7th European Symposium on
Computational Intelligence and Mathematics

# ESCIM 2015

**Cádiz, Spain**

**October 7th-10th, 2015**

**Proceedings**

Editors:
László Kóczy, Jesús Medina

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

Proceedings of ESCIM 2015


Edition: 1st
First published: 2015

ISBN: 978-84-608-2823-5

Published and printed by:
Universidad de Cádiz (Dept. Matemáticas), Spain

# Organization

## General Chairs

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

## International Program Commitee

| | |
|---|---|
| Pedro Cabalar | Universidad de A Coruña, Spain |
| Agata Ciabattoni | TU Wien, Austria |
| Davide Ciucci | University of Milano-Bicocca, Italy |
| Bernard De Baets | University of Gante, Belgium |
| Chris Cornelis | Universidad de Granada, Spain |
| Christian G. Fermueller | TU Wien, Austria |
| Péter Foldesi | Univ. Széchenyi István, Hungary |
| Lluis Godo | Artificial Intelligence Research Institute, Spain |
| László T. Kóczy | Univ. Széchenyi István, Hungary |
| Stanislav Krajci | UPJS Kosice, Slovakia |
| Ondrej Kridlo | UPJS Kosice, Slovakia |
| Piotr Kulczycki | Cracow University of Technology, Poland |
| Inmaculada Medina | Universidad de Cádiz, Spain |
| Jesús Medina | Universidad de Cádiz, Spain |
| Manuel Ojeda-Aciego | Universidad de Málaga, Spain |
| David Pearce | Universidad Politécnica de Madrid, Spain |
| Jozef Pócs | Slovak Academy of Sciences, Slovakia |
| Claudiu Pozna | Transilvania University of Brasov, Romania |
| Alex Tormási | Univ. Széchenyi István, Hungary |
| Esko Turunen | Tampere University of Technology, Finland |
| Agustín Valverde | Universidad de Málaga, Spain |

## Organizing Commitee

| | |
|---|---|
| Jesús Medina | Universidad de Cádiz, Spain |
| Valentín Liñeiro-Barea | Universidad de Cádiz, Spain |
| María Eugenia Cornejo-Piñero | Universidad de Cádiz, Spain |
| Juan Carlos Díaz-Moreno | Universidad de Cádiz, Spain |
| Ignacio García-García | Universidad de Cádiz, Spain |
| Eloísa Ramírez-Poussa | Universidad de Cádiz, Spain |
| María José Benítez-Caballero | Universidad de Cádiz, Spain |

## Sponsoring Institutions

Universidad de Cádiz, Spain
Széchenyi István University (Gyor)
Hungarian Fuzzy Association

# Table of Contents

# Preface

Mathematics is an indispensable field for a lot of areas such as Engineering, Computer Science, Physics, Chemistry and Business, in which improves the current methodologies and solves new challenges.

An important branch in Computer Science is Computational Intelligence, whose aim is to provide methods to deal with complex real-world problems for which traditional approaches are not feasible. Some of the methods that Computational Intelligence encompasses are, among others, fuzzy logic, evolutionary computation, neural networks, as well as probabilistic and statistical approaches, such as Bayesian networks or kernel methods.

In the past few years Computational Intelligence has become one of the main research topics at the Széchenyi István University. The first six Györ Symposia on Computational Intelligence have been successfully organized from 2008 to 2014. The seventh Györ Symposium on Computational Intelligence is jointly held with the fourth International Workshop on Mathematics and Soft Computing and it is called the 7th European Symposium on Computational Intelligence and Mathematics (ESCIM 2015). The location has been changed but preserves the philosophy of the past Györ Symposia enriching 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.

ESCIM 2015 will be held in Cádiz from October 7th to 10th, 2015, and it is organized by members of the University of Cádiz, Spain.

This symposium proceedings volume contains the contributions presented during ESCIM 2015, which have been included in different sections:

- Decision-Making under Uncertainty and Data Mining
- Evolutionary Computation, Metaheuristics and Machine Learning
- Software Verification and Validation
- Computational Optimization
- Computational Intelligence
- Mathematics and Soft Computing
- Formal Concept Analysis
- Graded Algebras and Algebras Admitting Multiplicative Bases
- Generalized Convexity and Fuzzy or Interval Valued Applications

We would like to thank the plenary speakers for their outstanding contributions to research and leadership in their respective fields. There were six plenary lectures covering the different areas of the symposium in charge of prestigious researches such as László Kóczy, Manuel Ojeda-Aciego, David Pearce, Jozef Pocs, Sandra Sandri and Nagy Szilvia.

We would also like to thank all the participants for their contributions to the symposium program and all the authors for their submitted papers. We are

also indebted to the special session organizers and 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 (Györ) and the Hungarian Fuzzy Association.

October 2015

Jesús Medina
Conference Chair
ESCIM 2015

# Keynote Speakers

## László T. Kóczy

1. Department of Information Technology, Széchenyi István University, Györ, Hungary
2. Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Budapest, Hungary

TITLE: Classification and recognition of movement sequences

ABSTRACT: Signal processing and classification as a sub problem of signal processing are well researched areas, but new methods and concepts are presented still today. Handwritten characters are satisfying the definition of signal, if we consider it as a chronologically ordered list of two dimensional coordinate pairs. The recognition of such handwritten characters could be solved with methods known from signal processing and classification. This work outlines hand-printed (non cursive) character recognition from a signal processing aspect starting with the introduction of various types of handwritten gestures and characters, then a short overview of issues and challenges (e.g. input quality, segmentation, pre-processing etc) of handwriting recognition with common solutions. It is followed by a brief summary of concepts of some uni- and multi-stroke character recognizers to present some examples found in literature. After that a single- and multi-stroke recognizer family (so called Fuzzy-Based Recognizer or FUBAR) is shown in details based on the Ph.D. dissertation of Alex Tormasi with a high emphasis on the basic concept of the recognition method, the construction of initial fuzzy rule bases with statistical and metaheuristic (bacterial evolutionary algorithm, big bang-big crunch, imperial competitive algorithm, particle swarm optimization) methods. Finally the properties of the metaheuristic methods based on the experiences from the research of FUBAR are also summarized.

## Manuel Ojeda-Aciego

Department of Applied Mathematics, University of Málaga. Spain

TITLE: Adjoints and non-canonical reasoning

ABSTRACT: Computational intelligence must necessarily deal with reasoning mechanisms with are outside the realm of classical logic. Several logical approaches have been developed to reproduce different aspects of non-canonical reasoning. In this talk, we will argue on the usefulness of mathematical tools, in particular the use of adjoints, to formalize some approaches to reason under imprecision, uncertainty, and lack of information.

## Sandra Sandri

Brazilian National Institute for Space Research (INPE), São José dos Campos, SP, Brazil

Title: Current Trends on Computational Intelligence for Space Research in Brazil

Abstract: The main driver of innovation in Space Research in Brazil is INPE, the Brazilian National Institute for Space Research. It is focused in areas such as meteorology and climate change, atmospheric science, space science and space engineering. It also provides services such as weather and climate monitoring, satellite tracking and control, and measuring the amount of forest fires, deforestation, lightnings and air pollution occurring in Brazil. This talk is divided into two parts; first, I will briefly present INPE, followed by some of its Computational Intelligence applications under development. These applications include the use of Neuro-Fuzzy Systems for the prediction of regime change in chaotic systems (related to meteorological phenomena), and the use of evolutionary techniques, such as Genetic Algorithms and Particle Swarms, to learn parameters for radar imagery filters.

## Jozef Pócs

Palacky University Olomouc. Czech Republic

Title: Fuzzy concept analysis with preference relations

Abstract: Formal Concept Analysis and its various fuzzy (many-valued) modifications represent methods of data analysis for identifying conceptual structures among data sets. A preference relation, either on objects or attributes, can be seen as an additional information, which should be included to a creation process of a concept lattice. From an algebraic point of view, we discuss some possibilities to include preferences into fuzzy concept lattices. The main emphasis will be on the so-called one-sided concept lattices.

## David Pearce

Universidad Politécnica de Madrid, Spain

Title: On Logics for Trust and Honesty

Abstract: In this talk I present two extensions of the well-known modal logics of trust from (Liau, 2003) formed by adding further axioms. The idea is to interconnect the trust modality with the individual belief modalities of agents. In the first of our logics we capture the idea that if an agent i trusts an agent j with respect to a statement p then i believes that j does not disbelieve p; while in

the second logic if i trusts j concerning p, then i believes that j believes p. In this way we can explicate a type of trust that is linked to honesty or sincerity. These quite intuitive extensions of the logic of trust help to solve some unintuitive consequences that arise when the semantics of trust and belief are independent. As a technical result we establish the soundness and completeness of these logics with respect to semantics based on neighourhood frames.

**Nagy Szilvia**

Faculty of Engineering Sciences. Univ. Széchenyi István, Hungary

TITLE: Wavelets and their possibilities in computational intelligence

ABSTRACT: Wavelet transform is ideal for image and data processing in many ways. It is easy to calculate wavelet transforms by rather simple convolutional filters. Wavelets are organized in resolution levels, and the wavelet transformed values correspond to spatial (temporal) positions - the higher the resolution level, the denser the grid of these positions. It is easy to find edges, patterns, average behavior and fine-scale behavior in a function or image with wavelets. In many inference systems, the way of processing of data is the key, and wavelets are very good candidates for this purpose.

By omitting wavelet coefficients near zero, the compression of data is also possible. Selecting these non-important coefficients might be done by evolutionary algorithms.

Wavelets are also suitable for solving differential equations. It is usual to solve a problem at various resolutions, however, local refinements are also possible, moreover, prediction of the next finer resolution level coefficients is a recent advancement, where computational intelligence can have significant role.

# New research on the probabilistic *p*-center problem

Maria Albareda-Sambola[1], Luisa I. Martnez-Merino[2], and Antonio M. Rodríguez-Chía[2]

[1] Universitat Politècnica de Catalunya. BarcelonaTech, Barcelona, Spain,
maria.albareda@upc.edu
[2] Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, Spain,
luisa.martinez@uca.es, antonio.rodriguezchia@uca.es

**Abstract.** This work deals with the p-center problem, where the aim is to minimize the maximum distance between any user and his center taking into account that the demand occurs in any site with a specific probability. The problem is of interest for the location of emergency centers. We consider different formulations for the problem and extensive computational tests are reported, showing the potentials and limits of each formulation on several types of instances. Finally, some improvements on these formulations have been developed obtaining in some cases much better resolution times.

**Keywords:** Location theory, p-center problem, uncertainty.

## 1 Introduction

Facility location models have been extensively studied in the literature. Different kinds of facilities have been modeled, such as routers or servers in communication network, warehouses or distribution centers in a supply chain, hubs or transhipment nodes in passenger transport, and hospital or emergency facilities in a public service system, among others. In general, the goal of these types of problems is to locate the facilities among a set of candidate sites and assign customers to the facilities optimizing some effectivity measure that usually depends of the distances between the facilities and the customers, see for instance [1, 2] and the references therein.

The $p$-center problem (pCP) is a well-known discrete optimization location problem which consists of locating $p$ centers out of $n$ sites and assigning (allocating) the remaining $n - p$ sites to the centers so as to minimize the maximum distance (cost) between a site and the corresponding center, see [3, 4]. The uncertainties can be generally classified into three categories: provider-side uncertainty, receiver-side uncertainty, and in-between uncertainty. In this paper we focus on the receiver-side uncertainty.

## 2 The problem

Let $N = \{1, \ldots, n\}$ be the given set of sites or customers. Throughout the paper we assume, without loss of generality, that the set of candidate sites for centers is identical to $N$. Let $p \geq 2$ be the number of centers to be located. For each pair $(i, j)$, $i, j \in N$, let $d_{ij}$ be the distance (cost, travel time) from $i$ to $j$. We assume $d_{ii} = 0 \ \forall i \in N$ and $d_{ij} > 0 \ \forall i, j \in N : \ i \neq j$. We do not assume other special properties like satisfaction of triangle inequality, that is to say, strictly speaking $d$ is not necessarily a distance. But we need to do an additional assumption to deal with the case of ties among several distances from the same site. If this is the case, in order to break ties we suppose that there are preferences on the centers in such a way that sites undoubtedly will choose one of the centers before the others. In practice, ties can be broken by slightly perturbing the tied distances. Summarizing, we will also assume $d_{ia} \neq d_{ib} \ \forall i, a, b \in N : \ a \neq b$. Associated with each customer $i \in N$ is the probability of having demand $0 \leq q_i < 1$. The events of demand occurrence are assumed to be independent.

To describe a solution to the PpCP we will need to identify the set of $p$ sites where facilities are open, and the assignment to one of those facilities of each of the potential customers, since at the moment of making the decision we do not know which customers will place a demand and which will not. In what follows, we will distinguish between the *assignment* cost of a customer and its *service cost*. The assignment cost corresponds to the distance between the customer and the facility it is assigned to a priori, whilst the service cost takes this same value but only in the demand scenarios where the customer does have demand.

In case of tie between a client and several plants, this will be assigned to the plant with the largest index. In case of ties between two clients and their plants we consider the largest distance the one assigned to the client with the greatest index.

**Lemma 1.** *For any set of probabilities $(q_1, \ldots, q_n)$ with $0 \leqslant q_i \leqslant 1, i \in N$, any feasible solution of the PpCP can be associated with a $\pi$ vector satisfying:*

1. $|\{j \in N : \pi_{ij} \neq 0\}| = 1 \quad \forall i \in N$
2. *If $d_{(1)} \leqslant d_{(2)} \leqslant \cdots \leqslant d_{(n)}$ is a non-decreasing sequence of distances between each customer and the open facility it is assigned to, and $(1), \ldots, (n)$ is the corresponding sequence of customers,*

$$\sum_{j=1}^{n} \pi_{(i)j} = q_{(i)} \prod_{t=i+1}^{n} (1 - q_{(t)}).$$

3.

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij} = 1 - \prod_{i=1}^{n} (1 - q_i) \leqslant 1, \tag{1}$$

*that allows to compute the expected maximum service cost as*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij} d_{ij}.$$

**Theorem 1.** *The optimal value of the objective function above is achieved in a solution where every site is covered by its closest plant.*

**Corollary 1.** *The closest assignment constraints can be used as valid inequalities for any formulation of the considered problem even if they are not needed to formulate it.*

Suppose that we want to minimize the expected maximum service cost, taking into account only the $K \leqslant p$ largest allocated distances. This problem is called $K$ probabiliestic $p$-center problem. In this case, closest assignment constraints must be added to formulation. Otherwise we can't assure that the sites are assigned to their closest located facility. In the following section we develop different formulations for PpCP and K-PpCP.

## 3 A first Formulation. Four-index Formulation

In this section we present a first formulation of PpCP that uses two families of binary variables and a group of continuous variables. For $i, j \in N$, we define binary variables

$$y_j = \begin{cases} 1, \text{ if a plant is opened at site } j, \\ 0, \text{ otherwise.} \end{cases}$$

and continuous variables
$\pi_{ij}$= probability that $d_{ij}$ is the largest service distance.

Note that in case of distance ties $\pi_{ij}$ variables are defined in the sense of Lemma 1. Finally, for all $i, j, k, \ell \in N$ such that $d_{ij} > d_{k\ell}$ or if $d_{ij} = d_{k\ell}$ for all $i > k$, we also define the binary variables

$$x_{ijk\ell} = \begin{cases} 1, \text{ if } i \text{ is allocated to } j, \ k \text{ to } \ell \text{ and } d_{ij} \text{ is the first} \\ \quad \text{candidate service distance larger than } d_{k\ell}; \\ \quad (d_{ij} > d_{k\ell}), \\ 0, \text{ otherwise.} \end{cases}$$

$$\text{(F1)} \quad \min \sum_{i=1}^{n}\sum_{j=1}^{n} \pi_{ij} d_{ij}$$

$$\text{s.t.} \sum_{j=1}^{n} y_j = p, \tag{2}$$

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{ijk\ell} \leqslant y_j, \forall i,j \in N \tag{3}$$

$$\sum_{k=1}^{n}\sum_{l=1}^{n} x_{k\ell ij} \leqslant y_j, \forall i,j \in N \tag{4}$$

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{ijk\ell} \leqslant 1, \forall i \in N \tag{5}$$

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{k\ell ij} \leqslant 1, \forall i \in N \tag{6}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{ijk\ell} = n-1, \tag{7}$$

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n}\sum_{\substack{j'=1 \\ j'\neq j}}^{n} x_{ij'k\ell} + \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{k\ell ij} \leqslant 1, \forall i,j \in N \tag{8}$$

$$\pi_{k\ell} \geqslant \frac{1-q_i}{q_i} q_k \pi_{ij} - 1 + x_{ijk\ell}, \forall i,j,k,\ell \in N \tag{9}$$

$$\pi_{ij} \geqslant q_i \left( \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{ijk\ell} - \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_{k\ell ij} \right) \forall i,j \in N \tag{10}$$

$$y_j, x_{ijk\ell} \in \{0,1\}, \forall i,j,k,\ell \in N \tag{11}$$

$$\pi_{ij} \in [0,1], \forall i,j \in N. \tag{12}$$

Constraint (2) ensures that $p$ facilities are opened, and constraints (3) and (4) force that all assignments of customers are made to open facilities. The sorting of the used service distances is made through constraints (5)-(8), taking advantage of the variable definition (recall that $x_{ijk\ell}$ is not defined, or is fixed to zero, if $d_{ij} \lneqq d_{k\ell}$). In particular, constraints (5) and (6) ensure that the distance to cover site $i$ is at most once inmediately greater/smaller than another distance from a site and its plant. Constraints (8) together with (5) and (6) ensure that any site $i$ is cover by at most one plant. Constraints (9)-(10) are used to guarantee that $\pi$ and $x$ variables take consistent values. Finally, the last sets of constraints set the domains of the variables.

As shown in Corolary 1, Closest Assignment Constraints (CAC) are valid. In this work, we have adapted the CAC set presented in [5]:

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n}\sum_{a=1;d_{ia}>d_{ij}}^{n} x_{iak\ell} + y_j \leqslant 1, \quad \forall i,j \in N, \tag{13}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{a=1;d_{ka}>d_{k\ell}}^{n} x_{ijka} + y_\ell \leqslant 1, \quad \forall k,\ell \in N. \tag{14}$$

## 4 Four index formulation with K largest distances

We add variables:

$$z_{kl} = \begin{cases} 1, \text{ if } k \text{ is allocated to } l \text{ and the distance } d_{kl} \text{ is} \\ \quad \text{among the } n-K \text{ smallest distances,} \\ 0, \text{ otherwise.} \end{cases}$$

Taking only into account K largest distance, we obtain the following formulation,

$$(\mathrm{F1}^K) \quad \min \sum_{i=1}^{n}\sum_{j=1}^{n} \pi_{ij} d_{ij}$$

$$\text{s.t. constraints (2)-(7),(10)-(14),}$$

$$\pi_{k\ell} \geqslant \frac{1-q_i}{q_i} q_k \pi_{ij} - 1 + x_{ijk\ell} - z_{k\ell}, \forall i,j,k,\ell \in N, \tag{15}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k'=1,\ell'=1 \\ d_{k'\ell'}\geqslant d_{k\ell}}}^{n} x_{ijk'\ell'} \geq K z_{k\ell},\ \forall k,\ell \in N. \tag{16}$$

$$z_{k\ell} \in \{0,1\},\ \forall k,\ell \in N. \tag{17}$$

Constraints (16) are used to guarantee that the $z$ variables take consistent values. In addition, (16) is similar to (9) where term $-z_{kl}$ has been include to distinguish wether the distance $d_{kl}$ is among $K-$largest distances.

## 5 Variable neighborhood search for PpCP

Variable Neighborhood search is a metaheuristic to solve combinatorial problems proposed by [6] for the $p$-median problem. It is a very well-known technique often used to solve discrete facility location problems and it usually provides high quality solutions. Since $p$-median problem is a particular case of DOMP, [7] adapted VNS for solving the DOMP. We will adopt this heuristic to solve PpCP.

The basic idea of VNS is to implement a systematic interchange of neighborhoods within a local search algorithm. The algorithm remains in the same solution until a better solution is found and then, the solution moves there.

In order to test the quality of VNS we considered a group of random generated examples with different combinations of $n$, $p$ and $K$. To compare the solutions given by VNS with the optimal ones, the solutions of $F1^K$ have been used. The gap between the optimal solution and that one obtained by VNS is computed according to:

$$gap = 100 \cdot \frac{z_{VNS} - z_{opt}}{z_{opt}},$$

where $z_{VNS}$ is the solution obtained by VNS and $z_{opt}$ the optimal objective function. We also compare running times of VNS heuristic and four index formulation with K largest distances. VNS method and four index formulation were implemented using Mosel programming language and compiled by Xpress 7.7. Instances were run on a Intel(R) Core(TM) i7-4790K CPU 32 GB RAM.

For each value of $(n, p, K)$, five instances are created in the following way. We generate $n$ random points in the plane and we obtain their distances using $l1$-norm. After that, demand probability of each customer is generated randomly. Next table reports average gap, VNS running time and $F1^K$ running time of each group of instances. As we can see, VNS heuristic provides good results in very small times.

| n | p | K | Gap | VNS-time | $F1^K$-time | n | p | K | Gap | VNS-time | $F1^K$-time |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 2 | 2 | 0.00 | 0.00 | 0.88 | 15 | 10 | 4 | 0.00 | 0.03 | >7200 |
| 10 | 3 | 3 | 0.00 | 0.00 | 60.62 | 20 | 3 | 5 | 0.00 | 0.01 | >7200 |
| 10 | 5 | 3 | 0.00 | 0.01 | 40.23 | 20 | 8 | 5 | 0.00 | 0.06 | >7200 |
| 13 | 3 | 4 | 1.41 | 0.00 | 1678.14 | 20 | 10 | 5 | 0.00 | 0.07 | >7200 |
| 13 | 5 | 4 | 0.00 | 0.01 | 1945.77 | 25 | 3 | 6 | 0.16 | 0.02 | >7200 |
| 13 | 8 | 4 | 0.25 | 0.02 | 785.31 | 25 | 8 | 6 | 0.33 | 0.11 | >7200 |
| 15 | 3 | 4 | 0.59 | 0.01 | >7200 | 25 | 10 | 6 | 1.28 | 0.14 | >7200 |
| 15 | 8 | 4 | 0.00 | 0.02 | >7200 | | | | | | |

Comparisons between VNS and $F1^K$ results.

## References


1. Daskin, M.S.: Network and Discrete Location: Models, Algorithms, and Applications, Wiley, New York (1995)
2. Drezner, Z., Hamacher, H.: Facility Location: Applications and Theory, Springer Verlag, New York (2002)
3. Albareda-Sambola, M., Díaz, J.A., Fernández, E.: Lagrangean duals and exact solution to the capacitated $p$-center problem, European Journal Operational Research, 201, 71-81 (2010)
4. Kariv, O., Hakimi, S.L.: An algorithmic approach to network location problems I: The $p$-Centers, SIAM Journal on Applied Mathematics, 37 (3), 513-538, (1979)
5. Wagner, J.L., Falkson, L.M.: The optimal nodal location of public facilities with price-sensitive demand, Geographical Analysis, 7, 69-83 (1975)
6. Mladenović, N., Hansen,P.: Variable Neighbourhood Search, Computers and Operation Research, 24, 1097-1100 (1997)
7. Domínguez-Marín, P., Nicke, S., Hansen P., Mladenović, N: Heuristic procedures for solving the Discrete Ordered Median Problem, Annals of Operations Research, 136, 145-173, (2005)

# Lie algebras with a set grading II

Antonio J. Calderón Martín*

Department of Mathematics
Faculty of Sciences, University of Cádiz
Campus de Puerto Real, 11510, Puerto Real, Cádiz, Spain.
ajesus.calderon@uca.es

**Abstract.** We review our results on Lie algebras with a set grading and apply them to study Lie algebras admitting a multiplicative basis. We show that if $\mathfrak{g}$ is a Lie algebra admitting a multiplicative basis $\mathcal{B}$ then $\mathfrak{g}$ decomposes as the direct sum of well-described ideals. Under mild conditions, it is also shown that the above decomposition is actually by means of the family of its minimal ideals.

**Keywords.** Graded Lie algebra, multiplicative basis, infinite dimensional Lie algebra, structure theory.

## 1 Introduction

The interest on group gradings on Lie algebras has been remarkable in the last years, motivated in part by their application in physics, geometry and topology where they appear as the natural framework for an algebraic model [1, 2, 7–10, 12–16]. However gradings by means of an arbitrary set, not necessarily a group, have been considered in the literature just in a slightly way. A complete review of the state of the art can be found in the recent monograph [11].

In the paper [6] we study Lie algebras of arbitrary dimension and over an arbitrary base field $\mathbb{K}$ graded by means of an arbitrary set $I$, by focussing on its structure. We begin by reviewing the results in [6]. Finally we will apply these results to the study of the structure of Lie algebras admitting a multiplicative basis.

## 2 Review on set graded Lie algebras

**Definition 1.** *Let $\mathfrak{g}$ be a Lie algebra and $I$ an arbitrary (non-empty) set. It is said that $\mathfrak{g}$ has a* set grading*, by means of $I$, if $\mathfrak{g} = \bigoplus_{i\in I} \mathfrak{g}_i$ where any $\mathfrak{g}_i$ is a linear subspace satisfying that for any $j \in I$ either $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ or $0 \neq [\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_k$ for some (unique) $k \in I$.*

*We call the support of the grading to the set $\mathfrak{S} := \{i \in I : \mathfrak{g}_i \neq 0\}$.*

* Supported by the PCI of the UCA 'Teoría de Lie y Teoría de Espacios de Banach', by the PAI with project numbers FQM298, FQM7156 and by the project of the Spanish Ministerio de Educación y Ciencia MTM2010-15223.

The regularity conditions will be understood in graded sense, that is, a *subalgebra* $\mathfrak{a}$ of a Lie algebra with a set grading $\mathfrak{g}$ is a linear subspace satisfying $[\mathfrak{a},\mathfrak{a}]\subset\mathfrak{a}$ and such that splits as $\mathfrak{a}=\bigoplus_{i\in I}\mathfrak{a}_i$ with any $\mathfrak{a}_i=\mathfrak{a}\cap\mathfrak{g}_i$. A subalgebra $\mathfrak{i}$ of $\mathfrak{g}$ is an *ideal* if $[\mathfrak{i},\mathfrak{g}]\subset\mathfrak{i}$. A Lie algebra with a set grading $\mathfrak{g}$ will be called *simple* if its product is nonzero and its only ideals are $\{0\}$ and $\mathfrak{g}$.

Let $\mathfrak{g}$ be an arbitrary set-graded Lie algebra with the set $\mathfrak{S}$ as support of the grading. For each $i\in\mathfrak{S}$, a new variable $\overline{i}\notin\mathfrak{S}$ is introduced and we denote by $\overline{\mathfrak{S}}:=\{\overline{i}:i\in\mathfrak{S}\}$ the set consisting of all these new symbols. We will also denote by $\mathcal{P}(A)$ the power set of a given set $A$.

Next, we consider the following operation,
$\star:(\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\times(\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\to\mathcal{P}(\mathfrak{S})$, given by

- For $i,j\in\mathfrak{S}$,
$$i\star j=\begin{cases}\emptyset, & \text{if } [\mathfrak{g}_i,\mathfrak{g}_j]=0;\\ \{k\}, & \text{if } 0\neq[\mathfrak{g}_i,\mathfrak{g}_j]\subset\mathfrak{g}_k.\end{cases}$$
- For $i\in\mathfrak{S}$ and $\overline{j}\in\overline{\mathfrak{S}}$,
$$i\star\overline{j}=\overline{j}\star i=\{k\in\mathfrak{S}:0\neq[\mathfrak{g}_k,\mathfrak{g}_j]\subset\mathfrak{g}_i\}.$$
- For $\overline{i},\overline{j}\in\overline{\mathfrak{S}}$,
$$\overline{i}\star\overline{j}=\emptyset.$$

From now on, given any $\overline{i}\in\overline{\mathfrak{S}}$ we will denote $\overline{(\overline{i})}:=i$. Given also any subset $\mathfrak{U}$ of $\mathfrak{S}\cup\overline{\mathfrak{S}}$, we write by $\overline{\mathfrak{U}}:=\{\overline{i}:i\in\mathfrak{U}\}$ if $\mathfrak{U}\neq\emptyset$ and $\overline{\emptyset}:=\emptyset$.

In this moment we have to note that sometimes it is interesting to distinguish one element $\mathfrak{o}$ in the support of the grading, because the homogeneous space $\mathfrak{g}_\mathfrak{o}$ has, in a sense, a special behavior to the remaining elements in the set of homogeneous spaces $\mathfrak{g}_i$,$i\in\mathfrak{S}$. From here, we are going to feel free in our study to distinguish one special element $\mathfrak{o}$ in the support of the grading. Hence, let us now fix an element $\mathfrak{o}$ such that either $\mathfrak{o}\in\mathfrak{S}$ satisfying the property $\mathfrak{o}\star i\neq\{\mathfrak{o}\}$ for any $i\in\mathfrak{S}\setminus\{\mathfrak{o}\}$, or $\mathfrak{o}=\emptyset$. Note that the possibility $\mathfrak{o}=\emptyset$ holds for the case in which it is not wished to distinguish any element in $\mathfrak{S}$.

Finally, we need to introduce the following mapping:

$$\phi:\mathcal{P}((\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\setminus\{\mathfrak{o},\overline{\mathfrak{o}}\})\times(\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\to\mathcal{P}((\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\setminus\{\mathfrak{o},\overline{\mathfrak{o}}\}),$$

as

- $\phi(\emptyset,\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})=\emptyset$,
- For any $\emptyset\neq\mathfrak{A}\in\mathcal{P}((\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}})\setminus\{\mathfrak{o},\overline{\mathfrak{o}}\})$ and $a\in\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}$,
$$\phi(\mathfrak{A},a)=((\bigcup_{x\in\mathfrak{A}}(x\star a))\setminus\{\mathfrak{o}\})\cup\overline{((\bigcup_{x\in\mathfrak{A}}(x\star a))\setminus\{\mathfrak{o}\})}.$$

Note that for any $\mathfrak{A} \in \mathcal{P}((\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}) \setminus \{\mathfrak{o}, \overline{\mathfrak{o}}\})$ and $a \in \mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}$ we get that

$$\phi(\mathfrak{A}, a) = \overline{\phi(\mathfrak{A}, a)} \tag{1}$$

and

$$\phi(\mathfrak{A}, a) \cap \mathfrak{S} = (\bigcup_{x \in \mathfrak{A}} (x \star a)) \setminus \{\mathfrak{o}\}.$$

Also observe that for any $i \in \mathfrak{S}$ and $a \in \mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}$ we have that $i \in x \star a$ for some $x \in \mathfrak{S}$ if and only if $x \in i \star \overline{a}$; while $i \in k \star a$ for some $k \in \overline{\mathfrak{S}}$ if and only if $\overline{k} \in \overline{i} \star a$. These facts together with Equation (1) imply that for any $\mathfrak{A} \in \mathcal{P}((\mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}) \setminus \{\mathfrak{o}, \overline{\mathfrak{o}}\})$ such that $\mathfrak{A} = \overline{\mathfrak{A}}$ and $a \in \mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}$ we have

$$i \in \phi(\mathfrak{A}, a) \cap \mathfrak{S} \text{ if and only if}$$

$$i \in \mathfrak{S} \text{ and either } \phi(\{i\}, \overline{a}) \cap \mathfrak{A} \cap \mathfrak{S} \neq \emptyset \text{ or } \phi(\{\overline{i}\}, a) \cap \mathfrak{A} \cap \mathfrak{S} \neq \emptyset.$$

**Definition 2.** *Let $i$ and $j$ be two elements in $\mathfrak{S} \setminus \{\mathfrak{o}\}$. We say that $i$ is* connected *to $j$ if there exists a family*

$$\{a_1, a_2, ..., a_{n-1}, a_n\} \subset \mathfrak{S}\dot{\cup}\overline{\mathfrak{S}}$$

*satisfying the following conditions:*

*If $n = 1$.*

1. $a_1 = i = j$.

*If $n \geq 2$.*

1. $a_1 \in \{i, \overline{i}\}$.

2. $\phi(\{a_1\}, a_2) \neq \emptyset$,
   $\phi(\phi(\{a_1\}, a_2), a_3) \neq \emptyset$,
   $\phi(\phi(\phi(\{a_1\}, a_2), a_3), a_4) \neq \emptyset$,
   $\cdots\cdots\cdots$
   $\phi(\phi(\cdots(\phi(\{a_1\}, a_2), \cdots), a_{n-2}), a_{n-1}) \neq \emptyset$.

3. $j \in \phi(\phi(\cdots(\phi(\{a_1\}, a_2), \cdots), a_{n-1}), a_n)$.

*The family $\{a_1, a_2, ..., a_{n-1}, a_n\}$ is called a connection from $i$ to $j$.*

The next result is of straightforward verification.

**Proposition 1.** *The relation $\sim$ in $\mathfrak{S} \setminus \{\mathfrak{o}\}$, defined by $i \sim j$ if and only if $i$ is connected to $j$, is an equivalence relation.*

By the above Proposition we can consider the quotient set $(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim=\{[i]: i\in \mathfrak{S}\setminus\{\mathfrak{o}\}\}$, becoming $[i]$ the set of elements in $\mathfrak{S}\setminus\{\mathfrak{o}\}$ which are connected to $i$.

Our next goal in this section is to associate an ideal $\mathfrak{g}_{[i]}$ of $\mathfrak{g}$ to any $[i]$. Fix $i\in\mathfrak{S}\setminus\{\mathfrak{o}\}$, we start by defining the set $\mathfrak{g}_{\mathfrak{o},[i]}\subset\mathfrak{g}_{\mathfrak{o}}$ as follows $\mathfrak{g}_{\mathfrak{o},[i]}:=(\sum\limits_{j,k\in[i]}[\mathfrak{g}_j,\mathfrak{g}_k])\cap\mathfrak{g}_{\mathfrak{o}}$, where $\mathfrak{g}_{\mathfrak{o}}:=\{0\}$ whence $\mathfrak{o}=\emptyset$.

Next, we define $\mathfrak{v}_{[i]}:=\bigoplus_{j\in[i]}\mathfrak{g}_j$. Finally, we denote by $\mathfrak{g}_{[i]}$ the direct sum of the two subspaces above, that is, $\mathfrak{g}_{[i]}:=\mathfrak{g}_{\mathfrak{o},[i]}\oplus\mathfrak{v}_{[i]}$.

We can verify the next result.

**Proposition 2.** *For any $i\in\mathfrak{S}\setminus\{\mathfrak{o}\}$, the graded linear subspace $\mathfrak{g}_{[i]}$ is an ideal of $\mathfrak{g}$.*

**Corollary 1.** *If $\mathfrak{g}$ is simple, then there exists a connection between any two elements of $\mathfrak{S}\setminus\{\mathfrak{o}\}$.*

**Lemma 1.** *If $[i]\neq[j]$ for some $i,j\in\mathfrak{S}\setminus\{\mathfrak{o}\}$ then $[\mathfrak{g}_{[i]},\mathfrak{g}_{[j]}]=0$.*

**Theorem 1.** *A set-graded Lie algebra $\mathfrak{g}$ decomposes as $\mathfrak{g}=\mathfrak{u}\oplus(\sum\limits_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{g}_{[i]})$, where $\mathfrak{u}$ is a linear complement of $\sum\limits_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{g}_{\mathfrak{o},[i]}$ in $\mathfrak{g}_{\mathfrak{o}}$ and any $\mathfrak{g}_{[i]}$ is one of the ideals given in Proposition 2. Furthermore $[\mathfrak{g}_{[i]},\mathfrak{g}_{[j]}]=0$ whenever $[i]\neq[j]$.*

*Proof.* Since we can write $\mathfrak{g}=\mathfrak{g}_{\mathfrak{o}}\oplus(\bigoplus\limits_{i\in\mathfrak{S}\setminus\{\mathfrak{o}\}}\mathfrak{g}_i)$ and

$$\mathfrak{g}_{\mathfrak{o}}=\mathfrak{u}\oplus(\sum_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{g}_{\mathfrak{o},[i]}),\qquad\bigoplus_{i\in\mathfrak{S}\setminus\{\mathfrak{o}\}}\mathfrak{g}_i=\bigoplus_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{v}_{[i]}$$

we clearly have $\mathfrak{g}=\mathfrak{u}\oplus(\sum\limits_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{g}_{[i]})$ being each $\mathfrak{g}_{[i]}$ an ideal of $\mathfrak{g}$ satisfying $[\mathfrak{g}_{[i]},\mathfrak{g}_{[j]}]=0$ when $[i]\neq[j]$ by Proposition 2 and Lemma 1.

In case it is not distinguished any element $\mathfrak{o}$ in the support of the grading, that is $\mathfrak{o}=\emptyset$, we have as an immediate consequence of Theorem 1 the next result.

**Corollary 2.** *If $\mathfrak{o}=\emptyset$, then $\mathfrak{g}=\bigoplus\limits_{[i]\in\mathfrak{S}/\sim}\mathfrak{g}_{[i]}$, where any $\mathfrak{g}_{[i]}$ is one of the ideals given in Proposition 2.*

We recall that the *center* of $\mathfrak{g}$ is the set $\mathcal{C}(\mathfrak{g})=\{x\in\mathfrak{g}:[x,\mathfrak{g}]=0\}$, and that it is said that $\mathfrak{g}_{\mathfrak{o}}$ is *tight* whence $\mathfrak{g}_{\mathfrak{o}}=\{0\}$ or $\mathfrak{g}_{\mathfrak{o}}=\sum\limits_{i,j\in\mathfrak{S}\setminus\{\mathfrak{o}\},i\star j=\{\mathfrak{o}\}}[\mathfrak{g}_i,\mathfrak{g}_j]$.

**Corollary 3.** *Suppose $\mathfrak{g}$ is centerless and $\mathfrak{g}_0$ is tight, then the set-graded Lie algebra $\mathfrak{g}$ decomposes as the direct sum of the ideals given in Proposition 2, $\mathfrak{g}=\bigoplus_{[i]\in(\mathfrak{S}\setminus\{\mathfrak{o}\})/\sim}\mathfrak{g}_{[i]}$.*

Now we ask ourselves if any of the components in the decompositions given in Theorem 1, Corollary 2 and Corollary 3 is simple. Under mild conditions we give an affirmative answer.

**Definition 3.** *We say that* $\mathfrak{g}$ *is of* maximal length *if* $\dim \mathfrak{g}_i = 1$ *for any* $i \in \mathfrak{S} \setminus \{\mathfrak{o}\}$.

**Definition 4.** *We say that* $\mathfrak{g}$ *is* $\mathfrak{S}$-multiplicative *if given* $i, j \in \mathfrak{S}$ *such that* $i \in j \star a$ *for some* $a \in \mathfrak{S} \dot{\cup} \overline{\mathfrak{S}}$ *then* $\mathfrak{g}_i \subset [\mathfrak{g}_j, \mathfrak{g}_a + \mathfrak{g}_{\overline{a}}]$.

The proof of the next results can be found in [6, Section 3].

**Theorem 2.** *Let* $\mathfrak{g}$ *be a centerless* $\mathfrak{S}$*-multiplicative Lie algebra with a set grading of maximal length and with* $\mathfrak{g}_\mathfrak{o}$ *tight. Then* $\mathfrak{g}$ *is simple if and only if it has all of the elements in* $\mathfrak{S} \setminus \{\mathfrak{o}\}$ *connected.*

**Theorem 3.** *Let* $\mathfrak{g}$ *be a centerless* $\mathfrak{S}$*-multiplicative Lie algebra with a set grading of maximal length and with* $\mathfrak{g}_\mathfrak{o}$ *tight. Then* $\mathfrak{g}$ *is the direct sum of the family of its minimal ideals, each one being a simple Lie algebra with a set grading having all of the elements different to* $\mathfrak{o}$ *in its support connected.*

## 3 Lie algebras admitting a multiplicative basis

**Definition 5.** *A basis* $\mathcal{B} = \{e_i\}_{i \in I}$ *of a Lie algebra* $(\mathfrak{g}, [\cdot, \cdot])$ *is said to be* multiplicative *if for any* $i, j \in I$ *we have either* $[e_i, e_j] = 0$ *or* $0 \neq [e_i, e_j] \in \mathbb{F} e_k$ *for some (unique)* $k \in I$.

*Remark 1.* The definition of multiplicative basis given in Definition 5 is a little bit more general than the usual one in the literature ([3, 4]). In fact, in these references, a basis $\mathcal{B} = \{e_i\}_{i \in I}$ is called multiplicative if for any $i, j \in I$ we have either $[e_i, e_j] = 0$ or $0 \neq [e_i, e_j] = e_k$ for some $k \in I$.

**Definition 6.** *Let* $\mathfrak{g}$ *be a Lie algebra with a multiplicative basis* $\mathcal{B}$*. It is said that a subalgebra* $\mathfrak{C}$ *of* $\mathfrak{g}$ *admits a multiplicative basis* $\mathcal{B}_\mathfrak{C}$ inherited *from* $\mathcal{B}$ *if* $\mathcal{B}_\mathfrak{C}$ *is a multiplicative basis of* $\mathfrak{C}$ *satisfying* $\mathcal{B}_\mathfrak{C} \subset \mathcal{B}$.

Consider now a Lie algebra $\mathfrak{g}$ with a multiplicative basis $\mathcal{B} = \{e_i\}_{i \in I}$. Observe that we can write

$$\mathfrak{g} = \bigoplus_{i \in I} \mathbb{K} e_i$$

being the above decomposition a set grading of $\mathfrak{g}$ by means of the set $I$. We would have $\mathfrak{g}_i = \mathbb{K} e_i$ for any $i \in I$. From here, the results in Section 2 apply and we can assert the following results.

**Theorem 4.** *Let* $\mathfrak{g}$ *be an arbitrary Lie algebra with a multiplicative basis. Then one has*

$$\mathfrak{g} = \bigoplus_{[i] \in I/\sim} \mathfrak{g}_{[i]},$$

*being any* $\mathfrak{g}_{[i]}$ *an ideal of* $\mathfrak{g}$ *admitting a multiplicative basis inherited from the one of* $\mathfrak{g}$ *and satisfying* $[\mathfrak{g}_{[i]}, \mathfrak{g}_{[j]}] = 0$ *whenever* $[i] \neq [j]$.

**Definition 7.** *We say that a Lie algebra $\mathfrak{g}$ admits a* $\star$-multiplicative basis $\mathcal{B} = \{e_i\}_{i\in I}$ *if it is multiplicative and given either $i, j \in I$ such that $j \in i \star a$ for some $a \in I \dot{\cup} \overline{I}$ or $j \in k \star i$ for some $k \in I$ then $e_j \in [e_i, \mathfrak{g}]$.*

Examples of Lie algebras admitting $\star$-multiplicative bases are the semisimple finite-dimensional Lie algebras over algebraically closed fields of characteristic 0, the semisimple separable $L^*$-algebras, the semisimple locally finite split Lie algebras and the split Lie algebras considered in [5, §3].

**Definition 8.** *A Lie algebra $\mathfrak{g}$ admitting a multiplicative basis $\mathcal{B}$ is called* minimal *if its only nonzero ideal admitting a multiplicative basis inherited from $\mathcal{B}$ is $\mathfrak{g}$.*

**Theorem 5.** *Let $\mathfrak{g}$ be a Lie algebra admitting a $\star$-multiplicative basis $\mathcal{B} = \{e_i\}_{i\in I}$. Then $\mathfrak{g} = \bigoplus_k \mathfrak{I}_k$ is the direct sum of the family of its minimal ideals, each one admitting a $*$-multiplicative basis inherited from $\mathcal{B}$.*

## References


1. Bahturin, Y.A., Shestakov, I.P. and Zaicev, M.: Gradings on simple Jordan and Lie algebras, J. Algebra 283 (2005), no. 2, 849-868
2. Bahturin, Y.A. and Tvalavadze, M.: Group gradings on $\mathfrak{g}_2$. Comm. Algebra 37 (2009), no. 3, 885-893.
3. Bautista, R., Gabriel, P., Roiter, A.V. and Salmeron, L.: Representation-finite algebras and multiplicative basis. Invent. math. 81 (1985), 217–285.
4. Bovdi, V.: On a filtered multiplicative bases of group algebras. II. Algebr. Represent. Theory 6 (2003), no. 3, 353–368.
5. Calderón, A.J.: On split Lie algebras with a symmetric root system. Proc. Indian Acad. Sci. (Math. Sci.) 118 (2008), no. 3, 351–356.
6. Calderón, A.J.: Lie algebras with a set grading. Linear Algebra and its applications. 452(2014), 7–20.
7. Calderón, A.J. and Sánchez, J.M.: On the structure of graded Lie supealgebras. Modern Phys. Letters A. 27(25), 1250142, (2012), 18 pp.
8. Draper, C.: Models on the Lie algebra $\mathfrak{f}_4$. Linear Algebra Appl. 428 (2008), no. 11-12, 2813–2839.
9. Draper, C. and Martín, C.: Gradings on $\mathfrak{g}_2$. Linear Algebra Appl. 418(1) (2006), 85–111.
10. Draper, C. and Martín, C.: Gradings on the Albert algebra and on $\mathfrak{f}_4$. Rev. Mat. Iberoam. 25 (2009), no. 3, 841-908.
11. Elduque, A. and Kochetov, M.: Gradings on simple Lie algebras. Mathematical Surveys and Monographs 189, American Mathematical Society, 2013.
12. Elduque, A.: Fine gradings on simple classical Lie algebras. J. Algebra, 324 (2010), no. 12, 3532–3571.
13. Havlíček, M., Patera, J. and Pelantová, E.: On Lie gradings II. Linear Algebra Appl. 277 (1998), 97–125.
14. Havlíček, M., Patera, J. and Pelantová, E.: On Lie gradings III. Linear Algebra Appl. 314 (2000), 1–47.
15. Kochetov, M.: Gradings on finite-dimensional simple Lie algebras. Acta Appl. Math. 108 (2009), no. 1, 101–127.
16. Patera J. and Zassenhaus, H.: On Lie gradings I. Linear Algebra Appl. 112, (1989), 87–159.

# On strongly split regular Hom-Poisson algebras

María J. Aragón Periñán and Antonio J. Calderón Martín*

Department of Mathematics
Faculty of Sciences, University of Cádiz
Campus de Puerto Real, 11510, Puerto Real, Cádiz, Spain.
mariajesus.aragonperin@alum.uca.es; ajesus.calderon@uca.es

**Abstract.** We review some of our results on split regular Hom-Poisson algebras obtained in [2]. Later we introduce the class of strongly split regular Hom-Poisson algebras and apply the previous results on split regular Hom-Poisson algebras to prove that a Lie-centerless arbitrary strongly split regular Hom-Poisson algebra $\mathfrak{P}$ is the direct sum of the family of its split-simple split-ideals.

**Keywords:** Hom-associative algebra, Hom-Lie algebra, Poisson algebra, root, root space, strongly split, structure theory.

## 1 Review on split regular Hom-Poisson algebras

**Definition 1.** *A* Hom-Lie algebra $\mathfrak{P}$ *is a vector space over a base field* $\mathbb{K}$ *endowed with a bilinear product* $\{\cdot,\cdot\} : \mathfrak{P} \times \mathfrak{P} \longrightarrow \mathfrak{P}$ *and with a linear map* $\phi : \mathfrak{P} \longrightarrow \mathfrak{P}$ *such that*

1. $\{x, y\} = -\{y, x\}$,
2. $\{\{x, y\}, \phi(z)\} + \{\{y, z\}, \phi(x)\} + \{\{z, x\}, \phi(y)\} = 0$, *(Hom-Jacobi identity),*

*for any* $x, y, z \in \mathfrak{P}$.

**Definition 2.** *A* Hom-Poisson algebra *is a Hom-Lie algebra* $(\mathfrak{P}, \{\cdot,\cdot\}, \phi)$*, endowed with a Hom-associative product, that is, a bilinear product denoted by yuxtaposition such that*

$$(xy)\phi(z) = \phi(x)(yz)$$

*for any* $x, y, z \in \mathfrak{P}$*, and such that the following* Hom-Leibniz identity

$$\{xy, \phi(z)\} = \{x, z\}\phi(y) + \phi(x)\{y, z\}$$

*holds for any* $x, y, z \in \mathfrak{P}$.

*When* $\phi$ *furthermore is a Poisson automorphism, that is a linear bijection such that* $\phi(\{x, y\}) = \{\phi(x), \phi(y)\}$ *and* $\phi(xy) = \phi(x)\phi(y)$ *for any* $x, y \in \mathfrak{P}$*, it is said that* $\mathfrak{P}$ *is a* regular *Hom-Poisson algebra.*

* Supported by the PCI of the UCA 'Teoría de Lie y Teoría de Espacios de Banach', by the PAI with project numbers FQM298, FQM7156 and by the project of the Spanish Ministerio de Educación y Ciencia MTM2010-15223.

A *subalgebra* $A$ of $\mathfrak{P}$ is a linear subspace such that $\{A, A\} + AA \subset A$ and $\phi(A) = A$. A linear subspace $I$ of $\mathfrak{P}$ is called an *ideal* if $\{I, \mathfrak{P}\} + I\mathfrak{P} + \mathfrak{P}I \subset I$ and $\phi(I) = I$. We refer to [3] for a first approach to Hom-Poisson algebras.

In the paper [2] we introduce the class of split Hom-Poisson algebras $\mathfrak{P}$ formed for those Hom-Poisson algebras satisfying that their underlying Hom-Lie algebras are split. So we recall that given a Hom-Lie algebra $(\mathfrak{P}, \{\cdot,\cdot\}, \phi)$ and by denoting by $H$ a maximal abelian subalgebra of $\mathfrak{P}$. For a linear functional

$$\alpha : H \longrightarrow \mathbb{K},$$

we define the *root space* of $\mathfrak{P}$ (respect to $H$) associated to $\alpha$ as the subspace

$$\mathfrak{P}_\alpha = \{v_\alpha \in \mathfrak{P} : \{h, v_\alpha\} = \alpha(h)\phi(v_\alpha) \text{ for any } h \in H\}.$$

The elements $\alpha : H \longrightarrow \mathbb{K}$ satisfying $\mathfrak{P}_\alpha \neq 0$ are called *roots* of $\mathfrak{P}$ with respect to $H$ and we denote $\Lambda := \{\alpha \in (H)^* \setminus \{0\} : \mathfrak{P}_\alpha \neq 0\}$. It is said that $\mathfrak{P}$ is a *split Hom-Lie algebra*, with respect to $H$, if

$$\mathfrak{P} = H \oplus (\bigoplus_{\alpha\in\Lambda} \mathfrak{P}_\alpha).$$

It is also said that $\Lambda$ is the *roots system* of $\mathfrak{P}$.

For an easier notation, the mappings $\phi|_H, \phi|_H^{-1} : H \to H$ will be denoted by $\phi$ and $\phi^{-1}$ respectively.

We recall some properties of split regular Hom-Lie algebras that can be found in [1, Lemma 1.3, Lemma 1.4].

**Lemma 1.** *Let $(\mathfrak{P}, \{\cdot,\cdot\}, \phi)$ be a split regular Hom-Lie algebra. Then for any $\alpha, \beta \in \Lambda \cup \{0\}$ the following assertions hold.*

1. $\phi(\mathfrak{P}_\alpha) \subset \mathfrak{P}_{\alpha\phi^{-1}}$ *and* $\phi^{-1}(\mathfrak{P}_\alpha) \subset \mathfrak{P}_{\alpha\phi}$.
2. $\{\mathfrak{P}_\alpha, \mathfrak{P}_\beta\} \subset \mathfrak{P}_{\alpha\phi^{-1}+\beta\phi^{-1}}$.
3. *If $\alpha \in \Lambda$ then $\alpha\phi^{-z} \in \Lambda$ for any $z \in \mathbb{Z}$.*
4. $\mathfrak{P}_0 = H$.

**Definition 3.** *A* split Hom-Poisson algebra *is a Hom-Poisson algebra in which the Hom-Lie algebra $(\mathcal{P}, \{\cdot,\cdot\})$ is split respect to a MASA $H$ of $(\mathcal{P}, \{\cdot,\cdot\})$.*

From Hom-Leibniz identity we get.

**Lemma 2.** *Let $\mathfrak{P}$ be a split regular Hom-Poisson algebra. Then for any $\alpha, \beta \in \Lambda \cup \{0\}$ we have that $\mathfrak{P}_\alpha\mathfrak{P}_\beta \subset \mathfrak{P}_{\alpha\phi^{-1}+\beta\phi^{-1}}$.*

Throughout this paper we will denote by $\mathbb{N}$ the set of all non-negative integers and by $\mathbb{Z}$ the set of all integers. Finally, we would like to note that our split regular Hom-Poisson algebras are considered of arbitrary dimension and over an arbitrary base field $\mathbb{K}$.

In the following, $\mathfrak{P}$ denotes a split regular Hom-Poisson algebra and

$$\mathfrak{P} = H \oplus (\bigoplus_{\alpha \in \Lambda} \mathfrak{P}_\alpha)$$

the corresponding root spaces decomposition.

Given a linear functional $\alpha : H \to \mathbb{K}$, we denote by $-\alpha : H \to \mathbb{K}$ the element in $H^*$ defined by $(-\alpha)(h) := -\alpha(h)$ for all $h \in H$. We also denote by

$$-\Lambda = \{-\alpha : \alpha \in \Lambda\}.$$

**Definition 4.** *Let $\alpha$ and $\beta$ be two elements in $\Lambda$. We will say that $\alpha$ is* connected *to $\beta$ if either $\beta = \epsilon\alpha\phi^z$ for some $z \in \mathbb{Z}$ and $\epsilon \in \{\pm 1\}$ or there exists $\{\alpha_1, \alpha_2, ..., \alpha_k\} \subset \pm\Lambda$, $k \geq 2$ such that:*

1. $\alpha_1 \in \{\alpha\phi^{-n} : n \in \mathbb{N}\}$.
2. $\alpha_1\phi^{-1} + \alpha_2\phi^{-1} \in \pm\Lambda$,
   $\alpha_1\phi^{-2} + \alpha_2\phi^{-2} + \alpha_3\phi^{-1} \in \pm\Lambda$,
   $\alpha_1\phi^{-3} + \alpha_2\phi^{-3} + \alpha_3\phi^{-2} + \alpha_4\phi^{-1} \in \pm\Lambda$,
   .....
   $\alpha_1\phi^{-i} + \alpha_2\phi^{-i} + \alpha_3\phi^{-i+1} + \cdots + \alpha_{i+1}\phi^{-1} \in \pm\Lambda$,
   .....
   $\alpha_1\phi^{-k+2} + \alpha_2\phi^{-k+2} + \alpha_3\phi^{-k+3} + \cdots + \alpha_i\phi^{-k+i} + \cdots + \alpha_{k-1}\phi^{-1} \in \pm\Lambda$.
3. $\alpha_1\phi^{-k+1} + \alpha_2\phi^{-k+1} + \alpha_3\phi^{-k+2} + \cdots + \alpha_i\phi^{-k+i-1} + \cdots + \alpha_k\phi^{-1} \in \{\pm\beta\phi^{-m} : m \in \mathbb{N}\}$.

*In this case, we will also say that $\{\alpha_1, ..., \alpha_k\}$ is a* connection *from $\alpha$ to $\beta$.*

The proof of the next result is analogous to the one of [1, Proposition 2.4].

**Proposition 1.** *The relation $\sim$ in $\Lambda$, defined by $\alpha \sim \beta$ if and only if $\alpha$ is connected to $\beta$ is of equivalence.*

From Proposition 2 we can consider the quotient set $\Lambda/\sim = \{[\alpha] : \alpha \in \Lambda\}$, becoming $[\alpha]$ the set of nonzero roots $\mathfrak{P}$ which are connected to $\alpha$. Fix $[\alpha] \in \Lambda/\sim$, we define the linear space $I_{H,[\alpha]} \subset H$ as follows

$$I_{H,[\alpha]} := span_{\mathbb{K}}\{\{\mathfrak{P}_\beta, \mathfrak{P}_{-\beta}\} + \mathfrak{P}_\beta\mathfrak{P}_{-\beta} : \beta \in [\alpha]\} \subset H.$$

Next, we define

$$V_{[\alpha]} := \bigoplus_{\beta \in [\alpha]} \mathfrak{P}_\beta.$$

Finally, we denote by $I_{[\alpha]}$ the direct sum of the two subspaces above, that is,

$$I_{[\alpha]} := I_{H,[\alpha]} \oplus V_{[\alpha]}.$$

The reference [2, Section 2] allows us to assert the next results.

**Theorem 1.** *The following assertions hold.*

1. *For any* $[\alpha] \in \Lambda/\sim$*, the linear space* $I_{[\alpha]} = I_{H,[\alpha]} \oplus V_{[\alpha]}$ *of* $\mathfrak{P}$ *associated to* $[\alpha]$ *is an ideal of* $\mathfrak{P}$.
2. *If* $\mathfrak{P}$ *is simple, then there exists a connection from* $\alpha$ *to* $\beta$ *for any* $\alpha, \beta \in \Lambda$ *and* $H = \sum_{\alpha\in\Lambda} (\{\mathfrak{P}_\alpha, \mathfrak{P}_{-\alpha}\} + \mathfrak{P}_\alpha\mathfrak{P}_{-\alpha})$.

**Theorem 2.** *We have* $\mathfrak{P} = U + \sum_{[\alpha]\in\Lambda/\sim} I_{[\alpha]}$*, where* $U$ *is a linear complement in* $H$ *of* $span_{\mathbb{K}}\{\{\mathfrak{P}_\alpha, \mathfrak{P}_{-\alpha}\} + \mathfrak{P}_\alpha\mathfrak{P}_{-\alpha} : \alpha \in \Lambda\}$ *and any* $I_{[\alpha]}$ *is one of the ideals of* $\mathfrak{P}$ *described in Theorem 1-1, satisfying* $\{I_{[\alpha]}, I_{[\beta]}\} + I_{[\alpha]}I_{[\beta]} = 0$ *if* $[\alpha] \neq [\beta]$.

## 2 Strongly split regular Hom-Poisson algebras

**Definition 5.** A split regular Hom-Poisson algebra $\mathfrak{P}$ with set of nonzero roots $\Lambda$ is called a *strongly split regular Hom-Poisson algebra* if $H = \sum_{\alpha\in\Lambda} (\{\mathfrak{P}_\alpha, \mathfrak{P}_{-\alpha}\}+\mathfrak{P}_\alpha\mathfrak{P}_{-\alpha})$ and given $\alpha, \beta \in \Lambda$ such that $\alpha\phi^{-1} + \beta\phi^{-1} \in \Lambda$ then we have $\{\mathfrak{P}_\alpha, \mathfrak{P}_\beta\} + \mathfrak{P}_\alpha\mathfrak{P}_\beta = \mathfrak{P}_{\alpha\phi^{-1}+\beta\phi^{-1}}$.

Let us focuss for a while on the concept of split-ideal in the framework of split Hom-Poisson algebras. An ideal $I$ of a regular split Hom-Poisson algebra $\mathfrak{P}$ is called a *split-ideal* if $I \cap H \neq 0$. A regular split Hom-Poisson algebra $\mathfrak{P}$ will be called *split-simple* if $\{\mathfrak{P}, \mathfrak{P}\}, \mathfrak{P}\mathfrak{P} \neq 0$ and it has no proper split-ideals. Finally, we recall that a root system $\Lambda$ is called *symmetric* if it satisfies that $\alpha \in \Lambda$ implies $-\alpha \in \Lambda$. From now on $\Lambda$ will be supposed symmetric.

Consider then $\mathfrak{P}$ a strongly split regular Hom-Poisson algebra and

$$\mathfrak{P} = H \oplus (\bigoplus_{\alpha\in\Lambda} \mathfrak{P}_\alpha)$$

the corresponding root spaces decomposition. Since $\mathfrak{P}$ is in particular a split regular Hom-Poisson algebra, all of the results developed in Section 2 hold. By Theorem 1-1 we know that for any $[\alpha] \in \Lambda/\sim$, the linear space $\mathfrak{P}_{[\alpha]}$ is an ideal of $\mathfrak{P}$. Furthermore, it is an split-ideal as consequence of the facts $\{H_{[\gamma]}, V_{[\alpha]}\} = 0$ for any $[\gamma] \neq [\alpha]$, $H = \sum_{[\beta]\in\Lambda/\sim} H_{[\beta]}$ and $\alpha \neq 0$. From here, we can also assert that any $\mathfrak{P}_{[\alpha]}$, $[\alpha] \in \Lambda/\sim$, is a strongly split-ideal admitting the split decomposition

$$\mathfrak{P}_{[\alpha]} = H_{[\alpha]} \oplus (\bigoplus_{\beta\in[\alpha]} \mathfrak{P}_\beta).$$

**Proposition 2.** *If* $(\mathfrak{P}, \{\cdot,\cdot\})$ *is centerless then any* $\mathfrak{P}_{[\alpha]}$ *is split-simple.*

*Proof.* Consider a split-ideal $I$ of $\mathfrak{P}_{[\alpha]}$. Since $I$ is also an ideal of the centerless split regular Hom-Lie algebra $(\mathfrak{P}_{[\alpha]}, \{\cdot,\cdot\})$, [1, Lemma 4.3] gives us that we can write

$$I = (I \cap H_{[\alpha]}) \oplus (\bigoplus_{\beta\in[\alpha]} (I \cap \mathfrak{P}_\beta))$$

with $I \cap H_{[\alpha]} \neq 0$. For any $0 \neq h \in I \cap H_{[\alpha]}$, the fact $(\mathfrak{P}_{[\alpha]}, \{\cdot,\cdot\})$ is centerless gives us that there exists $\beta \in [\alpha]$ such that $\{h, \mathfrak{P}_\beta\} \neq 0$. From here we get $\{I \cap H_{[\alpha]}, \mathfrak{P}_\beta\} = \mathfrak{P}_\beta$ and so $0 \neq \mathfrak{P}_\beta \subset I$. Since $\phi(I) = I$ we get as consequence that

$$\mathfrak{P}_{\beta\phi^z} \subset I \text{ for any } z \in \mathbb{Z}.$$

Given now any $\delta \in [\alpha] \setminus \{\pm\beta\phi^z : z \in \mathbb{Z}\}$, the fact that $\beta$ and $\delta$ are connected allows us to take a connection $\{\alpha_1, \alpha_2, ..., \alpha_k\}$ from $\beta$ to $\delta$. Since

$$\alpha_1, \alpha_2, \alpha_1\phi^{-1} + \alpha_2\phi^{-1} \in \Lambda$$

we have

$$\{\mathfrak{P}_{\alpha_1}, \mathfrak{P}_{\alpha_2}\} + \mathfrak{P}_{\alpha_1}\mathfrak{P}_{\alpha_2} = \mathfrak{P}_{\alpha_1\phi^{-1}+\alpha_2\phi^{-1}} \subset I$$

as consequence of $\mathfrak{P}_{\alpha_1} = \mathfrak{P}_\beta \subset I$. In a similar way

$$\{\mathfrak{P}_{\alpha_1\phi^{-1}+\alpha_2\phi^{-1}}, \mathfrak{P}_{\alpha_3}\} + \mathfrak{P}_{\alpha_1\phi^{-1}+\alpha_2\phi^{-1}}\mathfrak{P}_{\alpha_3} = \mathfrak{P}_{\alpha_1\phi^{-2}+\alpha_2\phi^{-2}+\alpha_3\phi^{-1}} \subset I$$

and we finally get by following this process that

$$\mathfrak{P}_{\alpha_1\phi^{-k+1}+\alpha_2\phi^{-k+1}+\alpha_3\phi^{-k+2}+\cdots+\alpha_i\phi^{-k+i-1}+\cdots+\alpha_k\phi^{-1}} = \mathfrak{P}_{\epsilon\delta\phi^{-m}} \subset I$$

for some $m \in \mathbb{N}$ and $\epsilon \in \pm 1$. From here we have, taking into account $\phi(I) = I$, that $\mathfrak{P}_{\epsilon\delta} \subset I$ and conclude $H_{[\alpha]} \subset I$. Consequently, taking also into account that $\{H_{[\alpha]}, \mathfrak{P}_\delta\} = \mathfrak{P}_{\delta\phi^{-1}}$ for any $\delta \in [\alpha]$, we get $V_{[\alpha]} \subset I$. We have showed $I = \mathfrak{P}_{[\alpha]}$ and so $\mathfrak{P}_{[\alpha]}$ is split-simple.

**Theorem 3.** *Any strongly split regular Hom-Poisson algebra $\mathfrak{P}$ such that $(\mathfrak{P}, \{\cdot,\cdot\})$ is centerless is the direct sum of split-ideals, each one being a split-simple strongly split Hom-Poisson algebra.*

*Proof.* Since we can write the disjoint union $\Lambda = \bigcup_{[\alpha]\in\Lambda\setminus\sim} [\alpha]$ we have $\mathfrak{P} = \sum_{[\alpha]\in\Lambda\setminus\sim} \mathfrak{P}_{[\alpha]}$. Let us now verify the direct character of the sum: given $x \in \mathfrak{P}_{[\alpha]} \cap \sum_{\substack{[\beta] \in \Lambda/\sim \\ \beta \nsim \alpha}} \mathfrak{P}_{[\beta]}$, since we have $\{\mathfrak{P}_{[\alpha]}, \mathfrak{P}_{[\beta]}\} = 0$ for $[\alpha] \neq [\beta]$, we obtain

$$\left\{x, \mathfrak{P}_{[\alpha]}\right\} + \left\{x, \sum_{\substack{[\beta] \in \Lambda/\sim \\ \beta \nsim \alpha}} \mathfrak{P}_{[\beta]}\right\} = 0.$$

From here $\{x, \mathfrak{P}\} = 0$ and so $x = 0$, as desired. Consequently we can write

$$\mathfrak{P} = \bigoplus_{[\alpha]\in\Lambda\setminus\sim} \mathfrak{P}_{[\alpha]}.$$

Finally, Proposition 2 completes the proof.

## References

1. Aragón, M.J. and Calderón, A.J.: Split regular Hom-Lie algebras. Journal of Lie Theory. 25(3) (2015), 875–888.
2. Aragón, M.J. and Calderón, A.J.: The structure of split regular Hom-Poisson algebras. Colloquium Mathematicum. In press.
3. Makhlouf, A. and Silvestrov, S.: Notes on 1-parameter formal deformations of Hom-Lie and Hom-Lie algebras. Forum Math., 22(4) (2010), 715–739.

# Variational-like inequalities in a fuzzy and interval-valued context

Gabriel Ruiz-Garzón[1], Rafaela Osuna-Gómez[2]
Antonio Rufián-Lizana[2] and Yurilev Chalco-Cano[3]

[1] Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, Facultad de Ciencias Sociales y de la Comunicación, Campus de Jerez de la Frontera, Cádiz, 11405, Spain, e-mail: gabriel.ruiz@uca.es
[2] Departamento de Estadística e Investigación Operativa, Universidad de Sevilla, Facultad de Matemáticas, Sevilla, 41012, Spain
[3] Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile

**Abstract.** The aim of this work is to generalize classical concepts of generalized convexity and relationships between solutions of variational-like inequalities and mathematical programming problems given in the finite dimensional case, to a fuzzy and interval-valued context. Firstly, in this work we have shown that invex fuzzy mappings are more useful than differential convex fuzzy mapping for fuzzy optimization. Secondly, we relate solutions of Stampacchia Interval Variational-Like Inequality with the weak minimum of the Interval Continuous-Time Problem. Some results proved are related to the search for equilibrium points in the Walrasian equilibrium price model, in the Wardrop's principle for traffic equilibrium problem and in oligopolistic market equilibrium problem.

**Keywords:** Generalized convexity fuzzy mappings; Variational-like inequality; Fuzzy optimization; Interval continuous-time problems.

## 1 Introduction

The imprecision occurring in the optimization problems is categorized as the fuzzy optimization problems or interval optimization problems.

In fuzzy context, the study of generalized convex fuzzy mapping and its connection with fuzzy optimization has been studied by many authors. For instance, in 1992, Nanda and Kar [3] proposed a concept of convex fuzzy mapping and proved that a fuzzy mapping is convex if and only if its epigraph is a convex set. In their book, [6], Ramik and Vlach gave several types of generalized convex sets and generalized concave functions based on the support set of a fuzzy set. Syau [10] introduced the concepts of pseudoconvexity, invexity and pseudoinvexity for fuzzy mappings of one variable by using a notion of differentiability.

Recently, Wu and Xu [13] introduced the concepts of fuzzy pseudoconvex, fuzzy invex, fuzzy pseudoinvex and fuzzy preinvex mappings from $\mathbb{R}^n$ to the set of fuzzy intervals based on the concept of differentiability of fuzzy mapping due to Wang and Wu [12]. More recently, using the derivative for fuzzy mapping due to Seikkala [9] (or Buckley and Feuring [2]), Panigrahi [4] introduced

generalized convex fuzzy mappings such as quasiconvexity, strict quasiconvexity, strong quasiconvexity and pseudoconvexity and he studied its connection with fuzzy optimization. In addition, using the derivative for fuzzy mapping due to Seikkala [9], Wu and Xu [14] introduced the concepts of invex and incave fuzzy mappings.

In all the previous articles, restrictive concepts of differentiable fuzzy mappings were used. For instance, the concept of derivative for fuzzy mapping due to Seikkala [9], similar to the Hukuhara-derivative introduced by Puri and Ralescu [5], possesses the following property: the diameter of the support $supp(F(t))$ of a differentiable fuzzy mapping $F$ is nondecreasing as $t$ increases. Consequently, other concepts of derivative for fuzzy mappings more general than previous concepts were well introduced. For instance, Bede and Gal [1] introduced the concept of strongly generalized differentiable fuzzy mapping, and Stefanini and Bede [11] introduced the concept of generalized Hukuhara derivative.

We show that a differentiable and convex fuzzy mapping $F : K \to E$ does not necessarily verify the following inequality

$$F(x) \succcurlyeq \tilde{\nabla} F(y)(x-y) + F(y), \tag{1}$$

for all $x, y \in K$. However, in some cases the inequality (1) is valid only when $x-y$ is replaced by an adequate $\eta(x, y)$. Consequently, the introduction of invex fuzzy mappings is still needed. Further, solutions of a fuzzy variational-like inequality problem with solutions of fuzzy optimization problem are relating.

The interval-valued optimization problems may provide an alternative choice for considering the uncertainty into the optimization problems.

The aim of this work is define the Continuous-Time Problem in an interval context and obtain optimality conditions for this problem and find relations between solutions of Interval Continuous-Time Problem and solutions of Interval Variational-like Inequality Problems, such that the results of classic optimization problem are particular cases that here we get.

## 2 The necessity of invexity in fuzzy context

Let $E$ denote the family of fuzzy intervals, i.e, $E$ denotes the family of compact and convex fuzzy sets on $\mathbb{R}$. Obviously, $[\mu]^\alpha$ is a nonempty compact and convex subset of $\mathbb{R}$ (denoted $[\mu_*(\alpha), \mu^*(\alpha)]$) for any $\mu \in E$ and $\alpha \in [0, 1]$.

The fuzzy optimization problem consists of

$$(\mathbf{FOP}) \begin{cases} \min F(x) \\ \text{s.t. } x = (x_1, ..., x_n) \in K \subseteq \mathbb{R}^n. \end{cases}$$

We recall that if there exists a $\delta$-neighborhood $N_\delta(x^*)$ around $x^* \in K$, such that for no x $(\neq x^*) \in K \cap N_\delta(x^*)$, $F(x) \preccurlyeq F(x^*)$, then $x^*$ is called a strict local optimal solution (SLOS).

The following definition of convexity for fuzzy mapping is well-known in the literature, for instance see [3, 4, 6, 13, 14].

**Definition 1.** *A fuzzy mapping $F : K \to E$ is said to be convex on a convex set $K \subset \mathbb{R}^n$ if for any $x, y \in K$, $\lambda \in [0,1]$,*

$$F(\lambda x + (1-\lambda)y) \preccurlyeq \lambda F(x) \tilde{+} (1-\lambda)F(y).$$

Note that $F$ is convex if and only if the endpoint functions $F_*(\cdot)(\alpha)$, $F^*(\cdot)(\alpha)$ are convex functions for all $\alpha \in [0,1]$. It's to possible to find a differentiable convex fuzzy mapping does not necessarily verify the following inequality $F(x) \succcurlyeq F'(y)(x-y)\tilde{+}F(y)$, for all $x, y \in K$, where $F'$ is the strongly generalized differential given by Bede and Gal [1].

To solve this inconvenience other concepts of invex and incave fuzzy mappings are given.

**Definition 2.** *A G-differentiable fuzzy mapping $F : K \subset \mathbb{R}^n \to E$ is called fuzzy invex (FIX) with respect to a function $\eta : K \times K \to \mathbb{R}^n$, if for all $x, y \in K$*

$$F(x) \succcurlyeq F'(y)\eta(x,y)\tilde{+}F(y).$$

**Definition 3.** *A G-differentiable fuzzy mapping $F : K \subset \mathbb{R}^n \to E$ is called fuzzy incave (FIC) with respect to a function $\eta : K \times K \to \mathbb{R}^n$, if for all $x, y \in K$*

$$F(x) \preccurlyeq F'(y)\eta(x,y)\tilde{+}F(y).$$

**Definition 4.** *A G-differentiable fuzzy mapping $F : K \subset \mathbb{R}^n \to E$ is called fuzzy strictly incave (FSIC) with respect to a function $\eta : K \times K \to \mathbb{R}^n$, if for all $x, y \in K$*

$$F(x) \prec F'(y)\eta(x,y)\tilde{+}F(y), \quad \forall x \neq y.$$

## 3 Fuzzy variational-like inequalities

In this section, we can define the next problems:

**Definition 5.** *Let $K \subset \mathbb{R}$ be an open convex set, $x^* \in K$, and let $F : K \to E$ be a G-differentiable fuzzy mapping. The Stampacchia fuzzy variational-like inequality problem is to find $x^* \in K$, such that there does not exist another $x \in K$, so that*

$$(SFVLIP) \quad F'(x^*)\eta(x,x^*) \preccurlyeq 0$$

**Definition 6.** *Let $K \subset \mathbb{R}$ be an open convex set, $x^* \in K$, and let $F : K \to E$ be a G-differentiable fuzzy mapping. The Minty fuzzy variational-like inequality problem is to find $x^* \in K$, such that there does not exist another $x \in K$, so that*

$$(MFVLIP) \quad F'(x)\eta(x,x^*) \succcurlyeq 0$$

An example of fuzzy variational-like inequality problem is the fuzzy transportation equilibrium problem with the next notation:

- $G = (N, A)$ a network graph $G$ in which $A$ represents the set of arcs and N the set node;

- $W$ the set of the origin-destination $(O-D)$, $|W|$ represents number of $(O-D)$;
- $R^w$ the set of paths connecting $O-D$ pair $w \in W$
- $R$ represents the set of all paths, $|R|$ represents number of paths, $r$ represents $r \in R$;
- $d_w$ the demand of $O-D$ pair $w \in W$, $d$ the demand vector $W$;
- $f_w^r$ the flow of path $r$ connecting $O-D$ pair $w$, $f$ the flow vector of paths;
- $x_a$ the flow of section $a$, $x$ the flow vector of sections;
- $\widetilde{c}_a$ the fuzzy travel costs of section $a$; $\widetilde{c}$ the fuzzy travel cost vector;
- $\widetilde{C}_r^w$ the fuzzy travel cost of path $r$ connecting $O-D$ pair $w$; $\widetilde{C}$ the fuzzy travel cost vector of paths;
- $\widetilde{\lambda}_w^*$ the minimum fuzzy travel cost of $O-D$ pair;

$\delta_{ar}^w = \begin{cases} 1, \text{if path} \quad r \in R \quad \text{traverses arc } a \in A, \\ 0, \text{otherwise.} \end{cases}$

**Theorem 1 (The Wardrop first principle).** *For each pair of origin-destination, the paths used have cost less than or equal to unused paths.*

The mathematical formulation of previous principle are the fuzzy transportation equilibrium conditions:

$$\widetilde{C}_r^w \begin{cases} = \widetilde{\lambda}_w^*, f_w^* > 0; \\ \qquad\qquad \forall r \in R_w, \forall w \in W, \text{ where } \widetilde{C}_r^{w^*} = \sum_{a \in A} \widetilde{c}_a^* \delta_{ar}^w \\ \geq \widetilde{\lambda}_w^*, f_w^* = 0; \end{cases}$$

where $f$ satisfies three conditions:

(a) $\sum_{r \in R_w} f_r^w = d_w, \forall w \in W$
(b) $f_r^w \geq 0, r \in R_w$
(c) $x_a = \sum_{w \in W} \sum_{r \in R_w} f_r^w \delta_{ar}^w, \forall a \in A$

In a user-equilibrium network, no traveller can improve his/her travel cost by unilaterally changing routes. The fuzzy user equilibrium model is equivalent to obtain a solution of fuzzy variational inequality model: finding a vector $f^* \in \Omega_f = \{f | f \text{satifies:(a)-(c)}\}$, such that $\widetilde{C}(f^*)(f - f^*) \leq 0, \forall f \in \Omega_f$.

We will prove that:

**Theorem 2.** *Let $K \subset \mathbb{R}$ be an open convex set, $x^* \in K$, and let $F : K \to E$ be a G-differentiable fuzzy invex mapping (FIX) with respect to $\eta$. If $x^*$ is a solution of $(SFVLIP)$, then $x^*$ is a strict local optimal solution (SLOS) of (FOP).*

**Theorem 3.** *Let $K \subset \mathbb{R}$ be an open convex set, $x^* \in K$, and let $F : K \to E$ be a G-differentiable fuzzy incave mapping (FIC) with respect to $\eta$. If $x^*$ is a strict local optimal solution (SLOS) of (FOP), then $x^*$ is a solution of (SFVLIP).*

**Theorem 4.** *Let $K \subset \mathbb{R}$ be an open convex set, $x^* \in K$, and let $F : K \to E$ be a G-differentiable fuzzy invex mapping (FIX) with respect to $\eta$. If $x^*$ is a strict local optimal solution (SLOS) of (FOP), then $x^*$ is a solution of (MFVLIP).*

## 4 Interval variational-like inequalities

We denote by $\mathcal{K}_C$ the family of all bounded closed intervals in $\mathbb{R}$, i.e.,

$$\mathcal{K}_C = \{[\underline{a}, \overline{a}] \ / \ \underline{a}, \overline{a} \in \mathbb{R} \text{ and } \underline{a} \leq \overline{a}\}$$

The Interval Continuous-Time Programming Problem $(ICTP)$ is defined as,

$$\begin{array}{l} \min \phi(x) = \int_0^T F(x(t),t)dt = \left[\int_0^T \underline{f}(x(t),t)dt, \int_0^T \overline{f}(x(t),t)dt\right] \\ \text{s.t.} \quad x \in X \end{array}$$

Here $X$ is a nonempty convex subset of the Banach space $L^n_\infty$, $\phi : X \to \mathcal{K}_C$. Let V be an open subset of $\mathbb{R}^n$ containing the set $\{\underline{x}(t) \in \mathbb{R}^n : x \in X, t \in [0,T]\}$. Let $F : V \times [0,T] \to \mathcal{K}_C$. We assume that $\underline{f}, \overline{f}$, are real functions defined on $V \times [0,T]$ and this functions are assumed to be Lebesgue measurable and integrable for all $x \in X$ and continuously differentiable functions with respect to their first argument.

**Definition 7.** *Let $\phi$ be an interval-valued function defined on $X$. It is said that $\bar{x} \in X$ is a (local) weak minimum (WM) for (ICTP) if there does not exist $x \in X$ $(\exists \delta > 0,\ x \in B(x_0, \delta) \cap X)$ such that $\phi(x) \prec \phi(\bar{x})$.*

The variational inequality problem has a general formulation that includes, among others, nonlinear equations, optimization problems, complementarity problems and fixed point problems. Variational inequalities were originally developed as a tool for the study of certain partial differential equations classes such as those that arise in mechanics. We define the next interval variational-like inequality problems.

The Stampacchia Interval Variational-Like Inequality Problem $(SIVLIP)$ finds $x^* \in X$ such that there exists no $x \in X$ satisfying

$$\int_0^T \tilde{\nabla} F(x^*(t),t)\eta(x(t),x^*(t),t)edt \prec [0,0]$$

The Minty Interval Variational-Like Inequality Problem $(MIVLIP)$ finds $x^* \in X$ such that there exists no $x \in X$ satisfying

$$\int_0^T \tilde{\nabla} F(x(t),t)\eta(x^*(t),x(t),t)edt \succ [0,0]$$

In some recent contributions, the Minty variational inequality problem has been termed a "dual" variational inequality problem in order to indicate its close relationship to the classical "primal" Stampacchia Inequality.

Let us now look for the conditions under which we can relate the Stampacchia Interval Variational-Like Inequality Problem $(SIVLIP)$ solutions with the weak minimums of $(ICTP)$.

**Theorem 5.** *Let $F : V \times [0,T] \to \mathcal{K}_C$ be a gH-differentiable function.*

*If $\phi(x) = \int_0^T F(x(t),t)dt$ is a pseudoinvex $(PIX)$ function respect to $\eta$ and $x^*$ solves the Stampacchia Interval Variational-Like Inequality Problem $(SIVLIP)$ respect to the same $\eta$ then $x^*$ is a weak minimum for $(ICTP)$.*

## 5 Conclusions and future research

In this work, we have defined the Stampacchia and Minty Interval Variational-Like Inequality Problems and we relate their solutions with the weak minimum for the Interval Continuous-Time Problem and Fuzzy Optimization Problem, under invexity hypothesis.

These results generalize another obtained in Euclidean spaces with classical mathematical programming problems or Continuous-Time Problems. We can generalize the Walrasarian equilibrium price model, the Wardrop's principle for traffic equilibrium problem and the oligopolistic market equilibrium problem to an environment of interval-valued functions.

An open question could be the study of optimality conditions able to characterize psuedoinvex functions through critical points in Interval Continuous-Time Problems, that have not been studied in the literature up to now.

## References

1. Bede, B., Gal, S.G.: Generalizations of the differentiability of fuzzy number valued functions with applications to fuzzy differential equation. Fuzzy Sets and Syst. 151, 581-599 (2005).
2. Buckley, J.J., Feuring, T.: Fuzzy differential equations. Fuzzy Sets and Syst. 110 43-54 (2000).
3. Nanda, S., Kar, K.: Convex fuzzy mappings. Fuzzy Sets and Syst. 32 359-367 (1989).
4. Panigrahi, M. et al.: Convex fuzzy mapping with differentiability and its applicaction in fuzzy optimization. Eur. J. Oper. Res. 185 47-62 (2007).
5. Puri, M.L., Ralescu, D.A.: Differentials for fuzzy functions. J. Math. Anal. Appl. 91 552-558 (1983).
6. Ramik, J., Vlach, M.: Generalized Concavity in Fuzzy Optimization and Decision Analysis. Kluwer Academics Publishers, Dordrecht, Boston, London, (2002).
7. Rufián-Lizana, A., Chalco-Cano, Y., Osuna-Gómez R. and Ruiz-Garzón, G.: On invex fuzzy mappings and fuzzy variational-like inequalities. Fuzzy Sets and Syst. 200 84-98 (2012).
8. Rufián-Lizana, A., Chalco-Cano, Y., Ruiz-Garzón G. and Román-Flores, H.: On some characterizations of preinvex fuzzy mappings, TOP 22 771–783 (2014).
9. Seikkala, S.: On the fuzzy initial value problem. Fuzzy Sets and Syst. 24 319-330 (1987).
10. Syau, Y.R.: Invex and generalized convex fuzzy mappings. Fuzzy Sets and Syst. 115 455-461 (2000).
11. Stefanini, L., Bede, B.: Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Anal. 71 (2009) 1311-1328 (2009).
12. Wang, G.X., Wu, C.X.: Directional derivatives and subdifferential of convex fuzzy mappings and application in convex fuzzy programming. Fuzzy Sets and Syst. 138 559-591 (2003).
13. Wu, Z., Xu, J.: Nonconvex fuzzy mappings and the fuzzy pre-variational inequality. Fuzzy Sets and Syst. 159 2090-2103 (2008).
14. Wu, Z., Xu, J.: Generalized convex fuzzy mappings and fuzzy variational-like inequality. Fuzzy Sets and Syst. 160 1590-1619 (2009).

# Mutation Testing in Event Programming Language

L. Gutiérrez-Madroñal, J.J Domínguez-Jiménez, and I. Medina-Bulo

UCASE Software Engineering Research Group,
University of Cadiz, Spain
{lorena.gutierrez,juanjose.dominguez,inmaculada.medina,}@uca.es
https://ucase.uca.es/

**Abstract.** Event processing queries are intended to process continuous event streams. These queries are partially similar to traditional SQL queries, but provide the facilities to express rich features (e.g., pattern expression, sliding window of length and time). An error while implementing a query may result in abnormal program behaviors and lost business opportunities. Mutation testing has been found to be effective to assess test suites quality and generating new test cases. In this work, we propose mutation-based testing of the Event Processing Language (EPL) 4.9.0 (a domain-specific language for processing events), mutation operators which modify different features of EPL queries and a new mutation analysis testing tool for EPL (MuEPL). Moreover, to evaluate MuEPL, we have applied to EPL programs.

**Keywords:** Mutation Testing, Event Processing Queries, Event Processing Language

## 1 Introduction

Mutation testing is a fault-based testing technique that introduces simple syntactic changes in the original program by applying *mutation operators*. Unlike other fault-based strategies that directly inject artificial faults into the program, the mutation method generates syntactic variations, *mutants*, of the original program by applying mutation operators. Each mutation operator represents "typical" programming errors, that the developer could make.

All mutants need to be run against the test suite to determine whether they can be told apart from the original program in some of its test cases. When a mutant can be told apart from the original program, the mutant has been *killed* by the test suite.

As for real-time, mutation testing has been applied to many traditional programming languages that have been growing and now they can be applied to real-time systems: Java [1], C [2], Ada [3]. The present study uses the mutation testing as a way to test a non traditional programming language. We apply mutation testing to a programming language created to be used in real-time systems, more particularly in Complex Event Processing (CEP) systems. To achieve this

goal its mutation operators must be defined, a mutation tool for the CEP system must be developed and finally they have to be tested (operators and mutation tool) using examples. Moreover, the mutation tool must satisfy certain real-time system peculiarities that hinder its implementation.

The structure of the rest of the paper is as follows: An introduction about mutation testing and EPL language are respectively shown in Section 2 and Section 3. Section 4 defines and clasifies the EPL mutation operators. Section 5 presents MuEPL, mutation testing tool for EPL. Section 6 describes the examples where MuEPL will be applied and evaluates the obtained results. Section 7 presents conclusions and future work.

## 2 Mutation Testing

Mutation testing is a fault-based testing technique providing a test criterion: the *mutation score*. This criterion can be used to measure the effectiveness of a test suite in terms of its ability to detect faults. Mutation testing generates mutants from the program under test by applying mutation operators to it. These mutation operators introduce slight syntactical changes into the program that should be detected by a high-quality test suite. Each mutation operator represents "typical" programming errors, that the developer could make. Thus, if a program contains the instruction `a > 2000` and we apply the relational mutation operator (which replaces a relational operator with another), the resulting mutant could contain the instruction `a >= 2000` instead, for example. If a test case is able to distinguish between the original program and the mutant, i. e. their outputs are different, it is said that this test case kills the mutant. On the contrary, if no test case in the test suite is able to distinguish between the mutant and the original program, it is said that the mutant stays *alive*. An *equivalent mutant* always produces the same output as the original program, hence it cannot be told apart from the original program. At this point it is necessary to clarify that program is used to denote the software under test, which could be a complete program or some smaller unit, such as a query.

## 3 Event Processing Language

Esper [4] is an open-source Java-based software product for Complex Event Processing (CEP) and Event Stream Processing (ESP), that analyzes series of events for deriving conclusions from them. It offers a domain-specific language for processing events called Event Processing Language (EPL). EPL is a declarative programming language for analyzing time-based event data and detecting situations as they occur.

EPL is a SQL like query language. However, unlike SQL that operates on tables, EPL operates on continuous stream of events. As a result, a row from a table in SQL is analogous to an event present in an event stream. An EPL statement starts executing continuously during runtime. While the execution is

taking place, EPL queries will be triggered if the application receives pre-defined or timer triggering events.

*EPL 4.9.0 query example*

```
select A as temp1, B as temp2 from pattern
     [every temp1.temperature > 400 -> temp2.temperature > 400]
```

EPL 4.9.0 is one of the latest versions of this language. In the above example a "Central" needs to measure the temperature of its systems, its temperature gauges take a reading of the core temperature every second and send the data to a central monitoring system. The EPL query of the figure, warns us if we have 2 consecutive temperatures above a certain threshold (400). This is a situation where it is needed to react quickly to emerging patterns in a stream of data events.

## 4 EPL mutation operators

In a previous work [5], was defined a list of mutation operators, some of them have been re-defined, and a new one has been defined. After checking several EPL queries that where used in real examples, the number of operators turn into 17. Due to the decreased number of mutation operators, the operators have been re-classified in four categories depending on which kind of EPL query element the are related to. These are identified by uppercase letters: P (*Pattern expression operators*), W (*Windows operators*), R (*Replacement operators*)[1] and I (*SQL Injection attack operators*).

Operators are uniquely identified by three uppercase letters: the first one is the category identifier, and the last two letters indicate the operator within the category. The Table 1 lists their names and provides a short descriptions of each of them. Some operators are specific for the EPL language, they appear marked with ☆.

## 5 MuEPL architecture

MuEPL lets us apply mutation analysis to EPL 4.9.0 queries, and it is the first tool with this capability. Figure 1 shows the relations between these components.

The capturer obtains the queries while the original program is under execution. Next the analyser receives the original queries and generates the information about which mutation operators can be applied and in which locations. This information is delivered to the mutant generator, which generates every possible mutant. Then, the execution engine runs the original and its mutant against the set of test cases and compares their behaviours to determine whether the mutants have been killed or stay alive.

[1] The *Replacement operators* may appear in the pattern of the query, but we considered extend their definition and apply them not only to the pattern but also to the rest of the query.

Due to the data nature which MuEPL deals with (events in real-time), we must ensure that all the programs receive the same events. This is because we want study the behaviour of the original program and the generated mutants under the same conditions. To achieve this, the execution system includes a mechanism that can synchronise the execution threads (original and mutants). This mechanism builds a "barrier" where all threads must wait, until all threads reach it, before any of the threads can continue.

| Operator | Description |
|---|---|
| | Pattern Expression Mutation |
| PNE | ☆ Removes the `not` keyword of the negated conditional expressions in the pattern |
| PTI | ☆ Increases the timer value by one unit in the pattern observer (`timer:at`, `timer:interval`) |
| PTD | ☆ Decreases the timer value by one unit in the pattern observer (`timer:at`, `timer:interval`) |
| | Replacement Mutation |
| RLO | Replaces a logical operator (`and`, `or`) by another of the same kind |
| RTU | Replaces one time unit (`milliseconds`, `seconds`, `minutes`, `hours`, `days`) by another of the same kind |
| RAF | Replaces an aggregate function (`max`, `min`, `avg`, `sum`, `count`, `median`, `stddev`, `avedev`) by another of the same kind. The `distinct` keyword could be also added |
| RAO | Replaces an arithmetic operator (`+`, `-`, `*`, `/`, `%`) by another of the same kind |
| RRO | Replaces a relational operator (`=`, `<>`, `<`, `>`, `<=`, `>=`) by another of the same kind |
| RNO | Replaces a number `e` by `e + 1` and `e - 1` |
| | Windows Mutation |
| WLI | ☆ Increases the data window length by one |
| WLD | ☆ Decreases the data window length by one |
| WTI | ☆ Increases the time window by one |
| WTD | ☆ Decreases the time window by one |
| WBL | ☆ Turn a batch window length into a ordinary window length |
| WBT | ☆ Turn a batch window time into a ordinary window time |
| | Injection Attack Mutation |
| IRC | Removes "where" condition from a query |
| INC | Negates the condition expression of a query |

**Table 1.** EPL mutation operators

## 6 Applying Mutation Testing to EPL programs

All the experiments in this section have been run in a Intel Core i7 machine with 4GB RAM and a 2.00GHz x 4, running Ubuntu 14.04 LTS.

We have used two programs from EsperTech website [4] to experimentally evaluate the mutation operators for EPL and MuEPL. The chosen programs are: *Self-Service Terminal* and *Transaction 3-Event Challenge*.

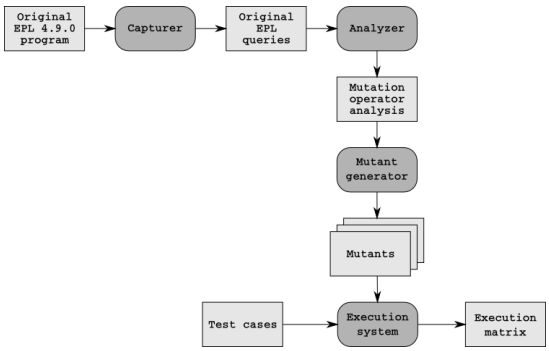



**Fig. 1.** MuEPL architecture

For each program it was needed to include a seed as input (and modify the random code which receive it). This is another undertaken measure to study the behaviour of the programs under the same circumtances.

**Self-Service Terminal** is about a J2EE-based self-service terminal managing system in an airport that gets a lot of events from connected terminals. The event rate is around 500 events per second. Some events indicate abnormal situations such as "paper low" or "terminal out of order". Other events observe activity as customers use a terminal to check in and print boarding tickets.

This program executes 6 queries, and for these queries 16 of the 17 mutation operators are applicable and we have used a test suite of 5 test cases. The non-applied mutation operator is: RAF, and are generated 79 mutants. The exhaustive execution of 80 threads (mutants + original) in parallel against each of the 5 test cases takes more than 3'15 hours in total.

| Test cases | | | Outputs | | |
|---|---|---|---|---|---|
| Iterations | Sec. | Seed | Killed | Live | Fail |
| 100 | 20 | 10 | 72 | 7 | 0 |
| 1000 | 50 | 25 | 76 | 1 | 2 |
| 2500 | 100 | 25 | 78 | 0 | 1 |
| 10000 | 250 | 15 | 76 | 1 | 2 |
| 25000 | 100 | 0 | 78 | 0 | 1 |

**Table 2.** Self-Service Terminal outputs

The live mutants for the first test case are from the operators: PTD (1), RAO (2), RRO (1), RTU (3). In the second test case is from RRO, and for the fourth test case is from RTU. It is worth remarkable how the outputs of different mutants, depending on the number of iterations and the sleep seconds, are failed.

**Transaction 3-Event Challenge** tracks three components of a transaction. The example uses at least three components, since some engines have different performance or coding for only two events per transaction. Each component comes to the engine as an event that are generated by the included "event generator". The transaction events come completely out of order; a bucket (with a modifiable size) is filled, and when it is full, it is shuffled. The larger the bucket size, the more events potentially come in between two events in a given transaction and so, the more the performance characteristics like buffers, hashes/indexes and other structures are put to the test as the bucket size increases.
This program executes 5 queries, and for these queries 9 of the 17 mutation operators are applicable. The applied mutation operators are: RLO, WTI, WTD, RTU, WBT, IRC, INC, RAF, RAO and RRO, and are generated 435 mutants. The machine was not able to run 436 threads (original + mutants) in parallel against the simplest test cases: bucket size "tiniest" (20), number of transactions 10, seed 10.

## 7 Conclusions and future work

The present study propose the mutation testing as a way to test a programming language developed to be used in real-time systems, EPL. In spite of being a query language, the execution of all the mutants have a high cost. This is because of the main drawbacks of mutation testing: commonly there is a large number of mutation operators that generate a wide numbers of mutants, each of them must be executed against the test suite. Under certain conditions described in an empirical study by Offut et al. [6], the number of mutants has quadratic complexity in program size.

Several techniques have been described to solve this problem; one of them is *mutant reduction technique*, which process only a subset of all the mutants. As a future work, MuEPL will be adapted to apply one of the mutant reduction techniques: *Evolutionary Mutation Testing* (EMT) [7], that finds mutants that help derive new test cases that improve the quality of the initial test suite. Applying EMT, we will obtain *strong mutants*: suviving mutants (which have not been killed by the test suite) and difficult to kill mutants (which have been killed by one and only one test case that kills no other mutant). According to our study, the mutations can be hard to find in quick executions (less number of iterations), so we have to focus our experiments on this kind of test cases.

## References

1. Ma, Y., Offutt, A.J., Kwon, Y.: MuJava: a mutation system for Java. Proceedings of the 28th international conference on Software engineering, ACM (2006), pp. 827–830
2. Agrawal, H., DeMillo, R., Hathaway, B., Hsu, W., Hsu, W., Krauser, E., Martin, R., Mathur, A. Spafford, E.: Design of mutant operators for the C programming language (1999)

3. Offutt, J., Voas, J., Payne, J.: Mutation operators for ada. Technical Report ISSE-TR-96-09, Information and Software Systems Engineering, George Mason University (1996)
4. EsperTech website, `http://www.espertech.com/esper/index.php`
5. Gutiérrez-Madroñal, L. and Shahriar, H. and Zulkernine, M. and Dominguez-Jimenez, J.J. and Medina-Bulo, I.: Mutation Testing of Event Processing Queries. Software Reliability Engineering (ISSRE), 2012 IEEE 23rd International Symposium on, (2012), pp. 21-30, ISBN:1071-9458
6. Offutt, A.J., Rothermel, G., Zapf, C.: An experimental evaluation of selective mutation. Software Engineering, 1993. Proceedings., 15th International Conference on, pp. 100-107, ISSN 0270-5257
7. Domínguez-Jiménez, J.J., Estero-Botaro, A., García-Domínguez, A., Medina-Bulo, I.: Evolutionary mutation testing. Information and Software Technology, Elsevier (2011), v. 53, n.10 pp. 1108–1123

# A new characterization of discrete Sugeno integrals on bounded distributive lattices

Radomír Halaš[1], Radko Mesiar[2], and Jozef Pócs[1,3]

[1] Department of Algebra and Geometry, Palacký University Olomouc, Czech Republic
[2] Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Slovakia
[3] Mathematical Institute, Slovak Academy of Sciences, Košice, Slovakia

**Abstract.** We study compatible aggregation functions on a general bounded distributive lattice $L$, where the compatibility is related to the congruences on $L$. Surprisingly, a new characterization of discrete Sugeno integrals is obtained.

## 1 Introduction

A deeper study of aggregation problems on bounded lattices was encouraged at a recent conference ABLAT 2014. Recall that an $n$-ary aggregation function $g\colon L^n \to L$, where $(L, 0_L, 1_L, \leq)$ is a bounded lattice (or, more generally, a bounded poset) is characterized by its non-decreasingness in each coordinate and by two boundary conditions $g(0_L, \ldots, 0_L) = 0_L$ and $g(1_L, \ldots, 1_L) = 1_L$. Typical aggregation functions on a bounded distributive lattice are lattice polynomials $p\colon L^n \to L$ given by

$$p(a_1, \ldots, a_n) = \bigvee_{I \in \mathcal{J}} \Big( \bigwedge_{i \in I} a_i \Big),$$

where $\mathcal{J} \subseteq 2^{\{1,\ldots,n\}}$ is a non-empty subset of the power set of $\{1, \ldots, n\}$.

The main aim of this short note is the study of $n$-ary aggregation functions on a bounded distributive lattice $L$ which preserve congruences on $L$. As a by-product, a new characterization of discrete Sugeno integrals on $L$ is obtained, covering also the basic case of the real unit interval $L = [0, 1]$. Therefore, in the next section some basic information concerning Sugeno integrals is given. Section 3 brings our main results, characterizing aggregation functions preserving the congruences on $L$. In Section 4 the impact of our new results to the standard Sugeno integral is given. Finally some concluding remarks are added.

## 2 Sugeno integral

Sugeno integral was introduced in 1972 by M. Sugeno in a paper written in Japanese, and it became well-known due to Sugeno's PhD. thesis [10]. For a

measurable space $(X, \mathcal{A})$ and a monotone measure $m\colon \mathcal{A} \to [0,1]$, ($m(\emptyset) = 0$, $m(X) = 1$), the Sugeno integral $\mathsf{Su}_m(f)$ of a measurable function $f\colon X \to [0,1]$ is given by

$$\mathsf{Su}_m(f) = \bigvee_{t\in[0,1]} \big(t \wedge m(\{x \in X \mid f(x) \geq t\})\big). \tag{1}$$

For a finite space $X = \{x_1 \ldots, x_n\}$, $\mathcal{A} = 2^X$, $f\colon X \to [0,1]$ can be identified with a vector $\mathbf{u} \in [0,1]^n$, $\mathbf{u} = (u_1, \ldots, u_n) = \big(f(x_1), \ldots, f(x_n)\big)$, and formula (1) can be rewritten into

$$\mathsf{Su}_m(f) = \bigvee_{i=1}^{n} \big(u_i \wedge m(\{x \in X \mid f(x) \geq u_i\})\big). \tag{2}$$

An alternative formula for the discrete Sugeno integral was proposed in [9]:

$$\mathsf{Su}_m(f) = \bigvee_{I\subseteq\{1,\ldots,n\}} \big(m(I) \wedge \big(\bigwedge_{i\in I} u_i\big)\big). \tag{3}$$

Observe that the Sugeno integral can be seen as a special instance of Ky Fan metric [6] as a distance of the function $f$ and the zero function $\mathbf{0}$. There are several properties of the discrete Sugeno integral and some of their settings yield an axiomatic characterization of this integral. First of all, for a fixed $m \in \mathbb{N}$, $\mathsf{Su}_m$ can be seen as an aggregation function [3], i.e., $\mathsf{Su}_m\colon [0,1]^n \to [0,1]$ is non-decreasing in each coordinate, and it satisfies two boundary conditions $\mathsf{Su}_m(\mathbf{0}) = 0$ and $\mathsf{Su}_m(\mathbf{1}) = 1$. Next, $\mathsf{Su}_m$ is

- comonotone maxitive, i.e., $\mathsf{Su}_m(f \vee g) = \mathsf{Su}_m(f) \vee \mathsf{Su}_m(g)$ whenever $f$ and $g$ are comonotone (meaning that they are measurable with respect to a single chain in $2^X$);
- min-homogeneous, i.e., $\mathsf{Su}_m(\mathbf{c} \wedge f) = c \wedge \mathsf{Su}_m(f)$ for any constant $c \in [0,1]$, $\mathbf{c} = (c, \ldots, c) \in [0,1]^n$;
- horizontally maxitive, i.e., $\mathsf{Su}_m(f) = \mathsf{Su}_m(\mathbf{c} \wedge f) \vee \mathsf{Su}_m(f_c)$ for any $c \in [0,1]$, where $f_c(x_i) = 0$ if $f(x_i) \leq c$ and $f_c(x_i) = f(x_i)$ otherwise (observe that $f_c$ is the smallest function on $[0,1]^n$ such that $f = (\mathbf{c} \wedge f) \vee f_c$ );
- $\mathsf{Su}_m(1_E) = m(E)$, where $1_E$ is the characteristic function of a set $E \subseteq X$;
- idempotent, i.e., $\mathsf{Su}_m(\mathbf{c}) = c$ for any $c \in [0,1]$.

For these and several other properties of the discrete Sugeno integral we refer to [1, 7] and [2]. Based on the above references, the Sugeno integral can be characterized as an $[0,1]^n \to [0,1]$ aggregation function which is comonotone maxitive and min-homogeneous. Observe that the comonotone maxitivity can be replaced by the horizontal maxitivity. For some other axiomatizations of the Sugeno integral see [2]. Marichal [8] has observed an important link between the lattice polynomials on $[0,1]$ and the Sugeno integral. More precisely, he has shown that the class of all Sugeno integrals on $X$ with cardinality $n$ coincides with the class of all polynomial functions $p\colon L^n \to L$, $L = [0,1]$, which are idempotent. This result applies to discrete Sugeno integral defined on any bounded chain

$L$, considering the formulae (1)–(3), and replacing $[0,1]$ by $L$. Also the above mentioned axiomatizations of the Sugeno integral can be extended to any chain $L$. However, in the case of a general bounded distributive lattice $(L, 0_L, 1_L, \leq)$, one can apply formula (1) or (3), but not (2), in general. Hence, we can consider a lattice valued measure $m\colon 2^X \to L$, $m(\emptyset) = 0_L$, $m(X) = 1_L$, $m(E_1) \leq m(E_2)$ whenever $E_1 \subseteq E_2 \subseteq X$, and for any $f\colon X \to L$ define a discrete $L$-valued Sugeno integral by

$$\mathsf{Su}_m(f) = \bigvee_{I \subseteq \{1,\ldots,n\}} \big(m(I) \wedge \big(\bigwedge_{i \in I} f(x_i)\big)\big). \tag{4}$$

For more details we recommend [2].

## 3 Compatible aggregation functions on distributive lattices

In this section we clarify the connection between monotone compatible functions on distributive lattices and their lattice polynomials. As the main result we will show that these functions can be identified with Sugeno integrals.

Recall that a lattice $L$ is distributive, if it satisfies one (or, equivalently, both) of the distributive identities

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

for all $a, b, c \in L$.

**Definition 1.** *Let $L$ be a lattice. A binary relation $R \subseteq L^2$ is compatible on the lattice $L$ if $(a,b),(c,d) \in R$ imply $(a \vee c, b \vee d) \in R$ and $(a \wedge c, b \wedge d) \in R$ for any $a, b, c, d \in L$. By a congruence on $L$ we understand any compatible equivalence on $L$.*

**Definition 2.** *Let $L$ be a lattice and $n \in \mathbb{N} \cup \{0\}$ be a non-negative integer. By an $n$-ary polynomial on the lattice $L$ we mean any function $p : L^n \to L$ defined inductively as follows:*

- *for each $i \in \{1,\ldots,n\}$, the $i$-th projection $p(x_1,\ldots,x_n) = x_i$ is a polynomial,*
- *any constant function $p(x_1,\ldots,x_n) = a$ for $a \in L$ is a polynomial,*
- *if $p_1(x_1,\ldots,x_n)$ and $p_2(x_1,\ldots,x_n)$ are polynomials, then so does the functions $p_1(x_1,\ldots,x_n) \vee p_2(x_1,\ldots,x_n)$ and $p_1(x_1,\ldots,x_n) \wedge p_2(x_1,\ldots,x_n)$,*
- *any polynomial is obtained by finitely many of the preceding steps.*

Informally, lattice polynomials are functions obtained by composing variables and constant functions by using of lattice operations. Note that polynomials defined in this way are called as weighted lattice polynomials in [8].

To simplify expressions, for any $n$-ary function $f : L^n \to L$ on a lattice $L$ and $\mathbf{x} = (x_1,\ldots,x_n), \mathbf{y} = (y_1,\ldots,y_n) \in L^n$, we put $f(\mathbf{x}) := f(x_1,\ldots,x_n)$, and $\mathbf{x} \leq \mathbf{y}$ iff $x_i \leq y_i$ for all $i \in \{1,\ldots,n\}$.

**Definition 3.** *Let $L$ be a lattice. A function $f : L^n \mapsto L$ is called compatible if for any congruence $\theta$ on $L$, if $(x_i, y_i) \in \theta$ for all $i \in \{1, \ldots, n\}$, then $\big(f(\mathbf{x}), f(\mathbf{y})\big) \in \theta$.*

It can be easily seen that for any lattice, its polynomials are always compatible functions. Recall, that compatible relations and compatible functions represent the well-known notions, investigated in connection with various algebraic structures. In particular, compatible functions on distributive lattices have been studied in deep by many authors, we refer the reader to [4] or [5].

Now, let $g$ be an $n$-ary compatible aggregation function on a bounded distributive lattice $(L, 0_L, 1_L, \leq)$ (not necessarily finite!). We associate with $g$ a function $\bar{g} : \{0_L, 1_L\}^n \mapsto L$ given by stipulation

$$\bar{g}(\mathbf{x}) := g(\mathbf{x}).$$

The function $\bar{g}$ is called a *characteristic function* of $g$.

Surprisingly, as shown in [4], $\bar{g}$ completely characterizes $g$:

**Lemma 1.** *[4] Let $g$ be a compatible aggregation function on a distributive lattice $L$. Then the characteristic function $\bar{g}$ determines $g$ uniquely.*

Let us note that Lemma 1 uses the fact that every distributive lattice $L$ can be (canonically) embedded into a Boolean algebra (using a set-theoretical representation of $L$). Then the representation of a compatible function $f$ on $L$ can be written in the form

$$f(x_1, \ldots, x_n) = \bigvee \Big\{ \big( f(\mathbf{a}) \wedge \bigwedge_{i \in \mathbf{a}^{-1}(1_L)} x_i \wedge \bigwedge_{i \in \mathbf{a}^{-1}(0_L)} x_i' \big) \mid \mathbf{a} \in \{0_L, 1_L\}^n \Big\},$$

where the complements $'$ refer to this Boolean algebra and for $\mathbf{a} \in \{0_L, 1_L\}^n$ the symbol $i \in \mathbf{a}^{-1}(1_L)$ denotes the fact that $a_i = 1_L$, while $i \in \mathbf{a}^{-1}(1_L)$ denotes $a_i = 0_L$, see [5].

To obtain the main result of the paper, we modify the approach developed in [5]. For any $\mathbf{a} \in \{0_L, 1_L\}^n$ consider the functions

$$G_{\mathbf{a}}(\mathbf{x}) := g(\mathbf{a}) \wedge \bigwedge \big\{ x_i \mid i \in \mathbf{a}^{-1}(1_L) \big\}. \tag{5}$$

**Theorem 1.** *For any monotone compatible function $g$ the following equality holds:*

$$g(\mathbf{x}) = \bigvee \big\{ G_{\mathbf{a}}(\mathbf{x}) \mid \mathbf{a} \in \{0_L, 1_L\}^n \big\}. \tag{6}$$

*Proof.* Since functions on the both sides are compatible, due to Lemma 1, it is enough to prove the above equality only for boolean inputs $\mathbf{x} \in \{0_L, 1_L\}^n$. Consider an arbitrary $\mathbf{a} \in \{0_L, 1_L\}^n$. We have the following possibilities:

(1) Let $\mathbf{a} \not\leq \mathbf{x}$. Then there is $j \in \{1, \ldots, n\}$ with $a_j = 1_L$ (i.e., $j \in \mathbf{a}^{-1}(1_L)$) and $x_j = 0_L$, which yields $\bigwedge\{x_i \mid i \in \mathbf{a}^{-1}(1_L)\} = 0_L$. Consequently, we obtain $G_{\mathbf{a}}(\mathbf{x}) = g(\mathbf{a}) \wedge 0_L = 0_L$.

(2) Let $\mathbf{a} = \mathbf{x}$. Then, evidently, $\bigwedge\{x_i \mid i \in \mathbf{a}^{-1}(1_L)\} = \bigwedge\{x_i \mid i \in \mathbf{x}^{-1}(1_L)\} = \bigwedge 1_L = 1_L$ whenever $\mathbf{a}^{-1}(1_L) \neq \emptyset$, and it equals $\bigwedge \emptyset = 1_L$ in case $\mathbf{a}^{-1}(1_L) = \emptyset$. In both cases we obtain $G_{\mathbf{a}}(\mathbf{x}) = g(\mathbf{a}) \wedge 1_L = g(\mathbf{a}) = g(\mathbf{x})$ since we assumed $\mathbf{a} = \mathbf{x}$.
(3) Assume $\mathbf{a} < \mathbf{x}$. Then as the function $g$ is monotone and $\mathbf{a} < \mathbf{x}$, we conclude $G_{\mathbf{a}}(\mathbf{x}) \leq g(\mathbf{a}) \leq g(\mathbf{x})$.

The above discussion leads to the desired equality

$$g(\mathbf{x}) = \bigvee \big\{G_{\mathbf{a}}(\mathbf{x}) \mid \mathbf{a} \in \{0_L, 1_L\}^n\big\}.$$

□

Consequently, we obtain

**Corollary 1.** *Compatible aggregation functions on distributive lattices are just their weighted idempotent lattice polynomials.*

Let us stress that this statement does not depend on the cardinality of a lattice $L$, and hence it holds also in a classical case when $L = [a, b]$ is any bounded interval of reals.

## 4 Sugeno integral as a compatible aggregation function

Consider a lattice $([0, 1], 0, 1, \leq)$, where $L = [0, 1]$ is the real unit interval. Then each element $\mathbf{a} \in \{0, 1\}^n$ can be identified with a characteristic function of a subset $I$ of $\{1, \ldots, n\}$, $\mathbf{a} = 1_I$. Comparing formulae (3) and (6), the next result is obtained easily.

**Theorem 2.** *Let $A\colon [0, 1]^n \to [0, 1]$ be an aggregation function. Then the following are equivalent:*

(i) *$A$ is a compatible function*
(ii) *there is a monotone measure $m$ on $\{1, \ldots, n\}$ so that $A = \mathsf{Su}_m$, i.e., $A$ is the Sugeno integral with respect to the measure $m$.*

Note that the monotone measure $m$ in Theorem 2 is given by $m(I) = A(1_I)$. Our result brings a new characterization of the classical Sugeno integral in discrete setting. Evidently, due to (4), Theorem 2 can be extended to any bounded distributive lattice $(L, 0_L, 1_L, \leq)$.

## 5 Concluding remarks

We have introduced a new property which characterizes the discrete Sugeno integral not only in its original form, when $[0, 1]$-valued functions and measures are considered, but also in the case of general bounded lattices. This property, compatibility, has an important impact for decision procedures which will be the

topic of our next investigations. Here we recall only the next fact: in multicriteria decision problems based on $n$ criteria and dealing with alternatives described by score vectors from $[0,1]^n$, often the exact numerical scores are replaced by some ordinal scale, e.g. by a linguistic scale. The transition from numerical inputs to linguistic values is done by means of interval partitions of the original scale $[0,1]$. When looking for normed utility functions (i.e., aggregation functions) where the output recommendation based on linguistic values does not depend on the numerical values of score vectors, then due to Theorem 2 only Sugeno integrals (i.e., idempotent polynomials) can be considered.

## Acknowledgements

The first author was supported by the international project Austrian Science Fund (FWF)-Grant Agency of the Czech Republic (GAČR) number I 1923-N25; the second author by the Slovak VEGA Grant 1/0420/15; the third author by the ESF Fund CZ.1.07/2.3.00/30.0041 and by the Slovak VEGA Grant 2/0028/13.

## References

1. Benvenuti, P., Mesiar, R., Vivona, D.: Monotone set functions-based integrals, Handbook of measure theory, Elsevier, 2002.
2. Couceiro, M., Marichal, J.-L.: Characterizations of discrete Sugeno integrals as polynomial functions over distributive lattices, Fuzzy Sets and Systems 161, 694–707 (2010).
3. Grabisch, M., Marichal, J.-L., Mesiar, R., Pap, E.: Aggregation Functions, Cambridge University Press, Cambridge, 2009.
4. Grätzer, G.: Boolean functions on distributive lattices, Acta Math. Hungarica 15(1-2), 195–201 (1964).
5. Haviar, M., Ploščica, M.: Congruence preserving functions on distributive lattices, Algebra Universalis 59, 179–196 (2008).
6. Ky Fan: Entfernung zweier zufälligen Größen und die Konvergenz nach Wahrscheinlichkeit, Mathematische Zeitschrift 49(1), 681–683 (1943).
7. Marichal, J.-L.: Aggregation Operators for Multicriteria Decision Aid, Ph.D. Thesis, Institute of Mathematics, University of Liège, Liège 1998.
8. Marichal, J.-L.: Weighted lattice polynomials, Discrete Mathematics 309(4), 814–820 (2009).
9. Mesiar, R.: $k$-order pan-additive discrete fuzzy measures, Proc. 7th IFSA Congress, Academia, Prague, 488–490(1997).
10. Sugeno, M.: Theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokyo Institute of Technology, Tokyo 1974.

# A Multiobjective Evolutionary Algorithm for Infrastructure Location in Vehicular Networks

Renzo Massobrio[1], Jamal Toutouh[2], and Sergio Nesmachnow[1]

[1] Universidad de la República, Montevideo, Uruguay
{renzom,sergion}@fing.edu.uy
[2] Universidad de Málaga, Málaga, Spain
jamal@lcc.uma.es

**Abstract.** This article describes a multiobjective evolutionary algorithm applied to locating roadside infrastructure for vehicular networks over realistic urban areas. A multiobjective formulation of the problem is introduced, considering quality-of-service and cost objectives. In the experimental analysis performed over a real map of Málaga, using real traffic information and antennas, the proposed multiobjective evolutionary algorithm computes accurate trade-off solutions for the problem.

## 1 Introduction

Vehicular ad hoc networks (VANETs) comprise a set of communicating nodes (*vehicles*) equipped with on-board units and roadside units (RSUs) installed beside the roads. RSUs act as network access points with higher communication capabilities than the vehicles. Thus, if two mobile nodes cannot directly exchange information because they are out of range, they can use RSUs to relay information between each other via vehicle-to-infrastructure communications. Using a fixed infrastructure of RSUs is an efficient alternative in order to improve the communication capabilities of VANETs.

Deploying such an infrastructure is a challenge because designers must decide about the number, type, and location of RSUs to maximize quality-of-service (QoS), while satisfying the deployment cost requirements.

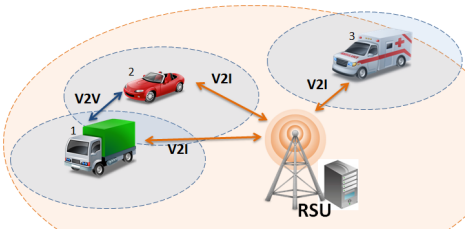



**Fig. 1.** Global VANET architecture.

The RSU Deployment Problem (RSU-DP) consists of placing a set of RSU terminals along the roads of a given area, maximizing the network QoS and minimizing the deployment costs. This is a hard-to-solve optimization problem on city-scaled areas, as the number of possible solutions is very large [4]. Heuristics and metaheuristics are promising methods to deal with the RSU-DP because they allow computing *good* infrastructure designs in reduced

J. Toutouh is supported by Grant AP2010-3108 of the Spanish Ministry of Education. University of Malaga, International Campus of Excellence Andalucía Tech.

execution times [1, 5]. Evolutionary algorithms (EAs) have emerged to successfully deal with complex optimization problems. In this study, we propose using the NSGA-II evolutionary algorithm [2] to optimally design the RSU infrastructure within a city-scaled road network in Málaga (Spain). In order to obtain realistic results, we consider real information about road traffic (traffic flows and road map) and hardware (network capabilities and costs).

Our main contributions are: *i*) introducing a fully multiobjective evolutionary approach to solve the RSU-DP; *ii*) considering realistic scenarios, larger than those solved in the related literature and accounting for real traffic data; *iii*) reporting accurate results for cost and QoS for the problem instances considered.

The article is organized as follows. Section 2 introduces the multiobjective version of the RSU-DP and reviews related work on the topic. Section 3 introduces evolutionary computation and the proposed MOEA to solve the problem. Section 4 reports the experimental evaluation, including a comparison against two intuitive greedy heuristics to solve the problem. Finally, Section 5 formulates the conclusions and the main lines for future work.

## 2 The RSU Deployment Problem

The RSU-DP considers the following elements:

- A set of *road segments* $S = \{s_1, \ldots, s_n\}$ for placing RSUs along the streets. Each segment $s_i$ is defined by a pair of points $p_j, p_k \in P = \{p_1, \ldots, p_m\}$. Each point $p_j$ is identified by its geographical coordinates (latitude, longitude). RSUs can be placed at any location within each segment $s_i$.
- An estimation of the number of vehicles per time period across each segment $s_i$, $VN(s_i)$, and the average vehicle speed for each segment $sp(s_i)$.
- A set of RSU types $T = \{t_1, \ldots, t_k\}$, each one with a given gain and transmission power that determines the covering area and the cost of the RSU.

The multiobjective version of the problem proposes to find a set of locations and the type of each RSU to deploy in each location, with the goal of maximizing the *number of vehicles* served by the RSU infrastructure (considering the coverage, number of vehicles, and speed per each road segment), while simultaneously minimizing the *total cost* of deployment.

*Related works.* Some works have applied EAs to RSU-DP variants. An early work studied applying a genetic algorithm (GA) that uses a VANET simulator to evaluate the QoS of computed solutions for a given area of Brunswick, Germany [3], considering 100 possible predefined locations for RSUs. The results show that a *good* cost/utility trade-off is obtained using between 10 and 30 RSUs. Cavalcante et al. [1] compared a GA against a greedy approach to solve the maximum coverage with time threshold problem using data form four Swiss regions. The GA obtained better vehicle coverage than the greedy approach.

In the review of related works, we did not find articles using explicit multiobjective methods for this specific problem. Furthermore, the use of MOEAs to solve the problem has not been studied. Thus, there is room to contribute in this line of work by studying efficent and accurate MOEAs to solve the RSU-DP.

## 3 A Multiobjective Evolutionary Algorithm for RSU-DP

This section presents the details of the proposed MOEA for RSU placement.

*Solution encoding.* Solutions are represented as real arrays of length $n = \#S$. Each position on the array holds the RSU information: the type is given by the integer part of the real number (0 stands for no RSU, and integers $1 \dots k$ represent types $t_1 \dots t_k$); position within the segment is given by the fractional part of the real number, mapping the interval $[0, 1)$ to points in the segment $[p_j, p_i)$. For instance, a value of 3.5 in position 5 of the tuple, means that in segment number 6 a RSU of type 3 is placed at the middle of the segment.

*Evolutionary operators.* Population is randomly initialized, using reals from the interval $[0,\ k + r]$ being $k$ the number of RSU types in $T$, and $r \in [0, 1)$. Given that one of the extremes of the ideal Pareto front is known (the solution that places no RSU has cost 0), we add that solution to the initial population. Future work includes adding solutions computed by greedy algorithms to the initial population as well. The crossover operator is *Intermediate Recombination*; offspring of parents $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfy $\alpha_i x_i + (1 - \alpha_i) y_i$ and $\beta_i y_i + (1 - \beta_i) x_i$ with $\alpha_i, \beta_i$ randomly chosen in $[-p, 1 + p]$ for a given $p$. An ad-hoc mutation operator was designed to provide diversity to the search: with probability $\pi_A$ we remove the RSU (if any) from the segment, with probability $\pi_B$ we change the type of the RSU (if any) to a random type picked uniformly in $T$, and with probability $1 - \pi_A - \pi_B$ we apply a Gaussian Mutation with a standard deviation of $\sigma$.

*Computing the objective functions.* Computing the total cost is straightforward, by adding the cost (according to the type) of each RSU placed in the scenario. For computing the QoS, we consider the distances and values in Figure 2: the RSU placed in the point "×" covers the subsegments $c_1$ (in $s_1$), $c_2$ (in $s_2$), in street A, and $c_3$ (in $s_3$), and $c_4$ (in $s_4$) in street B. The number of effective vehicles attended is computed by $\sum_{i=1}^{i=4} NV(s_i) \times \frac{c_i}{sp(s_i)}$. This requires computing the intersections between the road segments and the circle representing the coverage of the RSU. Given that the distances involved in the problem are relatively small, we use straight lines in the latitude-longitude space as an estimation, with negligible error. This approximation makes computation faster, thus improving the overall performance of the algorithm. Since the distance of a degree of longitude depends on the latitude, it is necessary to adjust for that by multiplying the longitude by the cosine of the latitude.

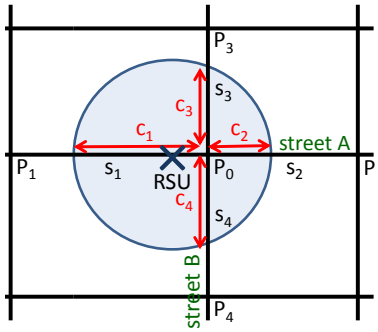



**Fig. 2.** Calculation of the vehicles attended by a RSU.

*Parametric configuration.* We performed an analysis to find the best values for NSGA-II parameters. In the parameter setting experiments, the best results were obtained using the configuration: *population size*=72, *crossover probability*=0.95, *mutation probability*=0.01, $\pi_A$=0.5, $\pi_B$=0.25, and $\sigma$=0.25.

## 4 Experimental Analysis

*Problem instances.* We defined a real world problem instance based on a real map of Málaga, real road traffic data, and real RSU network interfaces/antennas.

The map covers an area of 42.557 $km^2$, including a number of 106 points, defining 121 segments with lengths between 55 and 1556 m. The RSUs hardware are equipped with a IEEE 802.11p network interface, connected to an external antenna to improve the communication range according to a given antenna gain. Three types of IEEE 802.11p antennas are considered, according to three commercial omni-directional antennas from *Cetacea Wireless shop* (https://shop.cetacea.com/, see Table 1).

**Table 1.** General information about the antennas used to define different RSUs.

| *type* | *commercial model* | *gain* | *ERR* | *cost* |
|---|---|---|---|---|
| $t_1$ | Echo Series Omni Site Antenna | 6 dBi | 243.12 m | 121.70 $ |
| $t_2$ | Echo Series Omni Site Antenna | 9 dBi | 338.70 m | 139.20 $ |
| $t_3$ | Echo Series Omni Site Antenna | 12 dBi | 503.93 m | 227.50 $ |

In order to define the *effective radio range* (ERR) of each RSU, we evaluated the average percentage of data packets delivered correctly (*packet delivery ratio*, PDR), at different distances (from 0 to 650 m) for each RSU. Finally, to ensure a competitive QoS, we defined the ERR of each RSU as the distance at which the average PDR is equal or higher than 66.667%.

*Comparison against two greedy strategies.* We compare the results achieved by the proposed MOEA against two greedy heuristics. For the QoS objective, the greedy strategy places RSUs sequentially over non-covered segments starting with those with the higher ratio between number of vehicles and average speed. A segment is considered covered if it has a portion of $\lambda$ inside the coverage area of any RSU. The greedy strategy for cost is analogous, but stops when the QoS of the solution is equal to $\alpha \cdot Q$ where $Q$ is the best QoS value achieved by the greedy algorithm for QoS using $\lambda = 0.75$ and $\alpha \in [0, 1]$. For the experimental analysis the greedy algorithm for QoS was executed using $\lambda \in \{0.90, 0.95, 1.0\}$ and the greedy algorithm for cost was executed using $\alpha \in \{0.70, 0.75, 0.80\}$.

*Execution environment.* The experimental analysis was performed using 24 cores on an AMD Opteron 6172 2.10 GHz with 24 GB RAM at Cluster Fing: the high performance computing facility at Universidad de la República. Since computing the fitness of an individual is highly CPU-intensive, the evaluation of the population is done in parallel using 24 threads, thus each thread evaluates 3 individuals of the population. For each probem instance, we performed 20 independent runs of the MOEA and of both greedy algorithms.

*Numerical results.* In the experimental analysis the proposed MOEA has shown a good solving capability. NSGA-II significantly outperforms the two greedy heuristics while computing accurate Pareto fronts. Table 2 reports the best improvement of the proposed MOEA over the greedy strategies. NSGA-II is able to improve the QoS achieved by the greedy heuristic for cost in up to 73.9%

while keeping the same cost and improve up to 365.1% the cost achieved by the greedy heuristic for QoS while keeping the same QoS. Table 3 shows average, standard deviation and best results for standard multiobjective optimization metrics, where the small generational distance and spread values suggest both good convergence to an hypothetical ideal pareto front as well as good distribution among the non-dominated solutions. Finally, Figs. 3–4 show the global Pareto fronts achieved by the MOEA against the best results obtained by the greedy heuristics combining all 20 executions on normal and low traffic scenarios.

**Table 2.** Improvements over greedy heuristics.

| | | *instance* | | |
|---|---|---|---|---|
| | | *normal* | *low* | *high* |
| *improvement in cost* | RG_cost_70 | 70.2% | 67.7% | 68.6% |
| | RG_cost_75 | 67.1% | 68.1% | 66.5% |
| | RG_cost_80 | 73.9% | 69.7% | 70.7% |
| *improvement in QoS* | RG_QoS_90 | 299.3% | 307.5% | 333.4% |
| | RG_QoS_95 | 337.9% | 277.2% | 344.7% |
| | RG_QoS_100 | 365.1% | 331.4% | 341.8% |

**Table 3.** Multiobjective optimization metrics.

| | *normal* | *low* | *high* |
|---|---|---|---|
| generational distance | 2.7±0.3 (2.0) | 2.7±0.4 (1.5) | 2.7±0.3 (2.0) |
| spacing | 792.5±88.1 (615.5) | 739.8±68.9 (587.4) | 897.7±102.0 (719.1) |
| spread | $0.4\pm4.9\times10^{-2}$ (0.4) | $0.4\pm4.3\times10^{-2}$ (0.3) | $0.4\pm4.1\times10^{-2}$ (0.3) |
| relative hypervolume | $1.0\pm1.7\times10^{-2}$ (1.0) | $0.9\pm1.9\times10^{-2}$ (1.0) | $0.9\pm1.9\times10^{-2}$ (1.0) |

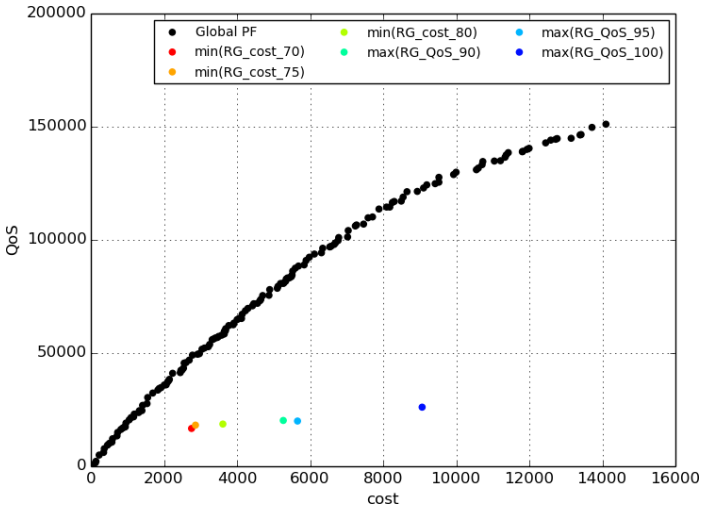



**Fig. 3.** Global Pareto front and heuristics results (normal traffic instance)

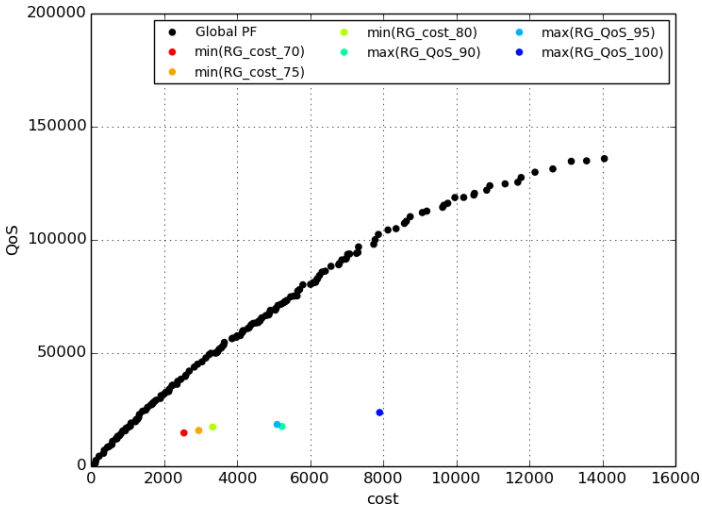



**Fig. 4.** Global Pareto front and heuristics results (low traffic instance)

## 5 Conclusions and Future Work

This article reports the advances on applying a multiobjective evolutionary approach to the problem of locating roadside infrastructure for vehicular networks over realistic urban areas. In the experiments performed, the proposed MOEA has shown good problem solving capabilities, computing accurate Pareto fronts and significantly improving over two greedy heuristics for the problem: up to 73.9% in cost and 365.1% in QoS. We are working now on extending the experimental analysis to other areas and considering additional information, such as accidents, in order to model a more realistic scenario for the problem.

## References

1. E. Cavalcante, A. Aquino, G. Pappa, and A. Loureiro. Roadside unit deployment for information dissemination in a VANET: An evolutionary approach. In *$14^{th}$ Genetic and Evolutionary Computation Conference*, pages 27–34, 2012.
2. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolut. Comput.*, 6(2):182–197, 2002.
3. C. Lochert, B. Scheuermann, C. Wewetzer, A. Luebke, and M. Mauve. Data aggregation and roadside unit placement for a vanet traffic information system. In *Proc. of $5^{th}$ ACM Int. Workshop on Vehicular Inter-Networking*, pages 58–65, 2008.
4. A. Reis, S. Sargento, F. Neves, and O. Tonguz. Deploying roadside units in sparse vehicular networks: What really works and what does not. *IEEE Trans. on Vehicular Technology*, 63(6):2794–2806, 2014.
5. C. Wang, X. Li, F. Li, and H. Lu. A mobility clustering-based roadside units deployment for vanet. In *$16^{th}$ Asia-Pacific Network Operations and Management Symposium*, pages 1–6, 2014.

# Different notions for fuzzy optima. Optimality conditions associated.

R. Osuna-Gómez[1], B. Hernández-Jiménez[2], Y. Chalco-Cano[3], G. Ruiz-Garzón[4]

[1] Departamento I.O., Universidad de Sevilla, Spain
[2] Universidad Pablo Olavide, Spain
[3] Instituto de Alta Investigación, Universidad de Tarapacá, Casilla 7D, Arica Chile
[4] Departamento de Estadística e I.O., Universidad de Cádiz, Spain

**Abstract.** Fuzzy numbers have been applied on decision and optimization problems in uncertain or vagueness environments. In these problems, the necessity of defining optima notions for decision-maker's preferences as well as prove necessary and sufficient optimality conditions for these optima are essential steps in order to deal with fuzzy optimization problems.

**Keywords:** Fuzzy numbers; crisp order relation; interval order relation; differentiable fuzzy mappings; stationary fuzzy point; fuzzy optimization.

## 1 Introduction

In conventional mathematical programming, the coefficients of problems are assumed to be deterministic and fixed in value. But there are many situations where this assumption is not valid because of uncertain environments. The fuzzy set theory, and particularly the concept of fuzzy number, provides an appropriate theoretical framework to model quantities that are imprecise because their own nature or some faults in measurement.

Fuzzy numbers have been applied on decision and optimization problems. In these problems, the necessity of procedures to rank fuzzy numbers is obvious. Ranking fuzzy numbers is a complex problem. All the proposed methods can be classified as corresponding to two different approaches:

1. Ranking fuzzy numbers using crisp relations (see [1], [2]). These are the procedures based on a ranking function and they provide a crisp total order relation between fuzzy numbers.
2. Using ordering relations between compact intervals in $\mathbb{R}$, using the fuzzy numbers characterization by their level sets (see [3], [4]).

Since results of comparison in real problems affect implicated individuals, their subjectivity should be reflected in the method for ranking.

We present the optimum definitions for fuzzy functions using the different ordering relations defined; and we present and relate the optimality conditions for the different optimum definitions given.

A fuzzy set on $\mathbb{R}^n$ is a mapping $\mu : \mathbb{R}^n \to [0,1]$. For each fuzzy set $\mu$, we denote $[\mu]^\alpha = \{x \in \mathbb{R}^n | \mu(x) \geq \alpha\}$ for any $\alpha \in (0,1]$ its $\alpha$-level set. By $supp\,\mu$ we denote the support of $\mu$, i.e. $\{x \in \mathbb{R}^n |\, \mu(x) > 0\}$. By $[\mu]^0$ we define the closure of $supp\,\mu$.

**Definition 1.** *A compact and convex fuzzy set $\mu$ on $\mathbb{R}^n$ is a fuzzy set with the following properties:*

1. *$\mu$ is normal, i.e. there exists $x_0 \in \mathbb{R}^n$ such that $\mu(x_0) = 1$;*
2. *$\mu(\lambda x + (1-\lambda)y) \geq min\{\mu(x), \mu(y)\}$, $x, y \in \mathbb{R}^n$, $\lambda \in [0,1]$;*
3. *$\mu$ is upper semicontinuous, i.e., $\{x : \mu(x) \geq \alpha\}$, is a closed set for all $\alpha \in [0,1]$*
4. *the closure of the set $\{x : \mu(x) > 0\}$ is compact.*

Let $\mathcal{F}_C$ denotes the family of all compact and convex fuzzy sets on $\mathbb{R}$, called fuzzy numbers or fuzzy intervals. By definition the $\alpha$-level sets of a fuzzy number are closed real intervals

$$u \in \mathcal{F}_C \Rightarrow [u]^\alpha = [\underline{u}_\alpha, \bar{u}_\alpha]$$

**Theorem 1.** *[5, 6] A fuzzy interval is completely determined by any pair $u = (\underline{u}, \overline{u})$ of functions $\underline{u}, \overline{u} : [0,1] \to \mathbb{R}$, defining the endpoints of the $\alpha$-level sets, satisfying the following three conditions:*
• *$\underline{u}(\alpha) = \underline{u}_\alpha \in \mathbb{R}$ is a bounded nondecreasing left-continuous function in $(0,1]$ and it is right-continuous at $0$;*
• *$\overline{u}(\alpha) = \overline{u}_\alpha \in \mathbb{R}$ is a bounded nonincreasing left-continuous function in $(0,1]$ and it is right-continuous at $0$;*
• *$\underline{u}(\alpha) \leq \overline{u}(\alpha)$, for all $\alpha \in [0,1]$.*

We denote by $\mathcal{F}_C^C$ the family of all level-continuous fuzzy intervals. Thus $u \in \mathcal{F}_C^C$ if the application $\alpha \mapsto [u]^\alpha$ is continuous.

**Proposition 1.** *Let $u = (\underline{u}_\alpha, \overline{u}_\alpha) \in \mathcal{F}_C^C$ be a fuzzy interval. Then, $u \in \mathcal{F}_C^C$ if and only if $\underline{u}_\alpha$ and $\overline{u}_\alpha$ are continuous functions of $\alpha$.*

## 2 Ordering relations based on average functions

In [2] a ranking function is defined to compare fuzzy numbers. This function was called "average index" because it can be interpreted as a weighted average in the following way: first, the decision-maker chooses a subset $Y$ of the unit interval, so that the associated level sets contain the information which is considered outstanding about the imprecise quantity. Next, he assigns a weight, represented by a probability distribution $P$, to the different elements or measurable subsets of $Y$. Also, the decision-maker determines a position function, $f_u(\alpha)$ giving to each associated level sets a real number. Finally, the index is defined as an average of positions of level sets in $Y$ using $P$.

**Definition 2.** *Let $u \in \mathcal{F}_C$ be a fuzzy number. Let $f_u(\alpha) : [0,1] \to \mathbb{R}$ such that $f_u \in \mathcal{R}(P)$. The real number*

$$V_P(u) = \int_Y f_u(\alpha) dP(\alpha)$$

*is called average index of $u$.*

By means of $V_P(\cdot)$ a comparison relation on $\mathcal{F}_C$ is built:

**Definition 3.** $\forall u, v \in \mathcal{F}_C$

- $u \leq_V v \Leftrightarrow V_P(u) \leq V_P(v)$
- $u <_V v$ *if* $u \leq_V v$ *and* $V_P(u) \neq V_P(v)$

In [2] the authors propose to choose one point included in each level set of $u$ as value for $f_u$:

$$f_u^\lambda : Y \to \mathbb{R}$$

$$f_u^\lambda(\alpha) = \lambda \overline{u}(\alpha) + (1-\lambda)\underline{u}(\alpha)$$

where $\lambda \in [0,1]$ is an optimism-pessimism degree, which must be selected by the decision maker: when the most advantageous decision is to choose the greatest quantity, an optimistic person would think of the upper extreme of the interval $\overline{u}$ ($\lambda = 1$), which reflects the greatest profit. On the contrary, a pessimistic person would prefer the lower extreme of the interval $\underline{u}$ ($\lambda = 0$), which represents the least he can win.

When the most advantageous decision is to choose the least quantity, the interpretation is the opposite, with $\lambda = 0$ for the optimism and $\lambda = 1$ for pessimism. Thus, if the optimism-pessimism degree of the decision-maker is $\mu \in [0,1]$, the parameter $\lambda$ for the function $f_u^\lambda$ is

$$\lambda = \begin{cases} \mu & \text{if the "best" is the "greatest"} \\ 1-\mu & \text{if the "best" is the "least"} \end{cases}$$

Between the two extreme values $\lambda = 0$ and $\lambda = 1$ there is an attitudes scale for the uncertainty for each decision-maker.

## 3 Ordering relations based on intervals

In [3] the authors propose order relations between alternatives which represent the decision-maker's preference when the cost of each alternative is known only to lie in an interval.

**Definition 4.** *Let $A = [\underline{a}, \overline{a}], B = [\underline{b}, \overline{b}]$ be two closed intervals in $\mathbb{R}$. The center and the width of an interval may be calculated as $A_C = (\overline{a} + \underline{a})/2$, $A_W = \overline{a} - \underline{a}$. Let us define order relations $\preceq_{LW}$, $\preceq_{LR}$ and $\preceq_{CW}$*

1. – $A \preceq_{LW} B \Rightarrow \underline{a} \leq \underline{b}$ *and* $A_W \leq B_W$.
   – $A \preceq_{LW} B \Rightarrow A \preceq_{LW} B$ *and* $A \neq B$, *i.e.,* $\underline{a} \leq \underline{b}$ *and* $A_W \leq B_W$, *with some strict inequality.*
   – $A \prec_{LW} B \Rightarrow \underline{a} < \underline{b}$ *and* $A_W < B_W$.
2. – $A \preceq_{LR} B \Rightarrow \underline{a} \leq \underline{b}$ *and* $\bar{a} \leq \bar{b}$.
   – $A \preceq_{LR} B \Rightarrow A \preceq_{LR} B$ *and* $A \neq B$, *i.e.,* $\underline{a} \leq \underline{b}$ *and* $\bar{a} \leq \bar{b}$, *with some strict inequality.*
   – $A \prec_{LR} B \Rightarrow \underline{a} < \underline{b}$ *and* $\bar{a} < \bar{b}$.
3. – $A \preceq_{CW} B \Rightarrow A_C \leq B_C$ *and* $A_W \leq B_W$.
   – $A \preceq_{CW} B \Rightarrow A \preceq_{CW} B$ *and* $A \neq B$, *i.e.,* $A_C \leq B_C$ *and* $A_W \leq B_W$, *with some strict inequality.*
   – $A \prec_{CW} B \Rightarrow A_C < B_C$ *and* $A_W < B_W$.

Basing on Theorem 1 we can order fuzzy numbers using the previous definition. We denote $\preceq_*$ any of the orders defined.

**Definition 5.** *For* $u, v \in \mathcal{F}_C$

1. $u \preceq_* v$ *if* $[u]^\alpha \preceq_* [v]^\alpha, \quad \forall \alpha \in [0,1]$
2. $u \preceq_* v$ *if* $[u]^\alpha \preceq_* [v]^\alpha, \forall \alpha \in [0,1]$ *and* $\exists \alpha_0 / [u]^{\alpha_0} \preceq_* [v]^{\alpha_0}$
3. $u \prec_* v$ *if* $[u]^\alpha \preceq_* [v]^\alpha \quad \forall \alpha \in [0,1]$

The order $\preceq_{LR}$ represents the decision-maker's preference for the alternative with lower minimum cost and maximum cost and the order $\preceq_{LW}$ represents the decision-maker's preference for the alternative with lower minimum cost and less uncertainty since the width of an interval can be regarded as an uncertainty risk or a type of variance and; $\preceq_{CW}$ represents the preference for the alternative with lower expected value and less uncertainty.

## 4 Fuzzy optimization. Minimum definitions

A mapping $F : K \subset \mathbb{R}^n \to \mathcal{F}_C$ is said to be a fuzzy mapping. Each $F(x)$ is a fuzzy interval, that is uniquely determined by two functions such that

$$[F(x)]^\alpha = [\underline{f}_\alpha(x), \bar{f}_\alpha(x)] = [\underline{f}(\alpha, x), \bar{f}(\alpha, x)] \quad \forall \alpha \quad \forall x \in K$$

Then for $F$, we define the family of interval-valued functions $F_\alpha : K \to \mathcal{K}_C$ given by $F_\alpha(x) = [F(x)]^\alpha$, for any $\alpha \in [0,1]$. Here, for each $\alpha \in [0,1]$, the endpoint functions $\underline{f}_\alpha, \bar{f}_\alpha : K \to \mathbb{R}$ are called lower and upper functions of $F$, respectively. Let us consider fuzzy functions $F : K \subseteq \mathbb{R} \to \mathcal{F}_C^C$ so that $V_P(F(x))$ is well defined.

Associated with the average index ordering relation and intervals ordering relation we give the following minimum definitions for fuzzy minima:

**Definition 6.** *It is said that* $\bar{x}$ *is a (strict) minimum*$^V$ *for* $F$ *if* $F(\bar{x}) \preceq_L F(x)$ $(F(\bar{x}) \prec_L F(x))$, $\forall x \in K$.

**Definition 7.** *It is said that $\bar{x}$ is a (strict, weak) minimum$^*$ for $F$ if there does not exist $x \in K$ such that $F(x) \preceq_* F(\bar{x})$ $(F(x) \underline{\preceq}_* F(\bar{x}), F(x) \prec_* F(\bar{x}))$.*

Next we establish the relations among the different minimum types:


## 5 Necessary fuzzy optimality conditions

In classical optimization methods, it is well-known that the stationary point concept (the one that cancels the derivative) plays a crucial role as a necessary optimality condition for problems defined by differentiable functions, since it allows to identify the potential candidates to be optimums. Different fuzzy stationary points definitions [8–11] can we find in the literature.

In this section we give necessary optimality conditions based on appropriate stationary point definitions for minimum concepts defined in the previous section and they attend to the decision-maker's attitude to the problem of choose the favourite alternative.

Let us consider hereafter fuzzy functions $F : S \subseteq \mathbb{R} \to \mathcal{F}_C^C$ where $S$ is an open set and consider they are level-wise differentiable, [4].

**Theorem 2.** *A fuzzy function $F : S \subseteq \mathbb{R} \to \mathcal{F}_C^C$ where $S$ is an open set is a level-wise differentiable fuzzy function if and only the endpoints functions associated are differentiable and $[F'(x)]^\alpha = [\min\{\underline{f}'_\alpha(x), \overline{f}'_\alpha(x)\}, \max\{\underline{f}'_\alpha(x), \overline{f}'_\alpha(x)\}]$.*

*Remark 1.* [7] If $\lambda = \frac{1}{2}$ we can relax the hypotheses considering gH-differentiable functions.

In order to guarantee that $\nabla V_P^\lambda(F(\cdot))$ exists we suppose that $\nabla f_F^\lambda(\cdot)$ is continuous with respect to $\alpha$, and together with $P$ monotonicity we can ensure the existence of the integral.

**Definition 8.** *It is said that $\bar{x} \in S$ is a* **stationary point**$^V$ *for $F$ if $\nabla V_P^\lambda(F(\bar{x})) = 0$ for some $\lambda$ and $P$.*

**Theorem 3.** *If $\bar{x}$ is a local minimum$^V$ then it is a stationary point$^V$.*

**Definition 9.** *It is said that $\bar{x} \in S$ is a* **stationary point**$^{LR}$ *for $F$ if $0 \in [F'(\bar{x})]^0$.*

**Theorem 4.** *If $\bar{x}$ is a weak minimum$^{LR}$ then it is a stationary point$^{LR}$.*

Now we give the relations between the stationary point definitions:

**Theorem 5.** *$\bar{x}$ is a stationary point$^V$ for any $\lambda \in [0, 1]$ and $P$, then $\bar{x}$ is a stationary point$^{LR}$.*

## References

1. R.R. Yager, A procedure for ordering fuzzy subsets of the unit interval, Inform. Sci. 24 (1981) 143-161
2. L.M. Campos-Ibáñez, A. González-Muñoz, A subjective approach for ranking fuzzy numbers, Fuzzy Sets and Systems 29 (1989) 145-153
3. H. Ishibuchi, H. Tanaka, Multiobjective programming of the interval objective function, European Journal of Operational Research, 48 (1990) 219-225.
4. H.C. Wu, The Karush-Kuhn-Tucker optimality conditions in optimization problem with interval-valued function, European Journal of Operational Research 176 (1) (2007) 46-59.
5. R. Goestschel, W. Voxman, Elementary fuzzy calculus, Fuzzy Sets and Systems 18 (1986) 31-43.
6. L. Stefanini, B. Bede, Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, Nonlinear Analysis 71 (2009) 1311-1328.
7. Y. Chalco-Cano, A. Rufián-Lizana, R. Osuna-Gómez, H. Román-Flores, A note on generalized convexity for fuzzy mappings through a linear ordering. Fuzzy Sets and Systems, 231 (2013) 70-83.
8. M. Panigrahi, G. Panda, S. Nanda, Convex fuzzy mapping with differentiability and its application in fuzzy optimization, European Journal of Operational Research 185 (2008) 47-62.
9. H.C. Wu, The Karush-Kuhn-Tucker optimality conditions for the optimization problem with fuzzy-valued objective function, Math. meth. Oper. Res. 66 (2007) 203-224.
10. H.C. Wu, The Karush-Kuhn-Tucker optimality conditions for multi-objective programming problems with fuzzy-valued objective functions, Fuzzy Optimization and Decision Making, 8 (2009) 1-28.
11. H.C. Wu, The optimality conditions for optimization problems with convex constraints and multiple fuzzy-valued objective functions, Fuzzy Optimization and Decision Making, 8 (2009) 295-321.

# Linear ordering on space of fuzzy intervals

Manuel Arana-Jiménez[1] and Aris Burgos-Pintos[2]

[1] Department of Statistics and Operational Research, University of Cádiz, Spain.
manuel.arana@uca.es
[2] Department of Production, AtSistemas, Spain. arisbp@gmail.com

**Abstract.** In this article, we present the results given by [5], in which it is studied generalized convexity for fuzzy mappings defined through a linear ordering on the space of fuzzy intervals. It is shown that invexity is the minimal property on the objective function so that a stationary point is an optimal solution. Furthermore, and based on [1], we provide an extension of this optimization result given in scalar case to multiobjective programming on fuzzy sets.

*Keywords:* Fuzzy mathematical programming; generalized Hukuhara differentiability; generalized convexitity.

## 1 Introduction

The purpose of this article is to present results on the study of optimization on the space of fuzzy intervals, focused on finding a linear ordering which allows us to characterize the solutions. We start with some preliminaries on fuzzy intervals, following with the definition of the order relation for a ranking value function $\tau$. Next we have gathered the results on differentiability of fuzzy mappings, the generalized differentiability, convexity and optimality conditions in fuzzy optimization. Finally, we present the recent results on extending the ranking function to multiobjective problem.

## 2 Preliminaries

Let $\mathcal{K}_C$ denote the family of all bounded closed intervals in $\mathbb{R}$, i.e.,

$$\mathcal{K}_C = \left\{ \left[\underline{a}, \overline{a}\right] \mid \underline{a}, \overline{a} \in \mathbb{R} \text{ and } \underline{a} \leq \overline{a} \right\},$$

Given two intervals $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$, we define the distance between $A$ and $B$ by

$$H(A, B) = \max\left\{ \left|\underline{a} - \underline{b}\right|, \left|\overline{a} - \overline{b}\right| \right\}.$$

It is well known that $(\mathcal{K}_C, H)$ is a complete metric space [2].

A fuzzy set on $\mathbb{R}^n$ is a mapping $u : \mathbb{R}^n \to [0, 1]$. For each fuzzy set $u$, we denote its $\alpha$-level set as $[u]^\alpha = \{x \in \mathbb{R}^n \mid u(x) \geq \alpha\}$ for any $\alpha \in (0, 1]$. The support of $u$ is denoted by $supp(u)$ where $supp(u) = \{x \in \mathbb{R}^n \mid u(x) > 0\}$. The closure of $supp(u)$ defines the 0-level of $u$, .i.e. $[u]^0 = cl(supp(u))$ where $cl(M)$ means the closure of the subset $M \subset \mathbb{R}^n$.

Let $\mathcal{F}_C$ denote the family of all fuzzy intervals. So, for any $u \in \mathcal{F}_C$ we have that $[u]^\alpha \in \mathcal{K}_C$ is a nonempty compact and convex subset on $\mathbb{R}$ for all $\alpha \in [0,1]$. Thus the $\alpha$-levels of a fuzzy interval are given by $[u]^\alpha = [\underline{u}_\alpha, \overline{u}_\alpha]$, $\underline{u}_\alpha, \overline{u}_\alpha \in \mathbb{R}$ for all $\alpha \in [0,1]$.

For fuzzy intervals $u, v \in \mathcal{F}_C$ represented by $[\underline{u}_\alpha, \overline{u}_\alpha]$ and $[\underline{v}_\alpha, \overline{v}_\alpha]$, respectively, and for any real number $\lambda$, addition $u + v$ and scalar multiplication $\lambda u$ are defined in [5].

Given $u, v \in \mathcal{F}_C$ a distance between $u$ and $v$ is defined in [5], such that $(\mathcal{F}_C, D)$ is a complete metric space.

We denote by $\mathcal{F}_C^C$ the family of all level-continuous fuzzy intervals [13]. It is well known that $(\mathcal{F}_C^C, D)$ is a separable and complete metric space [13]. Moreover, $\mathcal{F}_C^C$ is a closed subspace of $\mathcal{F}_C$.

**Definition 1.** *Let $u = (\underline{u}, \overline{u})$ be a fuzzy interval. We say that $u$ is a nonnegative fuzzy interval (nonpositive fuzzy interval, respectively) if $\underline{u}(0) \geq 0$ ($\overline{u}(0) \leq 0$, repectively).*

A crucial concept to obtain an useful definition of derivative for fuzzy functions is the difference between two fuzzy intervals. Toward this end we have the following definition[5].

**Definition 2.** *([18]) Given two fuzzy intervals $u$, $v$, the generalized Hukuhara difference (gH-difference for short) is the fuzzy interval $w$, if it exists, such that*

$$u \ominus_{gH} v = w \Leftrightarrow \begin{cases} (i)\ u = v + w, \\ or\ \ (ii)\ v = u + (-1)w. \end{cases}$$

It is easy to show that $(i)$ and $(ii)$ are both valid if and only if $w$ is a crisp number. Note that the case (i) is coincident to Hukuhara difference (see [11]) and so the concept of $gH$-difference is more general than $H$-difference.

If $u \ominus_{gH} v$ exists then, in terms of $\alpha$-levels, we have

$$[u \ominus_{gH} v]^\alpha = [u]^\alpha \ominus_{gH} [v]^\alpha = \left[\min\{\underline{u}_\alpha - \underline{v}_\alpha, \overline{u}_\alpha - \overline{v}_\alpha\}, \max\{\underline{u}_\alpha - \underline{v}_\alpha, \overline{u}_\alpha - \overline{v}_\alpha\}\right],$$

for all $\alpha \in [0,1]$, where $[u]^\alpha \ominus_{gH} [v]^\alpha$ denotes the gH-difference between two intervals (see [17, 18]).

## 3 A linear ordering on the space of fuzzy intervals

During the study of the different methods for ranking fuzzy intervals, most of the authors suggest mapping each fuzzy interval into the real line to define a ranking function (see for instance [6, 7, 19]). In [6] a ranking function called the Average Value (A.V.) was introduced. The A.V. was defined as dependent on several parameters, allowing flexibility in the final classification. The following definition of ranking function, introduced by Tsumura et al in [19], is a particular case of A.V. considering a mean optimism degree. For more details see [4, 6].

**Definition 3.** *Let $\tau : \mathcal{F}_C \to \mathbb{R}$ be a function defined by*

$$\tau(u) = \int_0^1 \alpha \left[\underline{u}(\alpha) + \overline{u}(\alpha)\right] d\alpha, \tag{1}$$

*for all fuzzy interval $u \in \mathcal{F}_C$. Then the function $\tau$ is called ranking value function.*

In this case $\tau(u)$ represents a mean value of the different $\alpha$-levels positions. In fact, the $\tau(u)$ represent a mean value of the fuzzy interval $u$. For more details see [4, 8, 10].

*Remark 1.* In the relation (1) we are considering the Lebesgue integral.

From ranking value function $\tau$, we consider the following order relation, $\preceq$, on $\mathcal{F}_C$ which was exhaustively studied by many authors [8, 9, 12, 15, 16, 20].

**Definition 4.** *Suppose that u and v are two fuzzy intervals. Then u precedes v ($u \preceq v$) if and only if $\tau(u) \leq \tau(v)$.*

*And u strictly precedes v ($u \prec v$) if [15, 16]*

$$u \preceq v, \quad \text{and } \tau(u) \neq \tau(v).$$

Note that the order relation $\preceq$ is reflexive and transitive. Moreover, any two elements of $\mathcal{F}_C$ are comparable under the ordering $\preceq$. For more details see [8, 9, 15, 16]

As consequence of Definition 3 and 4 we have the following.

**Lemma 1.** *For $u, v \in \mathcal{F}_C$,*
*(a) $u \preceq v \Leftrightarrow \tau(u) \leq \tau(v)$;*
*(b) $u \prec v \Leftrightarrow \tau(u) < \tau(v)$;*
*(c) if $u \preceq v$ and $v \preceq u$, then $\tau(u) = \tau(v)$.*

## 4 Differentiable fuzzy mappings

Henceforth, $K$ denotes an open subset of $\mathbb{R}^n$ an let $T = (a, b)$ be an open interval in $\mathbb{R}$.

A mapping $F : K \to \mathcal{F}_C$ is said to be a fuzzy mapping. For each $\alpha[0, 1]$, associated to $F$, we define the family of interval-valued functions $F_\alpha : K \to \mathcal{K}_C$ given by $F_\alpha(x) = [F(x)]^\alpha$. For any $\alpha[0, 1]$, we denote

$$F_\alpha(x) = [\underline{f_\alpha}(x), \overline{f^\alpha}(x)].$$

Here, for each $\alpha \in [0, 1]$, the endpoint functions $\underline{f_\alpha}, \overline{f^\alpha} : K \to \mathbb{R}$ are called lower and upper functions of $F$, respectively.

Next we present the concept of $gH$-differentiable fuzzy mappings in the one dimensional case.[1]

**Definition 5.** *([3]) Let $K \subset \mathbb{R}$ with $F : K \to \mathcal{F}_C$ a fuzzy function and $x_0 \in K$ and $h$ be such that $x_0 + h \in K$. Then the generalized Hukuhara derivative (gH-derivative, for short) of $F$ at $x_0$ is defined as*

$$F'(x_0) = \lim_{h \to 0} \frac{F(x_0 + h) \ominus_{gH} F(x_0)}{h}. \tag{2}$$

*If $F'(x_0) \in \mathcal{F}_C$ satisfying (2) exists, we say that $F$ is generalized Hukuhara differentiable (gH-differentiable, for short) at $x_0$.*

The $gH$-derivative for an interval-valued function [18] is similar to Definition 5. More precisely, an interval-valued function $F : K \to \mathcal{K}_C$ is $gH$-differentiable at $x_0 \in K$, with $gH$-derivative $F'(x_0) \in \mathcal{K}_C$, if (2) exists with respect to the limit in the metric space $(\mathcal{K}_C, H)$ where the difference is given by the $gH$-difference between intervals (see [18]).

**Theorem 1.** *Let $F : K \to \mathcal{F}_C$ be a fuzzy function. If $F$ is $gH$-differentiable then the interval-valued function $F_\alpha : K \to \mathcal{K}_C$ is $gH$-differentiable for each $\alpha \in [0, 1]$. Moreover*

$$[F'(x)]^\alpha = F'_\alpha(x). \tag{3}$$

*Proof.* The proof is a consequence of the definition of $gH$-differentiability. □

Motivated by Goetschel and Voxman [9], in [5] is given the following definition

**Definition 6.** *For each fuzzy mapping $F : K \to \mathcal{F}_C$, the ranking function $T_F : K \to \mathbb{R}$ associated to $F$ is defined by*

$$T_F(x) = \int_0^1 \alpha \left[\underline{f}(\alpha, x), \overline{f}(\alpha, x)\right]$$

Note that the real-valued function $T_F$ can be rewritten as being $T_F(x) = \tau(F(x))$.

The following results shows the connection between the continuity and $gH$-differentiability of a fuzzy mapping $F$ and the continuity and differentiability of the ranking function $T_F$, respectively.[5]

**Proposition 1.** *(Syau and Stanley Lee[15]). If $F : K \to \mathcal{F}_C$ is continuous, then $T_F : K \to \mathbb{R}$ is also continuous.*

Now we show the connection between the $gH$-differentiability of $F$ and differentiability of the ranking function $T_F$.

**Theorem 2.** *Let $K \subset \mathbb{R}^n$ be an open set. If $F : K \to \mathcal{F}_C^C$ is $gH$-differentiable and $\alpha \mapsto \left(\partial\left(\underline{f} + \overline{f}\right)/\partial x_i\right)(\alpha, x)$ is continuous, for each $i = 1, ..., n$ and $x \in K$, then $T_F : K \to \mathbb{R}$ is differentiable.*

## 5 Generalized convexity and optimality conditions in fuzzy optimization

In [15, 16, 20] were introduced some concepts of convexity and generalized convexity for fuzzy mappings based on the ranking valued function $\tau : \mathcal{F}_C \to \mathbb{R}$, introduced in the preceding Section. In what follows, let $\eta : K \times K \to \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be a nonempty invex set w.r.t. to $\eta$.

From the definitions of convex, preinvex and prequasiinvex of a fuzzy mapping $F : K \to \mathcal{F}_C$ given in [15, 16, 20], it is shown that the concept of convexity of $F$ is linked to the convexity of $T_F$. More precisely from [5], we have the following result.

**Theorem 3.** *Let $F : K \to \mathcal{F}_C$ be a fuzzy mapping. Then, $F$ is convex (preinvex, prequasiinvex) if and only if $T_F$ is convex (preinvex, prequasiinvex) respectively.*

Inspired by Theorem 3, in [5] it is proposed the following definition.

**Definition 7.** *Let $F : K \to \mathcal{F}_C$ be a fuzzy mapping such that $T_F$ is differentiable. Then, $F$ is said to be invex if $T_F$ is an invex function, i.e. for $x, y \in K$*

$$T_F(x) - T_F(y) \geq \nabla T_F(y)\eta(x, y)$$

In addition, Y. Chalco-Cano et al. [5] give the following definition of stationary point.

**Definition 8.** *We say that $x^* \in K$ is a stationary point for a G-differentiable fuzzy mapping $F : K \to \mathcal{F}_C$ if $\nabla T_F(x^*) = 0$.*

These concepts of generalized convexity given were used in [14–16, 20] to establish diverse results on fuzzy optimization. In [5] it is presented the following definition on fuzzy optimization.

**Theorem 4.** *Let $F : K \to \mathbb{R}$ be a fuzzy mapping. Then $F$ is invex if and only if every stationary point is a minimum point of $F$.*

The recently published article [1], shows us how to extend the ranking function to multiobjective problem. In fact they focus their attention on the following vector fuzzy optimization problem

$$\begin{array}{ll} (VFP) & \text{Minimize } F(x) = (F_1(x), \ldots, F_p(x)) \\ & \text{subject to: } x \in K \end{array}$$

where $K \subseteq \mathbb{R}^n$ is an nonempty open set, and $F : K \subseteq \mathbb{R}^n \to (\mathcal{F}_C)^p$ a vector fuzzy mapping. $K$ is said to be the feasible set.

In the particular case $p = 1$, we have described an order relation, $\preceq$, on $\mathcal{F}_C$, introduced in [9], and associated to a ranking value function $\tau$ by Definition 4 (see also [15, 16, 20]).

**Theorem 5.** *Let $F : K \to \mathbb{R}$ be a fuzzy mapping. Then, $F$ is invex if and only if every stationary point is a minimum point of $F$.*

As an extension of this order relation to vector fuzzy sets, Arana et al.[1] have proposed the partial order relations given by $\prec$, $\leqq$ and $\preceq$, on $(\mathcal{F}_C)^p$. These new relations allow them to define some concepts of optimal solutions, based on efficiency, as well as to study new necessary and sufficient optimality conditions for (VFP). Again, we have a necessary optimality conditions defined in terms of stationary point for the vectorial case as an extension of that given for the scalar case, when $p > 1$.

Arana et al. proved in [1] that pseudoconvexity-I is the minimal condition for a stationary point of $F$ to be a weakly efficient solution for the vector fuzzy mapping.

**Theorem 6.** *Every stationary point of $F$ is a weakly efficient solution of (VFP) if and only if $F$ is pseudoinvex-I.*

Finally, they established a similar result for efficient solutions for (VFP).

**Theorem 7.** *Every stationary point of $F$ is an efficient solution of (VFP) if and only if $F$ is pseudoinvex-II.*

Thus, we can state that pseudoinvexity-I and II are not only sufficient conditions, but the minimal, to obtain weakly efficient and efficient solutions from stationary points.

## 6 Conclusions

A proper linear ordering on the space of fuzzy intervals is very useful to obtain efficient and weakly efficient solutions of a fuzzy optimization problem. Y. Chalco-Cano et al.[5] present the ranking value function $\tau$ and the concepts of generalized convexity, which derive on results on fuzzy optimization based on invexity. Arana et al.[1] extend these results to multiobjective fuzzy problem with pseudoinvexity.

## References


1. M. Arana-Jiménez, A. Rufián-Lizana, Y. Chalco-Cano, H. Román-Flores, Generalized convexity in vector fuzzy optimization through a linear ordering, Information Sciences 312 (2015) 13–24.
2. Aubin J.P. & Cellina A, Differential Inclusions, New York: Springer-Verleg (1984).
3. B. Bede, L. Stefanini, Generalized differentiability of fuzzy-valued functions, Fuzzy Sets and Systems, 230 (2013) 119–141.
4. L. Campos and A. Gonzalez, Further contributions to the study of average value for ranking fuzzy numbers, Int. J. Approximate Reasoning 10 (1994) 135–153.
5. Y. Chalco-Cano, A. Rufián-Lizana, H. Román-Flores, R. Osuna-Gómez, A note on generalized convexity for fuzzy mappings through a linear ordering, Fuzzy Sets and Systems 231 (2013) 70–83.
6. L. De Campos Ibañez and M. Gonzalez-Muñoz, A subjective approach for ranking fuzzy numbers, Fuzzy Sets and Systems 29 (1989) 145–154.
7. D. Dubois, H. Prade, Ranking of fuzzy numbers in the setting of possibility theory, Inform. Sci. 30 (1980) 183–224.
8. G. Facchinetti, S. Giove and N. Pacciarotti, Optimisation of a non linear fuzzy function, Soft Computing 6 (2002) 476–480.
9. R. Goestschel, W. Voxman, Elementary fuzzy calculus, Fuzzy Sets and Systems 18 (1986) 31–43.
10. A. González, A study of the ranking function approach through mean values, Fuzzy Sets and Systems 35 (1990) 29–43.
11. M. Hukuhara, Integration des applications mesurables dont la valeur est un compact convexe, Funkcialaj Ekvacioj 10 (1967) 205–223.
12. J. Li, M. A. Noor, On properties of convex fuzzy mappings, Fuzzy Sets and Systems 219 (2013) 113–125.
13. H. Román-Flores and M. Rojas-Medar, Embedding of level-continuous fuzzy sets on Banach spaces, Information Sciences 144 (2002) 227–247.
14. Y.R. Syau, Invex and generalized convex fuzzy mappings, Fuzzy Sets and Systems 115 (2000) 455–461.
15. Y.R. Syau, E. Stanley Lee, Fuzzy Weirstrass theorem and convex fuzzy mappings, Computers and Mathematics with Applications 51 (2006) 1741–1750.
16. Y.R. Syau,, E. Stanley Lee, Preinvexity and $\Phi_1$-convexity of fuzzy mappings through a linear ordering, Computers and Mathematics with Applications 51 (2006) 405–418.
17. L. Stefanini, B. Bede, Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, Nonlinear Analysis 71 (2009) 1311–1328.
18. L. Stefanini, *A generalization of Hukuhara difference and division for interval and fuzzy arithmetic*, Fuzzy sets and systems, 161 (2010), 1564–1584.
19. Y. TSumura, T.Terano and M. Sugeno, Fuzzy fault tree analysis, Summary of papers on general fuzzy problems, report n$^o$ 7, (1981) 21–25.
20. H. Yan, J. Xu, A class convex fuzzy mappings, Fuzzy Sets and Systems 129 (2002) 47–56.

# DEA production games with fuzzy output prices

M.A. Hinojosa[1]*, S. Lozano[2], A.M. Mármol[3]

[1] *Dpto. de Economía, Métodos Cuantitativos e Historia Económica, Universidad Pablo de Olavide.*

[2] *Dpto. de Organización Industrial y Gestión de Empresas I, Universidad de Sevilla.*

[3] *Dpto. de Economía Aplicada III, Universidad de Sevilla.*

## Extended abstract

Owen[12] considered linear production programming problems in which multiple decision-makers pool resources to make several products and analyzed these situations by using cooperative game theory. Recently, Lozano[7] has generalized the model to the more general framework of DEA production problems.

In these models the technology is assumed to be implicit in the input-output data given by a set of recorded observations. The objective function represents the total revenue obtained from selling certain kinds of products, and the problem is formulated as a linear programming problem in which the revenue is maximized in the production possibility set induced by the set of recorded observations.

However, very often, in real world-situations, the assumption of certainty with respect to the nature of the parameters is unrealistic and in many applications, the use of fuzzy logic[14] has proved to be advantageous to deal with the imprecise nature of the data involved. Particularly, in the analysis of efficiency by using DEA models, imprecision in the data is a main drawback and their representation as fuzzy numbers enables a more realistic assessment of the efficiency of the decision making units (see for instance, Lertworasirikul et al.[6], Hatami-Marbini et al.[4], and Lozano[9].

*Corresponding author. Ctra. de Utrera Km.1, 41013, Sevilla Spain. email: mahinram@upo.es.

In this paper we address the cooperative model arising from DEA production problems with uncertain parameters. The introduction of uncertainty into the cooperative model raises new and interesting issues, since coalitions can form prior to the resolution of uncertainty and they must discuss divisions of the uncertain revenue by taking into account their potential worths which may also be uncertain. We assume that the lack of precision in the parameters of the linear production problem is modeled via fuzzy logic, that is, some of the parameters involved in the objective function and/or in the constraints of the production game are represented by fuzzy numbers.

Several cooperative models involving fuzzyness can be found in the literature. The present investigation deals with models in which the fuzziness concerns the values that the coalitions can achieve. Niszhizaki and Sakawa[11] precede us in investigating solutions for these games. They addressed the special case of fuzzy cooperative games arising from linear production programming problems with fuzzy parameters for which they proposed an infinite family of cores, each of which consists of a set of non-fuzzy payoff vectors. Recently, in Hinojosa et al.[5] and in Monroy et al.[10], a different approach has been presented to analyse the solutions of cooperative games with fuzzy payoffs and applied to the cases of fuzzy linear production games and fuzzy assignment games.

On the other hand, Lozano et al.[8] have investigated vector-valued DEA production games and their results serve as a basis for the analysis in a fuzzy environment which is developed in this paper.

As a first step to analyze the problem in a fuzzy environment, a partial order has to be considered in the set of fuzzy numbers. Hence, the concept of maximization of fuzzy objective functions on a feasible set must be understood as the search for the maximal elements with respect to this partial order. As a consequence, the game arising from the production situation, when the pool of resources is controlled by several agents, is a set-valued game in which each element of the set is a fuzzy number. In this situation, since there is not a total order among the payoffs, the comparisons between the payoffs obtained by the players and by the coalitions are not straightforward as in scalar games and, therefore, classic solution concepts are not applicable.

Previous literature has addressed this difficulty by establishing a utility function in order to induce a scalar game and to obtain allocations of the associated total revenue based on different solution concepts. However, this approach seldom helps towards an accurate analysis of the situation, since the results are non-fuzzy payoffs.

This paper carries out an *ex-ante* analysis of the production situation and proposes a solution for the DEA production game with fuzzy prices, namely the preference least core, which is applicable before the fuzziness is resolved. In this solution the fuzzy nature of the allocations is preserved, and therefore, the quantity finally assigned to each agent is a fuzzy number. The preference least core has recently been introduced in [8] for set-valued DEA production games, and is based on the same idea as the least core in standard TU games. Its main drawback in the fuzzy environment is the difficulty involved in the effective computation of the fuzzy allocations.

We adopt standard fuzzy orders in the set of fuzzy numbers (see González and Vila [2], [3], and Ramík and Římánek [13]), and define the excess of the coalitions accordingly. For DEA production games the fuzzy allocations in the preference least core allocate the revenue obtained with one of the efficient production vectors which minimize the excess of the coalitions. The main contribution in the paper is the proposal of a procedure to compute allocations in the preference least core. The procedure requires solving a single linear programming model, which at the same time yields the efficient fuzzy revenue obtained by cooperation and the allocation to the agents of this fuzzy quantity.

We also show how our approach is applied in a case study, for which allocations in the preference least core are obtained, both for the case where uncertain prices are represented by triangular fuzzy numbers and trapezoidal fuzzy numbers. Specifically the data are a fuzzy version of those in Färe and Zelenyuk [1]. The example shows how the level of ambiguity of the fuzzy numbers which represent per-unit profits affects the precision of the fuzzy quantities which are finally allocated to the agents.

**Acknowledgments:** This research was carried out with the financial support of grants DPI2013-41469-P, ECO2011-29801-C02 of the Spanish Ministry of Science and Innovation and the European Regional Development Fund (ERDF). Support from the Andalusian Regional Government, under grant P11-SEJ-7782, is also aknowledged.

## References

[1] R. Färe and V. Zelenyuk, On aggregate Farrell efficiencies, *European Journal of Operational Research* 146 (2003) 615-620.



[2] A. González and M.A. Vila, A discrete method for studying indifference and order relations between fuzzy numbers, *Information Sciences* 56 (1991), 245-258.

[3] A. González and M.A. Vila, Dominance relation on fuzzy numbers, *Information Sciences* 64 (1992), 1-16.

[4] A. Hatami-Marbini, A. Emrouznejad, M. Tavana, A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making, *European Journal of Operational Research* 214 (2011) 457-472.

[5] M.A. Hinojosa, A.M. Mármol, L. Monroy, F.R. Fernández, A multi-objective approach to fuzzy linear production games, *International Journal of Information Technology and Decision Making* 12(5) (2013) 927-943.

[6] S. Lertworasirikul, S.C. Fang, H.L. Nuttle, J.A. Joines, Fuzzy BCC model for data envelopment analysis, *Fuzzy Optimization and Decision Making* 2 (2003) 337-358.

[7] S. Lozano, DEA production games, *European Journal of Operational Research* 231 (2013) 405-413.

[8] S. Lozano, M.A. Hinojosa, A.M. Mármol, Vector-Valued DEA Production Games, *OMEGA* (2014) Doi: 10.1016/j.omega.2014.10.002.

[9] S. Lozano, Computing fuzzy process efficiency in parallel systems, *Fuzzy Optimization and Decision Making* 13(1) (2014) 73-89.

[10] L. Monroy, M.A. Hinojosa, A.M. Mármol, F.R. Fernández, Set-valued cooperative games with fuzzy payoffs. The fuzzy assignment game, *European Journal of Operational Research* 225 (1) (2013) 85-90.

[11] I. Nishizaki, M. Sakawa, *Fuzzy and Multiobjective Games for Conflict Resolution*, Physica-Verlag. Heidelberg (2001).

[12] G. Owen, *Game Theory*, Academic Press (1995).

[13] J. Ramík and J. Římánek, Inequality relation between fuzzy numbers and its use in fuzzy optimization, *Fuzzy Sets and Systems* 16 (2) (1985) 123-138.

[14] L.A. Zadeh, Fuzzy sets, *Information and control* 8 (1965) 338-353.

# Towards generating fuzzy rules via Fuzzy Formal Concept Analysis

Valentín Liñeiro-Barea [1], Jesús Medina [1], and Inmaculada Medina-Bulo [2]

[1]Department of Mathematics, University of Cádiz. Spain,
[2]Department of Computer Engineering, University of Cádiz. Spain
{valentin.lineiro,jesus.medina,inmaculada.medina}@uca.es

**Abstract.** Nowadays, the interest about extracting knowledge from databases has increased in a wide variety of areas like stock market, medicine or census data, to name a few. Fuzzy Formal Concept Analysis plays a crucial role in the characterization in the sets of objects related to different sets of attributes. A compact representation of this is provided by the rule base, composed by fully and partially true implications between attributes. This paper shows the way to obtain the base of rules given a fuzzy formal context.

**Keywords:** Fuzzy rule, Fuzzy Formal Concept Analysis, Fuzzy set

## 1 Introduction

Extracting knowledge from databases is a critical technique in a wide variety of areas like stock market prediction [9], disease diagnosis [8] or census data analysis [4], among others. The knowledge extracted is usually represented as a set of rules which summarizes completely the information stored in the database. These rules are usually extracted via APRIORI algorithm [2], which exploits frequent itemsets to select the most frequent and condifent rules of the database.

Formal Concept Analysis (FCA) [5] is a technique which helps to discover relations between sets of attributes and objects inside a database, known as concepts. FCA fixes well to retrieve the main concepts that a database has, and that can be useful to obtain a rules set. In the classical FCA [2], only boolean attributes are considered, leading to a set of rules which consider attributes fully true or false. This approach may not be the most accurate in contexts where uncertainty and noise are present.

The main goal of this paper is to use the concept of fuzzy rule (that is a rule with attributes that might not be fully true or false) for defining the concept of fuzzy rule base as a minimum set of rules needed to summarize the whole information present in a database. For this, Fuzzy FCA [7] is used.

## 2 Preliminaries

In this section, we will present the required preliminary definitions. In all the definitions a complete lattice $(L, \preceq)$ and a finite universe $U$ are considered.

**Definition 1.** *A* spanning tree $\mathcal{S}$ *of a connected graph* $\langle V, E\rangle$ *is a subset of* $E$ *which interconnects all elements in* $V$*.*

**Definition 2 ([10]).** *Given a fuzzy set* $f\colon U \to L$*, the fuzzy cardinality of* $f$ *defined as* $card(f) = \sum_{i\in U} f(i)$ *is called* sigma count*.*

### 2.1 Fuzzy Formal Concept Analysis

In this subsection the main definitions regarding Fuzzy Formal Concept Analysis (FFCA) will be covered.

The first definition introduce the basic operators will be consider throughout the paper.

**Definition 3.** *Let* $(P_1, \leq_1)$*,* $(P_2, \leq_2)$*,* $(P_3, \leq_3)$ *be posets and* $\&\colon P_1 \times P_2 \to P_3$*,* $\swarrow\colon P_3 \times P_2 \to P_1$*,* $\nwarrow\colon P_3 \times P_1 \to P_2$ *be mappings, then* $(\&, \swarrow, \nwarrow)$ *is an* adjoint triple *with respect to* $P_1, P_2, P_3$ *if:*

$$\textit{Adjoint property}: \quad x \leq_1 z \swarrow y \quad \textit{iff} \quad x \,\&\, y \leq_3 z \quad \textit{iff} \quad y \leq_2 z \nwarrow x$$

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

In the FFCA environment, the posets $(P_1, \leq_1)$ and $(P_2, \leq_2)$ considered in the previous definition are actually complete lattices [7] and they will be denoted as $(L_1, \preceq_1)$ and $(L_2, \preceq_2)$. Now a definition of multi-adjoint frame is presented.

**Definition 4.** *A multi-adjoint frame* $\mathcal{L}$ *is a tuple*

$$(L_1, L_2, P, \preceq_1, \preceq_2, \leq, \&_1, \swarrow^1, \nwarrow_1, \ldots, \&_n, \swarrow^n, \nwarrow_n)$$

*where* $(L_1, \preceq_1)$ *and* $(L_2, \preceq_2)$ *are complete lattices,* $(P, \leq)$ *is a poset and* $(\&_i, \swarrow^i, \nwarrow_i)$ *is an adjoint triple with respect to* $L_1, L_2, P$*, for all* $i \in \{1, \ldots, n\}$*.*

Now the concept of context is presented.

**Definition 5.** *Let* $(L_1, L_2, P, \&_1, \ldots, \&_n)$ *be a multi-adjoint frame, a context is a tuple* $(A, B, R, \sigma)$ *such that* $A$ *and* $B$ *are non-empty sets (interpreted as attributes and objects, respectively),* $R$ *is a* $P$*-fuzzy relation* $R\colon A \times B \to P$ *and* $\sigma\colon A \times B \to \{1, \ldots, n\}$ *is a mapping which associates any element in* $A \times B$ *with a particular adjoint triple in the frame.*

$L_2^B$ and $L_1^A$ denote the set of fuzzy subsets $g\colon B \to L_2$, $f\colon A \to L_1$ respectively. On these sets a pointwise partial order can be considered from the partial orders in $(L_1, \preceq_1)$ and $(L_2, \preceq_2)$, which provides $L_2^B$ and $L_1^A$ the structure of complete lattice. Given a multi-adjoint frame and a context, the concept-forming operators we will use $^{\uparrow}\colon L_2^B \to L_1^A$ and $^{\downarrow}\colon L_1^A \to L_2^B$ will be defined as:

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

for all $g \in L_2^B$, $f \in L_1^A$, $a \in A$, $b \in B$. These two operators form a Galois connection [7]. The notion of concept can be defined as usual: A multi-adjoint concept is a pair $\langle g, f\rangle$ satisfying that $g \in L_2^B$, $f \in L_1^A$ and $g^{\uparrow} = f$, $f^{\downarrow} = g$, with $(^{\uparrow}, ^{\downarrow})$ being the Galois connection defined above. The fuzzy subsets $g$ and $f$ in a concept are usually known as the extent and intent of the concept, respectively.

**Definition 6.** *The multi-adjoint concept lattice associated with a multi-adjoint frame* $(L_1, L_2, P, \&_1, \ldots, \&_n)$ *and a context* $(A, B, R, \sigma)$ *is the set*

$$\mathcal{M} = \{\langle g, f\rangle \mid g \in L_2^B, f \in L_1^A \text{ and } g^{\uparrow} = f, f^{\downarrow} = g\}$$

*in which the ordering is defined by* $\langle g_1, f_1\rangle \preceq \langle g_2, f_2\rangle$ *if and only if* $g_1 \preceq g_2$ *(*$f_2 \preceq f_1$*). This ordering provides* $\mathcal{M}$ *the structure of complete lattice [7].*

## 3 Multi-adjoint rule base

From now on, a multi-adjoint frame $(L_1, L_2, P, \&_1, \ldots, \&_n)$ and a context $(A, B, R, \sigma)$ will be fixed. First, the definition of support, a key measure of a fuzzy rule, is introduced.

**Definition 7.** *The support of* $f \in L_1^A$ *in* $(A, B, R, \sigma)$ *is defined as* $supp(f) = \frac{card(f^{\downarrow})}{|B|}$*. This measure indicates the proportion of objects which are related to a given fuzzy subset of attributes.*

Once defined the support, in the following definition we can provide the fuzzy rule own one.

**Definition 8.** *Given two fuzzy subsets of attributes* $f_1, f_2 \in L_1^A$*, the* fuzzy rule over $A$ from $f_1$ to $f_2$ *is given by the expression* $f_2 \leftarrow_{(s,c)} f_1$*,* $s = supp(f_1)$ *and* $c$ *is defined by* $c = \frac{supp(f_1 \cup f_2)}{supp(f_1)}$*, called* the confidence of the rule*, that is the proportion in which the rule is true. If the confidence is* 1*, the fuzzy rule is called* fuzzy implication*, which is denotated as* $f_2 \Leftarrow f_1$*.*

### 3.1 Fuzzy rules base

Given a context, the set of fuzzy rules computed may be huge and it also may include redundant and not interesting rules for the purpose they are being obtained. In order to fix that, a base of rules is computed.

**Definition 9 ([3]).** *A set of fuzzy rules* $T$ *is a* minimal base *if the following properties hold:*

- *Completeness. Any rule* $f_2 \leftarrow_{(s,c)} f_1 \notin T$ *can be derived by the rules in* $T$*.*
- *Non-redundancy. No rule* $f_2 \leftarrow_{(s,c)} f_1 \in T$ *can be derived by the rest of rules in* $T$*.*

A minimal base of fuzzy rules $T$ is formed by two special subsets, depending on the confidence threshold considered: the set of fuzzy rules with confidence $c = 1$ is called *fuzzy implication base* and, if $c < 1$, the obtained set is called *fuzzy partial implication base.*

### 3.2 Fuzzy partial implication base

A fuzzy implication base contains all the fuzzy rules that have the greater confidence value ($c = 1$), however considering partial implication is also interesting since some noise in the data can be mitigated and outlier data are also considered. This section defines in the multi-adjoint concept lattice framework a partial implication base extending the classical case [6] and relates this base to the given multi-adjoint concept lattice.

First of all the notion of multi-adjoint partial implication is introduced.

**Definition 10.** *Given the set of all intents in the context* $Int(A, B, R, \sigma) = \{f \in L_1^A \mid f = f^{\downarrow\uparrow}\}$*, the fuzzy rule* $f_2 \leftarrow_{(s,c)} f_1$*, in which the fuzzy subset of attributes are intents, that is* $f_1, f_2 \in Int(A, B, R, \sigma)$*, and* $f_1 \prec f_2$ *is called* fuzzy partial implication*.*

The following result shows how new multi-adjoint partial implication can be derived from a base via transitivity. Based on the idea given in [6], the following result is obtained.

**Theorem 1.** *Let* $f_1, f_2, f_3 \in Int(A, B, R, \sigma)$*, where* $f_1 \prec f_2 \prec f_3$*, and the rules* $f_2 \leftarrow_{(s,c)} f_1$*,* $f_3 \leftarrow_{(s',c')} f_2$*,* $f_3 \leftarrow_{(s,c'')} f_1$*, we have that* $c \cdot c' = c''$*.*

As a consequence, given $f_1, f_2, f_3 \in \mathrm{Int}(A, B, R, \sigma)$, satisfying $f_1 \prec f_2 \prec f_3$, and the partial implications $f_2 \leftarrow_{(s,c)} f_1$ and $f_3 \leftarrow_{(s',c')} f_2$ the implication $f_3 \leftarrow_{(s,c'')} f_1$ can be straightforwardly derived considering $c'' = c \cdot c'$.

This result provides that only the rules between neighbor intents are considered in a fuzzy partial implication base and gives us a mechanism in order to compute a fuzzy partial implication base, only considering a minimal subset of relations between neighbor intents removing cycles in the concept lattice, in other words, obtaining a spanning tree of the Hasse diagram of the concept lattice. If we need to compute the confidence of a new rule, we just need to identify the path followed in the spanning tree and then multiply the confidences of the rules involved, inverting the value in the rules we consider in the reverse form. Example 1 will show how to derive new partial implications using this method.

*Example 1.* Given a spanning tree $\mathcal{S}$ of a concept lattice $\mathcal{M}$ formed by the partial implications $\{f_2 \leftarrow_{(0.5,0.4)} f_1, f_3 \leftarrow_{(0.4,0.3)} f_1, f_4 \leftarrow_{(0.2,0.25)} f_2\}$, the derived partial implication $f_4 \leftarrow_{(s,c)} f_3$ can be obtained transiting in the spanning tree in this form: $\{f_1 \leftarrow f_3, f_2 \leftarrow f_1, f_4 \leftarrow f_2\}$.

The support of the derived rule is trivially $supp(f_3)$ and the confidence is $\frac{1}{0.3} \cdot 0.4 \cdot 0.25 = 0.\overline{33}$.

### 3.3 Computing the whole multi-adjoint rules base

It has been shown how a fuzzy partial implication base can be obtained from a context given. However, that base is only one component of the whole multi-adjoint rules base. In order to compute the multi-adjoint rules base, it is needed to obtain additionally the implication base. The process can be summarized in the following points:

1. Computing a fuzzy implication base from the context (see [3] for a whole explanation of the process).
2. Obtaining a spanning tree of the concept lattice associated with the context, as shown in the previous subsection.

Now, an example of computing the fuzzy partial implication base of a context is provided.

*Example 2.* Let $(L, \preceq, \&_G)$ be a multi-adjoint frame, where $\&_G$ is the Gödel conjunctor with respect to $L = \{0, 0.25, 0.5, 0.75, 1\}$. The context $(A, B, R, \sigma)$ is formed by the sets $A = \{a_1, a_2, a_3\}$ and $B = \{b_1, b_2\}$, the relation $R$ defined from Table 1 and the constant mapping $\sigma$. For each $a \in A$, the expression $a/1$ will simply be written as $a$ and $a/0$ will be omitted. Figure 1 shows the concept

| $R$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $b_1$ | 0.5 | 0.25 | 1.0 |
| $b_2$ | 1.0 | 0.5 | 0.75 |

Table 1: Relation $R$ of example 2

lattice associated with the context given. The fuzzy rules computed from the concept lattice are marked in red. As we can see, the red arrows form a spanning tree of the Hasse diagram of the concept lattice. The computed fuzzy partial implication base is shown in figure 2.

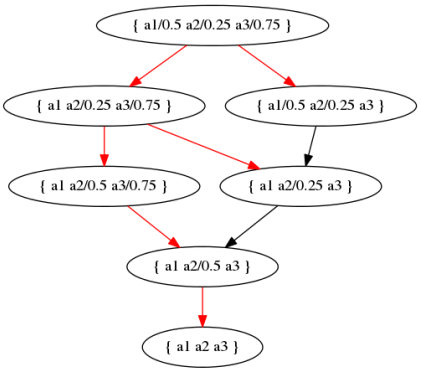



Fig. 1: Concept lattice associated with the context

$$\{a_1\} \leftarrow_{(0.75,0.75)} \{a_1/0.5, a_2/0.25, a_3/0.75\}$$
$$\{a_3\} \leftarrow_{(0.88,0.88)} \{a_1/0.5, a_2/0.25, a_3/0.75\}$$
$$\{a_2/0.5\} \leftarrow_{(0.62,0.83)} \{a_1, a_2/0.25, a_3/0.75\}$$
$$\{a_3\} \leftarrow_{(0.62,0.83)} \{a_1, a_2/0.25, a_3/0.75\}$$
$$\{a_3\} \leftarrow_{(0.5,0.8)} \{a_1, a_2/0.5, a_3/0.75\}$$
$$\{a_2\} \leftarrow_{(0.38,0.75)} \{a_1, a_2/0.5, a_3\}$$

Fig. 2: Fuzzy partial implication base

## 4 Conclusions and future work

In this paper FFCA has been utilized to provide an approximation of generating fuzzy rules of a database. First, fuzzy rules have been defined. With this defi-

nition, a fuzzy rules base has been characterized. Depending of the confidence threshold, two bases of rules compose the whole fuzzy rules base: fuzzy implication base ($c = 1$) and fuzzy partial implication base ($c < 1$). These properties help to obtain a strategy to generate both of them, as shown in the example given.

From now on a set of future work lines will be presented. First, it is critical to show the performance and the potential, in terms of information retrieved, of the procedure of generation, an experimental study is needed. There are another approaches in terms of rules bases, such the D-Basis [1]. This approach makes the consequents shorter and allows to compute a base in which an specific attribute is previously settled in the consequent, which may be a potential advantage to apply fuzzy rules in a regression scenario.

## References

1. K. V. Adaricheva and J. B. Nation. Discovery of the d-basis in binary tables based on hypergraph dualization. *CoRR*, abs/1504.02875, 2015.
2. R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In *ACM Sigmod International Conference On Management Of Data*, pages 207–216, 1993.
3. R. Belohlávek and V. Vychodil. Attribute dependencies for data with grades. *CoRR*, 2014.
4. O. Chertov and M. Aleksandrova. Using association rules for searching levers of influence in census data. *Procedia - Social and Behavioral Sciences*, 73(0):475 – 478, 2013. Proceedings of the 2nd International Conference on Integrated Information (IC-ININFO 2012), Budapest, Hungary, August 30 ? September 3, 2012.
5. B. Ganter and R. Wille. *Formal concept analysis: mathematical foundations*. Springer Science & Business Media, 2012.
6. M. Luxenburger. Implications partielles dans un contexte. *Mathématiques, Informatique et Sciences Humaines*, 29(113):35–55, 1991.
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. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, 40(4):1086 – 1093, 2013.
9. P. Paranjape-Voditel and U. Deshpande. An association rule mining based stock market recommender system. In *Emerging Applications of Information Technology (EAIT), 2011 Second International Conference on*, pages 21–24, Feb 2011.
10. L. A. Zadeh. Fuzzy logic. *Computer*, 21(4):83–93, 1988.

# Relating adjoint negations with strong adjoint negations

M. Eugenia Cornejo[1], Francesc Esteva[2], Jesús Medina[3], and Eloísa Ramírez-Poussa[1]

[1]Department of Statistic and O.R., University of Cádiz. Spain
mariaeugenia.cornejo@uca.es,eloisa.ramirez@uca.es
[3]Artificial Intelligence Research Institute (IIIA - CSIC), Bellaterra, Spain
esteva@iiia.csic.es
[2]Department of Mathematics, University of Cádiz. Spain
jesus.medina@uca.es

**Abstract.** Adjoint negations, whose definition is based on the implications of an adjoint triple, arise as a generalization of residuated negations. Recently, interesting properties of these negation operators have been introduced [5]. In this paper, a comparative survey with weak negations studied by Trillas, Esteva and Domingo [10, 13] is presented. Moreover, the relationship between weak and strong negations, introduced by these authors, is extended to adjoint negations. These technical developments lead us to increase the number of applications of adjoint negations.

**Key words:** residuated negations; weak and strong negations; adjoint triples.

## 1 Introduction

Negation operators play an important role in several frameworks and they have widely been studied in [8, 10, 20]. From residuated implications of a t-norm [4, 12, 19], it is defined the residuated negation defined from the residuated implication as $\neg x = x \to 0$. In addition, weak negations are one of the most general negation operators, which have heavily been studied by Trillas, Esteva and Domingo [10, 11, 13, 20]. In this paper, we will work with adjoint triples in order to consider more general negation operators.

Adjoint triples were firstly considered in [15, 18] taking into account the adjoint conjunctor and only one implication. They have been used as basic operators in Logic Programming [17], general substructural logics [3], Fuzzy Formal Concept Analysis [16], Fuzzy Relation Equations [9] and Rough Set Theory [7], providing more flexibility and increasing the range of applications.

From the implications of an adjoint triple, we define the generalization of the residuated negation which are called adjoint negations. Since they are associated with an adjoint triple with respect to three different posets, these negation operators are defined on two different posets. Dealing with this general structure is helpful in the applications as it has been highlighted in [1, 2, 9].

In this paper, we will compare adjoint negations with weak negations and we will show that adjoint negations are more general. Besides, a bijection between adjoint negations and strong adjoint negations will be presented, following the idea introduced by Trillas, Esteva and Domingo in [10, 13], in order to establish the relationship between adjoint negations and strong adjoint negations.

## 2 Adjoint negations and weak negations

Adjoint triples, which generalize triangular norms and their residuated implications [14], are considered to decrease the mathematical requirements of the basic operators used in several frameworks. In this paper, adjoint triples will be used in order to define adjoint negations. For that reason, we will start introducing the notion of adjoint triple.

**Definition 1.** *Let* $(P_1, \leq_1)$, $(P_2, \leq_2)$, $(P_3, \leq_3)$ *be posets and* $\&\colon P_1 \times P_2 \to P_3$, $\swarrow\colon P_3 \times P_2 \to P_1$, $\nwarrow\colon P_3 \times P_1 \to P_2$ *be mappings, then* $(\&, \swarrow, \nwarrow)$ *is an* adjoint triple *with respect to* $P_1, P_2, P_3$ *if:*

$$x \leq_1 z \swarrow y \quad \textit{iff} \quad x \,\&\, y \leq_3 z \quad \textit{iff} \quad y \leq_2 z \nwarrow x \tag{1}$$

*where* $x \in P_1$, $y \in P_2$ *and* $z \in P_3$. *The condition* (1) *is called* adjoint property.

If adjoint triples are used in environments that require finiteness such as Fuzzy Formal Concept Analysis to obtain a finite concept lattice [6, 16] and Fuzzy Relation Equations to guarantee the existence of minimal solutions [9], then it is important that adjoint triples are defined on regular partitions of the unit interval $[0, 1]$.

*Example 1.* Given $m \in \mathbb{N}$, the set $[0, 1]_m$ is a regular partition of $[0, 1]$ in $m$ pieces, for example $[0, 1]_2 = \{0, 0.5, 1\}$ divides the unit interval into two pieces.

A discretization of the Łukasiewicz t-norm is the operator $\&^*_{\mathrm{L}}\colon [0, 1]_{20} \times [0, 1]_8 \to [0, 1]_{100}$ defined, for each $x \in [0, 1]_{20}$ and $y \in [0, 1]_8$ as:

$$x \,\&^*_{\mathrm{L}}\, y = \frac{\lceil 100 \cdot \max(0, x + y - 1) \rceil}{100}$$

whose residuated implications $\swarrow^*_{\mathrm{L}}\colon [0, 1]_{100} \times [0, 1]_8 \to [0, 1]_{20}$, $\nwarrow^*_{\mathrm{L}}\colon [0, 1]_{100} \times [0, 1]_{20} \to [0, 1]_8$ are defined as:

$$z \swarrow^*_{\mathrm{L}} y = \frac{\lfloor 20 \cdot \min\{1, 1 - y + z\} \rfloor}{20} \qquad z \nwarrow^*_{\mathrm{L}} x = \frac{\lfloor 8 \cdot \min\{1, 1 - x + z\} \rfloor}{8}$$

where $\lceil \_ \rceil$ and $\lfloor \_ \rfloor$ are the ceiling and the floor functions, respectively. Hence, the triple $(\&^*_{\mathrm{L}}, \swarrow^*_{\mathrm{L}}, \nwarrow^*_{\mathrm{L}})$ is an adjoint triple. □

Now, we recall the definition of adjoint negations which is given from the implications of an adjoint triple and generalize the notion of residuated negation [4, 12, 19]. Adjoint negations are defined on two different posets since they are associated with an adjoint triple with respect to three different posets.

**Definition 2.** *Let $(P_1, \leq_1)$ and $(P_2, \leq_2)$ be two posets, $(P_3, \leq_3, \bot_3)$ be a lower bounded poset and $(\&, \swarrow, \nwarrow)$ an adjoint triple with respect to $P_1$, $P_2$ and $P_3$. The mappings $n_n \colon P_1 \to P_2$ and $n_s \colon P_2 \to P_1$ defined, for all $x \in P_1$, $y \in P_2$ as*

$$n_n(x) = \bot_3 \nwarrow x \qquad n_s(y) = \bot_3 \swarrow y$$

*are called* adjoint negations *with respect to $P_1$ and $P_2$.*

*The operators $n_s$ and $n_n$ satisfying that $x = n_s(n_n(x))$ and $y = n_n(n_s(y))$, for all $x \in P_1$ and $y \in P_2$, are called* strong adjoint negations.

Considering the adjoint triple $(\&_{\mathrm{L}}^*, \swarrow_{\mathrm{L}}^*, \nwarrow_{\mathrm{L}}^*)$ presented in Example 1, we introduce the next example of adjoint negations.

*Example 2.* The adjoint negations $n_s \colon [0,1]_8 \to [0,1]_{20}$ and $n_n \colon [0,1]_{20} \to [0,1]_8$ obtained from the adjoint triple $(\&_{\mathrm{L}}^*, \swarrow_{\mathrm{L}}^*, \nwarrow_{\mathrm{L}}^*)$ are defined as:

$$n_s(y) = \frac{\lfloor 20 \cdot (1-y) \rfloor}{20} \qquad n_n(x) = \frac{\lfloor 8 \cdot (1-x) \rfloor}{8}$$

Observe that the choice of the posets is fundamental. If the adjoint conjunctor is defined as $\&_{\mathrm{L}}^* \colon [0,1]_k \times [0,1]_t \to [0,1]_p$, the corresponding adjoint negations will be $n_s \colon [0,1]_t \to [0,1]_k$ and $n_n \colon [0,1]_k \to [0,1]_t$. Therefore,

(i) If $t = k$, then it is easy to verify that $n_s$ and $n_n$ are strong adjoint negations.
(ii) If $t \neq k$, the obtained adjoint negations are not strong adjoint negations, in general. □

One of the most general negation operators are weak negations, which have widely been studied by Trillas and Esteva et al [10, 11, 13, 20]. In order to compare adjoint negations with weak negations, we will remind the next definition.

**Definition 3 ([20]).** *Given a mapping $n \colon [0,1] \to [0,1]$ is said to be a* weak negation *if the following conditions hold, for all $x, y \in [0,1]$.*

1. *$n(1) = 0$;*
2. *if $x \leq y$ then $n(y) \leq n(x)$;*
3. *$x \leq n(n(x))$.*

*We will say that $n$ is a* strong negation *if the equality $x = n(n(x))$ holds, for all $x \in [0,1]$.*

The next theorem shows that adjoint negations are a generalization of weak negations.

**Theorem 1.** *If the mapping $n \colon [0,1] \to [0,1]$ is a weak negation, then there exists an adjoint triple $(\&, \swarrow, \nwarrow)$ with respect to the poset $([0,1], \leq)$ satisfying $n = n_s = n_n$.*

Once we have presented this result, we will study if the relation between weak and strong negations defined on a complete lattice studied in [10] can be extended to adjoint negations. This relationship ensures that weak negations can be defined uniquely from strong negations.

## 3 Relation between adjoint negations and strong adjoint negations

In this section, we will introduce the main result of this paper which proves that there exists an one to one correspondence between adjoint negations defined on two posets and strong adjoint negations defined on two complete meet-semilattices.

For that purpose, we will consider two posets $(P, \leq_P)$, $(Q, \leq_Q)$ and two complete meet-semilattices $(P', \preceq_{P'})$, $(Q', \preceq_{Q'})$ with maximum elements $\top_{P'}$ and $\top_{Q'}$, respectively, such that $P' \subseteq P$ and $Q' \subseteq Q$. From now on, the set of pair of adjoint negations $(n_s, n_n)$ with respect to $P$ and $Q$ satisfying that $n_s(P) = Q'$ and $n_n(Q) = P'$, will be denoted as $N_{(P',Q')}(P, Q)$ and the set of pairs of strong adjoint negations $(n'_s, n'_n)$ with respect to $P'$ and $Q'$ will be denoted as $SN(P', Q')$.

A bijection between $N_{(P',Q')}(P, Q)$ and $SN(P', Q')$ is obtained, as the following theorem shows.

**Theorem 2.** *There exists an one to one correspondence between $N_{(P',Q')}(P, Q)$ and $SN(P', Q')$.*

As a consequence, the next corollary is straighforwardly obtained.

**Corollary 1.** *Given a pair of strong adjoint negations $(n'_s, n'_n)$ with respect to $P'$ and $Q'$, there exists a pair of adjoint negations $(n_s, n_n)$ with respect to $P$ and $Q$ defined as:*

$$n_s(p) = n'_s(z_p) \quad with \quad z_p = \bigwedge_{P'}\{y \in P' \mid p \leq y\}$$
$$n_n(q) = n'_n(z_q) \quad with \quad z_q = \bigwedge_{Q'}\{x \in Q' \mid q \leq x\}$$

*such that $n_{s_{|P'}} = n'_s$, $n_{n_{|Q'}} = n'_n$, and $n_s(P) = Q'$, $n_n(Q) = P'$.*

There exist cases in which we can define only one pair of strong adjoint negations with respect to $P'$ and $Q'$. Then, applying the previous theorem and corollary, only one pair of adjoint negations can be defined with respect to $P$ and $Q$, as the following examples shows:

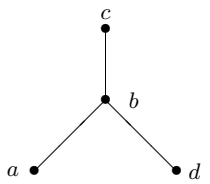
*Example 3.* Given $P' = \{p', \top_{P'}\}$ and $Q' = \{q', \top_{Q'}\}$. The unique pair of strong adjoint negations $(n'_s, n'_n)$ with respect to $(P', \preceq_{P'})$ and $(Q', \preceq_{Q'})$, is defined as $n'_s(p') = \top_{Q'}$, $n'_s(\top_{P'}) = q'$ and $n'_n(q') = \top_{P'}$, $n'_n(\top_{Q'}) = p'$. Then, there exists only one pair of adjoint negations $(n_s, n_n)$ with respect to $P$ and $Q$, being $(P, \leq_P)$ and $(Q, \leq_Q)$ two posets with maximum elements $\top_P \in P$ and $\top_Q \in Q$, such that $n_s(P) = Q'$ and $n_n(Q) = P'$. By Corollary 1, $n_s$ and $n_n$ are defined as follows:

$$n_s(p) = \begin{cases} \top_{Q'} & \text{if} \quad p \leq_P p' \\ q' & \text{otherwise} \end{cases} \qquad n_n(q) = \begin{cases} \top_{P'} & \text{if} \quad q \leq_Q q' \\ p' & \text{otherwise} \end{cases}$$

for all $p \in P$ and $q \in Q$.

*Example 4.* Let $(P' = \{a, b, c\}, \preceq_{P'})$ and $(Q' = \{x, y, z\}, \preceq_{Q'})$ two complete meet-semilattices such that $a \preceq_{P'} b \preceq_{P'} c$ and $x \preceq_{Q'} y \preceq_{Q'} z$. The pair $(n'_s, n'_n)$, defined as $n'_s(a) = z$, $n'_s(b) = y$, $n'_s(c) = x$ and $n'_n(x) = c$, $n'_n(y) = b$, $n'_n(z) = a$, is the unique pair of strong adjoint negations $(n'_s, n'_n)$ with respect to $P'$ and $Q'$. Therefore, applying Theorem 2, there exists only one pair of adjoint negations $(n_s, n_n)$ with respect to $P$ and $Q$, the posets given in Figure 1, such that $n_s(P) = Q'$ and $n_n(Q) = P'$. By Corollary 1, $n_s$ and $n_n$ are defined as follows:

$$n_s(p) = \begin{cases} z & \text{if} \quad p = a \\ y & \text{if} \quad p \in \{b, d\} \\ x & \text{if} \quad p = c \end{cases} \qquad n_n(q) = \begin{cases} c & \text{if} \quad q = x \\ b & \text{if} \quad q = y \\ a & \text{if} \quad q = z \end{cases}$$



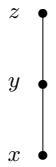




**Fig. 1.** The posets $(P, \leq_P)$ (left side) and $(Q, \leq_Q)$ (right side) of Example 4

## 4 Conclusions and further work

We have shown that adjoint negations are more general than weak negations studied by Trillas, Esteva and Domingo [10, 11, 13, 20]. Specifically, we have proven that every weak negation can be obtained from the implications of an adjoint triple. Moreover, an interesting generalization of the relation between weak and strong negations defined on a complete lattice studied in [10] has been presented. In this paper, a bijection between adjoint negations defined on two posets and strong adjoint negations defined on two complete meet-semilattices is shown.

As a further work, we will continue studying more properties of adjoint negations and possible applications of these operators. In addition, we will study the existence of an algorithm capable of computing the number of strong adjoint negations which can be defined on two complete meet-semilattices.

## References

1. L. Antoni, S. Krajci, O. Kridlo, B. Macek, and L. Pisková. On heterogeneous formal contexts. *Fuzzy Sets and Systems*, 234:22–33, 2014.

2. P. Butka, J. Pócs, and J. Pósová. On equivalence of conceptual scaling and generalized one-sided concept lattices. *Information Sciences*, 259(0):57–70, 2014.
3. P. Cintula, R. Horck, and C. Noguera. Non-associative substructural logics and their semilinear extensions: axiomatization and completeness properties. *Review of Symbolic Logic*, 6(3):394–423, 2013.
4. P. Cintula, E. P. Klement, R. Mesiar, and M. Navara. Residuated logics based on strict triangular norms with an involutive negation. *Mathematical Logic Quarterly*, 52(3):269–282, 2006.
5. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. General negations for residuated fuzzy logics. *Lecture Notes in Computer Science*, 8536:13–22, 2014.
6. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294(0):41 – 56, 2015.
7. 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.
8. M. E. Della Stella and C. Guido. Associativity, commutativity and symmetry in residuated structures. *Order*, 30(2):363–401, 2013.
9. J. C. Díaz and J. Medina. Multi-adjoint relation equations: Definition, properties and solutions using concept lattices. *Information Sciences*, 253:100–109, 2013.
10. F. Esteva. Negaciones en retculos completos. *Stochastica*, I:49–66, 1975.
11. F. Esteva and X. Domingo. Sobre funciones de negacin en [0,1]. *Stochastica*, IV:141–166, 1980.
12. F. Esteva, L. Godo, P. Hájek, and M. Navara. Residuated fuzzy logics with an involutive negation. *Archive for Mathematical Logic*, 39(2):103–124, 2000.
13. F. Esteva, E. Trillas, and X. Domingo. Weak and strong negation functions in fuzzy set theory. In *Proc. XI Int. Symposium on Multivalued Logic*, pages 23–26, 1981.
14. P. Hájek. *Metamathematics of Fuzzy Logic*. Trends in Logic. Kluwer Academic, 1998.
15. P. Julian, G. Moreno, and J. Penabad. On fuzzy unfolding: A multi-adjoint approach. *Fuzzy Sets and Systems*, 154(1):16–33, 2005.
16. 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.
17. 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.
18. 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.
19. W. San-Min. Logics for residuated pseudo-uninorms and their residua. *Fuzzy Sets and Systems*, 218(0):24–31, 2013. Theme: Logic and Algebra.
20. E. Trillas. Sobre negaciones en la teoría de conjuntos difusos. *Stochastica*, III:47–60, 1979.

# Some New Bivariate and Multivariate Dependence Measures

Alfonso J. Bello

University of Cadiz, Department of Stadistics and Operative Research

**Abstract.** Risk analysts, to build or hedge their portfolios, often have to decide among different assets, which is the best option to used. Then, they should evaluate the risk exposure of the new portfolio. In this paper we consider the important circumstances involved when the risk analysts are concerned with risks that exceed a certain threshold. Such conditions are well known to financial and insurance professionals, for instance in the context of hedging some asset or to extend his portfolio emphasizing the risk to exceed that threshold. In this framework, dependence measures are useful to make decisions and measure the dependence structure of the vectors under risk uncertainty. We propose new dependence measures for bivariant and multivariant random vectors. We study their bounds and properties. We also investigate the circumstances under which the existence of some stochastic orderings among their marginals, or some orderings among their dependence structure, imply an ordering among the corresponding conditional risk distributions.

## 1 Introduction

Random vectors can be used to describe financial and insurance portfolios. The risk analyst, to build or hedge his portfolio, often has to decide among differents assets which is the best option to use. Then, the risk analyst should evaluate the risk exposure to the new portfolio. Consider a portfolio $X$ and differents assets $Y_1, \ldots, Y_n$. To determine the distribution of the new portfolio $(X, Y_i)$, $i = 1, \ldots, n$, we have to know the joint distribution $H$ of $(X, Y_i)$. In practice however this often turns out to be very difficult. Anyway, in the joint distribution $H$, it is important the role that takes the marginal components and the role that takes the dependence structure (copula function). One way to simplify this problem is to simply neglect the dependence and assume that the risks are independent. Let $(X, Y_i)$ be a random vector with independent components, then $(X, Y_i)$ has cumulative distribution

$$H_i(x, y_i) = F(x)\ G_i(y_i).$$

Obviously, neglecting the dependence, we might underrate or overrate the risk of portfolio. Alternatively, one might consider the strongest positive dependence and assume that the risks are comonotonic. Let $(X, Y_i)$ be a random vector

with comonotonic components and let $F$ and $G_i$ be the corresponding marginal distribution functions, then $(X, Y_i)$ has cumulative distribution

$$H_i(x, y_i) = min\left\{F(x), G_i(y_i)\right\}.$$

In this case, we will likely be overrated the risk of portfolio.

Since the above approaches could be too restrictive assumptions, the dependence structure between the marginal components takes greater role. So, using dependence measures that consider the dependence structure and the behavior of marginal components prove to be useful. In this paper we introduce a bivariate and multivariate dependence measure that takes both aspects into account, and takes both extreme simplifications. These new measures differ from other multivariate dependence measures in e.g. Wolff (1980), Fernández Fernández and González-Barrios (2004), Taylor (2007), Behboodian et al. (2007), Schmid and Schmidt (2007), Koch and De Schepper (2011), Dhaene et al. (2013) or Cousin et al. (2014), as it focusses on the risk $X$ conditioning to $Y_i$, when $Y_i$ exceeds its $p-$quantile, rather than on the joint distribution function of $X$. In a finance context, this can be translated into a way for making decisions about the increasing or decreasing of a portfolio.

Distortion functions are very popular in financial and actuarial research. Denneberg (1990) introduced the idea of distortion and Wang (1996) developed it further. We can see in Sordo, Suarez and Bello (2014) how we can use distortion functions to describe differents conditional random variables.

**Remark**: At this time, the work is still in progress.

## References

1. Behboodian, J., Dolati, A., Úbeda Flores, M.: A multivariate version of Ginis rank association coefficient. Statistical Papers, 48: 295-304 (2007)
2. Cousin, A., Janon, A., Maume-Deschamps, V., Niang, I.: On the consistency of Sobol indices with respect to stochastic ordering of model parameters. Archive ouverte HAL. HAL Id: hal-01026373 (2014)
3. Dhaene, J., Linders, D., Schoutens, W., Vyncke, D.: A multivariate dependence measure for aggregating risks. Journal of Computational and Applied Mathematics, 263: 78-87 (2014)
4. Denneberg, D.: Distorted Probabilities and Insurance Premiums (Extended Abstract of Denneberg 1989). In Proceedings of 14. SOR, UIm Sept. 6-8, 1989. Frankfurt (Athen/:ium) (1990)
5. Fernández-Fernández, B., González-Barrios, J.M.: Multidimensional dependency measures. J. Multivar. Anal. 89(2), 351-370 (2004)
6. Koch, I., De Schepper, A.: Measuring comonotonicity in m-dimensional vectors. ASTIN Bulletin, 41(1): 191-213 (2011)
7. Schmid, F., Schmidt, R.: Multivariate extensions of Spearmans rho and related statistics. Statistics & Probability Letters, 77: 407-416 (2007)

8. Sordo, M.A., Suárez-Lloréns, A., Bello, A.J.: Comparison of conditional distributions in portfolios of dependent risks. Insurance: Mathematics and Economics 61, 62-69 (2015)
9. Taylor, M.: Multivariate measures of concordance. Annals of the Institute of Statistical Mathematics, 59(4): 789-806 (2007)
10. Wang, S.: Premium calculation by transforming the layer premium density. ASTIN Bull. 26, 71-92 (1996)
11. Wolff, E.F.: N−dimensional measures of dependence. Stochastica 4(3), 175-188 (1980)

# Using covers to characterize the solutions of fuzzy relation equations on linear carriers.

Jesús Medina[1], Esko Turunen[2], and J. C. Díaz-Moreno[1]

[1] Department of Mathematics. University of Cádiz, Spain.
Email: {jesus.medina,juancarlos.diaz}@uca.es
[2] Research Unit Computational Logic. Vienna University of Technology. Wien, Austria. Email: esko.turunen@tut.fi

**Abstract.** This paper studies a procedure to obtain the minimal solutions of fuzzy relation equations. From this study several results, based on the covering problem, are introduced generalizing other ones given in the literature.
**Keywords:** Fuzzy relation equations, minimal solutions, residuated structures.

## 1 Introduction

Supremum-$\odot$ fuzzy relation equations were introduced by E. Sanchez [4]. These and other many papers study the existence of solutions of these equations [1–3], and, in the affirmative case, they show that the set of solutions is a upper-preserving complete lattice in which the greatest solutions can easily be obtained. However, to know about minimal solutions is more difficult.

It is interesting to fix a general framework in which the minimal solutions of each solvable fuzzy relation equation exist and that each solution will be between the greatest solution and a minimal solution.

This paper considers a general setting, in which the operators may neither be commutative nor associative and they only need to be monotone and residuated inf-preserving mappings of non-empty sets on the right argument. The linearity of the carrier, together with the inf-preserving property, ensures the existence of minimal solutions whenever a solution exists.

Moreover, this paper provides a mechanism to obtain the minimal solutions.

## 2 General fuzzy relation equations

A complete linear lattice $(L, \preceq)$ is the carrier considered throughout this paper, hence, the bottom and the top elements exist in $L$ and are denoted as $0, 1$, respectively. Given a set $V$, the ordering $\preceq$ in the lattice induces a partial order on the set of $L$-fuzzy subsets of $V$, $L^V$. This ordering provides to $L^V$ the structure of a complete lattice.

Moreover, the general residuated operator used to define the fuzzy relation equation is $\odot\colon L \times L \to L$, such that it is order preserving and there exists an operator $\to\colon L \times L \to L$, satisfying the following adjoint property with $\odot$

$$x \odot y \preceq z \quad \text{if and only if} \quad y \preceq x \to z \tag{1}$$

for each $x, y, z \in L$. Note that this property is equivalent to $\odot$ preserves supremums in the second argument; $x \odot \bigvee\{y \mid y \in Y\} = \bigvee\{x \odot y \mid y \in Y\}$, for all $Y \subseteq L$. Hence, very few properties are assumed.

An important notion needed in this paper is the definition of *a cover*.

**Definition 1.** *Given an ordered set $(A, \preceq)$ and non-empty subsets $S_1, \ldots, S_n$ of $A$, an element $a \in A$ is* a cover of $\{S_1, \ldots, S_n\}$, *if for each $i \in \{1, \ldots, n\}$, there exists $s_i \in S_i$ such that $s_i \preceq a$. A cover $a \in A$ is called* minimal *if every element $d \in A$ satisfying $d \prec a$, is not a cover of $\{S_1, \ldots, S_n\}$.*

Note that when $(A, \preceq)$ is a complete lattice, minimal covers always exist in $A$.

**Definition 2.** *Given the pair $(\odot, \to)$, a* fuzzy relation equation *is the 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; and $R \circ X\colon U \times W \to L$ is defined, for each $u \in U$, $w \in W$, as*

$$(R \circ X)\langle u, w\rangle = \bigvee\{R\langle u, v\rangle \odot X\langle v, w\rangle \mid v \in V\}.$$

It is well known that the fuzzy relation equation (2) 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 [4, 5].

## 3 Minimal solutions generated by a given solution

**Definition 3.** *Given an operator $\odot\colon L \times L \to L$, we will say that it holds the* IPNE-condition *(making reference to that $\odot$ is Infimum Preserving of arbitrary Non-Empty sets), if it verify*

$$a \odot \bigwedge\{b_i \mid i \in \Gamma\} = \bigwedge\{a \odot b_i \mid i \in \Gamma\} \tag{3}$$

*for each element $a \in L$ and each non-empty subset $\{b_i \mid i \in \Gamma\} \subseteq L$.*

From now on, let us consider a general solvable fuzzy relation equation (2), where $R, X, T$ are finite, $U = \{u_1, \ldots, u_n\}$, $W = \{w_1, \ldots, w_m\}$, and $\odot$ satisfies the IPNE-condition.

Our first result characterizes the solutions of a solvable fuzzy relation equation $R \circ X = T$ by the covering elements of a family of subsets $S_{ij}$. Next, these

sets are defined. First of all, the auxiliary sets $V_{ij}$ need to be introduced, which are associated with the elements $u_i$, $w_j$ and the greatest solution $R \Rightarrow T$. Since for each $j = 1, \ldots, m$, $i = 1, \ldots, n$

$$\bigvee \{R\langle u_i, v\rangle \odot (R \Rightarrow T)\langle v, w_j\rangle \mid v \in V\} = T\langle u_i, w_j\rangle, \tag{4}$$

$L$ is linear and $V$ is finite, there exists at least one $v_s \in V$ validating the equation

$$R\langle u_i, v_s\rangle \odot (R \Rightarrow T)\langle v_s, w_j\rangle = T\langle u_i, w_j\rangle. \tag{5}$$

Therefore, the set $V_{ij} = \{v_s \in V \mid R\langle u_i, v_s\rangle \odot (R \Rightarrow T)\langle v_s, w_j\rangle = T\langle u_i, w_j\rangle\}$ is not empty and, for all $v \notin V_{ij}$, the strict inequality $R\langle u_i, v\rangle \odot (R \Rightarrow T)\langle v_s, w_j\rangle < T\langle u_i, w_j\rangle$ holds.

Each $v_s$ in $V_{ij}$ will provide a fuzzy subset $S_{ijs}$ as follows: Given $v_s \in V_{ij}$, we have that

$$\{d \in L \mid R\langle u_i, v_s\rangle \odot d = T\langle u_i, w_j\rangle\} \neq \varnothing$$

and the infimum $\bigwedge\{d \in L \mid R\langle u_i, v_s\rangle \odot d = T\langle u_i, w_j\rangle\} = e_s$ also satisfies the equality

$$R\langle u_i, v_s\rangle \odot e_s = T\langle u_i, w_j\rangle$$

by the IPNE-condition. These elements are used to define the fuzzy subsets $S_{ijs} \colon V \to L$ of $V$, defined by

$$S_{ijs}(v) = \begin{cases} e_s \text{ if } \ v = v_s \\ 0 \ \text{otherwise} \end{cases}$$

which form the set $S_{ij}$, that is $S_{ij} = \{S_{ijs} \mid v_s \in V_{ij}\}$, for each $i = 1, \ldots, n$, $j = 1, \ldots, m$. These sets will be used to characterize the set of solutions of Equation (2) by the notion of *covering*.

**Theorem 1.** *The L-fuzzy relation $X \colon V \times W \to L$ is a solution of a solvable Equation (2) if and only if $X \preceq (R \Rightarrow T)$ and, for each $j = 1, \ldots, m$, the fuzzy subset $X_j \colon V \to L$, defined by $X_j(v) = X\langle v, w_j\rangle$, is a cover of $\{S_{1j}, \ldots, S_{nj}\}$.*

As a consequence, the minimal solutions are characterized by the minimal covers.

**Corollary 1.** *$X \colon V \times W \to L$ is a minimal solution of Equation (2) if and only if, for each $j = 1, \ldots, m$, $X_j \colon V \to L$, defined by $X_j(v) = X\langle v, w_j\rangle$, is a minimal cover of $\{S_{1j}, \ldots, S_{nj}\}$.*

Hence, from the corollary above, minimal solutions of the fuzzy relation equation (2) are obtained from $R \Rightarrow T$ as follows:

**Procedure to obtain minimal solutions of Equation** (2)

- For each $j \in \{1, \ldots, m\}$ ($j$th element $w_j \in W$)
  - for each $i \in \{1, \ldots, n\}$ ($i$th element $u_i \in U$)
    - compute $V_{ij} = \{v_s \in V \mid R\langle u_i, v_s\rangle \odot (R \Rightarrow T)\langle v_s, w_j\rangle = T\langle u_i, w_j\rangle\}$
    - for each $v_s \in V_{ij}$ compute $e_s = \bigwedge\{d \in L \mid R\langle u_i, v_s\rangle \odot d = T\langle u_i, w_j\rangle\}$
    - construct the set $S_{ij}$ of corresponding characteristic mappings $S_{ijs}$.
  - Compute the minimal cover(s) $X_j$ of the set $\{S_{1j}, \ldots, S_{nj}\}$.
- Define the $L$-fuzzy matrix $X \colon V \times W \to L$ as $X\langle v, w_j\rangle = X_j(v)$.

**Example 1**

Let us assume 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 $x, y, z \in [0, 1]$.

Given $U = \{u_1, u_2, u_3\}$, $V = \{v_1, v_2, v_3\}$ $W = \{w_1, w_2, w_3\}$ 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 |

direct computation shows that the relation $R \Rightarrow T$, defined from the 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 |

is the greatest solution of Equation (2).

Given that

$$\begin{aligned} R\langle u_1, v_1\rangle \odot (R \Rightarrow T)\langle v_1, w_1\rangle &= 0.8 \\ R\langle u_1, v_2\rangle \odot (R \Rightarrow T)\langle v_2, w_1\rangle &= 0.2\,, \\ R\langle u_1, v_3\rangle \odot (R \Rightarrow T)\langle v_3, w_1\rangle &= 0.8 \end{aligned}$$

we have that $V_{1,1} = \{v_1, v_3\}$ and

$$S_{1,1} = \{\begin{pmatrix} 0.9 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix}\}.$$

Analogously, we have that $V_{2,1} = \{v_2, v_3\}$ and

$$S_{2,1} = \{\begin{pmatrix} 0 \\ 0.7 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix}\};$$

and $V_{3,1} = \{v_3\}$,

$$S_{3,1} = \{\begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix}\}$$

We observe that

$$X_1 = \begin{pmatrix} 0 \\ 0 \\ 0.9 \end{pmatrix} = S_{1,1} \cap S_{2,1} \cap S_{3,1},$$

so $X_1$ is the only minimal cover of $\{S_{1,1}, S_{2,1}, S_{3,1}\}$.

Next, we consider the second column of $R \Rightarrow T$. We can obtain that $V_{1,2} = \{v_1, v_3\}$ and

$$S_{1,2} = \{\begin{pmatrix} 0.5 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0.5 \end{pmatrix}\};$$

$V_{2,2} = \{v_2\}$ ,

$$S_{2,2} = \{\begin{pmatrix} 0 \\ 0.8 \\ 0 \end{pmatrix}\};$$

$V_{3,2} = \{v_2, v_3\}$,

$$S_{3,2} = \{\begin{pmatrix} 0 \\ 0 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0.8 \\ 0 \end{pmatrix}\}.$$

In this case, $\begin{pmatrix} 0 \\ 0.8 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.8 \\ 0 \end{pmatrix}$ are the only minimal covers of $\{S_{1,2}, S_{2,2}, S_{3,2}\}$.
Now, for the third column of $R \Rightarrow T$,

$V_{1,3} = \{v_1\}$

$$S_{1,3} = \{\begin{pmatrix} 0.8 \\ 0 \\ 0 \end{pmatrix}\};$$

$V_{2,3} = \{v_2, v_3\}$,

$$S_{2,3} = \{\begin{pmatrix} 0 \\ 0.4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0.6 \end{pmatrix}\};$$

$V_{3,3} = \{v_1\}$,

$$S_{3,3} = \{\begin{pmatrix} 0.8 \\ 0 \\ 0 \end{pmatrix}\}.$$

In this case, there are two minimal covers of $\{S_{1,3}, S_{2,3}, S_{3,3}\}$:

$$\begin{pmatrix} 0.8 \\ 0.4 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0.8 \\ 0 \\ 0.6 \end{pmatrix}.$$

This yields four fuzzy relations, defined as follows

| $X_4$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0 | 0.8 |
| $v_2$ | 0 | 0.8 | 0.4 |
| $v_3$ | 0.9 | 0.5 | 0 |

| $X_5$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0 | 0.8 |
| $v_2$ | 0 | 0.8 | 0 |
| $v_3$ | 0.9 | 0.5 | 0.6 |

| $X_6$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0.5 | 0.8 |
| $v_2$ | 0 | 0.8 | 0.4 |
| $v_3$ | 0.9 | 0 | 0 |

| $X_7$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|
| $v_1$ | 0 | 0.5 | 0.8 |
| $v_2$ | 0 | 0.8 | 0 |
| $v_3$ | 0.9 | 0 | 0.6 |

that solve Equation (2). By their construction and the properties of the Łukasiewicz conjunctor, they are minimal solutions.

## 4 Conclusion and future works

The main aim of this research is to define as generally as possible an algebraic structure that allows the existence of minimal solutions of the fuzzy relation equations defined based on this structure. For that, a general increasing operation $\odot$, which only satisfies the adjointness property, i.e. is residuated, and satisfies the IPNE-condition, has been considered to define a general fuzzy relation equation, which has minimal solutions whenever a solution exists. Moreover, a new algebraic characterization using the notion of covering is introduced, which provides a method to obtain the minimal solutions and, consequently, the whole set of solutions.

As future work, the obtained results will be applied to several problems in fuzzy logic, such as to abduction reasoning. It is well-known that implications in MV-algebras are infinitely distributive. A topic of future study is to characterize all structures where implication is infinitely distributivity. Algebraic structures that satisfy the INPE-condition are not studied much; also they will be a topic of future research.

## References

## 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. R. Bělohlávek. *Fuzzy Relational Systems: Foundations and Principles*. Kluwer Academic Publishers, 2002.
3. 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.
4. E. Sanchez. Resolution of composite fuzzy relation equations. *Information and Control*, 30(1):38–48, 1976.

5. 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.

# A combination of attribute reduction and size reduction strategies in multi-adjoint concept lattices

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

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

**Abstract.** Attribute and size reductions of concept lattices are key research topics in Formal Concept Analysis. This paper combines both strategies in the multi-adjoint concept lattice framework in order to simplify the information provided by the original context. For that purpose, we apply the attribute reduction and then the size reduction by means of an irreducible $\alpha$-cut concept lattice, analyzing the obtained properties.

**Keywords:** Formal concept analysis; fuzzy-attributes; concept lattice reduction

## 1 Introduction

Formal Concept Analysis is a tool in charge of extracting pieces of information from databases which contain a set of attributes $A$ and a set of objects $B$ together with a relation between them $R \subseteq A \times B$. These pieces of information are called concepts and they are hierarchized in order to obtain concept lattices.

It is well known that the computational complexity to obtain concept lattices decreases if the number of attributes is previously reduced. However, the process of attribute reduction is difficult in the fuzzy case [6, 10]. Although one of the main profits of the procedure presented in [6] is that the original concept lattice is conserved, this fact can become a drawback if the obtained concept lattice is very big and illegible. Therefore, it is also necesssary to study mechanisms for decreasing the size of concept lattices.

As well as the use of hedges in the concept-forming operators [1, 8, 9] and the methodology provided by granular computing [7], there exists another mechanism, based on the meet-irreducible elements of the lattice and a cut value given by the user, to reduce the size of concept lattices [3, 5]. This method provides a sublattice of the original concept lattice, called meet-irreducible $\alpha$-cut concept lattice. Consequently, the most representative knowledge is preserved.

In this paper, we combine both reduction mechanisms. Firstly, we will carry out the attribute reduction to obtain a concept lattice isomorphic to the original one and then we will apply the size reduction strategy. We study the influence of this combination on the size of the reduced concept lattices.

## 2 Preliminary notions

Some preliminary necessary notions and results to understand this work are introduced. In the concept lattice environment, we need to consider a multi-adjoint frame $(L_1, L_2, P, \&_1, \ldots, \&_n)$ where $(L_1, \leq_1)$ and $(L_2, \leq_2)$ are complete lattices, $(P, \leq)$ a poset and $(\&_i, \swarrow^i, \nwarrow_i)$ is an adjoint triple for all $i \in \{1, \ldots, n\}$, 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 in the frame. We 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. More information about these notions can be found in [6].

In order to recall the characterization of the $\wedge$-irreducible elements of a multi-adjoint concept lattice $(\mathcal{M}, \preceq)$, we introduce the following results [4, 6].

**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*

$$\phi_{a,x}(a') = \begin{cases} x & \text{if } a' = a \\ \bot_1 & \text{if } a' \neq a \end{cases}$$

*will be called* fuzzy-attributes*, where $\bot_1$ 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 ([6]).** *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 $b \in B$.*

### 2.1 Attribute classification in multi-adjoint concept lattices

Now, several theorems will be recalled in order to classify the set of attributes and thus a reduction in the number of attributes is obtained [6]. The first one characterizes the absolutely necessary attributes.

**Theorem 2.** *Given $a_i \in A$, we have that $a_i \in C_f$ if and only if there exists $x_i \in L_1$, such that $\langle \phi_{a_i,x_i}^{\downarrow}, \phi_{a_i,x_i}^{\downarrow\uparrow} \rangle \in M_F(A)$, satisfying that $\langle \phi_{a_i,x_i}^{\downarrow}, \phi_{a_i,x_i}^{\downarrow\uparrow} \rangle \neq \langle \phi_{a_j,x_j}^{\downarrow}, \phi_{a_j,x_j}^{\downarrow\uparrow} \rangle$, for all $x_j \in L_1$ and $a_j \in A$, with $a_j \neq a_i$.*

The next result, which characterizes the relatively necessary attributes, requires the use of the auxiliary sets $E_{a_i,x}$, where $a_i \in A$ and $x \in L_1$, defined as:

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

**Theorem 3.** *Given $a_i \in A$, we have that $a_i \in K_f$ if and only if $a_i \notin C_f$ and there exists $\langle \phi_{a_i,x_i}^{\downarrow}, \phi_{a_i,x_i}^{\downarrow\uparrow} \rangle \in M_F(A)$ satisfying that $E_{a_i,x_i}$ is not empty and $A \setminus E_{a_i,x_i}$ is a consistent set.*

Finally, it is shown the characterization of absolutely unnecessary attributes.

**Theorem 4.** *Given $a_i \in A$, 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.*

### 2.2 Meet-irreducible $\alpha$-cut concept lattice of $\mathcal{M}$

Considering the characterization given by Theorem 1, we recall a procedure to reduce the size of multi-adjoint concept lattices introduced in [5]. To apply this mechanism, given a value $\alpha$, for each attribute $a$, we only consider the following set of meet-irreducible elements of $(\mathcal{M}, \preceq)$:

$$\widehat{M_F}(A)_\alpha = \{\langle \phi^{\downarrow}_{a,x}, \phi^{\downarrow\uparrow}_{a,x} \rangle \in M_F(A) \mid \alpha \preceq_1 x\}$$

In order to get a complete lattice, we consider the concepts of $(\mathcal{M}, \preceq)$ obtained from the infimum of elements of $\widehat{M_F}(A)_\alpha$ and the greatest element in $(\mathcal{M}, \preceq)$, that is, $\langle g_\top, g^{\uparrow}_{\top} \rangle$.

**Definition 2.** *Given $\alpha \in L_1$, the set*

$$\widehat{\mathcal{M}}_\alpha = \{\langle g, f \rangle \in \mathcal{M} \mid g = \bigwedge_{i \in I} \phi^{\downarrow}_{a_i,x_i}, \text{ with } \langle \phi^{\downarrow}_{a_i,x_i}, \phi^{\downarrow\uparrow}_{a_i,x_i} \rangle \in \widehat{M_F}(A)_\alpha\} \bigcup \{\langle g_\top, g^{\uparrow}_{\top} \rangle\}$$

*is called* meet-irreducible $\alpha$-cut concept lattice of $\mathcal{M}$*, for short,* irreducible $\alpha$-cut CL.

The set $\widehat{\mathcal{M}}_\alpha$, with the ordering defined on $\mathcal{M}$, forms a complete sublattice of the original one. Therefore, this mechanism provides a reduction of the original concept lattice without modifying the information given by the concepts.

## 3 Combining both methods

An interesting option could be the combination of both reduction mechanisms. In this section, we will analyze the results of such combination, explaining the properties and the main advantages provided. Firstly, we need to introduce a new definition related to the new merging mechanism.

**Definition 3.** *Given a context $(A, B, R, \sigma)$, the frame $(L_1, L_2, P, \&_1, \ldots, \&_n)$, the concept lattice $(\mathcal{M}, \preceq)$, a value $\alpha \in L_1$ and $(\mathcal{M}^Y, \preceq)$ the concept lattice built from a reduct $Y \subseteq A$. The concept lattice obtained applying the irreducible $\alpha$-cut to the concept lattice $(\mathcal{M}^Y, \preceq)$, is called* r-irreducible $\alpha$-cut concept lattice *and it is denoted as $\widehat{\mathcal{M}}_\alpha{}^Y$.*

When the attribute classification satisfies that the set of relatively necessary attributes, $K_f$, is not empty then several reducts can be obtained. This fact will determine the subsequent size reduction i.e. depending on the choice of the starting reduct, we will obtain a major or minor reduction of the size. This is precisely what is shown in the following example.

*Example 1.* It will be considered the framework $\mathcal{L} = (L_1, L_2, L_3, \preceq, \&_G^*)$, where $L_1 = [0,1]_{10}$, $L_2 = [0,1]_4$ and $L_3 = [0,1]_5$ are the regular partitions of $[0,1]$ in 10, 4 and 5 pieces, respectively, and $\&_G^*$ is the Gödel conjunctor defined on $L_1 \times L_2$, see [2] for more details. The fixed context is $(A, B, R, \sigma)$, with $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, $B = \{b_1, b_2, b_3\}$, $R\colon A \times B \to L_3$ given by the table shown in the left side of Figure 1, and $\sigma$ is constantly $\&_G^*$. From this framework and

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

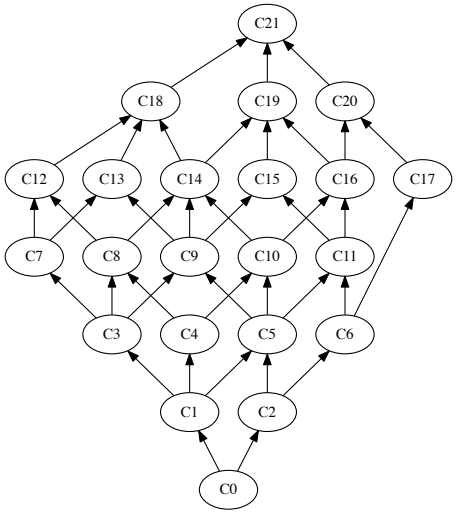



**Fig. 1.** The definition of the relation $R$ of Example 1 and the Hasse diagram of $(\mathcal{M}, \preceq)$

this context, we obtain a multi-adjoint concept lattice composed by 21 concepts.

The Hasse diagram of the concept lattice $(\mathcal{M}, \preceq)$ is presented in the right side of Figure 1. In this case, $C_{12}$, $C_{13}$, $C_{15}$, $C_{17}$, $C_{18}$, $C_{19}$ and $C_{20}$ are the meet-irreducible elements and the fuzzy-attributes associated with the meet-irreducible elements are listened below:

$$\begin{aligned}
&\langle \phi^{\downarrow}_{a_5,0.1}, \phi^{\downarrow\uparrow}_{a_5,0.1}\rangle = \langle \phi^{\downarrow}_{a_5,0.2}, \phi^{\downarrow\uparrow}_{a_5,0.2}\rangle = \langle \phi^{\downarrow}_{a_5,0.3}, \phi^{\downarrow\uparrow}_{a_5,0.3}\rangle = \langle \phi^{\downarrow}_{a_5,0.4}, \phi^{\downarrow\uparrow}_{a_5,0.4}\rangle =\\
&\langle \phi^{\downarrow}_{a_5,0.5}, \phi^{\downarrow\uparrow}_{a_5,0.5}\rangle = \langle \phi^{\downarrow}_{a_5,0.6}, \phi^{\downarrow\uparrow}_{a_5,0.6}\rangle = \langle \phi^{\downarrow}_{a_5,0.7}, \phi^{\downarrow\uparrow}_{a_5,0.7}\rangle = \langle \phi^{\downarrow}_{a_5,0.8}, \phi^{\downarrow\uparrow}_{a_5,0.8}\rangle =\\
&\langle \phi^{\downarrow}_{a_5,0.9}, \phi^{\downarrow\uparrow}_{a_5,0.9}\rangle = \langle \phi^{\downarrow}_{a_5,1.0}, \phi^{\downarrow\uparrow}_{a_5,1.0}\rangle = \langle \phi^{\downarrow}_{a_6,0.1}, \phi^{\downarrow\uparrow}_{a_6,0.1}\rangle = \langle \phi^{\downarrow}_{a_6,0.2}, \phi^{\downarrow\uparrow}_{a_6,0.2}\rangle =\\
&\langle \phi^{\downarrow}_{a_6,0.3}, \phi^{\downarrow\uparrow}_{a_6,0.3}\rangle = \langle \phi^{\downarrow}_{a_6,0.4}, \phi^{\downarrow\uparrow}_{a_6,0.4}\rangle = \langle \phi^{\downarrow}_{a_6,0.5}, \phi^{\downarrow\uparrow}_{a_6,0.5}\rangle = \langle \phi^{\downarrow}_{a_6,0.6}, \phi^{\downarrow\uparrow}_{a_6,0.6}\rangle = C_{12}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \langle \phi^{\downarrow}_{a_4,0.9}, \phi^{\downarrow\uparrow}_{a_4,0.9}\rangle = \langle \phi^{\downarrow}_{a_4,1.0}, \phi^{\downarrow\uparrow}_{a_4,1.0}\rangle = C_{13}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \langle \phi^{\downarrow}_{a_1,0.9}, \phi^{\downarrow\uparrow}_{a_1,0.9}\rangle = \langle \phi^{\downarrow}_{a_1,1.0}, \phi^{\downarrow\uparrow}_{a_1,1.0}\rangle = C_{15}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \langle \phi^{\downarrow}_{a_2,0.5}, \phi^{\downarrow\uparrow}_{a_2,0.5}\rangle = \langle \phi^{\downarrow}_{a_2,0.6}, \phi^{\downarrow\uparrow}_{a_2,0.6}\rangle = C_{17}\\
&\langle \phi^{\downarrow}_{a_4,0.5}, \phi^{\downarrow\uparrow}_{a_4,0.5}\rangle = \langle \phi^{\downarrow}_{a_4,0.6}, \phi^{\downarrow\uparrow}_{a_4,0.6}\rangle = \langle \phi^{\downarrow}_{a_4,0.7}, \phi^{\downarrow\uparrow}_{a_4,0.7}\rangle = \langle \phi^{\downarrow}_{a_4,0.8}, \phi^{\downarrow\uparrow}_{a_4,0.8}\rangle = C_{18}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \langle \phi^{\downarrow}_{a_1,0.7}, \phi^{\downarrow\uparrow}_{a_1,0.7}\rangle = \langle \phi^{\downarrow}_{a_1,0.8}, \phi^{\downarrow\uparrow}_{a_1,0.8}\rangle = C_{19}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \langle \phi^{\downarrow}_{a_2,0.3}, \phi^{\downarrow\uparrow}_{a_2,0.3}\rangle = \langle \phi^{\downarrow}_{a_2,0.4}, \phi^{\downarrow\uparrow}_{a_2,0.4}\rangle = C_{20}
\end{aligned}$$

According to the attribute classification theorems given in Section 2.1, we have that in this case $C_f = \{a_1, a_2, a_4\}$, $K_f = \{a_5, a_6\}$ and $I_f = \{a_3\}$. Considering this classification two reducts can be obtained, $Y_1 = \{a_1, a_2, a_4, a_5\}$ and $Y_2 = \{a_1, a_2, a_4, a_6\}$. From these reducts, we obtain two concept lattices $(\mathcal{M}^{Y_1}, \preceq)$ and $(\mathcal{M}^{Y_2}, \preceq)$ isomorphic to the original one, that is, $(\mathcal{M}, \preceq) \cong (\mathcal{M}^{Y_1}, \preceq) \cong (\mathcal{M}^{Y_2}, \preceq)$.

Now, we will obtain the r-irreducible $\alpha$-cut concept lattice of $(\mathcal{M}^{Y_1}, \preceq)$ with $\alpha = 0.7$. For that purpose, we need to consider the set of concepts $\langle \phi^{\downarrow}_{a,x}, \phi^{\downarrow\uparrow}_{a,x} \rangle$ belongs to $M_F(Y_1)$ such that $\alpha \preceq x$, with a maximal value $x$, that is:

$$\widehat{M_F}(Y_1)_{0.7} = \{\langle \phi^{\downarrow}_{a_5,1.0}, \phi^{\downarrow\uparrow}_{a_5,1.0} \rangle, \langle \phi^{\downarrow}_{a_4,1.0}, \phi^{\downarrow\uparrow}_{a_4,1.0} \rangle, \langle \phi^{\downarrow}_{a_1,1.0}, \phi^{\downarrow\uparrow}_{a_1,1.0} \rangle, \langle \phi^{\downarrow}_{a_4,0.8}, \phi^{\downarrow\uparrow}_{a_4,0.8} \rangle, \langle \phi^{\downarrow}_{a_1,0.8}, \phi^{\downarrow\uparrow}_{a_1,0.8} \rangle\}$$

Considering the infimum of elements of $\widehat{M_F}(Y_1)_{0.7}$ and the greatest element of $(\mathcal{M}^{Y_1}, \preceq)$, we compute the lattice $(\widehat{\mathcal{M}}^{Y_1}_{0.7}, \preceq)$ as usual, which can be seen[1] in the left side of Figure 2.

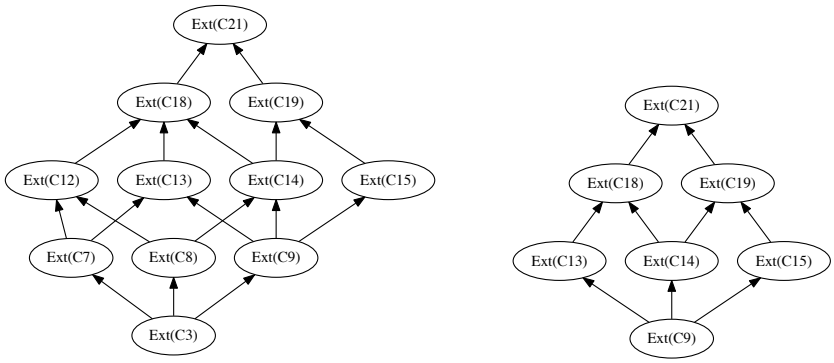



**Fig. 2.** The Hasse diagram of $(\widehat{\mathcal{M}}^{Y_1}_{0.7}, \preceq)$ (left) and $(\widehat{\mathcal{M}}^{Y_2}_{0.7}, \preceq)$ (right).

If we build the r-irreducible 0.7-cut concept lattice of $(\mathcal{M}^{Y_2}, \preceq)$, we need to consider:

$$\widehat{M_F}(Y_2)_{0.7} = \{\langle \phi^{\downarrow}_{a_4,1.0}, \phi^{\downarrow\uparrow}_{a_4,1.0} \rangle, \langle \phi^{\downarrow}_{a_1,1.0}, \phi^{\downarrow\uparrow}_{a_1,1.0} \rangle, \langle \phi^{\downarrow}_{a_4,0.8}, \phi^{\downarrow\uparrow}_{a_4,0.8} \rangle, \langle \phi^{\downarrow}_{a_1,0.8}, \phi^{\downarrow\uparrow}_{a_1,0.8} \rangle\}$$

Therefore, the reduced concept lattice is shown in the right side of Figure 2.

It is worth to note that the size reduction depends on the chosen reduct. For example, if we consider $Y_1$ we only remove the concepts $C_{17}$ and $C_{20}$, whereas

[1] Observe that, only the extension of the concept has been represented in each node since the intension does not coincide with the original intension of the concept because the number of attributes has been reduced.

taking into account the reduct $Y_2$ we also remove the concept $C_{12}$. This is due to the fact that $C_{12}$ is obtained from the fuzzy-attributes $\phi_{a_5,1.0}$ and $\phi_{a_6,0.6}$. Hence, if we regard $Y_2$ in the construction process then the fuzzy-attribute related to $C_{12}$ does not exceed the cut established by $\alpha = 0.7$.

Note that, if we had chosen a value for $\alpha$ less or equal to 0.6 then both reductions would have coincided.

## 4 Conclusions and further works

This paper has combined the attribute reduction mechanism given in [6] and the size reduction procedure presented in [5] in the multi-adjoint concept lattice environment. We have proven that if we reduce the number of attributes and then we carry out the meet-irreducible $\alpha$-cut concept lattice, the size of the reduced concept lattice depends on the reduct chosen from the attribute classification.

In the future, we will analyze the results of combining both reduction mechanism in order reversing and we will study how to obtain the relation between attributes and objects from the meet-irreducible $\alpha$-cut concept lattice in order to build a concept lattice isomorphic to the reduced concept lattice by the cut.

## References

## References

1. R. Bělohlávek and V. Vychodil. Reducing the size of fuzzy concept lattices by hedges. In *The 2005 IEEE International Conference on Fuzzy Systems*, pages 663–668, 2005.
2. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. A comparative study of adjoint triples. *Fuzzy Sets and Systems*, 211:1–14, 2013.
3. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Irreducible elements in multi-adjoint concept lattices. In *Intl Conference on Fuzzy Logic and Technology, EUSFLAT 2013*, pages 125–131, 2013.
4. 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.
5. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Fuzzy-attributes and a method to reduce concept lattices. *Lecture Notes in Computer Science*, 8536(0):189 – 200, 2014.
6. M. E. Cornejo, J. Medina, and E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294(0):41 – 56, 2015.
7. X. Kang, D. Li, and S. Wang. Research on domain ontology in different granulations based on concept lattice. *Knowledge-Based Systems*, 27(0):152–161, 2012.
8. J. Konecny. Isotone fuzzy galois connections with hedges. *Information Sciences*, 181(10):1804–1817, 2011.
9. J. Konecny, J. Medina, and M. Ojeda-Aciego. Intensifying hedges and the size of multi-adjoint concept lattices with heterogeneous conjunctors. In *The 9th International Conference on Concept Lattices and Their Applications*, pages 245–256, 2012.

10. L. Li and J. Zhang. Attribute reduction in fuzzy concept lattices based on the T-implication. *Knowledge-Based Systems*, 23(6):497–503, 2010.

# Proving termination with multiset orderings in PVS: theory, methodology and applications ⋆

José A. Alonso, María J. Hidalgo, and Francisco J. Martín–Mateos
{jalonso, mjoseh, fjesus}@us.es

Departamento de Ciencias de la Computación e Inteligencia Artificial.
Escuela Técnica Superior de Ingeniería Informática, Universidad de Sevilla
Avda. Reina Mercedes, s/n. 41012 Sevilla, Spain

**Abstract.** There exist a number of non-trivial termination proofs of functions (or algorithms) which are carried out more naturally and simpler using well-founded multiset orderings. We present in this paper a methodology to organize and simplify these kind of termination proofs in the PVS specification and verification system. This methodology uses a well–known result due to Dershowitz and Manna, which states that every well–founded relation on a set $T$ can be extended to a relation on finite multisets over $T$ which is also well-founded. Therefore, we also present a formalization of this theorem in PVS. We think this methodology can be very useful to develop non-trivial termination proofs in PVS. To illustrate this, we have applied our methodology to formalize in PVS some termination proofs, like an iterative version of the Ackermann's function and the McCarthy's 91 function.

## 1 Introduction

The use of well–founded orders for proving termination of recursive functions was suggested by Floyd in [4]. The idea is to find a set $T$, with a well–founded order $<$ and a measure function $m$ mapping the arguments of the function into the elements of $T$, such that the measure of the arguments is reduced in each recursive call. Due the well–foundedness of $<$, this measure can not decrease indefinitely and hence, the termination of the function is assured.

The most used well–founded order is the usual order on natural numbers and the lexicographic order on n–tuples of natural numbers. However, Dershowitz and Manna [3] showed that every well–founded relation on a set $T$ can be extended to a well–founded relation on the finite multisets over $T$. They also proved that the use of multiset orderings allows to construct simple and intuitive measure functions to carry out non–trivial proofs of termination. In particular, they showed that the multiset ordering can be used to prove the termination of Ackermann's function, McCarthy's 91 function and production systems, programs defined in term of rewriting rules.

⋆ This work was partially supported by TIC-6064 Excellence project (Junta de Andalucía) and TIN2013-41086-P project (Spanish Ministry of Economy and Competitiveness), cofinanced with FEDER funds.

On the other hand, in the field of formal verification, we often have to tackle the problem of proving termination of programs, logic reasoning systems or rewriting systems by using one of the current systems like ACL2, COQ, HOL, Isabelle, PVS,.... These proofs of termination are non–trivial and, in many cases, the use of multiset orderings can be very useful. Thus, in order to mechanize proofs of termination in a formal system by using multiset orderings, it would be appropriated to formalize these orderings in the corresponding system. In fact, when we formalized in PVS a tableaux algorithm for the $\mathcal{ALC}$ description logic [1], multisets were a key tool for proving its termination. Thus, we think it is useful to hold in PVS a theory of wellfoundeness of multiset orderings, in order to be able to carry out this kind of termination proofs in this system.

In this paper we present a methodology to systematize the proofs of termination of recursive functions using multiset orderings and we show two case studies where we have used this methodology for proving termination of different tail recursive functions. We use the Dershowitz and Manna theorem [3], that we have proved in an abstract way, allowing its instantiation to prove the well–foundedness of particular multiset relations.

We will remark the main features of the PVS system according as we will use them and we will explain both the definitions and the PVS proofs in a form that can be understood without being an expert in PVS. However, a detailed description of this system can be seen in [7]. Moreover, due to the lack of space, we will skip details of the proofs. Nevertheless, the whole formalization is available at `http://www.cs.us.es/~mjoseh/PROTEMO/`

## 2 Methodology for proving termination with multiset orderings

In order to define a recursive function in PVS with domain $D$ and range $R$, a measure function must be provided, along with an optional well-founded relation. The measure should be a function whose signature matches that of the recursive function, but with range type the domain of the order, which defaults to $<$ on $\mathbb{N}$ or on the ordinals. If an ordering $<_D$ is provided, then it must be a binary well-founded relation. Thus, when a recursive function is defined in PVS, several proof obligations (called TCCs) are generated to prove that $<_D$ is well-founded and to prove that the arguments of $f$ decrease with respect to $<_D$ in each recursive call.

According to the work of Dershowitz and Manna [3], multiset orders can be used to prove the termination of a recursive function $f$ in the following way: a measure function should be defined mapping each element of the domain $D$ to a finite multiset over a well-founded set $(T, <_T)$, and a relation on $D$ should be considered such that two elements are related if and only if their measures are related with respect to the multiset relation induced on $M(T)$ by $<_T$. The well-founded set $(T, <_T)$ will depend on each specific function $f$.

For the purpose of formalizing this kind of proofs in PVS, we have proved in PVS (following [3]) the well-foundedness of the multiset relation induced by a

well-founded relation, in a way that we can easily use it for any relation, simply by instantiating the parameters of the PVS theory.

**Theorem 1.** *Let $<$ be a transitive and well–founded relation on $T$. Then the relation $<_{mult}$ is a well–founded relation on $\mathcal{M}(T)$.*

Previously, and using the definition of well-foundedness based on the notion of minimal element, we have proved in PVS the well-foundedness of every relation which can be embedded in a well-founded relation by a monotonic function.

**Theorem 2.** *If $f : (T, <) \to (T', <')$ is monotone and $(T', <')$ is well–founded, then $(T, <)$ is well–founded.*

Thus, the idea presented above can be clarified in the following methodology: given a recursive function $f : D \to R$

**Step 1** Consider an appropiate set $T$ and build a well-founded relation on it. If $T$ is the set of natural numbers or the set of ordinals, then the relation could be the usual order $<$, whose well-foundedness property is ensured by the prelude of PVS. In other cases, we can prove that $(T, <_T)$ is well-founded by building a monotonic function over a known well-founded set $(T', <_{T'})$ and applying Theorem 2.

**Step 2** Apply the Dershowitz–Manna theorem (1), assuring that $(\mathcal{M}(T), <_{mult})$ is well–founded. To do this in PVS, it is enough to instantiate the parameters of the PVS theory `finite_bags_order` with $T$ and $<_T$. Then, the corresponding theorem `less_mult_is_wf` is automatically proved.

**Step 3** Define an adequate measure function `f_measure` : $D \to \mathcal{M}(T)$.

**Step 4** Define in $D$ the relation $<_D$ induced by the measure function:

$$x <_D y \Leftrightarrow \texttt{f\_measure}(x) <_{mult} \texttt{f\_measure}(x)$$

It should be noted that, in this way, the measure function is monotone. Therefore, Theorem 2 proves automatically that $<_D$ is a well–founded relation.

**Step 5** Use the relation $<_D$ as the well–founded relation needed to prove the termination of function $f$.

It should be noted that steps 2 and 4 can be considered totally mechanized, since it is only necessary to instantiate the parameters of a PVS theory. Nevertheless, steps 1, 3 and 5 are specific to each function.

## 3 Case studies

In this section, we present some examples in which the suggested methodology has been used in PVS to prove non–trivial termination properties.

### 3.1 Ackermann's function

A tail recursive function which computes Ackermann's function is $\texttt{A\_it}(m,n) = \texttt{A\_it\_aux}((m),n)$, where

$$\texttt{A\_it\_aux}(S,z) = \begin{cases} z & \text{if } S = () \\ \texttt{A\_it\_aux}((s_2,\ldots,s_k), z+1) & \text{if } S = (0,s_2,\ldots,s_k) \\ \texttt{A\_it\_aux}((s_1-1,s_2,\ldots,s_k),1) & \text{if } s_1 \neq 0 \wedge z = 0 \\ \texttt{A\_it\_aux}((s_1,s_1-1,s_2,\ldots,s_k), z-1) & \text{in other case} \end{cases}$$

where $S = (s_1,\ldots,s_k)$ is a stack such that in every step

$$\texttt{A\_it\_aux}(S,z) = \texttt{A}(s_k, \texttt{A}(s_{k-1},\ldots,\texttt{A}(s_1,z)))$$

The PVS specification of this function is the following

```
A_it_aux (p): RECURSIVE nat =
 LET S = proj_1(p), z = proj_2(p) IN
 IF null?(S) THEN z
 ELSE LET s1 = car(S), S2 = cdr(S) IN
      IF s1 = 0 THEN A_it_aux ((S2,z+1))
                ELSIF z = 0 THEN A_it_aux ((cons (s1-1,S2),1))
                            ELSE A_it_aux ((cons(s1,cons(s1-1,S2)),z-1))
      ENDIF
 ENDIF MEASURE p BY less_measure

A_it(m,n): nat = A_it_aux(((: m :), n))
```

Let us note that, in this specification, the relation `less_measure` is not yet determined. In [3], a proof of termination of this function using a multiset measure has been shown. In this case, the measure function maps a pair $((s_1,\ldots,s_k),z)$ into the multiset of pairs of natural numbers $\dot{\{}(s_1,z),(s_2+1,0),\ldots,(s_k+1,0)\dot{\}}$.

In order to build in PVS the appropriate relation `less_measure` and prove its properties, we carry out the following steps, according to the explained methodology.

**Step 1** In this case, $T = \mathbb{N}\times\mathbb{N}$ and $<_T$ is the lexicographic order. We prove that it is a well–founded relation by defining a monotone function into ordinals.

**Step 2** We prove that the extension to $\mathcal{M}(\mathbb{N}\times\mathbb{N})$ of the lexicographic order on $\mathbb{N}\times\mathbb{N}$ is a well–founded ordering, by instantiating the parameters of the PVS theory by $\mathbb{N}\times\mathbb{N}$ and the lexicographic order $lex$.

**Step 3** We define the *measure function* from the pairs $(S,z)$ in $\mathcal{M}(\mathbb{N}\times\mathbb{N})$

```
a_measure(p): finite_bag[[nat,nat]] =
  IF null?(S) THEN  emptybag
              ELSIF length(S)=1 THEN insert((car(S),z), emptybag)
              ELSE  insert((car(S),z), list_mult(cdr(S)))
  ENDIF
  WHERE S = proj_1(p), z = proj_2(p)
```

**Step 4** From step 3, we have directly proved, by Lemma 2, that the ordering induced by *measure function* between the arguments of `A_it_aux`, is well–founded

```
a_less(p1,p2): bool = less_mult(a_measure(p1),a_measure(p2))

a_less_wf: COROLLARY well_founded?[[list[nat],nat]](a_less)
```

**Step 5** So, this is the function we need as well-founded order in the definition of `A_it_aux`. We prove the proof obligations automatically generated by PVS to ensure that the arguments of the function decrease in each recursive call.

### 3.2 McCarthy's 91 function

McCarthy's 91 function is a recursive function defined by John McCarthy in [5] as

$$M(n) = \begin{cases} n - 10, & \text{if } n > 100 \\ M(M(n+11)), & \text{if } n \leq 100 \end{cases}$$

This function returns 91 for all $n \leq 101$ and $n - 10$ for $n > 101$. There are some papers [3, 2, 6] that address termination proofs of this function. Our goal here is to show how we use the methodology of multiset orderings in PVS to prove termination of an iterative version of McCarthy's 91 function. We follow the same steps as for Ackermann's function. First, we specify the function in PVS

```
mc_it_aux(p): RECURSIVE nat =
  LET n = proj_1(p), z = proj_2(p) IN
   IF n = 0 THEN z
            ELSIF z > 100 THEN mc_it_aux(n-1,z-10)
                          ELSE mc_it_aux(n+1,z+11)
   ENDIF MEASURE p BY less_measure

mc_it (x) :nat = mc_it_aux((1, x))
```

We will consider the termination problem of the function `mc_it_aux`. This task is not trivial due to the behavior of the second recursive call. In [3] a multiset measure is given to ensure termination of this function: every pair $(n, z)$ is measured by the finite multiset $\dot{\{}z, mc(z), mc^2(z), \ldots, mc^{n-1}(z)\dot{\}}$. Let us note that `mc_it_aux`$(n, z) = mc^n(z)$, and that the relation to compare multisets is the multiset relation induced by the following well–founded relation in $\mathbb{N}$:

$$m <_{mc} n \Leftrightarrow n < m \leq 111$$

In this case, we define in $\mathbb{N}$ the relation $m <_{mc} n = n < m \wedge m <= 111$ and we prove that $(\mathbb{N}, <_{mc})$ is well–founded, using Lemma 2. We prove that the multiset relation induced in $\mathcal{M}(N)$ by $<_{mc}$ is well-founded, by instantiating the parameters of the PVS theory `finite_bags_order` with $\mathbb{N}$ and $<_{mc}$. We define the *measure function* from $\mathbb{N} \times \mathbb{N}$ to $\mathcal{M}(\mathbb{N})$ and the order induced by this measure

function. Then, we have automatically proved that it is a well–founded relation. Finally, we use the relation `mc_it_less` as the well-founded relation needed to ensure the termination of the function `mc_it_aux` and we prove the proofs obligations generated to ensure that the arguments of the function decrease in each recursive call.

## 4 Conclusions and future work

We have presented a methodology to organize and simplify termination proofs which use well–founded multiset orderings. The main utility of this methodology is given by the easy way to prove the well–foundedness of a multiset relation. It is enough to instantiate the parameters of a PVS theory and, automatically, a corollary with the expected result is obtained. The non–mechanized part is to define the measure function, that it is specific for each function, and to prove that the arguments of the function decrease in each recursive call. In [1] we have used this methodology to prove in PVS the termination of a tableau algorithm for the $\mathcal{ALC}$ logic. This measure function is more complex than the previous ones, since in this case, when a kind of rule (universal rules) is applied, this one is not disabled forever, but it can be reapplied due to the introduction of new individuals by subsequent applications of another kind of rule (existential rules).

Finally, we would like to point out two lines for future work. First, in order to make the methodology more automatic, we would like to develop PVS strategies to increase the mechanization of the process. Second, we would like to apply this methodology to prove termination of rewriting systems or tableau algorithms, in which a measure in multisets was required.

## References


1. J. A. Alonso, J. Borrego, M. J. Hidalgo, F. J. Martín-Mateos, and J. L. Ruiz-Reina. A Formally Verified Prover for the ALC Description Logic. In *Theorem Proving in Higher Order*, volume 4732 of *Lecture Notes in Computer Science*, pages 135–150. Springer, 2007.
2. J. Cowles. Knuth's Generalization of McCarthy's 91 Function. In *Computer–Aided reasoning: ACL2 case studies*, pages 283–299. Kluwer Academic Publishers, 2000.
3. N. Dershowitz and Z. Manna. Proving Termination with Multiset Orderings. *Communications of the ACM*, 22(8):465–476, 1979.
4. R. W. Floyd. Assigning Meanings to Programs. *Mathematical Aspects of Computer Science*, 19:19–31, 1967.
5. Z. Manna and J. McCarthy. Properties of Programs and Partial Function Logic. *Machine Intelligence*, 5:27–37, 1970.
6. F. J. Martín-Mateos, J. A. Alonso, M. J. Hidalgo, and J. L. Ruiz-Reina. A Generic Instantiation Tool and a Case Study: A Generic Multiset Theory. In *Third International Workshop on the ACL2 Theorem Prover and its Applications (ACL2 '02)*, April 2002.
7. S. Owre, N. Shankar, J. M. Rushby, and D. W. J. Stringer-Calvert. PVS Language Reference. Technical report, Computer Science Laboratory, SRI International, 1999.

# Correct Application of Mutation Testing to the C++ Language

Pedro Delgado-Pérez, Inmaculada Medina-Bulo, Juan José Domínguez-Jiménez

Department of Computer Science and Engineering
University of Cadiz, Spain
{pedro.delgado,inmaculada.medina,juanjose.dominguez}@uca.es

**Abstract.** Success of mutation testing greatly depends on the mutation operators defined. As a white-box technique, selecting specific mutants for each language addressed is necessary, but it should be accompanied by an implementation focused on the particular details of the language. Only then we will be able to undertake a correct application of the technique, obtaining exactly the mutants that should be generated. This paper shows different C++-specific features that a mutation tool for this language should take into account with a twofold goal: creating valid but also useful mutants. Refining the implementation may reduce the computational cost of mutation testing application and enhance the effectiveness of mutation operators.

**Keywords**: Mutation testing, mutation operators, C++

## 1 Introduction

Mutation testing allows us to determine the adequacy of a test suite revealing several syntactic faults in our program. This technique can also be used to improve the test suite by introducing new test cases that are able to distinguish these faulty versions from the original program [7,9]. This difference can be found in the results of their executions. The faults, commonly known as mutations, are introduced in the program via the *mutation operators* defined for a certain language. For example, a mutation operator replacing relational operators may transform $x > 1$ into $x < 1$ to create a *mutant*. If a test suite detects the mutation, we say that the mutant is dead; otherwise, the mutant is either alive or is equivalent to the program under test.

Most works in literature deal with mutation operators from a high-level perspective based on their definition. On the contrary, there are few papers studying in depth their implementation as this is not such an interesting matter to research for being considered a technical detail. However, we claim that the correct implementation of operators is of a great importance in the application of mutation testing. Being this a white-box technique, automating mutation operators tailored to the specifics of a language is a key factor to produce the mutants as expected.

In this paper, we aim to show how, in addition to a suitable operator definition, the particularities of a mainstream language like C++ may affect the number of mutants generated and the mutation operator effectiveness. Section 2 addresses the mutation operators which can be created for this language and the method to put them into practice to achieve a high capacity for analysis of the code. Section 3 lists different C++ features impacting the number of mutants generated and how attending to them we can avoid mutants not useful for the purpose of mutation testing. The last section presents the conclusions and the future work to accomplish.

## 2 Mutation Operators in C++

### 2.1 Definition of Mutation Operators

Mutation operators mainly represent typical faults made by programmers. They are therefore obtained from the analysis of the most common mistakes in the development of applications in a certain language, representing the faults that the technique will treat within the code. Several faults are common to many general purpose languages, but each language possesses certain features making a specific study necessary. Thus, different works have been prepared to define a set of operators for a great range of languages [4]. In any case, we deem that the entire development of the technique should follow the same path so that it is feasible to compare every contribution in this field.

Mutation testing has been applied at different levels of the language. The level that the user chooses depends on the program characteristics and the type of testing to accomplish [4]. As a mainstream language, a set of mutation operators can be defined for each of these levels:

- **Unit level**: It refers to the traditional mutation applied to a single function or method to test its functionality. These operators are usually known as *traditional operators*.
- **Class level**: This level deals with the mutation of object-oriented features, modifying both declarations and expressions relating a class.
- **Integration level**: Intermediate level between the unit and the system levels. Function invocations are tested, making changes from parameters to values returned by the functions.
- **Multi-class or system level**: The operators at this level are intended to test a complete program, checking from interactions among classes to the user interface [6].

The mutation operators for C++ at the unit level take as starting point the research around C, taking into account the similarity of both languages. We can make the appropriate modifications to adapt the operators to C++ [1]. As for the class level, a complete set of class mutation operators were defined [2]; we adapted several operators from Java and C#, and also defined new operators regarding C++ features which had not been studied yet. In addition to these

levels, we can add new levels focused on specific properties of the language frequently used. Regarding C++, this level would mainly correspond to the C++ Standard Library, which provides a great amount of facilities (e.g. stream input/output). Hence, Kusano et al. [5] applied mutation testing for concurrency constructs in C++ applications.

### 2.2 Technique to Inject Mutations in C++

One of the traditional approaches when analyzing the code to inject mutations is by means of a pattern matching based on regular expressions over the code. However, this method presents some limitations and may not be sufficient to achieve the expected result. As a simple example, when using this approach we have to check that each mutation location is not within a line or block comment before applying the mutation. In addition, the analysis of C++ code is not easy, especially because of its huge grammar. In case of similar languages to C++ like Java, mutations have been inserted directly on the bytecode or introspection has been also used. However, we have to discard these options in the case of C++ as they are not available for the user.

Thus, the most robust and comprehensive approach to apply mutation testing to C++ is resorting to the abstract syntax tree (AST) generated with a full-fledged compiler for this language [2]. This tree structures the code and determines the relations among the elements. Therefore, the AST allows us to properly analyze and solve complex situations in this language, as the ones listed in Section 3.1. Moreover, we will be able to undertake a much more fine-grained control of the potential mutation locations as well as to create mutants complying with the grammar rules.

## 3 Implementation Criteria

### 3.1 Creating Tailored Mutants to C++ Particularities

The operator implementation mode explained in Section 2.2 is definitely challenging, as stated by Derezińska [3]: it is necessary to analyze the AST and establish a margin for each mutation operator, which sets the different situations where an operator can be applied or not. Notwithstanding, the insertion of the faults can be controlled in an accurate manner as we can grasp from the AST well-defined elements.

Nevertheless, we need to take into account several aspects related with the C++ characteristics so that mutations meet expectations. These features have a direct effect on the number and kind of mutants created. An illustrative enumeration of various considerations are described below:

- **Class and structure**: For compatibility with C, classes were introduced as an extension of the structures so that, in addition to data members, function members and operators could be added as well. However, members in a class are hidden by default while structures are visible from outside. This is, in

fact, the only difference between the keywords *struct* and *class*. Therefore, the set of operators applicable to classes can also perform on structures when possible.

- **Operator overloading property**: Most operators of this language can be redefined, giving them different semantics depending on the operand types. This feature needs to be considered in mutation testing as it implies the possibility of using operators of the language with user-defined types instead of defining new standard methods. At the same time, for instance, the traditional mutation operator replacing arithmetic operators may have changed its usual mathematical meaning when applied to user-defined objects. Nonetheless, the AST represents the invocations to an overloaded operator with a special kind of node, which is different from the one for traditional method callings. Hence, this aspect can be solved when implementing mutation operators related with this matter.
- **Typedefs and namespaces:** Evaluating types when searching for mutation locations in the code is an essential task. For instance, if a mutation operator replaces variables of the same type, a comparison of the types of the involved variables should be undertaken previously. The *typedef* keyword allows us to give a new name to a concrete type. Therefore, if we do not take this feature into account, some mutants may be overlooked.
  Also in this regard, several declarations can be grouped together in a *namespace*. We can declare similar elements in the code provided that they are in different namespaces. Thus, we need to properly qualify the references to declarations in namespaces to avoid confusion.
- **Definition and declaration:** Several elements in C++, as functions and methods, can be declared at any moment, but defined further in the code. This distinction is important in two aspects. Firstly, both the declaration and the definition should be modified if a mutation operator changes the signature or the value returned of a function or method. Secondly, we can invoke an element which has not been defined yet but has been declared previously.

### 3.2 Creating Useful Mutants

The operator implementation is usually a complex task that requires follow-up work. This implementation becomes more difficult when some conditions are imposed on the operators to prevent the creation of uninteresting mutants, i. e., mutants which do not help us assess the adequacy level of a test suite. However, this fact could allow for a reduction of mutants and, consequently, of the computational cost of the technique, which is a major concern when using mutation testing. Specific rules for several operators to cut out unnecessary mutants has been shown for Java [8].

We consider that the following kinds of mutants should be prevented as much as possible. Each of these kinds are accompanied by examples closely related to the C++ features:

- **Invalid mutants:** The executable programs created from the mutated code cannot be compiled or linked. An example of mutation that always produces an invalid mutant is deleting the initialization of reference type variables, as they must be initialized when declared.
- **Equivalent mutants:** There is not any input which is able to detect a difference between the original program and the mutant. For instance, the *PVI* operator [2] inserts the *virtual* keyword in methods which are not marked with this modifier. The method however may be already virtual although it is not marked as virtual. This happens when the method is overriding a virtual method in a base class. This mutation would produce an equivalent mutant (see Figure 1).
- **Trivial mutants:** The difference between the original and the mutant version is found by any input exercising the mutation. The usage of *templates* sometimes produces this type of mutants. For instance, some errors in a class template do not emerge until an object of the class template is created (whatever the type used), such as a constant data member which had not been initialized in a constructor.

**Original classes:**

```
class A{                    class B: public A{   class C: public B{
  ... ...                     ... ...              ... ...
  virtual void m(){...}       void m(){...}        void m(){...}
};                          };                   };
```

**Mutant:**

```
class B: public A{
  ... ...
  virtual void m(){...}
};
```

**Equivalent:**

The method 'm' in class B is virtual with or without including the *virtual* keyword because the method 'm' in A is already marked as *virtual*.

**Fig. 1.** Example of equivalent mutant in the *PVI* operator

## 4 Conclusion and Future Work

A correct implementation of the mutation operators beyond their narrow definitions is a decisive step towards a fruitful application of mutation testing. In this regard, the specifics of each language should be handled because the grammar rules may influence the kind of mutants generated with a mutation operator. The AST analysis is really useful to fulfil with the restrictions imposed on the application of the mutation operators. These restrictions, mainly based on the particular features of the language, may avoid generating mutants which are not useful, thereby reducing the high computational cost of the technique.

As future work, we intend to obtain a complete list of C++ features which can affect the mutation operators at the unit and the class level. This collection can be useful to develop mutation tools addressing this language so that the study of the results of different tools is somewhat comparable. Likewise, we aim to implement different mutation operators using rules to reduce equivalent, invalid and trivial mutants and then evaluate the impact of such improvement in the cost and effectiveness of the technique.

*Acknowledgement:* This paper was funded by the research scholarship PU-EPIF-FPI-PPI-BC 2012-037 of the University of Cadiz.

## References

1. Delgado-Pérez, P., Medina-Bulo, I., Domínguez-Jiménez, J.J.: Analysis of the development process of a mutation testing tool for the C++ language. In: The Ninth International Multi-Conference on Computing in the Global Information Technology, ICCGI 2014. Seville, Spain (2014)
2. Delgado-Pérez, P., Medina-Bulo, I., Domínguez-Jiménez, J.J., García-Domínguez, A., Palomo-Lozano, F.: Class mutation operators for C++ object-oriented systems. Annals of telecommunications 70(3-4), 137–148 (2015)
3. Derezinska, A., Kowalski, K.: Object-oriented mutation applied in common intermediate language programs originated from C#. In: Software Testing, Verification and Validation Workshops, IEEE 4th International Conference on. pp. 342–350 (2011)
4. Jia, Y., Harman, M.: An analysis and survey of the development of mutation testing. Software Engineering, IEEE Transactions on 37(5), 649 –678 (Oct 2011)
5. Kusano, M., Wang, C.: Ccmutator: A mutation generator for concurrency constructs in multithreaded C/C++ applications. In: Automated Software Engineering (ASE), 2013 IEEE/ACM 28th International Conference on. pp. 722–725. IEEE (2013)
6. Mateo, P.R., Usaola, M.P., Offutt, J.: Mutation at the multi-class and system levels. Science of Computer Programming 78(4), 364–387 (2013), special section on Mutation Testing and Analysis (Mutation 2010) and Special section on the Programming Languages track at the 25th ACM Symposium on Applied Computing
7. Offutt, J.: A mutation carol: Past, present and future. Information and Software Technology 53(10), 1098–1107 (2011), http://www.sciencedirect.com/science/article/pii/S0950584911000838, special Section on Mutation Testing
8. Offutt, J., Ma, Y.S., Kwon, Y.R.: The class-level mutants of MuJava. In: Proceedings of the 2006 International Workshop on Automation of Software Test. pp. 78–84. AST '06, ACM, New York, NY, USA (2006), http://doi.acm.org/10.1145/1138929.1138945
9. Woodward, M.R.: Mutation testing - its origin and evolution. Information and Software Technology 35(3), 163–169 (Mar 1993)

# Application of Fuzzy Signature State Machines in Renovation Process of Residential Blocks for Supporting Cost Optimization

Gergely I. Molnárka[1], Miklós F. Hatwágner[2], and László T. Kóczy[2]

[1] Department of Building Constructions and Architecture, Széchenyi István University, Egyetem tér 1, Győr, Hungary
[2] Department of Information Technology, Széchenyi István University, Egyetem tér 1, Győr, Hungary

**Abstract.** Due to the lack of periodic maintenance processes, several important renovation steps have to be realized in the near future in the pre-war residential houses in the Historic districts of Budapest, Hungary. Most of these steps causes enormous outcome for the owners, therefore these interventions have to be planned carefully both from engineering and economic aspects.
As a combination of the fuzzy signature structure and the principles of finite-state machines a new formal method is proposed for generating a tool for supporting the desired renovation planning, concerning the financial costs and reasonable sequence of work phases. With the support of information obtained by detailed building diagnostic analyses of the pre-war urban-type residential houses, generally accepted and applied technical instructions of the building renovation processes and the contractors' billing database the renovation procedure of the plinth zone (damp proof course, drainage connections and renderings) of a block of residential houses is presented as a case study.

**Keywords:** urban-type residential house, building renovation, fuzzy signature, fuzzy state machine

## 1 Introduction

The residential building stock of Budapest (Hungary) was nationalized after World War II, therefore theoretically the state administration took the responsibility for its maintenance. By the time of the re-privatization of the tenement houses it became clear that the designated state organizations were incompetent for this task because of three reasons: lack of financial funds, organizational deficiencies and lack of any real motivation and intentions (due to some ideological causes). For this reason, the owners (basically, the former tenants) inherited their residential houses in very bad condition that is still clearly visible and (tactile) in the historic districts. This problem particularly affects the tenement houses built before 1920 with traditional constructions and components.

The ownership communities of these houses have several alternatives for handling the current situation. Presuming that the owners feel themselves responsible for the maintenance of their properties, they may organize the renovation process in different ways. Considering the status of the given building, the main solutions can be the complete renovation without any cease, or step-by-step. Whereas it is generally assumed that the owners do not have adequate financial funding for for realizing any quick and proper renovation without external aid, the sequential intervention may be more appealing for the inhabitants. In this case any renovation procedure has to be planned properly in advance from the first work phase to the end, both financially and technically.

There are several factors influencing the schedule of the rehabilitation process. Without any details, the most important factors are: the *grade of danger* caused by the observed failure, the *interrelations* among deteriorated building components and decays, and the presence of a protocol for the repair, the *financial timing*, some complex *logistics aspects*, etc.

Due to the technical characteristics of the buildings, any house may be considered as a hierarchically ordered set of different building components where the state of each component and set of components influence the state of the complete building. This feature was advantageously used in a comparative analysis before [1].

## 2 A Model Proposed for Supporting the Renovation Process

As a summary of these statements the followings are ascertainable:

1. The physical condition of any building can be characterized by a continuous state scale.
2. The building is constituted by a hierarchically ordered component structure.
3. The state transition of the building is continuous in time; decay and renovation phases may slightly modify the transition.

The proposed approach to be introduced here for handling the above problem is a model and an attempt for a solution based on the principles of fuzzy signatures and fuzzy state machines combined into what we call *Fuzzy Signature State Machines (FSSM)*.

The most important reason for using fuzzy signatures here as the starting point is the fact that the structure of building diagnostic surveys follows the architectural and civil engineering common sense, where the sub-structures and components of each building are arranged in hierarchical tree-like structures. In this model the whole building might be presented by the root of the tree and each major sub-component is a first level branch, with further sub-branches describing sub-sub-components, etc. as it will be shown in the next sections.

At this point the definition of fuzzy signature should be recalled.

Vector Valued Fuzzy Sets (VVFS) [2] are a simple extension of the fuzzy sets that may be considered as a special case of L-fuzzy sets [3]:

$$A_n = \{X, \mu_{A_n}\}, \text{ where } \mu_A n : X \to [0,1]^n \tag{1}$$

Thus a membership degree is a multi-component value here, e.g. $[\mu_1, \mu_2, \cdots, \mu_n]^T$. Fuzzy signatures (FSig) represent a further extension of VVFS as here any component might be a further nested vector, and so on [5], [4]:

$$\begin{aligned} A_{fs} = \{X, \mu_{A_{fs}}\}, \text{where } \mu_{A_{fs}} : X \to M_1 \times M_2 \times \cdots \times M_n, \\ \text{where } M_i = [0,1] \text{ or } [M_{i_1} \times M_{i_2} \times \cdots \times M_{i_n}]^T \end{aligned} \tag{2}$$

The following simple example illustrates the structure of a simple fuzzy signature:

$$\mu_{A_fs} = [\mu_1, \mu_2, [\mu_{3_1}, \mu_{3_2}, [\mu_{3_{3_1}}, \mu_{3_{3_2}}, \mu_{3_{3_3}}]], \mu_4, [\mu_{5_1}, \mu_{5_2}], \mu_6]^T \tag{3}$$

The advantage of using fuzzy signatures is that here any closer grouping and subgrouping of fuzzy features may be given by the tree structure. Fuzzy signatures are associated with a set of aggregation. Each sub-component set may be aggregated by its respective aggregation operation, thus reducing the sub-component to one higher level.

The above example has the following associated aggregation structure: $\{a_0\{a_3\{a_{3_3}\}\{a_5\}\}$, where each $a_\circ$ denotes an aggregation, particularly the one associated with the child node $x_\circ$ associated with $\mu_\circ$, thus the following example signature might be reduced "upwards" to the root as follows:

$$\begin{aligned} \mu_{A_fs} \Rightarrow [\mu_1, \mu_2, [\mu_{3_1}, \mu_{3_2}, \mu_{3_3} = a_{3_3}(\mu_{3_{3_1}}, \mu_{3_{3_2}}, \mu_{3_{3_3}}), \mu_4, \mu_5 = \\ = a_5(\mu_{5_1}, \mu_{5_2}), \mu_6]^T \Rightarrow \\ \Rightarrow [\mu_1, \mu_2, \mu_3 = a_3(\mu_{3_1}, \mu_{3_2}, \mu_{3_3}), \mu_4, \mu_5, \mu_6]^T \Rightarrow \\ \Rightarrow \mu_0 = a_0(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) \end{aligned} \tag{4}$$

The operations among fuzzy signatures with partially different structure may be carried out, by finding the *largest common sub-structure* and reducing all signatures up to that substructure. This might be necessary if the surveys referred to this paper are considered as often their depth and detail are different. As an example, maybe in "Survey A" the plinth zone is considered as a single component of the house and is evaluated by a single linguistic quality label, while in "Survey B" this is done in detail, and the damp proofing system, casing and the plastering are are described separately.

In our previous work we applied VVFS and FSig [6] for describing sets of objects with uncertain features, especially when an internal theoretical structure of these features could be established. In [7] we presented an approach where the fuzzy signatures could be deployed for describing existing residential houses in order to support decisions of local authorities concerning when and how these buildings should be renovated involving non-measurable (and subjective) factors. In that research a series of theoretically arrangeable features were taken into

consideration and eventually a single aggregated fuzzy membership value could be calculated on the basis of available detailed expert evaluation sheets. In that model, however, the available information does not support any decision strategy concerning actual sequence of the measures leading to complete renovation; and it is also insufficient to optimize the sequence from the aspect of local or global cost efficiency. In the following section, the mathematical model of the proposed maintenance protocol will be introduced.

Finite State Machines are determined by the sets of input states $X$, internal states $Q$, and the transition function $f$ (Moore model). The latter determines the transition that will occur when a certain input state change triggers a state transition. For simplicity the following is assumed as the starting point of our new model:

$$A = \langle X, Q, f \rangle \tag{5}$$

$$f : X \times Q \to Q, \text{ where } X = \{x_i\} \text{ and } Q = \{q_i\} \tag{6}$$

Thus, a new internal state is determined by the transition function as follows:

$$q_{i+1} = f(x_i, q_i) \tag{7}$$

In matrix form:

$$F = \begin{bmatrix} f(x_1, q_1) & f(x_2, q_1) & \dots & f(x_n, q_1) \\ f(x_1, q_2) & & & f(x_n, q_2) \\ \vdots & & & \vdots \\ f(x_1, q_m) & f_(x_2, q_m) & \dots & f(x_n, q_m) \end{bmatrix} \tag{8}$$

The transition function/matrix maybe interpreted with help of a relation $R$ on $X \times Q^2$,where

$$R(x_i, q_j, q_k) = 1, \text{ if } f(x_i, q_j) = q_k \tag{9}$$

and

$$R(x_i, q_j, q_k) = 0, \text{ if } f(x_i, q_j) \neq q_k \tag{10}$$

In the present application an extension to fuzzy states is considered in the following sense. Every aspect of the phenomenon to model is represented by a state universe of sub-states $Q_i$. The states themselves are (fuzzy) subsets of the universe of discourse state sets, so that within $Q_i$ a frame of cognition is determined (its fineness depending on the application context and on the requirements toward the optimisation algorithm), so that typical states like "Totally intact", "Slightly damaged", "Medium condition", etc., up to "Dangerous for life" are considered. Any transition from one state to the other (improvement of the condition, refurbishment or renovation) involves a certain cost $c$. In the case of a transition from $q_i$ to $q_j$ it is expressed by a membership value $\mu_{i_j} = c(q_i, q_j)$. In our model the added cost $\Sigma\mu_{i_j}$ along a path $q_{i_1} \to q_{i_2} \to \dots \to q_{i_n}$ is not usually equivalent with the cost of the transition $\mu_{i_n}$ along the edge $q_{i_l} \to q_{i_n}$. This is in accordance with the non-additivity property of the fuzzy (possibility) measure and is very convenient in our application, as it is also not additive in the case of serial renovations.

As a simple example the Fig. 1 depicts the possible transitions among the states of a specific sub-state: $Q_0$ represents the initial (deteriorated) state, while $Q_{n_i}$ represents the acceptable (renovated) state.

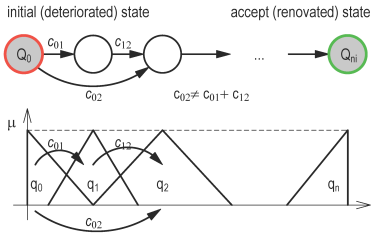



**Fig. 1.** The function diagram of fuzzy state machine

In the case of fuzzy signature machines each of the leaves contains a sub-automaton with the above property. The parent leave of a certain sub-graph is constructed from the child leaves, so that the sub-automaton

$A^i = A^{i_1} \times A^{i_2} \times \cdots \times A^{i_m}$, and thus the states of $A^i$ are $Q^i = Q^{i_1} \times Q^{i_2} \times \cdots \times Q^{i_n}$, so that the transition $Q^{j_1} \to Q^{j_2}$ in this case means the parallel (or subsequent) transitions $q_{j_{1_1}} \to q_{j_{1_2}} \times q_{j_{2_1}} \to q_{j_{2_2}} \times \cdots \times q_{j_{n_1}} \to q_{j_{n_2}}$. A special aggregation is associated with each leaf; similarly as it is in the fuzzy signatures, however, in this case the aggregation calculates the resulting cost $\mu_{j_{1_2}}$ of the transition $q_{j_1} \to q_{j_2}$, so that

$$\mu_{j_{1_2}} = c(q_{j_1}, q_{j_2}) = a_j(c(q_{j_{1_1}}, q_{j_{1_2}}), c(q_{j_{2_1}}, q_{j_{2_2}}), \cdots, c(q_{j_{n_1}}, q_{j_{n_2}})), \tag{11}$$

where $a_j$ stands for the respective (often non-symmetric) aggregation.

The selection of aggregation operator is a key issue that may determine the final result of the model; however, the signature structure makes the application of different aggregation methods possible for each node.

## 3 Case Study: Optimizing the Renovation of the Plinth Zone in a Block of Residential Houses

In the case study the "plinth zone" of the residential houses in the same block is represented. In practice, the building as a hierarchically ordered set of building components may have numerous state machines that are more complex in structure and in operation. However, the application of the proposed model may be easily understood on this simple unit, where three sub-state machines operate. The Fig. 2 illustrates the examined unit: in this model the observed conditions of the building components determine their initial state values (the accept state values express the reasonable and good condition of these components).

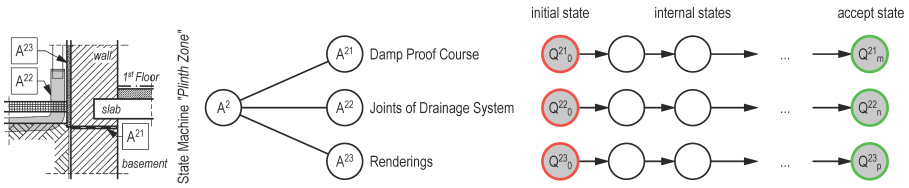



**Fig. 2.** The State Machine "Plinth Zone" ($A^2$) and its sub-State Machines ($A^{2_1} - A^{2_3}$)

The Table 1 summarizes the observations briefly in each building components in the examined sequence; the initial state values are assigned to each sub-state machine.

**Table 1.** Observations in the initial state of the sub-state machines of the "plinth zone"

| **Sub-state Machine** | **Observation** | **Initial state value** |
|---|---|---|
| $A^{2_1}$ Damp Proof Course | Missing or low performance of horizontal DPC → *the soil moisture flows up in the masonry structure* | $x_0^{2_1} = 0.25$ |
| $A^{2_2}$ Joints of Drainage System | The embedded sections below the ground are broken or cracked; the joints are not watertight → *the rainwater flows directly into the bottom wall without any hindrance* | $x_0^{2_2} = 0.45$ |
| $A^{2_3}$ Renderings | The formerly applied waterproof plaster prevents the water to evaporate towards the courtyard → *the water appears on the internal surface of walls* | $x_0^{2_3} = 0.30$ |

Due to some renovation processes, several internal states may be determined in the sub-sequences, where the interventions trigger the state change; the $c_i$ cost may be assigned to these state changes. With a rough approximation 2;1;2 internal states were determined respectively. The $c_i$ trigger were calculated based on the nationally accepted contractors' billing database.

The attributes of the building components and the simplicity of the method verify alike the application of Ordered Weighted Averaging Aggregation (OWA) operator as it was presented by Yager in [8]. In this case the $w$ weighting factors may represent the importance of the renovation steps that is also based on experts' statements. The Table 2 summarizes the basic data of each sub-state machines; their weighting factors are also indicated.

In the $A^2$ state machine $3 \times 2 \times 3 = 18$ states may be determined as a combination of the three sub-state machines. In the formed state space the accept state may be reached on numerous paths, where the optimum solution (the

**Table 2.** The initial, internal and accept states and the costs of transitions among state changes ($A^{2_1} - A^{2_3}$)

| State | Sub-state machines | | |
|---|---|---|---|
| | $\mathbf{A^{2_1}}$ | $\mathbf{A^{2_2}}$ | $\mathbf{A^{2_3}}$ |
| | $w^{2_1} = 0.75$ | $w^{2_2} = 0.55$ | $w^{2_3} = 0.25$ |
| $\mathbf{Q_0}$ | $x_0^{2_1} = 0.25$ | $x_0^{2_2} = 0.45$ | $x_0^{2_3} = 0.30$ |
| $\mathbf{Q_1}$ | $x_1^{2_1} = 0.65$ | $\mathbf{x_1^{2_2} = 0.78}$ | $x_1^{2_3} = 0.50$ |
| | $c_{0_1}^{2_1} = €\,51{,}650$ | $c_{0_1}^{2_2} = €\,8{,}064$ | $c_{0_1}^{2_3} = €\,8{,}064$ |
| $\mathbf{Q_2}$ | $\mathbf{x_2^{2_1} = 0.85}$ | | $\mathbf{x_2^{2_3} = 0.75}$ |
| | $c_{1_2}^{2_1} = €\,22{,}500$ | | $c_{1_2}^{2_3} = €\,6{,}850$ |
| | $c_{0_2}^{2_1} = €\,63{,}030$ | | $c_{0_2}^{2_3} = €\,11{,}930$ |

highest efficacy and the lowest cost in the same time) can not be found easily. Therefore a software was developed to evaluate the data: in the calculation the costs and the internal state values were also taken into consideration. The evaluation of the initial and the accept states of sub-state machines and their aggregation to the level of the $A^2$ state machine is illustrated in Fig. 3.

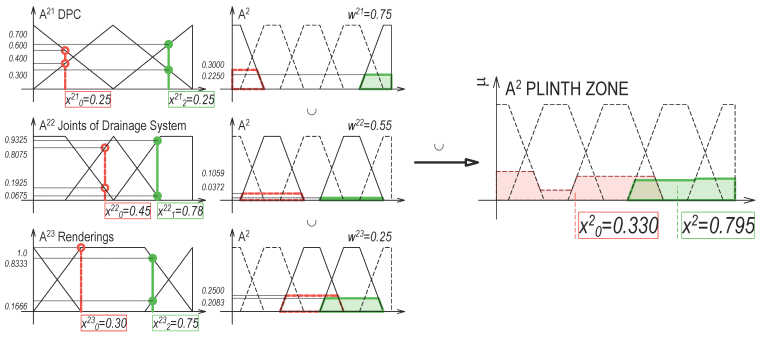



**Fig. 3.** The Initial and Accept States in $A^2$ State Machine

As a result, the $(Q_0^{21}, Q_0^{22}, Q_0^{23}) \rightarrow (Q_0^{21}, Q_0^{22}, Q_2^{23}) \rightarrow (Q_2^{21}, Q_1^{22}, Q_2^{23})$ path was calculated as the cost optimum solution in the renovation process of the plinth zone (as it is illustrated in Fig.4).

## 4 Conclusions and Future Work

It is needless to say that a complex renovation plan for the entire building may help the owners for developing the physical condition of their properties. In practice, the renovation of old residential houses is realized in sequences, so that several internal states may be distinguished. With the proposed model the cost

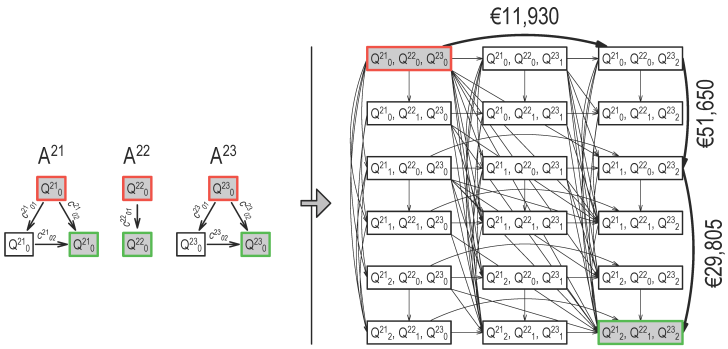



**Fig. 4.** The Cost-Optimized Path in the State Space of $A^2$ State Machine

effective renovation sequence may be found, were the costs and the state changes are taken into consideration in the same time.

This study revealed that the "plinth zone" renovation process may have 18 states with more than 90 state changes. The entire building has a limited number of such sequences; however this number might be rather high. Because of the computational complexity of the optimization problem, a meta-heuristics with reasonably low complexity but good convergence expectations is proposed, such as the Bacterial Evolutionary / Bacterial Memetic Algorithm, the Particle Swarm or the Imperialist Competitive Algorithm.

## References

1. Molnárka, G. I., Kovács, Sz., Kóczy, L. T.: Fuzzy Rule Interpolation Based Fuzzy Signature Structure in Building Condition Evaluation. Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on. IEEE, 2014.
2. Kóczy, L. T.: Vector Valued Fuzzy Sets, J. BUSEFAL 4, 41-57 (1980)
3. Goguen, J.: L-fuzzy sets, Journal of Mathematical Analysis and Applications 18, 145-174 (1967)
4. Pozna, C., Minculete, N., Precup, R.-E., Kóczy, L. T., Ballagi, Á.: Signatures: Definitions, operators and applications to fuzzy modelling, Fuzzy Sets and Systems 201, 86-104 (2012)
5. Kóczy, L. T., Vámos, T., Biró, G.: Fuzzy Signatures. In: EUROFUSE-SIC '99, pp. 210-217 (1999)
6. Vámos, T. and Kóczy, L.T. and Biro, G.: Fuzzy signatures in data mining. In: IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th, pp. 2842-2846 vol 5.
7. Molnárka, G. I., Kóczy, L. T.: Decision Support System for Evaluating Existing Apartment Buildings Based on Fuzzy Signatures. Int. J. of Computers, Communications & Control 6, 442-457 (2011)
8. Yager, R. R.: On ordered weighted averaging aggregation operators in multicriteria decision making. In: Systems, Man and Cybernetics, IEEE Transactions on 18, 183-190 (1988)

# Efficient Unfolding of Fuzzy Connectives for Multi-adjoint Logic Programs *

Pedro J. Morcillo and Ginés Moreno

Dept. Computing System, U. Castilla-La Mancha, Albacete 02071, Spain
Email: gines.moreno@uclm.es,pmorcillo@dsi.uclm.es

**Abstract** During the last decade we have designed the $\mathcal{FLOPER}$ tool for assisting the development of flexible software applications coded with a promising language in the fuzzy logic programming arena acomplishing with the the so-called *multi-adjoint logic programming approach*, where a set of (logic) PROLOG-like rules are accompanied with a set of (functional) HASKELL-like fuzzy connective definitions for manipulating truth degrees beyond the simpler case of *{true,false}*. Moreover, we have recently provided optimization techniques which reuse some variants of program transformation techniques based on unfolding which have been largely exploited in the pure functional -not fuzzy- setting for enhancing the behavior of such operators. In this paper we show how to improve the efficiency of the proper unfolding process by reusing the very well-known concept of dependency graph.

**Keywords:** Fuzzy Logic Programming, Connectives, Unfolding

## 1 Introduction

*Fuzzy Logic Programming* is an interesting and still growing research area that agglutinates the efforts for introducing fuzzy logic into *Logic Programming* [7], in order to provide techniques and constructs for dealing with uncertainty and approximated reasoning in a natural way. Most of these systems replace the classical inference mechanism of SLD–Resolution with a fuzzy variant which is able to handle partial truth. This is the case of *multi-adjoint logic programming* (MALP in brief, [8]), where programs are parametric to lattices modeling rich notions of truth degrees. In this framework, a program is a set of "weighted" rules together with a set of equations (*rewriting rules*) defining the repertoire of fuzzy connectives considered in a concrete lattice of truth degrees.

To solve a MALP goal, i.e, aquery to the system plus a substitution (initially the empty substitution, denoted by *id*), a generalization of the classical *modus ponens* inference rule called *admissible steps* are systematically applied on atoms in a similar way to classical resolution steps in pure logic programming, thus returning a state composed by a computed substitution together with an

* Work supported by the EU (FEDER), and the Spanish MINECO Ministry (*Ministerio de Economía y Competitividad*) under grant TIN2013-45732-C4-2-P.

expression where all atoms have been exploited. Next, this expression is interpreted under a given lattice, hence returning a pair $\langle$*truth degree*; *substitution*$\rangle$ which is the fuzzy counterpart of the classical notion of computed answer used in pure logic programming.

Moreover, in [9] we tried to reduce the complexity of connectives (also alleviating the computational cost of derivations) by safely removing all the intermediate calls performed on the equations defining the behavior of such connectives. In Section 2 we show that this process can be easily described in terms of "unfolding", a well-known, widely used, semantics-preserving program transformation operation which in most declarative paradigms is usually based on the application of computation steps on the body of program rules ([5,3,4] describe our experiences regarding the unfolding of fuzzy logic programms). The novelty of our present approach is that it is the first time that unfolding is not applied to program rules, but to connective definitions, maintaining the same final goal, i.e., generating more efficient code. Before concluding in Section 4, we reuse in Section 3 some techniques based on graphs coming from the field of program analysis to also improve the proper transformation process as much as possible, which conforms the main goal of the present paper.

## 2 MALP and Unfolding Connective Definitions

This section summarizes the main features of multi-adjoint logic programming (see [8] for a complete formulation of this framework). We work with a first order language, $\mathcal{L}$, containing variables, constants, function symbols, predicate symbols, and several (arbitrary) connectives to increase language expressiveness: implication connectives ($\leftarrow_1, \leftarrow_2, \ldots$); conjunctive operators (denoted by $\&_1, \&_2, \ldots$), disjunctive operators ($|_1, |_2, \ldots$), and hybrid operators (usually denoted by $@_1, @_2, \ldots$), all of them are grouped under the name of "aggregators" or directly "connectives". Aggregation operators are useful to describe/specify user preferences. An aggregation operator, when interpreted as a truth function, may be an arithmetic mean, a weighted sum or in general any monotone application whose arguments are values of a complete bounded lattice $L$. For example, if an aggregator $@$ is interpreted as $[\![@]\!](x, y, z) = (3x + 2y + z)/6$, we are giving the highest preference to the first argument, then to the second, being the third argument the least significant. Although these connectives are binary operators, we usually generalize them as functions with an arbitrary number of arguments. So, we often write $@(x_1, \ldots, x_n)$ instead of $@(x_1, \ldots, @(x_{n-1}, x_n), \ldots)$. By definition, the truth function for an n-ary aggregation operator $[\![@]\!] : L^n \to L$ is required to be monotonous and fulfills $[\![@]\!](\top, \ldots, \top) = \top$, $[\![@]\!](\bot, \ldots, \bot) = \bot$.

Additionally, our language $\mathcal{L}$ contains the values of a multi-adjoint lattice, $\langle L, \preceq, \leftarrow_1, \&_1, \ldots, \leftarrow_n, \&_n\rangle$, equipped with a collection of adjoint pairs $\langle \leftarrow_i, \&_i\rangle$, where each $\&_i$ is a conjunctor which is intended to the evaluation of *modus ponens* [8]. In general, $L$ may be the carrier of any complete bounded lattice but, for readability reasons, in the examples we shall select $L$ as the set of real numbers in the interval $[0, 1]$. A *L-expression* is a well-formed expression com-

posed by values and connectives of $L$, as well as variable symbols and *primitive operators* (i.e., arithmetic symbols such as $*, +, min$, etc...). In what follows, we assume that the truth function of any connective @ in $L$ is given by its corresponding *connective definition*, that is, an equation or *rewriting rule* of the form $@(x_1, \ldots, x_n) = E$, where $E$ is a $L$-expression not containing variable symbols apart from $x_1, \ldots, x_n$.

The use of connectives inside the definition of other connectives is a powerful expressive resource useful not only for programmers interested in describing complex aggregators, but it also plays an important role in fuzzy transformation techniques such as the fold/unfold framework we have described in [5,3,4]. The following definition recasted from [9], in essence describes a technique based on classical unfolding transformations for simplifying, when possible, connective definitions by "unnesting" unnecesary calls to other connectives.

**Definition 1 (C-Unfolding).** *Let $\langle L, \preceq \rangle$ be a multi-adjoint lattice containing the connective definitions $@(x_1, \ldots, x_n) = E$ and $@'(x'_1, \ldots, x'_m) = E'$, such that a call to $@'$ of the form $@'(t_1, \ldots, t_m)$ appears in $E$. Then, the unfolding of connective @ w.r.t. connective $@'$ or directly, the c-unfolding of @, is the new equation: $@(x_1, \ldots, x_n) = E[@'(t_1, \ldots, t_m)/E'']$, where $E''$ is obtained from the L-expression $E'$ by replacing each variable (formal parameter) $x'_i$ by its corresponding value (actual parameter) $t_i$, $1 \leq i \leq m$, that is $E'' = E'[x'_1/t_1, \ldots, x'_m/t_m]$.*

We assume here that the rules (equations) describing connective definitions are taken renamed apart (at least one of them) before applying an unfolding step, as it is also usual with program rules in many declarative transformation tasks.

*Example 1.* Given connective: $@^*(x_1, x_2) = \&_{\texttt{prod}}(|_{\texttt{luka}}(x_1, 0.6), x_2)$, and remembering that $|_{\texttt{luka}}(x'_1, x'_2) = min(1, x'_1 + x'_2)$, then, we can unfold connective $@^*$ w.r.t. connective $|_{\texttt{luka}}$ as follows:

- Firstly, we generate the "matcher" between the call $|_{\texttt{luka}}(x_1, 0.6)$ appearing in the "rhs" (right hand side) of the first rule and the "lhs" (left hand side) of the second rule $|_{\texttt{luka}}(x'_1, x'_2)$, thus producing links $x'_1/x_1$ and $x'_2/0.6$.
- Next, we apply both bindings to the rhs of the second rule, obtaining the L-expression $min(1, x_1 + 0.6)$.
- Then, this L-expression is used to replace the original call to $|_{\texttt{luka}}$ in the rhs of the first rule, producing $\&_{\texttt{prod}}(min(1, x_1 + 0.6), x_2)$.
- Finally, this last L-expression conforms the rhs of the new connective definition for $@^*$, that is: $@^*(x_1, x_2) = \&_{\texttt{prod}}(min(1, x_1 + 0.6), x_2)$.

Following the same method, but performing now the c-unfolding of $@^*$ w.r.t. $\&_{\texttt{prod}}$ whose connective definition is $\&_{\texttt{prod}}(x_1, x_2) = x_1 * x_2$, we obtain the final rule defining $@^*$ with the following shape $@^*(x_1, x_2) = min(1, x_1 + 0.6) * x_2$. Note that the new connective definition is just a simple arithmetic expressions involving primitive operators but no calls to other connectives, as wanted.

## 3 C-Unfolding and Call Graphs

Our experiences in fuzzy fold/unfold transformations [5,3,4], reveal us that drastic situations associated to degenerated transformation sequences might eventually produce highly nested definitions of connectives. For instance, assume the following sequence of (extremely inefficient) connective definitions:

$$\begin{array}{ll} @_{100}(x_1, x_2) & = @_{99}(x_1, x_2) \\ @_{99}(x_1, x_2) & = @_{98}(x_1, x_2) \\ @_{98}(x_1, x_2) & = @_{97}(x_1, x_2) \\ \ldots\ldots\ldots & \ldots\ldots\ldots \\ @_1(x_1, x_2) & = @_0(x_1, x_2) \\ @_0(x_1, x_2) & = x_1 * x_2 \end{array}$$

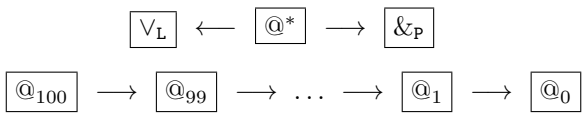
When trying to solve two expressions of the form $@_{100}(0.9, 0.8)$ and $@_0(0.9, 0.8)$, we obtain the same result 0.72, but the effort needed to solve the first expression is very high (due to the 100 avoidable calls to auxiliary connectives) compared with the second expression (which simply evaluates the arithmetic operator $*$).

Fortunately, by systematically performing c-unfolding on the previous connectives, this problem is successfully solved in a simple way: after applying a c-unfolding step on aggregator $@_{100}$ we obtain $@_{100}(x_1, x_2) = @_{98}(x_1, x_2)$, which admits a new c-unfolding process to become $@_{100}(x_1, x_2) = @_{97}(x_1, x_2)$, and following this trail, after applying the final one-hundredth c-unfolding step, we reach the desired connective definition $@_{100}(x_1, x_2) = x_1 * x_2$. Of course, the transformation process does not finish here, because we also need to rearrange the shape of all the remaining connective definitions. So, for each aggregator $@_i$, $0 \leq i \leq 100$, we need exactly $i$ c-unfolding steps to achieve the appropriate connective definition.

However, there exist a second, much more intelligent alternative to highly reduce the number of transformation steps needed to obtain the same final set of improved connective definitions. In our example, the idea is to proceed just in the inverse order than previously. So, since $@_0$ does not admit unfolding, we proceed with $@_1$, whose connective definition becomes $@_1(x_1, x_2) = x_1 * x_2$ after just a single c-unfolding step. Now, we take profit of this improved definition when unfolding $@_2$, since in just a unique (nor two) c-unfolding steps we obtain the optimal definition $@_2(x_1, x_2) = x_1 * x_2$. Note that the benefits of this last process, are also inherited when transforming $@_3$, $@_4$ and so on. So, the advantages obtained after applying each c-unfolding on a different connective, are "propagated" to the remaining connectives being improved, which implies that we simply need one hundred transformation steps to optimize the definitions of the whole set of connectives.

In order to identify in a systematic way the best ordering for performing c-unfolding operations on connectives, we firstly construct the *call graph* of a multi-adjoint lattice $L$ associated to a given program $\mathcal{P}$, i.e., a directed graph that contains the connective symbols as nodes and an edge from connective $@$ to aggregator $@'$ for each connective definition in $L$ of the form $@(x_1, \ldots, x_n) = E$,

where the $L$-expression $E$ contains a call to $@'$. Given an edge from node $@$ to node $@'$, we denote it as an *out-edge* of $@$ and as an *in-edge* of $@'$. For instance, the call graphs associated to all the connectives seen so far are:



As we are going to see, the use of call graphs will largely help us to decide when to unfold each connective in order to minimize the number of transformation steps. Anyway, before doing this, it is important to note that the construction of such graphs constitute a fast way to detect possibly abnormal connective definitions: if there exist cycles in the graph, all connectives involved on such cycles should be considered corrupt, since in most cases, their further evaluation might fall in infinite loop. For this reason, in what follows we only consider call graphs without cycles, as occurs with the ones depicted before.

When selecting a connective to apply c-unfolding, we give priority to those ones without out-edges, as occurs in our examples with nodes labeled with $\vee_{\mathtt{L}}$, $\&_{\mathtt{P}}$ and $@_0$, which in our particular case do not need c-unfolding because their definitions do not perform calls to other aggregators. Once a concrete connective has been selected and then unfolded as much as possible (and hence, its definition has been completely improved by removing all its auxiliary calls), then the proper node as well as all its in-edges (remember that it has not associated out-edges) are removed from the graph. The process is iterated as much as needed until the call graph becomes empty. For instance in our example, once removed nodes $\vee_{\mathtt{L}}$, $\&_{\mathtt{P}}$ and $@_0$, the new candidates are nodes $@^*$ and $@_1$. The first one is unfolded w.r.t. $\vee_{\mathtt{L}}$ and $\&_{\mathtt{P}}$ and then removed, whereas the second one is dropped out after being unfolded w.r.t. $@_0$. Then the process continues with $@_2$, next $@_3$ and so on, being $@_{100}$ the last connective whose definition is optimized by applying just a single c-unfolding step, thus accomplishing with the desired ordering and benefits reported along this section. This methodology can be formalized as follows:

1. Build a "call graph" G from a multi-adjoint lattice L, where:
   (a) each node in G is a connective defined in L.
   (b) each directed edge in G "means" a call from @ to @'.
2. While G is not empty:
   (a) select a node/connective @ in G without out-edges,
   (b) apply C-unfolding as much as possible on node/connective @,
   (c) remove @ as well as all its (in/out) edges from G.

We wish to finish this section by mentioning that there exists a wide tradition on the use of "graphs" (and many different extensions/variants of this formal concept) as an auxiliary data structure helping to analyze the behaviour of systems/programs at several levels, also taking profit in practice of its deep mathematical background. For instance, and simply focusing on termination topics in declarative programming (which is somehow influencing our recent

research interest), the notions of *dependency graphs* and *size-change graphs* have been well reported in [2,6]. Fortunately, the notion of "call graph" used in this paper is simpler than the two ones commented before, mainly due to the fact that our final goal is easier to achieve too. Anyway, further refinements of the present work perhaps will need a revision of such concepts.

## 4 Conclusions and Future Work

In this paper we were concerned with the optimization of fuzzy logic connectives whose artificial, inefficient definitions could have been automatically produced by previous transformation processes applied on fuzzy MALP programs. Our technique, inspired by rewriting-based unfolding, takes profit from clear precedents in pure functional programming. In this paper we have focused on the optimization of the proper unfolding process (initially presented in [9]) by making use of call graphs in order to decide the ordering in which several connective calls must be unfolded inside a concrete connective definition. For the immediate future, we plan to implement our technique inside the fuzzy logic programming environment $\mathcal{FLOPER}$ (visit `http://dectau.uclm.es/floper/`) we have designed for developing applications coded with the MALP language [1,10].

## References


1. J.M. Almendros-Jiménez, A. Luna, and G. Moreno. Fuzzy xpath through fuzzy logic programming. *New Generation Computing (in press)*, page 42, 2015.
2. T. Arts and J. Giesl. Termination of term rewriting using dependency pairs. *Theor. Comput. Sci.*, 236(1-2):133–178, 2000.
3. J.A. Guerrero and G. Moreno. Optimizing fuzzy logic programs by unfolding, aggregation and folding. *Electronic Notes in Theoretical Computer Science*, 219:19–34, 2008.
4. P. Julián, J. Medina, P.J. Morcillo, G. Moreno, and M. Ojeda-Aciego. An unfolding-based preprocess for reinforcing thresholds in fuzzy tabulation. In *Proc. of the 12th International Work-Conference on Artificial Neural Networks, IWANN'13*, pages 647–655. Springer Verlag, LNCS 7902, Part I, 2013.
5. P. Julián, G. Moreno, and J. Penabad. On Fuzzy Unfolding. A Multi-adjoint Approach. *Fuzzy Sets and Systems*, 154:16–33, 2005.
6. C. Lee, N. Jones, and A. Ben-Amram. The size-change principle for program termination. *SIGPLAN Not.*, 36(3):81–92, 2001.
7. J.W. Lloyd. *Foundations of Logic Programming*. Springer-Verlag, Berlin, 1987.
8. J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Similarity-based Unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146:43–62, 2004.
9. P.J. Morcillo and G. Moreno. Improving multi-adjoint logic programs by unfolding fuzzy connective definitions. In *Proc of the 13th International Work-Conference on Artificial Neural Networks, IWANN 2015 Mallorca, Spain, June 10-12*, pages 511–524. Springer Verlag, LNCS 9094, Part I, 2015.
10. C. Vázquez, G. Moreno, L. Tomás, and J. Tordsson. A cloud scheduler assisted by a fuzzy affinity-aware engine. In *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2015, Istanbul, Turkey, August 2-5*, page 8 (in press), 2015.

# Modelling the Uncertainty in the Condition Assessment of Residential Buildings

Ádám Bukovics[1], István Á. Harmati[2], László T. Kóczy[3]

[1] Department of Structural and Geotechnical Engineering, Széchenyi István University, Győr, Hungary
bukovics@sze.hu

[2] Department of Mathematics and Computational Sciences, Széchenyi István University, Győr, Hungary
harmati@sze.hu

[3]Department of Automation, Széchenyi István University, Győr, Hungary
koczy@sze.hu

[3]Department of Telecommunications and Media Informatics, BME, Budapest, Hungary
koczy@tmit.bme.hu

**Abstract.** With the aim of ranking the residential buildings, and supporting decision making for the purpose of interventions, expert opinions describing the status of buildings are often prepared. These, partly because of the subjectivity of the expert, and partly due to the quality and quantity of the available data, may include significant uncertainties and inaccuracy, the knowledge of which can influence the decisions on interventions. Our aim was to elaborate a method for determining the status of buildings, which can be used to model the uncertainties stemming from the circumstances of the examination. A fuzzy signature based method was elaborated to model the expert uncertainties. At the peaks of the tree structure the singleton values, provided by the expert, are transformed into membership functions subject to the uncertainties, then with the help of aggregation operators the status of the whole building, and the related uncertainties are determined. The fuzzy signature set based decision support method can achieve more adequate results than those achieved with the traditional statistical methods. In addition to the subjective evaluation of the expert, the reliability thereof can be well expressed with fuzzy membership functions. In order to make well-based decisions it is vital to know the reliability of the expert reports.

**Keywords:** fuzzy signatures, condition assessment, uncertainty, building diagnostics

## 1 Introduction

Determining the structural condition of a building, or analysing the status and ranking of the structures of various buildings with the aim of supporting the decision making for intervening purposes is a complicated task. Despite the uncertainties the condition of residential buildings can be modelled with appropriate accuracy, buildings can be ranked on the basis of their condition (considering various priorities) and decision supporting methods can be created in a tree-structure system using a fuzzy signature based model. In the course of our former research we have created decision support and ranking methods of this type. The aim of this work is to determine the status of certain building sub-structures and the buildings themselves. The singleton values and linguistic labels, provided by the experts for the status of building structures will be transformed into membership functions. When calculating the resulting membership function by the application of aggregation operators on the leaf "function" an additional operation is applied, to include the overall uncertainty of the evaluation of the building, thus the final result will be once more modified. The outcome will show to which extent the qualifications of the building structure, provided an expert report, are reliable.

## 2 Uncertainties of the status evaluations

When preparing a status evaluation analysis it is not sure whether two well-prepared experts would give the same evaluation of the status of structures. Professional experience and skill of the experts, completing the expert report may be different to a significant extent this is why the status of the examined structures, defined by certain experts may be inaccurate. The outcome of the evaluation analysis can be significantly influenced by the circumstances of the assessment. For example it may help to carry out this work, if the original design documentation of the building or a part of it is still available, or the discovered conditions have been documented during a later renovation or assessment.. There are construction materials which have been accepted and prevailing at the time of the construction of the building, later however disadvantageous features were detected and so these are not used anymore.

## 3 Grouping of uncertainties

Based on the above, expert uncertainties were sorted into two groups. Uncertainties in the first group (general uncertainties) have impact on the expert evaluation in each case, and the expert's subjectivity is integrated into the system.
General uncertainties of expert evaluation (GU):

- Partial subjectivity of expert evaluation
- Professional skill of the expert
- Features of the expert behaviour
- Elaborateness (detailedness) of the expert evaluation

Effects, independently from the expert are sorted into the second group (special uncertainties), when features of the membership function depend on the quality and quantity of input data available for the experts.. Such uncertainties are called special uncertainties of expert evaluation. Special uncertainties (SUP) are further sorted into two sub-groups,
Special uncertainties of expert evaluation, featuring certain building sub-structures:

- Method of checking (destruction or visual checking)
- Visual observability (easy to observe or hidden structural element)
- Presence of active deteriorations
- Quality of building materials
- Extent to which the construction materials are accepted

Special uncertainties of expert evaluation, featuring the whole building (SUT):

- Active vibration effects
- Former design documentations (availability)
- Date of construction

## 4 Effects of uncertainties exerted on membership functions

When modelling the expert uncertainties, to every expert estimation uncertainties are assigned so that the singleton membership value, provided by the expert, is transformed into a membership function. Triangular- and trapezoidal-shaped membership functions will be used to model expert uncertainties, because they can be easily handled, and due to the uncertainties of the building structures to be tested, membership functions of more complicated shape are not justified. The generation of the membership functions is as follows (Figure 1). The expert specifies a singleton value, which will be first normalised to [0,1] membership function will be applied around this value. It is supposed that the expert cannot certainly define the quality of the structure. The shape of the membership function will be modified step by step, while general and special uncertainties of expert evaluation are taken into account.

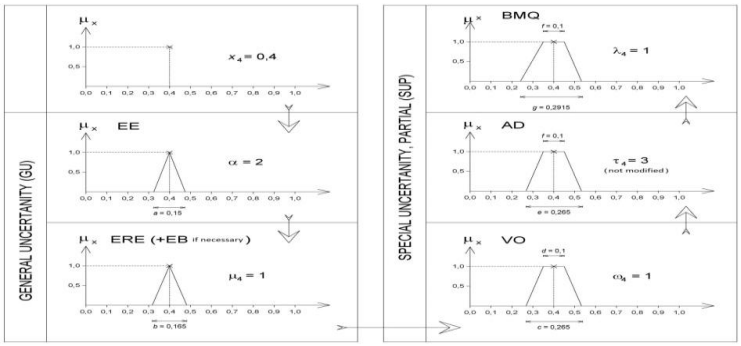



**Fig. 1.** Modelling the uncertainty with membership function

Every uncertainty modifies the shape of the membership functions. E.g. based on the quality of visual observation the triangular membership function may change in a way described in the Figure 2. The value of ω depends on the visual perceptibility of the structural element. In case of hidden structural elements it is e.g. ω=1.

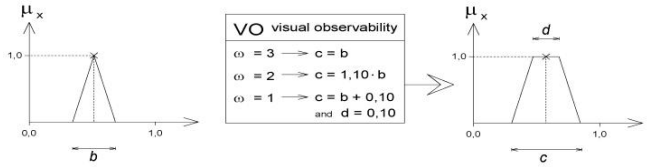



**Fig. 2.** The effect of the quality of visual observation to the membership value

## 5 The structure of the fuzzy signature

### 5.1 The basic structure of the fuzzy set signature

When modelling the status of residential buildings the components are well structured and a hierarchical tree-like structure can be built up from them, gaining significant additional information about the problem. At a higher level certain components of the structure are defined by a partial tree of the components. It is a basic requirement towards fuzzy membership functions on the leaves of the structure that the basic set thereof should be within the interval of [0,1]. In this problem a four-level fuzzy signature structure was used (Figure 3.).

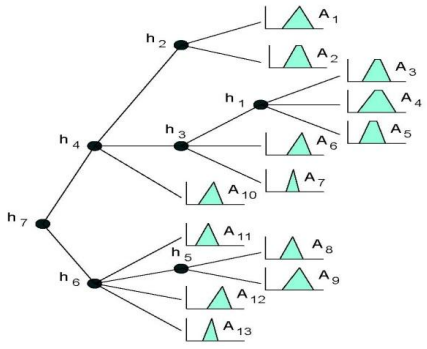



**Fig. 3.** Basic structure of the fuzzy set signature

The membership functions at the leaves of the structure is related to the following building structures: foundation structures ($A_1$), wall structures ($A_2$), cellar floor ($A_3$), intermediate floor ($A_4$), cover floor ($A_5$), side corridor structures ($A_6$), step structures ($A_7$), roof structures ($A_8$), roof covering ($A_9$), facade ($A_{10}$), footing ($A_{11}$), tin structures ($A_{12}$), insulation against soil moisture and ground water ($A_{13}$).

### 5.2 The applied aggregation operators

To every internal node of the structure aggregation operators were assigned. Thus it became possible to modify the structure so that a sub-tree of variables is reduced to the root of the sub-tree. Aggregation operators are calculating aggregated values from a set of values at the lower level nodes. The operation of aggregation can be specified with an n-variable function h: $[0,1]^n \rightarrow [0,1]$ [4].

The applied aggregation operators are piecewise combinations of the min and weighted mean operators, which depend on the special features of the building (number of storeys, extend of the cellar, side corridor). For example the $h_2$ aggregation operator, related to the status of the vertical load bearing structures and $h_7$ aggregation operator, related to the status of the building, were defined as follows (*n* is the number of the storeys of the building).

$$h_2 = \begin{Bmatrix} (0{,}50 - 0{,}05 \cdot (n-1)) \cdot x_1 + (0{,}50 + 0{,}05 \cdot (n-1)) \cdot x_2 \quad \text{if} \quad x_1 > 0{,}40 \\ (0{,}50 - 0{,}05 \cdot (n-1)) \cdot x_1 + (0{,}50 + 0{,}05 \cdot (n-1)) \cdot x_2 \quad \text{if} \quad 0{,}40 \geq x_1 \geq 0{,}20 \text{ and } x_2 < x_1 \\ (0{,}50 - 0{,}05 \cdot n) \cdot x_1 + (0{,}50 + 0{,}05 \cdot n) \cdot x_2 \quad \text{if} \quad 0{,}40 \geq x_1 \geq 0{,}20 \text{ and } x_2 \geq x_1 \\ \min(x_1; x_2) \quad \text{if} \quad x_1 < 0{,}20 \end{Bmatrix} \quad (1)$$

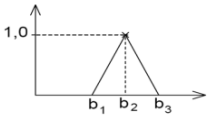
$$h_7 = \begin{Bmatrix} 0{,}75 \cdot h_4 + 0{,}25 \cdot h_6 \quad \text{if} \quad h_4 > 0{,}4 \\ \text{or} \quad 0{,}2 \leq h_4 \leq 0{,}4 \text{ and } h_6 < h_4 \\ 0{,}85 \cdot h_4 + 0{,}15 \cdot h_6 \quad \text{if} \quad 0{,}2 \leq h_4 \leq 0{,}4 \text{ and } h_6 \geq h_4 \\ \min(h_4; h_2) \quad \text{if} \quad h_4 < 0{,}2 \end{Bmatrix} \quad (2)$$

## 6 Testing the model

In order to test the model detailed technical-static expert reports on 340 residential buildings were available. Using them a database was created, which on the one hand contained the data of building structures and building diagnostics, and on the other hand it contained the data needed for determining the uncertainties of status evaluations. When creating the decision support model we have relied on these data as input data. The set of data, which belongs to the problem to be modelled, has a joint basic structure. Fine-tuning of membership functions, which handle the uncertainties is performed with the help of the database.

As an example we have chosen an old building which was built in 1894. The human expert's opinion was modelled by triangular shaped fuzzy numbers, according to the previous sections. The values of the shape parameters are shown in the table (Figure 4).

The final conclusion (the output of the fuzzy signature) is a triangular shaped fuzzy number with parameters (0.324; 0.376; 0.428). From this membership function a crisp conclusion can be obtained applying one of the well-known defuzzification methods, but the shape of the final fuzzy set provides lot of information about the uncertainty of the judgement, which should be taken into consideration before the decision.




| | $b_1$ | $b_2$ | $b_3$ |
|---|---|---|---|
| $A_1$ | 0,295 | 0,350 | 0,405 |
| $A_2$ | 0,200 | 0,250 | 0,300 |
| $A_3$ | 0,600 | 0,650 | 0,700 |
| $A_4$ | 0,600 | 0,650 | 0,700 |
| $A_5$ | 0,202 | 0,250 | 0,298 |
| $A_6$ | 0,350 | 0,400 | 0,450 |
| $A_7$ | 0,345 | 0,400 | 0,455 |
| $A_8$ | 0,190 | 0,250 | 0,310 |
| $A_9$ | 0,140 | 0,200 | 0,260 |
| $A_{10}$ | 0,300 | 0,350 | 0,400 |
| $A_{11}$ | 0,600 | 0,650 | 0,700 |
| $A_{12}$ | 0,190 | 0,250 | 0,310 |
| $A_{13}$ | 0,150 | 0,200 | 0,250 |

**Fig. 4.** Triangular shaped fuzzy numbers

From a mathematical point of view an interesting question is the sensitivity (or stability) of the signature-based decision support method. May a small change in the input variables, can result a large change of the final membership function, or not? The answer depends on the structure of the signature and on the applied aggregation operators. A similar question was analysed in details for the case of crisp inputs in [5] and [6] by using various type of inequalities [7], and the sensitivity was presented in terms of vector norms of change of the input. For the case of fuzzy variables a more sophisticated approach will be needed, this is part of our planned work.

## 7 Summary

Status evaluation with the aim of decision support interventions can be ideally performed with the use of fuzzy signatures. Within this the uncertainties of expert reports can be well modelled by transforming the membership values into membership functions. Taking into account the uncertainties important additional information will be available, which will promote appropriate decision making.

## Acknowledgement

This work was supported by TÁMOP-4.2.2.A-11/1/KONV-2012-0012, and by the Hungarian Scientific Research Found (OTKA) K105529 and K108405.

## References

1. Molnárka, G., Kóczy L. T.: *Decision Support System for Evaluating Existing Apartment Buildings Based on Fuzzy Signatures*, Int. J. of Computers, Communications & Control, 2011, No.3, pp. 442-457.
2. Bukovics, Á., Kóczy L. T.: *Fuzzy Signature-based Model for Qualification and Ranking of Residential Buildings*, XXXVIII. IAHS World Congress on Housing, Istanbul, Turkey, 2012. pp. 290-297
3. Kóczy, L. T., Vámos, T. and Biró, G., *Fuzzy Signatures*, Proceedings of EUROFUSE-SIC'99, pp. 25-28, 1999. Budapest.
4. Harmati, I. Á., Bukovics, Á., Kóczy, L. T.: *Sensitivity analysis of the weighted generalized mean aggregation operator and its applications to fuzzy signatures*, Proceedings of the WCCI 2014 (FUZZ-IEEE 2014), July 6-11, 2014, Beijing, China, pp. 1327-1332.
5. Goguen, J. A.: *L-fuzzy sets*, J.Math. Anal. Appl. 18(1) (1967), pp. 145-174.

# Pitting Corrosion Modelling of 316L Stainless Steel with Bayesian Neural Networks and ROC space

[a]M.J. Jiménez-Come, [b]I.J. Turias, [a]J.J Ruiz-Aguilar, [a]J.A. Moscoso

[a]Department of Civil and Industrial Engineering. [b]Department of Computer Science Engineering.
Intelligent Modelling of Systems Research Group. University of Cádiz.
Polytechnic School of Engineering (Algeciras)
Avda. Ramón Puyol s/n, 11202 Algeciras (Cádiz), Spain, Tel:34 956 028 000;
mariajesus.come@uca.es

**Abstract**
The knowledge of material corrosion behaviour is crucial in material design. The electrochemical techniques have been widely applied to study corrosion. However, these techniques involve a serial of human interpretation that may lead to poor quality results. This paper presents a model based on artificial neural networks (ANNs) considering Bayesian regularization (BNN) to predict pitting corrosion resistance of stainless steel automatically, without resorting to polarization tests. Three principal environmental factors are considered: chloride ion concentration, pH and temperature. The results demonstrate the utility of the models to simulate the relationships between the environmental conditions and the pitting behaviour of 316L stainless steel.

**Keywords.** stainless steel, pitting corrosion, modelling, Bayesian neural networks

## 1 Introduction

Stainless steel is the term used to describe alloys containing at least 10.5% of chromium. There is a wide variety of grades of stainless steels depending on the chemical composition and the heat treatment applied in the manufacturing process. One of the most common grades of stainless steel is the austenitic grade that contains 17-18% chromium and 8-11% of nickel [1]. These elements are the principal reason of the great corrosion resistance presented in this material. The corrosion resistance shown by this alloy is explained based on the thin oxide film formed on the surface in presence of oxygen. This layer is a result of the reaction between the oxygen in the environment and the chromium presented in the alloy. However, in aggressive environments, the passive film may be damaged and the corrosion resistance decreases [2]. Therefore, two may be the principal factors to be considered in the application of stainless steels: the selection of the suitable grades in addition to the knowledge about the corrosion resistance of the alloy, depending of the environmental conditions. The

right selection of the material for each application provides cost saving, reducing the maintenance costs during the service life of the structure.

Corrosion is one of the most critical aspects in stainless steel since it limits the application of this material. Among all types of corrosion, pitting corrosion is the major cause of failure of this material since it is the principal cause of fatigue cracking initiation leading to critical damage mechanisms in structures [3]. This is the reason why pitting corrosion remains as one of the most popular problem in electrochemistry and material science [4]. This type of corrosion usually occurs in the presence of chlorides. The susceptibility of stainless steel to pitting corrosion depends on three principal factors: the chemical composition of the alloy, the metallurgical conditions and the environmental factors. Many authors have tried to analyse pitting corrosion in order to control and reduce this problem in a wide range of engineering applications. In this way, the electrochemical tests have been widely used to evaluate the corrosion resistance [5]. In some cases, these techniques have to be followed by a microscopic analysis to determine the corrosion status of the material. This visual interpretation may include subjectivity in the results. Therefore, in order to get an efficient control of corrosion in stainless steel, a reliable prediction of the corrosion behaviour is necessary.

The use of computation techniques in material science has become more popular leading to benefits for the analysis of the properties of different materials. In the recent years, many authors have suggested applying them in order to understand the pitting corrosion behaviour of different materials. Cottis et al. [6] presented a model based on artificial neural networks (ANNs) to estimate electrochemical potentials as a function of solution composition and temperature. Due et. [7] used these techniques to analyse the effects of corrosion and fatigue cracks on the structural integrity. Boucherit et al. [8] applied these models to study the usefulness of inhibitors for the prevention of localised corrosion. Ramana et al. [9] presented ANN models to simulate the intricate inter-relationships between electrochemical potentials and different environmental factors whereas Cavanaugh et al. [10] applied ANNs to model pit growth as a function of chloride ion concentration, pH and temperature.

In this paper, with the aim to continue our studies about the prediction of the pitting corrosion behaviour as function of the environmental factors [11-12], an application of ANN using the Bayesian regularization is presented. The main goal of this work is understand pitting corrosion behaviour of 316L stainless steel. In this way, the model is presented as an effective tool to predict the pitting corrosion status of austenitic stainless steel automatically. The influence of the principal environmental factors on pitting corrosion is considered: chloride ion concentration, pH and temperature.

## 2 Materials and methods

### 2.1 The database

In order to predict the pitting corrosion resistance of austenitic stainless steel the model was built based on the experimental data obtained from the European project called “Avoiding catastrophic corrosion failure of stainless steel” – (RFSR-CT-2006-

00022). This project was developed in ACERINOX S.A. In this project, a total of 60 patterns of grade 316L stainless steel were subjected to polarization tests to determine the pitting corrosion status in different environmental conditions using NaCl as precursor salt. Values of the environmental conditions in addition to the pitting corrosion status took part into the data set for modelling as input and output variables, respectively.

The ranges for the tested conditions were chloride ion concentration (0.0025-0.1 mol/L), pH (3.5-8.5) and temperature (2-75 $^{o}$C). After each polarization test, all the samples were analysed microscopically for evidence of localized corrosion based on the formation of pits on the material surface (see Fig.1). Based on this analysis, all species were characterized by the environmental conditions in addition to the pitting corrosion status: 1 for samples suffering pitting corrosion and 0, otherwise.

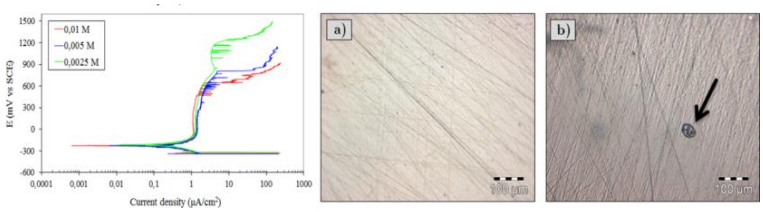



**Fig. 1.** Polarization curves in sodium chloride (T = 30ºC and pH = 3.5) and microscopic analysis of 316L stainless steel after polarization tests: a) no pitting attack b) pitting attack.

### 2.2 Classification model

The universal approximation property shown by the artificial neural networks makes this technique a promising alternative to be applied for pattern recognition purpose [13]. A neural network is made up of a number of processing elements called neurons organized in different layers: the first layer that receives the information from the external world is called the input layer, the last layer named the output layer and the layers between the input and the output ones called the hidden layers. All the neurons in each layer are connected to others by means of direct links. These connections are associated with the adjustable parameters called weights. The values of these parameters are adjusted in order to reduce the error between the output provided by the model and the target value for the entire data set [14], see equation (1). This process is known as supervised learning [15]

$$E_D(w) = \frac{1}{2}\sum_{i=1}^{N}(y_i - t_i) \tag{1}$$

Where $t_i$ and $y_i$ are the target and experimental outputs, respectively.

The multi-layer feedforward is the most common structure used for neural networks. Although the universal approximation property shown by ANNs, there may appear some problems in its implementation such as the selection of the optimal number of hidden neurons, the influence of the initial weight values selected randomly or

the way to avoid overfitting. To deal with these disadvantages, the Bayesian learning strategy was proposed. In this methodology, the weights involved in the model are settled down based on the Bayes' theorem where the posterior knowledge is obtained from the prior assumptions based on the training data. The evidence framework proposed by Mackay is one of the most popular example implementation of the Bayesian theorem [16-17]. This method adapts the prior probability distribution into posterior probability distribution based on the patterns taking part into the training set. The Bayesian regularization approach assigns a probabilistic nature to the weights in order to get the optimal structure for the model. This technique reduces the overfitting since too excessive complex models are penalized improving the generalization ability. In this case, the posterior distribution of the network weights is based on Gaussian approximation and the complexity of the neural network is controlled by means of regularization technique. Basically, Bayesian neural network can be defined as a back propagation network where an additional term in the error function is considered. These terms, introduced by Tikhonov [18] are called the regularization parameters and penalize the network complexity, see equation (2):

$$S(w) = \beta \cdot E_D(w) + \alpha \cdot E_w(w) \tag{2}$$

where the terms $\alpha$ and $\beta$ are called the regularization parameters. In order to get smoother network mappings, the penalty term ($E_w$) in the equation (3) penalizes the weight values that lead to an excessive curvature in the model.

The ratio $\alpha/\beta$ controls the effective complexity of the network. Thus, the correct selection of this ratio is critical within the regularization method. In the Bayesian framework, the weights of the networks are selected based on the maximization of the conditional probability. Assuming a Gaussian distribution for the probabilities, the terms $\alpha$ and $\beta$ can be defined as:

$$\alpha^{MP} = \frac{\gamma}{2E_w(x^{MP})} \tag{3}$$

$$\beta^{MP} = \frac{N-\gamma}{2E_D(x^{MP})} \tag{4}$$

In these equations, the term $x^{MP}$ represents the weights that maximize the posterior density and $\gamma$ is the effective number of parameters given by:

$$\gamma = n - 2\alpha^{MP} tr(H^{MP})^{-1} \tag{5}$$

where $H^{MP}$ is the Hessian matrix evaluated at $x^{MP}$ and $n$ is the total number of parameters.

## 3 Experimental procedure

In this paper, a three-layer feedforward artificial neural network based on the Bayesian regularization (BNN) was presented to predict the pitting corrosion behavior of austenitic stainless steel. The topology of the model consisted of three units in the

input layer corresponding to the environmental factors and two neurons in the output layer related to the pitting corrosion status for each sample. The number of neurons in the hidden layer was chosen empirically varying from 1 to 20. The values of the weight and bias were adjusted according to the optimization algorithm of Levenberg-Marquadt [15]. In order to improve the training process, the original data set was normalized within the range [-1,1]. In this way, the error was reduced by scaling the variables. On the other hand, 5-fold cross validation technique was applied to validate the model. This procedure divides the original data set into 5 different groups: 4 of the sets (80% of the original patterns) were used to create the model, whereas the remaining group (20% of the original patterns) was applied to test the model. The performance index values were obtained taking the average of all repetitions (20 times) for each configuration in order to select the model with the best generalization capability.

## 4 Results and discussion

In order to select the optimal structure for the model, different statistic metrics have been measured for all the configurations [18]:

$$Sensitivity = \frac{TP}{TP + FN} \tag{6}$$

$$Specificity = \frac{TN}{TN + FP} \tag{7}$$

$$\text{Precision} = \frac{TP}{TP + FP} \cdot 100 \tag{8}$$

$$\text{Accuracy} = \frac{TP + TN}{P + N} \cdot 100 \tag{9}$$

Where *TP* is the corrosion patterns classified correctly and *FP is* the corrosion patterns misclassified. *TN* is the no-corrosion patterns classified correctly and *FN*, the no-corrosion patterns wrong classified. The results were collected in Fig. 2.

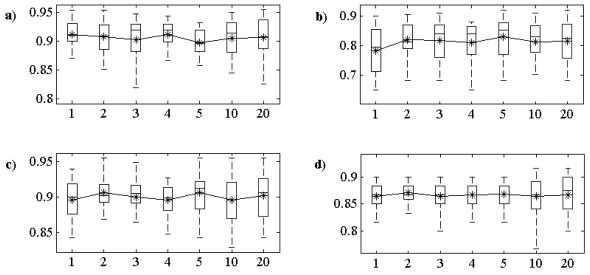



**Fig. 2.** Boxplot for the statistic metrics evaluated to analyze the influence of the number of hidden neurons on the classification performance: a) sensitivity, b) specificity, c) precision d) accuracy.

Based on the high values for the performance indices represented in the figure, it can be reflected the capacity of the BNN models to predict pitting status of 316L under different environmental conditions. However, in order to determine the optimal configuration model, the use of the Receiver Operating Characteristics (ROC) space was presented [19]. This analysis is evaluated by plotting the true positive fraction (sensitivity) versus the true negative fractions (1-specificity), see Fig.3. In this way, each model can be represented by a point on a 2-D plot. Therefore, the best models are those located closest to the upper left corner.

According to Fig.3, the best configurations may be the models with 2 or 5 hidden neurons since these models were the closest to the upper left corner. Considering both configurations, the optimal model can be selected according to the Occam razor's: the best model is the simplest one. Based on this principle, the optimal configuration for the BNN model to predict pitting corrosion status of austenitic stainless steel was obtained using two hidden neurons.

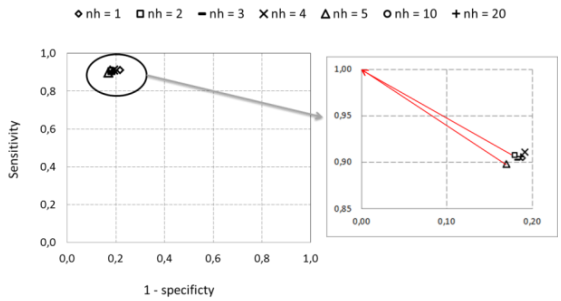



**Fig. 3.** Sensitivity and specificity plotted in ROC space for models to predict pitting corrosion behavior

This model provided excellent results with notable precision (90.7%). Fig. 4 depicts the pitting corrosion modelling of 316L stainless steel by using BNN with 2 hidden neurons.

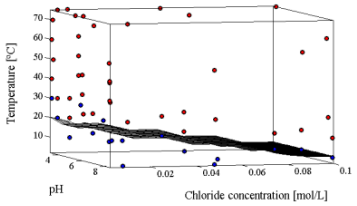



**Fig. 4.** Corrosion prediction modelling using BNN (3:2:2). Points in red are original corrosion patterns whereas points in blue are original no-corrosion patterns. The borderline between the corrosion and no corrosion patterns predicted by the model is represented.

## 5 Conclusions

The knowledge of the influence of the environmental conditions on the pitting corrosion behaviour of stainless steel is a critical task. The electrochemical tests are used to evaluate pitting corrosion susceptibility of the materials. However, the human interpretation in the microscopic analysis of the material surface, needed after the tests, leads to subjectivity in the results. To deal with this drawback, a model based on ANN using the Bayesian regularization was presented in this work in order to predict pitting behaviour. The results demonstrated the utility of the proposed method to be considered as alternative to electrochemical tests in order to predict corrosion behaviour of stainless steels automatically. The BNN with two hidden neurons allowed the prediction of the pitting corrosion behaviour by an automatic way with notable values of precision (90.7%). This model was presented an effective tool to analyse the pitting corrosion resistance of stainless steel according to the environmental conditions.

**Acknowledgments.**

Authors greatly appreciate the technical support provided by ACERINOX S.A in addition to Transfer Entity agent Andalusian Knowledge System (Fundación Campus Tecnológico Bahía de Algeciras).

**References**

[1] Peckner, D., Bernstein, I.M: Handbook of Stainless Steels, McGraw-Hill, 1977.
[2] Baddoo, N.: Journal of Constructional Steel Research. 64 (2008) 1199.
[3] McGuire, M.F.: Stainless Steels for Design Engineers, ASM International, 2008.
[4] Padhy, N., Ningshen, S., Kamachi Mudali, U., Raj, B.: Scr. Mater. 62 (2010) 45.
[5] Pletcher, D., Westcott, C.: Electrochemistry. 10 (2007) 85.
[6] Cottis, R.A., Li, Q., Owen, G. et al.: Mater. Des. 20 (1999) 169.
[7] Ok, D., Pu, Y., Incecik, A.: Ocean Eng. 34 (2007) 2222.
[8] Boucherit, M. Amzert, S., Arbaoui, F. et al.: Anti-Corros Methods Mater 55 (2008) 115.
[9] Ramana, K.V.S., Anita, T., Mandal, S., Kahappan, S. et al.: Mater Des. 30 (2009) 3770.
[10] Cavanaugh, M., Birbilis, N., Buchheit, R.: Electrochim Acta 59 (2012) 336.
[11] Jiménez-Come, M.J, Turias, I.J., Moscoso, J., Trujillo, F.: Materials and Corrosion 65 (2014) 881.
[12] Jiménez-Come, M.J., Muñoz, E., García, R. et al.: Journal of Applied Logic 10 (2012) 4.
[13] Hornik, K.: Neural Networks 2 (1989) 358.
[14] Bishop, C.M.: Pattern recognition and machine learning, Springer New York, 1995.
[15] Rumelhart, D.E., Hintont, G.E., Williams, R.J.: Nature 323 (1986) 533.
[16] MacKay, D.J.: Neural Computation. 4 (1992) 415.
[17] MacKay, D.J.: Neural Computation. 4 (1992) 448.
[18] Tikhonov, A.: Soviet Math. Dokl. 5 (1963) 1035.
[19] Fawcett, T.: Pattern Recog Lett. 27 (2006) 861.

# Convex Optimisation Problems in Bioengineering

Clemente Cobos Sanchez, Francisco Garcia-Pacheco, Jose Maria Guerrero Rodriguez, Angel Quiros Olozabal, and German Alvarez Tey

Depto. Ingeniería de Sistemas y Electrónica,
Depto. de Matemáticas Universidad de Cádiz, Cádiz, Spain
{clemente.cobos,garcia.pacheco,josem.guerrero,angel.quiros,german.alvarez}@uca.es

**Abstract.** *Many problems in engineering require to determine the spatial distribution of electric currents flowing on a conductive surface, which must satisfy some given requirements for the produced fields, electromagnetic energy, etc. The reconstruction of current distribution on the conducting surface subjected to these constraints is an inverse problem, which when formulated using boundary element methods can be posed as a convex optimisation. Here we present a convex optimisation framework to tackle problems in Bioengineering, that permits the prototyping of many different cost functions and constraints. Several examples of MRI gradients and TMS coils were designed and simulated to demonstrate the validity of the proposed approach.*

**Keywords:** Convex optimisation, Boundary element method, Bioengineering, Field synthesis

## 1 Introduction

Magnetic Resonance Imaging (MRI) has become an invaluable tool for diagnostic medicine. It is based on the use of well defined and controlled magnetic fields, as the magnetic field gradients, used to encode spatially the signals from the sample. These field gradients are generated by coils of wire, usually placed on cylindrical surfaces, although other geometries can be employed [1].

Transcranial Magnetic Stimulation (TMS) is a non-invasive technique to stimulate the brain [2], which is applied to studies of cortical effective connectivity, presurgical mapping, psychiatric and medical conditions, such as major depressive disorder, schizophrenia, bipolar depression, post-traumatic, stress disorder and obsessive-compulsive disorder, amongst others. In TMS, a strong, brief current pulse driven through a coil is used to induce an electric field stimulating neurons in the cortex.

The problem in MRI gradient coil and TMS coil design is to find optimal positions for the multiple windings of coils so as to produce fields with the desired spatial dependence and properties (low inductance, high gradient to current ratio, minimal resistance, good field gradient uniformity, high focality and field

penetration depth, etc.). We refer the reader to [1]-[3] for a wider perspective on these topics.

TMS and MRI coil design are then two examples of problems in bioengineering where is required to determine the spatial distribution of electric currents flowing on a conductive surface, which satisfies given requirements for the produced fields, electromagnetic energy, etc.

An appropriate and realistic formulation of this type of problems can be achieved by using a boundary element method (BEM), and incorporating the idea of stream function. The current density in the surface is then a vector field that is piecewise uniform (see [4, 6]). By using this current model, electromagnetic inverse problems, such as MRI and TMS coil design, can be formulated as a constrained optimization.

In this work, we present a convex optimisation framework for the solution of electromagnetic inverse problems in Bioengineering, such as MRI gradient and TMS coil design problem, allowing the prototyping of many different cost functions and constraints. Two examples of gradient and TMS coils were designed and simulated to demonstrate this method, and prototypes coils were built and tested to validate it.

## 2 Physical Model

A model of the current under search can be achieved by using a boundary element method (BEM), that allows the current distribution to be defined in terms of the nodal values of the stream function and elements of the local geometry (see [4]).

So let us assume that the surface, $S$, on which we want to find the optimal current, is divided into $T$ triangular elements with $N$ nodes, which are lying at each vertex of the element. We can then note the vector containing the nodal values of the stream function as $\psi \in \mathbb{R}^N$, which is going to be the optimization variable in this work.

The use of this current model allows the discrete formulation of all the magnitudes and physical properties of the coil involved in the design. All problems tackled here are convex and can be generalised as

$$\begin{cases} \text{minimise } f_0(\psi) \\ \text{subject to } f_i(\psi) \leq b_i, \;\; 1 \leq i \leq m \end{cases} \tag{1}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex functions for $i = 1, ..., m$.

## 3 MRI

### 3.1 Minimum inductance (stored energy) coil

The quality of the MRI images strongly depend on how linear the variation of the field is with position. Analogously, in order to improve the image formation process ideal coils should have a minimum inductance, which can be related to

the stored magnetic energy and dictates the speed at which current can be put into the coil.

The problem of designing a MRI gradient coil with good field gradient uniformity and minimum inductance can be posed as [4]

$$\begin{cases} \text{minimise } \psi^T L \psi \\ \\ \text{subject to } \dfrac{\|B_z \psi - b_t\|_\infty}{\|b_t\|_\infty} \leq D_{max} \end{cases}$$

where

- $H \in \mathbb{N}$, with $N > H$, is the number of points where the target field is defined.
- $B_z \in \mathbb{R}^{H \times N}$ is a known matrix, where the coefficient $B_z(i,j)$ is the z-component of the magnetic induction produced by the current element associated to the $j^{th}$-node in the prescribed $i^{th}$-point.
- $b_t \in \mathbb{R}^H$ is the target field, prescribed in the $H$ points.
- $L \in \mathbb{R}^{N \times N}$ is the inductance matrix, which is symmetric and positive-definite.
- The magnetic field is required to deviate by less than a given value from linearity, usually $D_{max} \sim 5\%$.

## 4 TMS

### 4.1 Minimum stored energy coil

Power requirements often limit the duration and frequency of repetitive TMS, for example, via coil heating. Thus, an ideal TMS coil should produce a strong stimulation of a prescribed region, and a minimum electric field in the rest of non target regions; and have a minimum stored magnetic energy.

The problem of designing a TMS coil that produces a maximum field in a target region with minimum stored energy can be posed as

$$\begin{cases} \text{maximise } \|B\psi\|_2 \\ \text{minimise } \psi^T L \psi \end{cases} \tag{2}$$

where

- $H \in \mathbb{N}$, with $N > H$, is the number of points where the target field is defined.
- $B \in \mathbb{R}^{H \times N}$ is a known matrix, where the coefficient $B(i,j)$ is the modulus of the magnetic induction produced by the current element associated to the $j^{th}$-node in the prescribed $i^{th}$-point.
- $b_t \in \mathbb{R}^H$ is the target field prescribed in the $H$ points, which in TMS can be considered constant, that is, $b_t(i) = b_t(j)$ for all $i, j = 1, ..., H$.
- $L \in \mathbb{R}^{N \times N}$ is the inductance matrix, which is symmetric and positive-definite.

This type of problem is also known as Tikhonov regularised minimisation. According to Subsection 5.6 and *1* of Corollary 1, this problem can be equivalently written as

$$\begin{cases} \text{maximise} \left\| \left(BC^{-1}\right) \psi \right\|_2 \\ \text{subject to} \left\| \psi \right\|_2 = 1 \end{cases}$$

where $L = C^T C$ is the Cholesky decomposition of the inductance matrix $L$ (which is symmetric and positive-definite).

Other admissible reformulations (see *2* of Corollary 1) are

$$\begin{cases} \text{minimise} \left\| \psi \right\|_2 \\ \text{subject to} \left\| \left(BC^{-1}\right) \psi \right\|_2 = \left\| \left(BC^{-1}\right) \right\|_2 \end{cases}$$

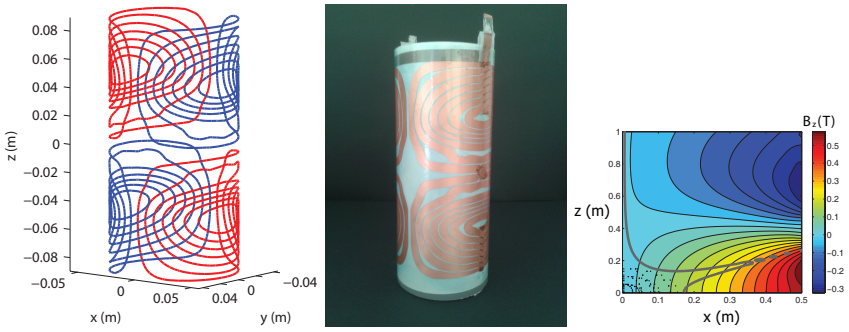
or more generally

$$\begin{cases} \text{minimise} \left\| \psi \right\|_2 \\ \text{subject to} \left\| \left(BC^{-1}\right) \psi - b_t \right\|_2 = \left\| \left(BC^{-1}\right) \right\|_2 \end{cases}$$

for $b_t$ constant.

## 5 Results

### 5.1 MRI




(a) (b) (c)

**Fig. 1.** a) equivalent numerical model of coil in Fig. 5.1 (red and blue colors are used to indicate wires in which there is a different sense of current flow). b) Photograph of the constructed prototype coil. c) Contours of the $B_z$ field produced by the wire arrangement Fig 5.1. The grey line delineates the region where the field deviates by less than 5% from linearity.

A prototype cylindrical transverse MRI gradient coil has been designed using the proposed convex framework (Fig. 5.1), it corresponds to a minimum inductance transverse cylindrical coil of radius 4.5 cm and height 18 cm, designed to produce a field gradient which deviates from linearity by less than 5% within a central, uniform distribution of 400 points spread over a sphere of radius of 3.5 cm. This prototype coil has been constructed using a variable track width produced in a flexible printed circuit board (PCB), where the copper thickness was 35 $\mu$m. The former on which the prototype tracks were laid was a cylinder of polyvinylchloride (PVC) with 3.4 mm thickness.

Figure 5.1 shows the gradient coil prototype, that when was connected to 6.0 A DC current supply (Agilent U8031A, USA), produces the magnetic field in Fig. 5.1. This coil when energized produce the $B_z$ field displayed in Fig. 5.1, which satisfies the initial requirements, as we can see the target region is within the grey line that delineates the volume where the field deviates by less than 5% from linearity.

### 5.2 TMS

The proposed approach has also been used to produce a TMS stimulator on a rectangular former of dimensions 20 cm $\times$ 10 cm, designed to have minimum stored energy and to maximize the magnetic field in a prescribed spherical volume of interest of radius 2.0 cm that is centred 4.0 cm below the center of the coil plane, as shown in Fig. 5.2 where the coil solution is also depicted.

In order to validate the proposed solution, we manufactured the corresponding prototype coil, which was wound with 1.5 mm thick continuous copper wire, Fig. 5.2. This TMS prototype when connected to 6.0 A DC current supply (Agilent U8031A, USA), produces the magnetic field in Fig. 5.2, which was measured using a magnetic flow sensor MAG 3100. The results obtained here indicates that the prototype TMS coil produces a remarkably high magnetic field in the target volume that decreases rapidly out of the volume.

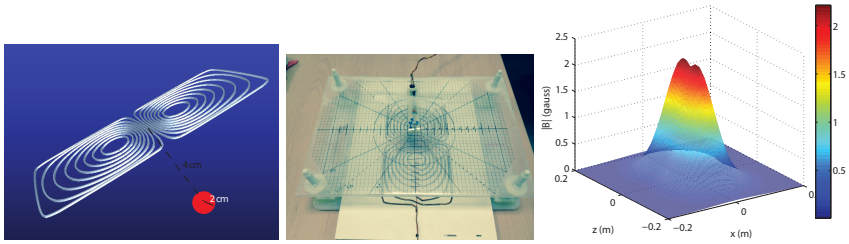
### 5.3 Numerical Implementation

Software was written in Fortran 90 to tackle the problems presented here, and the produced optimal solutions were found in good agreement with those generated using software for convex programming such us CVX [7] The Fortran 90 software also includes subroutine that allows the testing of the coil designs, as it calculates the field produced by the wire pattern via Biot-Savart integration.

### 5.4 Conclusion

Here we present a convex optimisation framework to tackle problems in Bioengineering. This is a powerful approach for designing of MRI gradient, and novel method for the generation of TMS coils wounds on arbitrarily shaped surface.

The method has been experimentally validated by constructing and testing prototype coils, where the magnetic fields produced show the accuracy of the proposed technique.




(a) (b) (c)

**Fig. 2.** a) Schematic diagram showing the TMS coil solution and along with the region of interest in which the desired magnetic field must be maximized. b) Photograph of the constructed TMS prototype coil and the experimental set up to measure the magnetic field with the flow sensor MAG 3100. c) Magnetic field modulus in a 20 cm $\times$ 20 cm plane which is centred 4.0 cm below the center of the coil plane.)

## References

1. Turner, R.: Gradient coil design: A review of methods, Magn. Reson. Imaging, 11, 903 – 920(1993).
2. Wassermann E. M., Epstein C. M., Ziemann U., Walsh V., Paus T., Lisanby S. H., editors. The Oxford handbook of transcranial magnetic stimulation. New York: Oxford University Press; 2008
3. Koponen L. M., Nieminen J. O., Ilmoniemi R.J. Minimum-energy coils for transcranial magnetic stimulation: application to focal stimulation. Brain Stimul 2014.
4. Cobos Sanchez, C., S. G. Garcia, L. D. Angulo, C. M. De Jong Van Coevorden, and A. Rubio Bretones. A divergence-free BEM method to model quasi-static currents: Application to MRI coil design. Progress In Electromagnetics Research B, Vol. 20, 187–203, 2010.
5. Poole M., Weiss P., Lopez H.S., Ng M., Crozier S., Minimax current density coil design, J. Phys. D: Appl. Phys. 43 (9) (2010) 095001. 13pp.
6. Poole M., Jon Shah N. Convex optimisation of gradient and shim coil winding patterns. Journal of Magnetic Resonance, Volume 244, 2014, Pages 36-45.
7. CVX Research, Inc., CVX: Matlab Software for Disciplined Convex Programming, version 2.0, (August 2012).

## Appendix: Mathematical foundations of the previous models

This appendix is devoted to provide the mathematical foundations needed to express the previous problems in the form given in Equation (1).

### 5.5 Brief introduction and background

Let $A$ be an $H \times N$ matrix and consider a norm $\|\cdot\|$ in $\mathbb{R}^N$. For every $\varepsilon \geq 0$ we can consider the closed $\|\cdot\|$-ball of center 0 and radius $\varepsilon$, which is a compact subset of $\mathbb{R}^N$:

$$\mathsf{B}_{\|\cdot\|}(0,\varepsilon) = \{\psi \in \mathbb{R}^N : \|\psi\| \leq \varepsilon\}.$$

When $\varepsilon = 1$ then we will simply write $\mathsf{B}_{\|\cdot\|}$ in lieu of $\mathsf{B}_{\|\cdot\|}(0,1)$. It is obvious that $\mathsf{B}_{\|\cdot\|}(0,\varepsilon) = \varepsilon\mathsf{B}_{\|\cdot\|}$ for all $\epsilon \geq 0$.

It can be proved that the sup of $A$ on the previous ball is attained at an element of its sphere. In other words,

$$\max\{\|A\psi\| : \|\psi\| \leq \varepsilon\} = \max\{\|A\psi\| : \|\psi\| = \varepsilon\} = \sup\{\|A\psi\| : \|\psi\| < \varepsilon\}.$$

We will denote by $\exp(A,\varepsilon)$ to the set of all those elements of the sphere of the ball above at which the previous max is attained. In other words,

$$\exp_{\|\cdot\|}(A,\varepsilon) := \left\{\varphi \in \mathsf{B}_{\|\cdot\|}(0,\varepsilon) : \|A\varphi\| = \max\{\|A\psi\| : \|\psi\| \leq \varepsilon\}\right\}.$$

Again, when $\varepsilon = 1$ we will write $\exp_{\|\cdot\|}(A)$ instead of $\exp_{\|\cdot\|}(A,1)$. On the other hand, note that $\exp_{\|\cdot\|}(A,\varepsilon) = \varepsilon \exp_{\|\cdot\|}(A)$ for all $\varepsilon \geq 0$.

The norm of the matrix $A$ is by definition $\|A\| := \max\{\|A\psi\| : \|\psi\| \leq 1\} = \max\{\|A\psi\| : \|\psi\| = 1\}$ and thus the elements of $\exp_{\|\cdot\|}(A)$ are precisely the vectors of $\mathbb{R}^N$ at which $A$ attains its norm (these vectors will be called the supporting vectors of $A$). It can be proved that $\|A\chi\| \leq \|A\|\|\chi\|$ for all $\chi \in \mathbb{R}^N$.

Recall that a (real) scalar product on $\mathbb{R}^N$ is defined by a positive-definite symmetric matrix $P$ as $(\varphi,\psi) := \varphi P\psi$. This scalar product makes $\mathbb{R}^N$ a Hilbert space whose norm is $\|\varphi\|_P := (\varphi,\varphi)^{\frac{1}{2}}$. It is well known among the functional analysts that all the Hilbert spaces of the same dimension are linearly isometric, which means that there exists a surjective linear isometry $T_P : \left(\mathbb{R}^N, \|\cdot\|_P\right) \to \left(\mathbb{R}^N, \|\cdot\|_2\right)$. If we keep denoting by $T_P$ to the matrix associated to the isometry $T_P$, then $\|T_P\psi\|_2 = \|\psi\|_P$ for all $\psi \in \mathbb{R}^N$. In matrix theory, the expression of $T_P$ is given by the Cholesky decomposition of the symmetric positive-definite matrix $P$, that is, $P = L^T L$. Indeed, notice that

$$\|T_P\psi\|_2^2 = \|\psi\|_P^2 = \psi^T P\psi = \psi^T(L^T L)\psi = (L\psi)^T(L\psi) = \|L\psi\|_2^2$$

which means that we can take $T_P := L$.

### 5.6 Uniformizing norms

Consider the optimization problem given in Equation 2

$$\begin{cases} \max \|B\psi\|_2 \\ \min \psi^T L\psi \end{cases} \quad \psi \in \mathbb{R}^N$$

If we assume that $L$ is symmetric and positive-definite, then there exists a matrix $T_L$ such that $\psi^T L\psi = \|T_L\psi\|_2$. Therefore, the above problem becomes

$$\begin{cases} \max \|B\psi\|_2 \\ \min \|T_L\psi\|_2 \end{cases} \quad \psi \in \mathbb{R}^N$$

Taking into consideration that $T_L$ is invertible, we can rewrite it as

$$\begin{cases} \max \|A\varphi\|_2 \\ \min \|\varphi\|_2 \end{cases} \quad \varphi \in \mathbb{R}^N$$

where $A := BT_L^{-1}$.

### 5.7 Turning maximization problems into minimization problems (without losing convexity)

The following results throws some light on the problem described in Equation (2) in the sense that conjugation of the two conditions of maximization and minimization makes it an unsolvable problem.

**Theorem 1.** *Let $A$ be an $H \times N$ matrix and consider a norm $\|\cdot\|$ in $\mathbb{R}^N$. Consider the optimization problem*

$$\begin{cases} \max \|A\varphi\| \\ \min \|\varphi\| \end{cases} \varphi \in \mathbb{R}^N$$

1. *There does not exist $\varphi \in \mathbb{R}^N$ such that, for all $\psi \in \mathbb{R}^N$, $\|A\varphi\| \geq \|A\psi\|$ and $\|\varphi\| \leq \|\psi\|$.*
2. *There are infinitely many $\varphi \in \mathbb{R}^N$ such that, for all $\psi \in \mathbb{R}^N$, either $\|A\varphi\| \geq \|A\psi\|$ or $\|\varphi\| \leq \|\psi\|$. These solutions are the elements of the set*

$$\bigcup_{\varepsilon \geq 0} \exp_{\|\cdot\|}(A, \varepsilon).$$

*Proof.*

1. Suppose to the contrary that there is $\varphi \in \mathbb{R}^N$ such that $\|A\varphi\| \geq \|A\psi\|$ and $\|\varphi\| \leq \|\psi\|$ for all $\psi \in \mathbb{R}^N$. Since $\|\varphi\| \leq \|\psi\|$ for all $\psi \in \mathbb{R}^N$ we must have that $\varphi = 0$ which then contradicts that $\|A\varphi\| \geq \|A\psi\|$ for all $\psi \in \mathbb{R}^N$ since $A\varphi = 0$.
2. In the first place, assume that there exists $\varphi \in \mathbb{R}^N$ such that $\|A\varphi\| \geq \|A\psi\|$ or $\|\varphi\| \leq \|\psi\|$ for all $\psi \in \mathbb{R}^N$. We will show that $\varphi \in \exp_{\|\cdot\|}(A, \varepsilon)$ for $\varepsilon := \|\varphi\|$. Indeed, let $\psi \in \mathbb{R}^N$ such that $\|\psi\| < \varepsilon\,(= \|\varphi\|)$. Then by assumption $\|A\varphi\| \geq \|A\psi\|$, which means that $\|A\varphi\| \geq \sup\{\|A\psi\| : \|\psi\| < \varepsilon\} = \max\{\|A\psi\| : \|\psi\| \leq \varepsilon\} \geq \|A\varphi\|$ since the open ball $\{\psi \in \mathbb{R}^N : \|\psi\| < \varepsilon\}$ is dense in the closed ball $\{\psi \in \mathbb{R}^N : \|\psi\| \leq \varepsilon\}$. As a consequence, $\varphi \in \exp_{\|\cdot\|}(A, \varepsilon)$.
   Conversely, we will show that every $\varphi \in \exp_{\|\cdot\|}(A, \varepsilon)$ verifies that $\|A\varphi\| \geq \|A\psi\|$ or $\|\varphi\| \leq \|\psi\|$ for all $\psi \in \mathbb{R}^N$. Indeed, fix any $\psi \in \mathbb{R}^N$. If $\varepsilon \leq \|\psi\|$, then we are done because $\|\varphi\| = \varepsilon$. If $\varepsilon > \|\psi\|$, then we conclude that

$$\|A\varphi\| = \max\{\|A\chi\| : \|\chi\| \leq \varepsilon\} \geq \|A\psi\|.$$

**Corollary 1.** *Let $A$ be an $H \times N$ matrix and consider a norm $\|\cdot\|$ in $\mathbb{R}^N$. The optimization problem*

$$\begin{cases} \max \|A\varphi\| \\ \min \|\varphi\| \end{cases} \varphi \in \mathbb{R}^N$$

*is equivalent to any one of the following:*

1. *The optimization problem*

$$\begin{cases} \max \|A\varphi\| \\ \|\varphi\| = 1 \end{cases} \varphi \in \mathbb{R}^N$$

*which consists of finding the elements of $\exp_{\|\cdot\|}(A)$, that is, the elements of $\mathbb{R}^N$ at which $A$ attains its norm.*

2. *The convex optimization problem*

$$\begin{cases} \min \|\varphi\| \\ \|A\varphi\| = \|A\| \end{cases} \varphi \in \mathbb{R}^N$$

*which again consists of finding the supporting vectors of $A$.*

*Proof.*

1. In accordance to Theorem 1, the solutions to the optimization problem

$$\begin{cases} \max \|A\varphi\| \\ \min \|\varphi\| \end{cases} \varphi \in \mathbb{R}^N$$

are the elements of the sets

$$\bigcup_{\varepsilon \geq 0} \exp_{\|\cdot\|}(A, \varepsilon).$$

Since $\exp_{\|\cdot\|}(A, \varepsilon) = \varepsilon \exp_{\|\cdot\|}(A)$ for all $\varepsilon > 0$, it suffices to find the supporting vectors of $A$, that is, the elements of $\exp_{\|\cdot\|}(A)$, which are precisely the solutions to the

$$\begin{cases} \max \|A\varphi\| \\ \|\varphi\| = 1 \end{cases} \varphi \in \mathbb{R}^N$$

2. All we need to show is that the solutions to the convex optimization problem

$$\begin{cases} \min \|\varphi\| \\ \|A\varphi\| = \|A\| \end{cases} \varphi \in \mathbb{R}^N$$

are the supporting vectors of $A$. Indeed, let $\varphi \in \exp_{\|\cdot\|}(A)$. If $\psi \in \mathbb{R}^N$ and $\|A\psi\| = \|A\|$, then we have that

$$\|\varphi\| = 1 = \frac{\|A\psi\|}{\|A\|} \leq \frac{\|A\|\|\psi\|}{\|A\|} = \|\psi\|$$

which implies that $\varphi$ is a solution of the convex minimization problem.
Conversely, let $\varphi$ a solution of the convex minimization problem. We will prove that $\varphi \in \exp_{\|\cdot\|}(A)$. By assumption, $\|A\varphi\| = \|A\|$ so all we need to show is that $\|\varphi\| = 1$. Indeed, it suffices to consider any $\psi \in \exp_{\|\cdot\|}(A)$. Since $\|A\psi\| = \|A\|$ we have that $\|\varphi\| \leq \|\psi\| = 1$. The other inequality follows from the fact that $\|A\| = \|A\varphi\| \leq \|A\|\|\varphi\|$.

# Formalization in ACL2 of Matrix Algebra Basic Concepts*

J.L. Pro–Martín , J.L. Ruiz–Reina and F.J. Martín–Mateos
jlpro@modinem.com, fjesus@us.es, jruiz@us.es

Computational Logic Group
Dept. of Computer Science and Artificial Intelligence, University of Seville
E.T.S.I. Informática, Avda. Reina Mercedes, s/n. 41012 Sevilla, Spain

**Abstract.** In this paper we present a formalization of basic operations on matrices in the ACL2 theorem prover, including addition, product, transpose, and inverse of matrices. We define these operations and give proofs of their main properties. The main result is an ACL2 implementation and formal verification of the Gauss-Jordan algorithm for computing the inverse of a matrix. Our formalization is based on *abstract stobjs*, an ACL2 feature allowing both convenient logical reasoning and execution efficiency. In fact, we get quite good execution time response for big matrices of several hundred thousands elements. The complete formalization can be found in [1].

## 1 Formalizing matrices in ACL2

The ACL2 system [4] is both a programming language, a logic for reasoning about programs defined in the language, and a mechanical theorem prover to assist in the proof process. The programming language is an extension of an applicative subset of Common Lisp and the logic is quantifier-free, first-order with equality, including a rule of inference of proof by structural induction. See [2] for a detailed description of ACL2.

ACL2 has been widely used in formal validation of hardware algorithms and in many math fields. As for matrix algebra, a previous work is [3], where a formalization of the Strassen algorithm for matrix multiplication is presented, using applicative *record structures*. Also, in [5] a number of operations (including inverse) are defined and some of their properties proved, using ACL2 bidimensional arrays. Since one of our main concerns is to improve execution efficiency, it is tempting to apply some interesting ACL2 features designed for that purpose, in a formalization of matrix algebra. Among these features are *stobjs* (for *Single Threaded OBJects*) and *abstract stobjs*.

In principle, if we use a list-based representation of matrices, accessing and updating cannot be done in constant time, because we need copying, as usual in an applicative setting. This feature is critical if we want to define algorithms

* This work was partially supported TIN2013-41086-P project (Spanish Ministry of Economy and Competitiveness), cofinanced with FEDER funds

dealing with big matrices. Fortunately, ACL2 stobjs allow lookups and (destructive) updates in constant time. When an object is declared to be single-threaded, ACL2 enforces certain syntactic restrictions on its use, ensuring that only one copy of the object is ever needed. In this way, this efficient data structure is consistent with the applicative semantics of ACL2.

Stobjs are structures composed by fields. These fields can store a single data or *unidimensional* arrays. But a matrix is intuitively represented by a two-dimensional array. For example, element $a_{ij}$ of matrix $A$ maps to element of the $i$th-row and $j$th-column. This use of two indices to access matrix elements is so extended in mathematics that we should preserve it in our formalization.

But here arises the very first problem: currently *stobj's* only allow one-dimensional arrays, so we define this stobj in ACL2 (the suffix "`...$c`" stands for *concrete* and will be explained later):

```
(defstobj matrix$c
 (m$c        :type (array rational (1)) :initially 0 :resizable t)
 (nrows$c    :type (integer 1 *) :initially 1)
 (ncolumns$c :type (integer 1 *) :initially 1))
```

Where `m$c` is the one-dimensional array that supports matrix elements, `nrows$c` and `ncolumns$c` gives us the number of rows and columns of the matrix defined above.

From the `:LOGIC` point of view a *stobj* will be represented by means of a list (made with *conses*) of elements. Also an array field of an stobj, from the logic point of view, is represented as a list of elements. For example, if we have $A$ the following $2 \times 3$ matrix, in the logic we would have the representation:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \longrightarrow \texttt{((1 2 3 4 5 6) 2 3)}$$

If we denote $c$ the number of columns of $A$, $r$ the number of rows of $A$ and $B$ the contents of the array field, it's easy to see the following mappings between elements in both representations:

1. $a_{ij} \mapsto b_{i \cdot c + j}$
2. $b_k \mapsto a_{\lfloor k/r \rfloor, k \mod r}$

So in the `:LOGIC` we would have to reason with this additional disadvantage of dealing with these mappings. This would increase the difficulty of our proofs and decrease their clarity. Fortunately, abstract stobjs allows us to use *stobj* in the `:EXEC` world but two-dimensional access in `:LOGIC` world.

### 1.1 *Abstract stobj's* in ACL2

*Abstract stobj's* (as opposed to *concrete stobj's*) is a relatively new ACL2 feature introduced in [6]. An abstract stobj provides an alternative logical representation of a concrete stobj. That is, we can define a *simpler* logical representation of

the concrete *stobj* in order to abstract its complexity in terms of reasoning. In our case, we will be able to give a natural and two-dimensional access matrix representation in the `:LOGIC` and a more complex but efficient one-dimensional array representation in the `:EXEC` for our matrices.

We introduce an *abstract stobj* using the `defabsstobj` event. Two functions are needed in this event: a *recognizer* and a *correspondence* function. In order to be finally admitted, some proof-obligations are generated:

1. `:correspondence`. Theorems establishing that abstract and concrete stobj's represent the same inner values in the terms of the correspondence function.
2. `:preserved`. Theorems that prove that updaters maintains the recognizer function.
3. `:guard-thm`. Theorem proving that guards are verified in every call in the `:EXEC` world.

These proof-obligations have to be proved once and for all, ensuring that the logical correspondence between the abstract stobj and its associated concrete stobj is preserved. So we can define the following abstract stobj:

```
(defabsstobj matrix
    :concrete matrix$c
    :recognizer (matrixp    :logic matrix$ap :exec matrix$cp)
    :creator (create-matrix :logic create-matrix$a
                            :exec create-matrix$c)
    :corr-fn matrix$corr
    :exports ((nrows    :logic nrows$a    :exec nrows$c)
              (ncolumns :logic ncolumns$a :exec ncolumns$c)
              (lookup   :logic lookup$a   :exec lookup$c)
              (update   :logic update$a   :exec update$c)
              (redim    :logic redim$a    :exec redim$c)))
```

Where the suffix "`...$a`" stands now for *abstract*. We can see here the primitives defined for our matrix object where their names (`nrows`, `ncolumns`, and so on) are some kind of self-explanatory about its behaviour. Let's see with some detail the function `(lookup A i j)`, that returns the $a_{ij}$ element. We must define the `:EXEC` and `:LOGIC` version of `lookup`:

```
(defun lookup$c (matrix$c i j)
    (m$ci (+ (* i (ncolumns$c matrix$c)) j) matrix$c))

(defun lookup$a (matrix$a i j)
    (nth j (nth i matrix$a)))
```

So to establish that `lookp$c` and `lookup$a` return the same value we must prove the following theorem in ACL2 (we have removed some not relevant conditions in the clause for the sake of clarity):

```
(defthm lookup{correspondence}
    (implies (matrix$corr matrix$c matrix)
             (equal (lookup$c matrix$c i j)
                    (lookup$a matrix i j))))
```

This theorem can be read as follows: "if `matrix$c` and `matrix` correspond to each other, `lookup` will return the same value". Once this is proved, we can use the primitives with the desired logic interface. The rest of the functions that use matrices must do it by means of those primitives.

## 2 Defining operations over matrices

We have defined some operations over matrices using, as explained, the given primitives in the `defabsstobj` event described above. Experience in the use of ACL2 prover tells us that the way one defines functions can dramatically change proofs complexity, so we tried to define functions in such a simple way that principle of induction could perform quite well.

For example, we can define the matrix addition like this:

```
(defun add-matrix-row (A B m n)
    (if (zp n)
        (update A m 0 (+ ( lookup A m 0) ( lookup B m 0)))
        (seq A
            (update A m n (+ (lookup A m n)
                             (lookup B m n)))
            (add-matrix-row A B m (1- n)))))

(defun add-matrix-rows (A B m n)
    (if (zp m)
        (add-matrix-row A B 0 n)
        (seq A
            (add-matrix-row A B m n)
            (add-matrix-rows A B (1- m) n))))

(defun add-matrix (A B)
    (add-matrix-rows A B (1- (nrows A))
                         (1- (ncolumns A))))
```

So, `add-matrix-row` adds the elements of `m` row, from 0 to `n` column. The next function, `add-matrix-rows`, note the plural, adds rows from 0 to `m`, and each of this, from column 0 to `n`. Finally, we have `add-matrix` function, that only makes first calling to last function with proper arguments to start the computation. This way of defining functions can be considered as a pattern that can be applied to other functions such as matrix equality, transposed matrix, matrix product and so on.

It turns out that this recursive pattern is specially well-suited when we prove properties of these functions by induction. That is, first we prove the property on only a given row of the matrix, by induction in the number of columns. Afterwards we prove it in a subset of rows of the matrix, to finally be able to prove it for the whole matrix. It is worth mentioning that almost every property (with some hard exceptions) can be proved in this way.

### 2.1 List of some proved properties

We now list some of the main properties proved in our formalization. For details, see the whole formalization in [1].

**Transposition**
$(A^T)^T = A$
$I_n^T = I_n$

**Opposite**
$-(-A) = A$
$(-A)^T = -(A^T)$

**Scalar product**
$\alpha \cdot (\beta \cdot A) = (\alpha \cdot \beta) \cdot A$
$0 \cdot A = \emptyset$
$\alpha \cdot \emptyset = \emptyset$
$1 \cdot A = A$
$-1 \cdot A = -A$
$(\alpha \cdot A)^T = \alpha \cdot A^T$
$(\alpha \cdot A)^T = \alpha \cdot A^T$
$\alpha \cdot (-A) = -(\alpha \cdot A)$

**Addition**
$A + B = B + A$
$(A+B)+C = A+(B+C)$
$A + \emptyset = A$
$\emptyset + A = A$
$A + (-A) = \emptyset$
$-(A + B) = -A + (-B)$
$(\alpha + \beta) \cdot A = \alpha A + \beta A$
$\alpha(A + B) = \alpha A + \alpha B$
$A + A = 2A$
$(A + B)^T = A^T + B^T$
$A + B = \emptyset \rightarrow A = -B$
$-A + A = \emptyset$

**Product**
$A \cdot \emptyset = \emptyset$
$\emptyset \cdot A = \emptyset$
$I_n \cdot A = A$
$A \cdot I_n = A$
$(A \cdot B) \cdot C = A \cdot (B \cdot C)$
$A \cdot (B + C) = A \cdot B + A \cdot C$
$(A + B) \cdot C = A \cdot C + B \cdot C$
$-A \cdot B = -(A \cdot B)$
$A \cdot (-B) = -(A \cdot B)$
$A \cdot (\alpha \cdot B) = \alpha \cdot (A \cdot B)$
$(\alpha \cdot A) \cdot B = \alpha \cdot (A \cdot B)$
$(A \cdot B)^T = B^T \cdot A^T$

**Row operations**
$F_{ii}A = A$
$F_{ij}I_n \cdot A = F_{ij}A$
$F_{ij}B \cdot A = F_{ij}(B \cdot A)$
$F_i(\alpha)I_n \cdot A = F_i(\alpha)A$
$F_i(\alpha)B \cdot A = F_i(\alpha)(B \cdot A)$
$F_{ij}(\alpha)I_n \cdot A = F_{ij}(\alpha)A$
$F_{ij}(\alpha)B \cdot A = F_{ij}(\alpha)(B \cdot A)$

Where:

- $A, B, C$: Matrices of arbitrary dimensions.
- $\alpha, \beta$: Arbitrary scalars.
- $\emptyset$: Matrix of arbitrary dimensions where all the elements are 0.
- $I_n$: The nth-order identity matrix.
- $F_{ij}A$: Swaps rows $i$ and $j$ of matrix $A$.
- $F_i(\alpha)A$: Multiply $i$ row of matrix $A$ by $\alpha$.
- $F_{ij}(\alpha)A$: Adds, to $i$ row of matrix $A$, the $j$ row multiplied by $\alpha$.

### 2.2 Gauss-Jordan algorithm implementation

Using the above row transformations $F_{ij}$, $F_i(\alpha)$ and $F_{ij}(\alpha)$, we can implement the Gauss-Jordan algorithm to get the inverse of a matrix. The main idea is that we begin with the (composed) matrix $(I_n|A)$ and we apply a sequence of row

transformations, trying to transform the matrix $A$ to $I_n$. We have proved that if the final result of these operations is the composed matrix $(B|C)$ and $C = I_n$, then $B \cdot A = I_n$, and, therefore, $B = A^{-1}$. That is we have defined and formally verified a Gauss-Jordan based algorithm for computing the inverse of a matrix.

Abstract stobjs turns out to be crucial for the whole formalization. They allow us to reason about the algorithm as if it were executed on a bidimensional and intuitive representation of matrices, although actually is executed using the more efficient (but more complex) unidimensional, stobj based, representation.

### 2.3 Execution efficiency

We have tested our implementation computing the main operations (addition, multiplication and inverse) on randomly generated matrices. For example, we compute the sum of matrices of dimensions up to $n = 1000$, product for dimensions up to $n = 300$ and inverse of matrices of dimensions up to $n = 100$. We compared our execution time with the execution time obtained using the implementation in [5]. We also compared memory allocation in both implementations.

Our implementation outperforms the implementation in [5] in all cases. In the case of addition and product, our implementation is more than three times faster. In the case of the inverse, although the execution times are more similar, our implementation is still about 20% faster. In the next tables we can see measured execution times for sum (left) and product (right) of matrices (all times in seconds):

| **Dimension** | **Stobj's** | **Gamboa** |
|---|---|---|
| 100 | 0,00 | 0,01 |
| 200 | 0,00 | 0,04 |
| 300 | 0,01 | 0,20 |
| 400 | 0,01 | 0,26 |
| 500 | 0,01 | 0,35 |
| 600 | 0,02 | 0,66 |
| 700 | 0,02 | 0,83 |
| 800 | 0,03 | 1,34 |
| 900 | 0,03 | 1,93 |
| 1000 | 0,04 | 2,67 |

| **Dimension** | **Stobj's** | **Gamboa** |
|---|---|---|
| 30 | 0,00 | 0,01 |
| 60 | 0,01 | 0,03 |
| 90 | 0,03 | 0,06 |
| 120 | 0,07 | 0,16 |
| 150 | 0,10 | 0,25 |
| 180 | 0,18 | 0,43 |
| 210 | 0,28 | 0,69 |
| 240 | 0,39 | 1,05 |
| 270 | 0,56 | 1,60 |
| 300 | 0,78 | 2,01 |

## References


1. `http://www.glc.us.es/acl2/matrix/acl2-matrix.tgz`.
2. M. Kaufmann, P. Manolios, and J S. Moore. *Computer-Aided Reasoning: An Approach.* Kluwer Academic Publishers, 2000.
3. F. Palomo, I. Medina and J.A. Alonso. Certification of Matrix Multiplication Algorithms (Strassen's Algorithm in ACL2) In *TPHOLs 2001.*
4. M. Kaufmann and J S. Moore. *ACL2 Version 7.1*, 2015.
   Homepage: `http://www.cs.utexas.edu/users/moore/acl2/`
5. J. Cowles, R. Gamboa and J.V. Baalen. Using ACL2 arrays to formalize matrix algebra. In *ACL2Workshop 2003.*
6. S. Goel, W. Hunt, and M. Kaufmann. Abstract Stobjs and Their Application to ISA Modeling. In *R. Gamboa and J. Davis (Eds.): ACL2 Workshop 2013*, pp 54-69.

# Preference-Based Genetic Algorithm for Solving the Bio-Inspired NK Landscape Benchmark

Christof Ferreira Torres, Sune S. Nielsen, Grégoire Danoy, and Pascal Bouvry

FSTC, University of Luxembourg
6, rue Richard Coudenhove-Kalergi,
Kirchberg, Luxembourg
christof.ferreira.001@student.uni.lu
{sune.nielsen,gregoire.danoy,pascal.bouvry}@uni.lu

**Abstract.** In molecular biology, the subject of protein structure prediction is of continued interest, not only to chart the molecular map of living cells, but also to design proteins with new functions. In this work a Preference-Based Genetic Algorithm (PBGA) is proposed aiming to optimise NK Landscape based benchmarks designed and shown to mimic the properties of the protein Inverse Folding Problem (IFP). The PBGA algorithm incorporates a weighted sum model in order to combine fitness and diversity into a single objective function scoring a set of individuals as a whole. By adjusting the sum weights, a direct control of the fitness vs. diversity trade-off in the algorithm population is achieved by means of a selection scheme iteratively removing the least contributing individuals. Experimental results demonstrate the better performance of the PBGA algorithm compared to other state-of-the-art algorithms both in terms of fitness and diversity.

## 1 Introduction

Protein engineering in general aims at designing molecules with desired properties. A method allowing to successfully design such molecules would find applications in a number of areas such as designing improved enzymes for biotechnology applications or new antibodies towards already known targets. However evaluating and therefore optimising real biological instances is very computationally demanding. A novel approach recently proposed by Nielsen et al. consists in an NK Landscape benchmark suite that mimics the properties of the Inverse Folding Problem (IFP) [6]. The IFP aims, given a protein sequence of $N$ amino acids, at finding other sequences that will result in the same 3D structure. The corresponding optimisation problem is highly *multi-modal* and the genetic algorithm proposed in this work addresses this aspect by adding a novel diversity controlling mechanism. The preference-based approach employs a Weighted Sum Model (WSM) in order to control the desired bias between fitness and diversity. The resulting WSM score allows to iteratively determine and remove the individual in the combined parent and offspring population, with the lowest overall fitness contribution with respect to the defined preferences. The remainder of

this article is organised as follows. First the current state-of-the-art is situated in the related literature in Section 2, then a detailed description of the problem and of the biological background is introduced in Section 3. In Section 4 the contribution of this work in terms of achieving an adjustable level of fitness and diversity as a Preference-Based Genetic Algorithm (PBGA) is presented. Section 5 describes the experiments conducted and provides the analysis of the results obtained for the NK benchmark suite. Finally the conclusion and perspectives are summarised in Section 6.

## 2 State-of-the-art

In meta-heuristics, the subject of exploration vs. exploitation characteristics has been thoroughly studied. In this aim, a number of works have sought to maintain and control diversity in population-based meta-heuristics, e.g. crowding methods by DeJong [2], fitness sharing by Goldberg and Richardson [3], cellular algorithms by Alba and Dorronsoro [1], diversity preserving selection strategies based on hamming distance Shimodaira [7] and on altruism by Laredo et al. [4].

Preference-based algorithms have been discussed in the literature [5, 8] and refer to algorithms where the user preference is incorporated in the choice of regions in the solution or objective space. Preference can be incorporated in a number of ways, e.g., by modifying the fitness evaluation or selection schemes. The Indicator Based Evolutionary Algorithm (IBEA) [9] is an example where an indicator that characterises the population as a whole is used to guide the algorithm by eliminating the least desired individuals of the parent and offspring population union. The proposed PBGA in this paper uses the same principle of iterative elimination, determining the overall most preferable subset directly rather than achieving it as an indirect effect of designed mechanisms.

## 3 Bio-Inspired NK Landscape Benchmark Problem

In the NK benchmark problem as well as in the Inverse Folding Problem (IFP), a single solution is represented as a sequence $A = \{aa_i\}$ and consists of $N$ residue positions, where $1 \leq i \leq N$ and $aa_i \in \{1, ..., 20\}$ corresponds to the set of 20 possible amino acids. The overall size and the number of local "hills and valleys" of the NK landscape model can be adjusted with two parameters, $N$ and $K$. This paper focuses on optimising two novel NK benchmark model instances[1] proposed by Nielsen et al. [6]. These consist in the combination of two NK models, $F^A(x)$ and $F^B(x)$, by a simple multiplication with different $K$ and different neighbourhood definitions as defined in the Table 1.

[1] The NK Landscape Protein IFP Benchmark Suite - `http://nk-ifp-bench.gforge.uni.lu/index.html`

Table 1: NK Landscape Protein IFP Benchmarks

| **Model** | **Setting** |
|---|---|
| *NK-IFP-1* | $F^A(x)$: a $K = 4$ semi-adjacent circular neighbourhood is designed as follows: $\{x_{i2}, x_{i1}, x_{i+1}, x_{i+2}\}$, omitting the central position $x_i$.<br>$F^B(x)$: a $K = 3$ neighbourhood of uniform random distribution. |
| *NK-IFP-2* | $F^A(x)$: a $K = 4$ semi-adjacent circular neighbourhood as *NK-IFP-1*.<br>$F^B(x)$: a $K = 5$ neighbourhood of uniform random + 20 positions wide triangular distribution. |

## 4 A Novel Preference-Based Approach

The main idea of the preference-based approach is to use a Weighted Sum Model (WSM) in order to constantly maintain a current population best fulfilling the defined preferences. In an iterative manner, the weakest individuals from the combination of parent and offspring populations are determined and removed until the desired population size is achieved.

**Algorithm 1** Preference-Based Genetic Algorithm

```
1: Initialise(P_0)
2: t ← 0
3: while t < t_max do
4:     Q_t ← makeNewOffspringPop(P_t)
5:     R_t ← P_t + Q_t
6:     while |R_t| > |P_t| do
7:         I ← getWeakestIndividual(R_t)
8:         R_t ← R_t − I
9:     end while
10:    P_t ← R_t
11:    t ← t + 1
12: end while
```

The procedure *getWeakestIndividual* of determining the weakest individual in Algorithm 1 is defined as follows:

1. Systematically remove one individual
2. Compute the weighted sum score according to Equation 1
3. Add the individual back to the population
4. Repeat from step 1. until all individuals have been tried once and the worst individual can be determined.

The weighted sum score of a given population $P$ is calculated as follows:

$$WSM_{score}(P) = -W_{fit} \cdot F_{fit}(P) + W_{div} \cdot F_{div}(P) \quad (1)$$

Note the negation of $W_{fit}$ in Equation 1 as we want to maximise diversity but also minimise fitness at the same time.

The population fitness $F_{fit}$ is computed by simply taking the average of the fitness of all $M$ individuals of the current population $P$:

$$F_{fit}(P) = \frac{1}{M} \sum_{i=1}^{M} F(x) \tag{2}$$

An effective and simple measure of distance between two sequences is the Hamming-distance. For two sequences $A = \{aa_i\}$ and $A' = \{aa'_i\}$ where $1 \leq i \leq N$, the normalised Hamming distance between them is defined as:

$$d_{Hamm}(A, A') = \frac{1}{N} \sum_{i=1}^{N} d_i \quad where \quad d_i = \begin{cases} 0 & \text{if } aa_i = aa'_i \\ 1 & \text{if otherwise} \end{cases} \tag{3}$$

The population diversity $F_{div}$ is computed by taking the average Hamming distance of each $M$ individuals to the remaining $M-1$ individuals of the population $P$:

$$F_{div}(P) = \frac{1}{M \cdot (M-1)} \sum_{i=1}^{M} \sum_{j=1}^{M} d_{Hamm}(A_i, A_j), \quad \forall i \neq j \tag{4}$$

## 5 Experimental Results

To study the performance of the PBGA with respect to fitness and diversity convergence, a number of experiments have been conducted to compare it against different Genetic Algorithms, i.e., the generational (gGA), the synchronous cellular (scGA) and the steady-state (ssGA). The PBGA was tested with the following six different weight ratio settings:

$$W_{(fit,div)} = \{(1.0, 0.0), (0.9, 0.1), (0.8, 0.2), (0.7, 0.3), (0.5, 0.5), (0.3, 0.7)\}.$$

Table 2 summarises the settings and parameters used to conduct the experiments.

Figure 1a illustrates the convergence of fitness for the best performing PBGA setting in comparison with the gGA, scGA and the ssGA. The gGA performs the worst and the PBGA with a weight setting of (0.9, 0.1) surpasses the ssGA and achieves better final fitness results than all of the other GAs. Figure 1b illustrates the diversity convergence for the same algorithms. It is noted that the PBGA achieves a higher diversity than the scGA and ssGA while at the same time having better fitness results. Similar graphs are obtained for the NK-IFP-2 model and are hence not shown here.

Table 3 summarises average fitness and diversity for all the algorithms tested highlighting best and worst algorithm results in light and dark grey respectively. With a weight setting of (0.9, 0.1) the PBGA achieves the best fitness for both

Table 2: Experimental settings.

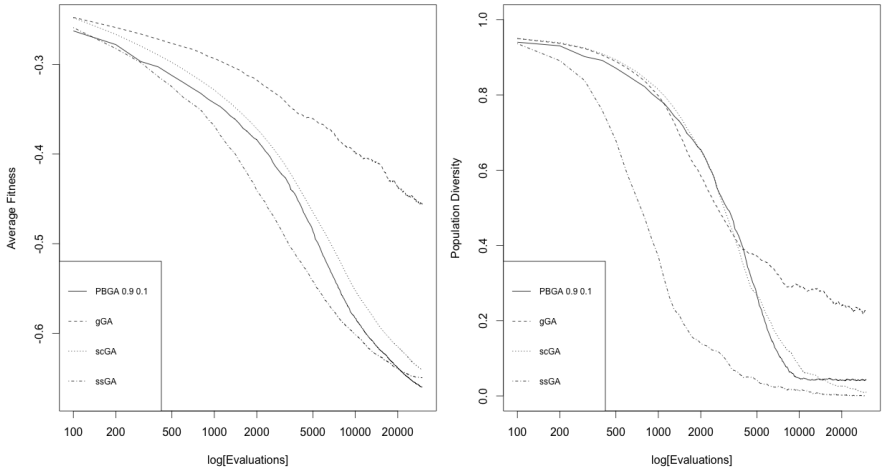
| **Setting** | **Value** |
|---|---|
| GAs | gGA, scGA, ssGA and PBGA |
| Population size | 100 |
| Termination condition | 30000 function evaluations |
| Number of independent runs | 30 |
| Selection | Binary tournament (BT) |
| Neighbourhood | C9 in scGA |
| Crossover operator | SPX, $p_c = 0.9$ |
| Mutation operator | Uniform, $p_m = \frac{1}{N}$ |
| Elitism | 2 individuals (for gGA) |

benchmark models with -0.662 for the best value and -0.660 on average for model 1 and with -0.632 for the best value and -0.631 on average for model 2. It is interesting to note that the PBGA with a weight setting of (0.5, 0.5) achieves better results than the gGA in terms of fitness as well as diversity for both models with -0.574 vs. -0.559 for the best fitness value and -0.511 vs. -0.456 on average for model 1 and with -0.550 vs. -0.545 for the best fitness value and -0.485 vs. -0.429 on average for model 2.

In order to provide statistical confidence, the Wilcoxon test indicator was applied with a 5% significance level. With a weight setting of (0.9, 0.1), the PBGA clearly outperforms the gGA and the scGA with statistical confidence for the average fitness with values -0.662 vs. -0.559 and -0.662 vs. -0.644 respectively for model 1 and with values -0.632 vs. -0.545 and -0.632 vs. -0.621 respectively for model 2, whereas in comparison with the ssGA the PBGA does not achieve as quick good results as the ssGA, but surpasses the ssGA in the end and achieves better average fitness values of -0.662 vs. -0.650 respectively for model 1 and with values -0.632 vs. -0.628 respectively for model 2. However, as seen in Figure 1a for model 1, the final slope is steeper than the ssGA, indicating better performance. The steeper final slope can be explained by the constantly high diversity as seen in Figure 1b for model 1, which allows for continued exploration while the other GAs suffer from premature convergence.

Table 3: Final values in terms of fitness and diversity averaged over 30 independent runs for the two NK benchmark models.

| | Model 1 | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | Fitness | | Diversity | | Fitness | | Diversity | |
| **Algorithm** | **Best** | **Average** | **Best** | **Average** | **Best** | **Average** | **Best** | **Average** |
| $PBGA_{1.0\ 0.0}$ | -0.649 | -0.648 $\pm 0.37E{-3}$ | 0.005 | 0.002 $\pm 1.78E{-3}$ | -0.628 | -0.628 $\pm 0.27E{-3}$ | 0.004 | 0.001 $\pm 1.56E{-3}$ |
| $PBGA_{0.9\ 0.1}$ | -0.662 | -0.660 $\pm 1.07E{-3}$ | 0.041 | 0.043 $\pm 0.84E{-3}$ | -0.632 | -0.631 $\pm 0.76E{-3}$ | 0.031 | 0.038 $\pm 3.25E{-3}$ |
| $PBGA_{0.8\ 0.2}$ | -0.652 | -0.627 $\pm 12.3E{-3}$ | 0.250 | 0.337 $\pm 43.8E{-3}$ | -0.621 | -0.594 $\pm 13.4E{-3}$ | 0.310 | 0.406 $\pm 48.1E{-3}$ |
| $PBGA_{0.7\ 0.3}$ | -0.629 | -0.582 $\pm 23.3E{-3}$ | 0.508 | 0.612 $\pm 52.1E{-3}$ | -0.602 | -0.557 $\pm 22.5E{-3}$ | 0.542 | 0.639 $\pm 48.7E{-3}$ |
| $PBGA_{0.5\ 0.5}$ | -0.574 | -0.511 $\pm 31.4E{-3}$ | 0.774 | 0.833 $\pm 31.4E{-3}$ | -0.550 | -0.485 $\pm 32.3E{-3}$ | 0.787 | 0.846 $\pm 29.6E{-3}$ |
| $PBGA_{0.3\ 0.7}$ | -0.527 | -0.458 $\pm 34.5E{-3}$ | 0.880 | 0.909 $\pm 14.6E{-3}$ | -0.503 | -0.440 $\pm 31.7E{-3}$ | 0.888 | 0.913 $\pm 12.4E{-3}$ |
| gGA | -0.559 | -0.456 $\pm 51.4E{-3}$ | 0.145 | 0.227 $\pm 40.9E{-3}$ | -0.545 | -0.429 $\pm 58.0E{-3}$ | 0.138 | 0.221 $\pm 41.1E{-3}$ |
| scGA | -0.644 | -0.641 $\pm 1.54E{-3}$ | 0.017 | 0.010 $\pm 3.36E{-3}$ | -0.621 | -0.619 $\pm 1.20E{-3}$ | 0.013 | 0.009 $\pm 2.18E{-3}$ |
| ssGA | -0.650 | -0.645 $\pm 0.18E{-3}$ | 0.005 | 0.001 $\pm 1.97E{-3}$ | -0.628 | -0.628 $\pm 0.14E{-3}$ | 0.001 | 0.001 $\pm 0.42E{-3}$ |




(a) Fitness convergence PBGA vs. GAs (b) Diversity convergence PBGA vs. GAs

Fig. 1: NK benchmark model NK-IFP-1 average fitness and diversity convergence.

## 6 Conclusion

In this paper a novel Preference-Based Genetic Algorithm (PBGA) was presented in combination with a weighted sum model, which allows to shift focus arbitrarily between diversity and fitness with a direct effect on the population as a whole without relying on secondary effects from added mechanisms or operators. The PBGA was evaluated on NK benchmark models and compared to state-of-the-art GAs. Final results were found comparable or better than the other GAs on average, while the diversity of found sequences remains higher at the same time. The best results were achieved using a weight setting of (0.9, 0.1) where 0.9 represents 90% of fitness and 0.1 represents 10% of diversity. In addition, the PBGA showed a better convergence, which promises even better solutions, given an evaluation budget beyond the computational limitations set in this work. Future work will focus on the development of a more advanced preference evaluation model using Fuzzy logic while adding more preferences such as crowding or elitism, and making the selection of preferences adaptive.

## References

1. E. Alba and B. Dorronsoro. The exploration/exploitation tradeoff in dynamic cellular genetic algorithms. *Evolutionary Computation, IEEE Transactions on*, 9(2):126–142, 2005.
2. K. A. De Jong. Analysis of the behavior of a class of genetic adaptive systems. 1975.

3. D. E. Goldberg and J. Richardson. Genetic algorithms with sharing for multimodal function optimization. In *Genetic algorithms and their applications: Proceedings of the Second International Conference on Genetic Algorithms*, pages 41–49. Hillsdale, NJ: Lawrence Erlbaum, 1987.
4. J. L. Jimenez Laredo, S. S. Nielsen, G. Danoy, P. Bouvry, et al. Cooperative selection: Improving tournament selection via altruism. In *The 14th European Conference on Evolutionary Computation in Combinatorial Optimisation*, 2014.
5. A. López-Jaimes and C. A. C. Coello. Including preferences into a multiobjective evolutionary algorithm to deal with many-objective engineering optimization problems. *Information Sciences*, 277:1–20, 2014.
6. S. S. Nielsen, G. Danoy, P. Bouvry, and E.-G. Talbi. Nk landscape instances mimicking the protein inverse folding problem towards future benchmarks. In *Proceedings of the Companion Publication of the 2015 on Genetic and Evolutionary Computation Conference*, GECCO Companion '15, pages 915–921, New York, NY, USA, 2015. ACM.
7. H. Shimodaira. Dcga: A diversity control oriented genetic algorithm. In *Tools with Artificial Intelligence, 1997. Proceedings., Ninth IEEE International Conference on*, pages 367–374. IEEE, 1997.
8. L. Thiele, K. Miettinen, P. J. Korhonen, and J. Molina. A preference-based evolutionary algorithm for multi-objective optimization. *Evolutionary Computation*, 17(3):411–436, 2009.
9. E. Zitzler and S. Künzli. Indicator-based selection in multiobjective search. In *Parallel Problem Solving from Nature-PPSN VIII*, pages 832–842. Springer, 2004.

# Stable models in normal residuated logic programs

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

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

**Abstract.** The existence of stable models for a normal residuated logic program defined on $[0, 1]$ and the uniqueness of these models in the particular case of the product t-norm, its residuated implication, and the standard negation have been recently studied by Madrid and Ojeda-Aciego [10]. In this paper, we introduce results which generalize the existence of stable models for normal residuated logic programs defined on any convex compact set of an euclidean space. In addition, we show which conditions are required in order to guarantee the uniqueness of a stable model for a normal residuated logic program defined on $\mathcal{C}([0, 1])$.

**Key words:** negation; normal residuated logic program; stable model.

## 1 Introduction

Searching for conditions guarantying the existence and uniqueness of fuzzy stable models in normal residuated logic programming has received a strong attention since the definition of this kind of programs [1].

However, the existence of stable models cannot be guaranteed for an arbitrary normal residuated logic program [2]. This is due to the fact that fuzzy framework includes two different dimmensions: the syntactic structure of the normal program (the syntaxis) and the choice of suitables connectives in the underlying lattice, the semantics of the program.

As the connectives are fixed in classical logic programming, we can only establish the syntactic conditions of the program. Nevertheless, we can choose many operators to use them as connectives in normal residuated logic programs, and this implies that semantics plays a crucial role in this framework.

Until now, only a few sufficient conditions have been found to ensure the existence of fuzzy stable models in some approaches. In [3], it has been proven that every normal logic program has stable models in the 3-valued Kleene logic. Furthermore, by [4–8], we know that every normal residuated logic program has stable models if the underlying residuated lattice has an appropriate bilattice structure [9]. Recently, it has been shown in [10] conditions to ensure the existence and unicity of stable models for a normal residuated logic program defined on $[0, 1]$.

In this paper, we will generalize the result of existence of stable models for programs defined on any convex compact set of an euclidean space. Moreover, we will introduce the conditions which guarantee the uniqueness of stable models for normal residuated logic programs defined on $\mathcal{C}([0,1])$.

## 2 Preliminaries

In this section, we will recall the main definitions and results which will be used in the paper. Firstly, we introduce the definition of residuated lattice.

**Definition 1.** *A* residuated lattice *is a tuple* $(L, \leq, *, \leftarrow)$ *such that:*

*(1)* $(L, \leq)$ *is a complete bounded lattice with* $\top$ *and* $\bot$ *the greatest and the least elements, respectively;*
*(2)* $(\leftarrow, *)$ *is an adjoint pair in* $(L, \leq)$*, that is, the equivalence:*

$$z \leq (x \leftarrow y) \quad \textit{if and only if} \quad y * z \leq x$$

*holds, for all* $x, y, z \in L$*.*
*(3)* $(L, *, \top)$ *is a commutative monoid.*

Note that, the adjoint pair is uniquely determined by the chosen operator $*$. Specifically, fixed a left-continuous operator $*$, its adjoint implication is defined as follows:

$$x \leftarrow y = \sup\{z \in L : y * z \leq x\}$$

Now, we will consider a residuated lattice enriched with a negation operator. A negation operator is any decreasing mapping $n : L \to L$ satisfying $n(\bot) = \top$ and $n(\top) = \bot$. The negation will model the notion of default negation.

**Definition 2.** *Given a residuated lattice with negation* $(L, \leq, *, \leftarrow, \neg)$*, a* normal residuated logic program $\mathbb{P}$ *is a finite set of weighted rules of the form:*

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

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

As usual, we denote the rules as $\langle p \leftarrow \mathcal{B}; \vartheta \rangle$, where $p$ is the *head* of the rule, $\mathcal{B}$ its *body* and $\vartheta$ its *weight*. A *fact* is a rule where no propositional symbols appear in the body.

The set of propositional symbols appearing in $P$ is denoted by $\Pi_{\mathbb{P}}$.

**Definition 3.** *A fuzzy* $L$-interpretation *is a mapping* $I : \Pi_{\mathbb{P}} \to L$ *which assigns a truth value to every propositional symbol appearing in* $P$*. We say that:*

*(1)* $I$ *satisfies a rule* $\langle p \leftarrow \mathcal{B}; \vartheta \rangle$ *if and only if* $\vartheta \leq I(p \leftarrow \mathcal{B})$*.*
*(2)* $I$ *is a model of* $\mathbb{P}$ *if it satisfies all rules in* $\mathbb{P}$*.*

The set of all $L$-interpretations will be denoted as $I_{\mathfrak{U}}$, where $\mathfrak{U}$ is the residuated algebra in which the lattice is defined. An ordering relation $\sqsubseteq$ can be defined in $I_{\mathfrak{U}}$ as follows: Given $I$ and $J$ two $L$-interpretations, $I \sqsubseteq J$ if and only if $I(p) \leq J(p)$, for all $p \in \Pi_{\mathbb{P}}$.

Given a finite normal residuated logic program $\mathbb{P}$ defined on $L$, the set of $L$-interpretations with the new ordering relation verifies some properties of the underlying lattice. Specifically, it inherits the properties of the cartesian product of several copies of the lattice. Indeed, each $L$-interpretation can be seen as an element of $L^n$, where $n$ is the cardinal of $\Pi_{\mathbb{P}}$.

**Theorem 1.** *If $\langle L, \leq \rangle$ is a complete lattice, then $\langle I_{\mathfrak{U}}, \sqsubseteq \rangle$ is a complete lattice.*

### 2.1 Immediate consequence operator and stable models

A generalization of the immediate consequence operator for normal residuated logic programs is given in the next definition.

**Definition 4.** *Let $\mathbb{P}$ be a normal residuated logic program. The* immediate consequence operator *is the mapping $T_{\mathbb{P}} : I_{\mathfrak{U}} \to I_{\mathfrak{U}}$ defined as*

$$T_{\mathbb{P}}(I)(p) = \sup\{I(\mathcal{B}) * \vartheta : \langle p \leftarrow \mathcal{B}; \vartheta \rangle \in \mathbb{P}\}$$

*where $p \in \Pi_{\mathbb{P}}$.*

If $\mathbb{P}$ is a positive program (without any negation), then $T_{\mathbb{P}}$ is a monotonic operator and we can characterize the models of the residuated program by the post-fix points of $T_{\mathbb{P}}$.

**Proposition 1.** *Let $\mathbb{P}$ be a positive residuated logic program. Then $M$ is a model of $\mathbb{P}$ if and only if $T_{\mathbb{P}}(M) \leq M$.*

Knaster-Tarski's fix point theorem ensures that the operator $T_{\mathbb{P}}$ has a least fix point. In addition, by the proposition above, this least fix point is actually the least model of $\mathbb{P}$. This fact leads us to define the least model semantics in positive residuated logic programs.

The main difference with respect to the case of normal residuated logic programs is that $T_{\mathbb{P}}$ is not necessarily monotonic. Therefore, we cannot guarantee the existence of the least model and we need another notion to define the semantics for a normal residuated logic program. A new mathematical object which generalizes the least model semantics to normal residuated logic programs is required. This object is the stable model of a program, which was defined in [11].

Let $\mathbb{P}$ be a normal residuated logic program and $I$ a fuzzy $L$-interpretation. First of all, we will build a positive residuated program $\mathbb{P}_I$ by substituting each rule in $\mathbb{P}$ such as

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

by the rule

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

Observe that, we can apply to $\mathbb{P}_I$ the known results to positive residuated program.

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

Thanks to the notion of reduct we can define a stable model of a program.

**Definition 6.** *Let $\mathbb{P}$ be a normal residuated logic program and let $I$ be a fuzzy $L$-interpretation. $I$ is said to be a stable model of $\mathbb{P}$ if and only if $I$ is a minimal model of $\mathbb{P}_I$.*

An important feature of stable models, which is also verified in our framework, is that a stable model is always a minimal fix point of $T_{\mathbb{P}}$.

**Proposition 2.** *Any stable model of $\mathbb{P}$ is a minimal fix point of $T_{\mathbb{P}}$.*

It is worth noting that the counterpart of Proposition 2 is not satisfied, in general, because the $T_{\mathbb{P}}$ operator is not necessarily monotonic.

## 3 On the existence and unicity of stable models

Our goal is to extend the obtained results by Madrid and Ojeda-Aciego about the existence and the unicity of stable models for normal residuated logic programs [10]. With this purpose, we need to consider an extension of Brouwer's fix point theorem.

**Theorem 2.** *Let $X$ an euclidean space and let $K$ be a a convex compact set not empty. Every continuous mapping $f\colon K \to K$ has a fix point.*

Note that, the set of all $L$-interpretations of a normal residuated logic program defined on a lattice with convex (compact, respectively) support is a convex (compact, respectively) set. This fact leads us to present the following result.

**Proposition 3.** *Let $\mathbb{P}$ be a normal residuated logic program defined on a lattice $(K, \leq, *, \leftarrow, \neg)$ where $K$ is a convex (compact, respectively) set in an euclidean space $X$. Then the set of $L$-interpretations of $\mathbb{P}$ is a convex (compact, respectively) set in the set if mappings defined on $X$.*

Applying Theorem 2 to the operator $R$ defined by $R(I) = T_{\mathbb{P}_I}$, we obtain that $T_{\mathbb{P}_I}$ have a fix point. As $\mathbb{P}_I$ is a positive residuated logic program, we obtain that this fix point is actually the minimal model of $\mathbb{P}_I$ and then it is a stable model of $\mathbb{P}$.

The continuity of the connectives $*$ and $\neg$ plays a key role in order to apply the Theorem 2 to the operator $R$.

**Theorem 3.** *Let $(K, \leq, *, \leftarrow, \neg)$ be a residuated lattice where $K$ is a convex compact non-empty set in an euclidean space. If $*$ and $\neg$ are continuous operators, then every finite normal program $\mathbb{P}$ defined on this lattice has at least a stable model.*

Finally, we present a result which ensure the uniqueness of the stable models for a normal program defined with the product adjoint pair and the standard negation on the set of subintervals of $[0,1]$, that is, $\mathcal{C}([0,1])$. This fact is interesting because not only one truth value can be assigned to each propositional symbol in $\mathbb{P}$, but we can assign a minimal truth value and a maximal truth value for the propositional symbol.

**Theorem 4.** *Let $\mathbb{P}$ be a normal residuated logic program defined on $\mathcal{C}([0,1])$, and let us write for each propositional symbol $p$ in $\mathbb{P}$, $\vartheta_p = \max\{\vartheta_j : \langle p \leftarrow \beta\ ; \vartheta_j\rangle \in \mathbb{P}\}$. Then, if for every rule $\langle p \leftarrow q_1 * \cdots * q_h * \neg q_{h+1} * \cdots * \neg q_k\ ; \vartheta\rangle \in \mathbb{P}$, the inequality below holds*

$$\left(\sum_{j=1}^{h} \vartheta_{q_1} \cdot \ldots \cdot \vartheta_{q_{j-1}} \cdot \vartheta_{q_{j+1}} \cdot \ldots \cdot \vartheta_{q_h} \cdot \vartheta\right) + (k-h)(\vartheta_{q_1} \cdot \ldots \cdot \vartheta_{q_h} \cdot \vartheta) < (1,1)$$

*then there exists only one stable model of $\mathbb{P}$.*

## 4 Conclusions

We have shown results which guarantee the existence of stable models for normal programs defined on a convex compact set, and which guarantee the uniqueness of stable models for normal programs defined on $\mathcal{C}([0,1])$.

## References


1. C., Damásio and L., Moniz Pereira. Monotonic and residuated logic programs. Sixth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU01), pp. 748–759. Springer Verlag (2001).
2. N., Madrid and M., Ojeda-Aciego. On coherence and consistence in fuzzy answer set semantics for residuated logic programs. Lect. Notes in Computer Science 5571, pp. 60–67, (2009).
3. T., Przymusinski. Well-founded semantics coincides with three-valued stable semantics. Fundamenta Informaticae 13, pp. 445–463, (1990).
4. M., Fitting. The family of stable models. The Journal of Logic Programming 17(2-4), pp. 197–225, (1993).
5. Y., Loyer and U., Straccia. Epistemic foundation of stable model semantics. Journal of Theory and Practice of Logic Programming 6, pp. 355–393, (2006).
6. U., Straccia. Query answering in normal logic programs under uncertainty. Lect. Notes in Computer Science 3571, pp. 687–700, (2005).
7. U., Straccia. Query answering under the any-world assumption for normal logic programs. Lect. Notes in Computer Science 3571, pp. 687–700, (2006).

8. U., Straccia. A top-down query answering procedure for normal logic programs under the any-world assumption. Proc. of the 10th Intl Conf on Principles of Knowledge Representation, pp. 329–339, (2006).
9. M. L., Ginsberg. Multivalued logics: a uniform approach to reasoning in arti cial intelligence. Computational Intelligence 4, pp. 265–316, (1988).
10. N., Madrid and M., Ojeda-Aciego. On the existence and unicity of stable models in normal residuated logic programs. International Journal of Computer, Mathematics (2012).
11. N., Madrid and M., Ojeda-Aciego. Towards a fuzzy answer set semantics for residuated logic programs. Web Intelligence/IAT Workshops, pp. 260–264, (2008).

# A Study of Multiple Sequence Alignment With Multi-Objective Metaheuristics

Antonio J. Nebro[1], Cristian Zambrano-Vega[2] Juan J. Durillo[3], and José Francisco Aldana Montes[1]

[1] Edificio de Investigación Ada Byron, University of Málaga, Spain,
[2] Facultad de Ciencias de la Ingeniería, Universidad Técnica Estatal de Quevedo, Quevedo - Los Ríos - Ecuador
[3] Distributed and Parallel Systems Group, University of Innsbruck, Austria

**Abstract.** Multiple sequence alignment (MSA) is a problem from the bioinformatics domain consisting in finding the best possible alignment for a set of three or more sequences. Different scores have been proposed to assess the quality of MSA solutions, so the problem can be formulated as a multi-objective optimization problem. In this paper we carry out a performance study involving five multi-objective metaheuristics which are representative of the state-of-the-art. The results when solving a number of instance problems reveals that the classical NSGA-II and SPEA2 algorithms can outperform more modern techniques.

**Keywords:** Multiple sequence alignment, multi-objective optimization, metaheuristics, performance comparison

## 1 Introduction

The alignment of multiple DNA, RNA and protein sequences (MSA) is a common task in Bioinformatics [1]. The aim of MSA is comparing different sequences in order to extract their shared information and their significant differences. The alignment of pair of sequences can be achieved by using dynamic programming techniques, but these strategies cannot be applied when dealing with three or more sequences because the search space grows exponentially with the number of sequences and it is also dependent on the sequence lengths. These reasons have lead to the use metaheuristics to deal with MSA problems [2].

An additional issue in MSA is that there exist different methods to measure the accuracy on an alignment, so the problem can be formulated as a multi-objective optimization problem [3][4]. The motivation of our work is that these studies rely on the use of the NSGA-II algorithm [5], so we are interested in determining whether other algorithms could be more adequate for solving MSA problems. In this paper, we elaborate our first approximation to this issue. Our main contribution is the comparison of a number of multi-objective metaheuristics: NSGA-II, SPEA2 [6], AbYSS [7], MOCell [8], and SMS-EMOA [9]. All these algorithms but NSGA-II are applied the first time to MSA to the best of our knowledge.

The rest of the paper is organized as follows. Section 2 includes a review of related work. The problem is described in Section 3. The experimentation details and an analysis of the obtained results are presented in Sections 4 and 5. Finally, the conclusions and lines of future work are commented in Section 6.

## 2 Related Work

In this section we briefly review some multi-objective approaches published in the literature to solve the MSA problem using multi-objective optimization techniques.

Ortuño *et al.* implemented a NSGA-II based multi-objective evolutionary algorithm to align multiple sequences and applied it to optimizing three objectives: STRIKE score, non-gaps percentage and totally conserved columns [3]. Soto and Becerra proposed a multi-objective evolutionary algorithm, also inspired in NSGA-II, to optimize pre-aligned sequences in [10]. They used two objectives functions to compare the quality of the MSA: the entropy and the MetAl metrics. A multi-objective genetic algorithm based in NSGA-II (MSAG-MOGA) is described in [4], where three objectives are considered: similarity, affine gap penalty and support.

The first two works take the approach of pre-computing alignments with existing tools (Muscle, ClustalW, Mafft, T-Coffee, etc.), in such a way that the initial populations contain aligned solutions. We use also this idea in this work.

It is worth nothing that these three papers consider different objectives, so there is no a consensus about how assess the quality of the alignments. This makes also makes difficult to compare new proposals against them.

## 3 Problem Description

Given a finite alphabet $\Sigma$ and a set $S = (s_1, s_2, ..., s_n)$ of $n$ sequences of varying length, an alignment is a matrix where all the symbols of the sequences appear in the same order and a special symbol or gap (typically represented with the character '-') can be inserted potentially at any position in such a way that all the sequences have the same length.

An example of alignment is shown below, representing four sequences with two aligned columns (marked with an asterisk).

```
APPSVFAEVPQ-AQPV
AKRS-V-E-PFR-IKM
-LISKRA-YP--I---
-SASTIGVEPC-RA-P
   *     *
```

The MSA problem consists then in inserting gaps in the proper places in order to maximize some scores. For example, in [3] two of the considered objectives are to maximize the non-gaps percentage and the percentage of completely aligned columns. These are very intuitive goals, but they are not fully contradictory from a multi-objective point of view: if after manipulating the sequences a column is

full of gaps then it can be removed, thus improving the number of non-gaps, but this does not implies a worsening in the percentage of aligned columns.

In this work we have select two objectives: the aforementioned percentage of aligned columns and the sum of pairs (SOP), which is computed by adding all the scores of the pairwise comparisons between each symbol in each column of the alignment. A scoring matrix is need to calculate the SOP; we have used the PAM250 matrix (with a gap penalty of -8).

## 4 Experimentation

In this section, we briefly describe the algorithms we have selected, the chosen benchmark, and experimentation methodology.

Our study includes five multi-objective metaheuristics. NSGA-II [5] and SPEA2 [6] are classical evolutionary algorithms which have been widely used since they were proposed. SMS-EMOA [9] and MOCell [8] are also evolutionary algorithms and they are representative of indicator-based and cellular techniques, respectively. The last method is AbYSS [7], a scatter search algorithm. All the metaheuristics have been implemented in the jMetal framework [11].

We have considered a set of parameter settings that have been adopted in other studies. This way, all the algorithms runs until 25000 function evaluations have computed, the population sizes have a size of 100 in the evolutionary algorithms (20 in AbYSS), and MOCell and AbYSS has an archive size of 100. All the metaheuristics include the same genetic operators: single-point crossover (applied with a probability of 1.0) and a multiple mutation operator which randomly selects one out of three mutations: one gap insertion (a gap is randomly inserted), one gap shifting (a random gap is selected and it is shifted with the symbol on the right or on the left), and gaps merging (a number of gaps are joined to appear consecutively in the sequence). This multiple mutation operator is applied with a probability of $1.0/L$, where $L$ is the number of sequences. The encoding used to represent the sequences consists of lists of characters.

As commented before, each problem has been previously aligned with a number of tools, namely Clustal Omega, T-Coffee, Mafft, and Muscle. The obtained alignments are included in the initial population of all the algorithms, and they are also used to create new solutions. The process consists on choosing an alignment and inserting a number of gaps (from 1 to 5) at random positions.

We have chosen five problems from the BAliBASE 3.0 library [12]. Concretely, we have taken five instances of the RV11 reference set, which are referred as to BB11001 to BB11005. They range from 4 to 14 sequences.

The experimentation methodology is described next. 20 independent runs have been carried of out of each combination algorithm-problem, and the Hypervolume quality indicator [13] has been computed to all the yielded Pareto front approximations. We report the obtained median and interquartile range (IQR) values. To check the significance of the differences between the algorithms we have applied the unpaired Wilcoxon rank-sum test with a confidence level

**Table 1.** Hypervolume quality indicator values. Median and IQR

| | NSGAII | SPEA2 | ABYSS | SMS-EMOA | MOCell |
|---|---|---|---|---|---|
| BB11001 | $0.00e+00_{0.0e+00}$ | $0.00e+00_{1.5e-01}$ | $1.46e-03_{5.7e-02}$ | $0.00e+00_{0.0e+00}$ | $0.00e+00_{0.0e+00}$ |
| BB11002 | $4.04e-01_{2.4e-01}$ | $4.97e-01_{1.2e-01}$ | $1.46e-01_{1.2e-01}$ | $4.99e-01_{1.2e-01}$ | $1.66e-01_{2.5e-02}$ |
| BB11003 | $2.28e-01_{1.6e-01}$ | $2.63e-01_{2.2e-01}$ | $0.00e+00_{0.0e+00}$ | $1.91e-01_{1.4e-01}$ | $0.00e+00_{0.0e+00}$ |
| BB11004 | $2.93e-01_{1.8e-01}$ | $2.80e-01_{9.3e-02}$ | $0.00e+00_{0.0e+00}$ | $2.03e-01_{1.5e-01}$ | $4.83e-02_{7.0e-02}$ |
| BB11005 | $4.44e-01_{1.3e-02}$ | $4.08e-01_{1.5e-02}$ | $0.00e+00_{0.0e+00}$ | $4.07e-01_{2.5e-02}$ | $3.90e-01_{3.5e-02}$ |

**Table 2.** Results of the Wilcoxon rank-sum test. Each symbol in the cells represents the five considered ploblems. The ▲ symbol indicates that the algorithm in the row is significantly better than the algorithm in the column, a ▽ means the opposite, and a '-' states that the differences are non-significant.

| | SPEA2 | ABYSS | SMSEMOA | MOCell |
|---|---|---|---|---|
| NSGAII | - - - - ▲ | ▽ ▲ ▲ ▲ ▲ | - - - - ▲ | - ▲ ▲ ▲ ▲ |
| SPEA2 | | - ▲ ▲ ▲ ▲ | - - - - - | ▲ ▲ ▲ ▲ ▲ |
| ABYSS | | | - ▽ ▽ ▽ ▽ | ▲ - - ▽ ▽ |
| SMSEMOA | | | | ▲ ▲ ▲ ▲ ▲ |

of 95% (i.e., significance level of 5% or p-value under 0.05), meaning that the differences are unlikely to have occurred by chance with a probability of 95%.

As the true Pareto fronts of the solved problems are unknown, we have built a reference front for each problem by joining all the fronts obtained by all the algorithms in all the independent runs and deleting the dominated solutions.

## 5 Results

The obtained results are included in Table 1, where the cells with dark and light gray background colors indicate respectively the best and second best values. We can observe than some cells contain a value of 0; this means that the obtained fronts are beyond the limits of the reference Pareto front, so the Hypervolume is not computed in those situations.

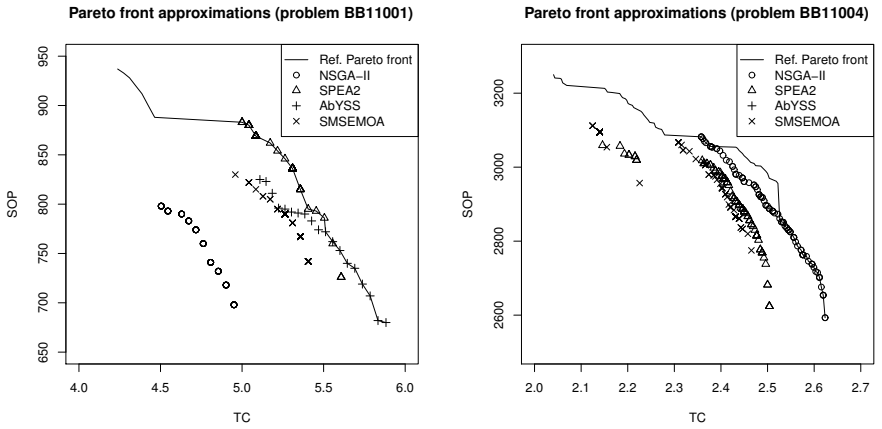
The values included in the table reveals that NSGA-II and SPEA2 have the best overall performance, outperforming the other algorithms. Both NSGA-II and SPEA2 are generational evolutionary algorithms, while SMS-EMOA and MOCell follow a steady-state scheme, so a first conclusion could be that the generational selection scheme can have a positive influence in searching for optimal alignments. However, the results of the Wilcoxon rank-sum test (see Table 2) show that most of the differences in the pairwise comparison between NSGA-II, SPEA2, and SMS-EMOA are not significant.

To illustrate the Pareto front approximations that are found by the compared algorithms (we have excluded MOCell), we include in Figure 1 the fronts having the best Hypervolume value for problems BB11001 and BB11004.

## 6 Conclusions and Future Work

We have presented a study of solving MSA problems with a number of multi-objective metaheuristics. Our main motivation has been that multi-objective

**Fig. 1.** Pareto front approximations having the best Hypervolume obtained by NSGA-II, SPEA2, SMS-EMOA, and AbYSS for problems BB11001 and BB11004.




optimization approaches for MSA are scarce in the literature, and practically all of them relies on the use of NSGA-II. So, we have selected five multi-objective metaheuristics and we have compared them again a benchmark of five problems.

The conclusion is that, according the parameter settings we have used and the chosen benchmark, the classic NSGA-II and SPEA2 algorithms outperform more recent proposals according to the Hypervolume quality indicator.

Our work is a first step in the open issue of MSA with multi-objective metaheuristics. First, we have used standard settings established in other studies (mainly on continuous optimization), so a parameter sensitivity study is needed to find more effective algorithm configurations to deal with MSA. Second, the benchmark must be augmented in a significant way to draw firmer conclusions. Finally, an analysis of the objectives to optimize must be carried out to decide which scores are really important; if more than four or five would be of interest, the MSA would become a many-objective problem, so a new set of metaheuristics should be needed.

**Acknowledgments.** This work is partially funded by Secretaría Nacional de Educación Superior Ciencia y Tecnología SENESCYT from Ecuador and by Grants TIN2014-58304-R (Ministerio de Ciencia e Innovación) and P11-TIC-7529 and P12-TIC-1519 (Plan Andaluz de Investigación, Desarrollo e Innovación).

## References

1. Pei, J.: Multiple protein sequence alignment. Current Opinion in Structural Biology **18**(3) (2008) 382 – 386 Nucleic acids / Sequences and topology.

2. da Silva, F., Pérez, J.S., Pulido, J.G., Rodríguez, M.V.: Alineagaa genetic algorithm with local search optimization formultiple sequence alignment. Applied Intelligence **32**(2) (2010) 164–172
3. Ortuno, F., Valenzuela, O., Rojas, F., Pomares, H., Florido, J., Urquiza, J., Rojas, I.: Optimizing multiple sequence alignments using a genetic algorithm based on three objectives: structural information, non-gaps percentage and totally conserved columns. Bioinformatics (Oxford, England) **29**(17) (September 2013) 2112–21
4. Kaya, M., Sarhan, A., Abdullah, R.: Multiple sequence alignment with affine gap by using multi-objective genetic algorithm. Computer methods and programs in biomedicine **114**(1) (April 2014) 38–49
5. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation **6**(2) (2002) 182–197
6. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm. In Giannakoglou, K., Tsahalis, D., Periaux, J., Papailou, P., Fogarty, T., eds.: EUROGEN 2001. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, Athens, Greece (2002) 95–100
7. Nebro, A., Luna, F., Alba, E., Dorronsoro, B., Durillo, J., Beham, A.: AbYSS: Adapting Scatter Search to Multiobjective Optimization. IEEE Transactions on Evolutionary Computation **12**(4) (August 2008)
8. Nebro, A., Durillo, J., Luna, F., Dorronsoro, B., Alba, E.: Design issues in a multiobjective cellular genetic algorithm. In Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T., eds.: Evolutionary Multi-Criterion Optimization. 4th International Conference, EMO 2007. Volume 4403 of Lecture Notes in Computer Science., Springer (2007) 126–140
9. Emmerich, M., Beume, N., Naujoks, B.: An emo algorithm using the hypervolume measure as selection criterion. In Coello, C., Hernández, A., Zitler, E., eds.: Third International Conference on Evolutionary MultiCriterion Optimization, EMO 2005. Volume 3410 of LNCS., Springer (2005) 62–76
10. Soto, W., Becerra, D.: A multi-objective evolutionary algorithm for improving multiple sequence alignments. In Campos, S., ed.: Advances in Bioinformatics and Computational Biology. Volume 8826 of Lecture Notes in Computer Science. Springer International Publishing (2014) 73–82
11. Durillo, J., Nebro, A.: jmetal: A java framework for multi-objective optimization. Advances in Engineering Software **42**(10) (2011) 760 – 771
12. Thompson, J., Koehl, P., Poch, O.: Balibase 3.0: latest developments of the multiple sequence alignment benchmark. Proteins **61** (2005) 127–136
13. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Transactions on Evolutionary Computation **3**(4) (1999) 257–271

# Clasification of RAP network using the Lipschitz's semidistance and the measure based on inducement

Sonia Pérez Plaza, Manuel Muñoz Márquez, Fernando Fernández Palacín, and Manuel Berrocoso Domínguez

Universidad de Cádiz
{sonia.perez,manuel.munoz,fernando.fernandez,manuel.berrocoso}@uca.es

**Abstract.** In this paper the location of the subduction zone between the African and Eurasian plates is studied. A method of classification of functional data is employed in order to determine the limits of this zone. The used data are provided by the RAP network and the method of classification is based on two new similarity measures defined in this work. Furthermore, in order to determine the optimal number of groups in each classification, Silhouette coefficient is employed. Finally, the results obtained are compared to the opinion's experts.

**Keywords:** FDA, classification, Lipschitz semi-distance, Silhouette coefficient

## 1 Introduction

The southern part of the Iberian peninsula is over a subduction zone that arise from the convergence of the Euroasian plate and African plate. The limits of this subduction zone are not well determined. In Figure 1 we can see different versions about the location of this zone. In this work we study the problem of location of this zone by using a method of classification.

This region is characterized by a complex seismotectonic pattern and moderate seismic activity associated with the convergence between Africa and Eurasia.

The Andalusian Positioning Network (RAP) is a permanent station network which cover Andalusian area. The stations make a geodetic frame to surveying and cartographic applications.

The seismic movements can be considered as external impulses that generate a displacement in the stations of RAP network. This displacement depends on the situation over the plates. Hence, when we classify the stations, it is important to use a measure that perceive the displacement of the coordinates.

The data employed in this work give the displacement of north and east coordinates between 2011 and 2013. In this paper, these data are considered as functional data.


**Fig. 1.** The limits beetween the Eurasian and the African plate.

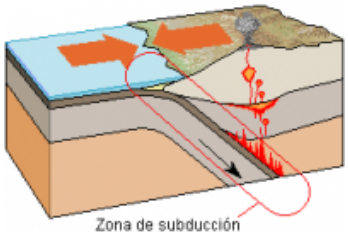



**Fig. 2.** Subduction zone.

## 2 FDA

Functional data analysis (FDA) extends the classical multivariate methods when data are functions or curves. According to [1], a functional random variable $X$ is a random variable with values in an infinite dimensional space.

The main source of difficulty when dealing with functional data consists in the fact that the observations are supposed to belong to an infinite dimensional space, whereas in practice one only has sampled curves observed into a finite set of time-points. Indeed, it is usual that we only have discrete observations $X_{ij}$ of each sample path $X_i(t)$ at a finite set of knots.

Because of this, the first step in FDA is often the reconstruction of the functional form of data from discrete observations. The most common solution to this problem is to consider that sample paths belong to a finite dimensional space spanned by some basis of functions ([2]).

An important choice to do when working with functional data is the basis of functions considered.

A basis in FDA is a set of independient functions such that any function can be approximated as a linear combination (of a sufficiently large number) of these functions. Hence, by using a basis, it is possible to aproximate the functional data (of infinite dimension) in a subspace of finite dimension.

The choice of a basis is essential and it must be made by considering how the studied functions are. That is, *Fourier basis* are used with periodic functions, *B-splines basis*, for smooth functions and *Wavelets basis*, for curves that are characterized by numerous local features like peaks or piecewise constants.

## 3 Similarity measures

Given two functions $f$ and $g$ in $L^2(\tau)$, where $\tau = [T_1, T_2] \subset \mathbb{R}$, the usual distance between these functions is given by:

$$d_{L^2}(f,g) = ||f-g||_{L^2} = \sqrt{(\int_{T_1}^{T_2} (f(t)-g(t))^2 dt)}$$

But this measure seems not to represent the intuitive idea of similitude between curves.

It seems straightforward to see that two parallel curves given by $f(t) = t$ and $g(t) = t + c$ represent two individuals with the same behaviour. Therefore, the distance between $f$ and $g$ should be 0. However, by using the previous distance,

$$d_{L^2}(f,g) = ||f-g||_{L^2} = \sqrt{|T_2 - T_1| * c^2} =$$

$$= c * \sqrt{(T_2 - T_1)} \neq 0$$

In order to avoid this problem, it is possible to use this distance but with the derivatives of the functions. However, in this case, the similitude between curves is not well measured. For this reason, it is necessary to introduced another measure when functional data are employed.

### 3.1 Lipschitz semi-distance

If we have a function defined between two euclidean spaces $(X, d_x)$ and $(Y, d_y)$, **Lipschitz measure** is defined in [3] as:

$$Lip(f) = \max\{\frac{||f(x)-f(y)||}{||x-y||} : x, y \in X\}$$

Based on this measure **Lipschitz norm** of a function $f$ is defined as follows:

$$||f||_{Lip} = Lip(f)$$

Finally, Lipschitz semi-distance is defined as the semi-distance induced by this norm, i.e.,

$$d_L(f,g) = ||f-g||_{Lip} = sup\{\frac{||(f-g)(x)-(f-g)(y))||}{||x-y||} : x, y \in X\}$$

### 3.2 Measure based on inducement

Now, we define a similarity measure based in the previously defined Lipschitz semi-distance. In Lipschitz semi-distance, the similarity between curves is given by the maximum in the difference of increasings. In this case, we consider all the relative maximums in the difference of increasings.

Hence, given two functions $f$, $g : \tau \to \mathbb{R}$, the measure based on inducement is defined as follows:

$$d_{est}(f,g) = \sum_{i=1}^{m} \{|h'(x_i)|, x_i \in \tau\}$$

where $h = f - g$ and $x_1, x_2, \ldots, x_m$ are the relative maximums of $|h'|$ in $\tau$.

## 4 Silhouette coefficient

Silhouette coefficient was proposed by Rousseeuw in 1987 ([4]) in order to provide an evaluation of clustering validity and to select the appropiate number of cluster after a partitioning technique.

Let us first take any object $i$ which is assigned to the cluster $A$ and we compute

$a(i)$ =average dissimilarity of $i$ to all other objects of $A$.

$$a(i) = \frac{1}{n_A} \sum_{r=1}^{n_A} d(i, a_r), a_r \in A \tag{1}$$

Let us now consider any cluster $C$ which is different from $A$, and compute

$d(i, C)$ =average dissimilarity of $i$ to all objects of $C$.

$$d(i, C) = \frac{1}{n_C} \sum_{r=1}^{n_C} d(i, c_r), c_r \in C \tag{2}$$

After computing $d(i, C)$ for all clusters $C \neq A$, we select the smallest of those numbers and denote it by $b(i) = min_{C \neq A} d(i, C)$

Now we take the **Silhouette coefficient** at $i$ as

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \tag{3}$$

And finally Rousseeuw define the Silhouette coefficient of the partitioning as

$$s = \frac{1}{n} \sum_{i=1}^{n} s(i) \tag{4}$$

where $n$ is the number of objects in the set.

## 5 Implementation and results

The data used in this study are the displacement in the north and east coordinates in the GPS stations between the years 2011 and 2013. These data have been modified in two phases: first, we complete the series by using a Kalman filter and, second, we eliminate the values that, according to the experts, are considered as outliers. These modified data have been fitted to Fourier basis.

After this process, we apply a hierarchical grouping procedure to the functional data. By considering the opinion of the experts, the displacements in the plane North-East have been used instead of considering separately the measure of both coordinates. The optimal number of groups has been fixed by considering Silhouette coefficient introduced in this work.

Now, we apply this procedure with Lipschitz semi-distance, with the similarity measure based on inducement, and the distance $L^2$ over the derivatives.

From this study, we obtain that, with Lispchitz semi-distance, the best classification is obtained, according Silhouette coefficient, for three groups. In this case, we see that a group is only consisting of Melilla (see Figure 4).

In the case where the similarity measure based on inducements is used, the best classification is obtained for two groups (see Figure 3).

Finally, when the usual distance over the derivatives is employed, according to Silhoutte cofficient, the best classification is also obtained for three groups. Therefore, this classification coincides with the classification obtained for Lipschitz semi-distance.

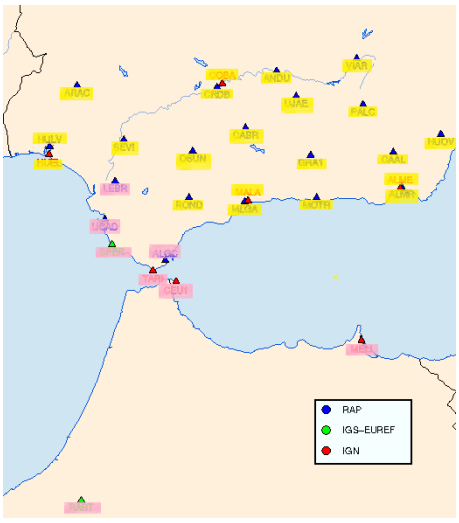



**Fig. 3.** Classification with $d_{est}$.

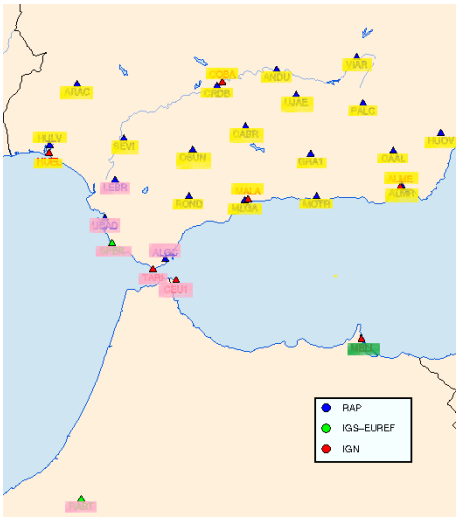



**Fig. 4.** Classification with $d_L$.

## 6 Conclusion

In this work we have studied the location of the subduction zone between the African and Eurasian plates. By using the data of RAP network, we have classified the stations in order to determine the limits of this zone.

In this classification, we have used functional data and for this type of data, we have introduced two new similarity measures. After the process of hierarchical grouping, we conclude that the best classification is the one obtained with the similarity measure based on inducements where two groups are distinguished.

According to the location of the subduction zone estimated by the experts, both groups are placed in both sides of this location and, thus, the obtained results are coherent.

## References

1. Ferraty, F., Vieu, P.: Nonparametric functional data analysis. Springer Series in Statistics. Springer, New York (2006)
2. Ramsay, J. O., Silverman, B. W: Functional data analysis. Springer Series in Statistics. Springer, New York, second edition (2005)
3. Weaver, N.: Lipschitz Algebras. World Scientific Publishing Co. Pte. Ltd. (1999)
4. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. 20, 53–65 (1987)

# Classical Symmetries and conservation laws of the semilinear damped beam equation

J.C. Camacho, M.S. Bruzón, M.L. Gandarias

Departamento de Matemáticas Universidad de Cádiz, Spain
Email: josecarlos.camacho@uca.es,m.bruzon@uca.es,gandarias@uca.es

**Abstract.** In this paper we study a nonlinear semilinear damped beam equation. We apply the Lie's group theory to make an analysis of the symmetry reductions of the equation. By the multipliers method we look for nontrivial conservation laws via integral formulae.

**Key words:** Classical Symmetries, Partial Differential Equation, Conservation laws

## 1 Introduction

In this paper we consider the semilinear damped beam equation:

$$\Delta \equiv u_{tt} + u_t + u_{xxxx} - \alpha\, u_{xx} = (F(u_x))_x, \tag{1}$$

where $t > 0$, $x \in \mathbb{R}$, $u = u(x,t) : (0,T) \times \mathbb{R} \to \mathbb{R}$ is an unknown function, $F$ is an nonlinear arbitrary function, differentiable, that depends of $u_x$, and $\alpha$ a positive constant. The equation depends on space, $x$, and time, $t$, $u = u(x,t)$ is the deflection of the roadbed and the nonlinear function $F$ models the force beam support.

Equation (1) can be written as a system of differential equations by using a new auxiliary variable $v$:

$$v = u_x, \qquad u_{tt} + u_t + v_{xxx} - \alpha\, v_x = f(v)\, v_x, \tag{2}$$

where $f(v) = F'(v)$.

In the past decade the theory of group of transformations has been used to arrive to new solutions of partial differential equations (PDEs). We use the classical method for finding symmetry reductions of PDEs, also called Lie group method of infinitesimal transformations, in short when PDEs or ordinary differential equations (ODEs) are invariant under a Lie group of transformations, a reduction transformation exists. In order to obtain all the solutions which are inequivalent with respect to the group it is sufficient to derive the solutions from the optimal system of subalgebras.

Conservation laws appear in many of physical, chemical and mechanical processes, such laws enable us solve problems in which certain physical properties

do not change in the course of time within an isolated physical system. The importance of conservation laws also embraces mathematics, for instance, the integrability of a PDE is strongly related with the existence of conservation laws. Furthermore, they can be used to obtain exact solutions of a PDE.

In this paper, we study the classical Lie symmetries of system (2) and we reduce it to systems of ODEs ([15],[1],[6],[7]). The structure is the following: First we find the point transformation group which leaves the system invariant. Next, from the optimal system of subalgebras we find the similarity variables and similarity solutions that reduce the equations to ODEs. In Section 4 we derive by using the multipliers method some nontrivial conservation laws.

## 2 Symmetry Reductions

Broadly speaking, to apply the classical method for finding symmetry reductions we seek fields of the form

$$v = p(x,t,u,v)\frac{\partial}{\partial x} + q(x,t,u,v)\frac{\partial}{\partial t} + r(x,t,u,v)\frac{\partial}{\partial u} + s(x,t,u,v)\frac{\partial}{\partial v} \qquad (3)$$

that leave the set of solutions of the system (2) invariant.
According to the Invariance Criterion ([15]) we obtain a relationship between the extended variables ($u$,$v$,$u_x$,$u_t$,$v_x$,...,$u_{xxxx}$). Taking into account that these variables are essentially independent, the coefficients in the equation must be equal to zero. This leads us to a system of differential equations in the infinitesimals $p$,$q$,$r$ and $s$ so called determining equations.

**Symmetries of the system (2)**

For system (2) we have that from the determining equations, we only find, finite dimensional algebras.
We obtain the following generators when $F(u)$ is a nonlinear function:

$$w_1 = \frac{\partial}{\partial x},\ w_2 = \frac{\partial}{\partial t},\ w_3 = \frac{\partial}{\partial u},\ w_4 = e^{-t}\frac{\partial}{\partial u}.$$

The optimal system is

$$\{\lambda w_1 + \mu w_2, \lambda w_1 + w_2 + \mu w_3, \lambda w_1 + w_2 + \mu w_4, \lambda, \mu \in \mathbb{R}\}.$$

**1.** For $\lambda w_1 + \mu w_2$, the symmetry transformation is given by

$$\begin{array}{l} z = \mu x - \lambda t, \\ u(x,t) = h(z) \\ v(x,t) = g(z). \end{array} \qquad (4)$$

Substituting (4) into equation (1) we obtain the ODE

$$\mu^4 \frac{d^4 h}{dz^4} + \left(\lambda^2 - \alpha\,\mu^2 - F'\left(\frac{d\,h}{dz}\right)\right)\frac{d^2 h}{dz^2} - \lambda\,\frac{d\,h}{dz} = 0. \qquad (5)$$

**2.** For $\lambda w_1 + w_2 + \mu w_3$, the symmetry transformation given by

$$\begin{aligned} z &= x - \lambda t, \\ u(x,t) &= \mu t + h(z) \\ v(x,t) &= g(z). \end{aligned} \tag{6}$$

leads to the ODE

$$\frac{d^4 h}{dz^4} + \left(\lambda^2 - \alpha - F'\left(\frac{d\,h}{dz}\right)\right)\frac{d^2 h}{dz^2} - \lambda\,\frac{d\,h}{dz} + \mu = 0. \tag{7}$$

**3.** For $\lambda w_1 + w_2 + \mu w_4$, the symmetry transformation given by

$$\begin{aligned} z &= x - \lambda t, \\ u(x,t) &= h(z) - \mu e^{-t} \\ v(x,t) &= g(z). \end{aligned} \tag{8}$$

leads to the ODE

$$\frac{d^4 h}{dz^4} + \left(\lambda^2 - \alpha - F'\left(\frac{d\,h}{dz}\right)\right)\frac{d^2 h}{dz^2} - \lambda\,\frac{d\,h}{dz} = 0. \tag{9}$$

## 3 Solutions

Some exact solutions can be obtained from equation (5) when $\mu = 1$: as the derivative of trigonometric, hyperbolic and exponential functions can be expressed in terms of themselves, we can choose $F$ as an algebraic function.

For example $h(z) = \tanh(z)$ is solutions of (9) for

$$F(h) = \frac{1}{2}\,(\lambda + 2)\,\log h - \frac{h\,\left(4\,\alpha - 4\,\lambda^2\right) + 6\,h^2}{4} + C.$$

So, in this case we obtain the solution of (1), $u(x,t) = \tanh(x - \lambda\,t)$

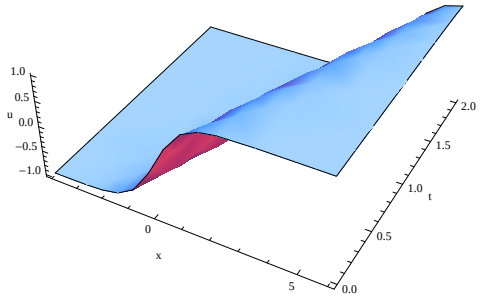



**Fig. 1.** Solutions $\mathbf{u} = \tanh(x - 2\,t)$

If we take $h(z) = \operatorname{sech}^2(z)$ is solutions of (5) when $\mu = 1$ and for

$$F'(h) = 1 - \frac{\sqrt{1-h}\,\sqrt{h+1}\,\left(12\,h^2-2\right)}{3\,h^2-2}$$

So, in this case we obtain the solution of (1), $u(x,t) = \operatorname{sech}^2(x - \lambda\, t)$

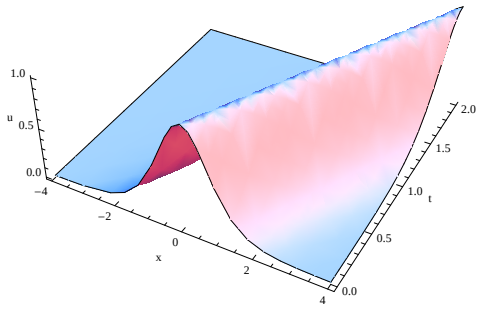



**Fig. 2.** Solutions $\mathbf{u} = \operatorname{sech}^2(x - 2\,t)$

Some of these solutions are soliton or kink solutions.

## 4 Multipliers method

Given a PDE a conservation law is a relation of the form

$$D_t(\Phi^t) + D_x(\Phi^x) = 0 \tag{10}$$

where $\mathbf{\Phi} = (\Phi^t, \Phi^x)$ represents the conserved density and flux, respectively, and $D_x$, $D_t$ denote the total derivative operators with respect to $x$ and $t$ respectively.

In [5] Anco and Bluman gave a general treatment of a direct conservation law method for partial differential equations expressed in a standard Cauchy-Kovaleskaya form in particular for evolution equations

$$u_t = G(x, u, u_x, u_{xx}, \ldots, u_{nx}).$$

The nontrivial conservation laws are characterized by a multiplier $\lambda$ with no dependence on $u_t$ satisfying

$$\hat{E}[u]\,(\Lambda u_t - \Lambda G(x, u, u_x, u_{xx}, \ldots, u_{nx})) = 0.$$

Here

$$\hat{E}[u] := \frac{\partial}{\partial u} - D_t\frac{\partial}{\partial u_t} - D_x\frac{\partial}{\partial u_x} + D_x^2\frac{\partial}{\partial u_{xx}} + \ldots.$$

The conserved current must satisfy

$$\Lambda = \hat{E}[u]\Phi^t$$

and the flux $\Phi^x$ is given by [11]

$$\Phi^x = -D_x^{-1}(\Lambda G) - \frac{\partial \Phi^t}{\partial u_x} G + G D_x \left( \frac{\partial \Phi^t}{\partial u_{xx}} \right) + \dots .$$

For equation (1) we get the following multipliers.

$$1, e^x, e^x u_x$$

Each multiplier determines a corresponding conserved density and flux:

$$\begin{aligned} \Lambda &= 1, \\ \phi^t &= u_t + u, \\ \phi^x &= u_{xxx} - \alpha u_x - F(u_x) \end{aligned}$$

$$\begin{aligned} \Lambda &= e^x, \\ \phi^t &= e^x(u_t + u), \\ \phi^x &= e^x(u_{xxx} - \alpha u_x - F(u_x)) - \int e^x(u_{xxx} - \alpha u_x - F(u_x))dx \end{aligned}$$

$$\begin{aligned} \Lambda &= \mathrm{e}^x u_x, \\ \phi^t &= \mathrm{e}^x u_t u_x, \\ \phi^x &= \int -\mathrm{e}^x u_x \left( a + \tfrac{d}{du_x} f\left(u_x\right) \right) du_x + 1/2 \left( 2\, u_x u_{xxx} - {u_{xx}}^2 - {u_t}^2 \right) \mathrm{e}^x. \end{aligned}$$

## 5 Conclusions

We have applied the Lie classical method to semilinear damped beam equation. Using the characteristic equation, the similarity variables are found. Then, the reduced form of the original nonlinear partial differential equation is obtained as a nonlinear ordinary differential equation. In order to obtain exact solutions we apply a direct method. By this method we have derived some travelling wave solutions. By the multipliers method we have obtained nontrivial conservation laws via integral formulae.

## Acknowledgments

The support of DGICYT project MTM2009-11875, FEDER funds and Junta de Andalucía FQM-201 group, is gratefully acknowledged.

## References


1. M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1972.
2. S. Anco, *Conservation laws of scaling-invariant field equations*, J. Phys. A, **36**, (2003), 8623–8638.
3. S. Anco and G. Bluman, *Direct construction of conservation laws from fields equations*, Phys. Rev. Letters, **78(15)**,(1997), 2869–2873.
4. S. Anco and G. Bluman, *Direct construction method for conservation laws of partial differential equations. Part I: Examples of conservation law classifications*, Euro. J. Appl. Math., **15(5)**, (2001), 547–568.
5. S. C. Anco, G.Bluman, *Direct constrution method for conservation laws for partial differential equations Part II: General treatment*, Euro. Jnl of Appl. Math. 2002, 41, 567-585.
6. G.W. Bluman, J.D. Cole, *The general similarity solution of the heat equation*, J.Math. Mech. 1969, 18, 1025-1042.
7. G.W. Bluman, S. Kumei, *Symmetries and differential equations*, Springer-Verlag, 1989.
8. J.C. Camacho, M.S. Bruzón, J. Ramirez, *Classical and Nonclassical Symmetries of the Beam Equation* pp. 2006.
9. B. Champagne, W. Hereman, P. Winternitz, *The computer calculation of Lie point symmetries of large systems of differential equations. Com. Phys. Comm.*, **66**, (1991), 319-340.
10. A.R. Champneys, P.J. McKenna,*On solitary waves of a piecewise linear suspended beam model*, Nonlinearity 10, (1997), 1763-1782.
11. N. Euler, M. Euler, *On nonlocal symmetries, nonlocal conservation laws and nonlocal transformations of evolution equations: two linearisable hierachies.* J. of Nonl. Math. Phys. 2009, 6, 489-504.
12. M.L. Gandarias, M.S. Bruzón, *Classical, and Nonclassical Symmetries of a Generalized Boussinesq Equation*, Journal of Nonlinear Mathematical Physics, Vol. 5, No. 1, 1998, pp. 8-12.
13. A.C. Lazer, P. J .McKenna, *Large Scale OscilIation Behavior in Loaded Asymnmetric Systems.* Ann. Inst. . H-Poincare, Analyse Nonlineaire: Vol. 4, pp.244-274, 1987.
14. P.J. McKenna, W. Walter, *Nonlinear Oscillation in a Suspension Bridge.* Arch. Rational Mech. Anal. 98. pp.167-177, 1987.
15. P.J. Olver,*Applications of Lie groups to differential equations*, Springer-Verlag, 1986.
16. H. Sano,*Finite-dimensional H8 control of flexible beam equation systems*, IMA Journal of Mathematical Control and Information, 19, 477-491. 2002
17. H. Takeda, *Global existence of solutions for higher order nonlinear damped wave equations*, Dynamical Systems and Differential Equations, Proceedings of the 8th AIMS International Conference, Dresden, Germany, 1358-1367 (2011).
18. H. Takeda, S. Yoshikawa, *On the initial value problem of the semilinear beam equation with weak damping I: smoothing effect*, J. Math. Anal. Appl. 401, 244-258 (2013)
19. Takeda, H, Yoshikawa, S: *On the initial data of the semilinear beam equation with weak damping II: asymptotic profiles.* J. Differ. Equ. 253 3061-3080 (2012).
20. H. Takeda, *Large time behavior of global solutions of the semilinear damped beam equation with slowly decaying data* , J. Math. Anal. Appl. 423, No. 2, 898-912 (2015). ISSN 0022-247X

# Automatic Detection of Metamorphic Relations: A Challenge for WS-BPEL

C. Castro-Cabrera and I. Medina-Bulo

UCASE Software Engineering Research Group,
University of Cadiz, Spain {maricarmen.decastro,inmaculada.medina,}@uca.es
https://ucase.uca.es/

**Abstract.** Web services have nowadays great impact on society due to numerous internet transactions existing. WS-BPEL is a Bussiness Process Enterprise Language that allows implement compositions as web services. This type of software requires to be tested to avoid errors and fatal consequences. In a previous work, the authors proposed to apply the Metamorphic Testing technique to WS-BPEL compositions through a particular architecture. That approach has some steps as the identification and implementation of properties to be used. This paper focuses on the composition analysis for obtaining information to automate that step.

**Keywords:** Metamorphic Testing, WS-BPEL, oracle, Metamorphic Relations, follow-up test cases

## 1 Introduction

Web Services Business Process language, WS-BPEL 2.0 [9] was standardized at the request of some TIC companies (HP, IBM, Oracle, Microsoft, etc.). This language allows us to develop a new Web Service (WS) designing more complex business processes from pre-existing WS, and there is a widely support software for them. However its development has not gone along with improvements on testing techniques to this type of software [1]. A deficient testing in a system could cause errors with negative consequences both economical and also human. Consequently, good testing methods to test correctness of compositions are required. Progresses in this sense are described in [10].
Metamorphic Testing (MT) [5] is a software testing technique using *metamorphic relations* (MRs). MRs are existing or expected relations defined on a set of inputs and their corresponding outputs for multiple executions of a function under test. The underlying concept is simple and its automation is not difficult. In fact, it has proved successful in testing and improving the quality of traditional imperative programs [14].
Regarding the cost effectiveness of MT, Zhang [13] conducted an experiment where the fault detection capabilities and time cost of MT were compared to the standard assertion checking method. Results showed that MT has the potential to detect more faults than the assertion checking method.

This paper discusses how to use MT to test WS compositions in WS-BPEL. Although MT has not been previously applied to this area, promising results have been obtained in a number of different applications. A component diagram for a testing framework implementing this approach is included as well as alternatives to automate the analysis step for the MR identification and implementation.
The structure of the rest of the paper is as follows: An introduction about metamorphic testing and WS-BPEL language are respectively shown in Section 2 and Section 3. Section 4 includes the particular architecture. Section 5 describes different alternatives for the analysis step and, finally, Section 6 presents the conclusions and future work.

## 2 Metamorphic testing

There are properties associated to some functions or applications, such that, if the inputs are changing (i.e.increase or decrease in a quantity), it should be possible calculate the new output through the output generated from the original input, without need to run the program. Therefore, if the new input is executed by the program, the output must be the same that had been calculated previously. The operations to apply to the inputs constitute properties, which relate the initial tests cases with the follow-up tests cases, they are called Metamorphic Relations (MR). MT is based on this notion and easily carried out in practice. The original test cases and their corresponding follow-up test cases are constructed based on these MRs. Both of them are executed using the program under test, to verify it. If any test case does not satisfy a MR, an error is detected.
For instance, given a program implementing the arithmetic mean of a set of numbers, permuting the order of the elements should not affect to the mean calculation; If other operations are applied such as, multiplying or increasing each value by a number, the resulting mean (follow-up test case output) should be easily predicted, multiplying or increasing (respectively) the original mean by that number. If the different outputs for their corresponding inputs are not as expected, then there must be a error in the implementation. Other example, based on lists of numbers it is showed in figure 1. The program $f$ plays the role of the *reverse order* function and $t$, multiply by 3, the MR:

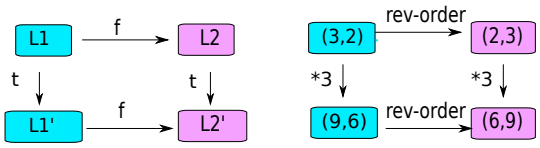



**Fig. 1.** Metamorphic Relation Example

## 3 WS-BPEL Composition Language

WS-BPEL is a programming language based on XML that is used to generate business processes from services defined previously. The resulting business process can be then reused as a WS in higher level compositions. A WS-BPEL composition contains four sections including declarations of: the relationships to the external partners, the variables, handlers, and description of the business process behavior. The major building blocks in WS-BPEL are the *activities*. Furthermore, WS-BPEL provides concurrency and synchronization primitives. Here is an example:

```
<flow>  ← Structured activity
 <links>  ← Container
  <link name="checkFl-BookFl"/> ← Element
 </links>
 <invoke name="checkFlight" ... > ← Basic activity
  <sources>  ← Container
   <source linkName="checkFl-BookFl"/> ← Element
  </sources>
 </invoke>
 <invoke name="checkHotel" ... />
 <invoke name="checkRentCar" ... />
 <invoke name="bookFlight" ← Attribute ...>
  <targets>  ← Container
   <target linkName="checkFl-BookFl" />
  </targets>
 </invoke>
</flow>
```

## 4 MT implementation and architecture

Firstly, it is mentioned a general implementation of MT, presented in [4]. The sequence is: Choose the initial test suite, select adequate MRs, generate the follow-up test suite applying MRs to initial test suite, execute the program with initial and follow-up test cases, compare the result and finally, improve the program correcting the detected errors, select new test cases and/or new MRs enhanced to successive iterations.

The goal is to implement MT to test WS-BPEL compositions. Therefore, it is necessary to take into account the peculiarities of this language and the compositions. Figure 2 describes the particular architecture lightly improved to include MuBPEL [6], a mutation tool to validate the technique. MuBPEL is a mutation testing tool for WS-BPEL 2.0. It can be used to evaluate the quality of a test suite by checking if it can tell apart a mutant from the original program. Mutants are slightly modified (mutated) versions of the original program in which a single syntactical change has been made: for example, "$4 + 5$" may have been changed to "$4 - 5$" *or* "$4 + 6$". In this way, MuBPEL is used to validate the MT technique. If a test case does not satisfy a MR, an error is detected (a mutant is killed).

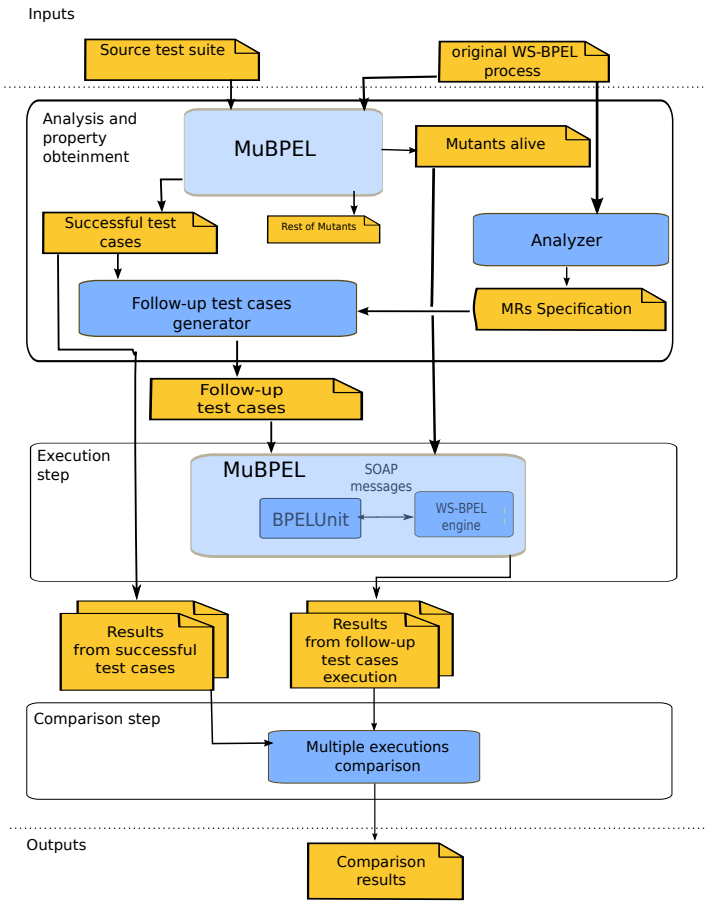



**Fig. 2.** Architecture to apply MT in WS-BPEL

A prototype was built following this arquitecture and analysing some compositions to obtain the properties. The MRs were designed and implemented by hand for each composition. Besides in previous works [2] and [3] that prototype was applied to some case studies with promising results. However, both the composition analysis and the obtained MRs were processed by hand.

Some MT applications on the literature are about problems or program whose properties are known previously. For instance, there are numerous works about machines learning applications or mathematical functions [8], [12]and [7]. This step is a challenge for WS-BPEL compositions. Several alternatives are proposed in the following section.

## 5 Alternatives to analyse the composition

Our goal is scanning every composition and to extract relevant information that assesses the framework to design and implement MRs. So somehow, this step makes easy generating MRs. For this purpose, we offer some possibilities (not exclusive):

1. Up the bussiness level (bussines rules)
2. Use Takuan [11] (invariant generator to WS-BPEL compositions)
3. Create a new analyzer to extract other information

The first option does not appear simple or applicable (The usual approach is the inverse step). The natural step goes in the inverse order. This requires working on the bussiness level to extract the composition behaviour, implementing properties and applying them.
With respect to the second option, Takuan is an open-source WS-BPEL dynamic invariant generator which can infer invariants from WS-BPEL process definitions. It generates relevant information, but perhaps too simple for our purpose. For instance, there are some invariants from Loan Approval composition as follow:

loanApprovalProcess._process1_sequence1:::EXIT
request.amount == orig(request.amount)
request.amount one of { 1500, 15000, 150000 }
risk.level one of { "high", "low" }

It would be necessary to combine them to obtain information useful for MR implementation. So, they could be complemented with the third option, developing a new application to extract information to asses the MR implementation. The goal of the third option is to locate key values and expressions in every composition to determine the properties to build. For example, a numeric constant in a condition could lead to an arithmetic property or a logical expression could lead to a logical relation between some test cases values.

## 6 Conclusions and future work

WS-BPEL business processes are considerably increasing in last years. For this reason, it is important the development of techniques that allow to test this type of software. Due to the language nature, specific to WS, it is necessary to implement alternatives tecniques to test this kind of compositions.
In addition, MT has been implemented in different languages efficiently and applications have been tested on various study fields such as medicine or bioinformatics. Actually, more than 80 papers have been published about this subject. Selection of adequate MRs is an important issue to this technique, so we ought to consider the problem context and the structure of the program under test.
A testing framework architecture to apply MT in WS-BPEL compositions is being built. Further some possibilities to guide the property identification and implementation have been presented. The future work includes to implement all steps and compare with other techniques.

## References


1. Bozkurt, M., Harman, M., Hassoun, Y.: TR-10-01: testing web services: A survey. Tech. Rep. TR-10-01, King's College, London (2010)
2. Castro-Cabrera, C., Medina-Bulo, I.: Application of Metamorphic Testing to a Case Study in Web Services Compositions, Communications in Computer and Information Science, vol. 314, pp. 168–181. Springer Berlin Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-35755-8_13
3. Castro-Cabrera, C., Medina-Bulo, I., Camacho, A.: Metasearch services composition in ws-bpel - an application of metamorphic testing. In: ICSOFT (2012)
4. Castro-Cabrera, M.d.C., Camacho-Magriñàn, A., Medina-Bulo, I., Palomo-Duarte, M.: Una arquitectura basada en pruebas metamòrficas para composiciones de servicios ws-bpel. In: Actas de las VII JCIS. pp. 9–22. Servizo de publicaciòns da Universidade da Coruña, A Coruña, España (Sep 2011)
5. Chen, T.Y.: Metamorphic testing: A new approach for generating next test cases. HKUSTCS98-01 (1998)
6. García-Domínguez, A., Estero-Botaro, A., Medina-Bulo, I.: MuBPEL: una herramienta de mutación firme para WS-BPEL 2.0. In: Actas de las XVI JISBD (2012)
7. Mahmud, S., Elahe, M.F., Jahan, H.: Verifying mathematical function using metamorphic testing. International Journal of Research in Computer Engineering & Electronics 3(5) (2014)
8. Murphy, C., Kaiser, G., Hu, L., Wu, L.: Properties of machine learning applications for use in metamorphic testing. In: Proc. of the 20th International Conference on SEKE (SEKE). pp. 867–872 (2008)
9. OASIS: Web Services Business Process Execution Language 2.0. http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html (2007), Organization for the Advancement of Structured Information Standards
10. Palomo-Duarte, M.: Service composition verification and validation. In: Jonathan Lee, S.P.M., Liu, A. (eds.) Service Life Cycle Tools and Technologies: Methods, Trends and Advances, pp. 200–219. IGI Global (2011)
11. Palomo-Duarte, M., Garcia-Dominguez, A., Medina-Bulo, I.: Automatic dynamic generation of likely invariants for ws-bpel compositions. Expert Systems with Applications 41(11), 5041 – 5055 (2014), http://www.sciencedirect.com/science/article/pii/S095741741400061X
12. Xie, X., Ho, J.W., Murphy, C., Kaiser, G., Xu, B., Chen, T.Y.: Testing and validating machine learning classifiers by metamorphic testing. Journal of Systems and Software 84(4), 544 – 558 (2011), the Ninth ICQS
13. Zhang, Z.Y., Chan, W.K., Tse, T.H., Hu, P.F.: An experimental study to compare the use of metamorphic testing and assertion checking. Journal of Software 20(10), 2637–2654 (2009)
14. Zhou, Z.Q., Huang, D.H., Tse, T.H., Yang, Z., Huang, H., Chen, T.Y.: Metamorphic testing and its applications. In: Proceedings of the 8th ISFST(ISFST 2004). Software Engineers Association (2004)

# Semandal: Extracting knowledge from city councils *

Daniel Albendín Moya[1], Gonzalo A. Aranda-Corral[1], Joaquín Borrego-Díaz[2], and Angel Cantó Vicente[1]

[1] Universidad de Huelva. Department of Information Technology. Crta. Palos de La Frontera s/n. 21819 Palos de La Frontera. Spain
[2] Universidad de Sevilla. Department of Computer Science and Artificial Intelligence. Avda. Reina Mercedes s/n. 41012 Sevilla. Spain

**Abstract.** One of the main goals of Semantic Web is to make all available information machine-readable and understood by other machines. For these, ontologies are key elements that will enable us to exploit all the advantages. Ontologies try to model the world in order to represent all web information. But, for general purposes, this is too broad and ambitious goal for only a single ontology or platform. In order to easy the creation of the ontology we reduce the scope only to news extracted from city councils web pages.
In this paper, we present a project where information and data are collected from webs and processed by means of Formal Concept Analysis to align it to an "ad-hoc" ontology.

**Keywords:** Knowledge extraction, Formal concept analysis, Semantic Web

## 1 Introduction

As the W3C establish in its web page: "*The Semantic Web[2] provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries...* In order to achieve these goals, traditional web should be translated into data or machine readable documents, located by URIs, and they are also be related to others. Semantic Web technologies can be used in a variety of application areas. Our interest is focused on information integration, where information from different sources can be organized to enhance its access, organization, etc.

Nowadays, Open Government Data (OGD) is other important emerging trend that merges the Open Data foundations with Public entities. OGD are data produced by or commissioned by government and its intended for freely used, reused and redistributed by anyone. It has great benefits as transparency, social and commercial value and let participatory Governance.

* Partially supported by TIC-6064 Excellence project (Junta de Andalucía) and TIN2013- 41086-P (Spanish Ministry of Economy and Competitiveness), cofinanced with FEDER funds

Semandal is a platform that tries to apply all concepts about Semantic Web and Open Government Data on the municipalities of Andalucía, Spain. This scope was chosen to reduce the dimensions of vocabularies, ontologies, and, even, databases. Semandal extracts the information from the traditional web pages and transform it into knowledge and republish it by means of an API and a structured format (machine readable). Even this, Semandal provides a mobile app to let user to access this information, but this is out of scope of this paper.

To transform information into knowledge we use Formal Concept Analysis[1] (FCA). FCA is a non supervised clustering technique which, from formal definitions, can construct concept lattices and set of rules which represents information.

Semandal's architecture is depicted in fig 1, where main modules can be distinguished and will lead the structure of our paper (only two firsts): Extraction, Classification and re-publishing.

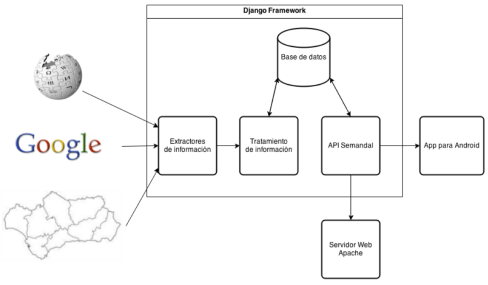



**Fig. 1.** Semandal's Architecture

### 1.1 Formal Concept Analysis

A useful bridge between Semantic Web and knowledge extraction could be *Formal Concept Analysis* (FCA) [1].

According R. Wille, FCA mathematizes the philosophical understanding of a concept as a unit of thought composed by two parts: the extension and the intension. The extension covers all objects (documents) belonging to this concept, while the intension comprises all common attributes (tags) valid for all the objects under consideration. It also allows the computation of concept hierarchies out of data tables, and it is also used for ontology mining from folksonomies. Several applications from FCA to tagging, folksonomies and semantic tasks have been developed (see [6]).

We represent a formal context as $M = (O, A, I)$, which consists of two sets, $O$ (objects) and $A$ (attributes) and a relation $I \subseteq O \times A$. Finite contexts can be represented by a 1-0-table (representing I as a Boolean function on $O \times A$). Basic FCA logical expressions are implications between attributes, that is, a pair of sets of attributes, written as $Y_1 \rightarrow Y_2$, which is true with respect to

$M = (O, A, I)$ and is defined as follows. A subset $T \subseteq A$ *respects* $Y_1 \to Y_2$ if $Y_1 \not\subseteq T$ or $Y2 \subseteq T$. We say that $Y_1 \to Y_2$ holds in $M$ ($M \models Y_1 \to Y_2$) if for all $o \in O$, the set $o'$ respects $Y_1 \to Y_2$. In that case, we say that $Y_1 \to Y_2$ is an implication of $M$.

Every implication has also associated some properties, e.g. support. Support is defined as the number of objects that contain all attributes from $Y_1$ and holds the implication.

## 2 Knowledge extraction and integration

One of the main goals of Semantic Web is information integration from different sources and, in our case, we start with a file of municipalities, obtained from INE[3], written in XLS format. From these, we develop software to connect to several APIs, from Google, Wikipedia, etc.. From Google Search we locate the web page address for all municipalities and, in some cases, where disambiguation is needed, we use Wikipedia to achieve it. Also we used other APIs to extract the geolocation, population, extension, etc...

After all web addresses were collected we found a big challenge, news extractions. All web pages were organized in really different ways which makes this task really hard. Based on most common patterns, we found up to 7 different clusters of similar organizations, and we built an extractor for each one.

Some of them, based on page's structure or news content, can extract the category associated to news, classified by publishers (supposed to be valid). Other extractors were not able to do it. From these, we obtained a set of classified and other unclassified news, which are quite appropriate to apply machine learning techniques.

## 3 Categories selection

From the set of classified news, we study the set of categories to remove any possible mistake.

First task was to remove typos, plurals or abbreviations, grouping words based on Levenshtein distance which are closer than 3, obtaining a set of categories not too big and working.

In order to semantize as much content as possible, we searched for a hierarchy or ontology where map our set of categories soundly, but we did not find any one which fits with our needs about municipalities news. To solve this, we created an "ad hoc" categories' hierarchy (fig. 2) where also some of extracted categories are removed (out of meaning in our scope) and merging some of them with the same meaning. Needed super-classes are added too.

From this hierarchy, we reclassify all news applying it: aligning categories to hierachy's concepts and adding all super-classes.

[3] Instituto Nacional de Estadística.(National Institute of Statistics)

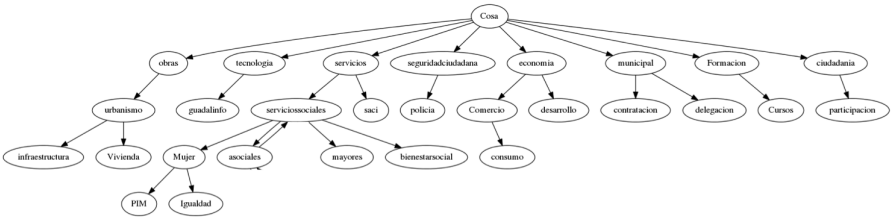



**Fig. 2.** Hierarchy of categories

## 4 Classification of information

Now, we are able to classify all news which were previously unclassified from its own content and other classified news. From these news, we planed to build a classifier based on words contained into the text. Firstly, we remove non significant words, as prepositions, articles, own names, numbers, etc...

Secondly, we calculated the relative weights of each word in the categories where they appear to obtain the significance of it into the category. After, we calculated the average of weights for each category and, if this number is high means that the word is too common in all categories and we should remove it. Other considered option, but with similar results we based on [3]

If we create a graph with resulting words connecting them and categories, setting each edge's weight to calculated frequency value. This graph is shown in fig. 3.

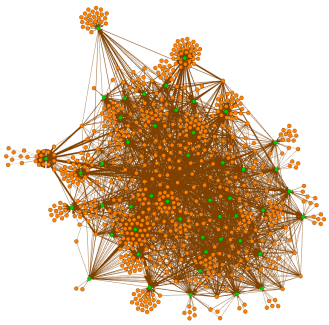


**Fig. 3.** Categories - Words

As we can see, there are words that define categories fairly well, but others remain poorly relevant. Nevertheless, we think this set of words is good enough to implement the classifier.

### 4.1 Formal Concept Analysis

Construction of classifier we made by means of a set of rules obtained applying Formal Concept Analysis technique. As above, we need to build a formal context to get a concept's lattice, as first step, which we can take as our first emergent ontolgy prototype about news, and , finally, a set of association rules which will be the responsible to classify news.

The formal context considered classified news as objects and relevant words and extracted categories as attributes. This will infer a set of attribute rules (with support and confidence) which, applied to unclassified news it will infer a new set of categories. Each new set of categories will be the new assigned to each news. An example of this context is shown in fig. 4.

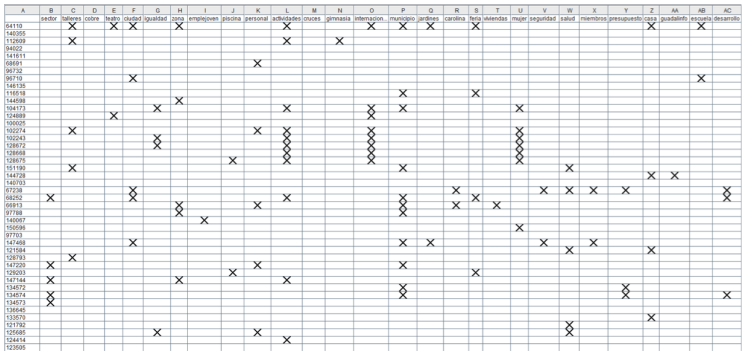


**Fig. 4.** Formal context

## 5 Experiments

For our experimentation, we had to choose a proper subset of news, since the total number of news is huge. We prepared some experiments to find out an affordable amount of size or time in order to obtain sound results.

We built 3 contexts, A, B and C, with different number of objects and attributes obtaining a number of rules which grow exponentially, as shown in table 1. In order to reduce this number of rules, we only considered as valid rules (Classif) that ones which have a $support > 0$ and have, at least, one category on the right side of the rules.

To build the expert system to generate new categories, we translate the rules into a CLIPS[4] format and we run it by means of a Jess[5] program, also developed by us.

Finally, we test the expert systems with 2 random news to check its soundness. News are in spanish, because of all database is focused on Spanish municipalities.

| Context | Attributes | Objects | Rules | Classif. |
|---|---|---|---|---|
| A | 145 | 133 | 7412 | 2650 |
| B | 208 | 258 | 26344 | 9813 |
| C | 257 | 372 | 66910 | 53898 |

**Table 1.** Size of experiments

| | |
|---|---|
| [Noticia 1] "*El novillero de Écija Antonio David, proclamado triunfador de la V feria de novilladas de promoción la granada de plata*" [4] | [Noticia 2] "*El ayuntamiento da luz verde para la construcción de otras 75 viviendas protegidas*"[5] |
| A: Turismo, Juventud<br>B: Turismo, Cultura<br>C: Festejos | A: Vivienda<br>B: Turismo, Servicios sociales<br>C: Servicios sociales, Obras |

## 6 Remarks and future works

Building a platform for extracting and processing information is a really hard task for machines, even when it is limited in scope, that still needs some work. Nonetheless, we have shown, in this paper, that FCA could be a useful tool to build classifiers that would allow us to transform extracted information into knowledge in an affordable way with soundness.

There are still other tasks within this platform which are really relevant for final success, as the acquisition of municipality's general information, not only news, it means, information about organizational structure, contacts, plenary sessions, etc... . From knowledge point of view, next step should be a deep work on words, using networks of concepts, as WordNet, to align them to the network and integrate in some RDF Open Data catalog.

## References

1. B. Ganter and R. Wille. Formal Concept Analysis - Mathematical Foundations. Springer, 1999.
2. T. Berners-Lee, J. Hendler, O. Lassila, *The Semantic Web*, Scientific American, 284(5), pp. 34–43, 2001
3. Gerard Salton and Michael J. McGill. 1986. Introduction to Modern Information Retrieval. McGraw-Hill, Inc., New York, NY, USA.
4. CLIPS (2009). CLIPS: A tool for building expert systems.
5. Ernest Friedman Hill. 2003. Jess in Action: Java Rule-Based Systems. Manning Publications Co., Greenwich, CT, USA.
6. R. Jaschke, A. Hotho, C. Schmitz, B. Ganter, and G. Stumme. Discovering shared conceptualizations in folksonomies. Journal of Web Semantics, 6(1):pp 38-53, 2008.

# Dichotomous sets of implications and directness

E. Rodríguez-Lorenzo, P. Cordero, M. Enciso, and A. Mora

University of Málaga, Andalucía Tech, Spain,
e-mail: {estrellarodlor,amora}@ctima.uma.es, {pcordero,enciso}@uma.es

**Abstract.** In this paper, we study the directness property of implications in formal concept analysis. We show how dichotomously split the implications in two subsets according to their premise closure. Thus, we define a new directness paradigm strongly based on a separated treatment of the two implication subsets and how to compute the proposed dichotomous direct implicational system from a set of implications.

## 1 Introduction

In Formal Concept Analysis (FCA), closure operators are one of the most basic notions. These operators allow to solve important exponential problems in different areas such as formal concept analysis, AI, databases, etc.

Moreover, closure operators are directly related to implications. K. Bertet points to a good direction in [3] where Implicational Systems (IS) are highlighted as convenient tools to handle a closure system. So, it makes sense that the search for the efficiency in the set closure computation is a major challenge [1].

In [4] K. Bertet et al. establish specific properties to achieve the mentioned goal. The properties that they considered is the directness and optimality, that is, the computation of the closure of an attribute set can be performed in one traversal of the implicational set and none implication can be removed without losing this property. Then, progressing this line, K. Adaricheva et al. [1] propose the so called *D-basis* as a subset of the basis proposed in [4], which is direct as well and has less number of implications.

In this paper, we are working in the design of new IS definitions suitable to describe closure system. We pay attention to the closure of each premise to make a separate treatment of those ones whose closure is the total set of attributes to provide an improvement in the performance of closure methods. Moreover, the premise of these implications fit exactly with the notion of keys in database and generators in FCA.

Here, we propose a new definition of IS named dichotomous IS whose main characteristics is the separate treatment of implications depending on the closure of its premise. Moreover, we introduce the notion of direct dichotomous basis (DD-basis) and illustrate its advantages.

## 2 Formal Concept Analysis

Formal Concept Analysis (FCA) is a formal framework oriented to data analysis and knowledge discovering. FCA extracts knowledge from the information

presented in a formal context providing equivalent ways of representation of knowledge: concept lattices and implicational systems. We focus on the second ones because they can be managed and depurated by using logic.

The relationship between a set of objects and a set of attributes are described using a formal context as follows:

**Definition 1.** *A formal context is a triple* $\mathbf{K} = (G, M, I)$ *where* $G$ *is a finite set whose elements are named objects,* $M$ *is a finite set whose elements are named attributes and* $I \subseteq G \times M$ *is a binary relation. Thus,* $(g, m) \in I$ *means the object* $g$ *has the attribute* $m$*.*

The concept of implication is a central point in this work together with a special case of implication introduced as follows:

**Definition 2.** *Let* $\mathbf{K} = (G, M, I)$ *be a formal context and* $A, B \in 2^M$*. The implication* $A \to B$ *holds in* $\mathbf{K}$ *if every object* $g \in G$ *satisfies the following:* $(g, a) \in I$ *for all* $a \in A$ *implies* $(g, b) \in I$ *for all* $b \in B$*.*

*Given a subset* $X \subseteq M$*,* $X$ *is a key if the implication* $X \to M$ *holds in* $\mathbf{K}$*. An implication* $X \to Y$ *such that* $X$ *is a key will be named a key implication.*

For this work, we only present a brief summary (only the rules) of the Simplification Logic ($\mathbf{SL}_{\text{FD}}$), an equivalent logic to Armstrong's Axioms [2] but more adequate to develop automated method for implications. See [5, 8] for a more detailed presentation of $\mathbf{SL}_{\text{FD}}$ its semantic, and how remove redundancy and compute closures using directly the $\mathbf{SL}_{\text{FD}}$.

It is assumed to be familiar with notions of the derivation of an implication from an IS, the semantic entailment and the equivalence between two ISs.

**Definition 3 (Rules of $\mathbf{SL}_{\text{FD}}$).**
*Reflexivity as axiom scheme and the following inference rules named fragmentation, composition and simplification are considered in* $\mathbf{SL}_{\text{FD}}$*.*

$$\texttt{[Ref]}\ \frac{}{A \to A}\quad \texttt{[Frag]}\ \frac{A \to BC}{A \to B}\quad \texttt{[Comp]}\ \frac{A \to B,\ C \to D}{AC \to BD}\quad \texttt{[Simp]}\ \frac{A \to B,\ C \to D}{A(C\text{-}B) \to D}$$

**Definition 4.** *Let* $\Sigma \subseteq \mathcal{L}_S$ *be an IS and* $X \subseteq S$*. The closure of* $X$ *wrt* $\Sigma$ *is the largest subset of* $S$*, denoted* $X^+_\Sigma$*, such that* $\Sigma \vdash X \to X^+_\Sigma$*.*

Finally, the notion of key, which plays a central role in this paper, can be characterized in terms of ISs. Thus, we are going to say that an attribute set $X$ is key with respect to an IS $\Sigma$ if it is a key with respect to any model of $\Sigma$.

**Proposition 1.** *Let* $\Sigma \subseteq \mathcal{L}_S$ *be an IS and* $X \subseteq S$*. The following conditions are equivalent:*

1. $X$ *is a key with respect to any model of* $\Sigma$*.*
2. $\Sigma \vdash X \to S$*.*
3. $X^+_\Sigma = S$*.*

## 3 Dichotomous set of implications

Fist, we emphasize that although some closure algorithms to solve these problems have a linear cost, due to its exhaustive use in some NP algorithms, a minor gain in the closure performance entails a major advantage for these complex methods.

Now, the goal is to remain the advantage of the directness which establishes that the closure of an attribute set may be computed with just one traverse of the set of implications. Thus, our approach is to provide an alternative direct IS such that it can be obtained with less cost.

The study of directness demands the design of closure operators for attribute sets. Thus, other authors who have studied different kinds of direct basis define closure operators for these bases: the direct-optimal basis [4] and the D-basis [1].

Now, we introduce the notion of dichotomous set of implications and a two-fold operator suitable for its management.

**Definition 5 (Dichotomous set of implications).** *A pair of implicational sets $\langle \Sigma^*, \Sigma^k \rangle$ is named a dichotomous set of implications if all $A \to B \in \Sigma^k$ are key implications.*

We define the $\sigma$ operator for dichotomous ISs as a composition of two underlaying operators: $\sigma_{\langle \Sigma^*, \Sigma^k \rangle} = \kappa_{\Sigma^k} \circ \pi_{\Sigma^*}$ [1] where

$$\kappa_{\Sigma}(X) = \begin{cases} M & \text{if } A \subseteq X \text{ for some } A \to B \in \Sigma \\ X & \text{Otherwise} \end{cases}$$

Assuming $\langle \Sigma^*, \Sigma^k \rangle$ being a dichotomous set of implications, we have that $\sigma_{\langle \Sigma^*, \Sigma^k \rangle}$ is isotone and extensive. The directness property can also be considered in this framework by means of the idempotence of the operator $\sigma_{\langle \Sigma^*, \Sigma^k \rangle}$.

**Definition 6.** *A dichotomous set of implications $\langle \Sigma^*, \Sigma^k \rangle$ is said to be direct if $\sigma_{\langle \Sigma^*, \Sigma^k \rangle}$ is a closure operator.*

The following proposition leads the way to characterize a direct dichotomous set of implications and the next corollary characterizes the directness in our dichotomous approach.

**Proposition 2.** *Let $\langle \Sigma^*, \Sigma^k \rangle$ be dichotomous set of implications. For all $n > 0$ we have that $\sigma^n_{\langle \Sigma^*, \Sigma^k \rangle} = \kappa_{\Sigma^k} \circ \pi^n_{\Sigma^*}$.*

**Corollary 1.** *A dichotomous set of implications $\langle \Sigma^*, \Sigma^k \rangle$ is direct if and only if $\Sigma^*$ is a direct IS.*

In this section we have focused on the directness property. Nevertheless, this is not an isolate property and it is usually related to other properties leading to the notion of basis, which constitutes the main issue of the following section.

[1] $\pi_{\Sigma}$ is the operator defined in [3]: $\pi_{\Sigma}(X) = X \cup \{b \in B | A \subseteq X \text{ and } A \to B \in \Sigma\}$

## 4 DD-basis

In this section we are interested in the improvement of ISs demanding some optimality or minimality properties. The property related to the notion of basis is minimality, which means that if any implication is removed from the set of implications, it is not equivalent to the initial one. But in this paper, the property that we need is the optimality: an IS is optimal if there is not an quivalent IS with less size, where the size is: $\|\Sigma\| = \sum_{A \to B \in \Sigma}(|A| + |B|)$.

The above properties are used in the literature to introduce different notions of basis, but another property must be introduced. The well-known Duquenne-Guiguess basis was introduced in [7] being a minimum (there is not another equivalent IS with less cardinality), but no direct, IS. In this paper we are interested in those basis strongly related with the directness property. In this way, Bertet et al. propose the following definition adding the optimality property.

**Definition 7 (Direct-optimal basis [3]).** *An IS $\Sigma$ is named a direct-optimal basis if $\Sigma$ is direct and any other equivalent direct IS has a greater size.*

Alternatively, K. Adaricheva et al. [1] introduce another basis related to the directness property and taking into account the order of the implications: the D-basis (see [1] for more details). A relevant issue in the area of direct basis is the cost of its computation. That is why we provide here an alternative direct basis definition in the framework of the dichotomous ISs.

**Definition 8 (DD-basis).** *Let $\langle \Sigma^*, \Sigma^k \rangle$ be a dichotomous set of implications, we say that it is a dichotomous direct basis, briefly DD-basis, if it is minimal and $\sigma_{\langle \Sigma^*, \Sigma^k \rangle}$ is a closure operator.*

The main advantage of the proposed DD-basis with respect to both alternative direct bases, is that our approach reduces the size of the subset of implications withstanding the exponential cost of the basis construction process, as the following section shows. This reduction comes from the removal of the key implications in the exponential task.

## 5 A method to compute the DD-basis

Most of the knowledge discovering methods in FCA returns a Duquenne-Guigues basis. In this section we focus on the design of an efficient method to compute a DD-basis equivalent to a given Duquenne-Guigues basis.

The use of the dichotomous set of implications is motivated by the idea of reducing the input of the costly task in the basis computation method. Corollary 1 establishes that a dichotomous set of implications is direct if and only if the first component is direct. Thus, we begin this section with the description of the transformation method of the first component $\Sigma^*$ into the corresponding equivalent direct-optimal basis by executing a modification of the algorithm we presented in [9]. Now, we briefly describe this algorithm. In a first stage the method simplifies implications with redundant attributes in both left and

right hand sides, transforming $\Sigma^*$ into its equivalent reduced one, denoted $\Sigma^*_r$, requires just the application of the rules of the $\mathbf{SL}_{\text{FD}}$.

Later, the algorithm exhaustively applies the inference rule called *strong simplification*, **[sSimp]**, that covers directness without losing reduceness:

$$\texttt{[sSimp]} \quad \frac{A \to B, C \to D}{AC\text{-}B \to D\text{-}(AB)},\ B \cap C \neq \emptyset \neq D \smallsetminus (A \cup B)$$

The thoroughly application of the **[sSimp]** rule to the set $\Sigma^*_r$ provides an equivalent direct-reduced IS, named $\Sigma^*_{dr}$, being the smallest IS fulfilling the following conditions: $\Sigma^*_r \subseteq \Sigma^*_{dr}$ and $\Sigma^*_{dr}$ is closed under the **[sSimp]** rule.

Once we have got a direct-reduced IS, we can further depurate it by removing extra-attributes and extra-implications thanks to the application of the rules of $\mathbf{SL}_{\text{FD}}$. The target IS for such a depuration step is said to be simplified holding the following conditions: for all $A, B, C, D \subseteq M$,

1. $A \to B$, $A \to C \in \Sigma$ implies $B = C$.
2. $A \to B$, $C \to D \in \Sigma$ and $A \subsetneq C$ imply $C \cap B = \emptyset = D \cap B$.

Thus, for the treatment of $\Sigma^*$, the Algorithm *DObasis* has three main stages, each one consisting in the transformation of a previous IS into an equivalent one fulfilling directness and optimality at the end of the process: the direct-optimal basis (see [9] for the proof of this assertion).

In the following, we are going to use this function to transform the first component of a dichotomous IS and compute its equivalent direct-optimal basis.

We begin with the transformation of the original set of implications into a dichotomous one splitting off the treatment for key implications and the others, which provides a better performance of the basis construction method. Note that in Duquenne-Guigues basis key implications are those $A \to B$ that satisfy $A \cup B = M$. These implications have to belong to the second component of the dichotomous set of implications.

Algorithm 1 below structures the above transformation in two consecutive stages: the splitting process (discerning what implications are keys) and then, the direct-optimal transformation for the first component. The following example illustrates the execution of Algorithm 1.

*Example 1.* Let $\Sigma = \{a \to d, ce \to g, cg \to e, de \to g, bg \to acde, cd \to abeg, abd \to ceg, adeg \to bc\}$ be an IS. In the first stage, we separate the key implications with a linear cost rendering: $\langle \Sigma^*, \Sigma^k \rangle = \langle \{a \to d, ce \to g, cg \to e, de \to g\}, \{bg \to acde, cd \to abeg, abd \to ceg, adeg \to bc\} \rangle$.

In the second stage, we apply Function DObasis to get a direct-optimal basis equivalent to $\Sigma^*$: $\Sigma^*_{do} = \{a \to d, ae \to g, ce \to g, cg \to e, de \to g\}$.

Finally, we joint both components of the dichotomous set of implications to get a DD-basis: $\Sigma_{DD} = \langle \{a \to d, ae \to g, ce \to g, cg \to e, de \to g\}, \{bg \to acde, cd \to abeg, abd \to ceg, adeg \to bc\} \rangle$.

## 6 Conclusions and future works

In this work, we have presented a new definition for ISs in which we characterize two sets of implications with specific properties. In this way, a new direct

**Algorithm 1: DD-basis**

**input** : A Duquenne-Guigues basis $\Sigma_{DG}$ on $M$
**output**: The DD-basis $\Sigma_{DD}$ on $M$
**begin**
  /* Stage 1: Generation of $\Sigma^*$ and $\Sigma^k$ by disjointing of $\Sigma_{DG}$ */
  $\Sigma^* = \emptyset, \Sigma^k = \emptyset$
  **foreach** $A \rightarrow_{\Sigma_{DG}} B$ **do**
    **if** $A \cup B = M$ **then** add $A \rightarrow B$ to $\Sigma^k$;
    **else** add $A \rightarrow B$ to $\Sigma^*$;
  /* Stage 2: Generation of $\Sigma^*_{do}$ by executing the above function of $\Sigma^*$ */
  $\Sigma^*_{do} := DObasis(\Sigma^*)$
  /* Output preparation: Generation of $\Sigma_{dd}$ by jointing $\Sigma^*_{do}$ and $\Sigma^k$ */
  $\Sigma_{DD} := \langle \Sigma^*_{do}, \Sigma^k \rangle$
  **return** $\Sigma_{DD}$

basis and the algorithm that renders this new basis, called DD-basis, have been proposed. The main goal we have achieved is the reduction of the cost of computing a direct basis focusing in one subset of the IS: the first component of the dichotomous set. As future work we are going to extend the proposed algorithm when the input was any IS and make a comparative among algorithms related to directness property.


### Acknowledgment

Supported by grant TIN11-28084 and TIN2014-59471-P of the Science and Innovation Ministry of Spain.


## References


1. K. V. Adaricheva and J. B. Nation and R. Rand, *Ordered direct implicational basis of a finite closure system*, International Symposium on Artificial Intelligence and Mathematics, ISAIM 2012.
2. W W. Armstrong, *Dependency structures of data base relationships*, Proc. IFIP Congress. North Holland, Amsterdam: 580–583, 1974.
3. K. Bertet, M. Nebut, *Efficient algorithms on the Moore family associated to an IS*, DMTCS, 6(2): 315–338, 2004.
4. K. Bertet, B. Monjardet, *The multiple facets of the canonical direct unit implicational basis*, Theor. Comput. Sci., 411(22-24): 2155–2166, 2010.
5. P Cordero, A. Mora, M. Enciso, I.Pérez de Guzmán, *SLFD Logic: Elimination of Data Redundancy in Knowledge Representation*, LNCS, 2527: 141–150, 2002.
6. B. Ganter, *Two basic algorithms in concept analysis*, Technische Hochschule, Darmstadt, 1984.
7. J.L. Guigues and V. Duquenne, *Familles minimales d'implications informatives résultant d'un tableau de données binaires*, Math. Sci. Humaines: 95, 5–18, 1986.
8. A. Mora, M. Enciso, P. Cordero, and I. Fortes, *Closure via functional dependence simplification*, IJCM, 89(4): 510–526, 2012.
9. E. Rodríguez Lorenzo, K. Bertet, P. Cordero, M. Enciso, A. Mora, *The Direct-optimal Basis via Reductions*, CLA: 145-156, 2014.

# A new functional measure of skewness based on the convex transform order

Antonio Arriaza-Gómez, Miguel A. Sordo, and Alfonso Suárez-Llorens

Dpto. Estadística e Investigación Operativa, Universidad de Cádiz,
Facultad de Ciencias, Campus Universitario Río San Pedro s/n,
11510 Puerto Real, Cádiz, Spain
{antoniojesus.arriaza,mangel.sordo,alfonso.suarez}@uca.es

**Abstract.** The tail behaviour of a probability distribution has been widely studied in order to provide robust tools to deal with risk in different fields, such as financial or insurance risk, best known as actuarial theory. In this paper, a new functional skewness measure from the comparative study of the left and right tails of a distribution is provided. The new measure is based on the convex transform order, which let us compare whenever one distribution has heavier tail than another. We study the properties of the functional measure and we shall prove that it allows to detect a tail property called *symmetry in tails.*

**Keywords:** skewness, asymmetry, heavy-tailed, convex transform order, risk.

## 1 Introduction

Asymmetry of a continuous distribution is commonly described as skewness, and it has been widely studied in order to measure meaningful differences in the behaviour of the distribution in respect to some location parameter as the mean, median, mode, etc. It is a general practice to make assertions about the symmetry or asymmetry of a probability density function based on scalar measures. Since most of them use all the information of the distribution to summarize in a single number, they do not capture all the meaning of being a symmetric distribution. There are several scalar measures used to quantify the degree of skewness of a distribution, some of most known are [1], [4], [9] and [11].

Asymmetry of probability distribution has been also studied applying a functional approach. Since the symmetry is a functional concept, it seems suitable to describe asymmetry using asymmetry functions. A partial list of the most important measures taking a functional approach include [2], [3], [5], [6], [7] and [8].

Since asymmetry is essentially influenced by the tail behavior of density function, this lead us to study the symmetry of a distribution from the comparative study of its tails. The tail of a distribution is the portion of distribution corresponding to large or small values of the random variable and its study is relevant

in actuarial theory, insurance risk, financial risk, etc. In these fields, those distributions that tend to assign higher probabilities to larger values are specially important, they are known as heavy-tailed. The weight of a tail is a property that can be interpreted as a relative concept (the $F$ distribution has a heavier tail than another $G$) or as an absolute concept (if $F$ verifies a certain property then $F$ is classified as heavy-tailed).

In this paper, we propose a new functional measure that let us compare the left and the right tail of a distribution $F$. This functional is based on the convex transform order defined by Van Zwet in [10] and it measures which tail is heavier than the other.

The convex transform order is closely related to skewness and shape of the tail distribution. It arises from the need to state when a non-negative distribution $G$ is more skewed to the right than another non-negative distribution $F$. Given two non-negative distributions $F$ and $G$ we say that $F$ is smaller than $G$ in the convex transform order, written $F \leq_c G$ if, and only if $G^{-1}F(x)$ is a convex function in its domain. This means that the $G$ distribution is obtained throughout "stretching" the $F$ distribution and thereby there exist a change of shape between both distributions. This change of shape involves in how the probability is distributed, in this case, the $G$ distribution displace more probability than the $F$ distribution to the right therefore it is accepted that the $G$ distribution is heavier than the $F$ distribution.

The main idea is to define a skewness function for probability distributions from the comparative study of its tails, interpreting them as non-negative variables. Before starting, we need to define the left tail distribution and the right tail distribution associated to a distribution $F$ from a quantile $F^{-1}(u)$. From here to forward we will denote $F^{-1}(u) = x_u$.

## 2 Definitions and properties of the functional measure of skewness

**Definition 1.** *Let $X$ be a random variable which follows a distribution $F$, and let $u$ be a number in $(0, \frac{1}{2})$, then we define the left tail distribution from the quantile $x_u$ of $X$ as*

$$(X - x_u)^- = \begin{cases} x_u - X & X \leq x_u \\ 0 & \textit{other case} \end{cases}$$

*and the right tail distribution from the quantile $x_{1-u}$ of $X$ as the variable*

$$(X - x_{1-u})^+ = \begin{cases} X - x_{1-u} & X \geq x_{1-u} \\ 0 & \textit{other case} \end{cases}$$

Let $X$ be a random variable with probability distribution $F$ (see Figure 1), and let $L = (X - x_u)^-$ and $R = (X - x_u)^+$ be its left and right tails from the quantile $x_u$ and $x_{1-u}$, respectively. Both variables are non-negative (see Figure 2) and their cumulative distribution function take the value $1 - u$ when these variables are null.

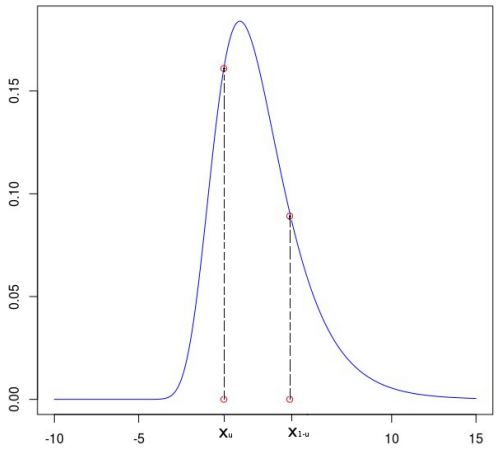



**Fig. 1.** Probability density function of $X \sim F$

We propose to compare $L$ and $R$ using the convex transform order. However, the convex transform order is a hard condition to be verified, thereby we will use a condition which is implied by the convex transform order. This condition appears in a natural way from a equivalent condition of the convex transform order. We denote $F_L$ and $F_R$ as the probability distribution function of $L$ and $R$, respectively. If $F_L^{-1}(F_R(x)) \in C^1(\mathbb{R}^+)$, then

$$L \leq_c R \Leftrightarrow (l_v - l_p) f_L(l_v) \geq (r_v - r_p) f_R(r_v) \quad \forall p, v \in (0, 1).$$

The above characterization implies the following condition,

$$\int_{1-u}^{1} (l_v - l_{1-u}) f_L(l_v) dv \geq \int_{1-u}^{1} (r_v - r_{1-u}) f_R(r_v) dv.$$

$$\Leftrightarrow \int_0^u (x_u - x_t) f_X(x_t) dt - \int_{1-u}^{1} (x_v - x_{1-u}) f_X(x_v) dv \geq 0.$$

We are now in condition to define a new functional measure of skewness.

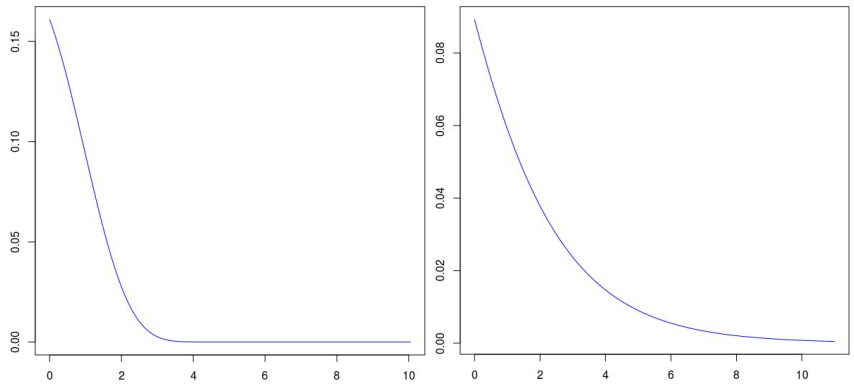



**Fig. 2.** Plots of probability density function of $(X - x_{1-u})^-$ and $(X - x_{1-u})^-$, respectively

**Definition 2.** *Let $X$ be a random variable and $F$ its cumulative distribution function, then*

$$S_X(u) = \varphi_X^-(u) - \varphi_X^+(1-u), \quad \forall u \in (0, \frac{1}{2}).$$

*where* $\varphi_X^-(u) = \int_0^u (x_u - x_v) f_X(x_v) dv$ *and* $\varphi_X^+(1-u) = \int_{1-u}^1 (x_v - x_{1-u}) f_X(x_v) dv$ *are called the left and right skewness measures, respectively.*

**Definition 3.** *We define the $\mathbb{H}$ set as the set of all continuous random variables $X$, with probability density function $f_X$, which verify that:*

1. $\mathbb{E}[f_X(X)] = \int_{-\infty}^{+\infty} f_X^2(x) dx < +\infty.$
2. $\mathbb{E}[X f_X(X)] = \int_{-\infty}^{+\infty} x f_X^2(x) dx < +\infty.$

It has been generally accepted that any measure $\gamma$ of skewness should satisfy the followings conditions:

1. $\gamma(F) = \gamma(aF + b)$ for all $a > 0$ and all b.
2. $\gamma(F) = -\gamma(-F)$.
3. If $F \leq_c G$, then $\gamma(F) \leq \gamma(G)$. Here, $\leq_c$ denotes the convex transform order.

**Proposition 1.** *Let $X \in \mathbb{H}$ be a random variable, the functional measure of skewness $S_X$ verifies the three previous conditions.*

**Proposition 2.** *Let $X \in \mathbb{H}$ be a random variable and $f_X$ its probability density function, then*

*If $f_X$ is a decreasing function in its domain* $\Longrightarrow S_X(u) \geq 0 \ \ \forall u \in (0, \frac{1}{2})$.
*If $f_X$ is an increasing function in its domain* $\Longrightarrow S_X(u) \leq 0 \ \ \forall u \in (0, \frac{1}{2})$.
*If $f_X$ is a symmetric function* $\Longrightarrow S_X(u) = 0 \ \ \forall u \in (0, \frac{1}{2})$.

**Corollary 1.** *Let $X, Y \in \mathbb{H}$ be two random variable, then*

$$\left.\begin{array}{c} \text{If } X \text{ is symmetric} \\ X \leq_c Y \end{array}\right\} \Rightarrow S_Y(u) \geq 0 \quad \forall u \in (0, \frac{1}{2}).$$

The new skewness measure compares both tails of a distribution from any quantile $x_u$ and its symmetric $x_{1-u}$. Thereby, it let us detect an interesting property called symmetry in tails, see Figure (3). Also the following result shows that the functional measure of skewness is bounded from a certain value $u_0 \in (0, 1)$ for unimodal distributions.

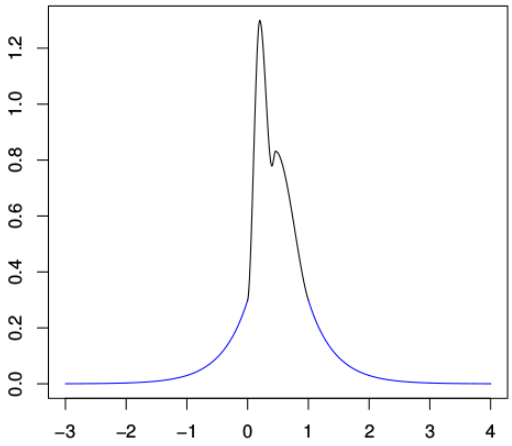



**Fig. 3.** Probability density function of a tail symmetric distribution

**Proposition 3.** *Let $X \in \mathbb{H}$ be a random variable with $F$ an strictly increasing function. If $X$ is an unimodal distribution then there exist $u_0 \in (0, \frac{1}{2})$ such that*

$$-\frac{u^2}{2} \leq S_X(u) \leq \frac{u^2}{2} \quad \forall u \leq u_0.$$

## References

1. Arnold, B. C. and Groeneveld, R. A. Measuring Skewness with Respect to the Mode. The American Statistician, 49, 34–38 (1995)
2. Benjamini, Y. and Krieger, A.M. Concepts and Measures for Skewness with Data-Analytic Implications. The Canadian Journal of Statistics, La Revue Canadienne de Statistique, 24, 131–140 (1996)
3. Boshnakov, G.N. Some measures for asymmetry of distributions. Statistics & Probability Letters, 77, 1111 - 1116 (2007)
4. Charlier, C.V.L. ber das Fehlergesetz. Arkiv fir Mathematik, Astronomi och Fysik, 2.8. (1920)
5. David, F.N. and Johnson, N.L. Some Tests of Significance with Ordered Variables, J.R. Statist. Soc. Ser. B, 18,1–20 (1956)
6. Doksum, K.A. Measures of Location and Asymmetry, Scand. J. Statist., 2, 11–22 (1975)
7. Groeneveld, R. A. and Meeden, G. Measuring Skewness and Kurtosis, Journal of the Royal Statistical Society. Series D (The Statistician), 33, 391– 399 (1984)
8. MacGillivray, H.L. Skewness and Asymmetry: Measures and Orderings, Ann. Statist., 14, 994–1011 (1986)
9. Pearson, K. Contributions to the Mathematical Theory of Evolutions, II: Skew Variations in Homogeneous Material. Transactions of the Royal Philosophical Society, Ser. A, 186, 343–414 (1985)
10. Van Zwet, W.R. Convex Transformations of Random Variables, in Mathematical Centre Tracts 7, Mathematisch Centrum, Amsterdam (1964)
11. Yule, G.W. An Introduction to the Teory of Statistics (2nd ed.), London: Griffin (1912)

# An LLVM-based Approach to Generate Energy Aware Code by means of MOEAs

Sébastien Varrette[1], Bernabé Dorronsorro[2] and Pascal Bouvry[1]

[1] Computer Science and Communication (CSC) Research Unit
University of Luxembourg, 6, rue Richard Coudenhove-Kalergi
L-1359 Luxembourg, Luxembourg
[2] University of Cadiz
Avenida de la Universidad, 10,
11519 Puerto Real, Cadiz, Spain

**Abstract.** Moderating the energy consumption and building eco-friendly computing infrastructure is of major concerns in the implementation of High Performance Computing (HPC) system, especially when a world-wide effort target the production of an Exaflop machine by 2020 within a power envelop of 20 MW. Tracking energy savings can be done at various levels and in this paper, we investigate the automatic generation of energy aware software with the ambition to keep the same level of efficiency, testability, scalability and security.
To this end, the Evo-LLVM framework is proposed. Based on the modular LLVM Compiler Infrastructure and exploiting various evolutionary heuristics, our scheme is designed to optimize for a given input source code (written in C) the sequence of LLVM transformations that should be applied to the source code to improve its energy efficiency without degrading its other performance attributes (execution time, parallel or distributed scalability). Measuring this capacity is based on the combination of several metrics optimized simultaneously with Multi-Objective Evolutionary Algorithms (MOEAs). In this position paper, the NSGA-II algorithm is implemented within the Evo-LLVM yet the analysis of more advanced heuristics is in progress. In all cases, the experimental validation of the framework over a pedagogical code sample reveal a drastic improvement of the energy consumed during the execution while maintaining (or even improving) the average execution time.

**Keywords:** Performance evaluation, Energy-efficiency, HPC, Evolutionary Algorithm, Fault-Tolerance Result-Checking

## 1 Introduction

Energy management has become a key challenge in the area of computing systems today. For large scale systems, such as data centers, energy efficiency has proven to be the key for reducting all kind of costs related to capital, operational expenses and environmental impact. Power drainage of a system is closely related to the type and characteristics of workload that the device is running. These

characteristics refer to the way the workload utilizes different resources and components of the system, such as CPU, memory, disc etc. Modern system design now includes components that support energy management at various level, for instance through a dynamic scaling of the power (or frequency) allocated to its usage (DVFS for the CPU etc.) and/or an integrated way to handle idle state for a more or less long period of time. In this paper, we take advantage of these techniques and the corresponding sensors embedded within the Linux kernel to estimate the average power consumption induced by the execution of a given process. Combined with other metrics quantifying the inherent performance of the execution, it is thus possible to design a Multi-Objective Evolutionary Algorithm (MOEA) system able to evolve a given source code (called the *reference* source in the sequel) to produce an set of energy-aware versions able to compete from the pure execution time point of view with the initial performance of the reference source. This idea led to the design of the Evo-LLVM framework presented in this paper.

This article is organized as follows: section 2 details the background of this work and reviews related works. Then, the Evo-LLVM framework is presented in the section 3. Implementation details of the proposed framework are provided in the section 3. The validation of the approach on a concrete benchmarking code is expounded in the section 4 which details and discusses the experimental results obtained. Finally, the section 5 concludes the paper and provides some future directions and perspectives opened by this study.

## 2 Context & Motivations

Since the advent of high-level programming language, research in the compilation domain have always seek to automate and find novel optimization techniques to produce a compiled code that improve the running time. In this context, many previous studies identified a large number of transformations that could be applied to the different section of a source code to generate different and hopefully improved version of the compiled executable. A reference summary of these transformations, their effects and their respective application context is described in [1]. Determining the optimal sequence of transformations to apply to a given source code that would minimize the execution time over a given computing system is proven to be an NP-complet [9] problem. It follows that all modern compilers such as GCC (the the GNU Compiler Collection) or LLVM rely on static heuristics involving a subset of transformations applied in an order that grant, in general, good results n general while ensuring a bounded compilation time [10]. Because of all these factors, the optimization operated by compilers hardly produces "optimal" output in any sense, and may even impede performances in some cases. It follows a considerable optimization work so as to try a set of transformation potentially valuable. This time-consuming process is generally performed by hand and requires expert engineering skills. The current state-of-the-art tries to address optimization problem from a transverse way, *i.e.* by means of automatic analysis schemes generally based on Evolution-

ary heuristics. For instance, a genetic approach is done in [2] to optimize the size of the output binaries. Also, the Acovea [8] framework (Analysis of Compiler Options via Evolutionary Algorithm) for `gcc` or Cole [7] investigates in an automatic manner the best combination of compiler options leading to the fastest executable program from a given source code. Complementary, recent advances over a new kind of software development environment inspired from Search Based Software Engineering (SBSE) [6] led to the definition of GISMOE challenge (Genetic Improvement of Software for Multiple Objective Exploration) [5]. The general idea is that it is possible to combine the recent advances in software test data generation, genetic programming and multi objective optimization to build a development environment capable of producing a Pareto program surface that would help the software designer to navigate between different version of the same program (typically the execution time, the memory usage and the energy efficiency). The work proposed in this paper definitively offers the basic building block able to propose a concrete answer to this challenge.

### 2.1 The LLVM Compiler Infrastructure

The LLVM compiler infrastructure project (formerly Low Level Virtual Machine) is a compiler infrastructure designed to be a set of reusable libraries with well-defined interfaces. It is written in C++ and is designed for compile-time, link-time, run-time, and "idle-time" optimization of programs written in arbitrary programming languages. LLVM was originally written to be a replacement for the existing code generator in the GCC stack and many of the GCC front ends have been modified to work with it. Widespread interest in LLVM has led to a number of efforts to develop entirely new front ends for a variety of languages. The one that has received the most attention is Clang, a new compiler supporting C, Objective-C and C++ supported by Apple. The core of LLVM is the intermediate representation (IR), a low-level programming language similar to assembly. IR is a strongly typed RISC instruction set which abstracts away details of the target.

In this article, we propose to exploit the flexibility offered by LLVM to manipulate the IR modelization of a given source code to check the opportunity of applying a sequence of supported transformations and evaluating the impact on the energy efficiency of the produced executable. The choice of the transformation to apply and their order shall be governed by an evolutionary heuristic. The validation of the approach shall be performed on a relevant set of benchmark applications.

### 2.2 Evolutionary Algorithms (EAs)

EA is a class of solving techniques based on the Darwinian theory of evolution [3] which involves the search of a *population* $X_t$ of solutions. Members of the population are feasible solutions and called *individuals*. Each iteration of an EA involves a competitive selection that weeds out poor solutions through the *evaluation* of a fitness value that indicates the quality of the individual as a solution

to the problem. The evolutionary process involves at each generation a set of stochastic operators that are applied on the individuals, typically recombination (or cross-over) and mutation.

## 3 The Evo-LLVM Compiler Framework

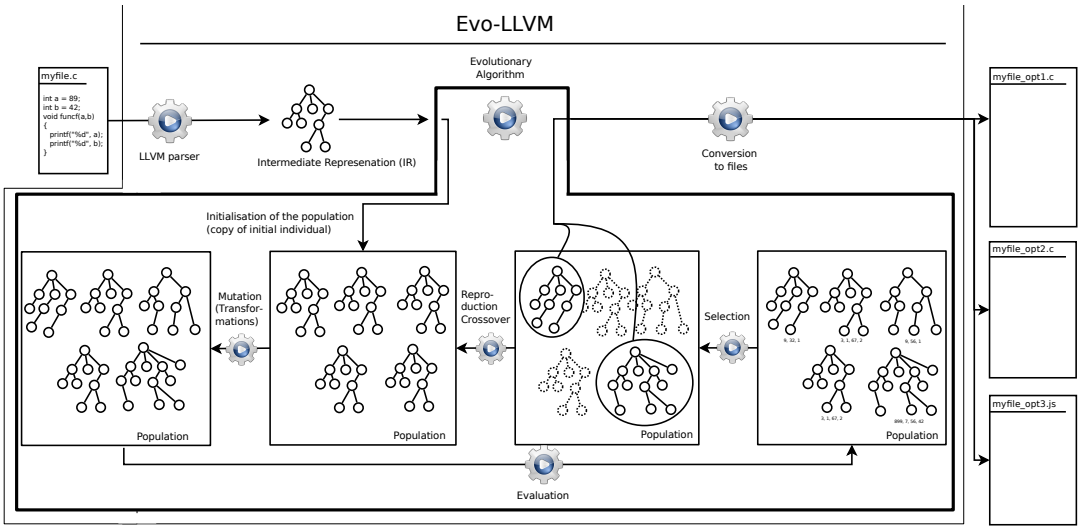



**Fig. 1.** Overview of Evo-LLVM framework, describing the full process of the generation of new representations of the code.

This section briefly review the Evo-LLVM framework as a natural extension of the Shadobf framework proposed in [**?**]. The general code optimization process operated by Evo-LLVM is illustrated in the figure 1: from the initial program $\mathcal{P}$ to be analysed (`myfile.c` in the figure), a reference individual (in the EA sense) $I_{\text{ref}}$ is generated that represents $\mathcal{P}$. Then, a complete population of $n$ individuals is generated by randomly applying a mutation (*i.e.* an LLVM transformation) on the reference individual $I_{\text{ref}}$. The MOEA process (NSGA-II in the current state of the implementation) then intervene to explore the search space induced by the different objectives seek for the produced binaries, evaluate the population and apply the genetic operators (mutation and cross-over). This permits to exhibit at each generation non-dominated Pareto solutions, each of them representing a set of derived (and hopefully more performant for all considered metrics) versions of the program $\mathcal{P}$ that propose a good trade-off between each objectives *i.e.* metrics.

The key characteristics of the Evo-LLVM framework are as follows:

– The C code is parsed using LLVM to produced the intermediate representation of the program (IR).

– LLVM has 54 built-in transformations. These range from tail call elimination (a method to optimize some recursive functions) to loop unwinding (reducing loop overhead). The order these transforms are applied can matter: for example, dead instruction elimination might not find an unused instruction in an unoptimized program, but after a few passes of other transforms, some instructions may be superfluous. This is important to keep in mind when designing the evolutionary algorithm, specifically when deciding on the crossover methods. Some of these may split up two transformations that only work well in tandem. In all cases, the LLVM transforms are randomly applied within Evo-LLVM individuals during the evolutionary operators;
– Throughout NSGA-II [4] (one of the reference selection algorithm for MOEAs considered in our initial implementation), *individuals* are selected by taking into account the non-domination criteria and the distance from one to the others to guarantee a good diversity as well as the leading *individuals* of the population. The concept of dominance is the following (in the case of minimization): an *individual* $I$ with the objectives values $f_{obj}(I)$ is said to be dominated by $J$ if

$$\forall obj \in objectives, f_{obj}(J) < f_{obj}(I)$$

  An *individual* is said to be non-dominated if it is not dominated by any other *individuals* in the population. All the non dominated solutions of a population are the approximated Pareto front of the problem. NSGA-II is selecting all the non-dominated solutions of the population, and if the size of the new population is lower than the maximum size, NSGA-II is selecting again all the non-dominated solutions of the old population but this time excluding the already selected *Individuals*.
– The performance metrics permits to evaluate each individuals.
– The IR model might be helpful to compute additional static metrics (sequential work, number of instructions).

## 4 Validation and Experimental Results

We have validated our approach a simple pedagogical example *i.e.* a quicksort algorithm. It was chosen as it involves many sections in the code that are worth optimizing: memory allocation, iterations, recursion and branching, all intertwined. However, the program is independent from the chosen algorithm, a few changes in the configuration make it possible to run the optimization on any algorithm.

For the sake of simplicity, we show here on short runs involving the simultaneous evolution of the power consumption along with the execution time when the number of benchmarks per individual is set to 100. A brief overview of the generated Pareto front is proposed.

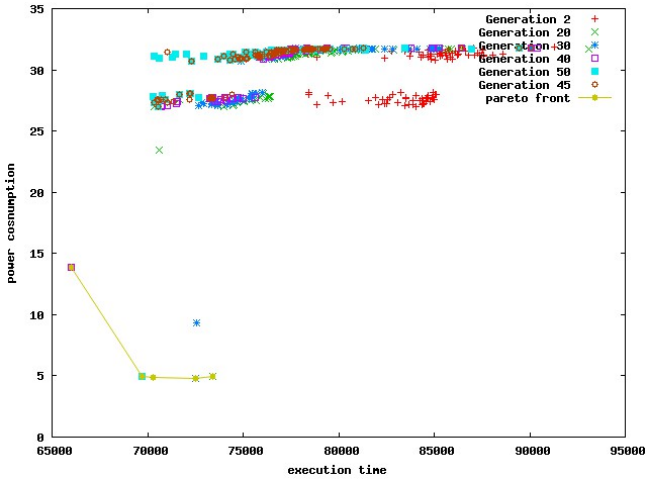



**Fig. 2.** Set of the 2D Pareto fronts approximation for the quicksort program using NSGA-II and Evo-LLVM

The table 1 shows the characteristics of a set of individuals selected after application of Evo-LLVM. In practice, we selected the individual which is the closest to the median values of every individuals which are in the Pareto front, allowing to have a good trade off between all the objectives. In all cases we see that the selected individuals upon successive generations demonstrate an interesting improvement in terms of power consumption while not degrading the execution time.

| Generation | Power | Execution time |
|---|---|---|
| $2^{nd}_{reference}$ | 31.0262 | 70554.0 |
| $20^{th}_{best}$ | 9.3593 | 72573.0 |
| $45^{th}_{best}$ | 4.9078 | 70266.0 |
| $50^{th}_{best}$ | 4.9183 | 69694.0 |

**Table 1.** Energy and execution time comparison for generation 20,45,50 for selected individuals after Evo-LLVM run.

## 5 Conclusion

The main objective of this work was to proceed to the automatic generation of energy aware software while maintaining the same level of efficiency, testability and scalability. As it is well-known, power drainage of a system is not a static property that depends solely on hardware characteristics. Using energy aware software will lead to significant reduction to the overall energy consumption. The benefits will reach not only large scale computing systems or data canters

but also home users that aim to a more efficient energy management. To this end, the Evo-LLVM framework is proposed. Based on the modular LLVM Compiler Infrastructure and exploiting various evolutionary heuristics, our scheme is designed to optimize for a given input source code (written in C) the sequence of LLVM transformations that should be applied to the source code to improve its energy efficiency without degrading its other performance attributes (execution time, parallel or distributed scalability). Measuring this capacity is based on the combination of several metrics optimized simultaneously with Multi-Objective Evolutionary Algorithms (MOEAs). In this position paper, the NSGA-II algorithm is implemented within the Evo-LLVM yet the analysis of more advanced heuristics is in progress. Experimental results on a simple pedagogical program demonstrated an 84.20% improvement on the average consumed energy while not degrading the execution time.

**Acknowledgments:** The experiments presented in this paper were carried out using the HPC facility of the University of Luxembourg.

## References

1. D. F. Bacon, S. L. Graham, and O. J. Sharp. Compiler transformations for high-performance computing. *ACM Comput. Surv.*, 26(4), 1994.
2. K. D. Cooper, P. J. Schielke, and D. Subramanian. Optimizing for reduced code space using genetic algorithms. In *Workshop on Languages, Compilers, and Tools for Embedded Systems*, pages 1–9, 1999.
3. C. Darwin. *On the Origin of Species by Means of Natural Selection*. Londres, 1859.
4. K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii. pages 849–858. Springer, 2000.
5. M. Harman, W. B. Langdon, Y. Jia, D. R. White, A. Arcuri, and J. A. Clark. The gismoe challenge: Constructing the pareto program surface using genetic programming to find better programs (keynote paper). In *Proceedings of the 27th IEEE/ACM International Conference on Automated Software Engineering*, ASE 2012, pages 1–14, New York, NY, USA, 2012. ACM.
6. M. Harman, P. McMinn, J. de Souza, and S. Yoo. Search based software engineering: Techniques, taxonomy, tutorial. In B. Meyer and M. Nordio, editors, *Empirical Software Engineering and Verification*, volume 7007 of *Lecture Notes in Computer Science*, pages 1–59. Springer Berlin Heidelberg, 2012.
7. K. Hoste and L. Eeckhout. Cole: compiler optimization level exploration. In *CGO*, pages 165–174, 2008.
8. S. Ladd. Acovea: Using Natural Selection to Investigate Software Complexities. [Online] see `www.coyotegulch.com/products/acovea/`, 2007.
9. A. Nisbet. GAPS: A Compiler Framework for Genetic Algorithm (GA) Optimised Parallelisation. In *HPCN Europe*, pages 987–989, 1998.
10. R. Stallman et al. *Using GCC: The GNU Compiler Collection Reference Manual*. FSF, 2005.

# Detecting User Influence in Twitter: PageRank vs Katz, a case study*

Hugo Rosa[1], Joao P. Carvalho[1,2], Ramon Astudillo[1], and Fernando Batista[1,3]

[1] INESC-ID
[2] Instituto Superior Técnico, Universidade de Lisboa
[3] ISCTE-IUL - Instituto Universitário de Lisboa
{hugo.rosa,joao.carvalho,ramon.astudillo,fmmb}@inesc-id.pt

**Abstract.** Microblogs, such as Twitter, have become an important socio-political analysis tool. One of the most important tasks in such analysis is the detection of relevant actors within a given topic through data mining, i.e., identifying who are the most influential participants discussing the topic. Even if there is no gold standard for such task, the adequacy of graph based centrality tools such as PageRank and Katz is well documented. In this paper, we present a case study based on a "London Riots" Twitter database, where we show that Katz is not as adequate for the task of important actors detection since it fails to detect what we refer to as "indirect gloating", the situation where an actor capitalizes on other actors referring to him.

**Keywords:** Page Rank, Katz, User Influence, Twitter, Data Mining

## 1 Introduction

Nowadays, there are 288 million active users on Twitter and more than 500 million tweets are produced per day [16]. The impact of Twitter on the Arab Spring [5] and how it beat the all news media to the announcement of Michael Jackson's death [14], are just a few examples of Twitter's role in society. When big events occur, it is common for users to post about it in such fashion, that it becomes a trending topic, all the while being unaware from where it stemmed or who made it relevant. The question we wish to answer is: "Which users were important in disseminating and discussing a given topic?".

Determining user relevance is vital to help determine trend setters [15]. The user's relevance must take into account not only global metrics that include the user's level of activity within the social network, but also his impact in a given topic [17]. Empirically speaking, an influential person can be described as someone with the ability to change the opinion of many, in order to reflect his own.

* This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) under project PTDC/IVC-ESCT/4919/2012 and funds with reference UID/CEC/50021/2013.

While [12] supports this statement, claiming that "a minority of users, called influentials, excel in persuading others", more modern approaches [4] seem to emphasize the importance of interpersonal relationships amongst ordinary users, reinforcing that people make choices based on the opinions of their peers. In [2], three measures of influence were taken into account: "in-degree is the number of people who follow a user; re-tweets mean the number of times others forward a user's tweet; and mentions mean the number of times others mention a user's name.". It concluded that while in-degree measure is useful to identify users who get a lot of attention, it "is not related to other important notions of influence such as engaging audience". Instead "it is more influential to have an active audience who re-tweets or mentions the user". In [7], the conclusion was made that within Twitter, "news outlets, regardless of follower count, influence large amounts of followers to republish their content to other users", while "celebrities with higher follower totals foster more conversation than provide retweetable content". The authors in [11] created a framework named "InfluenceTracker", that rates the impact of a Twitter account taking into consideration an Influence Metric, based on the ratio between the number of followers of a user and the users it follows, and the amount of recent activity of a given account. Much like [2], it also shows that "that the number of followers a user has, is not sufficient to guarantee the maximum diffusion of information (...) because, these followers should not only be active Twitter users, but also have impact on the network".

With the previous definitions of influence in mind, we propose a graph representation of user's influence based on "mentions". Whenever a user is mentioned in a tweet's text, using the @*user* tag, a link is made from the creator of the tweet, to the mentioned user, regardless of it being a retweet or a conversation. For example, the tweet "*Do you think we can we get out of this financial crisis,* @*userB?*", from @*userA*, creates the link: @*userA* $\longrightarrow$ @*userB*.

## 2 Network Analysis Algorithms

In graph theory and network analysis, the concept of centrality refers to the identification of the most important vertices's within a graph, i.e., most important users. We therefore define a graph $G(V, E)$ where V is the set of users and E is the set of directed links between them. Arguably the most well known centrality algorithm is PageRank [8]. It is one of Google's methods to its search engine and uses web pages as nodes, while back-links form the edges of the graph. It is defined by Equation 1 as $PR(v_i)$ of a page $v_i$.

$$PR_{vi} = \frac{1-d}{N} + d \sum_{v_j \in M(v_i)} \frac{PR(v_j)}{L(v_j)} \tag{1}$$

In Equation 1, $v_j$ is the sum ranges over all pages that has a link to $v_i$, $L(v_j)$ is the number of outgoing links from $v_j$, $N$ is the number of documents/nodes in the collection and $d$ is the damping factor. The PageRank is considered to be a random walk model, because the weight of a page $v_i$ is "the probability that

a random walker (which continues to follow arbitrary links to move from page to page) will be at $v_i$ at any given time. The damping factor corresponds to the probability of the random walk to jump to an arbitrary page, rather than to follow a link, on the Web. It is required to reduce the effects on the PageRank computation of loops and dangling links in the Web." [10]. The true value that Google uses for damping factor is unknown, but it has become common to use $d = 0.85$ in the literature. A lower value of $d$ implies that the graph's structure is less respected, therefore making the "walker" more random and less strict.

Another well known method is the Katz algorithm [6]. It is a generalization of a back-link counting method where the weight of each node is "determined by the number of directed paths that ends in the page, where the influence of longer paths is attenuated by a decay factor" and "the length of a path is defined to be the number of edges it contains" [10]. It is defined by Equation 2 "where $N(v_i, k)$ is the number of paths of length $k$ that starts at any page and ends at $v_i$ and $\alpha$ is the decay factor. Solutions for all the pages are guaranteed to exist as long as $\alpha$ is smaller than $\lambda > 1$, where $1/\lambda$ is the maximum in-degree of any page" [10].

$$I_{vi} = \sum_{k=0}^{\infty} [\alpha^k N(v_i, k)] \tag{2}$$

## 3 Experiments and Results

In order to test the network analysis methods presented above, a database from the London Riots in 2011 [3] was used. The Guardian Newspaper made public a list of tweets from 200 influential twitter users, which contains 17795 riot related tweets and an overall dataset of 1132938 tweets. Using a Topic Detection algorithm [1], we obtained an additional 25757 unhastagged tweets about the London Riots. It consists of a Twitter Topic Fuzzy Fingerprint algorithm [13] that provides a weighted rank of keywords for each topic in order to identify a smaller subset of tweets within scope. The sum of posting and mentioned users is 13765 (vertices) and it has 19993 different user mentions (edges), achieving a network connectivity ratio of $\frac{edges}{vertices} = 1.46$.

The remainder of this section presents the results of each algorithm's ranking for most influential users. An empirical study of the users is made, in order to ascertain their degree of influence. The graphs and ranking were calculated using *Graph-Tool* [9].

Table 1 shows how both network analysis algorithms behave with our graph representation, while highlighting the changes in rank between them, as shown by the arrows in the last column. Figure 1 provides a visual tool to the graph, as provided by PageRank. There is a relation between the number of mentions and the ranking in both algorithms, since these users are some of the most mentioned users in our dataset.

When comparing PageRank with Katz, several differences arise, but the top two users are agreed upon: i) @guardian, Twitter account of the world famous

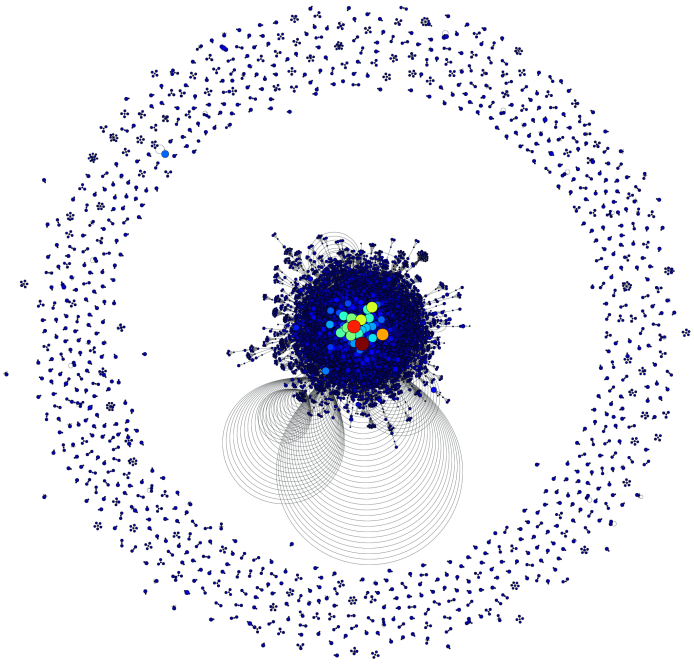


**Fig. 1.** User influence Page Rank Graph - larger circles indicate larger user influence.

**Table 1.** Most influential users according to Page Rank, and comparison with Katz.

| User | Mentions | | PageRank | | Katz | | |
|---|---|---|---|---|---|---|---|
| | # | rank | score | rank | score | rank | |
| @guardian | 160 | 2 | 0.0002854 | 1 | 0.022157 | 2 | |
| @skynewsbreak | 178 | 1 | 0.0002512 | 2 | 0.023479 | 1 | |
| @gmpolice | 122 | 4 | 0.0002128 | 3 | 0.019009 | 4 | |
| @riotcleanup | 107 | 6 | 0.0001767 | 4 | 0.017992 | 6 | ↘ |
| @prodnose | 67 | 14 | 0.0001761 | 5 | 0.014022 | 15 | ↘↘↘ |
| @metpoliceuk | 116 | 5 | 0.0001494 | 6 | 0.018709 | 5 | |
| @marcreeves | 69 | 11 | 0.0001476 | 7 | 0.014195 | 12 | ↘↘ |
| @piersmorgan | 78 | 8 | 0.0001465 | 8 | 0.014959 | 9 | |
| @scdsoundsystem | 69 | 12 | 0.0001442 | 9 | 0.014190 | 13 | ↘↘ |
| @subedited | 70 | 10 | 0.0001337 | 10 | 0.014278 | 11 | |
| @youtube | 48 | 20 | 0.0001257 | 11 | 0.012424 | 20 | ↘↘↘ |
| @bbcnews | 94 | 7 | 0.0001256 | 12 | 0.016426 | 8 | ↗↗ |
| @mattkmoore | 62 | 15 | 0.0001237 | 13 | 0.013614 | 16 | ↘ |
| ... | | | | | | | |
| @paullewis | 129 | 3 | 0.0000954 | 20 | 0.019602 | 3 | ↗↗↗↗ |
| ... | | | | | | | |
| @juliangbell | 61 | 16 | 0.0000275 | 188 | 0.0166597 | 7 | ↗↗↗↗↗↗↗ |

newspaper "The Guardian"; ii) @skynewsbreak, Twitter account of the news team at Sky News TV channel. This outcome agrees with [7] previous statement, that, "news outlets, regardless of follower count, influence large amounts of followers to republish their content to other users". Other users seem to fit the profile, namely @gmpoliceq and @bbcnews. Most of the other users are either political figures, political commentators or jornalists (@marcreeves, @piersmorgan, and @mattkmoore).

However, Katz's third and seventh top ranked users, are not in PageRank's top users. These are two very different cases: i) @paullewis, ranked 3rd by Katz shows up at 20th according to PageRank; ii) @juliangbell, ranked 7th by Katz shows up at 188th according to PageRank. The reason behind @paullewis high placement in the Katz rank is the number of mentions. As said previously, Katz is a generalization of a back-link counting method, which means the more back-links/mentions a user has, the higher it will be on the ranking. This user has 129 mentions, but PageRank penalizes it, because it is mentioned by least important users, which means a less sum weight is being transfered to it in the iterative process. This logic also applies to user @bbcnews. Additionally, @paullewis is also an active mentioning user, having mentioned other users a total of 14 tweets, while @skynewsbreak and @guardian have mentioned none. As a consequence, Paul Lewis transfers its influence across the network while the others simply harvest it. There are several users that drop in ranking from PageRank to Katz for the very same reason. Users such as @prodnose, @marcreeves and @youtube do not have enough mentions for Katz to rank them higher. User @juliangbell, despite mentioned often (61 times), is down on the PageRank because of indirect gloating, i.e., he retweets tweets that are mentioning himself: *"@LabourLocalGov #Ealing Riot Mtg: @juliangbell speech http://t.co/3BNW0q6"* was posted by @juliangbell himself. The user is posting somebody else's re-tweet of one of his tweets. As a consequence a link/edge was created from @juliangbell to @LabourLocalGov, but also from @juliangbell to himself, since his username is mentioned in his own tweet. Julian Bell is a political figure, making it acceptable that he would have a role in discussing the London Riots, but the self congratulatory behavior of re-tweeting other people's mentions of himself, is contradictory with the idea of disseminating the topic across the network. While Katz is not able to detect this effect, PageRank automatically corrects it. Contrary to what is mentioned in previous works, it is our comprehension that Katz is not adequate to detect a user's importance in social media such as Twitter.

## 4 Conclusions and Future Work

With this study, we have shown that in the context of user influence in Twitter, PageRank and Katz are not equal in performance, thus disproving previous claims. PageRank has proved a more robust solution to identify influential users in discussing and spreading a given relevant topic, specially when considering how it deals with indirect gloating, an item Katz fails to penalize.

## References


1. 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)
2. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.P.: Measuring user influence in twitter: The million follower fallacy. In: in ICWSM '10: Proceedings of international AAAI Conference on Weblogs and Social (2010)
3. Crockett, K, S.R.: Twitter riot dataset (tw-short) (2011)
4. Domingos, P., Richardson, M.: Mining the network value of customers. In: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 57–66. KDD '01, ACM, New York, NY, USA (2001), http://doi.acm.org/10.1145/502512.502525
5. Huang, C.: Facebook and twitter key to arab spring uprisings: report. http://www.thenational.ae/news/uae-news/facebook-and-twitter-key-to-arab-spring-uprisings-report (June 2011), accessed: 2014-05-02
6. Katz, L.: A new status index derived from sociometric analysis. Psychometrika 18(1), 39–43 (March 1953), http://ideas.repec.org/a/spr/psycho/v18y1953i1p39-43.html
7. Leavitt, A., Burchard, E., Fisher, D., Gilbert, S.: The influentials: New approaches for analyzing influence on twitter (2009)
8. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web (1999)
9. Peixoto, T.: https://about.twitter.com/company
10. Phuoc, N.Q., Kim, S.R., Lee, H.K., Kim, H.: Pagerank vs. katz status index, a theoretical approach. In: Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology. pp. 1276–1279. ICCIT '09, IEEE Computer Society, Washington, DC, USA (2009), http://dx.doi.org/10.1109/ICCIT.2009.272
11. Razis, G., Anagnostopoulos, I.: Influencetracker: Rating the impact of a twitter account. CoRR (2014), http://arxiv.org/abs/1404.5239
12. Rogers, E.M.: Diffusion of innovations (1962)
13. Rosa, H., Batista, F., Carvalho, J.P.: Twitter topic fuzzy fingerprints. In: WCCI2014, FUZZ-IEEE, 2014 IEEE World Congress on Computational Intelligence, International Conference on Fuzzy Systems. pp. 776–783. IEEE Xplorer, Beijing, China (July 2014)
14. Sankaranarayanan, J., Samet, H., Teitler, B.E., Lieberman, M.D., Sperling, J.: Twitterstand: News in tweets. In: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. pp. 42–51. GIS '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1653771.1653781
15. Tinati, R., Carr, L., Hall, W., Bentwood, J.: Identifying communicator roles in twitter. In: Proceedings of the 21st International Conference Companion on World Wide Web. pp. 1161–1168. WWW '12 Companion, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2187980.2188256
16. Twitter: https://about.twitter.com/company
17. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twitterrank: Finding topic-sensitive influential twitterers. In: Proceedings of the Third ACM International Conference on Web Search and Data Mining. pp. 261–270. WSDM '10, ACM, New York, NY, USA (2010), http://doi.acm.org/10.1145/1718487.1718520

# Simplification of Inference Problems Based on High Dimensional Vectors by Wavelet Transformation and Fuzzy Rule Interpolation

Ferenc Lilik[1], Szilvia Nagy[1], and László T. Kóczy[1,2]

[1] Széchenyi István University
H-9026 Győr, Hungary,
lilikf@sze.hu
[2] Budapest University of Technology and Economics
H-1117 Budapest, Hungary

**Abstract.** A new approach for inference based on treating sampled functions is presented. Sampled functions can be transformed into only a few points by wavelet analysis, thus the complete function is represented by these several discrete points. The finiteness of the teaching samples and the resulting sparse rule bases can be handled by fuzzy rule interpolation methods, like, e.g., KH interpolation. Using SHDSL transmission performance prediction as an example, the simplification of inference problems based on large, sampled vectors by wavelet transformation and fuzzy rule interpolation applied on these vectors are introduced in this paper.

**Keywords:** Fuzzy inference, performance prediction, fuzzy rule interpolation, wavelet analysis, telecommunications access networks

## 1 Introduction

Due to the great number of input values, making inference on phenomena which can be described by large-sized vectors are difficult and expensive. In order to construct efficient inference systems, simplification of the input space is needed. This simplification makes the process of the inference easier, however, it unavoidably rises the system's level of uncertainty and inaccuracy. During our previous research on performance prediction of physical links of telecommunications access networks, we had to encounter such problems in two ways. Horizontally, making decisions by the observation of only a part of the physical reality resulted in sparse fuzzy rule bases. Vertically, drastically lowering the number of the measured frequency dependent input values caused an inaccuracy in the final results.

In Section 2 the primary technical problem underlying the research on performance prediction is briefly reviewed. In Section 3 wavelet transformation and fuzzy rule interpolation as the algorithmic techniques applied in a combined way for handling the problems of simplification are outlined, and in Section 4 we present the test results of the new approach based on these techniques.

## 2 SHDSL Performance Prediction of physical links in telecommunications access networks

In our previous work [1, 2] we have laid the foundations of a Mamdani-type fuzzy inference method for pre-qualification of telecommunication access network links based on measured insertion loss and noise values of the given lines. We applied fuzzy rule bases of two types, the one was generated from the measured data's statistical properties using triangular sets, the other was generated by an evolutionary algorithm using trapezoidal sets [3]. Examples of the resulting rules can be seen in Fig. 1.

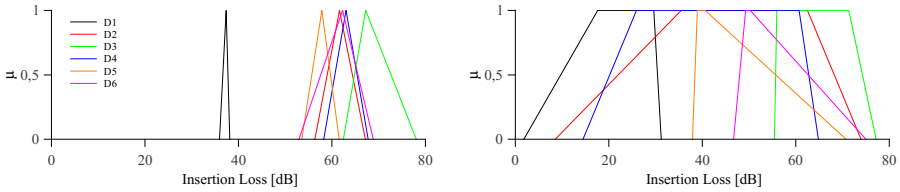



**Fig. 1.** Examples of rule antecedents from our previous predicting methods [2].

The above two rule bases were tested by the measurements of more than 60 wire pairs in operating access networks and there were no relevant differences between their respective results. In most of the cases, where all measured values belonged to insertion loss areas covered by antecedent sets, the predictions were successful. Only 13 lines out of 65 could be evaluated, and the predictions were correct in case of 12 lines form these 13.

## 3 Methods for handling the vertical and horizontal sparseness

The reason for the insufficient performance of the pre-qualification method is the two-dimensional sparseness of the inference system.

Vertical sparseness of the rule bases was derived from the partial usage of the possible input data. It was needed in order to decrease the dimensionality of the applied fuzzy inference system, however, a lot amount of information of the measured insertion loss functions was wasted. Finding a method which keeps the simplicity of the fuzzy system and the information of the used insertion loss functions was needed. As wavelet transformation is efficient in reducing the size of any continuous or discrete functions down to a required level, it seemed to be successfully applicable in the problem.

Horizontal sparseness of the fuzzy system, namely the sparseness of the rule bases, can be handled by the techniques of fuzzy rule interpolation. Stabilized KH interpolation fits continuous and mathematically stable functions to all $\alpha$-cuts of the membership functions in the rules, which can tackle the observations

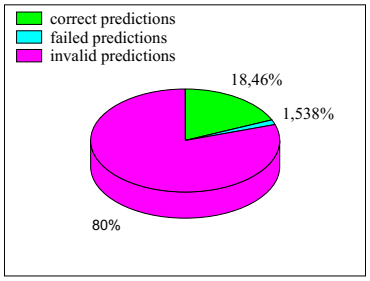



**Fig. 2.** Success rate of the rule bases.

in the gaps and out of the domains of the rules too (in this way performing also extrapolation).

Basics of wavelet transformation and stabilized KH interpolation are briefly overviewed in the followings.

### 3.1 On wavelet analysis

In data processing in general wavelet theory [5] has proved to be a very useful tool. The largest part of the methods use wavelets is the image compression [6] and data analysis, but it can also be used for solving differential equations [7].

Wavelet transform and of a function provides data about the function's fine-scale and rough-scale behavior. Wavelet analysis can be carried out by a series of filter pairs. There is a high-pass and a low-pass filter in all of the pairs, the high-pass ones (after a downsampling) giving the wavelet (detail) components and the low-pass ones being transformed further, as it can be seen in Fig. 3.

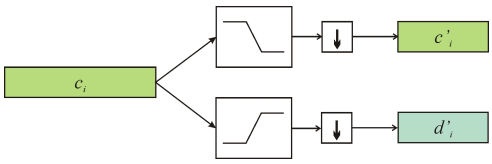



**Fig. 3.** One filter pair of the discrete wavelet transform. After the high pass and low pass convolutional filters and the downsamplings the transformed vectors $c'_i$ and $d'_i$ arise, their size is about half of the size of the original $c_i$.

In data analysis – also in our case – the starting point is a sampled function and the end result is the lowest resolution level low pass vector and the high pass vectors. Our starting vector is a series of insertion loss values measured at consecutive frequency points, and the resulting vectors give information about the large-scale behavior of the insertion loss vs. frequency function. In the following considerations Haar's [8] and Daubechies's [5] wavelet and scaling function sets

are used with 2 and 4 nonzero filter coefficients, respectively. Transformations of the starting sampled insertion loss functions were carried out until only 5 vector elements remained.

### 3.2 Stabilized KH rule interpolation

In case of sparse rule bases, KH interpolation [9, 10] is a mathematically stable and widely applicable fuzzy rule interpolation method. Its improved version is the stabilized KH interpolation. In our work we used this improved technique in order to eliminate the problems originating from the sparseness of the rule bases.

The method is based on the distances between the examination vector and the antecedent sets of the rule base. The closures of the $\alpha$-cuts of the interpolated resolution are given in [11].

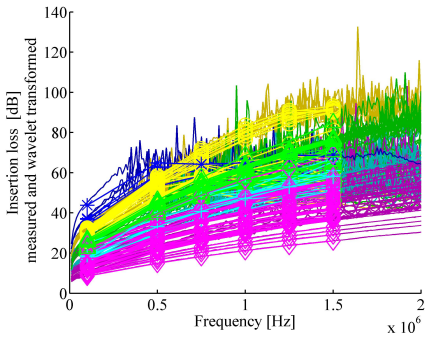



**Fig. 4.** Insertion loss values and the corresponding wavelet transforms. Different performance classes are indicated by different colors.

## 4 A new prediction method based on the combination of wavelet transformation and stabilized KH rule interpolation

In order to avoid the problems reviewed in Section 1, the techniques of Section 3.1 and 3.2 were used.

First, the wavelet transformed version of the insertion loss values used in rule base construction were calculated. Daubechies-2 (Haar) [8] and Daubechies-4 wavelets were used and the transformations were performed down to 5 points resolution. Fig. 4 shows the original and the Haar wavelet transformed insertion loss values as an example. As a matter of course, wavelet transformation results in discrete values, however, to make the corresponding points visible, they are graphically linked in Fig. 4.

The rule base using Daubechies wavelets did not give better results than the ones without any wavelet transformation, moreover, several additional errors occured. On the contrary, in case of Haar wavelets, accurate results arose for each of the 13 lines that produced valid results and one further line could be assessed, too, as it can be seen in the left hand side of Fig. 5.

In order to evaluate those lines that were previously not to be assessed, the new, Haar wavelets-type rule base was applied together with the stabilized KH rule interpolation. The 65 test lines were re-processed, thus the predictions became feasible in case of all lines. The predictions for the 13 wire pairs which were correctly evaluated previously remained valid, moreover, results of the predictions of 33 from the other 52 were correct, and 19 acceptable (in this contribution, results with a deviation of -1 from the correct values are considered as acceptable ones, all the others as incorrect) and there were no incorrect results.

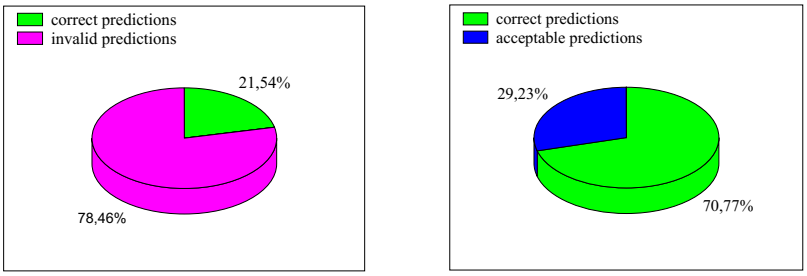



**Fig. 5.** Efficiency of the Haar wavelets based rule base alone (left) and supplemented with the stabilized KH rule interpolation (right).

The simplified "algorithm" of the construction of the predicting system is summarized as follows.

- Collection of insertion loss and bit rate data of wire pairs.
- Dividing the whole bit rate domain into groups (the more the number of the measured lines, the finer is the possible resolution) and clustering the measured values into these groups.
- Generation of several discrete values (6 in this case, however, other resolutions are examined by our ongoing investigations) from measured insertion loss functions by wavelet transformation (Haar wavelets are now recommended, though investigating other types of wavelets with other resolution levels are being in progress).
- Construction of fuzzy rule bases by clustered and wavelet transformed values.
- Wavelet transformation of the insertion loss function of the wire pair to be predicted.
- Prediction making by stabilized KH interpolation (can be made even if the input values can be found within the areas covered by antecedent fuzzy sets).

## 5 Conclusions

A novel performance prediction method based on interpolated fuzzy inference for telecommunications transmission lines and wavelet transformation of the values of the physical parameters influencing the performance was presented. The combination of the fuzzy rule interpolation and wavelet transformation was proposed in this paper in the first time. Wavelet transform was used for generating a coarse-grained view of the measured data, whereas the interpolation is applied for treating the sparseness of the rule bases. The method performed very well for the model system of the SHDSL connections, 52 predictions from 65 test cases were correct, and the other 19 were acceptable.

## References


1. F. Lilik, J. Botzheim: Fuzzy based Prequalification Methods for EoSHDSL Technology, Acta Technica Jauriensis, Series Intelligentia Computatorica, Vol. 4. No. 1., pp.135-145, 2011.
2. 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
3. K. Balázs, L. T. Kóczy: Constructing Dense, Sparse and Hierarchical Fuzzy Systems by Applying Evolutionary Optimization Techniques, Applied and Computational Mathematics, Vol. 11, No. 1, 2012, pp. 81-101
4. 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.
5. I. Daubechies: Ten Lectures on Wavelets, CBMS-NSF regional conference series in applied mathematics 61, SIAM, Philadelphia, 1992.
6. Ch. Christopoulos, A. Skodras, T. Ebrahimi: The JPEG2000 Still Image Coding System: An Overview, IEEE Trans. Consumer Electronics, Vol. 46, pp. 1103-1127 (2000). doi:10.1109/30.920468
7. Sz. Nagy, J. Pipek: On an economic prediction of the finer resolution level wavelet coefficients in electron structure calculations Phys. Chem. Chem. Phys., 2015, Accepted Manuscript doi:10.1039/C5CP01214G
8. A. Haar, Zur theorie der orthogonalen funktionen systeme, Math. Ann., Vol. 69, pp. 331-371 (1910).
9. L. T. Kóczy, K. Hirota: Approximate reasoning by linear rule interpolation and general approximation, International Journal of Approximate Reasoning, vol. 9, 1993, pp. 197-225, doi: 10.1016/0888-613X(93)90010-B
10. L. T. Kóczy, K. Hirota: Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases, Information Sciences, vol. 71, 1993, pp. 169-201, doi: 10.1016/0020-0255(93)90070-3
11. D. Tikk, I. Joó, L. T. Kóczy, P. Várlaki, B. Moser and T. D. Gedeon: Stability of interpolative fuzzy KH-controllers, Fuzzy Sets and Systems, vol. 125, 2002, pp. 105-119, doi: 10.1016/S0165-0114(00)00104-4

# State reduction methods for Fuzzy Cognitive Map to model regional waste management systems

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

[1]Department of Information Technology, Széchenyi István University, Győr, Hungary
miklos.hatwagner@sze.hu
[2]Department of Environmental Engineering, Széchenyi István University, Győr, Hungary
buruzs@sze.hu
[3]Department of Automation, 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.** The authors have investigated the sustainability of Integrated Waste Management Systems (IWMS). These systems were modeled by Fuzzy Cognitive Maps (FCM), which are known as adequate fuzzy-neural network type models for multi-component systems with a stable state. The FCM model was designed of thirty-three factors to describe the real world processes of IWMS in as much detailed and as much accurately as possible. Although, this detailed model meets the requirements of accuracy, the presentation and explanation of such a complex model is difficult due to its size.

While there is a general consensus in the literature about a very much simplified model of IWMSs, detailed investigation lead to the assumption that a much more complex model with considerably more factors (components) would more adequately simulate the rather complex real life behavior of the IWMS.

As the starting point we used the thirty-three component model based on the consensus of a workshop of experts coming from all areas of the IWMS (operation, regulation, management, etc.) and the set goal was to find the most accurate real model that could be obtained by analyzing and properly reducing this – very likely too much detailed, or atomized – model.

In this paper, a new state reduction approach is presented. The practical aspects of the results gained by these methods are evaluated.

**Keywords:** fuzzy cognitive maps, integrated waste management system, state reduction methods.

## 1 Introduction

During the previous investigations [1] the method of FCM was applied to model regional waste management systems which are determined by six factors. As a validation of the simulation results [2] data were collected based on the relevant literature to set up a time series. This time series served as an input to the Bacterial Evolutionary Algorithm (BEA) which generated an optimal connection matrix producing the possibly

adfa, p. 1, 2011.

most similar time series to the original one obtained from the literature. Despite the expectations, the six-factor FCM model proved to be rather inaccurate in practice [3], and this is why a refined, more detailed model, containing thirty-three factors was developed [6] with the support of a group of experts.
After the thorough examination of both the basic (6 factors) and the detailed (33 components) models it became apparent that the two models were very different, in their respective complexities and concepts. For this reason, we assumed that an intermediate model containing less than thirty-three but more than six factors would be presumably able to describe the mechanism and action of a real IWMS with sufficient accuracy.

Table 1 introduces the main factors of the basic model and the thirty-three sub-factors of the detailed model.

Table 1. The identified sub-factors of the main factors and the concept IDs (CID) of them

| 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 | | | |

On the basis of the detailed model, we might be able to support the strategic decision making process of the stakeholder in order to ensure the long-term sustainability of IWMS.

## 2 The investigated state reduction method

The idea of state reduction is similar to clustering but it can also be considered as a special, strongly generalized version of the state reduction technique of sequential circuits or finite state machines (see e.g. [10]). The methods construct clusters of factors and these clusters can be used later as factors of the reduced model. The members of clusters are selected based on their 'similarity'. Two factors are considered similar, if

their ‘distance’ is low. The distance of them have to be measured by an appropriate metric, and the applied metric differentiate the methods from each other. Different metrics and different distance values can also result in different clusters and reduced models. Several different metrics have been proposed e.g. in [7], but the basic idea of them is always the same and all versions use only the connection matrix and a threshold value of the maximum allowed distance. In this paper only one of the best solutions will be presented. Two factors $C_i$ and $C_j$ are considered similar, if the connections originating from them and leading to a third factor $C_k$, and also in the opposite direction have almost the same weights for all $C_k$, where $1 \leq i < j \leq n$, $n$ is the number of factors, and $i \neq k, j \neq k, 1 \leq k \leq n$. At first, all clusters contain only one of the factors, but as soon as a similar factor is found, they will be merged. During the next steps, the similarity of all current cluster members must be measured to the next candidate factor.

The main properties of similarity are the following: 1) all factor is similar to itself (reflexivity), 2) if factor $C_i$ is similar to $C_j$, then $C_j$ is similar to $C_i$ as well (symmetry). 3) But if $C_i$ is similar to $C_j$ and $C_j$ is similar to $C_k$, then $C_i$ is not always similar to $C_k$ (non-transitive). It means that the state reduction method is a fuzzy tolerance relation [5, 8].

After the presentation of the basic idea, the precise description of the methods are given. First, the clusters are disjoint sets of factors, and each one of them contains only one factor. $K_i = \{C_i\}$ for every $i = 1 \ldots n$ where $K_i$ is the $i$th cluster, $C_i$ is the $i$th factor (factors are often called ‘concepts’ in the FCM theory) and $n$ is the number of factors in the model (thirty-three in the IWMS model). In the next steps all clusters will be appended by other factors, if possible. The ‘distance’ between the next cluster candidate and all current cluster members are measured by the chosen metric.

The presented metric calculates the normalized, squared Euclidean distance of the connections starting from factors $C_i$ and $C_j$ to $C_k$, where $i \neq j \neq k, i, j, k = 1 \ldots n$. If this difference is below the threshold value (ε), the current factor is added to the cluster. The determined distance is normalized to [0, 1]. The applied metric is described more precisely by the following C-style pseudo-code (see **Fig. 1** and **Fig. 2**).

```
function isNear(i, j, eps, c)
  sum = 0;  // i, j = factor indexes, eps = ε
  for(k=0; k<n; k++) // n = number of factors
    if(k!=i and k!=j and !elementOf(k, c))
      dout = w(i, k)-w(j, k) // w(i, k) = w_ik
      sum = sum + dout * dout
      din = w(k, i)-w(k, j)
      sum = sum + din * din
  if(sum / ((n-2)*8) < eps)
    return true
  else
    return false
```

**Fig. 1.** Calculation of the distance of two concepts

```
function buildCluster(initialFactor, eps)
  c = {initialFactor}
  for(i=0; i<n; i++)
    if(i != initialFactor)
      member = true
      while(member and hasNextElement(c))
        j = nextElement(c)
        member = isNear(j, i, eps, c)
        if(member)
          c = c + {i}
  return c

function buildAllClusters(eps)
  clusters = {}
  for(i=0; i<n; i++)
    k = buildCluster(i, eps)
    if(!isElementOf(k, clusters))
      clusters = clusters + {k}
  return clusters
```

**Fig. 2.** Pseudo-code of the state reduction algorithm, Part 1

The state reduction is started by the `buildAllClusters` function (see **Fig. 2**). It requests the creation of each clusters by consecutive calling of the `buildCluster` function. The latter function sometimes produces the same clusters in different order, but `buildAllClusters` keeps only one of them. The distance of two factors are measured by `isNearA` or some of the other functions implementing different metrics.

When all the clusters are defined, the weights of the interconnections are defined by function `getWeight`. This function accepts two cluster arguments and provides the weight between these clusters. The return value is the average weight of connections among the factors of the specified clusters. The weight of self-loops are always zero according to the original FCM definition (see **Fig. 3**).

```
function getWeight(a, b)
  count = 0
  sum = 0
  while(hasNextElement(a))
    i = nextElement(a)
    while(hasNextElement(b))
      j = nextElement(b)
      if(i != j)
        count = count + 1
        sum = sum + w(i, j)
  if(count == 0)
    return 0
  else
```

```
return sum/count
```

**Fig. 3.** Pseudo-code of the state reduction algorithm, Part 2

The value of ε must be in the [0, 1] interval and must be chosen appropriately in every single case, because it plays an important role in the reduction process. Too low values do not lead to models containing significantly fewer factors (clusters), thus they are not useful. But if the value of ε is too high, the model will be oversimplified and will not have the required accuracy. For example in an extreme case, when ε is 1, the whole model collapses and only one big sole cluster remains. The knowledge and experience of experts are needed to specify a meaningful ε value. In order to show the connection between ε and the number of factors (clusters) some interesting value pairs are collected in Table 2.

Table 2. The number of factors in the reduced connection matrix

| ε | No. of factors |
|---|---|
| 0.015 | 29 |
| 0.023 | 24 |
| 0.024 | 23 |
| 0.037 | 22 |
| 0.070 | 21 |
| 0.080 | 17 |

It must be emphasized here that all model reduction activities necessarily cause information loss, and the accuracy of simplified models are always lower. In the suggested method, there are three root causes of information loss:

1. The connections among concepts inside the same cluster are neglected. The representation of these connections would result in self loops which is not allowed according to Kosko's original idea.
2. Every causal relation needs one time step before their effect can be observed. Long pathes of concepts and interconnections may cause long delays. If more or less elements of these pathes become inside the same cluster, these delays partially disappear.
3. Since the model reduction method is based only on the connection matrix, the *getWeight* function cannot take into account the effect of possibly different (source) concept values on the connected (destination) concepts, because all such connections are represented by a single connection in the reduced model.

Despite all these possible problems, the proposed state reduction method performed well in several practical problems [9]. Furthermore, an exhaustive investigation is under fulfillment in order to analyze the behavior of the proposed method on statistical basis.

## 3 Results

In the next, the authors give an overview about and shortly analyze the driving forces and impact of IWMS upon the results of the state reduction method (Table 3, Table 4).

Table 3. An example of clusters as a result of state reduction (ε = 0.06)

| Cluster ID | Reduced concepts |
|---|---|
| Q1 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C4.6 + C6.4 |
| Q2 | C1.1 + C1.3 + C2.1 + C2.3 + C2.5 + C4.2 + C4.3 + C4.4 + C4.5 + C4.7 |
| Q3 | C1.2 + C1.5 + C2.2 + C2.3 + C3.3 + C3.5 + C3.6 + C5.1 |
| Q4 | C2.4 + C2.5 + C2.6 + C4.4 + C4.5 + C4.6 + C5.1 |
| Q5 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C2.5 + C3.3 + C3.4 + C3.5 + C4.3 + C4.5 + C6.4 |
| Q6 | C1.1 + C2.1 + C2.4 + C2.5 + C2.6 + C4.2 + C4.4 + C4.5 + C4.6 + C5.1 |
| Q7 | C1.1 + C1.2 + C1.3 + C1.4 + C2.5 + C3.1 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C4.6 + C6.1 + C6.2 + C6.3 + C6.4 |
| Q8 | C1.1 + C1.2 + C1.4 + C1.5 + C2.5 + C3.2 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C4.6 + C6.1 + C6.2 + C6.3 + C6.4 |
| Q9 | C1.1 + C1.2 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C5.3 + C6.4 |
| Q10 | C1.1 + C1.2 + C1.4 + C1.5 + C2.5 + C4.1 + C4.2 + C4.4 + C4.5 + C4.6 + C4.7 + C5.2 + C5.3 + C6.4 |
| Q11 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.5 + C4.2 + C4.4 + C4.5 + C4.6 + C4.7 + C5.3 + C6.4 |
| Q12 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C2.5 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C4.6 + C6.4 |
| Q13 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C4.7 + C6.4 |
| Q14 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.5 + C4.3 + C4.4 + C4.5 + C4.6 + C5.1 |
| Q15 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.5 + C3.3 + C3.5 + C4.4 + C4.5 + C4.6 + C5.2 + C5.3 + C6.4 |
| Q16 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.5 + C5.3 + C6.4 |
| Q17 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.5 + C5.4 + C6.4 |
| Q18 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C2.5 + C3.3 + C3.4 + C3.5 + C4.3 + C4.5 + C6.1 + C6.4 |
| Q19 | C1.1 + C1.2 + C1.3 + C1.4 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C6.1 + C6.2 + C6.3 + C6.4 + C6.5 |
| Q20 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C6.3 + C6.4 |
| Q21 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C6.4 |
| Q22 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C3.3 + C3.4 + C3.5 + C4.3 + C4.4 + C4.5 + C6.3 + C6.4 + C6.5 |

Table 4. An example of clusters as a result of state reduction (ε = 0.08)

| Cluster ID | Reduced concepts |
|---|---|
| Q1 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.1 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C4.7 + C5.2 + C5.3 |
| Q2 | C1.1 + C1.2 + C1.3 + C1.5 + C2.1 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.1 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C4.7 + C5.2 + C5.3 |
| Q3 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C4.7 + C5.1 + C5.3 + C5.4 |
| Q4 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C5.2 + C5.3 + C5.4 + C6.2 + C6.5 |
| Q5 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.5 + C4.1 + C4.3 + C4.4 + C4.5 + C4.6 + C5.2 + C5.4 |

| | |
|---|---|
| Q6 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C4.1 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C4.7 + C5.2 + C5.3 + C6.4 |
| Q7 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.4 + C3.5 + C3.6 + C4.1 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C5.2 + C5.3 + C6.4 + C6.5 |
| Q8 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C4.6 + C5.3 + C5.4 |
| Q9 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.3 + C3.4 + C3.5 + C3.6 + C4.1 + C4.2 + C4.3 + C4.4 + C4.5 + C4.6 + C5.1 + C5.2 + C5.3 |
| Q10 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.7 + C5.1 + C5.2 + C5.3 + C5.4 + C6.5 |
| Q11 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.2 + C2.3 + C2.4 + C2.5 + C2.6 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C4.6 + C4.7 + C5.3 + C6.5 |
| Q12 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.7 + C5.1 + C5.3 + C5.4 |
| Q13 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.7 + C5.1 + C5.4 + C6.1 |
| Q14 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.7 + C5.3 + C5.4 + C6.1 + C6.2 + C6.4 |
| Q15 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.3 + C2.5 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C4.6 + C4.7 + C5.1 + C6.2 + C6.3 |
| Q16 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.3 + C2.5 + C2.6 + C3.1 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C5.1 + C5.3 + C5.4 + C6.4 |
| Q17 | C1.1 + C1.2 + C1.3 + C1.4 + C1.5 + C2.1 + C2.2 + C2.3 + C2.5 + C2.6 + C3.2 + C3.3 + C3.4 + C3.5 + C3.6 + C4.3 + C4.4 + C4.5 + C5.1 + C6.5 |

As it can be seen from the above tables, there are several overlaps among the clusters. According to this, some of the factors are presented many times in the new models. The role of factors are described in Table 5 and Table 6.

Table 5. Appearance of factors in the clusters – the possible framework of a new IWMS ($\varepsilon = 0.06$)

| CID | Role of factors | Nomination of factor | CID | Role of factors | Nomination of factor |
|---|---|---|---|---|---|
| **C1.1** | 21 | Engineering knowledge | **C1.4** | 19 | Technical requirements in the EU and national policy |
| **C2.5** | 21 | Wildlife (social acceptance) | **C3.3** | 19 | Culture |
| **C4.3** | 21 | Political and power factors | **C3.5** | 19 | Financing |
| **C1.5** | 20 | Technical level of equipment | **C1.2** | 18 | Technological system and its coherence |
| **C1.3** | 19 | Local geographical and infrastructural conditions | **C2.3** | 18 | Geographical factor |

Table 6. Appearance of factors in the clusters – the possible framework of a new IWMS ($\varepsilon = 0.08$)

| CID | Role of factors | Nomination of factor | CID | Role of factors | Nomination of factor |
|---|---|---|---|---|---|
| **C1.1** | **17** | Engineering knowledge | **C2.5** | **17** | Wildlife (social acceptance) |
| **C1.2** | **17** | Technological system and its coherence | **C3.5** | **17** | Financing |
| **C1.3** | **17** | Local geographical and infrastructural conditions | **C4.3** | **17** | Political and power factors |

| | | | | | |
|---|---|---|---|---|---|
| **C1.5** | **17** | Technical level of equipment | **C4.4** | **17** | Education |
| **C2.3** | **17** | Geographical factor | **C1.4** | **16** | Technical requirements in the EU and national policy |

Integrated modeling requires not only the consideration of the technical and economic system elements, but also social, environmental, legal and institutional factors, furthermore their sub-factors. In these cases, to deal with the situations where data at hand are often insufficient for an entire quantitative analysis and the uncertainty is high, a series of non-quantifiable elements become important.

IWMSs are organized along spatial and temporal scales. While modeling the system, it leads to the appearance of the connections and interaction between its factors and sub-factors. The factors of these systems are connected via material, energy, money and information flows and form a complex phenomenon through legal regulation.

During studying the results of the state reduction methods, it could be recognized that the most commonly occurring factor in the new models (Table 5, Table 6) is the 'Engineering knowledge'. Based on international experience it might be still surprising that the most important element in the system is 'Engineering knowledge'. The factor has a determining role in the design and operation of the systems according to technical-economic-environmental considerations.

This combination of factors tells about the relationship of the sub-factors as parts of the systems and highlights the question 'What is important in this system?'.

## 4 Conclusions

A new state reduction approach was introduced to make the otherwise too complex connection matrix of IWMS model easier to handle and understand. The factors of two simplified connection matrices was presented and evaluated. The presented reduction method will be further investigated in regard to its known shortcomings.

The state reduction approach proved to be good to combine different type of factors and create clusters. It thereby provides a comprehensive and more thoroughly understanding of an IWMS as a technical-economic-social-environmental system.

The conclusions based on the results of state reduction should be viewed together with existing scientific knowledge. In the next period, it is the authors' intention to study further the assumptions, but also be open to insights gained from a systemic approach to deliver a method for decision making on sustainable regional waste management.

## 5 Future research

The authors' purpose is to apply a modified version of the FCM. In the suggested method several connections that existed in the original model were neglected as their source and sink factors were represented by a single cluster of factors. The representation of these internal connections may lead to self-loops, which means non-zero elements in the main diagonal of the connection matrix. Usually this property of the matrix

is not accepted [4], but real life systems query the justification of this theory. The usage of clusters also removes some delays of the original models, which can cause different simulation results or limit cycle behavior. The authors intention is to analyze the effect caused by self-loops and missing delays, then provide advanced model reduction techniques.

**Acknowledgement.** The authors would like to thank to TÁMOP-4.2.2.A-11/1/KONV-2012-0012, TÁMOP-4.1.1.C-12/1/KONV-2012-0017, to the Hungarian Scientific Research Fund (OTKA) K105529 and K108405 for the support of the research.

## References

1. Buruzs, A., Pozna, R. C. and Kóczy, L. T. (2013). Developing Fuzzy Cognitive Maps for Modeling Regional Waste Management Systems. In Proceedings of the Third Int. Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering, Paper 19, Tsompanakis, Y. (ed), Civil-Comp Press, Stirlingshire, UK.
2. Buruzs, A., Hatwágner, M. F., Pozna, R. C. and Kóczy, L. T. (2013). Advanced Learning of Fuzzy Cognitive Maps of Waste Management by Bacterial Algorithm. In Proceedings of IFSA World Congress and NAFIPS Annual Meeting, IEEE, pages 890-895.
3. Buruzs, A., Hatwágner, M. F. and Kóczy, L. T. (2014) Modeling integrated sustainable waste management systems by fuzzy cognitive maps and the system of systems concept, in: Technical Transactions series Automatic Control, accepted for publication, 2014
4. Carvalho, J. P. (2010) On the Semantics and the Use of Fuzzy Cognitive Maps in Social Sciences, WCCI 2010 IEEE World Congress on Computational Intelligence, July, 18-23, 2010 - CCIB, Barcelona, Spain
5. 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.
6. Hatwágner, M. F., Buruzs, A., Földesi, P. and Kóczy, L. T. (2014) Strategic decision support in waste management systems by state reduction in FCM models. ICONIP2014 in prep.
7. Hatwágner, M. F., Kóczy, L. T. (2015) Parameterization and concept optimization of FCM models, submitted to FUZZ-IEEE 2015.
8. Klir, G. J., Folger, T. A. (1987) Fuzzy Sets, Uncertainty and Information, Prentice Hall.
9. Papageorgiou, E.I., Hatwágner, M.F., Buruzs, A., Kóczy, L.T., A new clustering approach for designing reduced Fuzzy Cognitive Map models for decision making: application on Hungarian Waste Management System, submitted to Journal of Applied Soft Computing
10. Paull, 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, pages 356 – 367.

# On the Use of F-transform on the Reduction of Concept Lattices

Petra Hodáková and Nicolás Madrid

University of Ostrava, Centre of Excellence IT4Innovations,
Institute for Research and Applications of Fuzzy Modeling,
30. dubna 22, 701 03 Ostrava 1, Czech Republic
{nicolas.madrid,petra.hodakova}@osu.cz

**Abstract.** In this paper, we show that F-transform can be used to reduce relational databases. Subsequently, we show that the respective concept lattice is reduced significantly as well. Moreover, we present a clarifying example of the procedure.

**Keywords:** F-transform, Fuzzy Sets, Fuzzy Concept Analysis, Knowledge Reduction.

## 1 Introduction

Fuzzy Formal Concept Analysis deals with the processing of imprecise knowledge in information systems [2, 3]. In this theory, the information of a relational databases is represented in terms of a complete lattice where its elements are called concepts. However, despite the information represented by the concept lattice is valuable, the complexity and the size (which increases exponentially with respect to the size of the relational database) makes the use of this theory impractical in many applied tasks. For this reason, recent approaches have dealt with Knowledge Reduction in relational databases to simplify the formal concept analysis of them [1].

On the other hand, F-transforms [4] is a theoretical tool that has shown its effectiveness on representing the information of signals (like temporal series, images, etc.) to a vector of few components. This paper applies F-transforms (based on residauted lattice) to Knowledge Reduction. Specifically, we begin by showing that objects (or attributes) in a relational database can be grouped in a new set of objects (or attributes). Then, we transfer the information of the original database to another where objects are given by the grouping previously mentioned. The transfer of information is given by F-transforms and therefore, there are two possible new relational databases.

This paper has the following structure. In Section 2 we recall briefly the theories of fuzzy property-oriented concept lattices and F-transforms. Then, in Section 3 we describe the reduction of relational tables by means of F-Transforms. Moreover, we illustrate the consequences of the reduction in concept lattices with an example. Finally, in Section 4 we present conclusions and future work.

## 2 Preliminaries

### 2.1 F-transforms on residuated lattices

In this section, we briefly recall the basic definitions and the main principles of F-transforms based on operations of a residuated lattice [4]. Let $(L, \leq, \&, \to)$ be a residuated lattice. A fuzzy partition of a finite set $\mathcal{U}$ is a set of $L$-fuzzy sets on $\mathcal{U}$ $\mathcal{A}_1, \ldots, \mathcal{A}_n$ fulfilling the covering property namely, for all $x \in \mathcal{U}$ there exists $k \in \{1, \ldots, n\}$ such that $\mathcal{A}_k(x) > 0$. The membership functions $\mathcal{A}_k(x)$, $k = 1, \ldots, n$ are called the *basic functions*.

**Definition 1.** *Let $f \colon \mathcal{U} \to L$ be a function and $\mathcal{A}_1, \ldots, \mathcal{A}_n$, with $n \leq |\mathcal{U}|$, be basic functions which form a fuzzy partition of $\mathcal{U}$. We say that the $n$-tuple of real numbers $\mathbf{F}_n^{\uparrow}[f] = [F_1^{\uparrow}, \ldots, F_n^{\uparrow}]$ is the (direct) $F^{\uparrow}$-transform of $f$ w.r.t. $\mathcal{A}_1, \ldots, \mathcal{A}_n$ if*

$$F_k^{\uparrow} = \bigvee_{x \in \mathcal{U}} (\mathcal{A}_k(x) \,\&\, f(x)). \tag{1}$$

*Moreover, we say that the $n$-tuple of real numbers $\mathbf{F}_n^{\downarrow}[f] = [F_1^{\downarrow}, \ldots, F_n^{\downarrow}]$ is the (direct) $F^{\downarrow}$-transform of $f$ w.r.t. $\mathcal{A}_1, \ldots, \mathcal{A}_n$ if*

$$F_k^{\downarrow} = \bigwedge_{x \in \mathcal{U}} (\mathcal{A}_k(x) \to f(x)). \tag{2}$$

The elements $F_1^{\uparrow}, \ldots, F_n^{\uparrow}$ and $F_1^{\downarrow}, \ldots, F_n^{\downarrow}$ are called *components* of the $F^{\uparrow}$-transform and $F^{\downarrow}$-transform, respectively.

The following lemma ([4]) shows that the components of the $F^{\uparrow}$-transform ($F^{\downarrow}$-transform) are lower mean values (upper mean values) of an original function which give least (greatest) element to certain sets.

**Lemma 1.** *Let $f \colon \mathcal{U} \to L$ be a function and $\mathcal{A}_1, \ldots, \mathcal{A}_n$, with $n \leq |\mathcal{U}|$, be basic functions which form a fuzzy partition of $\mathcal{U}$. Then the $k$-th component of the $F^{\uparrow}$-transform is the least element of the set*

$$S_k = \{a \in L |\ \mathcal{A}_k(x) \leq (f(x) \to a) \text{ for all } x \in \mathcal{U}\}$$

*and the $k$-th component of the $F^{\downarrow}$-transform is the greatest element of the set*

$$T_k = \{a \in L |\ \mathcal{A}_k(x) \leq (a \to f(x)) \text{ for all } x \in \mathcal{U}\}$$

*where $k = 1, \ldots, n$.*

### 2.2 Fuzzy property-oriented concept lattices

In this section we recall briefly a simplification of property-oriented concept lattices introduced in [2, 3]. So, because of the lack of space, here we restrict to residuated lattices instead of adjoin triples. The notion of fuzzy property-oriented context is defined below.

**Definition 2.** *Let* $(L, \leq, \&, \rightarrow)$ *be a residuated lattice. A* context *is a tuple* $(A, B, R)$ *such that* $A$ *and* $B$ *are non-empty sets (usually interpreted as attributes and objects, respectively),* $R$ *is an* $L$*-fuzzy relation* $R\colon A \times B \rightarrow L$.

From now on, we fix a context $(A, B, R)$. The mappings ${}^{\uparrow_\Pi}\colon L^B \rightarrow L^A$ and ${}^{\downarrow^N}\colon L^A \rightarrow L^B$ are defined, for $g \in L^B$ and $f \in L^A$ as, $g^{\uparrow_\Pi}$ and $f^{\downarrow^N}$, where

$$g^{\uparrow_\Pi}(a) = \bigvee_{b \in B} R(a,b) \,\&\, g(b)$$
$$f^{\downarrow^N}(b) = \bigwedge_{a \in A} R(a,b) \rightarrow f(a)$$

It is not difficult to prove that $({}^{\uparrow_\Pi}, {}^{\downarrow^N})$ forms an isotone Galois connection (also known as adjunction) and, therefore, ${}^{\uparrow_\Pi\downarrow^N}\colon L^B \rightarrow L^B$ is a closure operator and ${}^{\downarrow^N\uparrow_\Pi}\colon L^A \rightarrow L^A$ is an interior operator. A concept is a pair of mappings $\langle g, f\rangle$, with $g \in L^B, f \in L^A$, such that $g^{\uparrow_\Pi} = f$ and $f^{\downarrow^N} = g$, which will be called *fuzzy property-oriented concept*. In that case, $g$ is called the *extent* and $f$, the *intent* of the concept. The set of all these concepts will be denoted as $\mathcal{F}_{\Pi N}$.

**Definition 3.** *The associated* fuzzy property-oriented concept lattice *to the context* $(A, B, R)$ *is defined as the set*

$$\mathcal{F}_{\Pi N} = \{\langle g, f\rangle \in L^B \times L^A \mid g^{\uparrow_\Pi} = f \text{ and } f^{\downarrow^N} = g\}$$

*in which the ordering is defined by* $\langle g_1, f_1\rangle \preceq \langle g_2, f_2\rangle$ *iff* $g_1 \preceq_2 g_2$ *(or equivalently* $f_1 \preceq_1 f_2$*).*

## 3 Reducing the size of Relational Tables.

Throughout this section we consider a frame $(A, B, R)$ and a residuated lattice $(L, \leq, \&, \rightarrow)$. The idea underlying in the reduct is the creation of two smaller relational tables $R^\uparrow$ and $R^\downarrow$ that keep as much information from $R$ as possible. In order to reduce the size of the table is needed to reduce either the number of attributes or the number of objects. In this paper we focus on objects. In this way, we define a new set of objects $\overline{B}$ that can be considered as a set of fuzzy sets that group objects according to certain attributes in $A$. For instance, consider a relational table where objects are people and attributes are physical features of them. Then, we could group people according to their high and then, to define the following set of "new" objects $\{B_1 = VerySmall, B_2 = QuiteSmall, B_3 = Medium, B_4 = QuiteTall, B_5 = VeryTall\}$. To conclude the reduction, we only need to define the relations $R^\uparrow$ and $R^\downarrow$ between the new set of objects and the original set of attributes. For such a task we consider direct F-transforms. In this framework, each basic function from the chosen fuzzy partition determines a new object and the value assigned to it by the direct F-transform determines the value of the relation.

To define the basic functions (and then also the set of new objects) let us consider firstly, a fuzzy partition of $L$ given by fuzzy sets $\{\mathcal{L}_k \colon k \in \{1, \ldots, n\}\}$ and secondly, a subset of attributes $\overline{A} \subseteq A$. Then the fuzzy partition $\overline{B} = \{B_{k\overline{a}} \colon k \in \{1, \ldots, n\}$ *and* $\overline{a} \in \overline{A}\}$ of $B$ is defined by

$$B_{k\overline{a}}(b) = \mathcal{L}_k(R(b, \overline{a})), \quad b \in B. \tag{3}$$

Note that the fuzzy partition $\overline{B}$ groups original objects in fuzzy sets according to their relation with attributes in $\overline{A}$. Moreover, note the number of basic functions (i.e., the number of new objects) is $k \cdot |\overline{A}|$. So the size of the new set of objects depends on the number of attributes considered to define the partition. Once the fuzzy partition is fixed, we can define the following two $L$-fuzzy relational tables $R^{\uparrow}$ and $R^{\downarrow}$ between $\overline{B} = \{B_{k,\overline{a}} \colon k \in \{1, \ldots, n\}$ *and* $\overline{a} \in \overline{A}\}$ and $A$ as follows:

$$\begin{aligned} R^{\uparrow} &\colon \overline{B} \times A \to L \\ &(B_{k,\overline{a}}, a) \mapsto \bigvee_{b \in B} B_{k\overline{a}}(b) \,\&\, R(a, b) \\ R^{\downarrow} &\colon \overline{B} \times A \to L \\ &(B_{k,\overline{a}}, a) \mapsto \bigwedge_{b \in B} B_{k\overline{a}}(b) \to R(a, b) \end{aligned} \tag{4}$$

Note that original objects are used to define the values of the new ones. Finally, the reduction of the concept lattice given by the original frame $(A, B, R)$ is the pair of concept lattices associated to the frames $(A, \overline{B}, R^{\uparrow})$ and $(A, \overline{B}, R^{\downarrow})$. Below we show how the procedure works in a simple example.

| | *HighPower* | *BigSpace* | *HighConsume* | *Expensive* | *Sport* | *Familiar* |
|---|---|---|---|---|---|---|
| $b_1$ | 1 | 0.2 | 1 | 0.8 | 1 | 0 |
| $b_2$ | 1 | 1 | 0.8 | 1 | 0.6 | 1 |
| $b_3$ | 0.6 | 0.8 | 0.4 | 0.6 | 0.2 | 0.6 |
| $b_4$ | 0.8 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| $b_5$ | 0.6 | 0.4 | 0.2 | 0.6 | 0.2 | 0.2 |
| $b_6$ | 0 | 0.2 | 0 | 0.2 | 0 | 0 |
| $b_7$ | 0.8 | 0.2 | 0.8 | 0.8 | 0.8 | 0 |
| $b_8$ | 1 | 1 | 1 | 1 | 0 | 1 |
| $b_9$ | 0.6 | 1 | 0.4 | 0.6 | 0 | 1 |
| $b_{10}$ | 0.6 | 1 | 0.6 | 0.6 | 0 | 1 |
| $b_{11}$ | 0.6 | 0.6 | 0.4 | 0.4 | 0 | 0.6 |
| $b_{12}$ | 0.2 | 0.4 | 0.4 | 0.2 | 0 | 0.2 |
| $b_{13}$ | 0.8 | 0 | 0.8 | 1 | 0.8 | 0 |

**Fig. 1.** A car relational database.

*Example 1.* Let us consider the $L$-relational table in Figure 1 that relates types of cars (objects) with features (attributes). For the sake of simplicity, let $L$ be the unit interval $[0, 1]$ represented by finite set $L = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$, and the adjoint pair considered for the reduction and the construction of the concept lattice is the one given by the Gödel connectives. Let us consider the partition $\{\mathcal{L}_1, \mathcal{L}_2\}$ of $L$ given by:

| $x$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
|---|---|---|---|---|---|---|
| $\mathcal{L}_1(x)$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |

| $x$ | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
|---|---|---|---|---|---|---|
| $\mathcal{L}_2(x)$ | 1 | 0.8 | 0.6 | 0.4 | 0.2 | 0 |

Now, for the sake of simplicity let us consider just the attribute $Familiar \in A$ to make the reduct. Then, from Equation (3), we have that the partition of the set of objects $B$ with respect to the attribute $Familiar$ and the partition $\{\mathcal{L}_1, \mathcal{L}_2\}$ of $L$ is given by the following two fuzzy sets

| $b$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | $b_6$ | $b_7$ | $b_8$ | $b_9$ | $b_{10}$ | $b_{11}$ | $b_{12}$ | $b_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $B_{1\overline{a}}(b)$ | 0 | 1 | 0.6 | 0.6 | 0.2 | 0 | 0 | 1 | 1 | 1 | 0.6 | 0.2 | 0 |

| $b$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | $b_6$ | $b_7$ | $b_8$ | $b_9$ | $b_{10}$ | $b_{11}$ | $b_{12}$ | $b_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $B_{2\overline{a}}(b)$ | 1 | 0 | 0.4 | 0.4 | 0.8 | 1 | 1 | 0 | 0 | 0 | 0.4 | 0.8 | 1 |

Note that partitions $B_{1\overline{a}}$ and $B_{2\overline{a}}$ above represent the fuzzy sets of cars that are familiar and non familiar, respectively. Thus, it has sense that the two new objects in the new tables are denoted by $FamCars$ and $NonFamCars$. The new relational tables are given by F-transforms (4) as follows

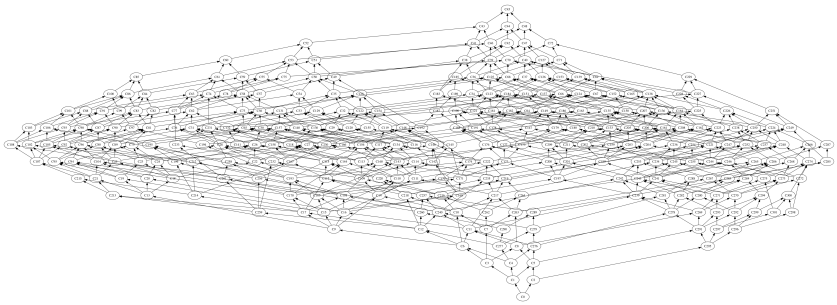
| $R^{\uparrow}$ | $HighPower$ | $BigSpace$ | $HighConsume$ | $Expensive$ | $Sport$ |
|---|---|---|---|---|---|
| $FamCars$ | 1 | 1 | 0.8 | 1 | 0.6 |
| $NonFamCars$ | 1 | 0.4 | 1 | 1 | 1 |

| $R^{\downarrow}$ | $HighPower$ | $BigSpace$ | $HighConsume$ | $Expensive$ | $Sport$ |
|---|---|---|---|---|---|
| $FamCars$ | 0.6 | 1 | 0.4 | 0.6 | 0 |
| $NonFamCars$ | 0 | 0 | 0 | 0 | 0 |

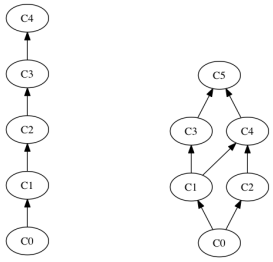
The tables above can be interpreted as follows. Tables $R^{\uparrow}$ and $R^{\downarrow}$ represent the possibility and necessity, respectively, of a familiar car (in some degree) to have a certain attribute. So $R^{\uparrow}(FamCars, a)$ and $R^{\downarrow}(FamCars, a)$ represent an upper and a lower bound, respectively, of the value $R(b, a)$ for any familiar car $b \in B$, i.e., $R^{\uparrow}(FamCars, a) \geq B_{1\overline{a}}(b) \,\&\, R(b, a)$ and $R^{\downarrow}(FamCar, a) \leq B_{1\overline{a}}(b) \rightarrow R(b, a)$.

It is interesting to mention that from the interpretability above, we can infer from the tables $R^{\uparrow}$ and $R^{\downarrow}$ that a familiar car must have a big space because $R^{\uparrow}(FamCars, BigSpace) = R^{\downarrow}(FamCars, BigSpace) = 1$. Moreover, the familiar cars are quite powerful and expensive as well as $R^{\downarrow}(FamCars, HighPower) = R^{\downarrow}(FamCars, Expensive) = 0.6$.

The concept lattice of the original relational table of Example 1 has 302 concepts and it is given by the following Hasse diagram



However, the concept lattices of the tables reduced by our procedure have only 5 and 6 concepts, respectively.




## 4 Conclusion and Future Works

In this paper we have presented the reduction of relational tables aimed to keep as much information from the original table as possible. Our future work is to apply the reduct based on the ordinary F-transforms and measure, determine and/or bound the information which is on one hand lost by the reduction and on the other hand kept by the reduction.

**Acknowledgments.** The research was supported by the European Regional Development Fund by projects (CZ.1.05/1.1.00/02.0070) and (TIN12-39353-C04-04).

Aditionally, we want to thank Eloisa Ramírez-Poussa for her help.

## References

1. M E. Cornejo, J. Medina, E. Ramírez-Poussa. Attribute reduction in multi-adjoint concept lattices. *Information Sciences*, 294: 41-56, 2015.
2. J. Medina. Towards multi-adjoint property-oriented concept lattices. *Lecture Notes in Artificial Intelligence*, 6401:159–166, 2010.
3. J. Medina. Multi-adjoint property-oriented and object-oriented concept lattices. *Information Sciences*, 190:95–106, 2012.
4. I. Perfilieva. Fuzzy transforms: Theory and applications. *Fuzzy Sets and Systems*, 157: 993–1023, 2006.

# Possible Applications of Fuzzy Relational Calculus for Some General Problems of Recommender Systems

Alex Tormási[1], Brigitta Szi[1], Péter Földesi[2], Dávid Zibriczky[3] and László T. Kóczy[1,4]

[1]Department of Information Technology, Széchenyi István University, Győr, Hungary
{szi.brigitta,tormasi,koczy}@sze.hu
[2]Department of Logistics and Forwarding, Széchenyi István University, Győr, Hungary
foldesi@sze.hu
[3]ImpressTV, Budapest, Hungary
david.zibriczky@impresstv.com
[4]Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Budapest, Hungary
koczy@tmit.bme.hu

**Abstract.** The main contribution of this paper is to overview and discusses possible applications of fuzzy relational calculus to solve some issues and challenges of recommender systems. The presented ideas are targeting the most essential aspects of these problems, the knowledge representation and handling.

**Keywords.** fuzzy relations; recommender systems; implicit and explicit feedback; cold start problem; hybrid filtering

## 1 Introduction

Recommender systems are none other, than information filtering algorithms, that help users in the discovery of items in the multitude of choices. Personalized recommendations reduce the time the user spent for looking for relevant items and increase the likelihood of meeting the user's expectations. Recommender systems could be considered as cognitive info communication systems [1], which decrease the cognitive load and increases the mathability [2] of the users, extends the users' ability to filter out and access relevant content.

By the assistance of content discovery the user satisfaction may increase, consequently recommender systems have also impact from the business point of view. Because of this, recommender systems became more and more popular among both the businesses and end-users in the last decade. Recommender systems shouldn't consider the maximization of key performance indicators for the businesses only, but finding the trade off between accuracy, coverage, diversity and serendipity.

This paper is organized as follows: after the Introduction in Section 2 some of the main difficulties of recommender systems are outlined. In Section 3 there is a short overview of fuzzy relational calculus with possible applications in recommender systems. Fuzzy methods in recommender systems are briefly summarized in Section

4. This is followed by some ideas and aspects of fuzzy relational calculus in recommender systems. Summary of the paper and the main phases of future work are outlined in Section 6.

## 2 Issues and Challenges of Recommender Systems

There are several influencing factors that make difficult to fulfil all of the requirements for an effective recommender system. One of the key challenges is the proper interpretation of user activities for user profiling. In practice two types of user interaction is distinguished. We consider an interaction as "explicit feedback", when the user expresses his preference over an item intentionally (e.g. he gives a rating 4 for a movie). Explicit feedbacks have significant information about the preference since it can be quantified in the algorithms. The typical examples of explicit feedbacks are ratings, likes, dislikes and adding contents to favourites. The other type of interactions is the "implicit feedback" that refers to all kind of interactions that cannot be interpreted precisely in terms of preference. For example, a profile view interaction is considered as implicit feedback because it has no explicit meaning about how much the user likes or dislikes it. The typical examples of implicit feedbacks are viewing a profile, buying a product or watching a movie. Based on another approach, the difference between implicit and explicit feedback is that implicit feedbacks are generated before consuming the items while explicit feedbacks are made after the interaction with the item. In practice the collection of explicit data is more difficult, because it requires the intention and some efforts from the user to express its preference over the items. By contrast of that implicit feedbacks are much easier, because it is just a tracking of user browsing on the site. The consequence of these properties is that on one hand explicit data is more meaningful but it is less provided, on the other hand implicit data is less meaningful but it is provided with a higher level of magnitude.

Main properties of implicit and explicit feedback according to [3] and [4] are summarized as follows: It could be easy to collect implicit feedback from user interactions, but it is noisy, difficult to interpret and has a low accuracy. Explicit feedback has lower availability, possibly as a result of the increased cognitive information processing it requires, but has no such a reward, which would directly motivate the users to be involved. It is possible to determine both positive and negative preferences from explicit feedback, but it is dependent from the context [3] and could be noisy [5-7]. Since explicit and implicit feedbacks are representing the same consumer preferences, there must be a relation between them; D. Parra et al. used logistic regression in order to extract explicit feedback from implicit feedback [8].

Another practical problem is that the preference of users is changing over time, the interpretation of an interaction strongly depends on the context (e.g. the time of the day or the device that the user is using). The deeper understanding of the problem depends on the knowledge extracted from both the behaviour and cognitive processes of the users and the processes from industrial or commercial point of view including

the advantages and disadvantages of applied technologies. In order to recommend items for the users, their profiles should be known. Usually it is possible through collecting and processing implicit and explicit feedback, which could be considered as different projections of the same user preferences.

Another challenge is solving the cold-start problem. A recommendation problem is considered as cold-start problem if neither explicit feedbacks nor implicit feedbacks are provided for an item or user. There are two kinds of cold-start problems [9-11], first is when a recommendation should be generated for a new user (with limited or without any previous knowledge of his preferences or patterns) and the other is when a new item appears in the system. To overcome these problems, content-based filtering (CBF) methods were introduced [9-12]. Metadata is essential to enhance the user and item models for better recommendation. At this stage basic data (such as gender, age group) are available for the system, but those could be ambiguous and it does not conclude that the user will have the typical preferences built from the data bank. A similar case is when a new item gets available and only basic information is ready such as categories, product descriptions or tags, which could be still misclassified. Content-based filtering techniques are targeting to use meta data to create more acceptable recommendation, but if the mate data is not well structured, the system will not perform well. The missing metadata might also lead to problems; when an item does not have proper description and/or tags some systems consider it not as missing data, but connects it with a negative property, which means that the missing tags are considered as features that do not apply to the item.

As the user interacts with the system, implicit or explicit feedbacks are collected about his behaviour. Analogously to new items, the users start to consume it and generating feedbacks for that. The increasing number of interactions improves the accuracy of user or item models. However it is generally true that more feedbacks results better models, there is a theoretical saturation point where an additional increment of the number of feedbacks doesn't result significant improvement in accuracy. The period between cold-start and saturation point is called "warm-up" period. It depends on the algorithm whether the user is in "warm-up" period, the key challenge is to reduce the amount of data required by the recommender algorithm.

The user interactions are not only used for personalization, but the extraction of user behavioural patterns, that called collaborative filtering (CF) [13]. Collaborative filtering methods recommend items based on what the similar users consumed. An advantage of collaborative filtering against the content-based filtering is that it is capable to extract behavioural patterns that cannot be explained by metadata. Conventional collaborative filtering can be powerful when a clear separation of user preferences is observed in consumption patterns. Usually it is not the case because the user preferences are usually mixed or fuzzy. The disadvantage of collaborative filtering methods is that they are not capable to solve the cold-start problem and performs.

To combine the advantages of collaborative- and content-based filtering, hybrid filtering was introduced [14]. Hybrid filtering methods are more complex than single collaborative- or content-based filtering methods, but offers better accuracy by solving cold-start problem and extracting consumption patterns at a time. Another

advantage of hybrid filtering is that misclassified metadata can be easier detected (e.g. a movie labelled "action" mainly consumed by "romantic" movie fans), furthermore similarity between different tags can be evaluated. However the conventional hybrid filtering has many advantages, it is still difficult to handle missing information and less meaningful implicit data sets.

The properties of computational intelligence techniques (like fuzzy methods, meta heuristics) enable them to properly handle some of the described problems. There are number of works related to the application fuzzy methods in recommender systems and the aim of this work is to give a brief overview of these and look for some possible new perspectives.

## 3 Fuzzy Relations and Basic Operations

Similarly to crisp and fuzzy sets [15], the fuzzy relations could be interpreted as generalized form of crisp relations, where the connection between items of two or more (discrete or continuous) sets could be expressed by a membership degree [16–17]. If we consider relation $R$ between sets $X_1$, $X_2$, …, $X_n$, then the formal description of the fuzzy relations is as follows:

$$R(X_1,\ldots,X_n) \subseteq X_1 \times \cdots \times X_n, \tag{1}$$

$$R(x_1,\ldots,x_n) = \mu R\langle x_1,\ldots,x_n\rangle. \tag{2}$$

Let $R$ be a fuzzy relation over the $X_1$, …, $X_n$, then $[R \downarrow Y]$ is the projection of the relation on the $Y$ multi sets formally:

$$[R \downarrow Y](\underline{y}) = \max_{\underline{y} \prec \underline{x}} R(\underline{x}). \tag{3}$$

The cylindrical extension could be considered as some sort of inverse operation of the above defined projection operator. It is marked as $[R \uparrow X - Y]$, where $R$ is a fuzzy relation and $X$, $Y$ are multi sets and the values are calculated for each $x$, where $\underline{y} \prec \underline{x}$:

$$[R \uparrow X - Y](\underline{x}) = R(\underline{y}). \tag{4}$$

The cylindrical closure (5) is similar to cylindrical extension, but it uses the intersection of multiple projections.

$$\mathrm{cyl}\{P_i\}(\underline{x}) = \min_{i \in I} [p_i \uparrow X - Y_i](\underline{x}), \tag{5}$$

where $P_i$ is a projection defined over $Y_i$ multi set. $\{P_i | i \in I\}$ is a set of projections of $R$ fuzzy relation defined over set $X$.

There is possibility for that nor do can the cylindrical extension and closure restore the original fuzzy relation, which situation can be illustrated as in Fig. 1.

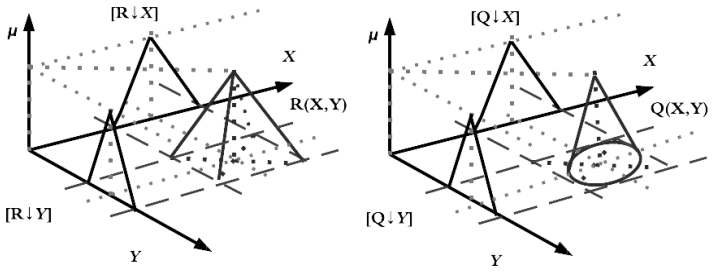



**Fig. 1.** Example of information loss during projection by two distinct fuzzy relations

The max-min composition of $P(X, Y)$ and $Q(Y, Z)$ binary fuzzy relations returns with $R(X, Z)$ binary fuzzy relation, which is associative, its inverse is identical with the inverse relations' reverse composition, but does not satisfy the conditions of commutativity. It is easy to see it can be generalized to any t-norm and t-conorm (or s-norms) pairs. Formally described:

$$R(x, z) = [P \circ Q](x, z) = \max_{y \in Y} \min[P(x, y), Q(y, z)]. \tag{6}$$

The relational connection operator is similar to the max-min composition; it produces an $R(X, Y, Z)$ ternary fuzzy relation from the (relational) connection of $P(X, Y)$ and $Q(Y, Z)$ binary fuzzy relations and it also could be generalized to any t-norm and t-conorm pairs, formally:

$$R(x, y, z) = [P * Q](x, y, z) = \min[P(x, y), Q(y, z)]. \tag{7}$$

The fuzzy similarity measure of vectorvalued fuzzy (VVF) sets proposed and detailed in [18], [19] and [20] could be also applied for fuzzy relations. The similarity relation $S(R, Q)$ of $R$ and $Q$ relations over $X$ multi set is:

$$S_{RQ} = M_{R \equiv Q}, \tag{8}$$

where

$$R \equiv Q = (R \cap Q) \cup (\overline{R} \cap \overline{Q}). \tag{9}$$

## 4 Fuzzy Methods in Recommender Systems

In the last couple of years, several solutions were proposed for the application of fuzzy methods for recommendation problems. In the evaluation of recommender system methods hybrid filtering seemed to be the most effective approach to overcome cold-start problem and exploit behavioural patterns that couldn't be explained solely metadata. Cornelis et al. applied fuzzy relations first in user and item similarities to improve the accuracy of conventional hybrid filtering methods [21].

Later, additional various fuzzy neighbour methods with the combination of collaborative and content-based filtering were introduced in [22-25].

One of the key challenges of recommender systems is to overcome the lack of information for user profiling. To address uncertainty due to vagueness Perez worked out a method by applying fuzzy preference relation that is capable to provide better recommendations for users with a few events [26], Zenebe proposed a general framework for discovering, interpreting and visualizing user preferences with fuzzy set theories [27]. The improvement of user warm-up period were also studied by Porcel and Herrera-Viedma, in [28] they presented a fuzzy linguistic recommendation strategy to improve the acquisition of user profile. Nilashi discussed the usability of fuzzy techniques in multi-criteria recommendation problem to provide better profile models [29].

In order to reduce the uncertainty of preference modelling fuzzy theories were also used for clustering methods. Nadi proposed a fuzzy clustering technique that captures user's behaviours on websites and provides more dynamic recommendations [30]. Liu and Gao studied the interpretation of user intentions with low amount of user actions, they proposed a recommendation solution by the application of fuzzy cluster analysis and cognitive maps [31]. Birtolo and Ronca published a study about two clustering collaborative filtering algorithms with the application of fuzzy logic. They measured a significant improvement in coverage of recommendations while the accuracy remained the same [32].

Several fuzzy-based recommendation methods were addressed to practical problems. Lu proposed a framework that helps students to find learning materials. For that he applied a multi-attribute evaluation method to capture the students' preferences and a fuzzy matching method to find the most suitable materials [33]. For telecommunication domain Wu designed a solution that deals with tree-based structure of contents by using fuzzy similarity measure [34]. Castro-Schez introduced a prototype of recommender system for B2C e-commerce businesses. They proposed a method that capable to deal with vague search preferences and provide fuzzy rule-based personalized recommendations of products [35]. Another application of fuzzy logic in e-commerce was proposed by Ramkumar, who introduced an automatic scoring for the reviews on products for spam detection [36]. Cornelis et al. addressed a method for the modelling of "one-and-only" items (the items that cannot be repetitive sold, e.g. houses). They applied fuzzy logic to extend existing collaborative filtering method and overcome the lack of collaboration [37]. García-Crespo applied fuzzy logic to provide personalized portfolio recommendations considering both financial attributes of investments and psychological aspects [38]. For the recommendation of candidates of political elections, Terán introduced a fuzzy clustering based method [39] and Dyczkowski proposed a voter preference modelling by intuitionistic fuzzy sets [40].

## 5 Possible Applications of Fuzzy Relations in Recommender Systems

The most trivial situation when a news site tries to categorize its viewers: some users tend to read latest news frequently no matter what kind of topic does it have, while older articles are read only according to his more specific preferences, which is suppressed by the data generated during reading the latest news. The main issues with profiling users and items were described in Section 2. In some situations there is no obvious way to define distinct clusters. The properties of fuzzy relations may help to overcome these problems. The fuzzy tolerance and equivalence relations have different mathematical properties, which makes it reasonable to investigate them in the context of recommendation, more specifically user and item preferences. An expression of the users/items and the clusters with fuzzy mathematics may have good results, since it is possible for a user (or item) to be part of various overlapping clusters. All the collected and non-collected (recommendation process related) data from the user are expressing his/her opinion and preferences over item or items. Both implicit, explicit feedback and meta data could be considered as sets of projections of a high-dimensional fuzzy relation; From this point of view it is obvious to assume that there hidden preferences of the user, which are not expressed in the collected data, but it does not require it for further processes. The composition and connection operators (or their generalized forms) defined over binary fuzzy relations (described in Section 3) could be also used to determine or estimate connections between various spaces of feedbacks and meta data.

Another possible application of fuzzy relations in user profiles are the following: consider a (global) fuzzy relation describing the users' basic data (e.g. gender, age group multi set) and their item consumption data (e.g. views, tags multi set). Assume a new user without any stored data; after this user interacts the system starts to observe the user, which could be considered as an imprecise projection of his preference relation. With the global and observed (user) relations it could be possible to determine the best fitting user preference models for the new user only by investigating the similarities. The same principles could be used to speed up the "warm up" period and to overcome the cold-start problem; the user preferences are determined according to the relation between the fuzzy relations expressed by his/her consumption patterns through feedback (or meta data etc) and the fuzzy relation representing the space of known user profiles.

Equivalence and/or similarity measure of fuzzy relations could be used to find the best fitting items with a specific (fuzzy relation) model in a global (fuzzy relation) space. It is easy to see that the computational requirements of the method could be decreased by limiting the calculation of similarity measure of the more specific fuzzy relation and the part of the global fuzzy relation, which is limited by the support of the more specific fuzzy relation. A simple illustration of fuzzy relations describing a user and global preferences can be seen in Fig. 2.

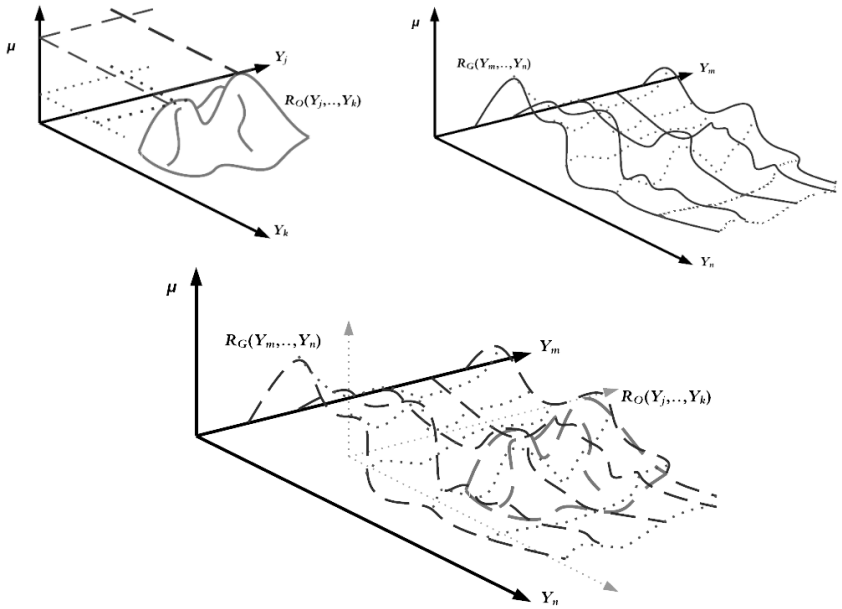



**Fig. 2.** Illustration of "specific" and "global" fuzzy relations describing items or users

Misclassified or purely detailed items without or with only a limited number of feedbacks are difficult to recommend, and the accuracy of these hugely depends on the quality of meta data. Some of the methods use meta data enrichment to overcome this problem. The projections and cylindrical extensions or closures of fuzzy relations representing the user preferences, feedbacks or even statistical parameters could be used to reduce the computational requirements of the methods by decreasing the complexity of the models, but it is also possible to apply these methods to fill out missing data or preferences of a single item or user by typical values of other similar preference relations with the same concepts shown above.

It is easy to see from the previous and this section that, there is a need for a comprehensive and detailed overview of fuzzy methods in recommender systems and it is reasonable to investigate these problems from mathematical aspects, but these are well over the limits of this work.

## 6 Summary and Future Work

In this paper the common problems of recommender systems were addressed from industrial point of view, basics of fuzzy methods and some of their applications were summarized. The mathematical backgrounds of these fuzzy relational concepts in recommender systems should be researched and developed in detail, implemented and compared with actual systems. In order to achieve the proposed object, the following main phases should be executed: (1) a representative detailed data set should be

collected (on both users and items), (2) detailed mathematical research of the application of fuzzy relational calculus (might include the development of new operators and methods in the field), (3) design and implementation of a recommendation system based on the output of previous phases, (4) comparison of actual and the developed systems and (5) integrating and testing selected methods in real life environment.

## Acknowledgement

This paper is partially supported by the GOP-1.1.1-11-2012-0172, the TÁMOP-4.2.2.A-11/1/KONV-2012-0012 and Hungarian Scientific Research Fund (OTKA) grants: K105529, K108405.

## References

1. P. Baranyi and A. Csapo, "Definition and Synergies of Cognitive Infocommunications", Acta Polytechnica Hungarica, vol. 9, no. 1, pp. 67-83, 2012.
2. B. Szi and A. Csapo, "An outline of some human factors contributing to mathability research", In the Proceedings of 5th IEEE Conference on Cognitive Infocommunications, Vietri sul Mare, pp. 583-586, 2014.
3. G. Jawaheer, M. Szomszor, P. Kostkova: Comparison of implicit and explicit feedback from an online music recommendation service, In Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec '10). ACM, USA, pp. 47-51, 2010.
4. X. Amatriain, J. Pujol, N. Tintarev, N. Oliver: Rate it again: increasing recommendation accuracy by user rerating. Proceedings of the third ACM conference on Recommender systems, ACM, pp. 173–180, 2009.
5. S.S. Anand, P. Kearney, M. Shapcott: Generating semantically enriched user profiles for Web personalization. ACM Transactions on Internet Technology 7:4, p. 22, 2007.
6. Y. Hu, Y. Koren, C. Volinsky: Collaborative Filtering for Implicit Feedback Datasets, 2008 Eighth IEEE International Conference on Data Mining, pp. 263-272, 2008.
7. X. Amatriain, J. Pujol, N. Oliver: I like it... I like it not: Evaluating User Ratings Noise in Recommender Systems, User Modeling, Adaptation, and Personalization. Springer, Berlin, pp. 247-258, 2009.
8. D. Parra, A. Karatzoglou, X. Amatriain, I. Yavuz: Implicit feedback recommendation via implicit-to-explicit ordinal logistic regression mapping, Proceedings of the CARS-2011, USA, 5 pages, 2011.
9. R. Van Meteren and M. Van Someren, "Using content-based filtering for recommendation", In Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop, pp. 47–56, 2000.
10. A. M. Rashid, G. Karypis, and J. Riedl. Learning Preferences of New Users in Recommender Systems : An Information Theoretic Approach. In ACM SIGKDD Explorations Newsletter, volume 10, page 90, Dec. 2008.
11. G. Shaw, Y. Xu, and S. Geva. Using Association Rules to Solve the Cold-Start Problem in Recommender Systems. In Advances in Knowledge Discovery and Data, pages 21-24, 2010.

12. H. Sobhanam, and a. K. Mariappan. Addressing cold start problem in recommender systems using association rules and clustering technique. In ICCCI, pp. 1-5, 2013.
13. R. Burke, "Hybrid recommender systems: Survey and experiments", In User Modeling and User-Adapted Interaction, 12(4), pp. 331–370, 2002.
14. J. Basilico, and T. Hofmann, Unifying collaborative and content-based filtering. In the Proceedings of The Twenty-First International Conference on Machine Learning, AXM p. 9, 2004.
15. L. A. Zadeh: Fuzzy Sets, in Information and Control, Vol. 8, pp. 338-353, 1965.
16. L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I, II, III. Information Science, 8:199–251, 301–357, 9:43–80, 1975.
17. L.A. Zadeh. Fuzzy logic and approximate reasoning. Synthese, 30(1):407–428, 1975.
18. L.T. Kóczy, and M. Hajnal, "Cluster analysis in the caryometry applying a new fuzzy algebra ", In W.J. Perkins (ed.), Biomedical Computing, Pitmans Medical Publishing Co. Ltd., Tunbridge Wells, pp. 183-190, 1977.
19. L.T. Kóczy, "Vector valued fuzzy sets", BUSEFAL- BULL STUD EXCH FUZZIN APPL, pp. 41-57, 1980.
20. L.T. Kóczy, M. Hajnal, Classification of textures by vectorial fuzzy sets, In: M M Gupta, E Sanchez (ed.) Approximate Reasoning in Decision Analysis, Amsterdam; Oxford; New York; Tokyo: North-Holland Publishing Company, pp. 157-164, 1982
21. Cornelis, C., Guo, X., Lu, J., & Zhang, G. (2005). A Fuzzy Relational Approach to Event Recommendation. Proceedings of 2nd Indian International Conference on Artificial Intelligence (IICAI'05).
22. Al-Shamri, M. Y. H., & Bharadwaj, K. K. (2008). Fuzzy-genetic approach to recommender systems based on a novel hybrid user model. Expert Systems with Applications, 35(3), 1386–1399.
23. Kant, V., & Bharadwaj, K. K. (2012). Enhancing recommendation quality of content-based filtering through collaborative predictions and fuzzy similarity measures. Procedia Engineering, 38, 939–944.
24. Zhang, Z., Lin, H., Liu, K., Wu, D., Zhang, G., & Lu, J. (2013). A hybrid fuzzy-based personalized recommender system for telecom products/services. Information Sciences, 235, 117–129.
25. Son, L. H. (2014). HU-FCF: A hybrid user-based fuzzy collaborative filtering method in Recommender Systems. Expert Systems with Applications, 41(15), 6861–6870.
26. Perez, L. G., Barranco, M., & Martinez, L. (2007). Building user profiles for recommender systems from incomplete preference relations. In Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International (pp. 1–6). IEEE.
27. Zenebe, A., Zhou, L., & Norcio, A. F. (2010). User preferences discovery using fuzzy models. Fuzzy Sets and Systems, 161(23), 3044–3063.
28. Porcel, C., & Herrera-Viedma, E. (2010). Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. Knowledge-Based Systems, 23(1), 32–39.
29. Nilashi, M., Ibrahim, O. bin, & Ithnin, N. (2014). Hybrid recommendation approaches for multi-criteria collaborative filtering. Expert Systems with Applications, 41(8), 3879–3900.
30. Nadi, S., Saraee, M., & Davarpanah-Jazi, M. (2010). A fuzzy recommender system for dynamic prediction of user's behavior. In Internet Technology and Secured Transactions (ICITST), 2010 International Conference for (pp. 1–5). IEEE.
31. Liu, W., & Gao, L. (2014). Recommendation System Based on Fuzzy Cognitive Map. Journal of Multimedia, 9(7), 970–976.

32. Birtolo, C., & Ronca, D. (2013). Advances in Clustering Collaborative Filtering by means of Fuzzy C-means and trust. Expert Systems with Applications, 40(17), 6997–7009.
33. Lu, J. (2004). Personalized e-learning material recommender system. In International conference on information technology for application (pp. 374–379).
34. Wu, D., Zhang, G., & Lu, J. (2013). A fuzzy tree similarity measure and its application in telecom product recommendation. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 3483–3488). IEEE.
35. Castro-Schez, J. J., Miguel, R., Vallejo, D., & López-López, L. M. (2011). A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals. Expert Systems with Applications, 38(3), 2441–2454.
36. Ramkumar, V., Rajasekar, S., & Swamynathan, S. (2010). Scoring products from reviews through application of fuzzy techniques. Expert Systems with Applications, 37(10), 6862–6867.
37. Cornelis, C., Lu, J., Guo, X., & Zhang, G. (2007). One-and-only item recommendation with fuzzy logic techniques. Information Sciences, 177(22), 4906–4921.
38. García-Crespo, Á., López-Cuadrado, J. L., González-Carrasco, I., Colomo-Palacios, R., & Ruiz-Mezcua, B. (2012). SINVLIO: Using semantics and fuzzy logic to provide individual investment portfolio recommendations. Knowledge-Based Systems, 27(0), 103–118.
39. Terán, L., & Meier, A. (2010). A fuzzy recommender system for eelections. In Electronic Government and the Information Systems Perspective (pp. 62–76). Springer.
40. Dyczkowski, K., & Stachowiak, A. (2012). A recommender system with uncertainty on the example of political elections. In Advances in Computational Intelligence (pp. 441–449). Springer.

# A method to analyse road risks using artificial vision and fuzzy logic

Juan Giralt, Juan Moreno-Garcia, Luis Jimenez-Linares, and Luis Rodriguez-Benitez

Information Systems and Technologies Department
University of Castilla-La Mancha (Spain)
{juan.giralt,juan.moreno,luis.jimenez,luis.rodriguez}@uclm.es
http://webpub.esi.uclm.es/investigacion/grupos/oreto

**Abstract.** This paper presents a system to detect abnormal movements of a vehicle in a road. This technique compares the trajectory of the vehicle with the information obtained from the lane marks of the road and detects lane changes and vehicle skids. Then, a process to obtain one single value representing the road's shape (left bend, straight-away, right bend) is done, to do this, each of the slopes of the two lines marking the edges of traffic lanes are computed. Then a comparison process to detect those frames where there is a logical correspondence between the vehicle displacement and the road shape is performed. The proposed system takes only as input information from the H264/AVC motion vectors and the videos are captured from a moving vehicle. As output the time intervals in which the vehicle displacement corresponds with a risky situation are obtained.

**Keywords:** motion vectors, H264/AVC, lane mark segmentation, linguistic comparison

## 1 Introduction

Intelligent Transport Systems use advanced technologies to improve vehicle's safety, for example, systems using computer vision techniques to segment the lane marks in a road [6], perform monitoring [8], and alerting to deviations in the path of the vehicle detected from information highway lines [1, 4]. There are also works dealing with this problem using fuzzy logic, such as Wang et al. [7] and Obradovic et al. [5]. One of the main features of the proposed method is its efficiency, that is because it works with very little input information. This is possible by taking as input data the motion vectors of the video compression standard H264/AVC. This standard uses motion compensation based on macroblocks, pixel arrays, etc. and exploits duplicate information in successive frames. Similar information present in a frame and another, called reference frame, are not stored, only a motion vector representing the displacement between macroblocks is used to code the macroblock spatial differences. The motion vector field obtained by H264/AVC can be considered as the sparse and imprecise approximation of the optical flow.

The rest of the paper is organized as follows. Section 2 presents the mechanism that allows to represent by means of a fuzzy value the information obtained from the segmentation of the lane marks of the road. Later, in Section 3 an algorithm to compare the vehicle displacement with the road shape is detailed. Finally, in Sections 4 and 5 the experimental results, the conclusions and the future works are shown.

## 2 Road's shape detection and representation

Algorithm 1 shows how to obtain a list containing the road's shape changes all along the video. The first step of this algorithm is based on a technique presented in [2]. This method detects the lane marks of the road that are represented using a set of statistical attributes. In this paper only is needed the information from the left and right lane marks slopes named $s(LL)$ and $s(RL)$, respectively.

**Algorithm 1** Computing the list $L'$

1: $RS \leftarrow$ detection of line geometry using [2]
2: $FRSList \leftarrow$ fuzzification of each $RS$ and computation of the list
3: $L \leftarrow$ Grouping consecutive elements of $FRSList$
4: $L' \leftarrow$ Grouping consecutive elements of $L$

Figure 1 shows the values of the left and right slopes of the first 1000 frames of a sample video. The road turns left when the line slopes decrease, and the road turns right when the slopes increase their values. If it is not possible to detect a line in a concrete frame the method assumes the last value detected as it happens in sharp turns, for example, in roundabouts [2].

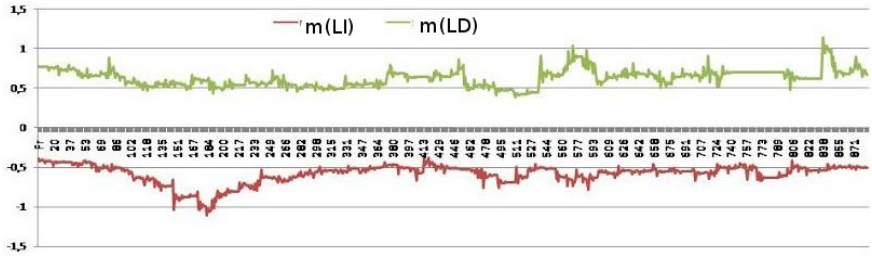



**Fig. 1.** Slopes of the left and right lines of the lane

A single measure called $s(m)$ is calculated to automatically obtain the road shape combining the slopes $s(LL)$ and $s(RL)$ for every frame (Equation 1).

$$s(m) \leftarrow s(LL) + s(RL) \tag{1}$$

Equation 2 assigns a description in natural language to the road's shape in every frame of the video. Two threshold values are needed ($U_{min}$ and $U_{max}$). These values are obtained empirically based on the minimal and maximal values of $s(m)$ in a frame sequence while the vehicle is moving in a straight line. This sequence could be considered as a "training sequence". $U_{min}$ and $U_{max}$ are $-0.1$ and 0.18 in Figure 2. The list containing the representative values of each frame is called *Road Shape* ($RS$).

$$text \leftarrow \begin{cases} Left\ bend, & If\ s(m) < U_{min} \\ Straight, & If\ U_{min} \leq s(m) \leq U_{max} \\ Right\ bend, & If\ s(m) > Umax \end{cases} \tag{2}$$

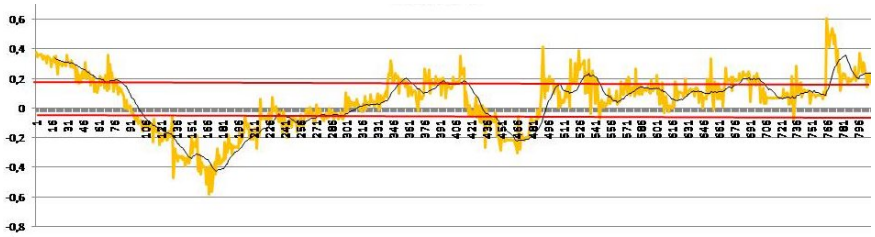



**Fig. 2.** Joint slopes and threshold values.

A linguistic variable [9] called *Road Shape* ($RSV$) is used now. The reason to translate the representation to the fuzzy domain is to manage the inherent noise in the H264 motion vectors and the use of a mobile camera that also add noise to the captured data (step 2 of Algorithm 1). This variable allows to add fuzzy boundaries around $U_{min}$ and $U_{max}$, and it is composed of three trapezoidal fuzzy sets corresponding to the linguistic labels *Left Bend* ($LB = \{-0.5, -0.5, U_{min} - k, U_{min} + k\}$), *Right Bend* ($RB = \{U_{min} - k, U_{min} + k, U_{max} - k, U_{max} + k\}$) and *Straight* ($S = \{U_{max} - k, U_{max} + k, 0.5, 0.5\}$) respectively. The parameter $k$ is determined empirically and it is based on the maximum and minimum values of $s(m)$ in each concrete video. $RSV$ support is defined between $[-0.5, 0.5]$ for the videos used in the experimentation.

The fuzzification process consists of the generation of tuples named $RSFuzzy$ in every frame and their structure is shown in Equation 3.

$$RSFuzzy = (frame, \mu_{LC}(s(m)), \mu_S(s(m)), \mu_{RC}(s(m))) \tag{3}$$

Each one of these tuples are stored in a list ordered by the number of frame (Equation 4).

$$FRSList \leftarrow FRSList + \{RSFuzzy\} \tag{4}$$

After that (step 3 of Algorithm 1), $SetSize$ consecutive elements are processed to obtain a single value that represents these elements. $SetSize$ is empirically defined, and it must be proportional to the modulus of the motion vectors. The more speed is reached the less the value of $SetSize$ is, because more distance is covered in less time and there can be more rapid changes in the shape of the road. Now, a linguistic value is going to represent the road's shape in $SetSize$ consecutive elements of $FRSList$. This value is named $Label$ and it is obtained as the linguistic variable with maximum membership once all the memberships to the same variable are added for the $SetSize$ frames. Then, in Equation 5 the new linguistic representation for $SetSize$ frames is shown.

$$MaxFCFuzzy = (frame\_ini,\ frame\_fin,\ Label) \tag{5}$$

Finally, in the last step of the Algorithm 1, a new clustering process is done and it obtains $L'$. This process is needed to joint consecutive elements in $L$ with the same value for the attribute $Label$. Equation 6 shows the structure of the tuple containing the union of consecutive elements in $L$ satisfying this condition.

$$MaxFCFuzzy = (frame\_ini,\ frame\_fin + SetSize,\ Label) \tag{6}$$

In brief, $L'$ is a list that contains the road's shape through the time using $MaxFCFuzzy$ tuples to represent it. For example, Equation 7 details how a video of 3420 frames is represented.

$$L' = \{(0,\ 139,\ S),\ (140\ ,389\ ,RB)...,(3300\ ,3419\ ,S)\} \tag{7}$$

## 3 Detection of risky situations

The detection of risky situations is done by means of a comparison process between the information contained in $L'$ and the results obtained in [3] about the characterization of the vehicle displacement. This method is similar to the one proposed in Section 2. It generates a tuple called *Video Vehicle Displacement* ($VVD$) with a similar structure to $L'$, where the linguistic labels used to represent the displacement are *Turning Left* ($TL$), *Straight* ($S$) and *Turning Right* ($TR$). Table 1 shows an example of the elements to be compared. The time intervals are different for each one of the lists since they are obtained using different processes. Then a mechanism to establish common time intervals that represent the same interval times must be done. This process is described in Section 3.1, after that, the comparison process is detailed (Section 3.2).

### 3.1 Obtaining the common intervals

Two ordered lists are obtained to identify the common intervals. These lists contain the initial frame of the video and the attributes that define the final of each temporal interval in each one of the lists. This attribute is named $frame\_fin$ in

**Table 1.** $L'$ (left) and $VVD$ (right) to compare.

| $f\_ini$ | $f\_fin$ | **Label** | $f\_ini$ | $f\_fin$ | **Label** |
|---|---|---|---|---|---|
| 0 | 109 | S | 0 | 119 | S |
| 110 | 283 | LB | 120 | 270 | TL |
| 284 | 327 | S | 271 | 329 | S |
| 328 | 458 | RB | 330 | 470 | TR |
| 459 | 479 | S | 471 | 482 | S |
| 481 | 490 | LB | 483 | 509 | TL |

Equation 5. Equations 8 and 9 show the two lists from the information shown in Table 1.

$$L_1 \leftarrow \{0, 109, 283, 327, 458, 479, 490\} \tag{8}$$

$$L_2 \leftarrow \{0, 119, 270, 329, 470, 482, 509\} \tag{9}$$

After that, the two lists are merged in order to obtain an ordered set named *Common Interval* ($CI$). As mathematical sets, it does not allow duplicate elements. For example, from the lists of the Equations 8 and 9 the $CI = \{0, 109, 119, 270, 283, 327, 329, 458, 470, 479, 482, 490, 509\}$ is obtained. Each two consecutive values of $CI$ is now a new interval. Both $L'$ and $VVD$ are reorganized using these intervals as it is shown in Table 2.

**Table 2.** New temporal values in $L'$ (left) and $VVD$ (right).

| $f\_ini$ | $f\_fin$ | **Etiq.** | $f\_ini$ | $f\_fin$ | **Etiq.** |
|---|---|---|---|---|---|
| 0 | 109 | S | 0 | 109 | S |
| 110 | 119 | LB | 110 | 119 | S |
| 120 | 270 | LB | 120 | 270 | TL |
| 271 | 283 | LB | 271 | 283 | S |
| 284 | 327 | S | 284 | 327 | S |

### 3.2 Comparison process

The elements of $L'$ and $VVD$ are now compared in the same temporal interval in order to detect the correspondences between the lane marks and the displacement of the vehicle. More concretely, the attribute *Label* of $L'$ and $VVD$ is used and from Equation 10 risky situations can be obtained. For example, this equation returns *TRUE* for the interval $[0, 109]$ and *FALSE* for the interval $[110, 119]$ (Table 2).

$$Comparison \leftarrow \begin{cases} True,\ Label(L') = LB\ and\ Label(VVD) = TL \\ True,\ Label(L') = S\ and\ Label(VVD) = S \\ True,\ Label(L') = RB\ and\ Label(VVD) = TR \\ False, In\ other\ case \end{cases} \quad (10)$$

The output of the comparison process is stored in a List of Differences ($LD$) containing the tuples with the following structure: $\{frame\_ini,\ frame\_fin,\ Label_{L'},\ Label_{VVD}\}$. These are the intervals with discrepancies between lane marks and the displacement of the vehicle.

## 4 Experimental results

In the experimentation, three videos are used. One is a video recorded by the authors whilst the two other videos were captured in a World Rally Car competition. The two drivers were Sebastian Loeb and Peter Solberg. We consider these two last videos very difficult for the aims of these tests since there are sudden changes in the direction in the speed and there a lot of vegetation in the ditch and in the road edge because the car drives along a rural minor road. This videos are identified as *Own*, *Loeb* y *Solberg*, respectively.

### 4.1 Evaluation of the detection of the road's shape

Table 3 exposes the obtained results for each one of the experiments in the process of road's shape recognition. The success rate is considerably lower for Solberg and Loeb tests and failures mainly occur when the car begins a sharp curve with a high speed. That is because the lane mark segmentation algorithm loses the line followed and this error affects to the high level recognition process here detailed.

**Table 3.** Road characterization.

| | **Own** | **Loeb** | **Solberg** |
|---|---|---|---|
| **Hits** | 89.6% | 67.7% | 60.2% |
| **Errors** | 10.4% | 32.3% | 38.8% |

### 4.2 Evaluation of risky situations

Table 4 shows the comparison process for one of the experiments. When $False$ is obtained as result, a new element to the list of differences is added.

Then, hits of the system occurs when an element in the list of differences corresponds with a risky situation in the video. The errors occur when the proposed system does not detect risky situations as it is shown in Table 5 (from the

**Table 4.** Comparison results in the experiment *Own*

| $f\_ini$ | $f\_fin$ | **VVD** | **L'** | **Comparison** |
|---|---|---|---|---|
| 0 | 26 | S | S | True |
| 27 | 47 | S | RB | False |
| 48 | 113 | S | S | True |
| 114 | 244 | TL | LB | True |
| 245 | 283 | TL | S | False |
| 284 | 327 | S | S | True |
| 328 | 375 | TR | S | False |
| 376 | 397 | TR | RB | True |
| 398 | 451 | TR | S | False |
| 452 | 462 | TR | RB | True |
| 463 | 495 | S | S | True |
| 496 | 527 | S | LB | False |
| 528 | 571 | S | S | True |
| 572 | 593 | S | RB | False |
| 594 | 710 | S | S | True |
| 711 | 752 | TR | S | False |
| 753 | 796 | TL | S | False |
| 797 | 807 | S | S | True |
| 808 | 839 | TR | S | False |
| 840 | 909 | S | S | True |
| 910 | 942 | S | LB | False |
| 943 | 999 | S | S | True |

column 1 to 4) where risky situations are not present in the list of differences. The percentage of the table refers to the number of frames. In *Loeb* experiment, our system reaches a 97.7% of hits, then the 2.3% corresponds to risky situations not detected.

**Table 5.** Risky detection in frames not present in *LD*

| **frames not present in *LD*** | | | | **frames present in *LD*** | | | |
|---|---|---|---|---|---|---|---|
| | **Own** | **Loeb** | **Solberg** | | **Own** | **Loeb** | **Solberg** |
| **Hits** | 100% | 97.7% | 76.3% | **Hits** | 32.5% | 30.2% | 29.9% |
| **Errors** | 0% | 2.3% | 23.7% | **Errors** | 67.5% | 69.8% | 70.1% |

Table 5 (from the column 5 to 8) details the results of the hits of our system. A 29.9% detects the risky situations in the experiment *Solberg* (70.1% can be considered errors). This percentage of errors are mainly due to the difficulty of this video and the fact that the car continuously drives from one lane to the another and other unexpected behaviours.

## 5 Conclusions and future works

In this work we have presented a new technique to represent in a linguistic way the road shape from a video sequence using as input data the H264/AVC motion vectors. The use of fuzzy logic allows to work with linguistic representations and to process the information from several frames simultaneously. The use of a very little amount of data allows to obtain a no time-consuming method. The use of linguistic variables to represent the vehicle displacement and the road shape makes interpretable the comparison process proposed. Despite the difficulty of the selected videos for the experimentation, acceptable results have been obtained in this first approach. As future works, it should be developed mechanisms to analyse in more detail each one of the differences obtained as output of the comparison process. Issues as their duration or the degree of differentiation between vehicle displacements and road's shape are factors that should be taken into account for future developments.

## Acknowledgement

This work has been funded by the Information Systems and Technologies Department of the University of Castilla-la Mancha.

## References

1. A. Borkar, M. Hayes, and M. T. Smith. A new multi-camera approach for lane departure warning. In *Proceedings of the 13th international conference on Advanced concepts for intelligent vision systems*, Springer-Verlag, Berlin, Heidelberg, pp. 58–69. 2011.
2. J. Giralt, L. Rodriguez-Benitez, J. Moreno-Garcia, C.J. Solana-Cipres, L. Jimenez: Lane mark segmentation and identification using statistical criteria on compressed video. *Integrated Computer-Aided Engineering* 20, 2, pp. 143–155, 2013.
3. J. Giralt, J. Moreno-Garcia, L. Jimenez-Linares L., E. Del Castillo, L. Rodriguez-Bentez. A Fuzzy Representation of Vehicle Trajectories using Motion Data from H264/AVC Video. In *Proceedings of the 14th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE 2014*, J. Vigo-Aguiar Ed., 2, pp. 635–646, 2014.
4. J. Lee and U. Yi. A lane-departure identification based on LBPE, Hough transform, and linear regression. *Comput. Vis. Image Underst* 99, 3 , pp. 359–383, 2005.
5. I. Obradovic, Z. Konjovic, E. Pap, and I. J. Rudas. Linear fuzzy space based road lane model and detection. *Know.-Based Syst*, pp. 37–47. 2013.
6. T. Tran, J. Son, B. Uk, J. Lee, and H. Cho. An adaptive method for detecting lane boundary in night scene. In *Proceedings of the Advanced intelligent computing theories and applications*. Springer-Verlag, Berlin, Heidelberg, pp. 301–308. 2010.
7. J. Wang, C. Lin, and S. Chen. Applying fuzzy method to vision-based lane detection and departure warning system. *Expert Syst. Appl.*, 37, 1, pp. 113–126. 2010.
8. Y. Wang, N. Dahnoun, and A. Achim. A novel system for robust lane detection and tracking. *Signal Process.*, 92, pp. 319–334. 2012.
9. L. A. Zadeh: Similarity relations and fuzzy orderings. *Information Sciences*, pp. 177–200. 1971.

# Using genetic algorithms to learn a fuzzy based pseudometric for k-NN classification

F. Martins[1], J. C. Becceneri[1], L. Dutra[1], D. Lu[2], and S. Sandri[1]

[1] Instituto Nacional de Pesquisas Espaciais
12201-970, São José dos Campos, SP, Brazil
E-mail: flavinha@dpi.inpe.br, becce@lac.inpe.br, luciano.dutra@dpi.inpe.br, sandra.sandri@inpe.br
[2] Michigan State U. - Center for Global Change and Earth Observations, East Lansing, MI, USA
E-mail: ludengsh@msu.edu

**Abstract.** We address the derivation of pseudometric based on fuzzy relations for classification applications, by the use of genetic algorithms to learn the fuzzy relations. We present an experiment for the classification of land use in an area of the Brazilian Amazon region.

**Keywords:** k-NN classification, fuzzy partitions, genetic algorithms

## 1 Introduction

In a previous work [2], we proposed a a function called $f^+$, based on fuzzy relations, which are themselves derived from fuzzy partitions, for use in classification applications. This function is the complement in $[0, 1]$ of a particular kind of fuzzy relation, called an Order Compatible Fuzzy Relation (OCFR$_\preceq$), defined using a total order $(\Omega, \preceq)$ [10]. An OCFR$_\preceq$ itself is derived from a type of fuzzy partition (a collection of fuzzy sets), called Convex Fuzzy Partitions (CFP$_\preceq$). The creation of OCFR$_\preceq$ was motivated by the need to ease the burden of creating suitable relations for use the particular fuzzy case-based reasoning classification approach proposed in [8]. In [2], we proved that $f^+$ function is i) a pseudometric, when obtained from a specific type of CFP$_\preceq$, called 2-Ruspini, and, in particular, a ii) metric, when this CFP$_\preceq$ is moreover composed solely of triangular fuzzy sets. The same happens in the case of multidimensional domains, for function $f^+_{(\mu)}$ that aggregates the results obtained for $f^+$ in each domain, using the arithmetic means as aggregation operator $\mu$.

Here we address the derivation of $f^+$ for k-NN classification applications [11], by the use of fuzzy genetic algorithms [1] to learn the fuzzy relations. We describe an application in the classification of land cover and use in an area of the Brazilian Amazon region.

## 2 Fuzzy relation based pseudometrics $f^+$ and $f^+_{(\mu)}$

Let $S : \Omega^2 \to [0, 1]$ be a fuzzy binary relation and $(\Omega, \preceq)$ be a total order. Formally, $S$ is an Order Compatible Fuzzy Relation with Respect to a Total Order $(\Omega, \preceq)$ (*OCFR*$_\preceq$ or *OCFR*, for short), when it obeys the following properties [10]:

– $\forall x, y, z \in \Omega$, $S(x, x) = 1$ (*reflexivity*)
– $\forall x, y, z \in \Omega$, $S(x, y) = S(y, x)$ (*symmetry*)
– $\forall x, y, z \in \Omega$, if $x \preceq y \preceq z$, then $S(x, z) \leq \min(S(x, y), S(y, z))$ (*compatibility with total order* $(\Omega, \preceq)$, or $\preceq$*-compatibility* for short).

Let $(\Omega, \preceq)$ be a total order and let $\mathbf{A} = \{A_1, ..., A_t\}$ be fuzzy partition (a collection of fuzzy sets) in $\Omega$; here $A_i$ denotes a fuzzy set but also its associated membership function. Let the core and support of a fuzzy set $A$ be defined as $core(A) = \{x \in \Omega \mid A(x) = 1\}$ and $supp(A) = \{x \in \Omega \mid A(x) > 0\}$), respectively [3]. Formally, $\mathbf{A}$ is a Convex Fuzzy Partition with Respect to a Total Order $(\Omega, \preceq)$ (*CFP*$_\preceq$ or *CFP*, for short), if it obeys the following properties [10]:

1. $\forall A_i \in \mathbf{A}, \exists x \in \Omega, A_i(x) = 1$ (*normalization*),
2. $\forall x, y, z \in \Omega, \forall A_i \in \mathbf{A}$, if $x \preceq y \preceq z$ then $A_i(y) \geq \min(A_i(x), A_i(z))$ (*convexity*),
3. $\forall x \in \Omega, \exists A_i \in \mathbf{A}, A_i(x) > 0$ (*domain-covering*),
4. $\forall A_i, A_j \in \mathbf{A}$, if $i \neq j$ then $core(A_i) \cap core(A_j) = \emptyset$ (*non-core-intersection*).

Let $\mathcal{A}_{(\Omega,\preceq)}$ denote the set of all CFPs that can be derived considering a total order $(\Omega, \preceq)$. CFP $\mathbf{A} \in \mathcal{A}_{(\Omega,\preceq)}$ is said to be a $n$*-CFP* if each element in $\Omega$ has non-null membership to at most $n$ fuzzy sets in $\mathbf{A}$ ($n \geq 1$). In particular, a 2-CFP$_\preceq$ $\mathbf{A}$ is called a *2-Ruspini* partition, when it obeys additivity:

– $\forall x \in \Omega$, $\sum_i A_i(x) = 1$ (*additivity*)

In [10], the authors propose to generate OCFR$_\preceq$ $S^+ : \Omega^2 \to [0, 1]$ from a CFP$_\preceq$ $\mathbf{A}$ as

$$S^+(x, y) = \begin{cases} 0, \text{if } S^*(x, y) = 0 \\ S_L(x, y), \text{otherwise} \end{cases}$$

$$\forall x, y \in \Omega, S^*(x, y) = \sup_i \min(A_i(x), A_i(y))$$

$$\forall x, y \in \Omega, S_L(x, y) = \inf_i \; 1 - \mid A_i(x) - A_i(y) \mid$$

Note that $S_L$ is constructed based on the Lukasiewicz biresiduated operator [9].

In [2], the following function was proposed for tasks in which metrics and pseudo-metrics are employed[1]:

$$\forall x, y \in \Omega, f^+_{\mathbf{A}}(x, y) = 1 - S^+_{\mathbf{A}}(x, y).$$

This formula can be written directly as:

$$\forall x, y \in \Omega, f^+_{\mathbf{A}}(x, y) = \begin{cases} 1, \text{if } \forall i, \min(A_i(x), A_i(y)) = 0, \\ \sup_i \; \mid A_i(x) - A_i(y) \mid, \text{otherwise.} \end{cases}$$

[1] A metric satisfies non-negativity, symmetry and the triangle inequality and the identity of indiscernibles properties. Pseudometrics obey the same properties, except for the identity of indiscernibles, that is substituted by anti-reflexivity, a weaker property.

When no confusion is possible, we denote $f^+_{\mathbf{A}}$ as simply $f^+$.

Let $O = \Omega_1 \times ... \times \Omega_m$, where $\forall i, (\Omega_i, \preceq)$ is a total order. Let $\mathbf{A}_i$ be a 2-Ruspini CFP$_\preceq$ on $\Omega_i$ and $f^+_i$ be derived from $\mathbf{A}_i$. Let $f^+_{(\mu)} : O \to [0,1]$ be the extension of function $f^+$ to multidimensional domains, defined as

$$f^+_{(\mu)}(x,y) = \mu(f^+_1(x_1,y_1), ..., f^+_m(x_m,y_m)),$$

where $\mu : [0,1]^m \to [0,1]$ is the arithmetic mean, i.e., $\mu(a_1,...,a_m) = \frac{\sum_{1 \le i \le m} a_i}{m}$.

In [2], it is proved that $f^+_{\mathbf{A}}$ is a pseudometric, in general, and a distance when all fuzzy sets in $\mathbf{A}$ are triangular. Function $f^+_{(\mu)}$ trivially satisfies symmetry, anti-reflexitivity and non-negativity. The same result holds for $f^+_{(\mu)}$. In the same work, function $f^+_{(\mu)}$ was tested in a real-world application and yielded very good results when compared to both the Euclidean and Mahalanobis distances.

## 3 Learning $f^+_{(\mu)}$ using genetic algorithms for k-NN classification

We propose to use genetic algorithms to learn the fuzzy partitions necessary for function $f^+_{(\mu)}$, which is also our fitness function. Here we consider classification by k-NN but other methods could be used.

Let $X = \{x_1, ...x_m\}$ be a set of variables, each of which defined in domain $\Omega_i = [l_i, u_i], i \in \{1, m\}$. We encode each chromosome as a sequence of $m$ genes, each of which related to a variable in $X$. The i-th gene is a sequence of parameters $< p_1, ..., p_s >$, representing points in domain $\Omega_i$ for a Ruspini partition. The sequence is such that $p_i \le p_{i+1}, 1 \le i \le s-1$. In a trapezoidal partition, the first (respec. last) fuzzy term will have $[l_i, p_1]$ (respec. $[p_s, u_i]$) as core and $[l_i, p_2]$ (respec. $[p_{s-1}, u_i]$) as support. In a triangular partition, the first (respec. last) fuzzy term will have $l_i$ (respec. $u_i$) as core and $[l_i, p_1]$ (respec. $[p_s, u_i]$) as support.

Crossover consists in choosing a cutting place in two selected chromosomes $c_1$ and $c_2$, and generating two new chromosomes $c_{12}$ and $c_{21}$. Let chromosome $c_i$ be described as $< p_{i,1}, ..., p_{i,s} >$ and let the cutting happen between the (k)-th and (k+1)-th genes. The crossover between any two chromosomes $c_1$ and $c_2$ would be generate two new chromosomes $c_{12}$ and $c_{21}$, respectively described as $< p_{1,1}, ..., p_{1,k}, p_{2,k+1}, , ..., p_{2,s} >$ and $< p_{2,1}, ..., p_{2,k}, p_{1,k+1}, , ..., p_{1,s} >$

If one of the generated chromosomes does not satisfy the condition on the $p_i$s, we reorganize the parameters. For example, let us suppose we have two chromosomes with 3 trapezoidal fuzzy sets Let $c_1$ and $c_2$ be described as $< 10, 20, 30, 40 >$ and $< 31, 32, 33, 34 >$, respectively, and that the cutting point is between $p_2$ and $p_3$. We obtain a valid chromosome, $c_{12} =< 10, 20, 33, 34 >$, and an invalid one, $c_{21} =< 31, 32, 30, 40 >$. We then rearrange the invalid chromosome as $c_{21} =< 30, 31, 32, 40 >$.

In this work we use $n$-fold cross-validation. First of all, a data set $T$ is partitioned in $n$ (approximately) equal parts (folds) $T_i$, such that $T = \cup_i T_i$. Then, for a given fold $i$, training is performed using the elements of all folds, except for those in $i$, and testing is performed the elements of fold $i$ itself, making $Train_i = \bigcup_{T_j \in T, j \neq i} T_j$, and $Test_i = T_i$.

## 4 Experiments

In the following, we briefly describe an experiment that illustrates the use of function $f^+_{(\mu)}$ in a land use classification task in the Brazilian Amazon region. The area of interest covers approximately 411 km$^2$ and in the municipality of Belterra, state of Pará, in the Brazilian Amazon region, partially contained in the National Forest of Tapajós. An intense occupation process occurred in the region along the BR-163 highway (Cuiabá-Santarém), with opening of roads to establish small farms, after deforestation of primary forest areas [4]. As a result, there are mosaics of secondary vegetation in various stages, with pastures and cultivated areas embedded in a forest matrix [5].

In this application, 14 attributes have been considered, derived from either radar or optical satellite images, with 6 classes: forest, initial or intermediate regeneration, advanced regeneration or degraded forest, cultivated area, exposed soil, and pasture. The samples consist of 138 ground information based hand-made polygons. The attribute value for each polygon is the average of the values for the pixels composing it. The experiments have been done using 6 folds (5 for training and 1 for testing).

To obtain the lower (respec. upper) bound for a variable domain, we took the smallest (respec. largest) value from the elements in the fold, less (respec. plus) 20%. We have tested two types of partition for each variable, a triangular and a trapezoidal one, each of which with 3 fuzzy terms. In the triangular experiment, each partition is described by $< p_1 >$, where $p_1$ is the core of the middle triangular fuzzy term. In the trapezoidal experiment, each partition is described by $< p_1, p_2, p_3, p_4 >$, where $[p_2, p_3]$ is the core of the middle trapezoidal fuzzy term.

In our experiments, for each fold, the candidate population has 10 chromosomes. Each chromosome has 3 genes, each of which describing a partition corresponding to one of 3 variables used here. We have used an elitist genetic algorithm, keeping the best 6 elements and combining the 3 first elements to generate the new candidates that replace the worst 4 elements. We used a mutation rate of .2 and 400 generations.

We have used two kinds of population in the initial generation for each fold: "random" and "selected". In the selected first population for the fuzzy terms, the points are obtained from a fixed set of percentage vectors. Considering all domains to be normalized to [0,1], the selected population for the trapezoidal fuzzy sets corresponds to the set of 10 quadruples $< .20, .40, .60, .80 >$, $< .05, .28, .52, .76 >$, $< .23, .47, .71, .95 >$, $< .23, .47, .52, .76 >$, $< .23, .28, .52, .76 >$, $< .23, .47, .71, .76 >$, $< .4, .55, .7, .85 >$, $< .15, .55, .7, .85 >$, $< .15, .3, .7, .85 >$ and $< .15, .3, .45, .85 >$. The selected population for the triangular fuzzy sets is obtained by taking the arithmetic means between $p_2$ and $p_3$ from the trapezoidal fuzzy terms. It corresponds to $< .50 >$, $< .40 >$, $< .59 >$, $< .49 >$, $< .40 >$, $< .59 >$, $< .62 >$, $< .62 >$, $< .50 >$ and $< .37 >$.

Figure 4 brings the accuracy results for this application, considering k-NN with 1 to 6 neighbours, using the several versions of function $f^+_{(\mu)}$: trapezoid-based and triangle-based, considering selected and random initial populations (kNN_dFtz_s, kNN_dFtz_r, kNN_dFtg_s, kNN_dFtg_r). For comparison, the figure also brings the Euclidean distance (kNN_dE).

We see from the figures that all methods had high accuracy and that the best average results in the 6 folds were obtained with the use of $f^+_\mu$ for the triangular partitions. The best individual results, considering all folds, were the same methods for 1, 2 and 3

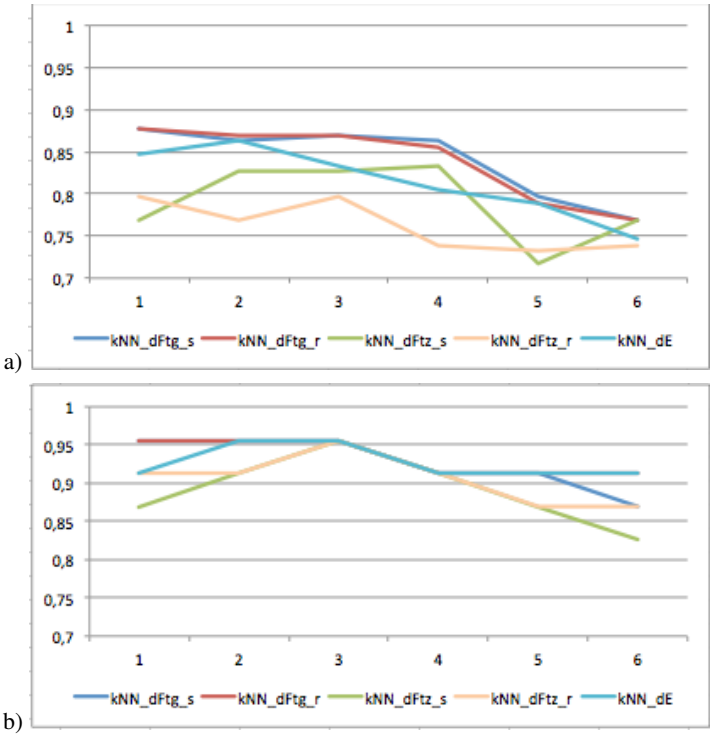



**Fig. 1.** Classification accuracy results for: a) k-NN average and b) k-NN maximum.

neighbours and the Euclidean distance for 2 and 3 neighbours. In particular, $f_{\mu}^{+}$ for the triangular partitions with the initial population obtained at random yielded the same results for the maximum as the Euclidean distance, except for 1 neighbour, when $f_{\mu}^{+}$ fares better. All methods fare better with a small number of neighbours. In particular, the best results for the triangular partitions, considering both the average and the maximum, is obtained already with a single neighbour. The worst results have been obtained with the trapezoidal partitions, for both types of initial populations.

## 5 Conclusions

In this work, we have proposed to use of genetic algorithms to learn fuzzy relations, that are parameters for a pseudometric $f_{(\mu)}^{+}$. We describe a classification application of land cover and use in an area of the Brazilian Amazon region, using k-NN. The results have shown that the triangular partitions produced the best results.

Future work includes experimenting with other data sets. We also intend to verify alternatives to reduce the computational cost, without a decrease in accuracy or adequately reducing the training data Another alternative consists in learning the partition for each variable separately; in order to calculate accuracy the distance relative to the

other variables would be fixed (e.g. Euclidean) and aggregated with the distance obtained from the partition.

This work is a first step towards using $f^{+}_{(\mu)}$ in [7], an extension to k-NN for image classification, in which there is the possibility of using multiple spaces, that can be originated from different data sources, having different ranges of values, as well as the geographical space itself, allowing the use of topological associations.

**Acknowledgements** The authors acknowledge financial support from Brazilian research funding agencies CAPES, CNPq and FAPESP.

## References

1. F. Herrera, Genetic Fuzzy Systems: Status, Critical Considerations and Future Directions, International Journal of Computational Intelligence Research, Vol.1, No.1, pp. 59-67 (2005)
2. Sandri, S., Martins-Bedê, F.T., Dutra, L.,: Using a Fuzzy Based Pseudometric in Classification. Proc. 14th Int. Conf. on Info. Proc. and Management of Uncertainty in Knowledge-Based Systems, IPMU 2014, Montpellier-Fr (2014)
3. Dubois, D., Prade, H.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York (1988)
4. Brazilian Institute of Environment and Renewable Natural Resources (IBAMA): Floresta Nacional do Tapajós Plano de Manejo. Vol I. (in Portuguese) Available at: <http://www.icmbio.gov.br/portal/images/stories/imgs-unidades-coservacao/flona_tapajoss.pdf>. Date accessed: 30 Dec. 2013 (2009)
5. Escada, M.I.S., Amaral, S., Rennó, C.D., Pinheiro, T.F.: Levantamento do uso e cobertura da terra e da rede de infraestrutura no distrito florestal da BR- 163. Repport INPE-15739-RPQ/824, INPE, S.J.Campos, Brazil. Available at: <http://urlib.net/sid.inpe.br/mtc-m18@80/2009/04.24.14.45>. Date accessed: 30 Dec. 2013 (2009)
6. Korsrilabutr, T., Kijsirikul, B.: Pseudometrics for Nearest Neighbor Classification of Time Series Data. Engineering Journal, Thailand, 13, May. 2009. Available at: <http://engj.org/index.php/ej/article/view/46>. Date accessed: 30 Dec. 2013 (2009)
7. Martins-Bedê, F.T.: Souza Reis, M., Pantaleão, E., Dutra, L., Sandri, S. An application of multiple space nearest neighbor classifier in land cover classification Proc. IGARSS'14, pp 1713–1716 (2014)
8. Mendonça, J.H., Sandri, S., Martins-Bedê, F.T., Guimarães, R., Carvalho, O.: Training strategies for a fuzzy CBR cluster-based approach, Mathware & Soft Computing, v. 20, pp 42–49 (2013)
9. Recasens, J.: Indistinguishability operators, modelling fuzzy equalities and fuzzy equivalence relations. Series: Studies in Fuzziness and Soft Computing, vol 260, Springer Verlag (2011)
10. Sandri, S., Martins-Bedê, F.T.: A method for deriving order compatible fuzzy relations from convex fuzzy partitions. Fuzzy Sets and Systems, pp 91–103 (2014)
11. Theodoridis, S. and Koutroumbas, K.: Pattern recognition, Academic Press, 3rd edition (2006)

# Reducing information systems considering similarity relations

María José Benítez[1], Jesús Medina[1⋆], Dominik Ślęzak[1]

[1]Departamento de Matemáticas Universidad de Cádiz, Spain
Email: mariajose.benitez, jesus.medina@uca.es
[2]Institute of Mathematics, University of Warsaw. Poland
Email: slezak@mimuw.edu.pl

**Abstract.** Attribute reduction is an important step in order to decrease the computational complexity to deriving information from databases. In this paper, we extend the notions of reducts and bireducts introduced in rough sets theory for attribute reduction purposes and let them work with similarity relations defined on attributes values. Hence, the related mathematical concepts will be introduced and the characterizations of the new reducts and bireducts will be given in terms of the corresponding generalizations of the discernibility function.

## 1 Introduction

Fuzzy Set Theory (FST) introduced by Zadeh [**?**] and Rough Set Theory (RST) proposed by Pawlak [7], are complementary approaches to treating imperfect knowledge: meanwhile the first one considers a certain degree of truth given, in the second one the available information is incomplete. Specifically, in the absence of exact information about a set, it is represented by a pair of sets, which are the lower approximation and the upper approximation of the set.

Although in the original version proposed by Pawlak, the considered approximations were classical sets, there have been introduced some new variants in which the approximations could be fuzzy sets. A first definition, the rough fuzzy sets, was given by Fariñas del Cerro and Prade in the eighties [3].

A very important part is to reduce the size of the database, without losing information or elements of judgment. To this end various types of so called reducts were presented and studied in the RST-related literature [1, 4, 6]

Bireducts extend classical RST-based notions of reducts in order to provide more flexibility in operating with subsets of attributes and subsets of objects that those attributes can efficiently describe [5, 9, 10]. The main objective of the bireducts is to reduce the original system preventing the occurrence of incompatibilities and eliminating existing noise in the original data.

In this paper we study representations of bireducts both in the classical case and in situations when the notion of equality is weakened towards similarity.

⋆ Corresponding author.

The organization of the paper is the following: Some basics concepts related to the notion of similarity relation, the notions of $\delta$-similar and $\delta$-discordant are called in Section 2. Section 3 presents the basic definitions with Boolean in the new similarity enviroment. Conclusions and propects for future work are given in Section 4.

## 2 Preliminaries

In this paper the classical theory of propositional logic will be considered in order to interpret the expression of the discernibility function.

**Definition 1.** *A WFF is said to be in* disjunctive normal form *(DNF) if it is* $\top$, $\bot$, *a cube or a disjunction (possibly empty) of cubes.*

*A WFF is said to be in* conjunctive normal form *(CNF) if it is* $\top$, $\bot$, *a clause or conjunction (possibly empty) of clauses.*

The above normal forms may be reduced using absorption laws until none of them can be further reduced, obtaining the reduced forms:

**Definition 2.** *A DNF is said to be* restricted *(briefly, RDNF), if it satisfies that any cube contains a literal or its complementary and it does not contain repeated literals, and other cubes.*

*A CNF is said to be* restricted *(briefly, RCNF), if it satisfies that any clause contains a literal or its complementary literal and it does not contain repeated literals, and other clauses.*

The previous definitions are critical to introducing and managing discernibility function used in RST and will be generalized in this work to consider similarity relations. Now, we will recall the basic definitions of RST, the notion of similarity relation and its use on a decision system, which provides when two objects are $\delta$-similar and $\delta$-discordant, with respect to a threshold $\delta$.

**Definition 3.** *An* information system $(U, \mathcal{A})$ *is a tuple, where* $U = \{x_1, \ldots, x_n\}$ *and* $\mathcal{A} = \{a_1, \ldots, a_m\}$ *are finite, non-empty sets of objects and attributes, respectively. Each* $a$ *in* $\mathcal{A}$ *corresponds to a mapping* $\bar{a}\colon U \to V_a$, *where* $V_a$ *is the value set of* $a$ *over* $U$. *For every subset* $B$ *of* $\mathcal{A}$, *the* $B$*-indiscernibility relation*[1] $I_B$ *is defined as the equivalence relation*

$$I_B = \{(x_i, x_j) \in U \times U \mid \text{ for all } a \in B,\ \bar{a}(x_i) = \bar{a}(x_j)\}, \tag{1}$$

*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* $U/I_B = \{[x]_B \mid x \in U\}$.

In RST, data is represented as an information system. Given $A \subseteq U$, its lower and upper approximations w.r.t. $B$ are defined by

$$I_B{\downarrow}A = \{x \in X \mid [x]_B \subseteq A\} \tag{2}$$

$$I_B{\uparrow}A = \{x \in X \mid [x]_B \cap A \neq \varnothing\} \tag{3}$$

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

**Definition 4.** *A* decision system $(U, \mathcal{A} \cup \{d\})$ *is a special kind of information system, in which* $d \notin \mathcal{A}$ *is called the decision attribute, and its equivalence classes* $[x]_d$ *are called decision classes.*

A well-known approach to generate all reducts of a decision system is based on its discernibility matrix and function [8]. The discernibility matrix of $(U, \mathcal{A} \cup \{d\})$ is the $n \times n$ matrix $O$, defined by, for $i$ and $j$ in $\{1, ..., n\}$,

$$O_{ij} = \begin{cases} \varnothing & \text{if } d(x_i) = d(x_j) \\ \{a \in \mathcal{A} \mid \bar{a}(x_i) \neq \bar{a}(x_j)\} & \text{otherwise} \end{cases} \tag{4}$$

The *discernibility function* of $(U, \mathcal{A} \cup \{d\})$ is the map $f\colon \{0,1\}^m \to \{0,1\}$, defined by

$$f(a_1^*, ..., a_m^*) = \bigwedge \left\{ \bigvee O_{ij}^* \mid 1 \leq i < j \leq n \text{ and } O_{ij} \neq \varnothing \right\} \tag{5}$$

in which $O_{ij}^* = \{a^* \mid a \in O_{ij}\}$. The Boolean variables $a_1^*, \ldots, a_m^*$ correspond to the attributes from $\mathcal{A}$. It can be shown that the prime implicants of $f$ constitute exactly all decision reducts of $(U, \mathcal{A} \cup \{d\})$.

We continue recalling the definition of similarity relationship, which extends the notion of equivalence relation and therefore the concept of equality.

**Definition 5.** *Given an arbitrary set* $V$*, the mapping* $E\colon V \times V \to [0,1]$*, is called* similarity relation *if it is reflexive, symmetric and transitive.*

In theory, we can define a similarity relation over the set of objects in an arbitrary way. However, in practice it is indeed resonable to refer to values of objects for available attributes.

There are several possibilities to define a similarity relation on the set of objects $U$. One of the most popular ways is as follows:

$$E_U(i,j) = \bigwedge_{a \in A} (E_a(a(i), a(j))) \tag{6}$$

**Definition 6.** *Given an information system* $\mathbb{A} = (U, \mathcal{A})$ *and a similarity relation family* $\mathcal{E} = \{E_a\colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}\}$ *we say that objects* $i, j \in U$ *are* $\delta$***-similar*** *if for all* $a \in \mathcal{A}$ *we have*

$$\delta \leq E_a(a(i), a(j))$$

*with* $\delta \in [0,1]$*. Otherwise, we say that objects* $i, j \in U$ *are* $\delta$***-discordant****, that is, if the following holds:* $\{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \neq \varnothing$*.*

## 3 Generalization of reducts and bireducts by similarities with Boolean decision attribute

In this section a threshold $\delta \in [0,1]$ is fixed, from which we will use the notions of $\delta$-similar and $\delta$-discordant pairs of objects to define the generalization of the

discernibility function using similarity relations. Hence, an information system $\mathbb{A} = (U, \mathcal{A})$ and a similarity relation family $\mathcal{E} = \{E_a \colon V_a \times V_a \to [0,1] \mid a \in \mathcal{A}\}$ will also be fixed. Moreover, a linear ordering $\leq$ will also be fixed in $U$. Since the specific definition of the ordering is not important, any one can be considered. Given $i, j \in U$, we will say that $i < j$, if $i \leq j$ and they are not the same object.

First of all, the definitions of information reducts and bireducts are introduced.

**Definition 7.** *The set $B \subseteq \mathcal{A}$ is called* $\delta$-information reduct *if and only if it is an irreducible subset such that every pair $i, j \in U$, which is $\delta$-discordant by $\mathcal{A}$, is also $\delta$-discordant by $B$.*

**Definition 8.** *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 $B$.*

In this paper, the results will be focused on the general case of decision reducts and bireducts. The cases of $\delta$-information reducts and bireducts arise as "particular cases" of them and similar results hold analogously. In this case we will have to make a distinction based on decision attribute because we have a definition whether the attribute is a Boolean decision or not. This section will handle decision systems with a Boolean decision attribute. i.e., Boolean decision systems.

**Definition 9.** *Let $\mathbb{A} = (U, \mathcal{A} \cup \{d\})$ be a Boolean decision system. The subset $B \subseteq A$ is called* $\delta$-decision reduct *if and only if it is an irreducible subset such that all pair $i, j \in U$ is $\delta$-discordant by $B$ where $d(i) \neq d(j)$.*

Note that $d(i)$ and $d(j)$ are Boolean values. Next, the notion of decision bireduct is given.

**Definition 10.** *Let $\mathbb{A} = (U, \mathcal{A} \cup \{d\})$ be a Boolean decision system. The pair $(B, X)$, where $B \subseteq \mathcal{A}$ and $X \subseteq U$, is called* $\delta$-decision bireduct *if and only if every pair $i, j \in X$ is $\delta$-discordant by $B$ when $d(i) \neq d(j)$ and the following properties hold:*

1. *There is no $C \subsetneq B$ such that all pair $i, j \in X$ are $\delta$-discordant by $C$, where $d(i) \neq d(j)$.*
2. *There is no $X \subsetneq Y$ such that all pair $i, j \in Y$ are $\delta$-discordant by $B$, where $d(i) \neq d(j)$.*

Now, we are going to introduce the discernibility function in this general framework in order to obtain both $\delta$-decision reducts and bireducts. Since for $\delta$-decision reducts only the attributes are needed we will call it unidimensional $\delta$-discernibility function (uni $\delta$-d function) and for $\delta$-decision bireducts, both attributes and objects are considered and so, we will call it bidimensional $\delta$-discernibility function (bi $\delta$-d function).

**Definition 11.** *Let $\mathbb{A} = (U, A \cup \{d\})$ be a decision system, the* unidimensional $\delta$-discernibility function of $\mathbb{A}$*, is defined as the following conjunctive normal form (CNF):*

$$\tau_{\mathcal{A}}^{uni} = \bigwedge \left\{ \bigvee \{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \mid i, j \in U,\ d(i) \neq d(j) \right\}$$

*where the elements of $\mathcal{A}$ are the propositional symbols of the language. Also, we can denote it as:*

$$\tau_{\mathcal{A}}^{uni} = \bigwedge_{\{i,j \mid E_d(d(i),d(j)) < \delta\}} \left( \bigvee_{\{a \mid d(i) \neq d(j)\}} a \right)$$

Note that, although the condition $i < j$ is not considered in the definition, this can be considered without loss of generality, since any proper closure is removed: If $i = j$, then $E_a(a(i), a(j)) = 1 \not< \delta$ and so, this case does not arises any clause. If $j < i$, then the same clause for $i < j$ is provided and so, this will be removed when the DNF will be computed. This remark can be applied to the rest of discernibility functions introduced in this paper.

Therefore, the unidimensional $\delta$-discernibility function of $\mathbb{A}$ can be written as:

$$\tau_{\mathcal{A}}^{uni} = \bigwedge_{\{i,j \mid i<j, E_d(d(i),d(j)) < \delta\}} \left( \bigvee_{\{a \mid d(i) \neq d(j)\}} a \right)$$

Next, the characterization of the $\delta$-decision reducts is given.

**Theorem 1.** *Given a Boolean decision system $\mathbb{A} = (U, A \cup \{d\})$. An arbitrary set $B$, where $B \subseteq \mathcal{A}$, is a $\delta$-decision reduct of $\mathbb{A}$ if and only if the cube $\bigwedge_{b \in B} b$ is a cube in the RDNF of $\tau_{\mathcal{A}}^{uni}$.*

The following definition is the natural extension of the discernibility function expression to $\delta$-decision bireducts.

**Definition 12.** *Let $\mathbb{A} = (U, A \cup \{d\})$ be a decision system, the conjunctive normal form*

$$\tau_{\mathcal{A}}^{bi} = \bigwedge \{i \vee j \bigvee \{a \in \mathcal{A} \mid E_a(a(i), a(j)) < \delta\} \mid i, j \in U, i < j,\ d(i) \neq d(j)\}$$

*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 characterize the $\delta$-decision bireducts.

**Theorem 2.** *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 $\tau_{\mathbb{A}}^{bir}$.*

## 4 Conclusion and future work

We have studied the reducts and bireducts in the classic environment of RST and considering similarity relations. We have generalized discernibility function, from which we could get the reducts and bireducts in these environments.

The inclusion of the similarity relations in theory provides a greater flexibility in these environments, dramatically increasing the range of possible applications. Moreover, we have also considered the $\delta$-information reducts and bireducts in FCA, providing a new reduction method based on RST, which very close to the FCA framework.

As future work, we will extend the theory to obtain bireducts in fuzzy environments, such as in fuzzy rough sets [1, 2]. Moreover, we will study in depth in the relation between concept lattice reduction and rough set reduction considering similarity relations and in the general fuzzy case. Furthermore, we apply the theory developed in both theories to practical cases.

## References


1. C. Cornelis, R. Jensen, G. Hurtado, and D. Ślęzak. Attribute selection with fuzzy decision reducts. *Information Sciences*, 180:209–224, 2010.
2. 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.
3. L. Fariñas del Cerro and H. Prade. Rough sets, twofold fuzzy sets and modal logic—fuzziness in indiscernibility and partial information. In A. D. Nola and A. Ventre, editors, *The Mathematics of Fuzzy Systems*, pages 103–120. Verlag TUV Rheinland, 1986.
4. A. Janusz and D. Ślęzak. Rough set methods for attribute clustering and selection. *Applied Artificial Intelligence*, 28(3):220–242, 2014.
5. 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.
6. J. Medina. Relating attribute reduction in formal, object-oriented and property-oriented concept lattices. *Computers & Mathematics with Applications*, 64(6):1992–2002, 2012.
7. Z. Pawlak. Rough sets. *International Journal of Computer and Information Science*, 11:341–356, 1982.
8. A. Skowron and C. Rauszer. The discernibility matrices and functions in information systems. In R. Słowiński, editor, *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, pages 331–362. Kluwer Academic Publishers, 1992.
9. 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.
10. S. Stawicki and D. Ślęzak. Recent advances in decision bireducts: Complexity, heuristics and streams. *Lecture Notes in Computer Science*, 8171:200–212, 2013.

# Reasoning on Molecular Interaction Maps

Jean-Marc Alliot, Robert Demolombe, Luis Farinas, Martín Diéguez and Naji Obeid

University of Toulouse, CNRS, IRIT
Toulouse, France

**Abstract.** Metabolic networks, formed by a series of metabolic pathways, are made of intra-cellular and extracellular reactions that determine the biochemical properties of a cell, and by a set of interactions that guide and regulate the activity of these reactions. Cancer, for example, can sometimes appear in a cell as a result of some pathology in a metabolic pathway. Most of these pathways are formed by an intricate and complex network of chain reactions, and are often represented in *Molecular Interaction Maps* (MIM), a graphical, human readable form of the cell cycle checkpoint pathways. In this paper, we present a logic, called Molecular Interaction Logic, which semantically characterizes MIMs.

## 1 Introduction

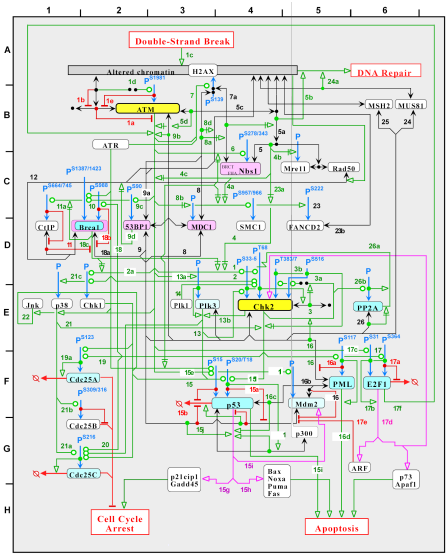
Metabolic networks, formed by series of metabolic pathways, are made of intra-cellular and extracellular reactions that determine the biochemical properties of a cell, and by a set of interactions that guide and regulate the activity of these reactions. These reactions can be positive (production of a new protein) or negative (inhibition of a protein in the cell). These reactions are at the center of a cell's existence, and are modulated by other proteins, which can either enable these reactions or, on the opposite, inhibit them.

Medical and pharmaceutical researches [11, 8] showed that the break of the double strand of DNA sometimes appear in a cell as a result of some pathology in a metabolic pathway, and double strand break (*dsb*) is a major cause of cancer.

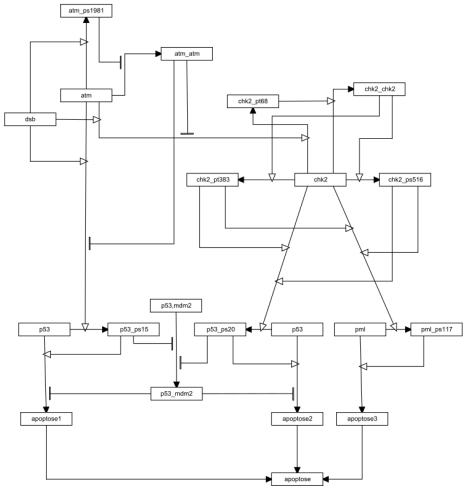
These pathways are used to investigate the molecular determinants of tumor response in cancers. The molecular parameters include the cell cycle checkpoint, DNA repair and apoptosis[1] pathways [15, 11, 8, 12, 14]. When DNA damage occurs, cell cycle checkpoints are activated and can rapidly kill the cell by apoptosis or arrest the cell cycle progression to allow DNA repair before cellular reproduction or division. Two important checkpoints that appear to function when parallel transduction cascades from DNA damage to the cell cycle checkpoint effectors are the $atm$-$chk2$ and the $atr$-$chk2$ pathways [15].

Most of these pathways are formed by an intricate and complex network of chain reactions, and are often represented in *Molecular Interaction Maps* (MIM), a human readable form of the cell cycle checkpoint pathways, such as the one in Figure 1(a), which represents the $atm$-$chk2$ and $atr$-$chk2$ pathways cited above.

[1] Apoptosis is the process of programmed cell death.




(a) $atm$-$chk2$/$atr$-$chk1$ molecular interaction map.




(b) $atm - chk2$ pathway

MIMs become increasingly larger and their density is constantly enriched with new information (references, date, authors, etc.). Although essential for knowledge capitalization and formalization, MIMs are difficult to use because of the very large number of elements involved as well as the inherent knowledge which, sometimes, is not formally described in the map.

In this paper we present a method to transform a MIM into a set of logical formulas. Subsets of Figure 1(a) will be used as examples, concentrating on the modelling of the $atm$-$chk2$ pathway leading to apoptosis.

The rest of this paper is organized as follows: section 2 introduces the concept of Molecular Interaction Maps and how they can be translated into a set of logical formulas. Section 3 describes Molecular Interaction Logic, a logic which is capable of describing and reasoning about general pathways and finally section 4 ends the paper with conclusions and future work.

## 2 Molecular Interaction Maps

A Molecular Interaction Map [10] (MIM) is a diagram convention which represents the interaction networks of multi-protein complexes, protein modifications and enzymes that are substrates of other enzymes. Although interactions between elements of a MIM can be complex, they can be represented using only three basic connectors: production

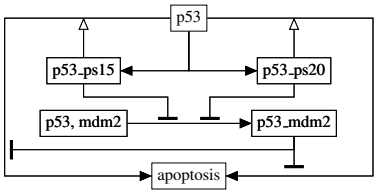
( $\rightarrow$ ), activation ($\multimap$) and inhibition ($\dashv$). Figure 1(b) presents the $atm$-$chk2$ pathway, an already pretty complex part of the large MIM of Figure 1(a), using only the aforementioned connectors.

A production relation means that a new substance is created as a result of a reaction on several primary components. For instance, the protein $atm$ can be dimerized to become the $atm\_atm$ protein or phosphorylated at serine 1981 resulting in the production of $atm\_ps1981$. These reactions can be triggered or blocked by other proteins or conditions. For example, in Figure 1(b), $atm\_ps1981$ blocks the dimerization of $atm$ into $atm\_atm$, while the double strand break ($dsb$) of DNA triggers the production of $atm\_ps1981$ by $atm$.

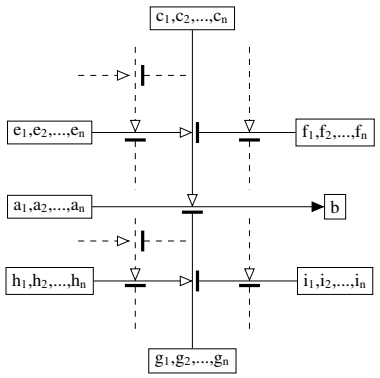
These interactions can be "stacked": for example, protein $p53$ can be phosphorylated at serine 15 to become $p53\_ps15$ (see Figure 1(b)). This reaction is triggered by $atm$, but the triggering itself has to be activated by $dsb$ and can be blocked by $atm\_atm$. Thus, the two main actions (production of a protein or inhibition of a protein) can be triggered or blocked by a stack of preconditions.

### 2.1 Translation of MIMs into formulas

Our first goal is to translate any MIM into a set of logical expressions in order to perform several automated reasoning tasks such as deduction or abduction. First, focusing on the diagram of Figure 1(c) (which corresponds to a sub-diagram of Figure 1(b)) will help getting an intuitive idea of how translation is performed.




(c) Apoptosis by $p53\_ps20$ and $p53\_ps15$ mediation.




(d) The general form of a basic production.

Here $apoptosis$ arises when protein $p53$ is phosphorylated at serine 20 or 15 (instances $p53\_ps20$ and $p53\_ps20$ respectively). However, $apoptosis$ would not happen if the dimer $p53\_mdm2$ is present. Thus the context would be *if* $p53$ *and either* $p53\_ps20$

*or* $p53\_ps15$ *are present and* $p53\_mdm2$ *is absent then* $apoptosis$ *is produced* (this example should of course be completed with the rules for producing the rest of objects in the diagram).

The general form of production relations is displayed in Figure 1(d).

Each arrow can be either an activation or an inhibition of the relation it applies to, and these activations/inhibitions can be stacked on any number of levels. The above examples give the idea behind the translation: it is a recursive process starting from the production relation and climbing up the tree. In order to formally describe the translation, the concept of *pathway context* is now defined:

**Definition 1 (Pathway context).** *Given a set of entities, a pathway context is formed by expressions defined by the following grammar:*

$$\alpha ::= \langle \alpha P \twoheadrightarrow, \alpha Q \dashv \rangle | \langle P \twoheadrightarrow, Q \dashv \rangle,$$

*where* $P$ *and* $Q$ *are sets (possibly empty) of propositional variables representing the conditions of activation (*$\twoheadrightarrow$*) or inhibition (*$\dashv$*) of the reaction. The first part of the pair is the activation context, the second part is the inhibition context. One, or both sets can be empty.* ■

For example, the $p53 \rightarrow apoptosis$ reaction of Figure 1(c) would lead to the following two pathway contexts:

$$\langle p53\_ps20 \twoheadrightarrow, p53\_mdm2 \dashv \rangle \quad (1)$$
$$\langle p53\_ps15 \twoheadrightarrow, p53\_mdm2 \dashv \rangle \quad (2)$$

**Definition 2 (Activation and inhibition expressions).** *Given a pathway context* $\alpha = \langle \alpha' P \twoheadrightarrow, \beta' Q \dashv \rangle$*, the* activation *and the* inhibition expressions *associated with the context* $\alpha$ *(denoted by* $A(\alpha)$ *and* $I(\alpha)$*) are defined recursively as:*

$$A(\alpha) = \bigwedge_{p \in P} p \wedge A(\alpha') \wedge (\bigvee_{q \in Q} \neg q \vee I(\beta')) \qquad I(\alpha) = \bigvee_{p \in P} \neg p \vee I(\alpha') \vee (\bigwedge_{q \in Q} q \wedge A(\beta'))$$

*The above expressions define the general forms of* $A(\alpha)$ *and* $I(\alpha)$*. If one part of the context* $\alpha$ *is empty, then the corresponding part is of course absent in* $A(\alpha)$ *and* $I(\alpha)$*.* ■

Following such definition, formulas associated with (1) are:

$$A((1)) = p53\_ps20 \wedge \neg p53\_mdm2 \qquad I((1)) = \neg p53\_ps20 \vee p53\_mdm2$$

**Definition 3 (Causal pathway formulas).** *A* causal pathway formula *is defined by the following grammar:*

$$F ::= [\alpha](p_1 \wedge \cdots \wedge p_n \rightarrow \mathbf{Pr}\ q) \mid [\alpha](p_1 \wedge \cdots \wedge p_n \rightarrow \mathbf{In}\ q) \mid F \wedge F,$$

*where $\alpha$ is a pathway context, $p_1, \cdots, p_n$, $q$ are propositional variables while $\mathbf{Pr}$ and $\mathbf{In}$ are modal concepts that qualify the process of activation or inhibition of proteins.* ■

Applied to the example of Figure 1(c), the causal pathway formula associated with the production rule $p53 \rightarrow apoptosis$ is

$$[(1)](p53 \rightarrow \mathbf{Pr}\ apoptosis) \wedge [(2)](p53 \rightarrow \mathbf{Pr}\ apoptosis)\,. \tag{3}$$

**Observation 1** *Each MIM can now be represented in terms of a causal pathway formula.* ■

## 3 Molecular Interaction Logic

In this section the semantics of the Molecular Interaction Logic (MIL) is formally introduced. This work extends a previous one [5, 4] where the MIMs were formalized via first order logic with equality, in which the pathway contexts were limited to one level of depth. From now on, $p$ means *protein p is present* and $\neg p$ means *protein p is absent*.

**Definition 4 (MIL interpretation).** *A MIL interpretation consists of a pair $(V_1, V_2)$ of classical evaluations i.e. $V : \mathcal{P} \rightarrow \{True, False\}$ where $\mathcal{P}$ is the set of propositional variables.* ■

The intuitive meaning behind these two evaluations correspond for $V_1$ to the protein present or absent, and for $V_2$ to the state of the protein resulting from the chemical reactions in the cell[2].

**Definition 5 (Satisfaction relation).** *Given a MIL interpretation $(V_1, V_2)$ and a formula $\alpha$, the satisfaction relation is defined as:*

*1)* $(V_1, V_2) \vDash p$ *iff* $V_1(p) = True$ *for* $p \in \mathcal{P}$

*2)* $\wedge$, $\vee$ *and* $\rightarrow$ *are satisfied as in classical logic.*

*3)* $(V_1, V_2) \vDash \mathbf{Pr}\ p$ *iff* $V_1(p) = V_2(p) = True$

*4)* $(V_1, V_2) \vDash \mathbf{In}\ p$ *iff* $V_1(p) = V_2(p) = False$

[2] If the semantics of the modal logic S5 is restricted to have at most two worlds then a strong normal form in which conjunctions and disjunctions are not in the scope of a modal operator can be found for this new logic [1]: the pathway causal formulas of MIL verify this condition.

*5)* $(V_1, V_2) \vDash [\alpha]F$ *iff* $(V_1, V_2) \nvDash A(\alpha)$ *or* $(V_1, V_2) \vDash F$

■

As usual, a formula $F$ is satisfiable if there is a model $(V_1, V_2)$ such that $(V_1, V_2) \vDash F$.

**Observation 2** *MIL can be characterized by the axioms of classical logic, plus the axioms:*

1. $[\alpha]F \leftrightarrow (A(\alpha) \rightarrow F)$
2. $\mathbf{Pr}\ p \rightarrow p$, *if* $p$ *is produced then* $p$ *is present*
3. $\mathbf{In}\ p \rightarrow \neg p$, *if* $p$ *is inhibited then* $p$ *is absent*

■

As a result of MIL semantics, the causal pathway formula (3) is logically equivalent to the conjunction of the following implications:

$$(p53 \wedge p53\_20 \wedge \neg p53\_mdm2) \rightarrow \mathbf{Pr}\ apoptosis \quad (4)$$

$$(p53 \wedge p53\_20 \wedge \neg p53\_mdm2) \rightarrow \mathbf{Pr}\ apoptosis \quad (5)$$

**Observation 3** *Any MIM can be transformed into a causal pathway formula, and every causal pathway formula is equivalent to a boolean composition of:*

- *propositional variables or their negation*
- *propositional variables qualified by* $\mathbf{Pr}$ *or* $\mathbf{In}$ *or their negation*

■

Axioms 2) and 3) of observation 2 have as consequence:

**Observation 4** *Given a MIM formula* $F$*, adding* $\mathbf{Pr}\ p \rightarrow p$ *and* $\mathbf{In}\ p \rightarrow \neg p$ *for each propositional variable* $p$ *in* $F$*, enables us to embbed MIL into classical logic.* ■

The notions of completion and production axioms, which are both important and implicit in MIMs, are presented first.

## 4 Conclusions and future work

We have presented a method to automatically translate MIMs into logical formulas, formalism that allows performing several kinds of reasoning such as deduction (in order to find inconsistencies in a representation) and abduction (which allows answering queries asked on MIMs). As a future work we want, on one hand, to enrich the language of MIL, with concepts like "aboutness" which are able to qualify, for example, proteins, allowing us to isolate the subgraph of a given MIM, regarding the qualified proteins. On the other hand, such enrichment of the language could include the introduction of temporal operators while to incorporate a temporal aspect to MIMs

## References


1. del Cerro, L.F., Herzig, A.: Contingency-based equilibrium logic. In: LPNMR'11. pp. 223–228 (2011)
2. Demolombe, R., Farinas, L.: An Inference Rule for Hypothesis Generation. In: IJCAI'91 (1991)
3. Demolombe, R., Farinas, L.: Information about a given entity: From semantics towards automated deduction. Journal of Logic and Computation 20(6), 1231–1250 (2010)
4. Demolombe, R., Farinas, L., Obeid, N.: Logical model for molecular interactions maps. In: Logical Modeling of Biological Systems, pp. 93–123. John Wiley & Sons (2014)
5. Demolombe, R., L. Fariñas, Obeid, N.: Translation of First Order Formulas into Ground Formulas via a Completion Theory. Journal of Applied Logic (), to appear
6. Erwig, M., Walkingshaw, E.: Causal reasoning with neuron diagrams. In: VLHCC '10. pp. 101–108 (2010)
7. Farinas, L., Inoue, K. (eds.): Logical Modeling of Biological Systems. John Wiley & Sons (2014)
8. Glorian, V., Maillot, G., Poles, S., Iacovoni, J.S., Favre, G., Vagner, S.: Hur-dependent loading of mirna risc to the mrna encoding the ras-related small gtpase rhob controls its translation during uv-induced apoptosis. Cell Death & Differentiation 18(11), 1692–70 (2011)
9. Inoue, K.: Linear resolution for consequence finding. Artificial Intelligence 56(2-3), 301 – 353 (1992)
10. Kohn, K.W., Aladjem, M.I., Weinstein, J.N., Pommier, Y.: Molecular interaction maps of bioregulatory networks: A general rubric for systems biology. Molecular Biology of the Cell 17(1), 1–13 (2006)
11. Kohn, K.W., Pommier, Y.: Molecular interaction map of the p53 and mdm2 logic elements, which control the off-on swith of p53 response to dna damage. Biochemical and biophysical research communications 331(3), 816–27 (2005)
12. Lee, W., Kim, D., Lee, M., Choi, K.: Identification of proteins interacting with the catalytic subunit of pp2a by proteomics. Proteomics 7(2), 206–214 (2007)
13. Muggleton, S., Bryant, C.H.: Theory completion using inverse entailment. In: ILP'00. pp. 130–146 (2000)
14. Pei, H., Zhang, L., Luo, K., Qin, Y., Chesi, M., Fei, F., Bergsagel, P.L., Wang, L., You, Z., Lou, Z.: MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. Nature 470(7332), 124–128 (2011)
15. Pommier, Y., Sordet, O., Rao, V.A., Zhang, H., Kohn, K.W.: Targeting chk2 kinase: molecular interaction maps and therapeutic rationale. Current pharmaceutical design 11(22), 2855–72 (2005)
16. Ray, O., Whelan, K., King, R.: Logic-based steady-state analysis and revision of metabolic networks with inhibition. In: CISIS'10. pp. 661–666 (2010)
17. Reiser, P.G., King, R.D., Kell, D.B., Muggleton, S., Bryant, C.H., Oliver, S.G.: Developing a logical model of yeast metabolism. Electronic Transactions in Artificial Intelligence 5, 233–244 (2001)
18. Rougny, A., Froidevaux, C., Yamamoto, Y., Inoue, K.: Analyzing SBGN-AF Networks Using Normal Logic Programs. In: Logical Modeling of Biological Systems, pp. 44–55. John Wiley & Sons (2014)