-squares g-inverse*, a $\{1, 4\}$-inverse is known as a *minimum-norm g-inverse*, a $\{1, 2, 3, 4\}$-inverse is known as a *Moore-Penrose inverse* or shortly a *MP-inverse* of $A$, and a $\{1, 2, 5\}$-inverse is known as a *group inverse* of $A$. If $A$ has at least one $\gamma$-inverse, then it is said to be $\gamma$-*invertible*. In particular, an element having the MP-inverse is called *MP-invertible*. It is worth noting that an element having a $\{1\}$-inverse is called a *regular element*.

If they exist, the Moore-Penrose inverse and the group inverse of a matrix $A$ are unique, and they are denoted respectively by $A^\dagger$ and $A^\sharp$.

It is easy to see that the zero matrix (matrix whose all entries are equal to 0) is a solution to equations (2), (3), (4) and (5), as well as to any system composed of some of these equations. However, we will show that these equations and related systems also have the greatest solutions, for an arbitrary fuzzy matrix $A$.

**Theorem 1.** *Let $A \in \mathbb{L}^{n\times n}$ be an arbitrary fuzzy matrix. Then the following is true*

(a) *the matrix $A$ has the greatest $\{2\}$-inverse;*
(b) *the matrix $A$ has the greatest $\gamma$-inverse, for each $\gamma \subseteq \{3, 4, 5\}$.*

*Proof.* (a) Let $\phi : \mathbb{L}^{n\times n} \to \mathbb{L}^{n\times n}$ be a mapping defined by $\phi(X) = XAX$, for every $X \in L^{n\times n}$. Then $\phi$ is an isotone mapping and the set of $\{2\}$-inverses of $A$ is equal to the set of fixed points of $\phi$. Since $\mathbb{L}^{n\times n}$ is a complete lattice, by the Knaster-Tarski theorem (Theorem 12.2 [8]) we obtain that there exists the greatest fixed point of $\phi$, i.e., the greatest $\{2\}$-inverse of $A$.

(b) We will prove the existence of the greatest $\{3, 4, 5\}$-inverse. All other cases can be proved in the same way.

It is clear that $(AX)^* = AX$ if and only if $AX \leqslant (AX)^*$, which is equivalent to $X \leqslant A\backslash(AX)^*$. Similarly we show that $(XA)^* = XA$ is equivalent to $X \leqslant (XA)^*/A$.

On the other hand, the equation (5) is equivalent to the system of inequalities $AX \leqslant XA$ and $XA \leqslant AX$, which are equivalent to $X \leqslant A\backslash(XA)$ and $X \leqslant (AX)/A$.

Therefore, the system consisting of equations (3), (4) and (5) is equivalent to the system of inequalities

$$X \leqslant A\backslash(AX)^*, \;\; X \leqslant (XA)^*/A, \;\; X \leqslant A\backslash(XA), \;\; X \leqslant (AX)/A,$$

which is equivalent to the single inequality

$$X \leqslant A\backslash(AX)^* \wedge (XA)^*/A \wedge A\backslash(XA) \wedge (AX)/A.$$

Define now a mapping $\phi : \mathbb{L}^{n\times n} \to \mathbb{L}^{n\times n}$ by

$$\phi(X) = A\backslash(AX)^* \wedge (XA)^*/A \wedge A\backslash(XA) \wedge (AX)/A.$$

Then $\phi$ is an isotone mapping and the set of all $\{3,4,5\}$-inverses of $A$ is the set of all post-fixed points of $\phi$, and again by the Knaster-Tarski theorem we obtain that there exists the greatest post-fixed point of $\phi$, i.e., there exists the greatest $\{3,4,5\}$-inverse of $A$. $\square$

Let us note that, by the Knaster-Tarski theorem, the greatest fixed point of the function $\phi$ defined in the proof (a) is also the greatest post-fixed point of this function. Consequently, the previous theorem also provides a method for computing the greatest $\{2\}$-inverse or the greatest $\gamma$-inverse, for $\gamma \subseteq \{3,4,5\}$, based on the Kleene's method for computing the greatest post-fixed point of an isotone mapping on a complete lattice. Namely, for any isotone mapping $\phi$ of $\mathbb{L}^{n\times n}$ into itself we define a sequence $\{X_k\}_{k\in\mathbb{N}}$ of matrices inductively, as follows:

$$X_1 = \phi(\mathbf{1}), \qquad X_{k+1} = \phi(X_k), \quad \text{for each } k \in \mathbb{N},$$

where $\mathbf{1}$ is the matrix whose all entries are 1 (the greatest matrix in $\mathbb{L}^{n\times n}$). If there exists $k \in \mathbb{N}$ such that $X_k = X_{k+1}$, then $X_k = X_{k+m}$, for each $m \in \mathbb{N}$, and $X_k$ is the greatest post-fixed point of $\phi$. In particular, this will happen whenever $\phi$ is defined as in the proof of (a) of Theorem 1 and $\mathbb{L}$ is a complete Heyting algebra, Łukasiewicz or Gödel structure. This will also happen whenever $\phi$ is defined as in the proof of (b) and $\mathbb{L}$ is the Gödel structure.

Next we aim our attention to the equation (1). For the sake of convenience, set $A^\diamond = (A\backslash A)/A = A\backslash(A/A)$.

**Theorem 2.** *A fuzzy matrix $A \in \mathbb{L}^{n\times n}$ is $\{1\}$-invertible if and only if $A^\diamond \in A\{1\}$.*

*If this is the case, $A^\diamond$ is the greatest $\{1\}$-inverse and $A^\diamond A A^\diamond$ is the greatest $\{1,2\}$-inverse of $A$.*

*Proof.* Clearly, if $A^\diamond \in A\{1\}$, then $A$ is $\{1\}$-invertible. Conversely, if $A$ is $\{1\}$-invertible and $B \in A\{1\}$, then $B$ is a solution to the the inequality $AXA \leqslant A$. According to the residuation property, $A^\diamond$ is the greatest solution to this inequality, whence $B \leqslant A^\diamond$. Now, $A = ABA \leqslant AA^\diamond A$, and thus, $A^\diamond$ is the gratest $\{1\}$-inverse of $A$.

It is easy to check that $A^\diamond AA^\diamond$ is a $\{1,2\}$-inverse of $A$, and if $B$ is an arbitrary $\{1,2\}$-inverse of $A$, then it is a $\{1\}$-inverse of $A$, whence $B \leqslant A^\diamond$, and thus, $B = BAB \leqslant A^\diamond AA^\diamond$. Therefore, $A^\diamond AA^\diamond$ is the greatest $\{1,2\}$-inverse of $A$. □

For Boolean matrices, a similar characterization of $\{1\}$-inverses and $\{1,2\}$-inverses can be derived from a theorem proved by B. M. Schein in [9] which concerns Boolean-valued relations. It is interesting to note that for a Boolean matrix $A$ the matrix $A^\diamond$ can be represented by $A^\diamond = (A^* A^c A^*)^c$, where $A^c$ is a matrix obtained from $A$ replacing every entry in $A$ by its complement in the two-element Boolean algebra (replacing 1 by 0 and 0 by 1).

**Theorem 3.** *Let $\gamma \subseteq \{3,4,5\}$, $\gamma_1 = \gamma \cup \{1\}$ and $\gamma_{1,2} = \gamma \cup \{1,2\}$, and let $A \in \mathbb{L}^{n\times n}$ be an arbitrary fuzzy matrix. Then the following is true:*

(a) *There exists the greatest $\gamma$-inverse $G$ of $A$ in the down-set $(A^\diamond]$;*
(b) *If $A \leqslant AGA$, then $G$ is the greatest $\gamma_1$-inverse and $GAG$ is the greatest $\gamma_{1,2}$-inverse of $A$;*
(c) *If $A \leqslant AGA$ does not hold, then $A$ does not have any $\gamma_1$-inverse nor $\gamma_{1,2}$-inverse.*

*Proof.* (a) We will prove that there exists the greatest $\{3,4,5\}$-inverse $G$ in $(A^\diamond]$. All other cases can be proved in the same way.

According to Theorem 1, a matrix $B \in \mathbb{L}^{m\times n}$ is a solution to the system consisting of equations (3), (4) and (5) and $B \leqslant A^\diamond$ if and only if

$$B \leqslant A\backslash(AB)^* \wedge (BA)^*/A \wedge A\backslash(BA) \wedge (AB)/A \wedge A^\diamond.$$

Define now a mapping $\phi : \mathbb{L}^{n\times n} \to \mathbb{L}^{n\times n}$ by

$$\phi(X) = A\backslash(AX)^* \wedge (XA)^*/A \wedge A\backslash(XA) \wedge (AX)/A \wedge A^\diamond.$$

Then $\phi$ is an isotone mapping and the set of all $\{3,4,5\}$-inverses $B$ of $A$ contained in $(A^\diamond]$ is the set of all post-fixed points of $\phi$, and by the Knaster-Tarski theorem we obtain that there exists the greatest post-fixed point $G$ of $\phi$, and therefore, $G$ is the greatest $\{3,4,5\}$-inverse of $A$ contained in $(A^\diamond]$.

(b) By $G \leqslant A^\diamond$ it follows that $AGA \leqslant AA^\diamond A \leqslant A$, and if $A \leqslant AGA$, then it is clear that $G$ is a $\gamma_1$-inverse of $A$ and it is easy to check that $GAG$ is a $\gamma_{1,2}$-inverse of $A$. Since every $\gamma_1$-inverse is also $\gamma$-inverse of $A$, and $G$ is the greatest $\gamma$-inverse of $A$, we conclude that $G$ is the greatest $\gamma_1$-inverse of $A$. On the other hand, if $H$ is an arbitrary $\gamma_{1,2}$-inverse of $A$, then it is a $\gamma$-inverse of $A$, so $H \leqslant G$, and hence, $H = HAH \leqslant GAG$, which means that $GAG$ is the greatst $\gamma_{1,2}$-inverse of $A$.

(c) As it was noted in the proof of (b), if $H$ is an arbitrary $\gamma_1$-inverse of $A$, then $H \leqslant G$, whence $A = AHA \leqslant AGA$. Therefore, if $A \leqslant AGA$ does not hold, then $A$ does not have any $\gamma_1$-inverse nor $\gamma_{1,2}$-inverse. □

**Corollary 1.** *Let $A \in \mathbb{L}^{n\times n}$ and let $G$ be the greatest $\{3,4\}$-inverse of $A$ in the down-set $(A^\diamond]$. Then the following is true:*

(a) *If* $A \leqslant AGA$*, then* $GAG$ *is the Moore-Penrose inverse of* $A$*;*
(b) *If* $A \leqslant AGA$ *does not hold, then* $A$ *does not have a Moore-Penrose inverse.*

**Corollary 2.** *Let* $A \in \mathbb{L}^{n\times n}$ *and let* $G$ *be the greatest* $\{5\}$*-inverse of* $A$ *in the down-set* $(A^{\diamond}]$*. Then the following is true:*

(a) *If* $A \leqslant AGA$*, then* $GAG$ *is the group inverse of* $A$*;*
(b) *If* $A \leqslant AGA$ *does not hold, then* $A$ *does not have a group inverse.*

## References


1. Cen, J-M.: Fuzzy matrix partial orderings and generalized inverses, Fuzzy Sets and Systems 105, 453–458 (1999)
2. Cui-Kui, Z.: On matrix equations in a class of complete and completely distributive lattices, Fuzzy Sets and Systems 22, 303–320 (1987)
3. Cui-Kui, Z.: Inverses of L-fuzzy matrices, Fuzzy Sets and Systems 34, 103–116 (1990)
4. Han, S.-C., Li, H.-X., Wang, J.-Y.: Resolution of matrix equations over arbitrary Brouwerian lattices, Fuzzy Sets and Systems 159, 40–46 (2008)
5. Hashimoto, H.: Subinverses of fuzzy matrices, Fuzzy Sets and Systems 12, 155–168 (1984)
6. Kim, K.H., Roush, F.W.: Generalized fuzzy matrices, Fuzzy Sets and Systems 4, 293–315 (1980)
7. Prasada Rao, P.S.S.N.V., Bhaskara Rao, K.P.S.: On generalized inverses of Boolean matrices, Linear Algebra and its Applications 11, 135–153 (1975)
8. Roman, S.: Lattices and Ordered Sets, Springer, New York, 2008.
9. Schein, B.: Regular elements of the semigroup of all binary relations, Semigroup Forum 13, 95-102 (1976)
10. Wang, Z.-D.: T-type regular L-relations on a complete Brouwerian lattice, Fuzzy Sets and Systems 145, 313–322 (2004)

# The IMBPC HVAC system: A Complete MBPC Solution for Existing HVAC Systems

António E. Ruano

Centre for Intelligent Systems, LAETA
Instituto Superior Técnico and Universidade do Algarve, Portugal
aruano@ualg.pt

## Abstract

According to recent studies, energy consumption of buildings (residential and non-residential) represents approximately 40% of total world energy consumption, half of this energy consumed by HVAC systems operation. It is therefore of fundamental importance to control efficiently the existing HVAC systems, in order to decrease energy usage and to increase compliance with the European Directives on the energy performance of buildings and energy efficiency.

Model Based Predictive Control (MBPC) is perhaps the most proposed technique for HVAC control, since it offers an enormous potential for energy savings. This talk will introduce the Intelligent MBPC (IMBPC) HVAC system, a complete solution to enable MBPC of existing HVAC installations in a building. The IMPBC HVAC minimizes the economic cost needed to maintain controlled rooms in thermal comfort during the periods of occupation. The hardware and software components of the IMBPC system are described, with a focus on the MBPC algorithm employed, and the design of Computational Intelligence predictive models.

The installation of IMBPC HVAC solution in a University building by a commercial company is described, and the results obtained in terms of economical savings and thermal comfort obtained are compared with standard, temperature regulated control.

# Wavelet Analysis and Structural Entropy Based Intelligent Classification Method for Combustion Engine Cylinder Surfaces

Szilvia Nagy and Levente Solecki

Széchenyi István University
H-9026 Győr, Hungary
{nagysz,solecki}@sze.hu

**Abstract.** Structural entropy is a good candidate for characterizing roughness of surfaces as it is sensitive not only to the general shape of the surface, but also to the rate of the high and low surface points. Wavelet analysis of the surface can separate the larger-scale behavior from the fine details, and together with the structural entropy it can define a behavior profile for the surface which is typically slightly different for new and for worn tribological surfaces.
An intelligent fuzzy classification scheme is introduced to characterize surfaces according to both their degree of wear and method of the surface measurement. The basis of the classification is the structural entropies of the original and the first wavelet transform of the height scan of the new and worn surfaces.

**Keywords:** Rényi entropy, wavelet analysis, surface classification, fuzzy classification.

## 1 Introduction

The structure of the cylinder surfaces of the combustion engines determine both the tribological and mechanical properties of the motor, thus the mileage and the emission of pollutants [1]. Measuring surface microgeometry or roughness can be carried out multiple ways, typically either with a touching needle (or ball) scanning the surface or optically (with a laser or light scanner). Both the processes can be executed either on the surface itself, or on a silicon replica, if the geometry of the object does not allow to access the surface itself. In case of the cylinder inner surface, if a direct (not replica-based) method is to be used, the cylinder has to be cut in order to make it scannable for the equipment.

The aim of this article is to develop a classification method that can determine whether a surface is worn or new, and which can serve later as a basis for determination of the grade of the wear. We have studied surfaces of worn and freshly grooved finish. In case of the worn surface we used two types of image acquiring methods: a touching scanner with needle diameter of 5 microns and an optical scanner applied on a silicon replica of the surface. The later method was used on the new surface, as for the touching roughness measurement the motor should have been cut. Two sample surfaces can be seen in Fig. 1.

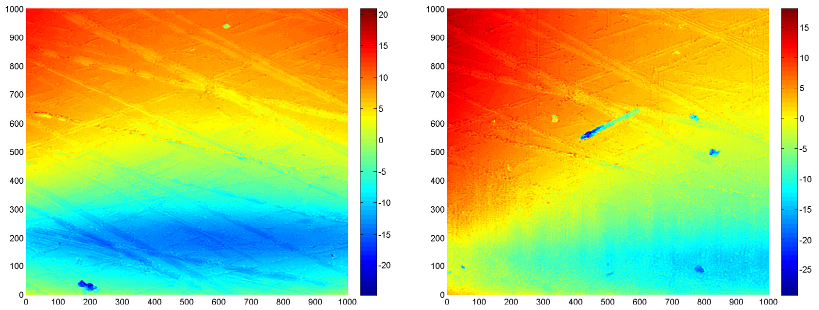


**Fig. 1.** Scanned surfaces before and after running of the motor. Optical scanner, from a silicon imprint of the surface. The units are microns.

In order to characterize the surface, multiple surface roughness measures were introduced from a simple difference of the highest and lowest points to the more complex topological measures. Rényi entropies are good candidates for measuring surface roughness [2]. The aim of this article is to find a method which can characterize surface roughness from images of the surfaces. As a first step, we generate a Rényi entropy based fuzzy characterization scheme, and next we improve it by using the entropies of the wavelet transforms of the pictures.

## 2 Structural entropy

The structural entropy was introduced in solid state physics [3] for characterizing various electron localization types. Later it was extended for microscopy applications [4, 2, 5]. For a probability distribution – which can be any shifted, rescaled surface – the structural entropy is defined the following way, Let the surface height $H_i$ of the $i$th point from the $N$ points fulfill

$$H_i \geq 0, \qquad \text{for } i = 1, \ldots, N \tag{1}$$

$$\sum_{i=1}^{N} H_i = 1. \tag{2}$$

Structural entropy is the difference of two Rényi entropies $S_i = \frac{1}{1-N} \ln \sum_{j=1}^{N} H_j^i$

$$S_{str} = S_1 - S_2. \tag{3}$$

Similarly, the difference of the zeroth and 2nd entropies can be related to a so called filling factor $q$ as

$$\ln(q) = S_0 - S_2. \tag{4}$$

The filling factor is also known form physics of electron structures as it shows the rate of the filled points of the electron distribution compared to all the points,

$$q = \frac{\left(\sum_{n=1}^{N} H_i^2\right)^{-1}}{N}. \tag{5}$$

$q$ fulfills the inequality

$$\frac{1}{N} \leq q \leq 1. \tag{6}$$

However, if the two quantities are combined, a more powerful tool is presented: any type of distribution forms a characteristic line in the $S_{str}(\ln q)$ map.

The structural entropy map of the worn and the new surfaces can be seen in Fig. 2. Each point represents a picture. It can be seen that the domains of the two variables very strongly overlap, thus these surfaces can hardly be characterized according to these measures. The results are similar for the images taken by the two different method as it can be learnt from the second subplot of Fig. 2.

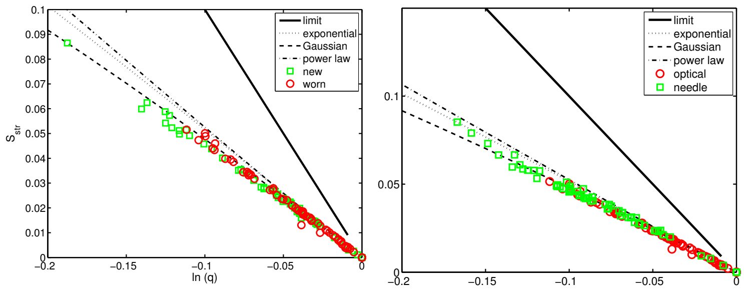



**Fig. 2.** First subplot: structural entropy vs. logarithm of the filling factor for the 2×64 scanned surfaces before and after running of the motor. Second subplot: the images of the worn surface taken according to 2 methods, by optical scanner from silicon mold samples, and by a touching scanner with needle point diameter of 5 microns. As a reference the curves of the exponential, Gaussian and 2nd order power law behaviours are also plotted. The theoretical limiting line of the possible $S_{str}(\ln q)$ points is shown with thick solid line.

## 3 Fuzzy classification scheme

64 pictures of each types were studied, half of them were selected to determine the fuzzy rules, the other half for testing the results. We have generated a simple set of fuzzy rules for both of the image pairs, similarly to [6–8]. We show only the results of the worn-new pairs in Fig. 3. The membership functions are different, thus there is a hope that using fuzzy inference, the classification can be carried out. However, the results are not promising: from the 64 pictures 1 was not classifiable as one of its antecedent parameters had 0 membership value, and half of the images were classified to a wrong group.

## 4 Wavelet analysis

In the electrical engineering and signal processing practice wavelet analysis is a series of high-pass and low-pass filter pairs. The only extraordinary features are the filter

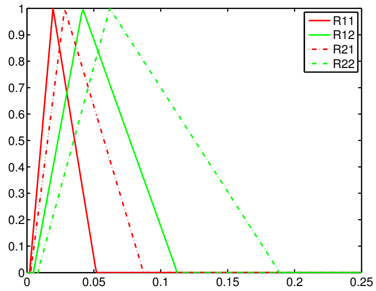



**Fig. 3.** Membership functions of the fuzzy rules, with the first index meaning the type of the image, i.e., new (solid line) or worn (dash-dotted line), the second the type of the function used, i.e., structural entropy (red) or logarithm of the filling factor (green).

coefficients, which can be adapted to the task, from a simple averaging of the two neighboring points to highly specialized image processing tools. After each filtering half of the resulting points is sorted out, i.e., the result is downsampled. These pairs of filters can be applied after each other, resulting in a series, where the first high-pass output gives the finest details, the second the next finest resolution details, while the low-pass outputs are filtered further. The last low-pass output gives the average behavior of the signal.

As the difference between the two types of surfaces lay probably in the fine details, we apply wavelet-analysis to the images and carry out the above, entropy-based fuzzy classification scheme with more antecedent variables.

In our case, as the signal is two-dimensional, i.e., a picture, both the rows and the columns of the input data are to be wavelet transformed. This results in 4 outputs, one, where low-pass filtering was applied in both dimensions, one where the first dimension was low-pass filtered, the second went through a high-pass step, etc. The structural entropy maps of all 4 resulting images after a Daubechies-2 [9] wavelet transform can be seen in Fig. 4. It is clear, that in case of the high-pass–high-pass part the domains on the two types of surfaces are different.

In the fuzzy classification scheme we selected the data of the original image, their low-pass–low-pass wavelet transform and the high-pass–high-pass transformed pictures. The two remaining data set does not give extra information according to the $S_{str}$ maps. Using the same 32–32 images for generating the rules as in the previous case, the resulting membership functions are given in Fig. 5. The results are significantly better: the number of false characterization decreased to 12 pictures from the previous 32, which suggests, that including another step in wavelet transform or another type of roughness measure will lead to a reliable method for our purposes.

## 5 Conclusions

We have developed a fuzzy classification scheme for surface roughness scanner images, that is able to distinguish the method of taking as well as the condition of the cylinder

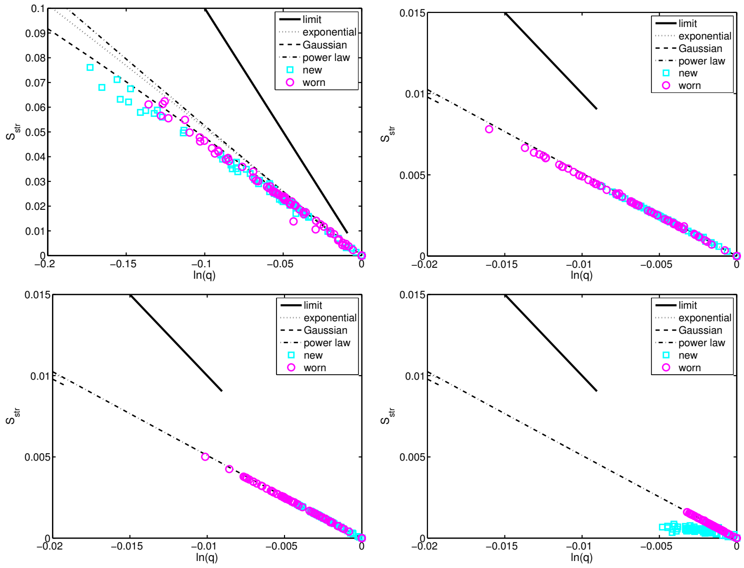



**Fig. 4.** Structural entropy vs. logarithm of the filling factor for the wavelet transforms of the $2 \times 64$ scanned surfaces before and after running of the motor. The upper left picture is the low-pass–low-pass component, the lower right one is the high-pass–high-pass component, and the other two subplots are the mixed low-pass–high-pass, and high-pass–low-pass ones.

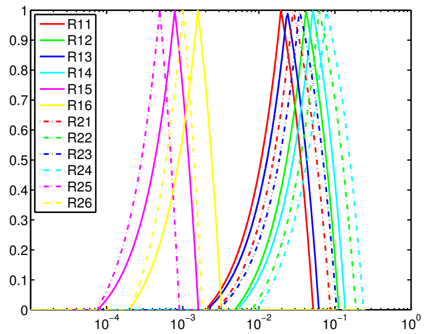



**Fig. 5.** Membership functions of the extended fuzzy rules. The first indices mean the type of the image (new (solid line) – worn (dash-dotted line)), the second indices and the color of the lines denote the type of the function used (original: $S_{str}$ (red) – $\ln(q)$ (green); wavelet transform, low-pass: $S_{str}$ (blue) – $\ln(q)$ (cyan); wavelet transform, high-pass: $S_{str}$ (magenta) – $\ln(q)$ (yellow). Logarithmic scale is used on the horizontal axis due to better visibility of the lower valued terms, however, the rules are simple triangle rules.

surface with a quite good accuracy. The method uses the structural entropy based surface characterization which can determine the general shape of the surface, together with wavelet analysis, which can separate the various scale behavior of the system. As the images have rather large fluctuation due to environmental effects and the precision demands of the grooving, the structural entropy maps of the different images overlap, however, applying a fuzzy classification with structural entropy and filling factor of the original image and the detail component of the wavelet transformed image being the antecedent parameters is able to distinguish the two types of images.


## Acknowledgements

The authors are grateful for Professor Lászlo T. Kóczy and Ferenc Lilik for their consultations about fuzzy inference systems, for Professor János Pipek for his advices in structural entropy related problems and Professor Veska Georgieva for her valuable comments on image processing related questions.


## References


1. M.R. Dreyer, L. Solecki, "Verschleissuntersuchungen an Zylinderlaufbahnen von Verbrennungsmotoren," 3. Symposium Produktionstechnik – Innovativ und Interdisciplinär, 6-7.04.2011, Zwickau, pp. 69-74.
2. A. Bonyár, L.M. Molnár, G. Harsányi, "Localization factor: a new parameter for the quantitative characterization of surface structure with atomic force microscopy (AFM),", *MICRON*, Vol. 43, pp. 305-310, 2012.
3. J. Pipek, and I. Varga, "Universal classification scheme for the spatial localization properties of one-particle states in finite d-dimensional systems," *Phys. Rev. A*, Vol. 46, 3148–3164, 1992.
4. I. Mojzes, Cs. Dominkonics, G. Harsányi, Sz. Nagy, J. Pipek, and L. Dobos, "Heat treatment parameters effecting the fractal dimensions of AuGe metallization on GaAs," *Appl. Phys. Lett.*, Vol. 91, No. 7, article No. 073107, 2007.
5. A. Bonyár, "AFM characterization of the shape of surface structures with localization factor,", *MICRON*, Vol. 87, pp. 1-9, 2016.
6. F. Lilik, J. Botzheim, "Fuzzy based Prequalification Methods for EoSHDSL Technology," *Acta Technica Jauriensis*, Series Intelligentia Computatorica, Vol. 4, pp. 135-145, 2011.
7. F. Lilik, L.T. Kóczy, "The Determination of the Bitrate on Twisted Pairs by Mamdani Inference Method," Issues and Challenges of Intelligent System and Computational Intelligence, Studies in Computational Intelligence, vol. 530, 2014, pp. 59-74, doi: 10.1007/978-3-319-03206-1_5
8. F. Lilik, Sz. Nagy, L.T. Kóczy, "Wavelet Based Fuzzy Rule Bases in Pre-qualification of Access Networks' Wire Pairs," IEEE Africon 2015, Addis Ababa, Ethiopia, 14-17 September 2015.
9. I. Daubechies, "Ten Lectures on Wavelets," CBMS-NSF regional conference series in applied mathematics 61, SIAM, Philadelphia, 1992.

# Improving Twitter Gender Classification using Multiple Classifiers *

Marco Vicente[1,2], Fernando Batista[1,2], and Joao P. Carvalho[1,3]

[1] L$^2$F – Spoken Language Systems Laboratory, INESC-ID Lisboa
[2] Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
[3] Instituto Superior Técnico, Universidade de Lisboa, Portugal
m.vicente.pt@gmail.com,
{fernando.batista,joao.carvalho}@inesc-id.pt

**Abstract.** The user profile information is important for many studies, but essential information, such as gender and age, is not provided when creating a Twitter account. However, clues about the user profile, such as the age and gender, behaviors, and preferences, can be extracted from other content provided by the user. The main focus of this paper is to infer the gender of the user from unstructured information, including the username, screen name, description and picture, or by the user generated content. Our experiments use an English labelled dataset containing 6.5M tweets from 65K users, and a Portuguese labelled dataset containing 5.8M tweets from 58K users. We use supervised approaches, considering four groups of features extracted from different sources: user name and screen name, user description, content of the tweets, and profile picture. A final classifier that combines the prediction of each one of the four previous partial classifiers achieves 93.2% accuracy for English and 96.9% accuracy for Portuguese data.

**Keywords:** gender classification, Twitter users, gender database, text mining.

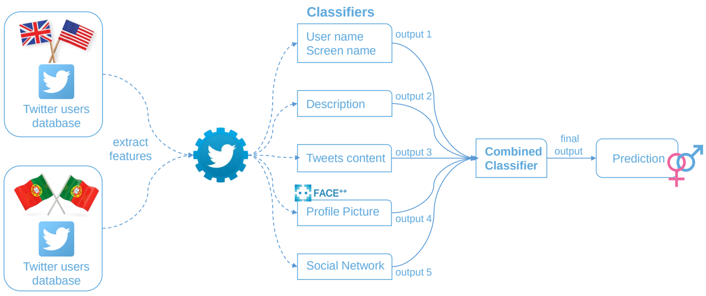
## 1 Introduction

Unlike other social networking services, the information provided by Twitter about a user is limited and does not specifically include relevant information. Such information is part of what can be called the user's profile, and can be relevant for a large spectra of social, demographic, and psychological studies about users' communities [6]. When creating a Twitter profile, the only required field is a user name. There are not specific fields to indicate information such as gender. Nevertheless, gender information is most of the times provided wittingly or unwittingly by the user. Knowing the gender of a Twitter user is essential for social networking studies, and useful for online marketing and opinion mining.

Our main goal is to automatically detect the gender of a Twitter user (male or female), based on features extracted from other profile information, profile picture, and the text content produced by the user. Previous research on gender detection is restricted

* Work supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) under project PTDC/IVC-ESCT/4919/2012 and reference UID/CEC/50021/2013.

**Fig. 1.** Combined classifier that merges the output of individual classifiers.




to features from the user generated content or from textual profile information. A relevant aspect of this study is that it involves a broader range of features, including automatic facial recognition from the profile picture. We have considered five different groups of features that were used in five separate classifiers. A final classifier, depicted in Fig. 1, combines the output of the other five classifiers in order to produce a final prediction.

This study was conducted for English and Portuguese users that produce georeferenced tweets. English is the most used language in Twitter, with 38% of the georeferenced tweets and, according to a study on 46 million georeferenced tweets [10], Portuguese is the third most used, with 6% of the georeferenced tweets. Portuguese is a morphologically rich language, contrarily to English, so interesting conclusions arise when comparing the performance achieved for both languages. Most of the previous research uses small labelled datasets, making it difficult to extract relevant performance indicators. Our study uses two large manually labelled datasets, containing 55K English and 57K Portuguese users. The proposed approach for gender detection is based on language independent features, apart from a language-specific dictionaries of first names, and can be easily extended to other Indo-European languages.

## Related work

The problem of gender detection has been previously applied to Twitter. The first gender detection study applied to Twitter users was presented by [14]. The features used for gender detection were divided in four groups: network structure, communication behavior, sociolinguistic features and the content of users' postings. They achieved an accuracy of 72.3% when combining ngram-features with sociolinguistic features using the stacked Support Vector Machine based classification model.

The state-of-the-art study of [5] collected a multilingual dataset of approximately 213M tweets from 18.5M Twitter users labelled with gender. The features were restricted to word and character ngrams from tweet content and three Twitter profile fields: *description*, *screen name* and *user name*. When combining tweet text with profile information (*description*, *user name* and *screen name*), they achieved 92% of accuracy, using Balanced Winnow2 classification algorithm. [1] proposes the use of features related to the principle of homophily. This means, to infer user attributes based on the immediate neighbors' attributes using tweet content and profile information. The experiments were performed using a Support Vector Machine-based classifier and the accuracy of their prediction model was of 80.2% using neighborhood data and 79.5% when using user data only. The improvement was not considerable. [2] studies gender detection suggesting a relationship between gender and linguistic style. The experiments were performed using a logistic regression classifier and the accuracy obtained was of 88.0%. Like [1], they also study gender homophily and have the same conclusion, the homophily of a user's social network does not increase minimally the accuracy of the classifier. [9] proposes the use of neural network models for gender identification. Their limited dataset was composed of 3031 manually labelled tweets. They applied both Balanced Winnow and Modified Balanced Winnow models. Using Modified Balanced Winnow with feature selection, 53 ngram features were chosen, they achieved an accuracy of 98.5%. In a consecutive work, [13] proposes the use of stream algorithms with ngrams. They manually labelled 3000 users, keeping one tweet from each user. They use Perceptron and Naïve Bayes with character and word ngrams. They report an accuracy of 99.3% using Perceptron when tweets' length is of at least 75 characters.

Recently, some studies suggest other possible features to infer gender. [3] studied the relationship between gender, linguistic style, and social networks. They reported an accuracy of 88%. [11] studies gender classification using celebrities the user follows as features combined with tweets content features. The accuracy achieved with Support Vector Machine-based classifiers using tweets content features is of 82%. When combined with the proposed features based on the followed celebrities, the accuracy increased to 86%. [12] proposes a method to extract user attributes from the pictures posted in Twitter. They created a dataset of 10K labelled users with tweets containing visual information. Using visual classifiers with semantic content of the pictures, they achieved an accuracy of 76%. Complementing their textual classifier with visual information features, the accuracy increased from 85% to 88%.

## 2 Data

Experiments here described use both Portuguese and English labelled datasets from a previous study [16]. The English dataset contains 65k labelled users and the Portuguese 58k labelled users. In order to be able to train and validate the classifiers, the datasets were divided into three subsets: training, development and test.

## 3 Features

Twitter does not provide gender information, though the gender can be inferred from the tweets' content and the profile information. In this section, we describe the features we extract from each group of attributes. Features are distributed in the following groups: *user name* and *screen name*, *description*, tweet content, profile picture and social network.

*User name and screen name.* We extracted features based in self-identified names found in the *user name* and *screen name* with gender association, as proposed in our previous work [15]. In order to associate names with the corresponding gender, we used a dictionary of English names and a dictionary of Portuguese names. Both dictionaries contain *gender* and *number of occurrences* for each of the names, and focus on names that are exclusively male or female. The English names dictionary contains 8444 names. It was compiled using the list of the most used baby names from the United States Social Security Administration. The dictionary is composed of 3304 male names and 5140 female names. The Portuguese names dictionary contains 1659 names, extracted from Baptista et al. [4]. The dictionary is composed of 875 male names and 784 female names. The *user name* and *screen name* are normalized for repeated vowels (e.g.: "eriiiiiiiiic"→"eric") and "leet speak" [8] (e.g.: "3ric"→"eric"). After finding one or more names in the *user name* or *screen name*, we extract the applicable features from each name by evaluating the following elements: "case", "boundaries", "separation" and "position". The final model uses 192 features.

*User description.* Users might provide clues of their gender in the description field. Having up to 160 characters, the description is optional. An example of user description is "I love being a mother.Enjoy every moment.". The word "mother" might be a clue to a possible female user. In order to extract useful information, we start by preprocessing the description and then we extract word unigrams, bigrams and trigrams from the preprocessed description field. We also use word count per tweet and smileys as features. Portuguese words tend to have suffixes to convey information such as gender or person and nouns inflect according to grammatical gender. For the Portuguese dataset, we also extract features related to these cases. Accordingly, if a description contains a female article followed by a word ending with the letter "a", the feature A_FEMALE_NOUN is triggered.

*Content of the tweets.* Features extracted from tweets' content can be divided in two groups: i) textual ngram features, like used in [5], or ii) content, style and sociolinguistic features, like emoticons, use of repeated vowels, exclamation marks or acronyms, as used in [14]. For both the textual ngram features and the style and sociolinguistic features, we only used the last 100 tweets from each labelled user. To extract textual features from tweets, we start by preprocessing the text. Retweets are ignored and the preprocessed text is used to extract unigrams, bigrams and trigrams based only on words. Though we only use word ngrams, it is advised to use character ngrams when analyzing tweets in languages like Japanese, where a word can be represented with only one character. In the study of [5], count-valued features did not improve significantly the performance. Accordingly, we also associate a boolean indicator to each feature, representing the presence or absence of the ngram in the tweet text, independently from the number of occurrences of each ngram. Besides word ngram features, we also extract content-based features, style features and sociolinguistic features that can provide gender clues. [7] suggests word-based features and function words as highly indicative of gender. We extract a group of features which include, social networks features, style features, character and word features.

*Profile picture feature.* Profile pictures have not been used in previous studies of gender detection of Twitter users, due to several reasons. One of the first reasons is that the profile picture is not mandatory. Also, many users tend to use profile pictures of celebrities or characters from movies and TV series. A third reason is because the picture may not be gender indicative. While the profile picture might not be good discriminating gender by itself, when combined with the other features, it might help increase significantly the accuracy of the prediction. Face++ (http://www.faceplusplus.com) is a publicly available facial recognition API that can be used to analyze the users' profile picture. We have used this tool through its API to extract the gender and the corresponding confidence. Such info was stored in our datasets. The API was invoked with the profile picture URL available on the last tweet of each user. In some cases, the API does not detect any face in the picture. 36% of the users in both datasets had no face detected. In the English dataset, more male users (34%) than female users (29%) have a profile picture with a recognizable face. In the Portuguese dataset, the opposite occurs, more female users (35%) than male users (30%) have a profile picture with a recognizable face.

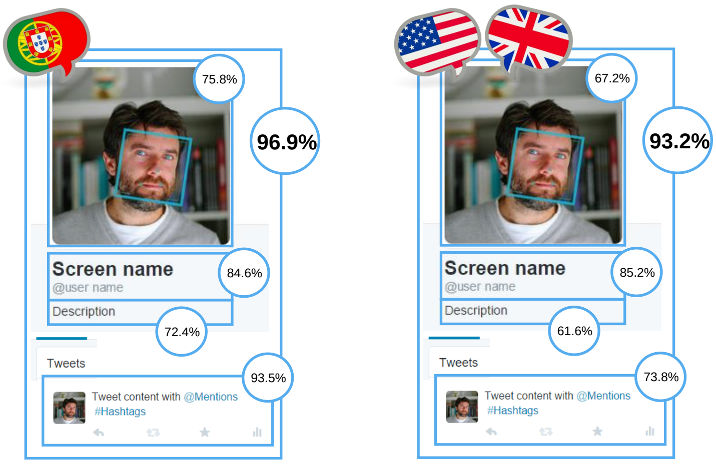
*Social network features.* Social network features consist in extracting the information related with the interaction between the user and other Twitter users. We extract the following attributes: Number of followers; Number of users followed; Follower-following ratio; Number of retweets; Number of replies; Number of tweets. These features alone might not be effective, but combined with the other features, could increment the global performance. We explored the extracted social network features, but we found out that these features were not indicative of gender. We observed no differences in the social network feature values between male and female. These results are consistent with the study of [14] that have analyzed users' network structure and communication behavior and observed the inability to infer gender from those attributes.

## 4 Experiments and results

Experiments here described use WEKA (http://www.cs.waikato.ac.nz/ml/weka) and the evaluation is performed using *Precision*, *Recall*, *F-Measure* and *Accuracy*. The combined classifier, shown in Fig. 1, receives as input the results obtained in the separate classifiers. The social network features were discarded. The separate classifiers are only used if information is available. E.g.: if a user has no description, the input from that classifier will be empty. Each classifier sends as output the confidence obtained in the classification. The values range from zero to one. If the confidence is of 100% in the class "Female," the value 1 is sent. If the confidence is of 100% in the class "Male," the value 0 is sent. If the confidence is not 100%, the values are adjusted accordingly. When the confidence received is of 0.5, we remove the input. We used an SVM to evaluate the combined classifier.

Fig. 2 summarizes the achieved accuracies per classifier for both datasets. In the Portuguese dataset we obtain 96.9% of accuracy. Only using tweets content, we already achieved an accuracy of 93.5%, but we improved the global accuracy. The experiments with the English dataset obtain an accuracy of 93.2%. With separate features, the best

**Fig. 2.** Classification accuracy per group of features for both datasets.




result was 85.2% using *user name* and *screen name* features. A good performance, since not all users self-assign a name in their profile information.

## 5 Conclusions

This study describes a method for gender detection using a combined classifier. We have used extended labelled datasets from our previous works [15, 17], partitioned into train, validation and test subsets. Instead of applying the same classifier for all features, we have grouped related features, used then in separate classifiers and then used the output of each classifier as input for the final classifier. In the Portuguese dataset, using only the tweet's text content achieves a baseline of 93.5% accuracy, but our combined classifier achieved an improved performance of 96.9% accuracy. The experiments with the English dataset achieve 93.2% accuracy. The features proposed, including the user name, screen name, profile picture and description, can be all extracted from a single tweet, except for the user text content. We successfully built two combined classifiers for gender classification of Portuguese and English users and, to our best knowledge, we provided the first study of gender detection applied to Portuguese Twitter users.

## References


1. Al Zamal, F., Liu, W., Ruths, D.: Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors. ICWSM 270 (2012)
2. Bamman, D., Eisenstein, J., Schnoebelen, T.: Gender in twitter: Styles, stances, and social networks. CoRR abs/1210.4567 (2012)
3. Bamman, D., Eisenstein, J., Schnoebelen, T.: Gender identity and lexical variation in social media. Journal of Sociolinguistics 18(2), 135–160 (2014)

4. Baptista, J., Batista, F., Mamede, N.J., Mota, C.: Npro: um novo recurso para o processamento computacional do português. In: XXI Encontro APL (Dec 2005)
5. Burger, J.D., Henderson, J., Kim, G., Zarrella, G.: Discriminating gender on twitter. In: EMNLP 2011. pp. 1301–1309. EMNLP '11, Association for Computational Linguistics, Stroudsburg, PA, USA (2011)
6. Carvalho, J.P., Pedro, V., Batista, F.: Towards intelligent mining of public social networks' influence in society. In: IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS). pp. 478 – 483. Edmonton, Canada (June 2013)
7. Cheng, N., Chandramouli, R., Subbalakshmi, K.: Author gender identification from text. Digital Investigation 8(1), 78–88 (2011)
8. Corney, M.W.: Analysing e-mail text authorship for forensic purposes. Ph.D. thesis, Queensland University of Technology (2003)
9. Deitrick, W., Miller, Z., Valyou, B., Dickinson, B., Munson, T., Hu, W.: Gender identification on twitter using the modified balanced winnow. Communications and Network 4(3) (2012)
10. Leetaru, K., Wang, S., Cao, G., Padmanabhan, A., Shook, E.: Mapping the global twitter heartbeat: The geography of twitter. First Monday 18(5) (2013)
11. Ludu, P.S.: Inferring gender of a twitter user using celebrities it follows. arXiv preprint arXiv:1405.6667 (2014)
12. Merler, M., Cao, L., Smith, J.R.: You are what you tweet... pic! gender prediction based on semantic analysis of social media images. In: Multimedia and Expo (ICME), 2015 IEEE International Conference on. pp. 1–6. IEEE (2015)
13. Miller, Z., Dickinson, B., Hu, W.: Gender prediction on twitter using stream algorithms with n-gram character features. International Journal of Intelligence Science 2(4A) (2012)
14. Rao, D., Yarowsky, D., Shreevats, A., Gupta, M.: Classifying latent user attributes in twitter. In: 2nd Int. Workshop on Search and Mining User-generated Contents. pp. 37–44. SMUC '10, ACM, New York, NY, USA (2010)
15. Vicente, M., Batista, F., Carvalho, J.P.: Twitter gender classification using user unstructured information. In: Proc. of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Istambul, Turkey (Aug 2015)
16. Vicente, M., Batista, F., Carvalho, J.P.: Creating extended gender labelled datasets of twitter users. In: IPMU 2016. Eindhoven, The Netherlands (June 2016)
17. Vicente, M., Carvalho, J.P., Batista, F.: Using unstructured profile information for gender classification of portuguese and english twitter users. In: SLATE'15. short papers, Madrid, Spain (June 2015)

# Exploiting Dynamics in Social Recommender Systems

Gergely Pósfai, Gábor Magyar, and Lászlo T. Kóczy

Department of Telecommunications and Media Informatics,
Budapest University of Technology and Economics
1111, Muegyetem rkp. 3, Budapest, Hungary
{posfai,magyar,koczy}@tmit.bme.hu
https://www.tmit.bme.hu

**Abstract.** In this paper we investigate the possibilities of improving social recommender systems through the exploitation of dynamic, temporal features. Dynamic effects are evaluated by empirically analyzing a variety of time-aware extensions of an existing information diffusion based fuzzy social recommender system method, called IDF-Social. We apply fuzzy subsets to model ratings and predictions in the recommendation process, where dynamic features influence the users' fuzzy membership values. Based on our rigorous experiments, we found that social recommender systems can benefit from incorporating dynamic features.

**Keywords:** dynamic fuzzy recommender systems, social network, information diffusion.

## 1 Introduction

In the past decade social and trust-aware recommender systems (RSs) have emerged. This group of RSs aim to leverage the users' social relations in order to improve recommendations [8, 4, 5, 1], which can be especially effective in certain cases, e.g. when producing recommendations for cold-start users [8, 4].

Most of the existing social and trust-aware RSs handle the users' social relations and preferences statically, assuming they do not change over time. In real life these networks rather than being static, they dynamically change: as time passes new friendships are made, but also existing friendships diminish or completely disappear. Additionally, user preferences also change over time.

In this paper we investigate the possibilities of taking dynamic effects into consideration during the recommendation process to enhance the performance of social RSs. We analyze the impact of temporal features via examining a range of dynamic extensions of an existing information diffusion based fuzzy social RS, called IDF-Social [9].

The rest of the paper is structured as follows. In Section 2 we briefly overview existing research on trust-aware and social RSs emphasizing dynamic aspects. In Section 3 we introduce the analyzed dataset. In Section 4 we describe how dynamic effects were incorporated in our methods. In Section 5 we present our experimental results. Finally, we conclude the paper in Section 6.

## 2 Related work

The main idea of trust-aware and social RSs is to incorporate the users' personal relations in the recommendation process [8, 4, 5, 1]. While trust-aware RSs make use of the users' web of trust, where connections between users are directed and indicate who trusts whom, social RSs leverage the users' undirected social network, in which bidirectional edges represent friendships. The two approaches are very similar, and sometimes not even distinguished [5].

Most of the existing trust-aware and social RSs handle the users' personal relations statically, without considering any temporal effects. However, many traditional RSs, i.e. RSs that ignore the users' friendships, have benefited from taking into account time-aware features [3, 7, 6]. Time-aware RSs can be grouped into two major classes [2]: *heuristic-based approaches*, which directly use the rating collection for computing predictions and *model-based approaches*, which use the rating collection to build models with which rating predictions are computed.

An important subgroup of *heuristic-based approaches* that relates most to our research is *time adaptive heuristics* [2], in which parameters and/or data are dynamically adjusted according to changes of some data characteristics through time [7, 6]. This type of RSs generally penalize older preferences that are presumed to be not/less valid at recommendation time, and usually utilize a continuous time representation. As opposed to our work, these existing RSs do not incorporate the users' social relations in the recommendation process. In our methods we apply *time adaptive heuristics* for weighting the users' relationships and preferences in the recommendation process.

## 3 Datasets

We conducted our experiments on the Flixster dataset[1] [5], which contains users' movie ratings (on a discrete scale between 0.5 and 5.0) and their freindship network. Movie ratings include user id, movie id, rating value and timestamp attributes, while the social network is given by a list of pairs of user ids representing bidirectional friendships. We reduced the original dataset and randomly selected 1500 movies for our analysis in order to be able to conduct enough number of tests in reasonable time. Also, we divided the dataset into a training set and a test set, consisting of the ratings between 06/01/2008 and 06/01/2009 and the ratings after 06/01/2009, respectively.

As described the only attribute in the dataset that contains temporal information is the timestamp of ratings, hence to incorporate dynamic effects in the recommendation process, we have to rely solely on this attribute.

## 4 Methodology

This section presents the dynamic extensions of the IDF-Social [9] method, which computes recommendations for users by propagating known ratings in the users' social network along friendships. We refer to the dynamic extensions of the IDF method as the D-IDF methods.

[1] http://www.cs.ubc.ca/ jamalim/datasets/

Formally, we have a set of users $U = \{u_1, u_2, ..., u_N\}$ and their friendship network transformed into a directed graph $G < U, E >$, where $E \subseteq U \times U$ is the set of edges derived by assigning directions to friendships in both ways. We denote the rating of user $u$ by $r_u$, which is a discrete value from the set of possible rating values $V = \{v_1, v_2, ..., v_M\}$. We have an initial set of users $U_0 = \{u_{r_1}, u_{r_2}, ..., u_{r_P}\}$ with observed ratings $R_0 = \{r_{u_{r_1}}, r_{u_{r_2}}, ..., r_{u_{r_P}}\}$. We also have another set of users $U_1 \subseteq U \setminus U_0$ with unknown ratings $R_1$. We aim to determine $R_1$ as accurately as possible based on knowledge of $R_0$ and $G < U, E >$.

We define fuzzy subsets on the set of users for every rating value:

$$W_{v_i} : U \to [0,1] \, , i = 1, 2, ..., M \; , \tag{1}$$

where $W_{v_i}(u_j)$ is the membership value of user $j$ in $W_{v_i}$ which is the fuzzy subset that belongs to the $v_i$ rating value. The membership values of the fuzzy subsets for a given user represent the user's – known or predicted – rating on the item, that is, how likely the user would give that rating value for the item. We calculate $W_{v_i}(u_j)$ as follows:

$$W_{v_i}(u) = \begin{cases} 0 & u \in U_0, r_u \neq v_i \; , \\ d_1(\Delta t(r_u)) & u \in U_0, r_u = v_i \; , \\ \vee^1_{q \in U_{v_i}} s(\vee^2_{p \in P_{q,u,d}} w(p)) & u \in (U \setminus U_0) \; , \end{cases} \tag{2}$$

where $\Delta t(r_u)$ represents the time between rating $r_u$ and the termination of the training dataset's period and $d_1$ is a dynamic function which weights known ratings based on their "age". $U_{v_i} \subseteq U_0$ is the set users who rated the item with value $v_i$, $P_{q,u,d}$ is the set of paths from user $q$ to user $u$ with length equal or less than $d$, which is an independent parameter. The $w$ function lowers membership values based on the propagation path, while $s$ denotes a transformation function which rescales membership values. The applied weighting function of (2) is

$$w(p) = \wedge(\alpha, w_{e_1}, \alpha, w_{e_2}, ..., \alpha, w_{e_l}) \, , p = (e_1, e_2, ..., e_l) \; , \tag{3}$$

$$w_{e_i} = d_2(\Delta t(r_{start(e_i)})) \; , \tag{4}$$

where $\wedge$ denotes a fuzzy t-norm aggregation, $\alpha$ is an independent parameter and $w_{e_i}$ denotes the weight of edge $e_i$, which is computed by the dynamic $d_2$ function with parameter $\Delta t(r_{start(e_i)})$, which is the time between the last known rating of the user the edge is starting from and the end of the training period.

The applied scaling function of (2) is

$$s(x) = \beta x \, , \beta \in [0,1] \; , \tag{5}$$

where $\beta$ is an independent parameter. Finally, the predictions are given as:

$$p_u = \frac{\sum_{i \in M} W_{v_i}(u) v_i}{\sum_{i \in M} W_{v_i}(u)} \; , \tag{6}$$

We also define a confidence value $c$ for each prediction, and keep only the ones that have higher confidence than a certain threshold, which controls the trade-off between prediction quality and quantity. We set the threshold for every method to achieve 40% coverage, i.e. being able to provide predictions for the 40% of the test cases. The confidence of the rating prediction for user $u$ is given as:

$$c_u = \sum_{i \in M} W_{v_i}(u) \; . \tag{7}$$

As can be seen, the $d_1$ and $d_2$ methods are responsible for taking into consideration temporal effects. While $d_1$ weights known ratings based on their "age", $d_2$ weights the outgoing edges of users in the social network based on the time of the users' ratings, assuming, that users with more recent ratings are more active, therefore their friendships are more valid, that is, we implicitly infer weights for friendships by examining the timestamps of ratings.

We defined 5 dynamic extensions of the IDF method, summarized in Table 1. The D-IDF-1 and D-IDF-2 methods weight ratings, the D-IDF-3 and D-IDF-4 methods weight friendships, while the D-IDF-5 method weights both ratings and friendships dynamically. $\Delta t$ is measured in months, and we set $\Delta t = 12$ for users without known ratings, as the training dataset spans only one year.

Based on extensive experiments we applied the following parameter settings in our final methods: $d = 2$, $\alpha = 0.2$, $\beta = 0.01$. The applied $\wedge$, $\vee^1$ and $\vee^2$ parameters are detailed in the next section.

Table 1: The applied time-aware functions.

| | **D-IDF-1** | **D-IDF-2** | **D-IDF-3** | **D-IDF-4** | **D-IDF-5** |
|---|---|---|---|---|---|
| $d_1(\Delta t)$ | $e^{-\lambda_1 \Delta t}$, $\lambda_1 = 0.001$ | $1.0 - \lambda_2 \Delta t$, $\lambda_2 = 0.001$ | 1 | 1 | $e^{-\lambda_1 \Delta t}$, $\lambda_1 = 0.001$ |
| $d_2(\Delta t)$ | 1 | 1 | $\lambda_3^{\Delta t}$, $\lambda_3 = 0.993$ | $1.0 - \lambda_4 \Delta t$, $\lambda_4 = 0.006$ | $\lambda_3^{\Delta t}$, $\lambda_3 = 0.993$ |

# 5 Experimental results

## 5.1 Comparisons with the Non-dynamic IDF-Social Method

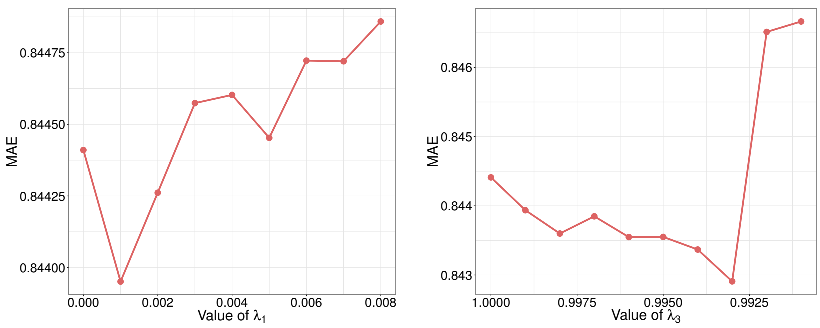
Table 2 shows the achieved *mean absolute error (MAE)* on the test set of the D-IDF methods compared to the MAE of the non-dynamic IDF-Social method. We applied the dynamic extensions for the top 4 fuzzy-operator combinations according to [9]. As can be seen, the dynamic extensions performed almost always better or equally good as the IDF-Social method. Interestingly the performance improved the most for the best fuzzy operator combination. The values in parentheses denote the applied hyperparameter of the corresponding fuzzy operator.

Table 2: The MAE of the most efficient fuzzy operator combinations at 40% coverage.

| $\wedge$ t-norm | $\vee^1$ t-conorm | $\vee^2$ t-conorm | IDF-S. | D-IDF-1 | D-IDF-2 | D-IDF-3 | D-IDF-4 | D-IDF-5 |
|---|---|---|---|---|---|---|---|---|
| Algebraic product | Schweizer & Sklar (-5.0) | Dombi (0.8) | 0.8444 | **0.8440** | **0.8440** | **0.8429** | **0.8432** | **0.8427** |
| Algebraic product | Yager (2.5) | Dombi (0.8) | 0.8448 | **0.8449** | **0.8449** | **0.8442** | **0.8441** | **0.8443** |
| Dombi (2.0) | Yager (2.5) | Dombi (0.8) | 0.8458 | **0.8457** | **0.8458** | **0.8458** | **0.8458** | **0.8457** |
| Schweizer & Sklar (-2.0) | Schweizer & Sklar (-5.0) | Dombi (0.8) | 0.8458 | **0.8458** | **0.8458** | **0.8456** | **0.8456** | **0.8456** |

## 5.2 Impact of dynamic parameters

Figure 1 shows the impact of dynamic parameters in the D-IDF-1 and D-IDF-3 methods. As can be seen, both extensions could decrease the MAE to a certain level, although the D-IDF-3 method achieved more improvement from time-aware features. It is important to note, that the dynamic parameters have to be adjusted carefully, as giving too much importance to time-decaying effects results in significant performance deterioration. The impact of $\lambda_2$ and $\lambda_4$ in the D-IDF-2 and D-IDF-4 methods are very similar to the impact of $\lambda_1$ and $\lambda_3$, respectively, however, due to space restrictions they are not included in the paper.




(a) Impact of parameter $\lambda_1$ in D-IDF-1. (b) Impact of parameter $\lambda_3$ in D-IDF-3.

Fig. 1: Impact of dynamic parameters. By the proper choice of their dynamic parameter the dynamic extensions of the IDF method could reduce the MAE to a certain level.

## 6 Conclusion

Trust-aware and social RSs have proved to be able to outperform traditional RSs in many cases. However, most of the existing trust-aware and social RSs lack the incorporation of dynamic, time-aware features. In this paper, we introduced and analyzed a range of dynamic extensions of the existing social RS method called IDF-Social, and showed that exploiting temporal effects can improve recommendation quality.

Our experiments showed promicing results, so we are planning to perform a deeper analysis of social dynamic effects in the future, by examining new datasets, where friendship related events have explicit timestamps.

**Acknowledgments.** The research was supported by National Research, Development and Innovation Office (NKFIH) K105529, K108405.

## References

1. Avesani, P., Massa, P., Tiella, R.: A trust-enhanced recommender system application: Moleskiing. In: Proceedings of the 2005 ACM symposium on Applied computing. pp. 1589–1593. ACM (2005)
2. Campos, P.G., Díez, F., Cantador, I.: Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols. User Modeling and User-Adapted Interaction 24(1-2), 67–119 (2014)
3. Ding, Y., Li, X.: Time weight collaborative filtering. In: Proceedings of the 14th ACM international conference on Information and knowledge management. pp. 485–492. ACM (2005)
4. Jamali, M., Ester, M.: Trustwalker: a random walk model for combining trust-based and item-based recommendation. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 397–406. ACM (2009)
5. Jamali, M., Ester, M.: A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of the fourth ACM conference on Recommender systems. pp. 135–142. ACM (2010)
6. Lathia, N., Hailes, S., Capra, L.: Temporal collaborative filtering with adaptive neighbourhoods. In: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. pp. 796–797. ACM (2009)
7. Lee, T.Q., Park, Y., Park, Y.T.: A time-based approach to effective recommender systems using implicit feedback. Expert systems with applications 34(4), 3055–3062 (2008)
8. Massa, P., Avesani, P.: Trust-aware collaborative filtering for recommender systems. In: On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE, pp. 492–508. Springer (2004)
9. Posfai, G., Koczy, L., Magyar, G.: Idf-social: an information diffusion-based fuzzy model for social recommender systems. In: The Congress on Information Technology, Computational and Experimental Physics (CITCEP 2015). pp. 106–112 (2015)

# Generalized Aggregation Functions and Quality Indicators in Overall Telecommunication Network Models: Some Open Issues

Stoyan Poryazov[1] and Emiliya Saranova[1,2]

[1] Bulgarian Academy of Science, Institute of Mathematics and Informatics, Sofia, Bulgaria
{stoyan,emiliya}@cc.bas.bg
[2] University of Telecommunication and Post, Sofia, Bulgaria
emilia.saranova@gmail.com

**Abstract.** In the present paper**,** an overall telecommunication system normalized performance model, which includes users, terminals and network equipment, has been considered. Two new overall network performance indicators have been proposed, expressed analytically and demonstrated numerically. The graphic presentations, as well as other considerations, show the necessity of a new aggregation function definition, have been presented accordingly. The results received are useful for Quality of Service prediction as a base for future Quality of Experience prognostication, using aggregation functions.

**Keywords:** overall telecommunication system, performance model, user model, quality of service, aggregation function.

## 1 Introduction

A telecommunication connection is establishing consecutively in many devices and stages, where there are many reasons for failures, with different probabilities. Users have to choose network provider, tariff scheme and terminal devices according their overall estimation of the Quality of Service (QoS). Simply put, users make aggregation of observed characteristics of the telecommunication services and make decisions. A very important task of the telecommunication researchers and designers is to predict the user estimation of QoS. In this paper, we try to analyze the aggregation functions approach, on the base of existing and two proposed indicators, aggregating a large number of failure probabilities, along with telecommunication connection and communication.

In Section 2, some Aggregation Function properties have been discussed. In Section 3, the used background model of the telecommunication system with QoS guarantees has been described on the conceptual level. In Section 4, four classical overall network QoS indicators have been mentioned, and two new ones have been proposed and expressed analytically. In Section 5, numerical predictions of the proposed indicators have been presented in three different cases. Section 6 contains a discussion of the numerical results and other interrelated considerations, showing the necessity of a new aggregation function definition, presented in Section 7. The conclusions can be found in Section 8.

## 2 On the aggregation function properties

Marichal in 2009 [4] has provided a great number of potential mathematical properties of the aggregation functions, but not a closed definition of such a function.

In tracing the history of aggregation function definitions, Mayor [6] notes the following properties: in 1984 – boundary conditions, monotonicity, symmetry; in ‘today’ – boundary conditions, monotonicity. In [5] they were – boundary conditions, monotonicity. We will demonstrate that the properties mentioned above are sometimes too restrictive and a more general definition of aggregation function may be more suitable.

## 3 Telecommunication system background model

This paper explores a model of overall telecommunication systems, including users, terminals, and a network with Quality of Service (QoS) guaranties [8]. Apart from GSM, BSDN and others, generalized virtual networks (VNET) with overall QoS guaranties, as an example of a telecommunication network, have also been considered. The conceptual model includes users' behavior, a limited number of homogeneous terminals; losses due to abandoned and interrupted dialing, blocked and interrupted switching, unavailable intent terminal, blocked and abandoned ringing and abandoned communication. In our approach, the network traffic, terminal traffic for A (calling) and B (called) terminals and users’ traffic have been divided and considered separately, in their interrelattionship.

### 3.1. Conceptual model

*3.1.1. Base Virtual Devices and Their Parameters*. We have used base virtual device types with names and graphic notation shown on Fig. 1. For every device we use the following notation for its parameters: $F$ stands for intensity of the flow [calls/second], $P$ = probability for directing the calls of the external flow to the device considered, $T$ = means service time, in the device, of a served call [seconds], $Y$ = intensity of the device traffic [Erlangs], $N$ = number of service places (lines, servers) in the virtual device (capacity of the device). In the normalized models [7], used in this paper, every virtual device, except the switch, has no more than one entrance and/or one exit. Switches have one entrance and two exits. For characterizing the intensity of the flow, we are using the following notation: $F$a for incoming flow, *dem.Fa*, and *rep.Fa* for demand and repeated flows respectively. The same characterization has been used for traffic intensity ($Y$).

*3.1.2. Virtual Base Device Names.* In the conceptual model each virtual device has a unique name, which is a concatenation of the corresponding first letters of the names of stages, branches end exits in the order shown in Fig. 1. For example, "*Yid*" means "traffic intensity in interrupted dialing” case.

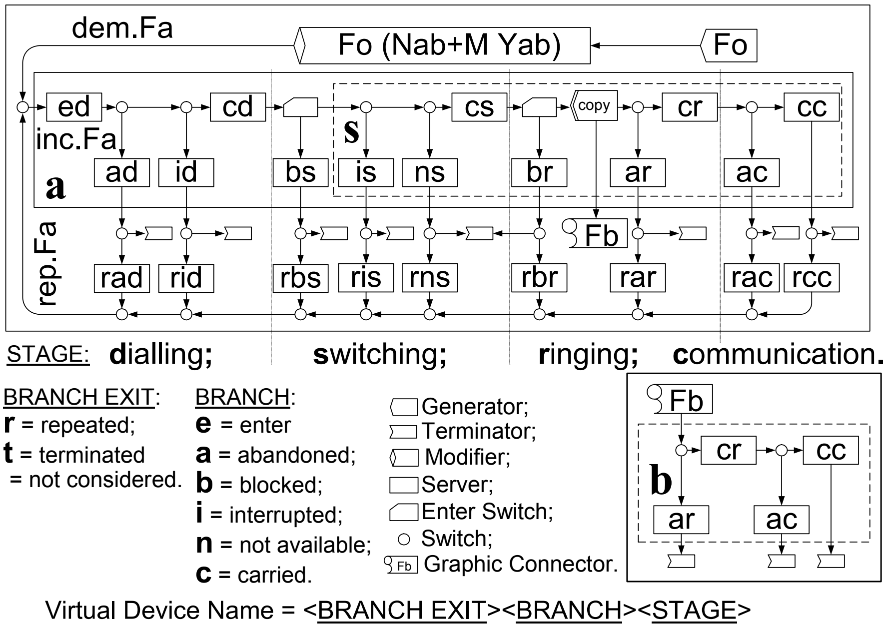



**Fig. 1.** Conceptual model of the telecommunication system, incl.: the paths of the calls, occupying A-terminals (**a**-device), switching system (**s**-device) and B-terminals (**b**-device); base virtual device types, with their names and graphic notation.

*3.1.3. The Paths of the Calls.* In this paper "call" means "call attempt" or "bid" according to [2]. Figure 1 shows the paths of the calls, generated from (and occupying) the A-terminals in the proposed network traffic model and its environment. *Fo* is the intent intensity of calls of one idle terminal; *M* is a constant, characterizing the BPP flow of demand calls (*dem.Fa*).

*3.1.4. The Comprising Virtual Devices and Their Names.* In this paper, we consider **b** = virtual device that comprises all the B-terminals (called) in the system (box with dashed line). From the important virtual devices, shown on Fig.1, comprising several base virtual devices, containing the paths of the calls occupying B-terminals and corresponding base virtual devices.

### 3.2. The analytical model

Based on the conceptual model, the analytical one [8] uses macro-state model of the system in stationary state, with: BPP input flow and repeated calls. Fourteen natural assumptions have been formulated. We consider the values of 10 basic dynamic parameters, which are mutually dependent: *Fo*, *Yab*, *Fa*, *dem.Fa*, *rep.Fa*, *Pbs*, *Pbr*, *ofd.Fs*, *Ts*, *ofd.Ys*. For these parameters, under assumptions made, we have a system of 9 equations with 6 generalized static parameters. We have chosen the intensity of the input calls flow *Fo* as the independent input variable. Thus, the considered system of equations has 9 equations and 9 output parameters with unknown values (the main output variables).

## 4 Overall network QoS indicators

### 4.1. Classical network efficiency indicators

The network efficiency classical indicators are ASR, ABR and NER, defined in [1]: Answer Seizure Ratio (ASR) = (number of seizures that result in an answer signal) / (the total number of seizures); Answer Bid Ratio (ABR) = (number of bids that result in an answer signal) / (total number of bids); ABR is similar to ASR except that it includes bids that do not result in a seizure; Network Effectiveness Ratio (NER) is designed to express the ability of networks to deliver calls to the far-end terminal; NER = (number of seizures) / (the sum of the number of seizures resulting in either an answer message, or a user busy, or a ring no answer, or in the case of ISDN a terminal rejection / unavailability). Thus, these indicators reflect network providers' attitude and exclude the possibilities of unsuccessful communication as well as the influence of repeated attempts.

### 4.2. Proposed network efficiency indicators

In this paper, we use output probabilities for blocked switching (*Pbs*) due to insufficient network equipment, and *Pbr* – probability for blocked ringing due to busy called B – terminal, as main input for analytical determination and numerical prediction of the two proposed overall network QoS indicators for call efficacy.

4.2.1. Successful Efficiency indicator

First, we have to determine the mean intensity of the input flow to the telecommunication system. This is the flow, occupying the calling (A) terminals, with frequency (rate) *Fa* which is a sum of demand (*dem.Fa*) and repeated attempts (*rep.Fa*):

$$Fa = dem.Fa + rep.Fa \tag{1}$$

On the other side:

$$rep.Fa = Fa\ Pr, \tag{2}$$

*Pr* (3) is aggregated probability of unsuccessful attempts to become repeated call attempts. It contains only probabilities:

$$\begin{aligned} Pr = {} & Pad\ Prad + (1 - Pad)\,[Pid\ Prid + (1 - Pid)[Pbs\ Prbs + \\ & + (1 - Pbs)[Pis\ Pris + (1 - Pis)[Pns\ Prns + (1 - Pns)[Pbr\ Prbr + \\ & + (1 - Pbr)[Par\ Prar + (1 - Par)[Pac\ Prac + (1 - Pac)Prcc]]]]]]]. \end{aligned} \tag{3}$$

The proof [9] is based on the direct description of the flows frequencies in Fig. 1 and the theorem of Little [3]. It follows from (1), (2) and (3) that:

$$dem.Fa = Fa\,(1 - Pr). \tag{4}$$

We define Successful Efficiency, extending ITU-T [2] successful call definition: "A call that has reached the wanted number and allows the conversation to proceed". Successful Efficiency (*Es*) (or Call Attempts Successfulness) is the ratio of all call attempts ended with successful finished communication (e.g. conversation) to all call attempts (*Fa*).

The frequency *Fcc* of the call attempts ended with successful finished communication corresponds to the flow in the base virtual device *cc* (carried communication) in Fig. 1. The analytical expression for Successful Efficiency (*Es*) is:

$$Es = (1 - Pad)(1 - Pid)(1 - Pis)(1 - Pbs)(1 - Pns)(1 - Par)(1 - Pbr)(1 - Pac) \qquad (5)$$

The main difference of Successful Efficiency (5) indicator from classical indicators mentioned, is including in the Successful Efficiency all possible factors making information transfer unsuccessful (e.g. uncompleted), after it had been initiated.

4.2.2. Demand Efficiency Indicator

We express the ratio, of the mean intensity of the generalized input flow (*Fa*) to the intensity of the demand calls (*dem.Fa*). This ratio is called 'Beta' in ITU recommendations and is used there as a parameter in calculations, not for indicator purposes:

$$\beta = \frac{Fa}{dem.Fa} = \frac{Fa}{Fa(1 - Pr)} = \frac{1}{1 - Pr} \qquad (6)$$

In normal conditions [9], when there is at least one successful call attempt in the system considered, $Pr < 1$.

Beta may be used as an independent indicator, but we propose a more general one: Demand Efficiency (*Ed*) or Demand Call Attempts Successfulness Indicator:

*Ed* = (Number of first (demand) call attempts cause successful communication) / (All call attempts made, for these successful demand call attempts):

$$Ed = \frac{Es}{\beta} = (1 - Pr)(1 - Pad)(1 - Pid)(1 - Pis)(1 - Pbs)(1 - Pns)(1 - Par)(1 - Pbr)(1 - Pac) \qquad (7)$$

In other words, Demand Efficiency (7) indicator presents the probability of a demand (first) call attempt to become a successful call attempt. It is a user-oriented indicator, compounding explicitly repeated attempts, connection and communication parameters.

## 5 Numerical prediction of the proposed indicators

In Figures 1, 2 and 3, numerical results are presented in the whole theoretical network load interval – terminal traffic (*Yab*) equals of 0% to 100% of the number of all terminals in the network. The input parameters are the same, excluding: (i) capacity of the network given as percentage of the number of all terminals in the system, causes blocking due to equipment insufficiency and (ii) the probability of repeated calls (*Pr*). The values of input parameters, in the presented numerical results, are typical for voice networks. Three cases have been considered:

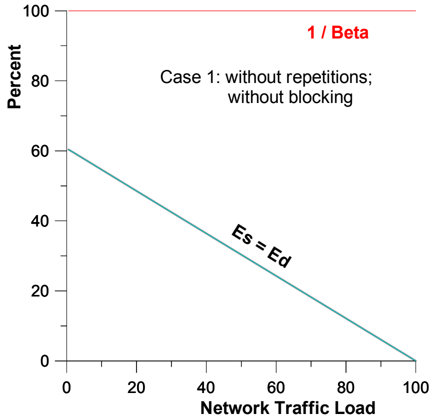



**Fig. 2.** Call Efficiency indicators 1/Beta, *Es* and *Ed* in Case 1. Beta = 1 and *Es* = *Ed*, because there are not repeated attempts in the system ($Pr = 0$). Beta is a constant, greater than any of its components. *Es* and *Ed* are decreasing monotonic functions.

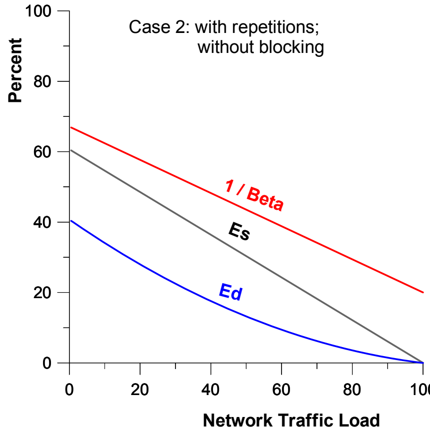



**Fig. 3.** Call Efficiency indicators 1/Beta, *Es* and *Ed* in Case 2. Repeated attempts make worse the performance considerably. *Ed*, *Es* and 1/Beta, are decreasing monotonic functions.

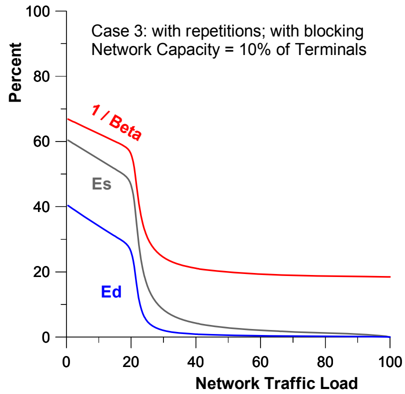



**Fig. 4.** Call Efficiency indicators 1/Beta, *Es* and *Ed*, in Case 3. Blocking sharply make worse the network performance. *Ed*, *Es* and 1/Beta, are decreasing functions.

## 6 Discussion of the results

From Fig. 2, 3 and 4, it is obvious that *Ed* is the most sensitive user-oriented connection and communication quality indicator, among those considered, of a telecommunication network with QoS guarantees.

- All indicators include a large number of probabilities, and they may be considered as aggregated probabilities.
- All indicators are continuous, decreasing and monotonic functions.
- Successful Efficiency is a symmetric function, but Demand Efficiency is not symmetric (*Pr* is not symmetric).

- Overall network performance indicators have the meaning and usage of aggregation functions, but they do not fulfill the formal aggregation properties, expressed in Section 2.
- “Such properties can be dictated by the nature of the values to be aggregated” [4], so the important conclusion of the paper is that the definition of aggregation functions have to be generalized.

Another example of very useful aggregation function in the telecommunication systems, is the mean occupation time of terminals, e.g. of the B-terminal (*Tb*). In the general case (with losses), it is aggregation of service times in the four base virtual devices (Fig. 1) [9]:

$$Tb = Par\,Tar + (1 - Par)[Tcr + Pac\,Tac + (1 - Pac)\,Tcc] \qquad (8)$$

In the case without losses *Tb* is a simple sum:

$$Tb = Tcr + Tcc \qquad (9)$$

**Note that:** It is unnecessary for *Tb* to be in the interval of [0,1] (the mean conversation is 3 minutes); the dimension of *Tb* is SI unit ‘Time’; *Tb* may be greater, equal or lower than some of the aggregated in it values, depending of the actual probabilities of losses: $Tb \in [Tar, Tcr + Tcc]$.

Other examples of aggregation by summation are ‘evaluation forms’ for experts reviewing the project proposals – the scores of every criterion are simply added, with or without weights.

## 7 Proposed definition of generalized aggregation function

We agree with Mayor [6]: “Aggregation is the process of combining several input values into a single representative output value, and the functions that carry out this process are called aggregation functions.”

Based on this general understanding, we propose the following definition:

Let $A(x_1,\dots,x_k, a_1,\dots,a_m)$ is a function of *k+m* arguments. *A* is Generalized Aggregation Function of the arguments $x_1, x_2,\dots,x_k$ if:

(i) The value of *A* and $x_1, x_2,\dots,x_k$ are of the same type (e.g. real) and of the same measurement unit; (ii) The value of *A* represents $x_1, x_2,\dots,x_k$, in any meaning.

Arguments $a_1, a_2\dots,a_m$ are not aggregated, but influence the aggregation. They are ‘aggregating’, and may be variables, not only fixed weights, see (7) and (8). A Generalized Aggregation Function must have at least one aggregated and one aggregating argument. An important example is the service time *T* of an aggregated device, containing: (i) one service device with demand service time *dem.T* and (ii) a cycle with infinity repetitions of probability *P<1* each:

$$T = \frac{dem.T}{1-P}$$

This aggregation example may be considered as infinity sum of weighted addends. This approach has been used many times in the practice, e.g. 9 times in Fig.1.

The proposed definition allows restrictive properties (for examples, see Section 2) and includes many aggregation functions, actually needed in the engineering and scientific practice (e.g. presented in this paper).

## 8 Conclusion

A model of Overall Telecommunication Systems, including users, terminals and network with Quality of Service (QoS) guarantees has been used for analyzing and development of overall network performance indicators. Two indicators have been proposed: Successful Efficiency and Demand Efficiency. Demand Efficiency is the most sensitive user-oriented connection performance quality indicator, among those considered, of a telecommunication network with QoS guaranties. It is a suitable QoS indicator for Quality of Experience (QoE) prediction.

The proposed indicators are aggregation functions, but they do not fulfill some restrictions in the usual aggregation function definitions. A definition of Generalized Aggregation Function has been proposed. It is more suitable for the engineering and scientific practice. The results received are useful for QoS prediction as a base for future Quality of Experience prognostication, using aggregation functions. The development of Generalized Aggregation Function Theory opens many research opportunities.

**Acknowledgments.** This work is coordinated under EU COST Action IC 1304 and partially financed by Bulgarian NSF Project DCOST 01/9

## References

1. ITU-T Rec. E.425 (03/2002). Internal automatic observations.
2. ITU-T Rec. E.600 (03/93). Terms and definitions of traffic engineering.
3. Little J. D. C., A Proof of the Queueing Formula L=λW. Operations Research, 9, 1961, 383-387.
4. Marichal, J.-L. Aggregation functions for decision making, arXiv:0901.4232v1, 2009.
5. Martines-Panero, M., J. L. Garsia-Lapresta, L. C. Meneses. Multidistances and Dispersion Measures. In: Tomasa Calvo Sánchez, Joan Torrens Sastre (Editors) – 2016 Fuzzy Logic and Information Fusion: To commemorate the 70th birthday of Professor G. Mayor. Studies in Fuzziness and Soft Computing, Springer 2016, 123-134.
6. Mayor, G., E. Trillas. One of the First Spaniards Interested in the Study of Aggregation functions. In: R. Seising, L. Arguelles Mendes (Eds). Accuracy and Fuzziness – A Life in Science and Politics, Studies in Fuzziness and Soft Computing, Springer 2015, 141-148.
7. Poryazov S. A., 2001. On the Two Basic Structures of the Teletraffic Models. Conference "Telecom'2001" - Varna, Bulgaria, 10-12 October 2001, 435-450.
8. Poryazov, S., Saranova, E. Some General Terminal and Network Teletraffic Equations in Virtual Circuit Switching Systems. Editors: N. Ince, E. Topuz, Springer Sciences + Business Media, Chapter 24, LLC 2006, pp. 471-505.
9. Poryazov S., E. Saranova. Models of Telecommunication Networks with Virtual Channel Switching and Applications. Prof. Marin Drinov Academic Publ. H., Sofia, 2012, pp. 238 (in Bulgarian).

# Author Index