Advances in Intelligent Systems and Computing 643

Janusz Kacprzyk
Eulalia Szmidt
Slawomir Zadrożny
Krassimir T. Atanassov
Maciej Krawczak *Editors*

Advances in Fuzzy Logic and Technology 2017

Proceedings of: EUSFLAT-2017 – The 10th Conference of the European Society for Fuzzy Logic and Technology, September 11–15, 2017, Warsaw, Poland IWIFSGN'2017 – The Sixteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, September 13–15, 2017, Warsaw, Poland, Volume 3

Springer

Advances in Intelligent Systems and Computing

Volume 643

About this Series

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within “Advances in Intelligent Systems and Computing” are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/11156

Janusz Kacprzyk · Eulalia Szmidt
Slawomir Zadrożny · Krassimir T. Atanassov
Maciej Krawczak
Editors

Advances in Fuzzy Logic and Technology 2017

Proceedings of: EUSFLAT-2017 – The 10th Conference of the European Society for Fuzzy Logic and Technology, September 11–15, 2017, Warsaw, Poland IWIFSGN'2017 – The Sixteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, September 13–15, 2017, Warsaw, Poland, Volume 3

Springer

Editors
Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
Warsaw
Poland

Eulalia Szmidt
Systems Research Institute
Polish Academy of Sciences
Warsaw
Poland

Slawomir Zadrożny
Systems Research Institute
Polish Academy of Sciences
Warsaw
Poland

Krassimir T. Atanassov
Institute of Biophysics and Biomedical Engineering, Department of Bioinformatics and Mathematical Modelling
Bulgarian Academy of Sciences
Sofia
Bulgaria

Maciej Krawczak
WIT - Warsaw School of Information Technology
Warsaw
Poland

and

Systems Research Institute
Polish Academy of Sciences
Warsaw
Poland

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-66826-0 ISBN 978-3-319-66827-7 (eBook)
DOI 10.1007/978-3-319-66827-7

Library of Congress Control Number: 2017951306

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

This volume constitutes the proceedings of the two collocated international conferences. The main part includes the papers accepted, after a strict peer review process, for the presentation at, and for the inclusion in the proceedings of the 10th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-2017) held in Warsaw, Poland, on September 11–15, 2017. It is combined with the papers accepted, also after a strict peer review process, for the presentation at, and for the inclusion in the proceedings of the Sixteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN'2017) held in Warsaw, Poland, on September 13–15, 2017.

The EUSFLAT-2017 Conference was organized by the Systems Research Institute, Polish Academy of Science, Department IV of Engineering Sciences, Polish Academy of Sciences, and the Polish Operational and Systems Research Society. It is the 10th jubilee edition of the flagship conference of the European Society for Fuzzy Logic and Technology (EUSFLAT). The aim of the conference, in line with the mission of the EUSFLAT Society, is to bring together theoreticians and practitioners working on fuzzy logic, fuzzy systems, soft computing, and related areas and to provide for them a platform for the exchange of ideas, discussing newest trends and networking.

The papers included in the proceedings volume have been subject to a thorough review process by highly qualified peer reviewers. Comments and suggestion from them have considerably helped improve the quality of the papers but also the assignment of the papers to best suited sessions in the conference program. In the proceedings volume, the papers have been ordered alphabetically with respect to the name of the first author, and a convenient author's index is included at the end of the volume.

Thanks are due to many people and parties involved. First, in the early stage of the preparation of the conference general perspective, scope, topics, and coverage, we have received an invaluable help from the members of the International Committees of both conferences, notably the chairs responsible for various aspects of the conferences, as well as many people from the European Society for Fuzzy Logic and Technology (EUSFLAT). That help during the initial planning stage had resulted in a very attractive and up-to-date proposal of the scope and coverage that had clearly implied a considerable interest of the international research communities active in the areas covered who submitted a large number of very interesting and high-level papers. An extremely relevant role of the organizers of special sessions, competition, and other events should also be greatly appreciated. Thanks to their vision and hard work, we had been able to collect many papers on focused topics which had then resulted, during the conferences, in very interesting presentations and stimulating discussions at the sessions.

Though EUSFLAT-2017 is a subsequent edition of the main European conference on the broadly perceived fuzzy logic and technology, and an overwhelming majority of participants come from Europe, many people from other continents have also decided to submit their contributions. This has clearly resulted in a "globalization" of the EUSFLAT conferences which we have been able to increasingly notice since its founding. Of a particular importance in this respect is that among the plenary and keynote speakers, there are top researchers and scholars, as well as practitioners, not only from Europe but also from other continents.

The members of the Program Committee, together with the session organizers, and a group of other anonymous peer reviewers have undertaken a very difficult task of selecting the best papers, and they have done it excellently. They deserve many thanks for their great job for the entire community who is always concerned with quality and integrity. We also wish to thank the members of the EUSFLAT Board for their support throughout the organization process.

At the stage of the running of the conference, many thanks are due to the members of the Organizing Committee, chaired by Ms. Krystyna Warzywoda and Ms. Agnieszka Jóźwiak, and supported by their numerous collaborators.

And last but not least, we wish to thank Dr. Tom Ditzinger, Dr. Leontina di Cecco, and Mr. Holger Schaepe for their dedication and help to implement and finish this large publication project on time maintaining the highest publication standards.

June 2017 The Editors

Contents

Higher Degree Fuzzy Transform: Application to Stationary Processes and Noise Reduction

Linh Nguyen(✉) and Michal Holčapek

Institute for Research and Applications of Fuzzy Modelling, NSC IT4Innovations, University of Ostrava, 30. dubna 22, 701 03 Ostrava 1, Czech Republic
{Linh.Nguyen,Michal.Holcapek}@osu.cz

Abstract. In this contribution, we first elaborate the theory of the fuzzy transform of higher degree (F^m-transform, $m \geq 0$) applied to stationary processes that was initiated by Holčapek et al. in [5,6]. Then, we provide mathematical justification for its application to reduction of irregular fluctuations (noise) generated by specific stationary processes.

Keywords: Fuzzy transform · Stationary process · Noise reduction

1 Introduction

The fuzzy transform (F-transform) applied to stationary processes was first introduced in [5], and then generalized to the F^m-transform ($m \geq 1$) in [6]. These papers proposed techniques, based on the fuzzy transform of real-valued functions introduced by Perfilieva in [8,9], to analyze stationary processes. More precisely, under certain assumptions for stationary processes, the authors provided mathematical justification showing that the F^m-transform is a good technique for approximation as well as suppression of such stationary processes depending on specific settings of the used fuzzy partition. These abilities of the F^m-transform supports its application to reduction of noise in time series that is represented by a realization of a stationary process.

Recently, the theory of the F^m-transform has been widely investigated in [2,3,7] at which a new approach for representation of the direct F^m-transform was introduced based on monomial bases of approximation spaces. In this paper, we introduce the theory of the F^m-transform applied to stationary processes using this new approach. Especially, we prove approximation theorems that were not included in [6] for the higher degree fuzzy transform. We then consider application of the F^m-transform to reduction of specific types of noise usually exhibiting in time series. More precisely, under specific assumptions characterizing short-memory and long-memory stationary processes,[1] we prove that the irregular fluctuations generated by such stationary processes can be significantly reduced by the application of the F^m-transform with a reasonable adjustment of parameters of the used fuzzy partition.

[1] These are more natural than which were considered in the early investigations.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_1

2 Preliminaries

For any complex number $c \in \mathbb{C}$, we use $|c|$ to denote the absolute value of c. i.e., $|c| = (c \cdot \bar{c})^{\frac{1}{2}}$, where $\bar{c}$ is the complex conjugate of c.

2.1 Basic Concepts of Stationary Processes

Through out this paper, we restrict our analysis to the class of continuous-time, complex-valued stochastic processes $\xi(t)$, $t \in \mathbb{R}$ defined on a probability space $(\Omega, \mathscr{F}, P)$ where Ω is a sample space, $\mathscr{F}$ is a σ-algebra on Ω, and P is a probability measure defined on $\mathscr{F}$. Moreover, we use **E** and **Var** to denote the expectation and variance of a random variable, respectively. We also use **Cor** to denote the correlation between two random variables. Let X, Y be two complex-valued random variables. Then, $\mathbf{Var}(X) = \mathbf{E}\,|X - \mathbf{E}(X)|^2$ and $\mathbf{Cor}(X, Y) = \mathbf{E}\left(X \cdot \overline{Y}\right)$.

Definition 1. *A stochastic process $\xi(t)$ is said to be* a stationary process *if, for any $t \in \mathbb{R}$ it holds that $\mathbf{E}(\xi^2(t)) < \infty$, $\mathbf{E}(\xi(t))$ is constant and independent on t, and $\mathbf{Cor}(\xi(t), \xi(t+\tau))$ is independent on t for each τ.*

Below, we introduce a special function characterizing relation of random variables corresponding to different time moments of a stationary process.

Definition 2. *Let $\xi(t)$ be a stationary process. The* correlation function *of $\xi(t)$ is denoted by $\gamma(\cdot)$ and defined by $\gamma(\tau) = \mathbf{Cor}(\xi(t+\tau), \xi(t))$, $\tau \in \mathbb{R}$.*

Let us recall definitions of the mean-square convergence and the convergence in probability relating to stochastic processes.

Definition 3. *Let $\xi(t)$ and X be respectively a stochastic process and a random variable on the probability space $(\Omega, \mathscr{F}, P)$, and put $t_0 \in \mathbb{R}$*

(i) $\xi(t)$ converges in mean-square to X as t tends to t_0, denoted by $\mathbf{l.i.m}_{t \to t_0} \xi(t) = X$, if $\lim_{t \to t_0} \mathbf{E}|\xi(t) - X|^2 = 0$.

(ii) $\xi(t)$ converges in probability to X as t tends to t_0, denoted by $\lim_{t \to t_0} \xi(t) \overset{P}{=} X$, if, for any $\epsilon > 0$, $\lim_{t \to t_0} P\{|\xi(t) - X| > \epsilon\} = 0$.

Remark 1. *By the Chebyshev inequality that*

$$P\{|\xi(t) - X| > \epsilon\} \le \frac{\mathbf{E}\left(\xi(t) - X\right)^2}{\epsilon^2},$$

the mean-square convergence implies the convergence in probability.

Finally, to introduce the theory of the (higher degree) fuzzy transform applied to stationary processes, we need the mean-square integrals of the form

$$\int_a^b \xi(t) f(t) dt \tag{1}$$

where $f(t)$ is an arbitrary real function and $\xi(t)$ is a stationary process. For more details about the integral, we refer to [10]. A necessary and sufficient condition for the existence of (1) is the existence of the double integral

$$\int_a^b \int_a^b \gamma(t-s)f(t)f(s)dtds < \infty,$$

where γ is the correlation function of $\xi(t)$. In the sequel, when considering integrals of form (1), we always assume that the previous condition is fulfilled.

2.2 Generalized Uniform Fuzzy Partition

A fuzzy partition of an interval or the real line is one of the cores of the (higher degree) fuzzy transform. In this paper, we restrict ourselves to the generalized uniform fuzzy partitions (see [4]) formed by fuzzy sets determined by a generating function that are uniformly spread along the real line.

Definition 1. *A function $K : \mathbb{R} \to [0,1]$ is said to be a* generating function *if it is a continuous and even function that is non-increasing in $[0,\infty)$ and satisfies that $K(t) > 0$ iff $t \in (-1,1)$.*

Definition 2. *Let K be a generating function, and let h and r be positive real constants. Let $\mathcal{A} = \{A[h,r,k] \mid k \in \mathbb{Z}\}$ be a set of fuzzy sets on $\mathbb{R}$ determined by*

$$A[h,r,k](t) = K\left(\frac{t-kr}{h}\right), \qquad k \in \mathbb{Z}.$$

The set $\mathcal{A}$ is said to be a generalized uniform fuzzy partition of the real line determined by the triplet (K,h,r) *if the Ruspini condition is fulfilled, i.e.,*

$$\sum_{k\in\mathbb{Z}} A[h,r,k](t) = 1, \quad \text{for any } t \in \mathbb{R}. \tag{2}$$

The fuzzy set $A[h,r,k]$, $k \in \mathbb{Z}$ is called the k-th basic function *of the fuzzy partition $\mathcal{A}$.*

3 Higher Degree Fuzzy Transform Applied to Stationary Processes

Similarly to the (higher degree) fuzzy transform of (real or complex-valued) functions (see [2,3,7–9]), the fuzzy transform of higher degree applied to stationary processes consists in two phases: direct and inverse transformation. The former transforms a stationary process into a set of polynomial stochastic processes called the direct fuzzy transform components, and the latter provides a model for approximation of the original stationary process from its direct fuzzy transform.

3.1 Direct $\mathbf{F}^m$-Transform

We introduce the definition of the direct F^m-transform of a stationary process on the basis of the new representation devoted in [2,3,7] using monomial bases.

Definition 4. *Let $\xi(t)$ be a stationary process, $\mathcal{A} = \{A[h,r,k] \mid k \in \mathbb{Z}\}$ a generalized uniform fuzzy partition of the real line $\mathbb{R}$ determined by the triplet (K,h,r), and let $t_k = kr$ for any $k \in \mathbb{Z}$. The* direct F^m-transform, *$m \in \mathbb{N}$,* of *ξ with respect to $\mathcal{A}$ is the set of polynomial stochastic processes $\mathrm{F}^m_{\mathcal{A}}[\xi] = \{F^m_k[\xi](t) \mid k \in \mathbb{Z}\}$ where, for any $k \in \mathbb{Z}$,*

$$F^m_k[\xi](t) = C_{k,0} + C_{k,1}(t-t_k) + \ldots + C_{k,m}(t-t_k)^m, \quad t \in [t_k - h, t_k + h]$$

determined by

$$(C_{k,0}, C_{k,1}, \ldots, C_{k,m})^T = (\mathcal{H}_m)^{-1} \cdot (\mathcal{Z}_m)^{-1} \cdot \mathcal{Y}_{n,k}$$

where $\mathcal{H}_m = \mathrm{diag}(1, h, \ldots, h^m)$, $\mathcal{Z}_m = (Z_{ij})_{i,j=\overline{1,(m+1)}}$ defined by

$$Z_{ij} = \int_{-1}^{1} t^{i+j-2} K(t) dt, \quad i,j = 1, \ldots, m+1,$$

and $\mathcal{Y}_{m,k} = (Y_{k,1}, \ldots, Y_{k,m+1})^T$ defined by

$$Y_{k,\ell} = \int_{-1}^{1} \xi(th + t_k) \cdot t^{\ell-1} K(t) dt, \quad \ell = 1, \ldots, m+1. \tag{3}$$

*The stochastic process $F^m_k[\xi](t)$, $t \in [t_k - h, t_k + h]$ is called the k-*th component of the direct F^m-transform of *ξ.*

By the linearity property of the mean-square integral in (3) with respect to the stationary process $\xi(t)$, the direct F^m-transform satisfies the linearity property. Namely, for arbitrarily two stationary processes $\xi(t)$ and $\eta(t)$ and $a, b \in \mathbb{C}$. Then,

$$F^m_k[a\xi + b\eta] = aF^m_k[\xi] + bF^m_k[\eta], \quad \text{for any } k \in \mathbb{Z}. \tag{4}$$

In the sequel, we use $\boldsymbol{\mathcal{A}}[K]$ to denote a family of generalized uniform fuzzy partitions of the real line determined by the triplet (K,h,r) for $h, r \in (0,\infty)$ such that the ratio h/r is constant.

Below, we provide statistical properties of the direct F^m-transform components that need to prove subsequent approximation theorems.

Lemma 1. *Let $\xi(t)$ be a stationary process such that $\mathbf{E}(\xi(t)) = 0$. Let $\mathrm{F}^m_{\mathcal{A}}[\xi]$ be the direct F^m-transform of $\xi(t)$ with respect to a generalized fuzzy partition $\mathcal{A}$. Then, for any $k \in \mathbb{Z}$, $t \in \mathrm{Supp}\, A[K,h,r]$, it holds that $\mathbf{E}\left(F^m_k[\xi](t)\right) = 0$.*

Sketch of Proof. The proof is obviously obtained by Definition 4 and the assumption that $\mathbf{E}(\xi(t)) = 0$. □

Lemma 2. *Let $\xi(t)$ be a stationary process such that $\mathbf{E}(\xi(t)) = 0$, $\mathbf{Var}(\xi(t)) = \sigma^2$ and its correlation function is continuous at the origin. For any $h \in (0, \infty)$ and $k \in \mathbb{Z}$, let $F^m_{k,(K,h)}[\xi]$ be the k-th component of the direct F^m-transform of $\xi(t)$ with respect to a fuzzy partition $\mathcal{A} = \{A[h, r, k] \mid k \in \mathbb{Z}\} \in \boldsymbol{\mathcal{A}}[K]$. Then, for any $\epsilon > 0$,*

(i) there exists $h_0 > 0$ such that for any $0 < h \le h_0$, $k, k' \in \mathbb{Z}$, $|k-k'| < 2h/r$,[2] *then*

$$\left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right) - \sigma^2\right| < \epsilon,$$

for any $t \in \mathrm{Supp}\, A[h, r, k] \cap \mathrm{Supp}\, A[h, r, k']$,

(ii) there exists $h_0 > 0$ such that for any $0 < h \le h_0$, $k \in \mathbb{Z}$, then

$$\left|\mathbf{Var}\left(F^m_{k,(K,h)}[\xi](t)\right) - \sigma^2\right| < \epsilon,$$

for any $t \in \mathrm{Supp}\, A[h, r, k]$,

(iii) there exists $h_0 > 0$ such that for any $0 < h \le h_0$, $k \in \mathbb{Z}$, then

$$\left|\mathbf{Cor}\left(\xi(t), F^m_{k,(K,h)}[\xi](t)\right) - \sigma^2\right| < \epsilon,$$

for any $t \in \mathrm{Supp}\, A[h, r, k]$.

Sketch of Proof. (i) For any $h \in (0, \infty)$, $k, k' \in \mathbb{Z}$, $|k-k'| < 2h/r$, by Definition 4, we obtain that

$$\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right) = \sum_{i,j,p,\iota=0}^{m} \left(\frac{t-t_k}{h}\right)^i \cdot \left(\frac{t-t_{k'}}{h}\right)^j \cdot$$
$$V_{i+1p+1} \cdot V_{j+1\iota+1} \int_{-1}^{1}\int_{-1}^{1} \gamma\left((t-s)h + (k-k')r\right) \cdot t^p s^\iota K(t)K(s)dtds$$

for any $t \in \mathrm{Supp}\, A[h, r, k] \cap \mathrm{Supp}\, A[h, r, k']$ where $(V_{\ell j})_{\ell,j=\overline{1,m+1}} = (\mathcal{Z}_m)^{-1}$ and $\gamma(\cdot)$ is the correlation function of $\xi(t)$. The proof is then obtained by application of Lebesgue's dominated convergence theorem with respect to the fact that $\lim_{h\to 0} \gamma\left((t-s)h + (k-k')r\right) = \gamma(0) = \sigma^2$.

(ii) It is straightforward consequence of (i) corresponding to $k = k'$ and Lemma 1.

(ii) Similarly to the proof of (i), this proof is obtained by the fact that

$$\mathbf{Cor}\left(\xi(t), F^m_{k,(K,h)}[\xi](t)\right) = \sum_{i,j=0}^{m} \left(\frac{t-t_k}{h}\right)^i \cdot V_{i+1j+1} \int_{-1}^{1} \gamma(t - t_k - sh) \cdot s^j K(s)ds,$$

and the application of Lebesgue's dominated convergence theorem based on the fact that $\lim_{h\to 0} \gamma(t - t_k - sh) = \gamma(0) = \sigma^2$. □

[2] This inequality is equivalent to the fact that $\mathrm{Supp}\, A[h, r, k] \cap \mathrm{Supp}\, A[h, r, k'] \neq \phi$.

Lemma 3. *Let the assumptions of Lemma 2 be satisfied. Then, for any $\epsilon > 0$,*

(i) there exists $h_0 > 0$ such that for any $0 < h \le h_0$, $k, k' \in \mathbb{Z}$, $|k - k'| < 2h/r$, then
$$\left|\mathbf{Cor}\left(\xi(t) - F^m_{k,(K,h)}[\xi](t), \xi(t) - F^m_{k',(K,h)}[\xi](t)\right)\right| < \epsilon,$$
for any $t \in \operatorname{Supp} A[h, r, k] \cap \operatorname{Supp} A[h, r, k']$,

(ii) there exists $h_0 > 0$ such that for any $0 < h \le h_0$, $k \in \mathbb{Z}$, then
$$\mathbf{Var}\left(\xi(t) - F^m_{k,(K,h)}[\xi](t)\right) < \epsilon,$$
for any $t \in \operatorname{Supp} A[h, r, k]$.

Sketch of Proof. (i) The proof is obtained by application of Lemma 2 and the fact that for any $h \in (0, \infty)$, $k, k' \in \mathbb{Z}$, $|k - k'| < 2h/r$, then

$$\begin{aligned}&\left|\mathbf{Cor}\left(\xi(t) - F^m_{k,(K,h)}[\xi](t), \xi(t) - F^m_{k',(K,h)}[\xi](t)\right)\right| \\ &\quad\le \left|\sigma^2 - \mathbf{Cor}\left(\xi(t), F^m_{k',(K,h)}[\xi](t)\right)\right| + \left|\sigma^2 - \mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), \xi(t)\right)\right| \\ &\qquad + \left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right) - \sigma^2\right|\end{aligned}$$

for any $t \in \operatorname{Supp} A[h, r, k] \cap \operatorname{Supp} A[h, r, k']$.

(ii) is a straightforward consequence of (i) for which $k = k'$. □

3.2 Inverse F^m-Transform

In this subsection, we provide a model to approximate a stationary process from its direct F^m-transform components.

Definition 5. *Let $\mathcal{A} = \{A[h, r, k] \mid k \in \mathbb{Z}\}$ be a generalized uniform fuzzy partition of $\mathbb{R}$. Let $\mathrm{F}^m_{\mathcal{A}}[\xi] = \{F^m_k[\xi] \mid k \in \mathbb{Z}\}$ be the direct F^m-transform of a stationary process ξ with respect to $\mathcal{A}$. The stochastic process*

$$\widehat{\xi}^m_{\mathcal{A}}(t) = \sum_{k \in \mathbb{Z}} F^m_k[\xi](t) \cdot A[h, r, k](t), \quad t \in \mathbb{R} \tag{5}$$

is called the inverse F^m-transform *of ξ with respect to the direct F^m-transform $\mathrm{F}^m_{\mathcal{A}}[\xi]$ and the fuzzy partition $\mathcal{A}$.*

The following Corollary is a straightforward consequence of Lemma 1.

Corollary 4. *Let $\xi(t)$ be a stationary process with $\mathbf{E}(\xi(t)) = 0$. Let $\mathrm{F}^m_{\mathcal{A}}[\xi] = \{F^m_k[\xi] \mid k \in \mathbb{Z}\}$ be the direct F^m-transform of $\xi(t)$ with respect to a generalized uniform fuzzy partition $\mathcal{A}$ of the real line $\mathbb{R}$, and $\widehat{\xi}^m_{\mathcal{A}}$ the inverse F^m-transform of ξ with respect to $\mathrm{F}^m_{\mathcal{A}}[\xi]$ and $\mathcal{A}$. Then, $\mathbf{E}\left(\widehat{\xi}^m_{\mathcal{A}}(t)\right) = 0$, for any $t \in \mathbb{R}$.*

Additionally, by the linearity property of the direct F^m-transform showed in (4), the inverse F^m-transform preserves the linearity property as well. Namely, for arbitrarily two stationary processes ξ and η, and two complex numbers a, b, then $\widehat{(a\xi + b\eta)}_{\mathcal{A}}^{m} = a\widehat{\xi}_{\mathcal{A}}^{m} + b\widehat{\eta}_{\mathcal{A}}^{m}$.

In the following, we provide a very important theorem showing that a stationary process can be approximated with arbitrary precision by the inverse F^m-transform stochastic process.

Theorem 5. *Let $\xi(t)$ be a stationary process such that its correlation function is continuous at the origin. For any $h \in (0, \infty)$, let $\mathrm{F}^m_{(K,h)}[\xi]$ be the direct F^m-transform of $\xi(t)$ with respect to a fuzzy partition $\mathcal{A} = \{A[h, r, k] \mid k \in \mathbb{Z}\} \in \boldsymbol{\mathcal{A}}[K]$. Let $\widehat{\xi}^m_{(K,h)}$ be the inverse F^m-transform of $\xi(t)$ with respect to $\mathrm{F}^m_{(K,h)}[\xi]$ and $\mathcal{A}$. Then,*

$$\mathbf{l.i.m.}_{h \to 0}\, \widehat{\xi}^m_{(K,h)}(t) = \xi(t), \quad \textit{for any } t \in \mathbb{R}. \tag{6}$$

Sketch of Proof. Let $h \in (0, \infty)$ be arbitrarily, and assume that

$$\mathrm{F}^m_{(K,h)}[\xi] = \left\{ F^m_{k,(K,h)}[\xi](t) \mid k \in \mathbb{Z} \right\}.$$

For $t \in \mathbb{R}$, let $\Lambda_h(t) = \{k \in \mathbb{Z} \mid A[h, r, k](t) \neq 0\}$. By the assumption that the ratio h/r is a constant, the number of elements in $\Lambda_h(t)$ is limited (bounded). Moreover, $|k - k'| < 2h/r$ for any $k, k' \in \Lambda_h(t)$. From Definition 5, we find that

$$\mathbf{E}\left|\widehat{\xi}^m_{(K,h)}(t) - \xi(t)\right|^2 \leq \sum_{k,k' \in \Lambda_h} \left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t) - \xi(t), F^m_{k',(K,h)}[\xi](t) - \xi(t)\right)\right| \cdot A[h, r, k](t) A[h, r, k'](t).$$

The proof is then obtained by application of Lemma 3 to this inequality. □

4 Reduction of Noise

In this section, we devote theoretical justification for application of the F^m-transform to reduction of noise generated by specific types of stationary process (e.g. the noise in a time series is standardly assumed to be a realization of a stationary process). Let $\xi(t)$ be a stationary process and $\gamma(\cdot)$ be its correlation function. In the sequel, we restrict our analysis on two following assumptions. The first one supposes that $\xi(t)$ is a "short-memory" stationary process characterized by the quick decay of $\gamma(\cdot)$. This assumption is formalized by the following statement:

$$\lim_{h \to \infty} \frac{1}{h} \cdot \int_0^h |\gamma(\tau)| d\tau = 0. \tag{7}$$

The second one assumes that

$$\gamma(\tau) = \sum_{j=1}^{\kappa} A_j e^{\mathrm{i}\lambda_j \tau} \tag{8}$$

where $A_j \in \mathbb{C}$, i is the imaginary unit, and $\lambda_j \in \mathbb{R}$. This assumption characterizes the periodicity property of the correlation function. Moreover, a correlation function, which slowly decays, can be approximately represented by the form of (8) (see [10]). Therefore, this assumption can be considered as a characterization of "long-memory" stationary processes.[3]

In the following part, we prove that any stationary process satisfying one of two previous assumptions can be significantly reduced in the sense of reduction of its variability by application of the F^m-transform.

Lemma 6. *Let $\xi(t)$ be a short-memory stationary process with $\mathbf{E}(\xi(t)) = 0$. For any $h \in (0, \infty)$ and $k \in \mathbb{Z}$, let $F^m_{k,(K,h)}[\xi]$ be the k-th component of the direct F^m-transform of ξ with respect to a fuzzy partition $\mathcal{A} = \{A[h, r, k] \mid k \in \mathbb{Z}\} \in \boldsymbol{\mathcal{A}}[K]$. Then, for any $\epsilon > 0$,*

(i) there exists $h_0 > 0$ such that for any $h \geq h_0$, $k, k' \in \mathbb{Z}$, $|k - k'| < 2h/r$, then

$$\left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right)\right| < \epsilon$$

for any $t \in \operatorname{Supp} A[h, r, k] \cap \operatorname{Supp} A[h, r, k']$,

(ii) there exists $h_0 > 0$ such that for any $h \geq h_0$, $k \in \mathbb{Z}$, then

$$\mathbf{Var}\left(F^m_{k,(K,h)}[\xi](t)\right) < \epsilon$$

for any $t \in \operatorname{Supp} A[h, r, k]$.

Sketch of Proof. (i) Let $(V_{ij})_{i,j=\overline{1,m+1}} = (\mathcal{Z}_m)^{-1}$, and let $h \in (0, \infty)$ be arbitrarily and let $k, k' \in \mathbb{Z}$ such that $|k - k'| < 2h/r$. By the similar arguments used in the proof of Lemma 2 together with the changing of variables of integrals, for any $t \in \operatorname{Supp} A[h, r, k] \cap \operatorname{Supp} A[h, r, k']$, we find that

$$\left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right)\right| \leq 4 \cdot \sum_{i,j,p,\iota=0}^{m} |V_{i+1p+1} \cdot V_{j+1\iota+1}| \cdot \int_0^{2h+|k-k'|r} \frac{|\gamma(u)|}{h} du.$$

Since $\xi(t)$ is a short-memory stationary process satisfying the assumption in (7), it is easy to prove the desired statement.

(ii) is a straightforward consequence of (i). □

[3] We use the name "short-memory" and "long-memory" stationary process only to distinguish two considered assumptions. And these names, somehow, characterize the behavior of the correlation function. By the aim of this paper, we do not further investigate the same concepts used in probability and statistic (see [1]).

Lemma 7. *Let the assumptions of Lemma 6 be satisfied. The only one difference is that $\xi(t)$ is now a long-memory stationary process. Then, both of consequences of Lemma 6 are satisfied.*

Sketch of Proof. By the same analysis used in the proof of Lemma 2, we find that

$$\left|\mathbf{Cov}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right)\right| \leq \sum_{i,j,p,\iota=0}^{m} |V_{i+1p+1} \cdot V_{j+1\iota+1}| \cdot |I_{p,\iota}(h)|, \quad (9)$$

where $I_{p,\iota}(h) = \int_{-1}^{1}\int_{-1}^{1} \gamma\left((t-s)h + (k-k')r\right) \cdot t^p s^\iota K(t)K(s)dtds$, for any $p, \iota = 0, 1, \ldots, m$. By the assumption (8), it is easy to prove that $\lim_{h\to\infty} |I_{p,\iota}(h)| = 0$. The proof is then a straightforward consequence of these results. □

Theorem 8. *Let $\xi(t)$ be a stationary process with $\mathbf{E}(\xi(t)) = 0$. For any $h \in (0, \infty)$, let $\mathrm{F}^m_{(K,h)}[\xi]$ be the direct F^m-transform of $\xi(t)$ with respect to a fuzzy partition $\mathcal{A} = \{A[h, r, k] \mid k \in \mathbb{Z}\} \in \boldsymbol{\mathcal{A}}[K]$. Let $\widehat{\xi}^m_{(K,h)}$ be the inverse F^m-transform of $\xi(t)$ with respect to $\mathrm{F}^m_{(K,h)}[\xi]$ and $\mathcal{A}$. Suppose that correlation function $\gamma(\cdot)$ of $\xi(t)$ satisfies the assumption (7) or (8). Then, for any $t \in \mathbb{R}$, it holds that*

(i) $\lim_{h\to\infty} \mathbf{Var}\left(\widehat{\xi}^m_{(K,h)}(t)\right) = 0,$

(ii) $\mathbf{l.i.m}_{h\to\infty}\, \widehat{\xi}^m_{(K,h)}(t) = 0.$

Sketch of Proof. (i) Let $h \in (0, \infty)$ be arbitrarily, and assume that

$$\mathrm{F}^m_{(K,h)}[\xi] = \left\{F^m_{k,(K,h)}[\xi](t) \mid k \in \mathbb{Z}\right\}.$$

For any $t \in \mathbb{R}$, let $\Lambda_h(t) = \{k \in \mathbb{Z} \mid A[h, r, k](t) \neq 0\}$. Then, by the same analysis of Theorem 5 and from Corollary 4, we find that

$$\mathbf{Var}\left(\widehat{\xi}^m_{(K,h)}(t)\right) = \mathbf{E}\left|\widehat{\xi}^m_{(K,h)}(t)\right|^2 \leq \sum_{k,k'\in\Lambda_h(t)} \left|\mathbf{Cor}\left(F^m_{k,(K,h)}[\xi](t), F^m_{k',(K,h)}[\xi](t)\right)\right| \cdot A[h, r, k](t)A[h, r, k'](t).$$

By the correlation function of $\xi(t)$ satisfies the assumption (7) or (8), from Lemmas 6 and 7, we obtain statement (i).

(ii) is a straightforward consequence of (i). □

Let the assumptions of Theorem 8 be satisfied and let $R(t)$ be a realization of $\xi(t)$, i.e., $R(t) = \xi(\boldsymbol{\omega}, t)$ where $\boldsymbol{\omega}$ is an elementary event ($\boldsymbol{\omega} \in \Omega$) of a probability space $(\Omega, \mathscr{F}, P)$ on which ξ is modeled. It follows from Remark 1 that it is in an arbitrarily high probability value that $\lim_{h\to\infty} \widehat{R}^m_{(K,h)}(t) = 0$, for any $t \in \mathbb{R}$. In other words, the irregular fluctuations that are realizations of the stationary process $\xi(t)$ can be significantly reduced using F^m-transform technique by a reasonable setting of the bandwidth of the used fuzzy partition (i.e., the bandwidth is large enough).

5 Illustrative Examples

In this section, we provide examples demonstrating that the F^m-transform is an efficient technique for reduction of noise usually exhibiting in time series. For the sake of simplicity, we only use triangle type of the generalized uniform fuzzy partitions. Namely, we use the generating functions of the form $\alpha \cdot K^{tr}$ where $\alpha \in (0,1]$, and $K^{tr}(t) = \max\{0, 1 - |t|\}$.

Let us consider the following stationary processes:

1. $\xi_1(t) = \varepsilon(t) + 0.6\varepsilon(t-1) + 0.6\varepsilon(t-2) + 0.3\varepsilon(t-3) + 0.7\varepsilon(t-4)$, where $\varepsilon(t) \sim WN(0, 9)$,
2. $\xi_2(t) = \varepsilon_1 \sin\left(\frac{\pi}{6}t\right) + \varepsilon_2 \cos\left(\frac{\pi}{6}t\right) + \eta_1 \sin\left(\frac{2\pi}{3}t\right) + \eta_2 \cos\left(\frac{2\pi}{3}t\right)$, where $\varepsilon_1, \varepsilon_2, \eta_1, \eta_2$ are independent random variables having the same normal distribution with $\varepsilon_i \sim \mathcal{N}(0, 9)$ and $\eta_i \sim \mathcal{N}(0, 16)$, $i = 1, 2$.

It is easy to see that $\xi_1(t)$ and $\xi_2(t)$ satisfy the assumptions (7) and (8), respectively. Indeed, the correlation function of $\xi_1(t)$ cuts off after the order $p = 4$, and the correlation function of $\xi_2(t)$ is periodic. For each process $\xi_i(t)$, $i = 1, 2$, we generate three noise data $R_{ij}(t)$, $j = 1, 2, 3$ corresponding to $t = 0, \ldots, 100$. These are characterized by the mean and variance in Table 1. We apply the F^2-transform (direct and inverse phase) with respect to the generalized uniform fuzzy partition determined by the triplet $(0.5 \cdot K^{tr}, 20, 10)$ to each of R_{1j}, $j = 1, 2, 3$. The obtained results are depicted in Fig. 1. Additionally, we apply the F^1-transform with respect to the generalized uniform fuzzy partition determined by the triplet $(0.5 \cdot K^{tr}, 8, 4)$ to the noise R_{2j}, $j = 1, 2, 3$. The obtained results are depicted in Fig. 2. From Figs. 1 and 2, it is easy to see that the variabilities of these noise data are significantly reduced by the application of fuzzy transform technique. More evidences for this evaluation can be found in Table 2 at which the variance of the obtained results is powerfully decreased.

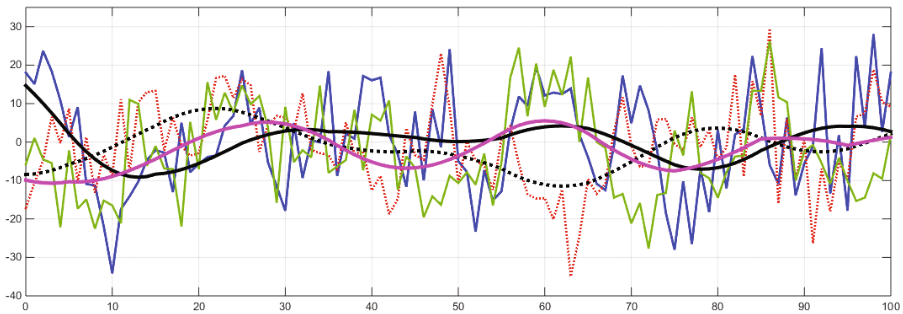


Fig. 1. Inverse F^2-transforms (black, dotted black, and magenta lines) of the realizations 1, 2, and 3 (blue, dotted red, and dark-green lines), respectively.

Table 1. Mean and variance of the noise data.

Process	$\xi_1(t)$			$\xi_2(t)$		
Noise data	R_{11}	R_{12}	R_{13}	R_{21}	R_{22}	R_{23}
Mean	0.1989	−0.9763	−2.4236	−0.2882	0.8262	−0.1278
Variance	171.3627	124.9096	143.4554	705.7320	3042.8487	244.9710

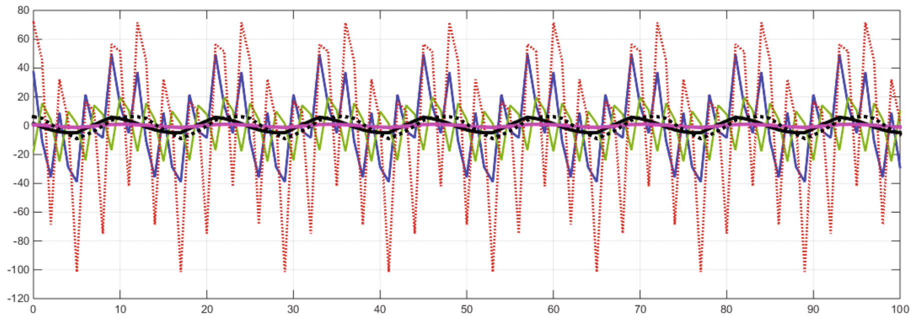


Fig. 2. Inverse F^1-transforms (black, dotted black, and magenta lines) of the realizations 1, 2, and 3 (blue, dotted red, and dark-green lines), respectively.

Table 2. Mean and variance of the inverse F^m-transforms.

Inverse F^m-transform	$\hat{R}^2_{11}$	$\hat{R}^2_{12}$	$\hat{R}^2_{13}$	$\hat{R}^1_{21}$	$\hat{R}^1_{22}$	$\hat{R}^1_{23}$
Mean	−0.1715	−1.3579	−1.9961	−0.1137	0.0356	−0.0067
Variance	17.6358	27.4762	22.2306	13.4436	27.2926	0.5136

6 Conclusions

In this paper, we provided a new approach for the representation of the direct F^m-transform of stationary processes. From the advantage of this approach, we carefully investigated essential properties of the F^m-transform. Furthermore, we proved that it is a good technique for reduction of certain types of noise usually exhibiting in time series that are generated by specific stationary processes. This makes it possible to apply the F^m-transform to time series analysis (or particularly, to time series decomposition, see [2]).

Acknowledgments. This work was supported by the project LQ1602 IT4Innovations excellence in science. The additional support was provided by the Czech Science Foundation through the project of No.16-09541S.

References

1. Beran, J., Feng, Y., Ghosh, S., Kulik, R.: Long-Memory Processes: Probabilistic Properties and Statistical Methods, 1st edn. Springer, Heidelberg (2013)
2. Holčapek, M., Nguyen, L.: Trend-cycle estimation using fuzzy transform of higher degree. Iranian J. Fuzzy Syst. (2017)
3. Holčapek, M., Nguyen, L., Tichý, T.: Polynomial alias higher degree fuzzy transform of complex-valued functions. Fuzzy Sets Syst. (2016). doi:10.1016/j.fss.2017.06.011, (submitted)
4. Holčapek, M., Perfilieva, I., Novák, V., Kreinovich, V.: Necessary and sufficient conditions for generalized uniform fuzzy partition. Fuzzy Sets Syst. **277**, 97–121 (2015)
5. Holčapek, M., Novák, V., Perfilieva, I.: Analysis of stationary processes using fuzzy transform. In: Proceedings of European Society for Fuzzy Logic and Technology, pp. 714–721 (2013)
6. Holčapek, M., Novák, V., Perfilieva, I.: Noise reduction in time series using F-transform. In: Proceedings of IEEE International Conference on Fuzzy Systems (2013). doi:10.1109/FUZZ-IEEE.2013.6622492
7. Nguyen, L., Holčapek, M., Novák, V.: Multivariate fuzzy transform of complex-valued functions determined by monomial basis. Soft Computing **21**(13), 3641–3658 (2017)
8. Perfilieva, I.: Fuzzy transforms: theory and applications. Fuzzy Sets Syst. **157**, 993–1023 (2006)
9. Perfilieva, I., Daňková, M., Bede, B.: Towards a higher degree F-transform. Fuzzy Sets Syst. **180**, 3–19 (2011)
10. Yaglom, A. M.: An introduction to the theory of stationary random functions. Revised English ed. Translated and edited by R.A. Silverman. Prentice-Hall, Inc. XIII, Englewood Cliffs, NJ (1962)

Sheffer Stroke Fuzzy Implications

Wanda Niemyska[1], Michał Baczyński[1(✉)], and Szymon Wąsowicz[2]

[1] Institute of Mathematics, University of Silesia in Katowice, Bankowa 14, 40-007 Katowice, Poland
{wanda.niemyska,michal.baczynski}@us.edu.pl

[2] Department of Mathematics, University of Bielsko-Biała, Willowa 2, 43-309 Bielsko-Biała, Poland
swasowicz@ath.bielsko.pl

Abstract. A new family of fuzzy implications, motivated by classic Sheffer stroke operator, is introduced. Sheffer stroke, which is a negation of a conjunction and is called NAND as well, is one of the two operators that can be used by itself, without any other logical operators, to constitute a logical formal system. Classical implication can be presented just by Sheffer stroke operator in two ways which leads to two new families of fuzzy implication functions. It turns out that one of them is mainly a subclass of QL-operations, while the other one, called in our paper as SS_{qq}-implications, is independent of other well-known families of fuzzy implications. Basic properties of Sheffer stroke implications are also analysed.

Keywords: Fuzzy connectives · Fuzzy implication · Sheffer stroke · NAND

1 Introduction

The basic fuzzy connectives that perform the role of generalized "And", "Or" and "Not" are t-norms, t-conorms and fuzzy negations, respectively, whereas fuzzy IF-THEN rules are usually managed through multivalued implications called in the literature as fuzzy implications or fuzzy implication functions. Many researchers have devoted their efforts to the study of this wide family of functions (see [1–3,8]). There are three main ways of defining fuzzy implication functions: from basic fuzzy logic operations like t-norms, t-conorms, fuzzy negations, uninorms, copulas, etc.; from unary functions on the unit interval; from other fuzzy implication functions. In this article we focus on yet another two classes of fuzzy implication functions defined from binary operations on the unit interval.

Sheffer stroke [4,9], denoted by "|", is a negation of a conjunction:

$$p|q \equiv \neg(p \wedge q).$$

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_2

Classical implication can be presented just by Sheffer stroke operator in two ways:

$$p \to q \equiv \; p|(p|q) \;\equiv\; \neg(p \wedge \neg(p \wedge q)), \tag{PQ}$$

$$p \to q \equiv \; p|(q|q) \;\equiv\; \neg(p \wedge \neg(q \wedge q)). \tag{QQ}$$

The above tautologies may lead to the following two formulas of multivalued implication:

$$I(x, y) = N(C(x, N(C(x, y)))),$$
$$I(x, y) = N(C(x, N(C(y, y)))),$$

for $x, y \in [0, 1]$, where C is some generalization of the classical conjunction and N is some generalization of the classical negation.

In this paper we analyse properties of these two new families of functions, based on the above two formulas. Section 2 contains basic notions, definitions and facts used in the sequel. In Sect. 3 we present basic examples of Sheffer stroke implications and we show when these families are fuzzy implications. In Sect. 4 we analyse basic properties of Sheffer stroke implications, in particular we examine the left neutrality property, the exchange principle, the identity principle and the ordering property.

2 Preliminaries

Definition 1 (see [6,7]). *A function $T\colon [0,1]^2 \to [0,1]$ is called a triangular norm (t-norm for short) if it is symmetric, associative and increasing function, and it satisfies $T(x, 1) = x$, for all $x \in [0, 1]$.*

Definition 2 (see [6,7]). *A function $S\colon [0,1]^2 \to [0,1]$ is called a triangular conorm (t-conorm for short) if it is symmetric, associative and increasing function, and it satisfies $S(x, 0) = x$, for all $x \in [0, 1]$.*

Definition 3 (see [2,7]). *A non-increasing function $N\colon [0,1] \to [0,1]$ is called a fuzzy negation, if $N(0) = 1$, $N(1) = 0$. Moreover, a fuzzy negation N is called*

(i) strict if it is strictly decreasing and continuous,
(ii) strong if it is an involution, i.e., $N(N(x)) = x$, for all $x \in [0, 1]$.

In the sequel we will use basic examples of t-norms and fuzzy negations with the notations from the monograph [2], i.e., $T_{\mathbf{M}}$ is the minimum t-norm, $T_{\mathbf{P}}$ is the algebraic product t-norm, $T_{\mathbf{LK}}$ is the Łukasiewicz t-norm, $T_{\mathbf{D}}$ is the drastic product t-norm, $T_{\mathbf{nM}}$ is the nilpotent minimum t-norm, $N_{\mathbf{C}}$ is the classical fuzzy negation, $N_{\mathbf{D1}}$ is the least fuzzy negation, while $N_{\mathbf{D2}}$ is the greatest fuzzy negation. Their formulas are collected in Table 1.

Definition 4 (see [2,6]). *A function $I\colon [0,1]^2 \to [0,1]$ is called a fuzzy implication, if it satisfies the following conditions:*

Table 1. Basic t-norms and fuzzy negations.

$T_{\mathbf{M}}(x,y) = \min(x,y)$	$T_{\mathbf{P}}(x,y) = xy$
$T_{\mathbf{D}}(x,y) = \begin{cases} 0, & \text{if } x,y \in [0,1) \\ \min(x,y), & \text{otherwise} \end{cases}$	$T_{\mathbf{nM}}(x,y) = \begin{cases} 0, & \text{if } x+y \leq 1 \\ \min(x,y), & \text{otherwise} \end{cases}$
$T_{\mathbf{LK}}(x,y) = \max(x+y-1,0)$	$N_{\mathbf{C}}(x) = 1-x$
$N_{\mathbf{D1}}(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \in (0,1] \end{cases}$	$N_{\mathbf{D2}}(x) = \begin{cases} 1, & \text{if } x \in [0,1) \\ 0, & \text{if } x = 1 \end{cases}$

(I1) *I is non-increasing with respect to the first variable,*
(I2) *I is non-decreasing with respect to the second variable,*
(I3) $I(0,0) = 1$,
(I4) $I(1,1) = 1$,
(I5) $I(1,0) = 0$.

The set of all fuzzy implications will be denoted by $\mathcal{FI}$.

Similarly as for t-norms and fuzzy negations, the notion for basic examples of fuzzy implications is the same as in the book [2]. We recall here the definition of a QL-operation, which is in use in the paper, while the definitions, examples and properties of other well-known families of implications, namely (S,N)-implications, R-implications, f-generated implications and g-generated implications, one may find in [2].

Definition 5. *A function $I\colon [0,1]^2 \to [0,1]$ is called a QL-operation if there exist a t-norm T, a t-conorm S and a fuzzy negation N such that*

$$I(x,y) = S(N(x), T(x,y)), \qquad x,y \in [0,1]. \tag{1}$$

The standard properties of fuzzy implications are listed below.

Definition 6 (see [2]). *We say that a fuzzy implication I satisfies*

(i) the identity principle, if

$$I(x,x) = 1, \qquad x \in [0,1], \tag{IP}$$

(ii) the left neutrality property, if

$$I(1,y) = y, \qquad y \in [0,1], \tag{NP}$$

(iii) the exchange principle, if

$$I(x, I(y,z)) = I(y, I(x,z)), \qquad x,y,z \in [0,1], \tag{EP}$$

(iv) the ordering property, if

$$x \leq y \Longleftrightarrow I(x,y) = 1, \qquad x,y \in [0,1]. \tag{OP}$$

Lemma 1 (see [2, Sects. 2 and 3]). *If I is an (S,N)-implication, then it satisfies* (NP) *and* (EP). *If I is an R-implication, then it satisfies* (NP) *and* (IP). *If I is a QL-operation, then it satisfies* (NP). *If I is an f-generated implication or a g-generated implication, then it satisfies* (NP) *and* (EP).

3 Sheffer Stroke Implications

As we have noted in Introduction, there are two representations of implications that use only Sheffer stroke operator (see Eqs. (PQ) and (QQ)), which leads us to two families of implications. We introduce them in the next two subsections and call them SS_{pq}-implications and SS_{qq}-implications, respectively.

3.1 $\mathbf{SS_{pq}}$-Implications

Now, we define new family of implications based on the equivalence (PQ).

Definition 7. *A function* $I\colon [0,1]^2 \to [0,1]$ *is called an* SS_{pq} *- implication if there exist a fuzzy negation* N *and a t-norm* T *such that*

$$I(x,y) = N(T(x, N(T(x,y)))), \qquad x, y \in [0,1]. \tag{2}$$

If I *is an* SS_{pq} *- implication generated from the couple* (N,T)*, then we will often denote it by* $I^{pq}_{N,T}$*.*

Example 1. Probably the best known SS_{pq}-implication would be the one generated from $N_{\mathbf{C}}$ and $T_{\mathbf{LK}}$, which is a fuzzy implication $I_{\mathbf{KD}}$, i.e.,

$$\begin{aligned} I^{pq}_{N_{\mathbf{C}},T_{\mathbf{LK}}}(x,y) &= N_{\mathbf{C}}(T_{\mathbf{LK}}(x, N_{\mathbf{C}}(T_{\mathbf{LK}}(x,y)))) \\ &= 1 - \max(x + (1 - \max(x + y - 1, 0)) - 1, 0) \\ &= 1 - \max(\min(1-y, x), 0) = 1 - \min(1-y, x) = \max(y, 1-x) \\ &= I_{\mathbf{KD}}(x,y), \end{aligned}$$

for all $x, y \in [0,1]$.

Let us now examine some basic properties of the function $I^{pq}_{N,T}$ to check whether it actually could serve as a fuzzy implication operation.

Lemma 2. *If* $I^{pq}_{N,T}$ *is an* SS_{pq}*-implication, then it satisfies (I2)–(I5).*

Proof. Property (I3) holds since

$$I^{pq}_{N,T}(0,0) = N(T(0, N(T(0,0)))) = N(T(0, N(0))) = N(T(0,1)) = N(0) = 1.$$

The next two properties are direct conclusions from Definition 7 as well. To prove (I2) it is enough to observe that N is decreasing and T is increasing with respect to the second variable. □

The immediate question arises if the requirement (I1), i.e., $I^{pq}_{N,T}$ is decreasing with respect to the first variable, is fulfilled as well. Unfortunately, this is not always true. Consider, e.g. the classical negation $N_{\mathbf{C}}$ and a product t-norm $T_{\mathbf{P}}$. Then we obtain

$$I^{pq}_{N_{\mathbf{C}},T_{\mathbf{P}}}(x,y) = 1 - x + x^2 y, \qquad x, y \in [0,1],$$

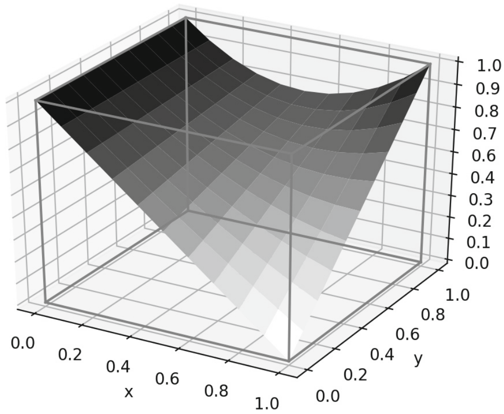


Fig. 1. The plot of $I^{pq}_{N_{\mathbf{C}},T_{\mathbf{P}}}(x,y) = 1 - x + x^2y$, which is a QL-operation, but not a fuzzy implication.

and this function is decreasing for $x \in [0, \frac{1}{2y}]$ and increasing for $x \in [\frac{1}{2y}, 1]$, for any specific $y \in (0.5, 1]$ (see Fig. 1). Therefore, however there are fuzzy implications among this new family of operations (e.g. $I^{pq}_{N_{\mathbf{C}},T_{\mathbf{LK}}} = I_{\mathbf{KD}} \in \mathcal{FI}$), an SS_{pq}-implication is not – in general – a fuzzy implication.

Theorem 1. *An SS_{pq}-implication $I^{pq}_{N,T}$ is a fuzzy implication if and only if it satisfies (I1).*

It is an open problem to characterize property (I1) in the terms of N and T. It turns out that the family of SS_{pq}-implications is basically a subclass of QL-operations, i.e., all SS_{pq}-implications generated from strong negations are QL-operations.

Theorem 2. *Let N be a strong fuzzy negation and T be a t-norm. The SS_{pq}-implication $I^{pq}_{N,T}$ is a QL-operation.*

Proof. Consider the t-conorm S which is N-dual to the t-norm T, i.e., let

$$S(x,y) := N(T(N(x), N(y))), \qquad x, y \in [0,1].$$

Then for all $x, y \in [0,1]$ we have

$$\begin{aligned} I^{pq}_{N,T}(x,y) &= N(T(x, N(T(x,y)))) = N(T(N(N(x)), N(T(x,y)))) \\ &= S(N(x), T(x,y)), \end{aligned}$$

which is the representation (1) of the QL-operation $I_{T,S,N}$. □

The following Table 2 introduces basic examples of SS_{pq}-implications and indicates which of them are fuzzy implications or QL-operations. The last four implications are not QL-operations, because they do not satisfy (NP), what will be proven in the next section, while due to Lemma 1 all QL-operations do have this property. Please note that the first implication is indicated as I_{-1} in [2], while the last implication is indicated as I_4 in [5].

Table 2. Basic SS$_{pq}$-implications.

Negation	T-norm	Sheffer stroke SS$_{pq}$-implication	$I \in \mathcal{FI}$	$I \in$QL-oper.
$N_{\mathbf{C}}$	$T_{\mathbf{M}}$	$I(x,y) = \max(1-x, \min(x,y))$	✗	✓
$N_{\mathbf{C}}$	$T_{\mathbf{LK}}$	$I_{\mathbf{KD}}$	✓	✓
$N_{\mathbf{C}}$	$T_{\mathbf{P}}$	$I(x,y) = 1 - x + x^2 y$	✗	✓
$N_{\mathbf{C}}$	$T_{\mathbf{D}}$	$I(x,y) = \begin{cases} 1, & \text{if } y = 1 \\ y, & \text{if } x = 1 \\ 1-x, & \text{otherwise} \end{cases}$	✗	✓
$N_{\mathbf{C}}$	$T_{\mathbf{nM}}$	$I(x,y) = \begin{cases} 1, & \text{if } x \leq y \text{ and } y > 1-x \\ y, & \text{if } x > y \text{ and } y > 1-x \\ 1-x, & \text{otherwise} \end{cases}$	✗	✓
$N_{\mathbf{D1}}$	$T_{\mathbf{M}}, T_{\mathbf{P}}$	$I(x,y) = \begin{cases} 0, & \text{if } y = 0 \text{ and } x > 0 \\ 1, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{LK}}, T_{\mathbf{nM}}$	$I(x,y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } x + y > 1 \\ 0, & \text{otherwise} \end{cases}$	✗	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{D}}$	$I(x,y) = \begin{cases} 0, & \text{if } x > 0 \text{ and} \\ & (y = 0 \text{ or } \max(x,y) < 1) \\ 1, & \text{otherwise} \end{cases}$	✗	✗
$N_{\mathbf{D2}}$	any T	$I(x,y) = \begin{cases} 0, & \text{if } x = 1 \text{ and } y < 1 \\ 1, & \text{otherwise} \end{cases}$	✓	✗

3.2 SS$_{qq}$-Implications

Now, we define the second family of Sheffer stroke implications, based on the equivalence (QQ).

Definition 8. *A function $I\colon [0,1]^2 \to [0,1]$ is called an SS$_{qq}$ - implication if there exist a fuzzy negation N and a t-norm T such that*

$$I(x,y) = N(T(x, N(T(y,y)))), \qquad x,y \in [0,1]. \tag{3}$$

If I is an SS$_{qq}$ - implication generated from the couple (N,T), then we will often denote it by $I^{qq}_{N,T}$.

Let us check if $I^{qq}_{N,T}$ is a fuzzy implication or if some additional requirements are necessary, as in the case of SS$_{pq}$-implications.

Theorem 3. *If $I^{qq}_{N,T}$ is an SS$_{qq}$ - implication, then it is a fuzzy implication, thus $I^{qq}_{N,T} \in \mathcal{FI}$.*

Proof. The conditions $I^{qq}_{N,T}(0,0) = 1$, $I^{qq}_{N,T}(1,1) = 1$, $I^{qq}_{N,T}(1,0) = 0$ are direct conclusions from Definition 8, the fact that $T(0,x) = 0, T(1,x) = x$ for all $x \in [0,1]$, and the property (N1) satisfied by N. The function $I^{qq}_{N,T}(x,\cdot)$, for any fixed $x \in [0,1]$, is increasing since T is increasing and N is decreasing. For the same reason, the function $I^{qq}_{N,T}(\cdot,y)$, for any fixed $y \in [0,1]$, is decreasing. □

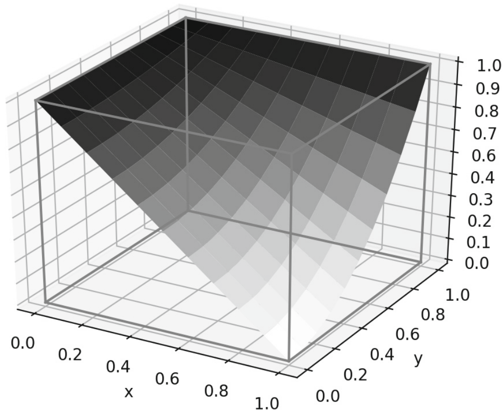


Fig. 2. The plot of $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}(x,y) = 1 - x + xy^2$ fuzzy implication.

Example 2. Let us consider the SS_{qq}-implication $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}$ generated from the classical negation $N_{\mathbf{C}}$ and the product t-norm $T_{\mathbf{P}}$:

$$I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}(x,y) = 1 - x + xy^2, \qquad x, y \in [0,1]. \tag{4}$$

The function $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}$ is a continuous fuzzy implication (see Fig. 2), which does not belong to any well-known family of fuzzy implications, i.e., it is none of an R-implication, an (S,N)-implication, a QL-operation, an f- or g-generated implication. This is because $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}$ does not satisfy (NP), i.e., $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}(1,y) = y^2 \neq y$ for $y \in (0,1)$, but due to Lemma 1, implications from all the above mentioned families of fuzzy implications do satisfy that property.

Remark 1. Actually we may consider more general family of functions

$$I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(x,y) = 1 - x + xy^p, \qquad x, y \in [0,1],$$

for any $p > 0$. When $p = 1$ we have $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1} = I_{\mathbf{RC}}$ the Reichenbach implication (see [2]). It turns out that all functions $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}$ are continuous fuzzy implications, but none of them satisfies (IP) or (OP), and $I_{\mathbf{RC}}$ is the only one that satisfies (NP) or (EP).

Indeed, firstly we have $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(1,y) = y^p$, which is equal to y for all $y \in [0,1]$ only if $p = 1$, thus (NP) is satisfied only for $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1} = I_{\mathbf{RC}}$. Let us consider $x = 1, y = 0.5, z = 0$ to examine (EP),

$$I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(1, I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(0.5,0)) = I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(1,0.5) = 0.5^p,$$
$$I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(0.5, I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(1,0)) = I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(0.5,0) = 0.5,$$

and $0.5^p = 0.5$ if and only if $p = 1$. Furthermore

$$\begin{aligned} I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1}(x, I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1}(y,z)) &= I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1}(x, 1-y+yz) \\ &= 1 - xy(1-z) \\ &= I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1}(y, I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1}(x,z)), \end{aligned}$$

thus $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},1} = I_{\mathbf{RC}}$ is the only one satisfying (EP) (cf. [2]).

Finally, assume that (IP) is satisfied, so $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}(x,x) = 1 - x + x^{p+1} = 1$, for all $x \in [0,1]$ and some $p > 0$. Then $x^{p+1} - x = 0$, thus $x(x^p - 1) = 0$, so either $x = 0$ or $x^p = 1$, for all $x \in (0,1]$. Thus $p = 0$, but we have an assumption that $p > 0$, so $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}},p}$ does not satisfy (IP) for any $p > 0$ and consequently also (OP) is not satisfied.

Example 2 with the function $I^{qq}_{N_{\mathbf{C}},T_{\mathbf{P}}}$ indicates that this new family of SS_{qq}-implications is independent from the other well-known families of fuzzy implications.

Table 3 presents basic examples of SS_{qq}-implications which in most cases are not QL-operations.

Table 3. Basic SS_{qq}-implications.

Negation	T-norm	Sheffer stroke SS_{qq}-implication	$I \in \mathcal{FI}$	$I \in$QL-oper
$N_{\mathbf{C}}$	$T_{\mathbf{M}}$	$I_{\mathbf{KD}}$	✓	✓
$N_{\mathbf{C}}$	$T_{\mathbf{LK}}$	$I(x,y) = \begin{cases} 1-x, & \text{if } y \le 0.5 \\ 1, & \text{if } x \le 2y-1 \\ 2y-x, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{C}}$	$T_{\mathbf{P}}$	$I(x,y) = 1 - x + xy^2$	✓	✗
$N_{\mathbf{C}}$	$T_{\mathbf{D}}$	$I(x,y) = \begin{cases} 1, & \text{if } y = 1 \\ 1-x, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{C}}$	$T_{\mathbf{nM}}$	$I(x,y) = \begin{cases} 1-x, & \text{if } y \le 0.5 \\ y, & \text{if } x > y > 0.5 \\ 1, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{M}}$, $T_{\mathbf{P}}$	$I(x,y) = \begin{cases} 0, & \text{if } y = 0 \text{ and } x > 0 \\ 1, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{LK}}$, $T_{\mathbf{nM}}$	$I(x,y) = \begin{cases} 0, & \text{if } x > 0 \text{ and } y \le 0.5 \\ 1, & \text{otherwise} \end{cases}$	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{D}}$	$I_0(x,y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } y = 1 \\ 0, & \text{if } x > 0 \text{ and } y < 1 \end{cases}$	✓	✗
$N_{\mathbf{D2}}$	any T	$I(x,y) = \begin{cases} 0, & \text{if } x = 1 \text{ and } y < 1 \\ 1, & \text{otherwise} \end{cases}$	✓	✗

Remark 2. Please note that $I_{\mathbf{KD}}$ is the only fuzzy implication which is both SS_{pq}-implication and SS_{qq}-implication, with the assumption that a negation used in both definitions is classical. Indeed, assume that N is a strong negation and put $x = 1$ in (2). Then we obtain that $I^{pq}_{N,T_1}(1, y) = N(N(y)) = y$, for all $y \in [0, 1]$. Now, putting $x = 1$ in (3) we have $I^{qq}_{N,T_2}(1, y) = N(N(T_2(y, y))) = T_2(y, y)$, for all $y \in [0, 1]$. Therefore t-norm T_2 has to be idempotent, which implies $T_2 = T_{\mathbf{M}}$, and in the case when $N = N_{\mathbf{C}}$ we obtain Kleene-Dienes implication $I_{\mathbf{KD}}$.

4 Basic Properties of Sheffer Stroke Implications

This section is devoted to study basic properties of both types of Sheffer stroke implications, namely (NP), (EP), (IP) and (OP). We introduce just partial results which we plan to complete in our future works.

Proposition 1. *An SS_{pq}-implication $I^{pq}_{N,T}$ satisfies* (NP) *if and only if N is strong.*

Proof. For any $y \in [0, 1]$ we have

$$I^{pq}_{N,T}(1, y) = N(T(1, N(T(1, y)))) = N(N(y)),$$

which is equal to y for all $y \in [0, 1]$ if and only if N is a strong negation. □

Proposition 2. *Let N be a strong negation. An SS_{qq}-implication $I^{qq}_{N,T}$ satisfies* (NP) *if and only if $T = T_M$.*

Proof. For any $y \in [0, 1]$ we have

$$I^{qq}_{N,T}(1, y) = N(T(1, N(T(y, y)))) = N(N(T(y, y))) = T(y, y),$$

which is equal to y for all $y \in [0, 1]$ if and only if T is an idempotent t-norm, thus $T = T_{\mathbf{M}}$. □

Proposition 3. *Let $T = T_{\mathbf{M}}$. An SS_{qq}-implication $I^{qq}_{N,T}$ satisfies* (NP) *if and only if N is strong.*

Proof. For any $y \in [0, 1]$ we have

$$I^{qq}_{N,T}(1, y) = N(T(1, N(T(y, y)))) = N(N(T(y, y))) = N(N(y)),$$

which is equal to y for all $y \in [0, 1]$ if and only if N is a strong negation. □

We suppose that implication $I^{qq}_{N,T}$ satisfies (NP) only if N is strong and T is the minimum t-norm. However, the following result states what we can prove for now.

Proposition 4. *If an SS_{qq}-implication $I^{qq}_{N,T}$ satisfies* (NP)*, then N is continuous, while T is not continuous and Archimedean simultaneously, but $T(x, x)$ is an injective function.*

Proof. Let us assume that $I^{qq}_{N,T}$ satisfies (NP), i.e.,

$$I^{qq}_{N,T}(1,y) = N(T(1,N(T(y,y)))) = N(N(T(y,y))) = y, \qquad y \in [0,1].$$

This directly implies that N is surjective, which, with its monotonicity, gives continuity of N. Continuous fuzzy negation N has a unique fixed point $e \in (0,1)$, i.e., $N(e) = e$. Let $f(x) := T(x,x)$, $x \in [0,1]$. If T is continuous, then f is continuous and from Darboux property there exists $x_e \in (0,1)$ such that $f(x_e) = e$. If T is continuous and Archimedean, then $f(x_e) = T(x_e,x_e) < x_e$ (see [2, Remark 2.1.4]), thus $e < x_e$. Then we obtain a contradiction by

$$x_e = I^{qq}_{N,T}(1,x_e) = N(N(T(x_e,x_e))) = N(N(e)) = N(e) = e \neq x_e.$$

Thus T cannot be continuous and Archimedean simultaneously.

Finally, let us assume that $T(x,x)$ is not an injective function, i.e., there exist $x_1,x_2 \in [0,1], x_1 \neq x_2$, such that $T(x_1,x_1) = T(x_2,x_2)$. Then

$$x_1 = I^{qq}_{N,T}(1,x_1) = N(N(T(x_1,x_1))) = N(N(T(x_2,x_2))) = I^{qq}_{N,T}(1,x_2) = x_2,$$

which is a contradiction. □

Finally, we present some partial results connected to (IP) and (OP).

Proposition 5. *Let $I_{N,T}$ be an SS_{pq}-implication or an SS_{qq}-implication, generated from a strong fuzzy negation N and a strict t-norm T. Then $I_{N,T}$ does not satisfy* (OP) *nor* (IP)*.*

Proof. Firstly, notice that if T is a strict t-norm, then $T(x,y) = 0$ implies $x = 0$ or $y = 0$. Indeed, because of the representation of a strict t-norm by an additive generator (see [7]) we have

$$\begin{aligned} T(x,y) =& f^{-1}(f(x)+f(y)) = 0 \;\Leftrightarrow\; f(x)+f(y) = \infty \\ &\Leftrightarrow\; f(x) = \infty \text{ or } f(y) = \infty \Leftrightarrow x = 0 \text{ or } y = 0. \end{aligned}$$

If $I_{N,T}$ satisfies (IP), then $I_{N,T}(x,x) = N(T(x,N(T(x,x)))) = 1$, for all $x \in [0,1]$. Applying the negation N to both sides of the above equation we obtain $T(x,N(T(x,x))) = 0$, for all $x \in [0,1]$, which due to the fact just showed above implies that $x = 0$ or $N(T(x,x)) = 0$ for all $x \in (0,1]$. Since T is continuous, the set of values of $T(x,x)$ for $x \in (0,1]$ consists the set $(0,1]$. Thus $N = N_{\mathbf{D1}}$, which is not a strong negation, a contradiction. Of course, if $I_{N,T}$ does not satisfy (IP), then it does not satisfy (OP), too. □

Example 3. An SS_{pq}-implication $I^{pq}_{N_{\mathbf{C}},T_{\mathbf{LK}}} = I_{\mathbf{KD}}$ satisfies (EP), while other SS_{pq}-implications generated from the classical negation $N_{\mathbf{C}}$ and the following t-norms $T_{\mathbf{M}}$, $T_{\mathbf{P}}$, $T_{\mathbf{D}}$, $T_{\mathbf{nM}}$ do not satisfy (EP). In the case of SS_{qq}-implications it seems that all ot them generated from the least or the greatest fuzzy negations $N_{\mathbf{D1}}, N_{\mathbf{D2}}$ do satisfy (EP), while most of implications generated from the classical fuzzy negation $N_{\mathbf{C}}$ do not (see Table 4). However, the full characterization of solutions to (EP) among Sheffer stroke implications remains an open problem.

Table 4. Sheffer stroke implications and their properties.

Negation	T-norm	(NP)	(EP)	(IP)	(OP)
SS_{pq}-implications					
$N_{\mathbf{C}}$	$T_{\mathbf{M}}, T_{\mathbf{P}}, T_{\mathbf{D}}, T_{\mathbf{nM}}$	✓	✗	✗	✗
$N_{\mathbf{C}}$	$T_{\mathbf{LK}}$	✓	✓	✗	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{M}}, T_{\mathbf{P}}$	✗	✓	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{LK}}, T_{\mathbf{nM}}, T_{\mathbf{D}}$	✗	✗	✗	✗
$N_{\mathbf{D2}}$	any T	✗	✓	✓	✗
SS_{qq}-implications					
$N_{\mathbf{C}}$	$T_{\mathbf{M}}$	✓	✓	✗	✗
$N_{\mathbf{C}}$	$T_{\mathbf{LK}}, T_{\mathbf{P}}, T_{\mathbf{D}}, T_{\mathbf{nM}}$	✗	✗	✗	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{M}}, T_{\mathbf{P}}$	✗	✓	✓	✗
$N_{\mathbf{D1}}$	$T_{\mathbf{LK}}, T_{\mathbf{nM}}, T_{\mathbf{D}}$	✗	✓	✗	✗
$N_{\mathbf{D2}}$	any T	✗	✓	✓	✗

5 Conclusions

According to their name Sheffer stroke implications have a basic genesis related to Sheffer stroke which is one of the two operators that can be used by itself, without any other logical operators, to constitute a logical formal system. Such a genesis is in and of itself a good enough reason to study this new family of implications. In the paper we introduced two subfamilies of Sheffer stroke implications, namely SS_{pq}-implications and SS_{qq}-implications, and we showed that while the first one is mainly a subfamily of QL-operations, the second one is in fact a new family of fuzzy implications. We introduced some basic properties of both families. However it seems that Sheffer stroke implications usually do not satisfy (NP), (EP), (IP) or (OP), we have got so far only partial results. The questions about other properties, among them t-conditionality, distributivity laws and laws of contraposition, are in our future plans. We believe that answering them in the future would show the potential in applications of Sheffer stroke implications.

Acknowledgments. The work on this paper for Wanda Niemyska and Michał Baczyński was supported by the research project No. 2015/19/B/ST6/03259 from the National Science Centre, Poland.

References

1. Baczyński, M., Beliakov, G., Bustince, H., Pradera, A.: Advances in Fuzzy Implication Functions. Studies in Fuzziness and Soft Computing, vol. 300. Springer, Heidelberg (2013)

2. Baczyński, M., Jayaram, B.: Fuzzy Implications. Studies in Fuzziness and Soft Computing, vol. 231. Springer, Heidelberg (2008)
3. Baczyński, M., Jayaram, B., Massanet, S., Torrens, J.: Fuzzy implications: past, present, and future. In: Kacprzyk, J., Pedrycz, W. (eds.) Springer Handbook of Computational Intelligence, pp. 183–202. Springer, Heidelberg (2015)
4. Church, A.: Introduction to Mathematical Logic, vol. 1. Princeton University Press, Princeton (1956)
5. Drewniak, J.: Selfconjugate fuzzy implications, In: De Baets, B., Fodor, J.C. (eds.) Proceedings of Joint Eurofuse-Soft and Intelligent Computing Conference, Budapest, Hungary, pp. 351–353, May 1999
6. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht (1994)
7. Klement, E., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
8. Mas, M., Monserrat, M., Torrens, J., Trillas, E.: A survey on fuzzy implication functions. IEEE Trans. Fuzzy Syst. **15**(6), 1107–1121 (2007)
9. Sheffer, H.M.: A set of five independenr postulates for Boolean algebras, with application to logical constants. Trans. Amer. Math. Soc. **14**, 481–488 (1913)

Towards Fuzzy Type Theory with Partial Functions

Vilém Novák(✉)

Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, NSC IT4Innovations, 30. dubna 22, 701 03, Ostrava 1, Czech Republic
Vilem.Novak@osu.cz
http://irafm.osu.cz

Abstract. This paper is a study of fuzzy type theory (FTT) with partial functions. Out of several possibilities we decided to introduce a special value "$*$" which represents "undefined". In the interpretation of FTT, this value lays outside of the corresponding domain. In the syntax, it is naturally represented by the description operator acting on the empty (fuzzy) set which, of course, has no element and so, choosing an element from its kernel gives no result, i.e., it is undefined. We will demonstrate that our approach leads to reasonable characterization of the undefinedness. We will also show that any consistent theory of FTT has a model.

Keywords: Partial functions · Higher-order fuzzy logic · Fuzzy type theory · EQ-algebra

1 Introduction

There are several reasons for considering partial functions in mathematics and its applications. We can meet them in recursion theory (no ending algorithm), in the analysis (division by zero), in computer science (mistake in the computer program), and elsewhere. In the logical analysis of natural language (cf. [2]), one can meet sentences that denote nothing, e.g., "The present French king is bald".

There are several ways how partial functions can be introduced. For example, W. Farmer in his paper [5] discusses 8 possibilities which appear in the literature: non-denoting expressions as non-well-formed terms, functions represented as relations, total functions with unspecified value, many-sorted language, error values, non-existent values, partial valuation for terms and formulas, partial valuation for terms but total valuation for formulas. Each of these approaches provides a certain kind of solution but has also drawbacks. It seems that fully satisfactory solution does not exist.

The powerful formal theory applied in many branches from linguistics and logic to computer science is type theory that formalized higher-order logic. A couple of years ago, the type theory has been generalized to the fuzzy one (FTT; see [9]). The syntax of this theory is extension of the λ-calculus.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_3

The logical system has more axioms than the classical type theory and the fundamental connective is that of *fuzzy equality*, i.e., this formalizes imprecise equality that characterizes relation between objects that do not need to be the same but only similar to each other in a certain degree. The crucial notion is that of *type* that can be understood as an index characterizing the kind of objects in concern. Hence, we distinguish basic types o of *truth values* and ϵ of *elements* and then we can form more complex types $\beta\alpha$ representing functions from a set M_α of elements of type α to a set M_β of elements of type β. Hence, semantics of FTT is based on the concept of *frame* that is a system of sets $(M^*_\alpha, \stackrel{\circ}{=}_\alpha)_{\alpha\in Types}$ for all types $\alpha \in Types$, each of them endowed by the fuzzy equality $\stackrel{\circ}{=}_\alpha$. We must also consider the algebra of truth values $\mathscr{E}_\Delta$.

In this paper, we introduce FTT with partial functions. Our solution is standard in the sense that we introduce a special value $*$ interpreted as "undefined". This approach has one important advantage. Namely, Tichý in [15] found a counterexample showing that when dealing with partial functions, we cannot use the important principle of type theory called λ-conversion. However, Lepage in [7] demonstrated that by introducing $*$ as a special value, the counterexample of Tichý can be overcome and so, the principle of λ-conversion is preserved. This is important because the latter principle is an important tool of type theory.

The idea presented in this paper follows the idea of [11] to use the *description operator* $\iota_{\alpha(o\alpha)}$. The solution there, however, was not satisfactory as the behavior of $*$ was close to falsity. Recall that due to [8], interpretation of the description operator is just a partial function giving a value from the kernel of the corresponding fuzzy set if the latter is normal, and giving nothing otherwise. This suggests the idea that $*_o$ of type o can be defined as the formula $\iota_{o(oo)} \cdot \lambda x_o \bot$, whose interpretation applies the description operator to the empty set. Of course, empty set contains no element and so, this formula gives no result. We thus get a natural interpretation of $*_o$. Similarly, the $*_\epsilon$ is defined as $\iota_{\epsilon(o\epsilon)} \cdot \lambda x_\epsilon *_o$, i.e., the description operator is applied to nowhere defined function. The latter principle is then applied also to $*_\alpha$ for arbitrary type.

We develop our theory as extension of FTT based on the EQ-algebra of truth values [10]. This is a special algebra introduced as natural algebra of truth degrees for higher-order fuzzy logics. Because of lack of space, we could not include proofs of the theorems. These are available in the full paper [12].

2 Truth Values and Fuzzy Equality

2.1 Truth Values

The truth degrees form a linearly ordered bounded good EQ$_\Delta$-algebra (see [10,13])

$$\mathscr{E}_\Delta = \langle E, \wedge, \otimes, \sim, \mathbf{0}, \mathbf{1}, \Delta\rangle \tag{1}$$

where for all $a, b, c, d \in E$:

(E1) $\langle E, \wedge, \mathbf{1}\rangle$ is a commutative idempotent monoid (i.e. $\wedge$-semilattice). We put $a \le b$ iff $a \wedge b = a$, as usual. Then $\mathbf{1}$ is the top and $\mathbf{0}$ the bottom element.

(E2) $\langle E, \otimes, \mathbf{1}\rangle$ is a monoid, $\otimes$ is isotone w.r.t. $\leq$.
(E3) $a \sim a = \mathbf{1}$ (reflexivity)
(E4) $((a \wedge b) \sim c) \otimes (d \sim a) \leq c \sim (d \wedge b)$ (substitution)
(E5) $(a \sim b) \otimes (c \sim d) \leq (a \sim c) \sim (b \sim d)$ (congruence)
(E6) $(a \wedge b \wedge c) \sim a \leq (a \wedge b) \sim a$ (monotonicity)
(E7) $a \sim \mathbf{1} = a$ (goodness)

We define $a \to b = (a \wedge b) \sim a$ (implication) and $\neg a = a \to \mathbf{0}$ (negation). It can be proved that $a \leq b$ iff $a \to b = 1$. The above algebra is extended by the *delta operation* $\Delta : E \longrightarrow E$ which keeps $\mathbf{1}$ and sends all smaller values to $\mathbf{0}$ (for the details, see [3,4]).

It should be emphasized that every residuated lattice is a good EQ-algebra with fuzzy equality being the biresiduation $a \sim b = (a \to b) \wedge (b \to a)$. An EQ-algebra $\mathscr{E}$ is *prelinear* if for all $a, b \in E$, $((a \to b) \to c) \leq ((b \to a) \to c) \to c$. If $\mathscr{E}$ is prelinear then it is *lattice ordered* where the join is defined by $a \vee b = ((a \to b) \to b) \wedge ((b \to a) \to a)$.

2.2 Extended Algebra of Truth Values

To deal with partial functions, we will consider a special "truth value" $*$ where $* \notin E$ and interpret it as *undefined*. Extended EQ$_\Delta$-algebra of truth values $\mathscr{E}^*_\Delta$ has the support $E^* = E \cup \{*\}$ and the operations $\sim, \wedge, \otimes$ and Δ are extended to the whole E^* as follows.

Let $a, b \in E$ and $\bigcirc \in \{\wedge, \otimes\}$. Then the following tables define the operations in the extended algebra $\mathscr{E}^*_\Delta$:

$\sim$	b	$*$
a	$a \sim b$	$\mathbf{0}$
$*$	$\mathbf{0}$	$\mathbf{1}$

$\bigcirc$	b	$*$
a	$a \bigcirc b$	$*$
$*$	$*$	$*$

$\to$	b	$*$
a	$a \to b$	$\mathbf{0}$
$*$	$\mathbf{1}$	$\mathbf{1}$

$\vee$	b	$*$
a	$a \vee b$	$\mathbf{0}$
$*$	$\mathbf{0}$	$\mathbf{0}$

x	Δx
a	Δa
$*$	$*$

x	$\neg x$
a	$\neg a$
$*$	$\mathbf{0}$

Note that $\mathbf{0}$ remains bottom element for all $a \in E$. We at the same time have $* \leq \mathbf{0}$ but $\mathbf{0} \not\leq *$.

Finally we introduce the following derived operations on E^*:

$$?x = x \sim *, \qquad !x = \neg ?x,$$
$$\downarrow x = x \sim \mathbf{1}, \qquad \uparrow x = \neg !x \vee \downarrow x.$$

	?	!	$\downarrow$	$\uparrow$
a	$\mathbf{0}$	$\mathbf{1}$	a	a
$*$	$\mathbf{1}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{1}$

$a \in E$.

The operation "?" is a test for *undefined*, "!" is a test for *defined*, $\downarrow$ and $\uparrow$ are *star-0* and *star-1 reinterpretation*, respectively.

2.3 Fuzzy Equality

A *fuzzy equality on truth values* is the operation $\sim$ from the EQ-algebra $\mathscr{E}$. This fuzzy equality is separated, i.e., $a \sim b = \mathbf{1}$ implies $a = b$. A fuzzy equality $\circeq$ on an arbitrary set M is a binary fuzzy relation $\circeq : M \times M \longrightarrow E$ that is reflexive, symmetric and $\otimes$-transitive. If $m, m' \in M$ then we usually write $[m \circeq m']$

instead of $\doteq(m, m')$. We will extend the fuzzy equality $\doteq$ to $\doteq: M^* \times M^* \longrightarrow E$ where $M^* = M \cup \{*\}$, $* \notin M$, as follows[1]:

$$\begin{array}{c|cc} \doteq & y & * \\ \hline x & x \doteq y & \mathbf{0} \\ * & \mathbf{0} & \mathbf{1} \end{array} \qquad x, y \in M. \tag{2}$$

Let $M \subseteq M_\beta^{M_\alpha}$ where M_α, M_β are sets endowed with the corresponding fuzzy equalities $\doteq_\alpha, \doteq_\beta$. We will add an element $*_{\beta\alpha} \notin M$ to M (the $*_{\beta\alpha}$ represents a *nowhere defined function* on M_α). Then

$$[h \doteq h'] = \bigwedge_{m \in M_\alpha^*} [h(m) \doteq_\beta h'(m)], \qquad h, h' \in M^* \tag{3}$$

is the fuzzy equality $\doteq: M^* \times M^* \longrightarrow E$ where $\doteq_\beta$ is the fuzzy equality on the set M_β. Note that (2) gives $\mathbf{0}$ in (3) whenever there is $m \in M_\alpha^*$ such that for $h, h' \in M$, either $h(m) = *$ or $h'(m) = *$. Consequently, also for functions $h \in M$ we obtain $[h \doteq *_{\beta\alpha}] = \mathbf{0}$ and $[*_{\beta\alpha} \doteq *_{\beta\alpha}] = \mathbf{1}$.

3 Syntax of Partial FTT

The basic syntactical objects of FTT are classical — see [1], namely the concepts of type and formula. The atomic types are ϵ (elements) and o (truth degrees). Complex types $(\beta\alpha)$ are formed from previously formed ones β and α. The set of all types is denoted by *Types*.

The *language* of FTT denoted by J consists of variables $x_\alpha, \ldots$, special constants $c_\alpha, \ldots$ ($\alpha \in \mathit{Types}$), auxiliary symbols λ and brackets. Formulas are formed of variables, constants (each of specific type), and the symbol λ. Thus, each formula A is assigned a type (we write A_α). The set of formulas of type α is denoted by Form_α, the set of all formulas by Form. Interpretation of a formula $A_{\beta\alpha}$ is a function from the set of objects of type α into the set of objects of type β. Thus, if $B \in \mathit{Form}_{\beta\alpha}$ and $A \in \mathit{Form}_\alpha$ then $(BA) \in \mathit{Form}_\beta$. Similarly, if $A \in \mathit{Form}_\beta$ and $x_\alpha \in J$, $\alpha \in \mathit{Types}$, is a variable then $\lambda x_\alpha\, A_\beta \in \mathit{Form}_{\beta\alpha}$ is a formula whose interpretation is a function that assigns to each object of type α an object of type β represented by the formula A_β.

The set of formulas of type α, $\alpha \in \mathit{Types}$, is denoted by Form_α. A set of all formulas of the language J is $\mathit{Form} = \bigcup_{\alpha \in \mathit{Types}} \mathit{Form}_\alpha$.

Specific constants always present in the language of FTT are the following: $\mathbf{E}_{(o\alpha)\alpha}$, $\alpha \in \{o, \epsilon\}$ (fuzzy equality), $\mathbf{C}_{(oo)o}$ (conjunction), $\mathbf{S}_{(oo)o}$ (strong conjunction), $\mathbf{D}_{oo}$ (delta) and $\iota_{\alpha(o\alpha)}$, $\alpha \in \{o, \epsilon\}$ (the description operator).

A variable x_α is *bound* in a formula A_δ if the latter has a well formed part $(\lambda x_\alpha B_\beta)$. Otherwise x_α is *free*. A formula A is *closed* if it does not contain free variables. A closed formula A_o of type o is called a *sentence*.

[1] Note that this "$*$" is a different element from "$*$" introduced for truth values.

Formal definitions

(i) *Basic fuzzy equality* $\equiv_{(o\alpha)\alpha} \equiv \lambda x_\alpha\, \lambda y_\alpha \cdot (\mathbf{E}_{(o\alpha)\alpha}\, y_\alpha) x_\alpha, \alpha \in \{o, \epsilon\}$.
(ii) *Strong conjunction* $\&_{(oo)o}$, *conjunction* $\wedge_{(oo)o}$, *delta* $\boldsymbol{\Delta}_{oo}$, *truth* $\top$, *falsity* $\bot$, *implication* $\Rightarrow_{(oo)o}$, *negation* $\neg_{oo}$ and *disjunction* $\vee_{(oo)o}$ are defined in the same way as in [10].
(iii) *General quantifier:* $(\forall x_\alpha) A_o \equiv (\lambda x_\alpha\, A_o \equiv \lambda x_\alpha\, \top)$
(iv) *Fuzzy equality between functions:*

$$\equiv_{(o(\beta\alpha))(\beta\alpha)} \equiv \lambda f_{\beta\alpha}\, \lambda g_{\beta\alpha} \cdot (\forall x_\alpha)(f_{\beta\alpha}\, x_\alpha \equiv g_{\beta\alpha}\, x_\alpha).$$

(v) *The values "undefined":*

$$*_o \equiv \iota_{o(oo)} \cdot \lambda x_o\, \bot, \tag{4}$$
$$*_\epsilon \equiv \iota_{\epsilon(o\epsilon)} \cdot \lambda x_\epsilon\, *_o, \tag{5}$$
$$*_{\beta\alpha} \equiv \lambda x_\alpha\, *_\beta, \qquad \alpha, \beta \in Types. \tag{6}$$

Note that the element $*_{\beta\alpha}$ represents the nowhere defined function.

(vi) *Undefined:* $?_{o\alpha} \equiv \lambda x_\alpha \cdot x_\alpha \equiv_\alpha *_\alpha$.
(vii) *Defined:* $!_{o\alpha} \equiv \lambda x_\alpha \cdot \neg ?x_\alpha$.
(viii) *Star-0 reinterpretation:* $\downarrow_{oo} \equiv \lambda x_o \cdot x_o \equiv \top$.
(ix) *Star-1 reinterpretation:* $\uparrow_{oo} \equiv \lambda x_o \cdot \neg !x_o \vee \downarrow x_o$.
(x) *Existential quantifier:* $(\exists x_\alpha) A_o \equiv (\forall y_o)((\forall x_\alpha) \boldsymbol{\Delta}(A_o \Rightarrow\uparrow y_o) \Rightarrow\uparrow y_o)$ (y_o does not occur in A_o).

Convention 1. *The fuzzy equality between formulas of the type* α *is denoted by* $\equiv_\alpha$. *The fuzzy equality between truth values is usually called* fuzzy equivalence.

3.1 Axioms and Inference Rules

Fundamental axioms

(FT-fund1) $\boldsymbol{\Delta}(x_\alpha \equiv_\alpha y_\alpha) \Rightarrow (f_{\beta\alpha}\, x_\alpha \equiv_\beta f_{\beta\alpha}\, y_\alpha), \qquad \alpha, \beta \in Types$,
(FT-fund2) $(f_{\beta\alpha} \equiv g_{\beta\alpha}) \Rightarrow (f_{\beta\alpha}\, x_\alpha \equiv_\beta g_{\beta\alpha}\, x_\alpha), \qquad \alpha, \beta \in Types$,
(FT-fund3) $(\lambda x_\alpha B_\beta) A_\alpha \equiv_\beta C_\beta, \qquad \alpha, \beta \in Types$,
where C_β is obtained from B_β by replacing all free occurrences of x_α in it by A_α, provided that A_α is substitutable to B_β for x_α (*lambda conversion*).
(FT-fund4) $(x_\epsilon \equiv_\epsilon y_\epsilon) \equiv_\epsilon (y_\epsilon \equiv_\epsilon x_\epsilon)$,
(FT-fund5) $(x_\epsilon \equiv_\epsilon y_\epsilon) \& (y_\epsilon \equiv_\epsilon z_\epsilon) \Rightarrow (x_\epsilon \equiv_\epsilon z_\epsilon)$.

Axioms of truth degrees As usual in fuzzy logic, we have two kinds of conjunction, namely the "ordinary" conjunction $\wedge$ and the *strong conjunction* $\&$. Let $\bigcirc \in \{\wedge, \&\}$.

(FT-tval1) $\top$
(FT-tval2) $(x_o \wedge y_o) \equiv (y_o \wedge x_o)$,

(FT-tval3) $(x_o \bigcirc y_o) \bigcirc z_o \equiv x_o \bigcirc (y_o \bigcirc z_o)$,
(FT-tval4) $!A_o \Rightarrow ((A_o \equiv \top) \equiv A_o)$,
(FT-tval5) $!A_\alpha \equiv (!A_\alpha \equiv \top)$, $\alpha \in Types$,
(FT-tval6) $(A_\alpha \equiv_\alpha A_\alpha) \equiv \top$, $\alpha \in Types$,
(FT-tval7) $!((A_\alpha \equiv_\alpha A_\alpha) \equiv \top)$, $\alpha \in Types$,
(FT-tval8) $(A_o \bigcirc \top) \equiv A_o$,
(FT-tval9) $((\top \,\&\, A_o) \equiv A_o)$,
(FT-tval10) $!A_o \Rightarrow ((A_o \wedge \bot) \equiv \bot)$,
(FT-tval11) $(A_o \wedge A_o) \equiv A_o$,
(FT-tval12) $((x_o \wedge y_o) \equiv z_o) \,\&\, (t_o \equiv x_o) \Rightarrow (z_o \equiv (t_o \wedge y_o))$,
(FT-tval13) $(x_o \equiv y_o) \,\&\, (z_o \equiv t_o) \Rightarrow (x_o \equiv z_o) \equiv (y_o \equiv t_o)$,
(FT-tval14) $(x_o \Rightarrow (y_o \wedge z_o)) \Rightarrow (x_o \Rightarrow y_o)$,
(FT-tval15) $\boldsymbol{\Delta}(x_o \Rightarrow y_o) \Rightarrow (x_o \,\&\, z_o \Rightarrow y_o \,\&\, z_o)$,
(FT-tval16) $\boldsymbol{\Delta}(x_o \Rightarrow y_o) \Rightarrow (z_o \,\&\, x_o \Rightarrow z_o \,\&\, y_o)$,
(FT-tval17) $((x_o \Rightarrow y_o) \Rightarrow z_o) \Rightarrow ((y_o \Rightarrow x_o) \Rightarrow z_o) \Rightarrow z_o$.

Axioms of delta

(FT-delta1) $(g_{oo}(\boldsymbol{\Delta} x_o) \wedge g_{oo}(\neg \boldsymbol{\Delta} x_o)) \equiv (\forall y_o) g_{oo}(\boldsymbol{\Delta} y_o)$,
(FT-delta2) $\boldsymbol{\Delta}(A_o \wedge B_o) \equiv \boldsymbol{\Delta} A_o \wedge \boldsymbol{\Delta} B_o$,
(FT-delta3) $\boldsymbol{\Delta}(A_o \vee B_o) \Rightarrow \boldsymbol{\Delta} A_o \vee \boldsymbol{\Delta} B_o$,
(FT-delta4) $(\boldsymbol{\Delta} \uparrow A_o \vee \neg \boldsymbol{\Delta} \uparrow A_o) \equiv \top$.

Axioms of star

(FT-B1) $(\bot \equiv *_o) \equiv \bot$,
(FT-B2) $(\top \equiv *_o) \equiv \bot$,
(FT-B3) $A_o \bigcirc *_o \equiv *_o$,
(FT-B4) $\boldsymbol{\Delta} *_o \equiv *_o$,
(FT-B5)]$((A_\alpha \equiv_\alpha *_\alpha) \vee \neg(A_\alpha \equiv_\alpha *_\alpha)) \equiv \top$, $\alpha \in \{o, \epsilon\}$.
(FT-B6) $!A_o \Rightarrow (!B_o \Rightarrow !(A_o \bigcirc B_o))$,
(FT-B7) $!A_o \Rightarrow (!B_o \Rightarrow !(A_o \equiv B_o)$.

Axioms of quantifiers

(FT-quant1) $\boldsymbol{\Delta}(\forall x_\alpha)(A_o \Rightarrow B_o) \Rightarrow (A_o \Rightarrow (\forall x_\alpha) B_o)$, x_α is not free in A_o,
(FT-quant2) $(\forall x_\alpha)(A_o \Rightarrow B_o) \Rightarrow ((\exists x_\alpha) A_o \Rightarrow B_o)$, x_α is not free in B_o,
(FT-quant3) $(\forall x_\alpha)(A_o \vee B_o) \Rightarrow ((\forall x_\alpha) A_o \vee B_o)$, x_α is not free in B_o.

Axioms of descriptions

(FT-descr1) $\iota_{\alpha(o\alpha)}(\mathbf{E}_{(o\alpha)\alpha}\, y_\alpha) \equiv_\alpha y_\alpha$, $\alpha = \{o, \epsilon\}$,
(FT-descr2) $(\forall x_\alpha)(\neg \boldsymbol{\Delta}(f_{o\alpha} x_\alpha \equiv \top)) \Rightarrow (\iota_{\alpha(o\alpha)} f_{o\alpha} \equiv *_\alpha)$, $\alpha \in \{o, \epsilon\}$.

The inference rules remain unchanged:

(R) *Let* $A_\alpha \equiv A'_\alpha \in Form_o$ *and* $B_o \in Form_o$ *be formulas. Then we infer from them a formula* B'_o *which comes from* B_o *by replacing one occurrence of* A_α *by* A'_α*, provided that the occurrence of* A_α *in* B_o *is not an occurrence of a variable immediately preceded by* λ.

(N) *Let* $A_o \in Form_o$ *be a formula. Then from* A_o *infer* $\boldsymbol{\Delta} A_o$.

A theory T is a set of formulas of type o (determined by a subset of special axioms, as usual). Provability is defined as usual. If T is a theory and A_o a formula then $T \vdash A_o$ means that A_o is provable in T. A theory T is *contradictory* if $T \vdash \bot$. Otherwise it is *consistent*.

The following theorem shows that we must be careful with definability of formulas. In syntax, this is internal definability that must be provable to be able to deal with it.

Theorem 1. *(a)* $(A_o \equiv \top) \vdash !A_o$.
(b) $A_o, (A_o \equiv B_o) \vdash B_o$. *(Rule (EMP))*
(c) $(A_o \equiv \top), (A_o \Rightarrow B_o) \vdash B_o$. *(Modus Ponens I)*
(d) Let $\vdash !A_o \equiv \top$. *Then* $\vdash A_o$ *iff* $\vdash A_o \equiv \top$.
(e) $!A_o, A_o, (A_o \Rightarrow B_o) \vdash B_o$. *(Modus Ponens II)*
(f) $!A_o \vdash (A_o \Rightarrow \bot) \Rightarrow \neg A_o$.

3.2 Logical Connectives and the "undefined"

In this subsection we will show that our fuzzy type theory with the concepts of defined and undefined behaves in accordance with the intuition.

Theorem 2. *Let* $\alpha \in Types$. *Then*

(a) $\vdash A_\alpha \equiv_\alpha A_\alpha$, $\vdash !(A_\alpha \equiv A_\alpha)$ *and* $\vdash !(*_\alpha \equiv *_\alpha)$, $\qquad \alpha \in Types$.
(b) $!A_\alpha, !B_\alpha \vdash !(A_\alpha \equiv B_\alpha)$,
(c) $\vdash ?*_\alpha$, $\quad \vdash !?*_\alpha$, $\quad \vdash \neg !*_\alpha$.

By (c), it is provable that the value "undefined" is internally undefined and also, not defined. However, the predicate undefined is itself defined.

Corollary 1. *If* A_o *is an axiom of FTT then* $\vdash !A_o$.

Theorem 3. *(a)* $A_o \equiv \top \vdash (\forall x_\alpha) A_o$. *(generalization I)*
(b) $!A_o, A_o \vdash (\forall x_\alpha) A_o$. *(generalization II)*
(c) $\Box A_o, A_o \equiv B_o \vdash \Box B$ $\quad$ *where* $\Box \in \{!, ?, \downarrow, \uparrow\}$.

Theorem 4. *Let* $\bigcirc \in \{\wedge, \&\}$. *Then*

(a) $!A_o, A_o, !B_o, B_o \vdash A_o \bigcirc B_o$,
(b) $\vdash (A_o \Rightarrow B_o), (B_o \Rightarrow A_o) \vdash A_o \equiv B_o$,
(c) $!A_o, !B_o \vdash !(A_o \wedge B_o)$,
(d) $!A_o, !B_o \vdash !(A_o \Rightarrow B_o)$,
(e) $\vdash (\uparrow A_o \Rightarrow *_o) \equiv \bot, \qquad \vdash (\downarrow A_o \Rightarrow *_o) \equiv \bot, \qquad \vdash (*_o \Rightarrow A_o) \equiv \top$,
(f) $\vdash !(*_o \vee *_o)$ *and* $\vdash (*_o \vee *_o) \equiv \bot$,
(g) $!A_o, B_o \vdash !(A_o \vee B_o)$.

Theorem 5.(a) $\vdash (\forall x_\alpha)B_o \Rightarrow (B_{o,x_\alpha}[A_\alpha] \equiv \top)$, *(substitution I)*
(b) $!B_o \vdash (\forall x_\alpha)B_o \Rightarrow B_{o,x_\alpha}[A_\alpha]$, *(substitution II)*
(c) $!B_{o\alpha}A_\alpha \vdash B_{o\alpha}A_\alpha \Rightarrow (\exists x_\alpha)B_{o\alpha}x_\alpha$. *($\exists$-substitution)*

Theorem 6. (Deduction theorem). *Let T be a theory and $A_o \in \mathrm{F}orm_o$ a closed formula such that $T \vdash !A_o$. Then*

$$T \cup \{A_o\} \vdash B_o \quad \textit{iff} \quad T \vdash \boldsymbol{\Delta} A_o \Rightarrow B_o$$

for every formula $B_o \in \mathrm{F}orm_o$ such that $T \vdash !B_o$.

The following simple theorem says that if "undefined" is defined in a theory T then it is contradictory.

Theorem 7. *If $T \vdash !*_o$ then T is contradictory.*

Theorem 8.(a) $\vdash (\forall x_\alpha)*_o \equiv \bot$ *and* $\vdash (\exists x_\alpha)*_o \equiv \bot$.
(b) *Let* $T \vdash (\exists x_\alpha)\boldsymbol{\Delta}?A_{o\alpha}x_\alpha$. *Then* $T \vdash (\forall x_\alpha)A_{o\alpha}x_\alpha \equiv \bot$.
(c) *Let* $T \vdash (\forall x_\alpha)\boldsymbol{\Delta}?A_{o\alpha}x_\alpha$. *Then* $T \vdash (\exists x_\alpha)A_{o\alpha}x_\alpha \equiv \bot$.

By this theorem, if there is an undefined functional value $A_{o\alpha}x_\alpha$ then it is not true that $A_{o\alpha}x_\alpha$ holds for all x_α. And vice versa, this it is undefined for all x_α then it is not true that it holds for some x_α.

Theorem 9. *For all types $\alpha, \beta, gamma \in Types$:*

(a) $\vdash *_{\beta\alpha} *_\alpha \equiv_\beta *_\beta$.
(b) $(\forall x_\gamma)(\neg\boldsymbol{\Delta}(A_{o\gamma}x_\gamma \equiv \top)) \vdash \iota_{\gamma(o\gamma)}A_{o\gamma} \equiv_\gamma *_\gamma$
(c) $\vdash \iota_{\alpha(o\alpha)}*_{o\alpha} \equiv_\alpha *_\alpha$.

By (b), the description operator applied to a subnormal fuzzy set gives "undefined". By (c), the nowhere defined function at the "undefined" argument gives again the value "undefined".

4 Semantics of Partial FTT

Let $\mathscr{M}$ be a frame for FTT. Then we will introduce an *extended general frame*

$$\mathscr{M}^* = \langle (M^*_\alpha, \mathring{=}_\alpha)_{\alpha \in Types}, \mathscr{E}^*_\Delta, I_o, I_\epsilon \rangle \tag{7}$$

so that the following holds:

(i) We put $M^*_o = E^*$ where the latter is a support of the extended EQ-algebra $\mathscr{E}^*_\Delta$. Furthermore, let $*_\epsilon \notin M_\epsilon$. Then we put $M^*_\epsilon = M_\epsilon \cup \{*_\epsilon\}$. For all $\alpha = \gamma\beta$, the function

$$*_{\gamma\beta} : M^*_\beta \longrightarrow M^*_\gamma$$

where $*_{\gamma\beta}(m_\beta) = *_\gamma$ for all $m_\beta \in M^*_\beta$, represents "undefined". For all types $\gamma\beta$ we put $M^*_{\gamma\beta} \subseteq (M^*_\gamma)^{M^*_\beta}$, where we require that $*_{\gamma\beta} \in M^*_{\gamma\beta}$.

(ii) The $\mathscr{E}^*_\Delta$ is an extended algebra of truth degrees (EQ$_\Delta$-algebra). We assume that the sets M^*_{oo}, $M^*_{(oo)o}$ contain all the operations discussed in Subsect. 2.2.

(iii) $\mathring{=}_\alpha$: $M^*_\alpha \times M^*_\alpha \longrightarrow L$ is a fuzzy equality on M^*_α for every $\alpha \in Types$. We define: $\mathring{=}_o := \sim$ and $\mathring{=}_\epsilon$ is the fuzzy equality on M^*_ϵ given explicitly. The fuzzy equality $\mathring{=}_{\beta\alpha}$ for complex types is defined in (3).

(iv) $I_o : \mathscr{F}(M_o) \longrightarrow M_o$, $I_\epsilon : \mathscr{F}(M_\epsilon) \longrightarrow M_\epsilon$ are partial functions interpreting the basic description operators. Let $B \underset{\sim}{\subset} M_\alpha$, $\alpha \in \{o, \epsilon\}$. Then

$$I_\alpha(B) = \begin{cases} a_B \in \mathrm{Ker}(B) & \text{if } B \text{ is normal,} \\ *_\alpha & \text{otherwise,} \end{cases}$$

Interpretation of formulas in the frame $\mathscr{M}$ is defined using an assignment p of elements from $\mathscr{M}$ to variables. By $p' = p\backslash x_\alpha$ we denote an assignment that equals to p for all variables except for x_α. The set of all assignments over $\mathscr{M}$ is denoted by Asg($\mathscr{M}$).

For arbitrary assignment $p \in$ Asg we define $\mathscr{M}^*_p(x_\alpha) = p(x_\alpha) \in M^*_\alpha$, $\mathscr{M}_p(\mathbf{E}_{(oo)o}) := \sim$, $\mathscr{M}_p(\mathbf{E}_{(o\epsilon)\epsilon} := \mathring{=}_\epsilon$, $\mathscr{M}_p(\mathbf{C}_{(oo)o}) := \wedge$, $\mathscr{M}_p(\mathbf{S}_{(oo)o}) := \otimes$, $\mathscr{M}_p(\mathbf{D}_{oo}) := \Delta$. Description operators[2]: $\mathscr{M}_p(\iota_{\epsilon(o\epsilon)}) = I_o$ and $\mathscr{M}_p(\iota_{o(oo)}) = I_\epsilon$.

Interpretation of $B_{\beta\alpha}A_\alpha$ is $\mathscr{M}_p(B_{\beta\alpha}A_\alpha) = \mathscr{M}_p(B_{\beta\alpha})(\mathscr{M}_p(A_\alpha))$. Interpretation of $\lambda x_\alpha\, A_\beta$ is a function $\mathscr{M}_p(\lambda x_\alpha\, A_\beta) = F : M_\alpha \longrightarrow M_\beta$ which assigns to each $m_\alpha \in M_\alpha$ the element $F(m_\alpha) = \mathscr{M}_{p'}(A_\beta)$ determined by an assignment p' such that $p' = p\backslash x_\alpha$ and $p'(x_\alpha) = m_\alpha$.

A model of a theory T is a general frame $\mathscr{M}$ for which $\mathscr{M}_p(A_o) = \mathbf{1}$ holds for all axioms A_o of T and all assignments $p \in$ Asg. A formula A_o is true in the theory T, $T \models A_o$ if it is true in the degree $\mathbf{1}$ in all its models.

[2] Recall that the description operator represents, in fact, the defuzzification operation (cf. [14, Chapt. 3]).

5 Canonical Model of Partial FTT

5.1 Extension of Theories

The canonical model of EQ-FTT with partial functions can be obtained by extension of the canonical model for the basic FTT (see [9,10]) where we define a special function $\mathscr{V}$, whose domain and range are equivalence classes of formulas $|A_\alpha|$ obtained using $A_\alpha \approx B_\alpha$ iff $T \vdash A_\alpha \equiv B_\alpha$, $\alpha \in Types$.

A theory T is *linear* if for every two formulas A_o, B_o such that $T \vdash !A_o$ and $T \vdash !B$, $\vdash A_o \Rightarrow B_o$ or $T \vdash B_o \Rightarrow A_o$. It is *extensionally complete* if for every closed formula of the form $A_{\beta\alpha} \equiv B_{\beta\alpha}$, $T \not\vdash A_{\beta\alpha} \equiv B_{\beta\alpha}$ it follows that there is a closed formula C_α such that $T \vdash !C_\alpha$ and $T \not\vdash A_{\beta\alpha} C_\alpha \equiv B_{\beta\alpha} C_\alpha$.

Theorem 10. *Every consistent theory T can be extended to a maximally consistent, extensionally complete and linear consistent theory $\overline{T}$.*

Now we will put

$$^T E = \{|A| \mid A \in Form_o, A \text{ closed}, T \vdash !A\}, \tag{8}$$

$$^T E^* = {}^T E \cup \{|*_o|\} \tag{9}$$

where $*_o$ is defined in (4). It follows from the definition of "≈" that

$$|*_o| = \{A \mid A \in Form_o, T \vdash ?A\}.$$

Theorem 11. *Let T be a consistent linear extensionally complete theory. Then the algebra*

$$^T\mathscr{E}^* = \langle {}^T E^*, {}^T\wedge, {}^T\otimes, {}^T\sim, {}^T\Delta, {}^T\mathbf{1}, {}^T\mathbf{0}\rangle \tag{10}$$

is an extended linearly ordered good EQ$_\Delta$-algebra.

5.2 Canonical Frame and Completeness

Let T be a consistent, linear and extensionally complete theory. Then the canonical frame is

$$\mathscr{M}^* = \langle (M^*_\alpha, \mathring{=}_\alpha)_{\alpha \in Types}, {}^T\mathscr{E}^*_\Delta, I_o, I_\epsilon\rangle \tag{11}$$

1. $M^*_\alpha = \{\mathscr{V}(A_\alpha) \mid A_\alpha \in Form_\alpha, A_\alpha \text{ closed}\}$, $\alpha \in Types$, where $\mathscr{V}$ is defined inductively:
   (i) If $\alpha = o$ then

$$\mathscr{V}(A_o) = |A_o|,$$
$$\mathscr{V}(*_o) = |*_o| = |\iota_{o(oo)} \cdot \lambda x_o \bot|.$$

   (ii) If $\alpha = \epsilon$ then $\mathscr{V}(A_\epsilon) = |A_\epsilon|$,

$$\mathscr{V}(A_\epsilon) = |A_\epsilon|,$$
$$\mathscr{V}(*_\epsilon) = |*_\epsilon| = |\iota_{\epsilon(o\epsilon)} \cdot \lambda x_\epsilon\ *_o|.$$

(iii) If $\alpha = \gamma\beta$ then we put $\mathscr{V}(A_{\gamma\beta}) \subseteq M^*_\beta \times M^*_\gamma$ which is a relation consisting of couples

$$\langle \mathscr{V}(B_\beta), \mathscr{V}(A_{\gamma\beta}B_\beta)\rangle$$

for all closed $B_\beta \in \mathit{Form}_\beta$ and $A_{\gamma\beta} \in \mathit{Form}_{\gamma\beta}$. As a special case,

$$\mathscr{V}(*_{\gamma\beta}) = \{\langle \mathscr{V}(A_\beta), \mathscr{V}(*_\gamma)\rangle \mid \mathscr{V}(A_\beta) \in M^*_\beta\}.$$

The fuzzy equality on M^*_α is defined by

$$[\mathscr{V}(A_\alpha) \doteq_\alpha \mathscr{V}(B_\alpha)] = |A_\alpha \equiv B_\alpha|. \tag{12}$$

2. The ${}^T\mathscr{E}^*_\Delta$ is the extended EQ$_\Delta$-algebra from Theorem 11.
3. Let $\alpha \in \{o, \epsilon\}$ and $A_{o\alpha}$ be a formula such that $T \vdash !A_{o\alpha}$. Let $\mathscr{V}(A_{o\alpha}) = \{\langle \mathscr{V}(B_\alpha), \mathscr{V}(A_{o\alpha}B_\alpha)\rangle \mid \mathscr{V}(B_\alpha) \in M^*_\alpha\}$ be a fuzzy set on M^*_α. Then we put

$$I(\mathscr{V}(A_{o\alpha})) = \begin{cases} |\iota_{\alpha(o\alpha)}A_{o\alpha}|, \text{ if } \vdash (\exists x_\alpha)\boldsymbol{\Delta}(A_{o\alpha}x_\alpha), \\ \mathscr{V}(*_\alpha), \text{ if } \vdash (\forall x_\alpha)(\neg\boldsymbol{\Delta}(A_{o\alpha}x_\alpha \equiv \top)). \end{cases}$$

Theorem 12. (Completeness of FTT with partial functions) *Let T be a theory, the special axioms of which have the form $A_o \equiv \top$. Then T is consistent iff it has a general model $\mathscr{M}$.*

6 Partial Functions

Let us briefly demonstrate that the results by Lapierre and Lepage in [6,7] can be easily expressed in FTT. We define:

(i) *Total function* $\mathrm{TotF}_{o(\beta\alpha)} \equiv \lambda f_{\beta\alpha} \cdot (\forall x_\alpha)(!f_{\beta\alpha}x_\alpha)$
(ii) *Partial function* $\mathrm{PartF}_{o(\beta\alpha)} \equiv \lambda f_{\beta\alpha} \cdot (\exists x_\alpha)(!x_\alpha \,\&\, ?f_{\beta\alpha}x_\alpha)$

A function $f_{\beta\alpha}$ is *strict* if $\vdash f_{\beta\alpha}*_\alpha \equiv *_\beta$. It is *non-strict* if $\vdash !f_{\beta\alpha}*_\alpha$.

We will also introduce a special ordering:

$$\lhd_{(o\alpha)\alpha} \equiv \lambda x_\alpha\, \lambda y_\alpha \cdot ?x \vee \boldsymbol{\Delta}(x_\alpha \equiv y_\alpha). \tag{13}$$

It can be proved that this relation is crisp and is indeed an ordering. On the basis of of it, we can define *monotonous functions:*

$$\mathrm{MonF}_{o(\beta\alpha)} \equiv \lambda f_{\beta\alpha} \cdot (\forall x_\alpha)(\forall y_\alpha)((x_\alpha \lhd y_\alpha) \equiv (f_{\beta\alpha}x_\alpha \lhd f_{\beta\alpha}y_\alpha)). \tag{14}$$

Lapierre and Lepage gave many arguments in favor of the idea to consider all the functions to be monotonous in the sense of (14).

Lemma 1. *Let T be a consistent theory in which $f_{\beta\alpha}$ is total and non-strict such that the following is provable:*

$$T \vdash (\exists x_\alpha)\boldsymbol{\Delta}(!x_\alpha \,\&\, \neg\boldsymbol{\Delta}(f_{\beta\alpha}x_\alpha \equiv f_{\beta\alpha}*_\alpha)). \tag{15}$$

Then $T \not\vdash$ MonF $f_{\beta\alpha}$.

The function $f_{\beta\alpha}$ in Lemma 1 extends the domain from M_α to M^*_α assigning $*_\alpha$ an element from M_α. Such a function is not monotonous.

7 Conclusion

In this paper, we studied possibility how partial function can be introduced in fuzzy type theory. We have chosen the most general kind of FTT based on the EQ-algebra of truth degrees introduced in [10]. For each type α we introduced a special value "undefined" (denoted by $*_\alpha$) laying outside of the given domain. In our construction, we used the fact that the description operator $\iota_{\alpha(o\alpha)}$ gives no result when applied to formulas representing subnormal fuzzy sets. Hence, we defined $*_o$ as the formula $\iota_{o(oo)} \cdot \lambda x_o \bot$ and accordingly the other values "undefined". This made it possible to have "undefined" inside the set Form of all the formulas without necessity to extend the language by new constant with special properties.

One may argue that, in fact, all our functions remain total and we only added a special interpretation to one specific functional value. This is true. But we must realize that what we provide is a *mathematical model*. And from this point of view, we argue that there is no principal difference when considering a function $f(x)$ to be undefined for a given x or, that $f(x) = *$ where $*$ is a special value laying outside of the original domain and range of f. Our theory makes it also possible to introduce other kinds of undefined values, for example "error" or "missing value".

Acknowledgment. This paper was supported by the grant 16-19170S of GAČR, Czech Republic.

References

1. Andrews, P.: An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof. Kluwer, Dordrecht (2002)
2. Duží, M., Jespersen, B., Materna, P.: Procedural Semantics for Hyperintensional Logic. Springer, Dordrecht (2010)
3. El-Zekey, M.: Representable good EQ-algebras. Soft Comput. **14**, 1011–1023 (2009)
4. El-Zekey, M., Novák, V., Mesiar, R.: On good EQ-algebras. Fuzzy Sets Syst. **178**, 1–23 (2011)
5. Farmer, W.M.: A partial functions version of Church's simple theory of types. J. Symb. Log. **55**, 1269–1291 (1990)
6. Lappiere, S.: A functional partial semantics for intensional logic. Notre Dame J. Form. Log. **33**, 517–541 (1992)
7. Lepage, F.: Partial functions in type theory. Notre Dame J. Form. Log. **33**, 493–516 (1992)
8. Novák, V.: Descriptions in the full fuzzy type theory. Neural Netw. World **5**, 559–565 (2003)
9. Novák, V.: On fuzzy type theory. Fuzzy Sets Syst. **149**, 235–273 (2005)
10. Novák, V.: EQ-algebra-based fuzzy type theory and its extensions. Log. J. IGPL **19**, 512–542 (2011)
11. Novák, V.: Fuzzy type theory, descriptions, and partial functions. In: Galichet, S., Montero, J., Mauris, G. (eds.) Proceedings of 7th International Conference EUSFLAT-2011 and LFA-2011, pp. 189–195. Atlantis Press, Amsterdam (2011)

12. Novák, V.: Fuzzy type theory with partial functions. Fuzzy Sets Syst. (submitted)
13. Novák, V., de Baets, B.: EQ-algebras. Fuzzy Sets Syst. **160**, 2956–2978 (2009)
14. Novák, V., Perfilieva, I., Dvořák, A.: Insight into Fuzzy Modeling. Wiley, Hoboken (2016)
15. Tichý, P.: Foundations of partial type theory. Rep. Math. Log. **14**, 59–72 (1982)

Dynamic Intuitionistic Fuzzy Evaluation of Entrepreneurial Support in Countries

Sezi Cevik Onar(✉), Basar Oztaysi, and Cengiz Kahraman

Industrial Engineering, ITU Management Faculty,
Istanbul Technical University, Macka, Besiktas, Istanbul, Turkey
{cevikse, oztaysib, kahramanc}@itu.edu.tr

Abstract. Entrepreneurship includes various activities including starting a new business from scratch, creating and developing new business areas for existing organizations. The countries should provide a supportive entrepreneurial environment since entrepreneurship is the key element for the sustainable growth of a country. In order to improve entrepreneurial support, the level of support should be measured with different dimensions in the various periods. In this study, a dynamic intuitionistic fuzzy evaluation method is developed for determining entrepreneurial support within a country. Five countries are evaluated with the proposed method, and a sensitivity analysis is conducted to show the robustness of the model.

Keywords: Dynamic intuitionistic fuzzy sets · Entrepreneurial support · Multi-criteria decision making

1 Introduction

Today's modern economy affected all sectors and companies in different ways. There exist newly introduced business processes and production techniques. Not only new technologies but also the entrepreneurs who have a pioneering role in implementing new technology and new business models. Entrepreneurship is a growing concept of this modern era; it receives interest from many disciplines such as economics, management, and sociology. Entrepreneurship becomes the trigger of economic developments in many countries. Since entrepreneurship creates new job opportunities and new sources of productivity, it is a critical method to cope with unemployment problems. Also, entrepreneurship takes its source from innovation, and it is a key for innovation implementations. Entrepreneurship enables to meet the demands of the customers better, and it ensures business growth. Entrepreneurship is a motivation source for change.

In order to be successful in the future, governments need to support entrepreneurship. The performance of this entrepreneurial support can be improved by measurement. This measurement involves uncertain and vague dimensions that should be measured at different periods. Intuitionistic fuzzy sets enable defining both membership and non-membership values and are excellent tools for dealing with imprecision. In this study, a dynamic intuitionistic fuzzy entrepreneurial support evaluation model is developed. The remaining of the paper is organized as follows: Sect. 2 explains the entrepreneurial support dimensions. In Sect. 3, the preliminaries of intuitionistic fuzzy

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_4

sets and the steps of dynamic intuitionistic fuzzy evaluation methods are given. In Sect. 4, the proposed model is used to evaluate entrepreneurial support of five countries. Conclusion Section concludes the paper and offers further suggestions.

2 Entrepreneurial Support

Entrepreneurship is a multidimensional concept that involves uncertainties. Schumpeter (1934) indicate the crucial role of entrepreneurship in economy and society. In Schumpeter's (1934) definition, the entrepreneur is declared as the person of the cause of economic development. In 1980's, Schultz (1982) criticize previous definitions since they ignore the entrepreneurs' behaviors, characteristics, and interactions with their environment. Drucker (1993) claim that innovation is the key element of entrepreneurship. Entrepreneurship is not only establishing a new business from the beginning, but also creating new ventures and developing new sectors of activity in established companies.

Entrepreneur is the one who seeks for opportunities and tries to achieve them by taking all the risks. Entrepreneurship involves all these risk taking, opportunity seeking, implementation and innovation processes of the entrepreneurs. Thus, establishing a new firm and making innovations are included in entrepreneurship process. The factors that have a negative impact on potential entrepreneurs can be listed as:

- Changes in government policies, exchange rates, inflation rate, loan interests and demand rates
- High finance charges
- Insufficient investment opportunities, labor force, incentives and audit
- High level of competition
- Ambiguous state policies
- Inadequate level of cooperation between industrial institutions and universities

Six main expectations of entrepreneurs can help compensating these disincentive factors:

- Easier bureaucratic procedures
- Implementing modern and permanent government policies
- Satisfying financial needs
- Increasing incentives
- Accepting standards and inspecting them
- Eliminating infrastructure inabilities

Attractive entrepreneurial environments enable compensating negative impacts. Education and experience level of the country, connections, cultural factors, government policies, technological developments and funds are the main determinants that show the entrepreneurial support within a country. Table 1 summarizes the entrepreneurial support dimensions.

Table 1. Entrepreneurial support dimensions

Factor	Definition	References
Education and experience level of the country (EE)	Investigating the effects of education levels and previous work experiences of the entrepreneurs	Singer et al. (2015), Cansız (2014), Jain and Ali (2013), Keat et al. (2011), Gerba (2012), Grilo and Thurik (2008), Karadeniz (2010), Van der Sluis et al. (2008), Othman et al. (2006), Kristiansen and Indarti (2004), Ghazali et al. (1995)
Connections (CN)	Identifying possible outcomes of having connections in chosen business area for entrepreneurial activity	Jain and Ali (2013), Bygrave and Zacharakis (2011), Eyal (2008), Xu et al. (2008), Baum et al. (2000), Preiss et al. (1996), Powell and Brantley (1992), Sacks et al. (2001), Helper (1990), Nohria and Eccles (1992), Klyver and Foley (2012), Zhao et al. (2006), Witt (2004)
Cultural Factors (CF)	Defining how entrepreneurs are affected from their nations' and close environments' manners towards them	Jain and Ali (2013), Tracy (2013), Karadeniz (2010), Singer et al. (2015), Ratten (2014), Klyver and Foley (2012), Morrison (2000), Begley and Tan (2001), Thomas and Mueller (2000), Davidsson and Wiklund (1997), Shane (1992), Ettlie et al. (1993), Tiessen (1997), Lee and Peterson (2000)
Government Policies (GP)	Identifying governmental policies applied to the entrepreneurs and ways to improve them in order to increase the level of entrepreneurial activity	Sebora et al. (2009), Jain and Ali (2013), Çetindamar et al. (2012); Karadeniz (2010), Singer et al. (2015), Cansız (2014), Minniti (2008), Smallbone et al. (2010), Tende (2014), Ratten (2014), Dana (2004); Soriano and Galindo-Martín (2012), Skica et al. (2013)
Technological Developments (TD)	Determining the necessary technological opportunities and related institutions especially for technology related start ups	Cansız (2014), Örnek & Danyal (2015), Robson et al. (2009)
Funds (FU)	Giving information about different ways of fund raising and concerning institutions, banks, venture capital markets, angel investors	Jain and Ali (2013), Cansız (2014), Kuratko (2014), De Bettignies and Brander (2007), Landier (2003), Becker-Blease and Sohl (2015), Calopa et al. (2014); Moy (2014), Khanin et al. (2012), Keuschnigg and Nielsen (2003), Rogers (2012), Ho and Wong (2007), Kreft and Sobel (2005), Black and Strahan (2002), Hein et al. (2005), Breit and Arano (2009)

3 Dynamic Intuitionistic Fuzzy Evaluation

3.1 Preliminaries

An intuitionistic fuzzy set $\tilde{I}$ in a given set X can be defined as in Eq. (1).

$$\tilde{I} = \left\{ \langle x, \mu_{\tilde{I}}(x), v_{\tilde{I}}(x), \rangle; x \in X \right\}, \tag{1}$$

where $\mu_{\tilde{I}}$ is the membership degree $\mu_{\tilde{I}} : X \to [0,1]$ and $v_{\tilde{I}} : X \to [0,1]$ $v_{\tilde{I}}$ is the non-membership degree satisfying the condition $0 \leq \mu_{\tilde{I}}(x) + v_{\tilde{I}}(x) \leq 1$, for every $x \in X$.

Let $\tilde{I}(t) = (\mu_{\widetilde{I(t)}}, v_{\widetilde{I(t)}}, \pi_{\widetilde{I(t)}})$ be a dynamic intuitionistic fuzzy variable is a variable with a time dimension. $\tilde{I}(t)$ satisfies Eq. (2) (Xu and Cai 2012):

$$\mu_{\alpha(t)} + v_{\alpha(t)} \leq 1, \pi_{\alpha(t)} = 1 - \mu_{\alpha(t)} - v_{\alpha(t)} \tag{2}$$

where t is a time variable, $\mu_{\widetilde{I(t)}} \in [0,1]$, $\mu_{\widetilde{I(t)}} \in [0,1]$.

Let $\widetilde{I(t_1)}, \widetilde{I(t_2)}, \ldots, \widetilde{I(t_n)}$ be the intuitionistic fuzzy numbers collected at n different periods. The addition and multiplication operations can be given as in Eqs. (3) and (4) (Xu and Cai 2012).

$$\widetilde{I(t_1)} \oplus \widetilde{I(t_2)} = (\mu_{\widetilde{I(t_1)}} + \mu_{\widetilde{I(t_2)}} - \mu_{\widetilde{I(t_1)}}\mu_{\widetilde{I(t_2)}}, v_{\widetilde{I(t_1)}} v_{\widetilde{I(t_2)},} \left(1 - \mu_{\widetilde{I(t_1)}}\right) + \left(1 - \mu_{\widetilde{I(t_2)}}\right) - v_{\widetilde{I(t_1)}} v_{\widetilde{I(t_2)}}) \tag{3}$$

$$\lambda\widetilde{I(t_1)} = (1 - (1 - \mu_{\widetilde{I(t_1)}})^{\lambda}, v^{\lambda}_{\widetilde{I(t_1)}}, \left(1 - \mu_{\widetilde{I(t_1)}}\right)^{\lambda} - v^{\lambda}_{\widetilde{I(t_1)}}), \lambda > 0 \tag{4}$$

$I_i^+ = (1, 0, 0)$ represents the largest intuitionistic fuzzy number where $I_i^- = (0, 1, 0)$ represents the smallest intuitionistic fuzzy number.

3.2 Dynamic Intuitionistic Fuzzy Evaluation Method

In this Section, the steps of dynamic intuitionistic fuzzy evaluation method developed by Xu and Yager (2008) is given.

Step 1. Define the criteria and alternatives.

Step 2. Evaluate alternatives at different periods using dynamic intuitionistic fuzzy numbers.

Step 3. Determine the weight vector $\omega(t) = \left(\omega(t_1), \omega(t_2), \ldots, \omega\left(t_p\right)\right)$ by using Eq. (5).

$$\omega(t_{k+1}) - \omega(t_k) = c, \ \omega(t_k) = \eta + (k-1)c \tag{5}$$

where t_k is the kth period, ω_{t_k} is the weight at the t_k^{th} period $(t_n) \in [0,1]$ and $\sum_{k=1}^{p} \omega(t_k) = 1$.

Step 4. Aggregate dynamic intuitionistic fuzzy numbers using Eq. (6).

$$\widetilde{DIFWA}_{\omega(t)}^{\left(\widetilde{I}(t_1),\widetilde{I}(t_2),\ldots,\widetilde{I}(t_n)\right)} = \left(1-\prod_{k=1}^{n}\left(1-\mu_{\widetilde{I}(t_k)}\right)^{\omega(t_k)}, \prod_{k=1}^{n}\nu_{\widetilde{I}(t_k)}^{\omega(t_k)}, \prod_{k=1}^{n}\left(1-\mu_{\widetilde{I}(t_k)}\right)^{\omega(t_k)} - \prod_{k=1}^{n}\nu_{\widetilde{I}(t_k)}^{\omega(t_k)}\right) \tag{6}$$

Step 5. Determine the intuitionistic fuzzy ideal $Y^+ = \left(I_1^+, I_2^+, \ldots, I_m^+\right)^T$ and negative ideal solutions $Y^+ = \left(I_1^-, I_2^-, \ldots, I_m^-\right)^T$.

Step 6. Calculate closeness coefficient for the i[th] alternative using Eq. (7).

$$C_i = \frac{\sum_{j=1}^{m} w_j(1-\nu_{ij})}{\sum_{j=1}^{m} w_j(1+\pi_{ij})}, \tag{7}$$

where i = 1,2,…,n and w_j is the weight of the j[th] attribute

4 Entrepreneurial Support Evaluation

In this Section, the last years' entrepreneurial support level of five countries are evaluated by using dynamic intuitionistic fuzzy evaluation. Three experts with academic background quarterly assess the entrepreneurial support of five countries. Education and

Table 2. Dynamic intuitionistic fuzzy evaluations

	EE	CN	CF	GP	TD	FU
t1						
Country 1	(0.1, 0.8, 0.1)	(0.2, 0.7, 0.1)	(0.7, 0.2, 0.1)	(0.1, 0.5, 0.4)	(0.6, 0.2, 0.2)	(0.2, 0.2, 0.6)
Country 2	(0.9, 0.1, 0)	(0.2, 0.1, 0.7)	(0.8, 0.1, 0.1)	(0.3, 0.6, 0.1)	(0.5, 0.2, 0.3)	(0.5, 0.1, 0.4)
Country 3	(0.2, 0.4, 0.4)	(0.4, 0.4, 0.2)	(0.5, 0.3, 0.2)	(0.1, 0.8, 0.1)	(0.7, 0.1, 0.2)	(0.2, 0.2, 0.8)
Country 4	(0.8, 0.1, 0.1)	(0.8, 0.1, 0.1)	(0.3, 0.3, 0.4)	(0.7, 0.2, 0.1)	(0.7, 0.1, 0.2)	(0.8, 0.1, 0.1)
Country 5	(0.2, 0.4, 0.4)	(0.5, 0.4, 0.1)	(0.6, 0.3, 0.1)	(0.6, 0.2, 0.2)	(0.2, 0.7, 0.1)	(0.8, 0.1, 0.1)
t2						
Country 1	(0.5, 0.3, 0.2)	(0.2, 0.4, 0.4)	(0.5, 0.3, 0.2)	(0.1, 0.2, 0.7)	(0.3, 0.4, 0.3)	(0.5, 0.2, 0.3)
Country 2	(0.5, 0.3, 0.2)	(0.3, 0.5,0.2)	(0.2, 0.1, 0.7)	(0.1, 0.1, 0.8)	(0.4, 0.3, 0.3)	(0.8, 0.1, 0.1)
Country 3	(0.6, 0.3, 0.1)	(0.7, 0.2, 0.1)	(0.3, 0.5, 0.2)	(0.7, 0.1, 0.2)	(0.3, 0.5, 0.2)	(0.6, 0.1, 0.3)
Country 4	(0.3, 0.2, 0.5)	(0.5, 0.1, 0.4)	(0.4, 0.3, 0.3)	(0.9, 0.1, 0)	(0.2, 0.3, 0.5)	(0.2, 0.3, 0.5)
Country 5	(0.6, 0.2, 0.2)	(0.4, 0.3, 0.3)	(0.4, 0.2, 0.4)	(0.8, 0.1, 0.1)	(0.3, 0.1, 0.6)	(0.6, 0.2, 0.2)
t3						
Country 1	(0.9, 0.1, 0.1)	(0.2, 0.5, 0.3)	(0.8, 0.1, 0.1)	(0.9, 0.1, 0)	(0.4, 0.3, 0.3)	(0.1, 0.4, 0.5)
Country 2	(0.8, 0.1, 0.1)	(0.8, 0.1, 0.1)	(0.3, 0.7, 0)	(0.4, 0.3, 0.3)	(0.5, 0.3, 0.2)	(0.2, 0.7, 0.1)
Country 3	(0.8, 0.1, 0.1)	(0.2, 0.5, 0.3)	(0.5, 0.5, 0)	(0.2, 0.2, 0.6)	(0.3, 0.5, 0.2)	(0.6, 0.4, 0)
Country 4	(0.8, 0.2, 0)	(0.1, 0.5, 0.4)	(0.2, 0.5, 0.3)	(0.3, 0.2, 0.5)	(0.9, 0.1, 0)	(0.5, 0.2, 0.3)
Country 5	(0.4, 0.4, 0.2)	(0.6, 0.2, 0.2)	(0.4, 0.5, 0.1)	(0.9, 0.1, 0)	(0.3, 0.2, 0.5)	(0.7,0.1, 0.2)
t4						
Country 1	(0.2, 0.2, 0.6)	(0.7, 0.2, 0.1)	(0.1, 0.8, 0.1)	(0.4, 0.2, 0.4)	(0.4, 0.2, 0.4)	(0.2, 0.7, 0.1)
Country 2	(0.2, 0.4, 0.4)	(0.2, 0.3, 0.5)	(0.7, 0.1, 0.2)	(0.5, 0.1, 0.4)	(0.8, 0.1, 0.1)	(0.2, 0.6, 0.2)
Country 3	(0.8, 0.1, 0.1)	(0.5, 0.1, 0.4)	(0.5, 0.4, 0.1)	(0.5, 0.2, 0.3)	(0.9, 0.1, 0)	(0.3, 0.1, 0.6)
Country 4	(0.7, 0.2, 0.1)	(0.8, 0.2, 0)	(0.3, 0.3, 0.4)	(0.9, 0.1, 0)	(0.4, 0.4, 0.2)	(0.7, 0.2, 0.1)
Country 5	(0.6, 0.2, 0.2)	(0.7, 0.1, 0.2)	(0.5, 0.3, 0.2)	(0.5, 0.2, 0.3)	(0.8, 0.1, 0)	(0.9, 0, 0.1)

experience level of the country (EE), connections (CN), cultural factors (CF), government policies (GP), technological developments (TD) and funds (FU) are the evaluation factors. The compromised dynamic intuitionistic evaluations are given in Table 2.

The dynamic evaluations are aggregated by using Eq. (6). The aggregated score of the countries is given in Table 3.

Table 3. Aggregated intuitionistic fuzzy evaluations

Aggregated			
	EE	CN	CF
Country 1	(0.605, 0.202, 0.193)	(0.46, 0.343, 0.197)	(0.543, 0.307, 0.15)
Country 2	(0.61, 0.217, 0.173)	(0.486, 0.214, 0.3)	(0.548, 0.179, 0.273)
Country 3	(0.736, 0.143, 0.121)	(0.471, 0.214, 0.315)	(0.465, 0.435, 0.1)
Country 4	(0.698, 0.187, 0.115)	(0.623, 0.214, 0.163)	(0.294, 0.35, 0.356)
Country 5	(0.516, 0.264, 0.22)	(0.605, 0.176, 0.219)	(0.464, 0.322, 0.214)
Aggregated			
	GP	TD	FU
Country 1	(0.604, 0.178, 0.218)	(0.406, 0.259, 0.335)	(0.246, 0.406, 0.348)
Country 2	(0.386, 0.166, 0.448)	(0.641, 0.186, 0.173)	(0.422, 0.367, 0.211)
Country 3	(0.449, 0.2, 0.351)	(0.705, 0.224, 0.071)	(0.464, 0.162, 0.374)
Country 4	(0.8, 0.132, 0.068)	(0.654,0.217, 0.129)	(0.591, 0.202, 0.207)
Country 5	(0.749, 0.141, 0.11)	(0.57, 0.15, 0.28)	(0.803, 0, 0.197)

The closeness coefficient for the countries is calculated by using Eq. (7) (Table 4).

Table 4. Closeness coefficient for the countries

	Closeness coefficient
Country 1	0.579
Country 2	0.616
Country 3	0.63
Country 4	0.668
Country 5	0.683

Country 5 provides the highest entrepreneurial support whereas Country 3 provides the least entrepreneurial support.

In order to check the robustness of the results, we apply one at a time sensitivity analysis. In this sensitivity analysis, the weight of the entrepreneurial support factors is changed. Figure 1 illustrates the results of sensitivity analysis.

The sensitivity analysis indicates that the change in the criteria weights has a significant impact on the entrepreneurial support. Countries should define their priorities and improve entrepreneurial support based on these preferences.

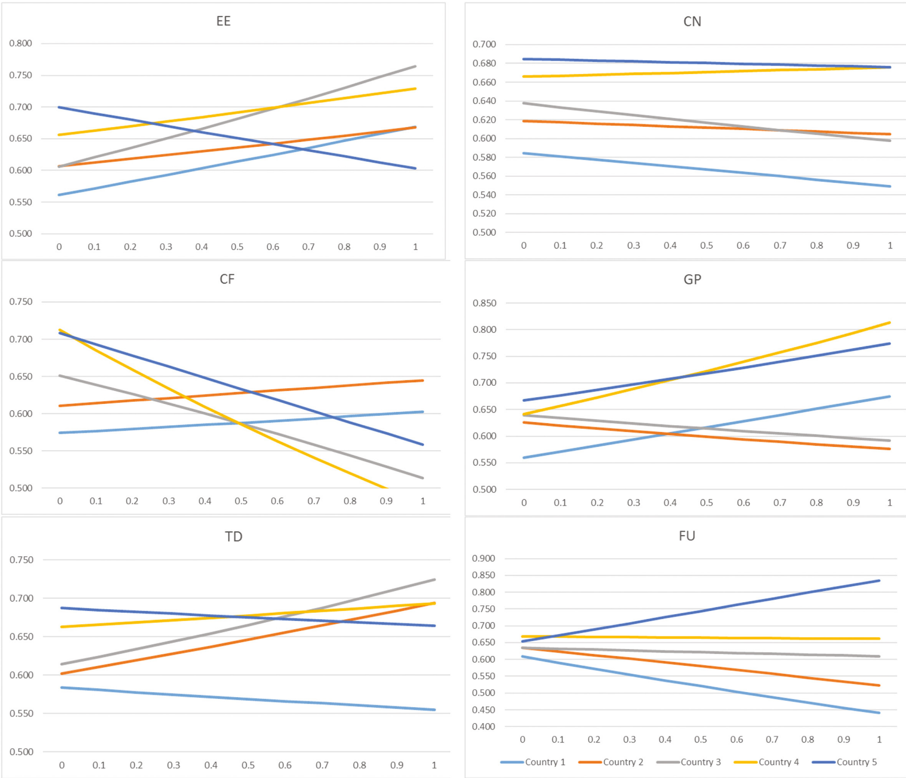


Fig. 1. Sensitivity analysis

5 Conclusion

In this study, dynamic intuitionistic fuzzy evaluation method is used to define the entrepreneurial support within a country. Education and experience level of the country, connections, cultural factors, government policies, technological developments, and funds are the main criteria that affect entrepreneurial support. This dynamic, multi-criteria approach enables us using different criteria and making the evaluation at various periods. The results indicate that the criteria weights have a significant impact on the final entrepreneurial support.

For the further studies, the weights of the criteria can be defined with a hierarchical approach. The other extensions of fuzzy sets such as Type-2 fuzzy sets, hesitant fuzzy sets or Pythagorean fuzzy sets can be used to model entrepreneurial support.

References

Baum, J.A., Calabrese, T., Silverman, B.S.: Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. Strateg. Manag. J. **21**(3), 267–294 (2000)

Becker-Blease, J.R., Sohl, J.E.: New venture legitimacy: the conditions for angel investors. Small Bus. Econ. **45**(4), 735–749 (2015)

Begley, T.M., Tan, W.L.: The socio-cultural environment for entrepreneurship: a comparison between East Asian and Anglo-Saxon countries. J. Int. Bus. Stud. **32**(3), 537–553 (2001)

Black, S.E., Strahan, P.E.: Entrepreneurship and bank credit availability. J. Financ. **57**(6), 2807–2833 (2002)

Breit, E., Arano, K.: Reduction of information asymmetries in small business lending: evidence from the survey of small business finances (SSBF). J. Financ. Econ. Pract. 71–85 (2009)

Bygrave, W., Zacharakis, A.: Entrepreneurship. Wiley, Hoboken (2011)

Calopa, M., Horvat, J., Lalić, M.: Analysis of financing sources for start-up companies. Manag. J. Contemp. Manag. Issues. **19**(2), 19–44 (2014)

Cansız, M.: Innovative Entrepreneurs of Turkey: The Case of Turkish Technoparks. Republic of Turkey, Ministry of Development, Turkey (2014)

Çetindamar, D., Gupta, V.K., Karadeniz, E.E., Egrican, N.: What the numbers tell: the impact of human, family and financial capital on women and men's entry into entrepreneurship in Turkey. Entrep. Reg. Dev. **24**(1–2), 29–51 (2012)

Dana, L.P. (ed.): Handbook of Research on International Entrepreneurship. Edward Elgar Publishing, Cheltenham (2004)

Davidsson, P., Wiklund, J.: Values, beliefs and regional variations in new firm formation rates. J. Econ. Psychol. **18**(2), 179–199 (1997)

De Bettignies, J.E., Brander, J.A.: Financing entrepreneurship: bank finance versus venture capital. J. Bus. Ventur. **22**(6), 808–832 (2007)

Drucker, P.F.: Managing in Turbulent Times. Routledge, London (1993)

Ettlie, J.E., Dreher, C., Kovacs, G.L., Trygg, L.: Cross-national comparisons of product development in manufacturing. J. High Technol. Manag. Res. **4**(2), 139–155 (1993)

Eyal, O.: Caught in the net: the network-entrepreneurship connection in public schools. Int. J. Edu. Manag. **22**(5), 386–398 (2008)

Gerba, D.: Impact of entrepreneurship education on entrepreneurial intentions of business and engineering students in Ethiopia. Afr. J. Econ. Manag. Stud. **3**(2), 258–277 (2012)

Ghazali, A., Ghosh, B.C., Tay, R.S.T.: The determinants of self-employment choice among university graduates in Singapore. Int. J. Manag. **12**, 26–35 (1995)

Grilo, I., Thurik, R.: Determinants of entrepreneurial engagement levels in Europe and the US. Ind. Corp. Change **17**(6), 1113–1145 (2008)

Hein, S.E., Koch, T.W., MacDonald, S.S.: On the uniqueness of community banks. Fed. Reserve Bank Atlanta Econ. Rev. **90**(1), 15–36 (2005)

Helper, S.: Comparative supplier relations in the US and Japanese auto industries: an exit/voice approach. Bus. Econ. Hist. **19**, 153–162 (1990)

Ho, Y.P., Wong, P.K.: Financing, regulatory costs and entrepreneurial propensity. Small Bus. Econ. **28**(2–3), 187–204 (2007)

Jain, R., Ali, S.W.: A review of facilitators, barriers and gateways to entrepreneurship: directions for future research. South Asian J. Manag. **20**(3), 122–163 (2013)

Karadeniz, E.: Entrepreneurship in Turkey 2010. The Global Entrepreneurship Monitor Report, Turkey (2010)

Keat, O.Y., Selvarajah, C., Meyer, D.: Inclination towards entrepreneurship among university students: an empirical study of Malaysian university students. Int. J. Bus. Soc. Sci. **2**(4), 206–220 (2011)
Keuschnigg, C., Nielsen, S.B.: Tax policy, venture capital, and entrepreneurship. J. Public Econ. **87**(1), 175–203 (2003)
Khanin, D., Ogilvie, K., Leibsohn, D.: International entrepreneurship, venture capital networks, and reinvestment decisions. J. Int. Entrep. **10**(1), 1–24 (2012)
Klyver, K., Foley, D.: Networking and culture in entrepreneurship. Entrep. Reg. Dev. **24**(7–8), 561–588 (2012)
Kreft, S.F., Sobel, R.S.: Public policy, entrepreneurship, and economic freedom. Cato J. **25**(3), 595–616 (2005)
Kristiansen, S., Indarti, N.: Entrepreneurial intention among Indonesian and Norwegian students. J. Enterp. Culture **12**(1), 55–78 (2004)
Kuratko, D.F.: Introduction to Entrepreneurship. Cengage Learning, Canada (2014)
Landier, A.: Start-up financing: from banks to venture capital. Unpublished working paper, University of Chicago, Chicago, IL (2003)
Lee, S.M., Peterson, S.J.: Culture, entrepreneurial orientation, and global competitiveness. J. World Bus. **35**(4), 401–416 (2000)
Minniti, M.: The role of government policy on entrepreneurial activity: productive, unproductive, or destructive? Entrep. Theory Pract. **32**(5), 779–790 (2008)
Morrison, A.: Entrepreneurship: what triggers it? Int. J. Entrep. Behav. Res. **6**(2), 59–71 (2000)
Moy, R.L.: Swimming with the sharks: case studies in venture capital entrepreneurship. J. Acad. Bus. Educ. **15**, 56–65 (2014)
Nohria, N., Eccles, R.C.: Networks and Organizations: Structure, Form and Action. Harvard Business School Press, Boston (1992)
Örnek, A.S., Danyal, Y.: Increased importance of entrepreneurship from entrepreneurship to techno-entrepreneurship (startup): provided supports and conveniences to techno-entrepreneurs in Turkey. Procedia-Social Behav. Sci. **195**, 1146–1155 (2015)
Othman, M.N., Ghazali, E., Sung, Y.S.: Graduate versus non-graduate entrepreneurs in urban Malaysia: some insights into entrepreneurial personality, company and family background differences. J. Int. Bus. Entrep. Dev. **3**(1–2), 57–76 (2006)
Powell, W.W., Brantley, P.: Competitive cooperation in biotechnology: learning through networks? In: Nohria, N., Eccles, R.G. (eds.) Networks and Oganizations: Structure, Form, and Action, pp. 366–394. Harvard Business School Press, Boston (1992)
Preiss, K., Goldman, S.L., Nagel, R.N.: Cooperate to Compete: Building Agile Business Relationships. John Wiley and Sons, Hoboken (1996)
Ratten, V.: Encouraging collaborative entrepreneurship in developing countries: the current challenges and a research agenda. J. Entrep. Emerg. Econ. **6**(3), 298–308 (2014)
Robson, P.J., Haugh, H.M., Obeng, B.A.: Entrepreneurship and innovation in Ghana: enterprising Africa. Small Bus. Econ. **32**(3), 331–350 (2009)
Rogers, T.M.: Bank market structure and entrepreneurship. Small Bus. Econ. **39**(4), 909–920 (2012)
Sacks, M., Ventresca, M., Uzzi, B.: Network sources of national competitive advantage: institutions and the contingent value of brokerage positions. Am. Behav. Sci. **44**(10), 1579–1601 (2001)
Schultz, T.W.: Investment in entrepreneurial ability. Scand. J. Econ. **82**(4), 437–448 (1982)
Schumpeter, J.A.: The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. Translated by Redvers Opie. Harvard University Press, Cambridge (1934)

Sebora, T.C., Lee, S.M., Sukasame, N.: Critical success factors for e-commerce entrepreneurship: an empirical study of Thailand. Small Bus. Econ. **32**(3), 303–316 (2009)
Shane, S.A.: Why do some societies invent more than others? J. Bus. Ventur. **7**(1), 29–46 (1992)
Shane, S., Venkataraman, S.: The promise of entrepreneurship as a field of research. Acad. Manag. Rev. **25**(1), 217–226 (2000)
Singer, S., Amoros, J.E., Moska, D.: GEM 2014 Global Report. The Global Entrepreneurship Monitor, London (2015)
Skica, T., Bem, A., Zygadlo, K.: The role of local government in the process of entrepreneurship development. Financ. Internet Q. e-Finanse **9**(4), 1–24 (2013)
Smallbone, D., Welter, F., Voytovich, A., Egorov, I.: Government and entrepreneurship in transition economies: the case of small firms in business services in Ukraine. Serv. Ind. J. **30** (5), 655–670 (2010)
Soriano, D.R., Galindo-Martín, M.A.: Government policies to support entrepreneurship. Entrep. Reg. Dev. **24**(9–10), 861–864 (2012)
Tende, S.: Government initiatives toward entrepreneurship development in Nigeria. Glob. J. Bus. Res. **8**(1), 109–120 (2014)
Thomas, A.S., Mueller, S.L.: A case for comparative entrepreneurship: assessing the relevance of culture. J. Int. Bus. Stud. **31**(2), 287–301 (2000)
Tiessen, J.H.: Individualism, collectivism, and entrepreneurship: a framework for international comparative research. J. Bus. Ventur. **12**(5), 367–384 (1997)
Tracy, M.: The state of entrepreneurship in Turkey. Thesis Ph.D. Pace University, USA (2013)
Van der Sluis, J., Van Praag, M., Vijverberg, W.: Education and entrepreneurship selection and performance: a review of the empirical literature. J. Econ. Surv. **22**(5), 795–841 (2008)
Witt, P.: Entrepreneurs' networks and the success of start-ups. Entrepreneurship Reg. Dev. **16**(5), 391–412 (2004)
Xu, Z., Cai, X.: Interval-Valued Intuitionistic Fuzzy Information Aggregation. Springer, Berlin (2012)
Xu, Z.S., Yager, R.: Dynamic intuitionistic fuzzy multiple attribute decision making. Int. J. Approx. Reason. **48**, 246–262 (2008)
Xu, Z., Lin, J., Lin, D.: Networking and innovation in SMEs: evidence from Guangdong Province, China. J. Small Bus. Enterp. Dev. **15**(4), 788–801 (2008)
Zhao, H., Seibert, S.E.: The big five personality dimensions and entrepreneurial status: a meta-analytical review. J. Appl. Psychol. **91**(2), 259–271 (2006)

Hesitant Fuzzy Evaluation of System Requirements in Job Matching Platform Design

Sezi Cevik Onar(✉), Basar Oztaysi, and Cengiz Kahraman

Industrial Engineering, ITU Management Faculty, Istanbul Technical University, Macka, Besiktas, Istanbul, Turkey
{cevikse, oztaysib, kahramanc}@itu.edu.tr

Abstract. System requirements are vital for software development. Defining the appropriate requirements and their importance levels and taking the necessary actions to fulfill the most crucial ones are the keys to a successful software program. However, prioritization of the requirements is a complex problem that involves fuzziness and ambiguities. In this study, we propose a multi-criteria decision-making approach based on HFLTS (Hesitant Fuzzy Linguistic Term Sets) to evaluate the system requirements. The proposed method is applied to G@together project that focuses on developing an electronic job-matching platform for disadvantaged people.

Keywords: Hesitant fuzzy sets · Requirements prioritization · Hesitant Fuzzy Linguistic Term Sets

1 Introduction

Defining the system requirements of a software system involves determining associated documentations, architectural design principles, coding and testing policies. It has a vital role in software system development processes. The end users and stakeholders accept an information system only when the requirements are well captured, analyzed and prioritized [1–3]. Prioritization of requirements enables the software to function as expected [4]. Requirements prioritization focus on identification of essential requirements as perceived by relevant stakeholders [5]. It is important to implement the core requirements of stakeholders with respect to cost, quality, available resources and delivery time [6, 7].

Determining which, among a set of requirements to be implemented first and the order of implementation is known as requirements prioritization. Prioritization of requirements provides many advantages before architecture design or coding. It helps to deal with the challenges associated with software development such as; limited resources, inadequate budget, insufficiently skilled programmers. It also helps planning software releases since not all the elicited requirements can be applied in a single release [8]. Perini et al. [3] propose that prioritizing requirements plays a significant role in a system development process as it enhances software release planning, budget control, and scheduling.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_5

Leffingwell [9] defines the prioritization process as a complex multi-criteria decision-making process. To fulfill the prioritization process, stakeholders will have to compare the set of requirements to determine their relative importance through a weighting or scoring system [10]. These comparisons may become complex as the number of requirements increase [11].

The fuzzy sets theory is a mathematical representation of uncertainty [12]. Fuzzy sets are frequently used in multi-criteria decision-making problems. Extensions of fuzzy sets such as intuitionistic fuzzy sets, Type-2 fuzzy sets, Pythagorean fuzzy sets and hesitant fuzzy sets are developed to represent uncertainty inherent in a system better. In the cases where more than one sources of vagueness exist, regular fuzzy sets may have shortcomings [13].

Hesitant fuzzy sets (HFSs) are developed by Torra [14] and enable solving problems in which a set of values are possible for membership of a single element. Based on HFS, Rodriquez et al. [15] propose using linguistic terms for problem modeling and solution. Hesitant fuzzy terms sets (HFLTS) enable the expert evaluations to be captured by linguistic terms flexibly.

In this paper, software requirements prioritization problem is handled by using HFLTS and a case study from a G@together project, an EU project supported by Urban Europe, is provided. In the proposed methodology, the obtained requirements are grouped, and a hierarchy of requirements are formed. Elements in each level of the hierarchy are pairwise compared using HFLTS.

The rest of the paper is organized as follows. Section 2 gives a literature review of software requirement prioritization studies. Section 3 first introduce hesitant fuzzy sets and then give the steps of the methodology. In Sect. 4, a real case study is provided, and the requirements are prioritized. Finally, Sect. 5 provides the conclusions and future research suggestions.

2 Current Studies on Requirements Prioritization

Prioritization of information systems requirements has been taking attention of academic researchers. In one of the recent studies, Achimugu et al. [4] present an extended literature review by examining 73 studies on the topic. The authors identify 49 different methods used for information systems requirements prioritization. Among these techniques, the most commonly used ones are determined as; Analytic Hierarchy Process (AHP), Quality Function Deployment (QFD), Planning Game (PG), Binary Search Tree (BST), and Cumulative Voting (CV). AHP, proposed by Thomas Saaty [16] is based on pair-wise comparison matrix to calculate the relative importance of each requirement. Quality Function Deployment is used to add the design quality into subsystems and component parts [17]. In PG (Planning Game) approach, the users categorize their requirements into three classes namely, essential, conditional and optional. The categorization process is based on two criteria: business value judged by the clients and technical risk judged by the developers [18]. BST (Binary Search Tree) technique analyzes all the elicited requirements and ranks them in a hierarchical order using a parent-child relationship [4]. In CV (Cumulative Voting) method, all stakeholders are given a fictional $100 and asked to spend it on the requirements.

As all stakeholders accomplish the process, the total numbers of stakeholders to prioritize requirements [9] divides the total expended money for each requirement.

3 Hesitant Fuzzy Multi-criteria System Requirements Evaluation

3.1 Preliminaries

A hesitant fuzzy set (HFS) on X, where X is a fixed set can be defined as follows:

$$E = \{ <x, h_E(x) > | x \in X \}, \tag{1}$$

where $h_E(x)$ denotes membership degrees of the element $x \in X$ to the set E and its values are in [0, 1].

A hesitant fuzzy linguistic term set (HFLTS) H_s is a finite subset of consecutive linguistic terms of a linguistic term set S which can be shown as $S = \{s_0, s_1, \ldots, s_g\}$ [15, 19].

An HFLTS, H_s, is an ordered finite subset of consecutive linguistic terms of a linguistic term set S which can be shown as $S = \{s_0, s_1, \ldots, s_g\}$.

3.2 Steps of the Methodology

In this study, the weights of the system requirements are defined with a hierarchical HFLTS approach [20].

- *Defining the hierarchy of system requirements.*
- *Computing the weights of system requirements.*
- Pairwise evaluations of system requirements with HFLTS given in Table 1 and the context-free grammar “between” and “is”.

Table 1. Linguistic scale for hesitant fuzzy AHP

Linguistic term	Abb.	Triangular fuzzy number
Absolutely High Importance	(AHI)	(7,9,9)
Very High Importance	(VHI)	(5,7,9)
Essentially High Importance	(ESHI)	(3,5,7)
Weakly High Importance	(WHI)	(1,3,5)
Equally High Importance	(EHI)	(1,1,3)
Exactly Low Importance	(EE)	(1,1,1)
Equally Low Importance	(ELI)	(0.33,1,1)
Weakly Low Importance	(WLI)	(0.2,0.33,1)
Essentially Low Importance	(ESLI)	(0.14,0.2,0.33)
Very Low Importance	(VLI)	(0.11,0.14,0.2)
Absolutely Low Importance	(ALI)	(0.11,0.11,0.14)

- Calculating fuzzy envelope for HFLTS by using the OWA operator given in Eqs. (2–8) [19, 21].

$$\boldsymbol{\alpha} = \min\left\{a_L^i, a_M^i, a_M^{i+1}, \ldots . a_M^j, a_R^j\right\} = a_L^i \tag{2}$$

$$\boldsymbol{\delta} = \max\left\{a_L^i, a_M^i, a_M^{i+1}, \ldots . a_M^j, a_R^j\right\} = a_R^j \tag{3}$$

$$\boldsymbol{\beta} = \begin{cases} a_m^i, \text{if } i+1 = j \\ OWA_{w^2}\left(a_m^i, \ldots . a_m^{\frac{i+j}{2}}\right), \text{if } i+j \text{ is even} \\ OWA_{w^2}\left(a_m^i, \ldots . a_m^{\frac{i+j-1}{2}}\right), \text{if } i+j \text{ is odd} \end{cases} \tag{4}$$

$$\boldsymbol{\gamma} = \begin{cases} a_m^{i+1}, \text{if } i+1 = j \\ OWA_{w^1}\left(a_m^j, a_m^{j-1} \ldots . a_m^{\frac{i+j}{2}}\right), \text{if } i+j \text{ is even} \\ OWA_{w^1}\left(a_m^j, a_m^{j-1}, \ldots . a_m^{\frac{i+j+1}{2}}\right), \text{if } i+j \text{ is odd} \end{cases} \tag{5}$$

$$OWA(a_1, a_2, \ldots, a_n) = \sum\nolimits_{j=1}^{n} w_j b_j \tag{6}$$

where b_j is the j^{th} largest of the aggregated arguments $a_1, a_2, \ldots, a_n$, and $W = (w_1, w_2, \ldots, w_n)^T$ is the associated weighting vector satisfying $w_i \in [0, 1]$, i = 1, 2, …, n and $\sum_{i=1}^{n} w_i = 1$.

$$w_1^1 = \alpha_2, w_2^1 = \alpha_2(1-\alpha_2), \ldots \ldots . w_n^1 = \alpha_2(1-\alpha_2)^{n-2} \tag{7}$$

$$w_1^2 = \alpha_1^{n-1}, w_2^2 = (1-\alpha_1)\alpha_1^{n-2}, \ldots \ldots . w_n^2 = 1-\alpha_1, \tag{8}$$

where $\alpha_1 = \frac{g-(j-i)}{g-1}$, $\alpha_2 = \frac{(j-i)-1}{g-1}$ g is the number of terms in the evaluation scale, j is the rank of highest evaluation and i is the rank of lowest evaluation value of the given interval.

- Obtaining aggregated pairwise comparison matrix ($\tilde{C}$).

$$\tilde{C} = \begin{vmatrix} 1 & \tilde{c}_{12} & \cdots & \tilde{c}_{1n} \\ \tilde{c}_{21} & 1 & \cdots & \tilde{c}_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \tilde{c}_{n1} & \tilde{c}_{n2} & \cdots & 1 \end{vmatrix} \tag{9}$$

where $\tilde{c}_{ij} = \left(c_{ij_l}, c_{ij_{m1}}, c_{ij_{m2}}, c_{ij_u}\right)$.

- Computing fuzzy geometric mean for each row $(\tilde{r}_i)$ of the matrix $\tilde{C}$:

$$\tilde{r}_i = (\tilde{c}_{i1} \otimes \tilde{c}_{i2} \ldots \otimes \tilde{c}_{in})^{1/n} \tag{10}$$

- Calculating the fuzzy weight $\left(\tilde{\mathrm{w}}_{\mathrm{i}}^{\mathrm{CR}}\right)$:

$$\tilde{\mathrm{w}}_{\mathrm{i}}^{\mathrm{CR}} = \tilde{r}_i \otimes (\tilde{r}_i \oplus \tilde{r}_2 \ldots \oplus \tilde{r}_n)^{-1} \tag{11}$$

- Calculating the fuzzy global weights of sub-customer requirements:

$$\tilde{\mathrm{w}}_{\mathrm{ij}}^{\mathrm{G}} = \tilde{\mathrm{w}}_{\mathrm{ij}}^{\mathrm{CR}} \times \tilde{\mathrm{w}}_{\mathrm{ij}}^{\mathrm{CR}} \tag{12}$$

where $\tilde{\mathrm{w}}_{\mathrm{ij}}^{\mathrm{G}}$ is the global weight of sub-customer requirement *ij*.
- Defuzzifying $\tilde{\mathrm{w}}_{\mathrm{ij}}^{\mathrm{G}}$ and normalizing the defuzzified values:

$$w_{ij}^{G} = \frac{\alpha + 2\beta + 2\gamma = \delta}{6} \tag{13}$$

$$w_{ij}^{N} = \frac{w_{ij}^{G}}{\sum_i \sum_j w_{ij}^{G}} \tag{14}$$

4 Application

The proposed method is applied to G@together project funded by JPI Urban Europe which an interdisciplinary project is aiming to build an e-platform facilitating the employment of qualified but disadvantaged people. People from lower income background, disabled people, and victims of family violence are considered as disadvantaged groups suffering social exclusion, having limited access to natural resources, education, and economic opportunities.

Although there are various studies in the field of job matching [22–26] the considered problem is different from other from various perspectives. First of all, in classical job matching job seekers and employers are looking for the best available alternative. However, in the given case, the job seekers may not be the best available because by definition they are disadvantaged. Another important distinction about systems is that policy makers are also stakeholders of them.

Elicited system requirements are grouped under five factors: (i) Awareness and Personal Improvement Content (APIC): This part contains the main theme and message of the organization and useful information and links related to the self-improvement of the job seekers. (ii) Tools and utilities for entering the job and job seeker profile data (PD): These are the system features that introducing job seekers and companies regarding their specific characteristics. (iii) Filtering/Matching (FM): This part includes search and advanced search features enabling to filter according to multiple profile criteria such as industry or location. (iv) Evaluation Tests (ET): The system features to

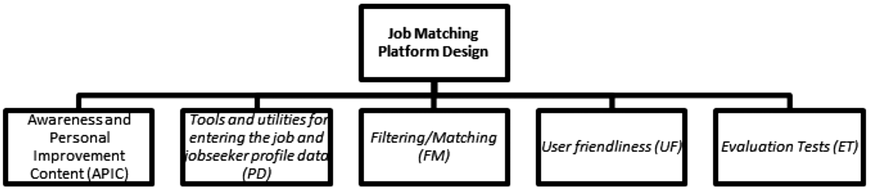


Fig. 1. System requirements hierarchy

evaluate the competencies, personal qualities, and suitability of the job seekers for the job. (v) User friendliness (UF): Includes features such as integration with corporate websites and informing about available employment opportunities through instant messaging. System requirements hierarchy is given in Fig. 1.

Three experts evaluate the system requirements. The compromised verbal evaluations are given in Table 2.

Table 2. Verbal evaluations

	APIC	PD	FM	ET	UF
APIC	EE	Between EHI and ESHI	Between ESLI and WLI	Between EE and EHI	Between EHI and ESHI
PD		EE	Between WLI and ELI	Between WLI and EE	ELI
FM			EE	Between ELI and EHI	Between EHI and WHI
ET				EE	Between ESHI and VSHI
UF					EE

Table 3 is obtained by using aggregation operations.

Table 3. Numerical evaluations

	APIC	PD	FM	ET	UF
APIC	(1,1,1,1)	(1,2.78,3.22,7)	(0.14,0.2,0.33,1)	(1,1,1,3)	(1,2.78,3.22,7)
PD	(0.14,0.31,0.36,1)	(1,1,1,1)	(0.2,0.33,1,1)	(0.2,0.93,1.07,1)	(0.33,1,1,1)
FM	(1,3,5,7)	(1,1,3,5)	(1,1,1,1)	(0.33,1,1,3)	(1,1,3,5)
ET	(0.33,1,1,1)	(1,0.93,1.08,5)	(0.33,1,1,3)	(1,1,1,1)	(3,5,7,9)
UF	(0.14,0.31,0.36,1)	(1,1,1,3)	(0.2,0.33,1,1)	(0.11,0.14,0.2,0.33)	(1,1,1,1)

The proposed method is applied to the importance levels in Table 4 are obtained.

Table 4. Importance levels

	Fuzzy importance	Defuzzified values
APIC	(0.06,0.17,0.27,0.94)	0.3610885
PD	(0.03,0.1,0.17,0.35)	0.1613799
FM	(0.07,0.2,0.45,1.21)	0.4834381
ET	(0.07,0.21,0.32,0.92)	0.382027
UF	(0.03,0.07,0.12,0.35)	0.1419731

Filtering/Matching (FM) is the most important whereas user friendliness (UF) is the least important system requirement.

5 Conclusion

The real world application shows that HFLTS base method provides an efficient and easy to use tool for requirements prioritization. The experts can easily make their evaluations using linguistic terms sets. The proposed method has a limitation since it only focuses only on user requirements. The technique can be improved by integrating developers' point of view showing the costs and technical complexity of the requirements, to better help system development.

For the further studies, the robustness of the paper can be analyzed with a sensitivity analysis.

References

1. Davis, A.M.: The art of requirements triage. Computer **36**(3), 42–49 (2003)
2. Mead, N.R.: Software Engineering Requirements Prioritization Introduction. Institute Web Publication, Carnegie Mellon University, Pittsburgh (2006)
3. Perini, A., et al.: An empirical study to compare the accuracy of AHP and CBRanking techniques for requirements prioritization. In: Fifth International Workshop on Comparative Evaluation in Requirements Engineering, CERE 2007 (2007)
4. Achimugu, P., et al.: A systematic literature review of software requirements prioritization research. Inf. Softw. Technol. **56**(6), 568–585 (2014)
5. Ruhe, G., Eberlein, A., Pfahl, D.: Trade-off analysis for requirements selection. Int. J. Softw. Eng. Knowl. Eng. **13**(4), 345–366 (2003)
6. Finkelstein, A., et al.: A search based approach to fairness analysis in requirement assignments to aid negotiation, mediation and decision making. Requirements Eng. **14**(4), 231–245 (2009)
7. Barney, S., Aurum, A., Wohlin, C.: A product management challenge: creating software product value through requirements selection. J. Syst. Archit. **54**(6), 576–593 (2008)
8. Berander, P., Khan, K.A., Lehtola, L.: Towards a research framework on requirements prioritization. In: Proceedings of Sixth Conference on Software Engineering Research and Practice in Sweden (SERPS 2006) (2006)

9. Leffingwell, D., Widrig, D.: Managing Software Requirements: A Unified Approach. Addison-Wesley Longman Inc., Reading (2000)
10. Kobayashi, A., Maekawa, M.: Need-based requirements change management. In: Proceedings of the Eighth Annual IEEE International Conference and Workshop on Engineering of Computer Based Systems, ECBS 2001 (2001)
11. Kassel, N.W., Malloy, B.A.: An approach to automate requirements elicitation and specification. In: Proceedings of the 7th IASTED International Conference on Software Engineering and Applications, Marina Del Rey, CA, USA (2003)
12. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
13. Mendel, J.M.: Uncertainty, fuzzy logic, and signal processing. Sig. Process. **80**(6), 913–933 (2000)
14. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
15. Rodriguez, R.M., Martinez, L., Herrera, F.: Hesitant fuzzy linguistic term sets for decision making. IEEE Trans. Fuzzy Syst. **20**(1), 109–119 (2012)
16. Saaty, T.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York (1980)
17. Akao, Y.: Development History of Quality Function Deployment. The Customer Driven Approach to Quality Planning and Deployment. Asian Productivity Organization, Tokyo (1994)
18. Karlsson, L., et al.: Pair-wise comparisons versus planning game partitioning—experiments on requirements prioritisation techniques. Empirical Softw. Eng. **12**(1), 3–33 (2007)
19. Rodríguez, R.M., Liu, H., Martínez, L.: A fuzzy representation for the semantics of hesitant fuzzy linguistic term sets. Adv. Intell. Syst. Comput. **277**, 745–757 (2014)
20. Cevik Onar, S., Büyüközkan, G., Öztayşi, B., Kahraman, C.: A new hesitant fuzzy QFD approach: an application to computer workstation selection. Appl. Soft Comput. J. **46**, 1–16 (2016)
21. Filev, D., Yager, R.R.: On the issue of obtaining OWA operator weights. Fuzzy Sets Syst. **94** (2), 157–169 (1998)
22. García-Sánchez, F., et al.: An ontology-based intelligent system for recruitment. Expert Syst. Appl. **31**(2), 248–263 (2006)
23. Lin, H.T.: A job placement intervention using fuzzy approach for two-way choice. Expert Syst. Appl. **36**(2), 2543–2553 (2009). Part 1
24. Faliagka, E., Tsakalidis, A., Tzimas, G.: An integrated e-recruitment system for automated personality mining and applicant ranking. Internet Res. **22**(5), 551–568 (2012)
25. Álvarez de Toledo, P., Núñez, F., Usabiaga, C.: An empirical approach on labour segmentation. Applications with individual duration data. Econ. Model. **36**, 252–267 (2014)
26. Chen, J., et al.: Recommendations for two-way selections using skyline view queries. Knowl. Inf. Syst. **34**(2), 1–28 (2013)

An Interval Valued Hesitant Fuzzy Clustering Approach for Location Clustering and Customer Segmentation

Sultan Ceren Öner(✉) and Başar Öztayşi

Industrial Engineering Department,
Istanbul Technical University, Maçka, İstanbul, Turkey
{csalkin, oztaysib}@itu.edu.tr

Abstract. Because of the irrepressible growth in information technologies and telecommunication infrastructure especially for mobile devices, people are more disposed to search proper products and find out attractive offers with lower prices. In order to reach potential customers, companies deal with offering personalized messages including special promotions and discounts. In this respect, recommender systems have begun to use as one of the essential tools for making appropriate selections considering diversified conditions and personal preferences. On the other hand, users' preferences could not be easily determined or predicted in some cases, as seen in visiting prediction of mobile users. Thus, the use of location based service applications enable the determination of users visiting patterns, except making predictions. In this study, an interval valued hesitant fuzzy clustering approach is adapted based on location similarity and fuzzy c means clustering is applied for user segmentation. After that, matching location groups and user segments is provided the representation of user visiting tendency. By using this approach, advertisers will be able to handle their advertisements considering location similarities and user groups that helps the implementation of personalized advertising recommender systems.

Keywords: Location similarity · Segmentation · Hesitant fuzzy clustering · Interval valued hesitant fuzzy clustering

1 Background

As a result of rapid changes in mobile devices and location tracking technologies, more and more people follow instant messages that present customized offers for specific locations. Location based systems or in another saying, location aware systems rely on geographical information determination of mobile users via global positioning systems (GPS) or Bluetooth (via Beacons) to pinpoint users' real time locations (Lee et al. 2015) .Today, location based systems are widely adapted in real life, especially seen for targeted advertisement. This adaptation can be seen as location based recommender systems which present an alternative channel that enables companies to send special offers based on consumers' preferences and previous location and present opportunities for personalization and increasing revenue (Xu et al. 2009). According to Junglas and Watson (2008), two main steps are necessary to conduct location based systems:

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_6

location determination and modeling to indicate the context of location, and “possible” future locations by geo spatial data processing and prediction algorithms and second, user preference modeling for specifying user properties from previously searching patterns, online profiles, previously visited places etc.

In recent years, location determination and the prediction of future visits have become important and integrated with recommender systems in targeted advertisement management, especially making personalized offers which can enhance companies’ communication channels for customers with limited advertisement budgets. These systems are relied on time and preference based services including monitoring of current position of mobile user and searching appropriate offers or advertisement according to possible visiting places extracted from previously visited places (Wu et al. 2015). However, customers generally avoid to read of follow the instant messages because of the substantial number of messages, irrelevant offers or redundancy of the message context (Shin and Lin 2016). The other reasons of avoiding these push up messages are location privacy which is a significant barrier to the penetration of location based services and accidentally keystrokes (Abbas et al. 2014; Pingley et al. 2012). To overcome this problem, continuous information feedback for incoming conditions might be beneficial. Thus, advertisers are prone to search and apply more efficient approaches that provide both personalized contexts and broad alternatives to their recommender systems for individual and commonly-held advertisements (Dao et al. 2012). Additionally, user interests and needs continuously change and the differentiation of these needs and interests can be problematic to perform efficient recommender systems. Thus, location prediction and customer segmentation exist as the most important topics in location based recommender systems for reflecting user movements as a characteristic of customer preferences or needs (Fan et al. 2015).

In this respect, the main research question of this study is the matching of customer movements with respect to previously visited locations and location characteristics the determination of alternative locations according to the similarities assigned by common characteristics. For this purpose, location clustering can be applied for grouping locations and alternative location detection can be conducted of in terms of the similarities appeared from the location clusters and customer segments. Since the nature of the problem contains insufficient data, vague environment and conflicting expectations from diverse users, the clustering problem varies as a fuzzy clustering problem, in particular, fuzzy partition of clusters (Aliahmadipour et al. 2017). In addition to that, location data could be presented as interval valued fuzzy numbers to provide the hesitancy of users’ movements which can enhance the recommendation of proper alternative locations. Therefore, in this study, location grouping is conducted by using interval valued hesitant fuzzy clustering. Then, the extracted location clusters are evaluated according to fuzzy c means based user segments. In this way, alternative location clusters for each characteristic are acquired.

The remainder of this paper is structured as follows: the second section explains the brief concepts of location based mobile advertising and location based clustering. The third part presents basic concepts of fuzzy clustering, hesitant fuzzy sets and interval valued hesitant fuzzy clustering. The proposed methodology is shown with a numerical application and the last section contains the conclusion and discussion part.

2 Related concepts

2.1 Location Based Mobile Advertising

Because of the widespread usage of mobile technologies such as GPS and Bluetooth 4.0, location based systems can adapt users' personalized interests in a specific location in a specific time (Fan et al. 2015). Thus, location based systems aim to provide content providers to send proper services to customers when they visit a specific location at a specific time (Tussyadiah 2012). The main concepts of these services are tagging, tracking, navigating and mobile commerce. In particular, tagging enables matching relevant information to a specific location. Tracking services provide information on the position of objects. In addition to that, navigating systems assist people to move from one place to a target destination. Mobile commerce systems are adopted as sending proper advertisements to a targeted customer in a specific location to facilitate location and event based service flow (Wu et al 2015). All these services are the results of consumer interests for the developing of targeted location based systems (Lin et al. 2016).

Location based services include both pull and push systems. Pull advertising reveals when customers search information about a specific topic or item. As distinct from pull advertising, advertisers send automatic messages to customers appeared a specific location in push advertising. All these mechanisms could be evaluated as a type of targeted advertising where vendors send customized messages to consumers' mobile devices when they are near a specific location (Unni and Harmon 2007). In this respect, permission based pull advertising orientation; real time positioning and personalization are fundamentals for location based advertising.

Although significant advantages appeared using location based systems, some challenges are also aroused. Two main challenges are related to the privacy issue and irritation of messages: customers do not want to share their location data with advertisers and do not expect to receive push messages due to the irrelevance of the message and disruption of instant messages (Xu et al. 2009). The other challenge for location based advertising is customers' location accuracy and the prediction of future routes that a large volume of both spatial and temporal data should be processed. Another research direction is the decision of the integration of different data sources especially for complicated business problems as seen in mobile marketing. In our case, selection of alternative locations requires collecting different data from diversified sources to provide conduction of the relationship between people, time, location, objects and their relationships between each other. The core business in these applications is to ensure necessary information to users, provide the automatic implementation of the services that satisfy customer expectations and collect information concerning about actual data. (Gavalas et al. 2014).

2.2 Location Based Clustering

The revealing of advancements in mobile technologies and the wide application of location based services has derived the integration of location based services with respect to user similarities. Today, advertisers are prone to deliver advertisements and promotion messages considering location information that is stated from mobile devices (Lin et al. 2016). The demand for location based systems integration with clustering is extracted as a need of understanding the process that customer makes the purchasing decision considering location, time and actual needs (Schilke et al. 2004). The needs can be derived from consumer life styles, demographical information, consumption behavior and the reaction to previously sent messages (Shin and Lin 2016). On the other hand, these factors don't solely reflect the entire purchasing decision. Thus, researchers try to search other indicators that can reflect customer characteristics such as geographical data, digital participation in social media and search history for products for better understanding of the changes in customers purchasing tendency (Gavalas et al. 2014).

Location based clustering is an essential tool for the use of recommender systems. Recommender systems have divided into two main filters: content based filtering and collaborative filtering. Content based filtering considers the degree of similarity between offers and interests. However, collaborative filtering investigates the degree of integration of recommendations and user preferences (Wei et al. 2010). According to Zhang et al. (2007), content based filtering is conducted by grouping similar items for offering new items according to previous preferences and collaborative filtering relies on finding the appropriate option based on users' preferences to decide new users' preferences (Zhang et al. 2007). In this respect, user preference similarities could be evaluated from geo data, visited location category, price level, rating, and number of stars, as seen from Foursquare for location based recommender systems. To conduct the similarity calculation, users should be grouped as segments and the recommendations should be given according to the diversification of these segments. The divergence between users can be accomplished either by categorization or by adapting analytical techniques such as cluster analysis, heuristic methods, regression, neural network, kNN, link analysis, decision tree, association rule mining etc. (Park et al. 2012). By using these methods, gathering user comments and reflections that the essential clue of customer disposition are provided to assess the changes in customer behavior and make realistic offers not to spend huge amount of data that means a vast of time, money and manpower. Lastly, runtime of the segmentation and recommendation techniques have a significant role for the implementation of real time location based clustering (Öztayşi et al. 2016).

3 Motivation

Because of the results of mobile technologies have been widely penetrated and improved to facilitate marketing operations and advertising, mobile-wireless technologies are increasingly applied to send proper messages to customers (Ortega et al. 2013). Consequently, mobile location based applications have been gradually adapted to make individualized recommendations using customers' geo data occurred from GPS or wireless indoor positioning systems (Liu 2007). As a consequence of increasing demand on mobile technologies and improvements of the integration of smartphones with location detection systems, personal positioning techniques based advertising systems have become emitted and these improvements facilitated the tracking customers' physical location and purchasing behavior in shopping fields. For this reason, clustering algorithms are widely conducted on the evaluation of the similarities stated from individuals' positions in shopping mall and provide the assignment of the customer groups. In other words, location-aware systems can be adapted to advertising activities by making location sensitive suggestions or promotions in accordance with customer's previous visits for focusing on the customers' future buying behavior.

The literature provides a wide range of algorithms for clustering, these algorithms are categorized according to the formation of clusters by using fuzzy data or crisp data and (Oztaysi and Onar 2014). Crisp clustering is applied for classical sets theory and includes the process of converting input data to mutually exclusive subsets. Based on this fact, crisp clustering algorithms assign input data to one certain cluster. On the other hand, fuzzy clustering algorithms assign an element to several clusters simultaneously with various degrees of membership that ranges between 0 and 1 and naturally, the belonging of the clusters are presented as degrees of each input data to allow the appearance of diversified distribution of assigned clusters (Oztaysi and Isik 2014). This property is useful for grouping customers to provide them wide range of recommendations.

Academicians and practitioners generally use clustering algorithms for dividing a dataset into different clusters. To reflect uncertainty for forming clusters, clustering tasks are handled as two points of view; i) considering uncertain data, ii) considering crisp data with uncertain clusters, in particular, fuzzy partition. (Aliahmadipour et al. 2017) Therefore, one of the most adapted clustering algorithm for fuzzy clustering is fuzzy c means clustering as agglomerative approach in which initial clusters should be determined in advance (Chen et al. 2014).

In this study, location grouping according to Foursquare ratings of each shopping mall, number of votes in Foursquare, monthly total number of visits, number of stores in shopping mall, transportation level (1 presents worst and 3 presents best), real estate index from Hurriyet Emlak and percentage of rent price variations in a year is achieved initially by using interval valued hesitant fuzzy c means clustering. Then, users are grouped according to visiting time zone (morning (1), afternoon (2), late afternoon (3) and night (4)), location, type of mobile application and location visit day using fuzzy c means clustering. Finally, user clusters are matched with location groups with some certain criteria such as number of visits per month and number of visited different places. This approach will ensure gathering location similarity and constitutes the matching of location groups and user groups and enhance various location alternatives that a specific user could visit. Thus, advertisers will be able to submit location related promotions and advertisements to relevant customer groups as a starting point of recommender systems.

4 Methodology

4.1 Preliminaries

Hesitant fuzzy sets are the extension of fuzzy sets that presents hesitancy degree when there is an uncertainty in terms of the determination of membership function. Some of the related definitions are given in the following:

Definition 1. A hesitant fuzzy set (*HFS*) on a reference set X is a function of h that returns to a subset of values in [0, 1] and h could be represented as follows:

h: $X \rightarrow \{[0,1]\}$

In this respect, a *HFS* can be expressed as the union of the membership functions.

Definition 2. Let M is a set of n number of membership function which could be represented as $M = \{\mu_1, \mu_2.., \mu_n\}$ and HFS with M could be defined as h_M: $M \rightarrow \{[0, 1]\}$ and $h_M(x) = \{\mu_1(x) \bigcup \mu_2(x) \bigcup .. \bigcup \mu_n(x)\}$.

Definition 3. Let S a linguistic term set as $\{s_0, s_1, s_2,.., s_f\}$ which has an order of terms as S: $s_i \leq s_j$ where $i \leq j$ and has a maximization and minimization operator as max $(s_i, s_j) = s_i$ and min $(s_i, s_j) = s_j$ where $i \geq j$. A hesitant fuzzy linguistic term set is represented as H_s which has an ordered finite subset of sequential linguistic terms using upper (H_{S^+}) and lower bounds (H_{S^-}) and could be also defined as $H_{S^+} = \max(s_i) = s_j$ and where $s_i \leq s_j$ and $H_{S^-} = \min(s_i) = s_j$ where $s_j \leq s_i$ for $s_i \in H_s; \forall i$

Definition 4. The HFLTS could be defined as the composition of upper and lower bounds which could be presented as H_S: $[H_{S^-}, H_{S^+}]$ where $H_{S^-} \leq H_{S^+}$.

Definition 5. Wang et al. (2014). Linguistic term s_i is involved in a linguistic term set S, if $\psi_i \in R^+$. Linguistic scale function φ is mapping as: $S \rightarrow R^+$ such that φ: $s_i \rightarrow \psi_i (i = 0, 1, \ldots, 2t)$, where $0 \leq \psi_1 < \psi_2 < \ldots < \psi_{2t} \leq 1$. In this respect, linguistic scale function used in this study is given as follows: $\varphi(s_i) = \begin{cases} \frac{t^\alpha - (t-i)^\alpha}{2t^\alpha} & i = 0, 1, \ldots, t \\ \frac{t^\beta + (i-t)^\beta}{2t^\beta} & i = t+1, t+2, \ldots, 2t \end{cases}$, where $\alpha, \beta \in (0, 1]$.

Definition 6. *Interval valued hesitant fuzzy sets.* (Wang et al. 2014). Let $X = \{x_1, x_2, \ldots, x_n\}$ be a reference set and $s_{\theta(X)} \in S$. An interval valued hesitant fuzzy linguistic set is defined as $H = \{x, s_{\theta(X)}, \Gamma_H(x), x \in X\}$

Zhang et al. (2016) defined some operations on the IVHFSs. Let $\alpha = \langle s_{\theta(\alpha)}, \Gamma_\alpha \rangle$ and $\beta = \langle s_{\theta(\beta)}, \Gamma_\beta \rangle$ be two interval valued hesitant fuzzy linguistic number (IVHFN).

(1) $\text{neg}(\alpha) = \bar{\varphi}^{-1}\left(\bar{\varphi}(s_{2t}) - \bar{\varphi}\left(s_{\theta(X)}\right)\right), \bigcup_{r_1 = [r_1^-, r_1^+] \in \Gamma_\alpha} \left\{\left[1 - r_1^+, 1 - r_1^-\right]\right\}$

(2) $\alpha \oplus \beta = \bar{\varphi}^{-1}\left(\bar{\varphi}\left(s_{\theta(\alpha)}\right) + \bar{\varphi}\left(s_{\theta(\beta)}\right) - \bar{\varphi}\left(s_{\theta(\alpha)}\right).\bar{\varphi}\left(s_{\theta(\beta)}\right)\right)$,
$\bigcup_{r_1 = [r_1^-, r_1^+] \in \Gamma_\alpha, r_2 = [r_2^-, r_2^+] \in \Gamma_\beta} \left\{\left[r_1^- + r_2^- - r_1^-.r_2^-, r_1^+ + r_2^+ - r_1^+.r_2^+\right]\right\}$

Definition 7. (Wang et al. 2014). Let $\alpha = s_{\theta(\alpha)}, \Gamma_\alpha = s_{\theta(\alpha)}, \bigcup_{r = [r^-, r^+] \in \Gamma_\alpha} \{[r^-, r^+]\}$ be an interval valued hesitant fuzzy linguistic number. Score function S(α) is defined for an interval valued hesitant fuzzy linguistic number (IVHFLN) as follows:

$$S(\alpha) = \varphi^{-}\left(s_{\theta(\alpha)}\right) \times \frac{\sum_{r=[r^-,r^+]\in\Gamma_\alpha}(r^- + r^+)}{2.\#\Gamma_\alpha}$$

where # Γ_α is the number of the interval numbers in Γ_α. For two *IVHFNs* α and β if $S(\alpha) > S(\beta)$, then $\alpha > \beta$; if $S(\alpha) = S(\beta)$, then $\alpha = \beta$.

Definition 8. Let $\alpha = s_{\theta(\alpha)}, \Gamma_\alpha$ and $\beta = s_{\theta(\beta)}, \Gamma_\beta$ are two *IVHFLN*s. Distance between two IVHFNs is defined in the following:

$$d(\alpha,\beta) = \sqrt{\frac{1}{2\times\#\Gamma_{\tilde{h}}} \times \sum_{r_1=[r_1^-,r_1^+]\in\Gamma_\alpha, r_2=[r_2^-,r_2^+]\in\Gamma_\beta} \left(\left(\bar{\varphi}\left(s_{\theta(\alpha)}\right)r_1^- - \bar{\varphi}\left(s_{\theta(\beta)}\right)r_2^-\right)^2 + \left(\bar{\varphi}\left(s_{\theta(\alpha)}\right)r_1^+ - \bar{\varphi}\left(s_{\theta(\beta)}\right)r_2^+\right)^2\right)}$$

where $\#\Gamma_{\tilde{h}} = max\left(\#\Gamma_\alpha, \#\Gamma_\beta\right)$ that # Γ_α and # Γ_β are the number of the interval numbers in Γ_α and $\Gamma_\beta$respectively.

4.2 Fuzzy c means clustering

Clustering is defined as the process of dividing a set of observations into subgroups which are entitled clusters. Various techniques can be used for clustering and the aim of these techniques is to organize input data so as to make similar objects in a cluster, and dissimilar objects in different clusters (Han et al. 2001). Clustering techniques are based on similarity term which is calculated by mathematical distance (Babuska 2009). Clustering is an unsupervised learning technique, in other words there are no predefined groups and a single correct solution (Theodoridis and Koutroumbas 2008).

In clustering problem, the input is a set of observations or objects each of which consists of different attributes. The result of cluster analysis produces the clusters and membership of each data point to these clusters. These outputs are represented by the partition matrix. Ruspini (1970) defines the conditions for a fuzzy partition matrix as follows:

$$\mu_{ik} \in [0,1], 1 \le i \le c, 1 \le k \le N \tag{1a}$$

$$\sum\nolimits_{i=1}^{c} \mu_{ik} = 1, 1 \le k \le N, \tag{1b}$$

$$0 < \sum\nolimits_{k=1}^{N} \mu_{ik} < N, 1 \le i \le c \tag{1c}$$

Equation (1b) constrains the sum of each column to 1, and thus the total membership of each equals one.

One of the most popular fuzzy clustering methods is fuzzy c-means (FCM) which is based on minimization of the following objective function:

$$J(Z,U,V) = \sum\nolimits_{i=1}^{c} \sum\nolimits_{j=1}^{N} \left(\mu_{ij}\right)^m z_j - v_i^2 \tag{2}$$

where Z is the data set to be partitioned, U is the fuzzy partition matrix, V is the vector of cluster centers. In the formula, N represents the number of observations, c is the number of clusters and μ shows the membership value, m is the parameter called fuzzifier which determines the fuzziness of the resulting clusters. The fuzzifier parameter can get values 1 and more. When the fuzzifier parameter equals to one, then the clusters are formed in crisp format. In the formula, z_k - v_i shows the distance between observation k and the center of cluster i.

The minimization of the mention objective function represents a nonlinear optimization problem that can be solved by using a variety of methods such as iterative minimization, simulated annealing or genetic algorithms. Babuska (2009) presents the steps of fuzzy c-means (FCM) algorithm as follows:

1. *Initialize* $U=[u_{ij}]$ *matrix,* $U^{(0)}$
2. *At k-step: calculate the centers vectors* $V^{(k)}=[v_i]$ *with* $U^{(k)}$

$$v_i = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} \cdot z_j}{\sum_{i=1}^{N} \mu_{ij}^{m}}$$

3. *Update* $U^{(k)}$, $U^{(k+1)}$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\left\| z_j - v_i \right\|}{\left\| z_j - v_k \right\|} \right)^{\frac{2}{m-1}}}$$

If $\left\| U^{(k+1)} - U^{(k)} \right\| < \delta$ *then STOP; otherwise return to step 2.*

4.3 Interval Valued Hesitant Fuzzy c means clustering

In some real life cases, data could obtain hesitancy or crisp data clustering could be presented including hesitation degree. This hesitation may be applied as considering uncertain data and hesitant fuzzy data or considering crisp data set with respect to uncertain clusters (Aliahmadipour et al. 2017). Thus, fuzzy c means clustering algorithm should be adapted to interval valued hesitant fuzzy sets. Before implementation of clustering, location value evaluation and comparison is conducted with three expert. The procedure is described in the following:

1. Initialize U = [u_{ij}] matrix and transform the data to pairwise interval valued hesitant fuzzy expression for each decision maker.
2. Aggregate pairwise interval valued hesitant fuzzy expression using IVHFN aggregation operator given in the following:

Let $\alpha_j = G_j(x) = s_{\theta(\alpha_j)}, \Gamma_{\alpha_j}$ (j = 1, 2, …, n) be a collection of IVHFLNs, then the IVHFLPWA operator can be defined as follows:

$$IVHFLPWA(\alpha_1, \alpha_2, \ldots, \alpha_n) = \frac{T_1}{\sum_{i=1}^{n} T_i}\alpha_1 \oplus \frac{T_2}{\sum_{i=1}^{n} T_i}\alpha_2 \oplus \ldots \oplus \frac{T_n}{\sum_{i=1}^{n} T_i}\alpha_n$$

$$= f^{-1}\left(\sum_{j=1}^{n} \frac{T_j}{\sum_{i=1}^{n} T_i} f\left(s_{\theta(\alpha_j)}\right)\right), \bigcup_{r_1=[r_1^-, r_1^+]\in\Gamma_{\alpha_1},\ldots,r_n=[r_n^-, r_n^+]\in\Gamma_{\alpha_n}} \left\{\left[\frac{\sum_{j=1}^{n} f\left(s_{\theta(\alpha_j)}\right) r_j^- T_j}{\sum_{j=1}^{n} f\left(s_{\theta(\alpha_j)}\right) T_j}, \frac{\sum_{j=1}^{n} f\left(s_{\theta(\alpha_j)}\right) r_j^+ T_j}{\sum_{j=1}^{n} f\left(s_{\theta(\alpha_j)}\right) T_j}\right]\right\}$$

where$T_1 = 1; T_j = \prod_{k=1}^{j-1} S(\alpha_k)$ for (j = 1, 2,…, n) and $S(\alpha_k)$ is a score function of α_k calculated from Definition 7.

3. Determine cluster number according to Xie Beni index.
4. At k-step: calculate the centers vectors $V^{(k)}=[v_i]$ with $U^{(k)}$

$$v_i = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} . z_j}{\sum_{i=1}^{N} \mu_{ij}^{m}}$$

5. Calculate distance between cluster centers and points using Definition 8.
6. Same as the Step 3 in fuzzy c means clustering.

5 Application

As mentioned before, location similarity and grouping is necessary as the initial step for location prediction of moving customers. Thus, location grouping is adapted using Foursquare ratings of each shopping mall, number of votes in Foursquare, monthly total number of visits, number of stores in shopping mall, transportation level as 1 presents worst and 3 presents best, real estate index from Hurriyet Emlak and percentage of rent price variations in a year. After that, pairwise comparison matrix of each shopping mall is gathered in the form of interval valued hesitant fuzzy numbers. Due to the page restrictions, a sample of crisp data and hesitant information computed from linguistic scale function are given in Tables 1 and 2. Hesitant fuzzy information matrix of each locations are determined from comparison of two locations with each other and interval valued hesitant fuzzy numbers are gathered from this comparison using linguistic scale function description given in Definition 5. Note that the lower side of pairwise comparison matrix is calculated from Definition 6 using neg (α).

After that, interval valued hesitant fuzzy c means clustering is adapted to this pairwise comparison matrix. In order to determine the number clusters, Xie-Beni index is used. The algorithm is run with different values for c parameter as presented in Table 3. The lower values of Xie-Beni index refer to better clustering results, thus for

Table 1. A sample of crisp data of shopping malls (7 from 31 locations)

Location ID	Location name	Foursquare rating	Votes in foursquare	Total number of visits (month)	Variety of stores	Transportation level	Real estate index	Variety of rent (%)
111	City's Nişantaşı	8.8	17,447	875,342	90	3	36	−0.14
60	212 AVM	6.6	8900	543,876	115	1	16	−0.22
62	Akasya AVM	9	29,469	1,134,129	237	3	35	0.23
66	Astoria	6.6	2813	798,645	61	2	31	−0.4
68	Buyaka	8.5	17,556	654,390	106	1	14	−0.04
69	Canpark	6.6	7206	876,522	180	2	14	−0.04
71	Capitol	8	14,827	1,245,176	137	3	18	0.04

Table 2. A sample of aggregated pairwise comparison of hesitant fuzzy information of each location

	111	60	62	66	68
111	[1, 1]	{[0.6,0.8], [0.7,0.8]}	{[0.3, 0.6], [0.7, 0.9]}	{[0.3, 0.4], [0.5, 0.7]}	{[0.3, 0.4], [0.4, 0.5], [0.5, 0.6]}
60		[1,1]	{[0.3, 0.4], [0.5, 0.6]}	{[0.3, 0.5], [0.6, 0.8]}	{[0.2, 0.3], [0.4, 0.5]}
62			[1,1]	{[0.7, 0.8], [0.8, 0.9]}	{[0.5, 0.7], [0.8,0.9]}
66				[1,1]	{[0.4, 0.6]}
68					[1,1]

Table 3. Xie-Beni index values for different values of c parameter

C	2	3	4	5
Xie-Beni	1.52×10^{-06}	10.28×10^{-07}	3.45×10^{-05}	6.2×10^{-05}

this study c value is selected as three, which means in the study three clusters are formed. Second, cluster centers are randomly determined and IVHFLS based distance of each location to each cluster center is computed from using Definition 8.

In order to understand the formed groups, the best way is to analyze the centroid table. Centroid table shows the typical characteristic of each cluster by calculating the cluster centers. Table 4 shows the resulting centroid table for the study when c is equal to three. A sample centroid table for locations 111, 60, 62, 66 and 68 is presented in Table 4.

Finally, clusters and locations in each cluster are represented in Figure 1. From this graph, one could conclude that Cluster 1 and Cluster 3 have intersections that constitute alternative locations of each other. Cluster 2 and Cluster 1 have less intersection and Cluster 2 and Cluster 3 don't have any intersection. According to the characteristics of

Table 4. Centroid table of fuzzy c-means algorithm (c = 3)

Locations	Clstr1	Clstr2	Clstr3
111	2.69	0.14	1.7
60	0.23	0.01	0.99
62	3.67	0.7	6.39
66	0.35	0.01	0.83
68	1.67	0.14	1.21

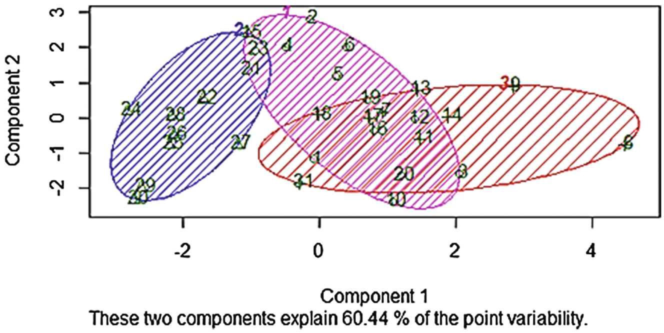


Fig. 1. Cluster plot for locations (shopping malls)

the locations, Cluster 1 is entitled as "Crowded shopping malls" due to their transportation availability and increasing level of sub-province development in their own right. Cluster 2 is named "Middle income focused shopping malls" as due to the variety of stores for each income level. Cluster 3 is identified as "Upstate shopping malls" because of their prestigious popularity and outstanding brand.

In the second phase, users are grouped according to time zone (morning (1), afternoon (2), late afternoon (3) and night (4)), location, type of mobile application and location visit day by implementing fuzzy c means clustering. Again, Xie-Beni index values for different values of c parameter are determined and four clusters are gathered to perform whole dataset of users. Second, cluster centers are randomly assigned and distance between user visits vector to cluster centers are calculated. Finally, customers are grouped as seen from Figure 2.

The most important result from this study is to identify the characteristics of the user segments. Based on the cluster plot given in Figure 2, the segments are entitled as "voyager visitors". The customers in cluster 1 are using airports and upstate shopping malls dramatically higher than the existing customers and generally visit in Fridays and Saturdays. In the second cluster, customers could be named as "weekday visitors" that generally involves workers that visits shopping malls in the afternoon breaks. Third cluster obtains "popular shopping mall visitors" that customers spend their leisure time in popular and overcrowded shopping malls. Final cluster is named as "vacation" focused hedonics" that generally visit prestigious shopping malls in the weekends.

The final step for location based clustering is the matching location segments and user segments. To perform the matching of visited location and user segments, location

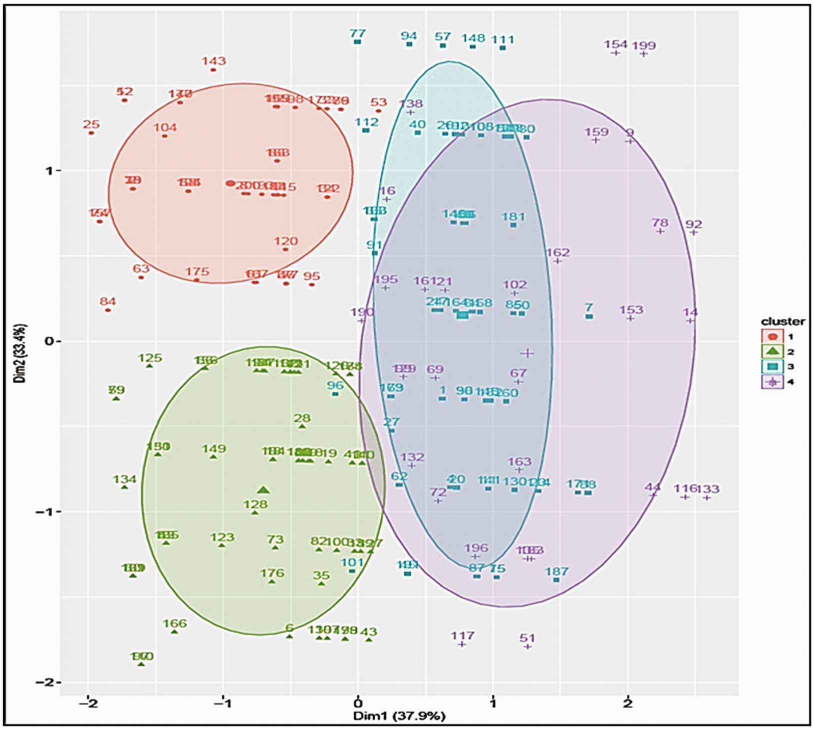


Fig. 2. Cluster plot for customer segments

Table 5. User segment-location group matching

	(1) Crowded shopping malls	(2) Middle income focused shopping malls	(3) Upstate shopping malls
Voyager visitors			x
Weekday visitors	x	x	
Popular shopping mall visitors			x
"Vacation" focused hedonics	x		

clusters and user clusters database are merged that rows present user segments and columns indicate location segmentation cluster considering some certain criteria such as number of visits per month and number of visited different places as presented in the following manner.

According to Table 5, "Vacation" focused hedonics" mainly visit crowded shopping malls to discover new shopping trends and make use of their spare time by activities. Similarly, "Weekday visitors" has the same reason but they generally use shopping malls in the afternoon break for eating lunch. Besides that, "Weekday visitors" also visits middle income focused shopping malls after hours to follow promotions. The motivation of this preference may be they do not want to allocate their spare time for going shopping. On the other hand, "Voyager visitors" and "Popular shopping mall visitors" generally prefer to go shopping malls to spend their time in their weekends.

6 Conclusion

Because of the results that mobile technologies have been widely penetrated and improved to facilitate marketing operations, mobile-wireless technologies are increasingly applied to send proper messages to customers. Before that, customer purchasing tendency should be properly analyzed considering diversified characteristics such as previous visits, location data etc. In this respect, location clustering and user segmentation can be applied for grouping locations before the determination of alternative locations.

To detect users' location, various services can be maintained such as an advertisement or navigation to a specific location. In this study, a novel use of this data is presented and initial result from a real world case study is conducted. To this end, data from a beacon network is collected, preprocessed and clustered for user segmentation. On the other side, location segmentation is implemented using interval valued hesitant fuzzy clustering.

Results indicate that such a location data from various locations has the potential to show customers' life style and interests. In traditional marketing segmentation, segmentation is generally based on demographics or customer value. However, the proposed segmentation approach is more accurate since it is based on real location data. As a result, market segmentation based on customer location data propose a high potential for segmentation and get insight about each individual customer before implementing personalized advertising recommender systems.

References

Aliahmadipour, L., Torra, V., Eslami, E.: On hesitant fuzzy clustering and clustering of hesitant fuzzy data, fuzzy sets, rough sets, multisets and clustering. In: Madani, K. (ed.) Computational Intelligence, vol. 671, pp. 157–168. Springer International Publishing, New York (2017)

Tussyadiah, I.P.: A concept of location-based social network marketing. J. Travel Tour. Mark. **29** (3), 205–220 (2012)

Xu, H., Oh, L.-B., Teo, H.-H.: Perceived effectiveness of text vs. multimedia location-based advertising messaging. Int. J. Mobile Commun. **7**(2), 154–177 (2009)

Lee, S., Kim, K.J., Sundar, S.S.: Customization in location-based advertising: effects of tailoring source, locational congruity, and product involvement on ad attitudes. Comput. Hum. Behav. **51**, 336–343 (2015)

Shin, W., Lin, T.T.C.: Who avoids location-based advertising and why? Investigating the relationship between user perceptions and advertising avoidance. Comput. Hum. Behav. **63** (2016), 444–452 (2016)

Junglas, I.A., Watson, R.T.: Location-based services. Commun. ACM **51**(3), 65–69 (2008)

Wu, C.H., Kao, S.C., Wu, C.C., Huang, S.: Location-aware service applied to mobile short message advertising: design, development, and evaluation. Inf. Process. Manag. **51**(5), 625–642 (2015)

Dao, T.H., Jeong, S.R., Ahn, H.: A novel recommendation model of location-based advertising: context-aware collaborative filtering using GA approach. Expert Syst. Appl. **39**(3), 3731–3739 (2012)

Abbas, R., Michael, K., Michael, M.G.: The regulatory considerations and ethical dilemmas of location-based services (LBS): a literature review. Inf. Technol. People **27**(2014), 2–20 (2014)

Pingley, A., Yu, W., Zhang, N., Fu, X., Zhao, W.: A context-aware scheme for privacy-preserving location-based services. Comput. Netw. **56**(2012), 2551–2568 (2012)

Fan, S., Lau, Raymond Y.K., Leon Zhao, J.: Demystifying big data analytics for business intelligence through the lens of marketing mix. Big Data Res. **2**(1), 28–32 (2015)

Gavalas, D., Konstantopoulos, C., Mastakas, K., Pantziou, G.: Mobile recommender systems in tourism. J. Netw. Comput. Appl. **39**, 319–333 (2014)

Lin, T.T.C., Paragas, F., Goh, D., Bautista, J.R.: Developing location-based mobile advertising in Singapore: a socio-technical perspective. Technol. Forecast. Soc. Chang. **103**, 334–349 (2016)

Schilke, S.W., Bleimann, U., Furnell, S.M., Phippen, A.D.: Multi-dimensional-personalization for location and interest-based recommendation. Internet Res. **14**(5), 379–385 (2004)

Unni, R., Harmon, R.: Perceived effectiveness of push vs. pull mobile location-based advertising. J. Interact. Advert. **7**(2), 28–40 (2007)

Wei, R., Xiaoming, H., Pan, J.: Examining user behavioral response to SMS ads: Implications for the evolution of the mobile phone as a bona-fide medium. Telemat. Inform. **27**(1), 32–41 (2010)

Zhang, X., Edwards, J., Harding, J.: Personalized online sales using web usage data mining. Comput. Ind. **58**, 772–782 (2007)

Park, D.H., Kim, H.K., Choi, Y., Kim, J.K.: A literature review and classification of recommender systems research. Expert Syst. Appl. **39**, 10059–10072 (2012)

Öztayşi, B., Gokdere, U., Simsek, E., Oner, C.: A novel approach to segmentation using customer locations data and intelligent techniques handbook of research on intelligent techniques and modeling applications. In: Kumar, A., Dash, M.K., Trivedi, S.K. (eds.) Marketing Analytics, pp. 21–39. IGI Global, Hershey (2016)

Ortega, F., Bobadilla, J., Hernando, A., Gutiérrez, A.: Incorporating group recommendations to recommender systems: alternatives and performance. Inf. Process. Manag. **49**, 895–901 (2013)

Liu, H.: Survey of wireless indoor positioning techniques and systems. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **37**(6), 1067–1080 (2007)

Oztaysi, B., Isik, M.: Supplier evaluation using fuzzy clustering. In: Kahraman, C., Oztaysi, B. (eds.) Supply Chain Management Under Fuzziness: Recent Developments and Techniques, pp. 61–80. Springer, New York (2014)

Oztaysi, B., Onar, S.C.: User segmentation based on twitter data using fuzzy clustering: the case of sports news. In: Bhatnagar, V. (ed.) Data Mining in Dynamic Social Networks and Fuzzy Systems, pp. 316–333. IGI Global, Hershey (2014)

Ruspini, E.H.: Numerical methods for fuzzy clustering. Inf. Sci. **2**, 319–350 (1970)

Babuska, R.: Fuzzy and neural control disc course lecture notes (2009). http://www.dcsc.tudelft.nl/~disc_fnc/transp/fncontrol.pdf. Accessed 5 Oct 2012

Han, J., Kamber, M.: Data Mining Concepts and Techniques, pp. 5–33. Morgan Kauffman Publishers, Burlington (2001)

Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 4th edn. Elsevier, Amsterdam (2008)
Chen, N., Xu, Z.S., Xia, M.M.: Hierarchical hesitant fuzzy K-means clustering algorithm. Appl. Math.-J. Chin. Univ. **29**, 1–17 (2014)
Zhang, W., Ju, Y., Liu, X.: Multiple criteria decision analysis based on Shapley fuzzy measures and interval-valued hesitant fuzzy linguistic numbers. Comput. Ind. Eng. Arch. **105**(C), 28–38 (2016)
Wang, J.Q., Wu, J.T., Wang, J., Zhang, H.Y., Chen, X.H.: Interval-valued hesitant fuzzy linguistic sets and their applications in multi-criteria decision making problems. Inf. Sci. **288**, 55–72 (2014)

Aggregation of Risk Level Assessments Based on Fuzzy Equivalence Relation

Pavels Orlovs[1(✉)] and Svetlana Asmuss[1,2]

[1] Department of Mathematics, University of Latvia, Zellu str. 25, Riga, Latvia
pavels.orlovs@gmail.com, svetlana.asmuss@lu.lv
[2] Institute of Mathematics and Computer Science, Raina blvd. 29, Riga, Latvia

Abstract. The paper deals with the problem of aggregation of risk level assessments. We describe the technique of a risk level evaluation taking into account values of the risk level obtained for objects which are in some sense equivalent. For this purpose we propose to use the construction of a general aggregation operator based on the corresponding fuzzy equivalence relation. Numerical example of the investment risk level aggregation using an equivalence relation obtained on the basis of different macroeconomic factors for countries of one region is considered.

Keywords: Aggregation operator · General aggregation operator · Fuzzy equivalence relation · Risk level assessment

1 Introduction

Our paper deals with the special construction of a general aggregation operator, which is based on a fuzzy equivalence relation. The need for such operator may appear dealing with different problems. For example, it may appear in decision making if fuzzy sets represent evaluation of some objects provided by several experts in the case when at the same time we have a fuzzy equivalence relation between these objects. In order to obtain the evaluation of some object it is important to take into account how the experts evaluated equivalent objects. So if we want to aggregate several experts' evaluations for this object taking into account the equivalence relation, it could be performed by the proposed operators.

Let us describe a particular real-world example where one could use a general aggregation operator based on a fuzzy equivalence relation. For example, we could consider an investment firm or a bank, which investigates investment opportunities in different countries. One of the key components of risk management in such institutions is a country risk evaluation. Management of the institution approves maximal limits for risk exposures to be taken in different countries. The evaluation of these risk limits is usually performed by risk analysts (experts) taking into account economical, financial, political and social background of the particular countries. There are general guidelines for aggregation of similar risk

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_7

factors [1,9]. An alternative approach for aggregating risk assessments using fuzzy methods is described in [5]. But none of the mentioned models gives an opportunity to take into account the existing equivalence between objects.

It is usually important to know, which limit was assigned previously to some similar in many aspects country (or equivalent with some degree). For example, considering Asian region, we assume that Japan and North Korea are equivalent with very low degree (close to 0), while Japan and South Korea could be considered equivalent with high degree by many factors. Therefore, to obtain the evaluation of some country for management approval, a risk manager could obtain the aggregated result, taking into account the evaluation for equivalent with high degree countries by using aggregation operators, which are defined by using the corresponding equivalence relation. For this purpose we propose to use the construction of a general aggregation operator introduced and developed in our previous papers [2–4] (i.e., the upper general aggregation operator based on a fuzzy equivalence relation).

The paper is organized as follows. Section 2 is devoted to the construction of an upper general aggregation operator based on a fuzzy equivalence relation; here we recall also the definition of a fuzzy equivalence relation and the definitions of ordinary and general aggregation operators. In Sect. 3 the model of application of upper general aggregation for risk level assessments is considered. In Sect. 4 the applied construction of a fuzzy equivalence relation based on a metric is described. Finally, Sect. 5 presents an example of aggregation of experts' evaluations of the countries risk level provided by the proposed technique and its analysis.

2 Upper General Aggregation Operator Based on a Fuzzy Equivalence Relation

Aggregation is the process of combining several numerical values into a single representative value. Mathematically aggregation operator is a function that maps multiple inputs from a set into a single output from this set. In the classical case [6,7,10] aggregation operators are defined on interval $[0,1]$.

Definition 1. *A mapping $A : \bigcup_n [0,1]^n \to [0,1]$ is called an aggregation operator if the following conditions hold:*

(A1) $A(0,\dots,0) = 0$;
(A2) $A(1,\dots,1) = 1$;
(A3) for all $n \in \mathbb{N}$ and for all $x_1,\dots,x_n, y_1,\dots,y_n \in [0,1]$:

$$x_i \le y_i, i = 1,\dots,n \Longrightarrow A(x_1,\dots,x_n) \le A(y_1,\dots,y_n).$$

The notion of general aggregation operator $\tilde{A}$ acting on $[0,1]^X$, where $[0,1]^X$ is the set of all fuzzy subsets of a set X, was introduced in 2003 by A. Takači [8]. We denote a partial order on $[0,1]^X$ by $\preceq$. In this paper we consider the case:

$$\mu \preceq \eta \text{ if and only if } \mu(x) \le \eta(x) \text{ for all } x \in X$$

(here $\mu, \eta \in [0,1]^X$). The least and the greatest elements of this order are denoted by $\tilde{0}$ and $\tilde{1}$, which are indicators of $\varnothing$ and X respectively, i.e., $\tilde{0}(x) = 0$ and $\tilde{1}(x) = 1$ for all $x \in X$.

Definition 2. *A mapping* $\tilde{A}\colon \bigcup_n([0,1]^X)^n \to [0,1]^X$ *is called a general aggregation operator if and only if the following conditions hold:*
($\tilde{A}1$) $\tilde{A}(\tilde{0},\ldots,\tilde{0}) = \tilde{0}$;
($\tilde{A}2$) $\tilde{A}(\tilde{1},\ldots,\tilde{1}) = \tilde{1}$;
($\tilde{A}3$) *for all* $n \in \mathbb{N}$ *and for all* $\mu_1,...,\mu_n,\eta_1,...,\eta_n \in [0,1]^X$:

$$\mu_1 \preceq \eta_1,\ldots,\mu_n \preceq \eta_n \Longrightarrow \tilde{A}(\mu_1,\ldots,\mu_n) \preceq \tilde{A}(\eta_1,\ldots,\eta_n).$$

Two widely used approaches to construct a general aggregation operator $\tilde{A}$ based on an ordinary aggregation operator A are the pointwise extension of A and the T-extension [8] of A, whose idea comes from the classical extension principle and uses a t-norm T. Our approach is to construct a general aggregation operator by using a fuzzy equivalence relation. Let us recall the definition of a fuzzy equivalence relation.

Definition 3. *Let* T *be a t-norm and* E *be a fuzzy relation on a set* X*, i.e.,* E *is a fuzzy subset of* $X \times X$*. A fuzzy relation* E *is called a* T*-fuzzy equivalence relation if and only if for all* $x, y, z \in X$ *it holds*
($E1$) $E(x,x) = 1$ *(reflexivity);*
($E2$) $E(x,y) = E(y,x)$ *(symmetry);*
($E3$) $T(E(x,y),E(y,z)) \leq E(x,z)$ *(T-transitivity).*

Using the idea of upper and lower approximation operators in [2] we have described upper and lower general aggregation operators based on a fuzzy equivalence relation. In this paper we are dealing with the upper general aggregation operator.

Definition 4. *Let* $A\colon [0,1]^n \to [0,1]$ *be an aggregation operator,* T *be a left continuous t-norm and* E *be a* T*-fuzzy equivalence relation defined on a set* X*. The upper general aggregation operators* $\tilde{A}_{E,T} : \bigcup_n([0,1]^X)^n \to [0,1]^X$ *is defined by*

$$\tilde{A}_{E,T}(\mu_1,\ldots,\mu_n)(x) = \sup_{x' \in X} T(E(x,x'), A(\mu_1(x'),\ldots,\mu_n(x'))), \quad (1)$$

where $x \in X$ *and* $\mu_1,\ldots,\mu_n \in [0,1]^X$.

In [2] it was shown that operator $\tilde{A}_{E,T}$ actually is a general aggregation operator. In this paper we will use constructions of $\tilde{A}_{E,T}$ based on minimum t-norm T_M, product t-norm T_P and Lukasiewicz t-norm T_L. We also use in (1) such ordinary aggregation operators as arithmetic mean (AVG), weighted arithmetic mean ($WAVG$, in which the weights for the smallest and the greatest elements are two times smaller than for the other elements, which weights are equal), and the modification of $WAVG$, where the weights for the smallest and the greatest elements are excluded ($ZAVG$).

3 Upper General Aggregation Operator in Risk Level Assessments

Now we will demonstrate how such construction could be used in risk level assessments. We consider the following problem. Let X be the set of all countries in the world or in some particular region. Let $\mu_i(x)$, where $\mu_i \colon X \to [0,1]$, be a normalized evaluation of country's $x \in X$ risk level by the i-th expert (a country is considered as more risky if this evaluation is closer to 1). As ordinary aggregation operator A one could take the arithmetic mean or the weighted arithmetic mean aggregation operator. It is important to define appropriate fuzzy equivalence relation $E \colon X \times X \to [0,1]$ between the objects of X.

We want to obtain an assessment of the risk level of some country by taking arithmetic mean of the experts evaluations of other countries taking into account fuzzy equivalence relation between these countries and to compare it with ordinary arithmetic mean operator.

Let us consider the discrete universe X, which consists of 8 countries from some region, and the following T_M-fuzzy equivalence relation E given in a matrix form:

$$E = \begin{pmatrix} 1 & 0.9 & 0.8 & 0.8 & 0.8 & 0.7 & 0.1 & 0.1 \\ 0.9 & 1 & 0.8 & 0.8 & 0.8 & 0.7 & 0.1 & 0.1 \\ 0.8 & 0.8 & 1 & 0.8 & 0.8 & 0.7 & 0.1 & 0.1 \\ 0.8 & 0.8 & 0.8 & 1 & 0.8 & 0.7 & 0.1 & 0.1 \\ 0.8 & 0.8 & 0.8 & 0.8 & 1 & 0.7 & 0.1 & 0.1 \\ 0.7 & 0.7 & 0.7 & 0.7 & 0.7 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 1 & 0.9 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.9 & 1 \end{pmatrix}.$$

Suppose, we have evaluations of the risk level for each country given by 5 experts and expressed in the form of fuzzy sets $\mu_i : X \to [0,1]$, $i = 1, \ldots, 5$:

$$\mu_1 = \begin{pmatrix} 1 \\ 0.9 \\ 0.8 \\ 0.6 \\ 1 \\ 0.6 \\ 0 \\ 0.9 \end{pmatrix}, \ \mu_2 = \begin{pmatrix} 0.8 \\ 0.8 \\ 0.7 \\ 0.8 \\ 0.4 \\ 0.1 \\ 0.1 \\ 0.8 \end{pmatrix}, \ \mu_3 = \begin{pmatrix} 1 \\ 0.6 \\ 0.7 \\ 0.6 \\ 0.6 \\ 0.5 \\ 0.1 \\ 0.9 \end{pmatrix}, \ \mu_4 = \begin{pmatrix} 0.9 \\ 1 \\ 0.9 \\ 0.8 \\ 0.6 \\ 0.7 \\ 0 \\ 0.8 \end{pmatrix}, \ \mu_5 = \begin{pmatrix} 0.7 \\ 0.5 \\ 0.7 \\ 0.6 \\ 0.6 \\ 0.6 \\ 0.1 \\ 0.7 \end{pmatrix}.$$

First, we calculate the results of aggregations $\widetilde{AVG}$, $\widetilde{WAVG}$, $\widetilde{ZAVG}$, where $\widetilde{AVG}$, $\widetilde{WAVG}$, $\widetilde{ZAVG}$ are the pointwise extensions of AVG, $WAVG$, $ZAVG$ respectively, thus obtaining the aggregated evaluations:

$$\widetilde{AVG} = \begin{pmatrix} 0.88 \\ 0.76 \\ 0.76 \\ 0.68 \\ 0.64 \\ 0.50 \\ 0.06 \\ 0.82 \end{pmatrix}, \ \widetilde{WAVG} = \begin{pmatrix} 0.44 \\ 0.38 \\ 0.39 \\ 0.34 \\ 0.33 \\ 0.24 \\ 0.03 \\ 0.41 \end{pmatrix}, \ \widetilde{ZAVG} = \begin{pmatrix} 0.45 \\ 0.38 \\ 0.37 \\ 0.33 \\ 0.30 \\ 0.28 \\ 0.03 \\ 0.42 \end{pmatrix}.$$

We can see, that operators $\widetilde{WAVG}$ and $\widetilde{ZAVG}$ decrease the evaluations of risk level comparing to $\widetilde{AVG}$, but give similar values.

Then we apply the upper general aggregation operator in order to obtain the upper approximation of AVG, $WAVG$, $ZAVG$ taking into account equivalence relation E. This result will give us the most conservative assessment of risk levels for each country, which is the goal of a risk manager.

First, let us take the strongest t-norm $T = T_M$ and apply the upper general aggregation operator. For example:

$$\widetilde{AVG}_{E,T_M}(\mu_1, \ldots, \mu_5)(x) = \max_{x' \in X} T_M(E(x, x'), AVG(\mu_1(x'), \ldots, \mu_5(x'))).$$

As a result we obtain the following fuzzy sets in a vector form:

$$\widetilde{AVG}_{E,T_M} = \begin{pmatrix} 0.88 \\ 0.88 \\ 0.80 \\ 0.80 \\ 0.80 \\ 0.70 \\ 0.82 \\ 0.82 \end{pmatrix}, \ \widetilde{WAVG}_{E,T_M} = \begin{pmatrix} 0.44 \\ 0.44 \\ 0.44 \\ 0.44 \\ 0.44 \\ 0.44 \\ 0.41 \\ 0.41 \end{pmatrix}, \ \widetilde{ZAVG}_{E,T_M} = \begin{pmatrix} 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.42 \\ 0.42 \end{pmatrix}.$$

As one can see, the risk level assessments by using T_M equalize. At the same time the assessments are more conservative. For example, the average risk level for the 7th country is 0.06, while the upper approximation $\widetilde{AVG}$ taking into account the equivalence relation gives us more conservative result 0.82, because this country is equivalent with the high degree to the much more risky 8th country. Operators $\widetilde{WAVG}$, $\widetilde{ZAVG}$ give us less conservative result.

Now let us consider the results obtained by using Lukasiewicz t-norm T_L:

$$\widetilde{AVG}_{E,T_L} = \begin{pmatrix} 0.88 \\ 0.78 \\ 0.76 \\ 0.68 \\ 0.68 \\ 0.58 \\ 0.72 \\ 0.82 \end{pmatrix}, \ \widetilde{WAVG}_{E,T_L} = \begin{pmatrix} 0.44 \\ 0.38 \\ 0.39 \\ 0.34 \\ 0.33 \\ 0.24 \\ 0.31 \\ 0.41 \end{pmatrix}, \ \widetilde{ZAVG}_{E,T_L} = \begin{pmatrix} 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.45 \\ 0.42 \\ 0.42 \end{pmatrix}.$$

These results are more trustful. Vectors $\widetilde{WAVG}_{E,T_L}$ and $\widetilde{ZAVG}_{E,T_L}$ have similar values, but the assessment obtained by $\widetilde{AVG}_{E,T_L}$ is more conservative and one could more probably choose this operator.

Finally, we use product t-norm T_P:

$$\widetilde{AVG}_{E,T_P} = \begin{pmatrix} 0.88 \\ 0.79 \\ 0.76 \\ 0.70 \\ 0.70 \\ 0.62 \\ 0.74 \\ 0.82 \end{pmatrix}, \ \widetilde{WAVG}_{E,T_P} = \begin{pmatrix} 0.44 \\ 0.39 \\ 0.39 \\ 0.35 \\ 0.35 \\ 0.31 \\ 0.37 \\ 0.41 \end{pmatrix}, \ \widetilde{ZAVG}_{E,T_P} = \begin{pmatrix} 0.45 \\ 0.41 \\ 0.37 \\ 0.36 \\ 0.36 \\ 0.32 \\ 0.38 \\ 0.42 \end{pmatrix}.$$

The results are similar to the previous case of Lukasiewcz t-norm T_L. To summarize we could make the conclusion, that operators $\widetilde{AVG}_{E,T_L}$ and $\widetilde{AVG}_{E,T_P}$ suit better for our purposes. One can see that for the 7th and the 8th elements values equalize. It means that with t-norms T_P and T_L equivalence relation E has more influence on the result of aggregation.

Depending on the problem specifics and the construction of T-fuzzy equivalence relation E one could choose different t-norms T and aggregation operators A, which will influence the result. Analysing not only this, but also several other examples, we conclude that for the risk level assessment it is better to use AVG, which give more conservative results.

4 Fuzzy Equivalence Relation Based on a Metric

Suppose we have a metric space (X, d). In order to construct a fuzzy equivalence relation we will use the following formula:

$$E(x, y) = \frac{1}{1 + d(x, y)}, \qquad x, y \in X. \tag{2}$$

By using formula (2) we could obtain relation E, which is not necessary T-transitive for an arbitrary t-norm T. In general, this condition could not fulfil for the minimum t-norm, but always holds for the product and Lukasiewicz t-norm. For example, let us consider the following fuzzy equivalence relation matrix

$$E = \begin{pmatrix} 1 & 0.5 & 0.7 \\ 0.5 & 1 & 0.6 \\ 0.7 & 0.6 & 1 \end{pmatrix},$$

which is obtained by using (2). This relation E is T-transitive with respect to T_L and T_P, but is not T_M-transitive.

Let us show that fuzzy equivalence relation E, which is based on the metric (2), is T_P- and T_L-transitive. First, we will check condition $(E3)$ for all elements

$x, y, z \in X$ in the case of product t-norm T_P:

$$\frac{1}{1+d(x,y)} \cdot \frac{1}{1+d(y,z)} \leq \frac{1}{1+d(x,z)}.$$

We rewrite the inequality in equivalent forms

$$(1+d(x,y)) \cdot (1+d(y,z)) \geq 1+d(x,z) \iff$$
$$\iff d(x,y)+d(y,z)+d(x,y)d(y,z) \geq d(x,z),$$

the last of which holds by the triangle inequality.

Now let us demonstrate that $(E3)$ holds for each $x, y, z \in X$ in the case of Lukasiewicz t-norm T_L

$$\min\left(\frac{1}{1+d(x,y)} + \frac{1}{1+d(y,z)} - 1, 0\right) \leq \frac{1}{1+d(x,z)}.$$

We should check that

$$\frac{1}{1+d(x,y)} + \frac{1}{1+d(y,z)} - 1 \leq \frac{1}{1+d(x,z)} \iff$$
$$\iff \frac{2+d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)+d(x,y)d(y,z)} \leq \frac{1}{1+d(x,z)} + 1.$$

Now we evaluate

$$\frac{2+d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)+d(x,y)d(y,z)} \leq \frac{2+d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)} =$$
$$= \frac{1}{1+d(x,y)+d(y,z)} + 1.$$

We rewrite the inequality in equivalent forms

$$\frac{1}{1+d(x,y)+d(y,z)} + 1 \leq \frac{1}{1+d(x,z)} + 1 \iff$$
$$\iff 1+d(x,y)+d(y,z) \geq 1+d(x,z),$$

the last of which holds by the triangle inequality.

Let us demonstrate by using the real-world data how a fuzzy equivalence relation could be constructed. We consider universe X of five objects. As the objects in this example we will choose 5 countries, and we will construct a fuzzy equivalence relation between them. The distances between countries are constructed on the basis of different macroeconomic factors:

(1) economic growth (GDP, annual variation in %);
(2) industrial production (annual variation in %);
(3) unemployment rate (in %);

Table 1. Normalized macroeconomic factors for five countries

	1	2	3	4	5
Economic growth (GDP, annual variation in %)	2.54	2.41	2.44	−1.17	−1.22
Industrial production (annual variation in %)	4.65	2.85	3.08	−2.67	−2.91
Unemployment rate (in %)	0.71	0.81	1.09	1.02	1.37
Public debt (% of GDP)	1.11	1.33	0.25	1.04	1.27
Current account (% of GDP)	3.23	0.75	1.09	1.49	−1.55

(4) public debt (% of GDP);
(5) current account (% of GDP).

The obtained factors, normalized by the average of each factor, are given in Table 1.

Let us denote the number of macroeconomic factors by K, then each object could be represented as a vector of dimension K. We use the Euclidean metric to evaluate distances between the objects:

$$d(x,y) = \sqrt{\sum_{k=1}^{K}(x_k - y_k)^2}, \qquad x,y \in \mathbb{R}^K. \tag{3}$$

In order to obtain fuzzy equivalence matrix E we apply formula (2):

$$E = \begin{pmatrix} 1 & 0.24 & 0.26 & 0.11 & 0.09 \\ 0.24 & 1 & 0.46 & 0.13 & 0.12 \\ 0.26 & 0.46 & 1 & 0.13 & 0.12 \\ 0.11 & 0.13 & 0.13 & 1 & 0.24 \\ 0.09 & 0.12 & 0.12 & 0.24 & 1 \end{pmatrix}.$$

As one can see, the values of degrees of equivalence are very small, but, for example, the distance between countries 1 and 2 is very small as well, which means, that equivalence degree should be high. In order to improve the equivalence relation, we introduce additional parameter c and modify formula (2) in the following way:

$$E(x,y) = \frac{1}{1 + cd(x,y)}, \qquad x,y \in X. \tag{4}$$

Such parameter c does not affect that E is T_P- and T_L-transitive.

To determinate the value of parameter c we consider the following function:

$$\phi_c(t) = \frac{1}{1 + ct}, \qquad t \geq 0. \tag{5}$$

As argument t of function (5) we will take metric $d(x,y)$. Let us consider the example, where metric $d(x,y)$ takes values from interval $[0,1]$, and calculate

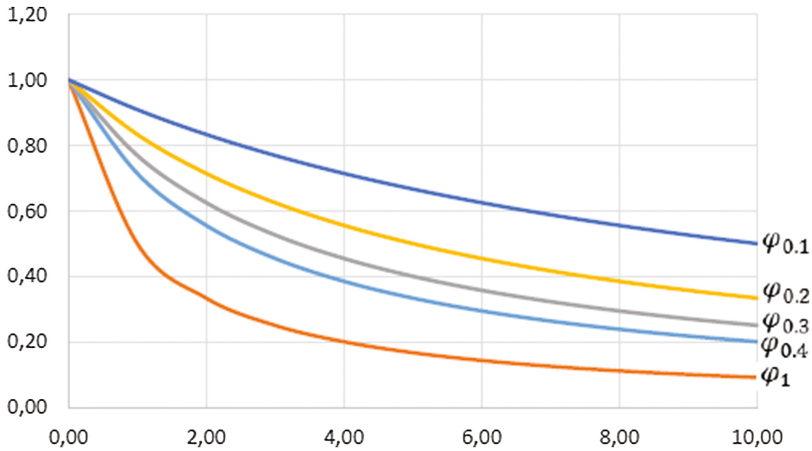


Fig. 1. Dependence of function ϕ on parameter c

the values of function ϕ for some values of parameter c. Figure 1 shows the dependence of the values of function ϕ on different values of parameter c.

In order to improve the constructed equivalence relation, we will choose $c = 0.2$. Then we will obtain the following equivalence relation matrix

$$E = \begin{pmatrix} 1 & 0.62 & 0.64 & 0.37 & 0.34 \\ 0.62 & 1 & 0.81 & 0.43 & 0.41 \\ 0.64 & 0.81 & 1 & 0.42 & 0.40 \\ 0.37 & 0.43 & 0.42 & 1 & 0.62 \\ 0.34 & 0.41 & 0.40 & 0.62 & 1 \end{pmatrix},$$

which is relevant to the real situation.

5 Aggregation of Experts' Evaluations of Countries Risk Level

In this section we apply the construction of the upper general aggregation operator for the problem of aggregation of several experts' evaluations of a country risk level, taking into account a fuzzy equivalence relation between countries. We consider 28 countries from the same region. For the construction of an equivalence relation we use the approach described in the previous section. We take the same macroeconomic factors for 5 years and take the average values of them. We calculate the distances between the countries on the basis of these macroeconomic factors, and then we construct the equivalence relation matrix by using $c = 0.2$.

Let us consider the case when we have four experts and their evaluations of each country's risk level (Table 2 shows evaluations for some of the countries).

Table 2. Experts evaluations

	μ_1	μ_2	μ_3	μ_4
1	0.1	0.6	0.4	0.3
3	0.2	0.7	0.7	0.8
4	0.9	0.7	0.4	0.8
8	0.6	0.2	0.5	0.3
12	0.8	0.9	0.6	0.7
13	0.7	0.7	0.8	0.5
14	0.1	0.2	0.3	0.5
19	0.2	0.5	0.5	0.4
26	0.4	0.3	0.2	0.5

Then we calculate the average AVG of these evaluations and the results of the upper general aggregation operator using ordinary aggregation operator AVG, and the product and Lukasiewicz t-norms. As the result, we obtain vectors $\widetilde{AVG}$, $\widetilde{AVG}_{E,T_P}$, $\widetilde{AVG}_{E,T_L}$.

Table 3. Results of aggregation

	$\widetilde{AVG}$	$\widetilde{AVG}_{E,T_P}$	$\widetilde{AVG}_{E,T_L}$
1	0.35	0.45	0.35
3	0.60	0.60	0.60
4	0.70	0.70	0.70
8	0.40	0.54	0.48
12	0.75	0.75	0.75
13	0.68	0.68	0.68
14	0.28	0.46	0.35
19	0.40	0.41	0.40
26	0.35	0.62	0.58

In order to analyse how the equivalence relation matrix influences the result of aggregation, we comment the results for each country from Table 3, and then we give a conclusion.

1. The 1st country is equivalent with the high degree to the 3rd (the equivalence degree is 0.75) and to the 13th (the equivalence degree is 0.66), but the 13th country has higher experts' evaluation. In this case we can see that both values (the equivalence relation and the expert evaluation) influence the result. The aggregation using the Lukasiewicz t-norm has no difference with ordinary aggregation AVG.

2. For the 3rd country the result doesn't change, however this country is equivalent with the high degree to the 1st country and to the 6th country. Experts' evaluations for these countries are quite low, therefore they have no influence on the result for the 3rd country.
3. The similar situation is for the 4th country: it is equivalent with the high degree 0.88 to the 26th country, which has low experts' evaluation 0.35.
4. For the 8th country the results of upper general aggregation are better than the ordinary aggregation obtained by using both t-norms, and the influence comes from the 17th country.
5. For the countries with numbers 12 and 13 the results of aggregations are the same, because for the high experts' evaluation the equivalence degrees are not high enough.
6. The 14th country has significant increase in the result of the upper general aggregation comparing to the ordinary aggregation. It influences by the high degree of equivalence with the 13th country, which has high experts' evaluations.
7. For the 19th country the result of the upper general aggregation obtained by using product t-norm T_P slightly differs due to the influence of the 3rd country, which is equivalent to the 19th with the degree 0.68.
8. The results of the upper general aggregation for the 26th country significantly increase in compare to the ordinary aggregation. These changes are influenced by the 4th country, which is equivalent to the 26th with the degree 0.88 and has very high experts' evaluations.

Summarizing, we can see that the result is highly influenced by the values of experts evaluations. If the evaluation is very high, then it influences other countries, for which the experts evaluation was not so high. For the most cases the result of the upper general aggregation differs from the ordinary aggregation using the product t-norm. In order to increase the influence of the Lukasiewicz t-norm, we should increase the influence of the equivalence relation matrix. It is possible by decreasing the value of parameter c. If it is necessary to decrease the influence of matrix E, then we should increase the value of parameter c. It is the way how we can control the influence of the equivalence relation.

6 Conclusion

The proposed technique was illustrated with a real-world example of the investment risk level aggregation for countries of one region using an equivalence relation obtained on the basis of different macroeconomic factors. The obtained results are applicable and could be modified using not only numerical parameters, but also other ordinary aggregation operator and other t-norm. The dipper analysis of the influence of all these factors on the result of aggregation will be a subject of further research.

References

1. Risk aggregation and diversification, Research paper, Committee on Risk Management and Capital Requirements (2016)
2. Orlovs, P., Asmuss, S.: General aggregation operators based on a fuzzy equivalence relation in the context of approximate systems. Fuzzy Sets Syst. **291**, 114–131 (2016)
3. Orlovs, P., Asmuss, S.: On some properties of general aggregation operators based on a fuzzy equivalence relation. Adv. Intell. Syst. Res. **89**, 1354–1361 (2014)
4. Orlovs, P., Asmuss, S.: Upper and lower generalized factoraggregations based on fuzzy equivalence relation. In: Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1772–1777 (2014)
5. Chan, H.K., Wang, X.: Hierarchical Model for Risk Assessment. Springer, London (2013)
6. Grabisch, M., Marichal, J.-L., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, New York (2009)
7. Calvo, T., Mayor, G., Mesiar, R.: Aggregation Operators. Physica-Verlag, Heidelberg (2002)
8. Takaci, A.: General aggregation operators acting on fuzzy numbers induced by ordinary aggregation operators. Novi Sad J. Math. **33**(2), 67–76 (2003)
9. Pezier, J.: Application-based financial risk aggregation methods. ISMA Center Discussion Papers 2003–11, pp. 1–16 (2003)
10. Detyniecki, M.: Fundamentals on Aggregation Operators, Berkeley (2001)

Six Sigma Project Selection Using Interval Neutrosophic TOPSIS

İrem Otay[1(✉)] and Cengiz Kahraman[2]

[1] Department of Industrial Engineering, Engineering Faculty, Okan University, 34959 Tuzla-Istanbul, Turkey
irem.otay@okan.edu.tr

[2] Department of Industrial Engineering, Management Faculty, Istanbul Technical University, 34367 Maçka-Istanbul, Turkey
kahramanc@itu.edu.tr

Abstract. Six Sigma approaches aim at providing almost defect-free products and/or services to customers. Six Sigma is a powerful and comprehensive management tool for meeting customer needs. Well-designed projects are capable to provide significant financial benefits, bring competitive advantage and increased customer satisfaction.

Well-designed projects having clear and concise descriptions and objectives are capable to provide significant financial benefits, increased customer satisfaction and bring competitive advantage. Selecting Six Sigma improvement projects has been one of the most challenging and frequently discussed issues in the literature. Selecting the most useful project/s is a key success factor in Six Sigma approach. Selecting Six Sigma projects is a multi criteria decision making problem involving many tangible and intangible criteria under uncertainty. In this paper, uncertainty will be handled by neutrosophic sets. "A neutrosophic set deals with the origin, nature, and scope of neutralities, as well as their interactions with different ideational spectra" [1]. In neutrosophic sets, truth-membership, indeterminacy-membership and falsity-membership are all together included. Neutrosophic sets are accepted as a super set of the other types of sets such as classical sets, ordinary fuzzy sets, hesitant fuzzy sets, intuitionistic fuzzy sets, and soft sets.

In this paper, we employ interval neutrosophic TOPSIS method to evaluate Six Sigma projects. By reviewing the literature, seven criteria e.g. total cost, required time and customer satisfaction are taken into account. To the best knowledge of the authors, this is the first study to evaluate Six Sigma projects using interval neutrosophic TOPSIS approach with group decision making.

Keywords: Interval neutrosophic sets · Multi criteria decision making · TOPSIS · Six sigma project selection

1 Neutrosophic Sets in Multicriteria Decision Making

Neutrosophic sets have been introduced by Florantin Smandarache in 2013. These sets aim at including the information for truthness, indeterminacy, and falsity all together in a set. Neutrosophic sets have been extensively used in decision making processes since

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_8

they appeared in 2013. In the following we present some literature review results using graphical illustrations. Figure 1 shows the journals publishing Neutrosophic based papers. Journal of Intelligent and Fuzzy Systems and Neural Computing and Applications are the leading journals in publishing the neutrosophic based papers.

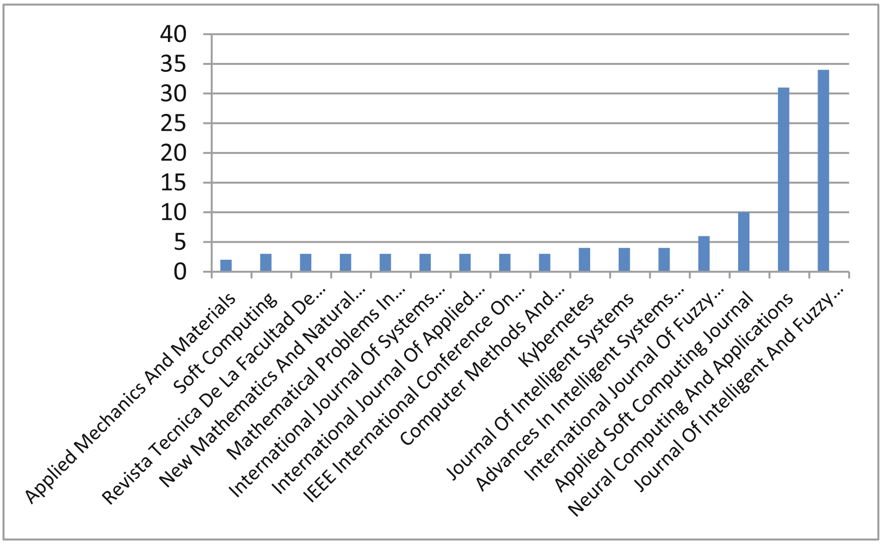


Fig. 1. Journals publishing Neutrosophic based papers

Figure 2 presents the authors who have most published neutrosophic based papers. J. Ye, F. Smarandache, Y. Guo, and P. Liu are the leading authors.

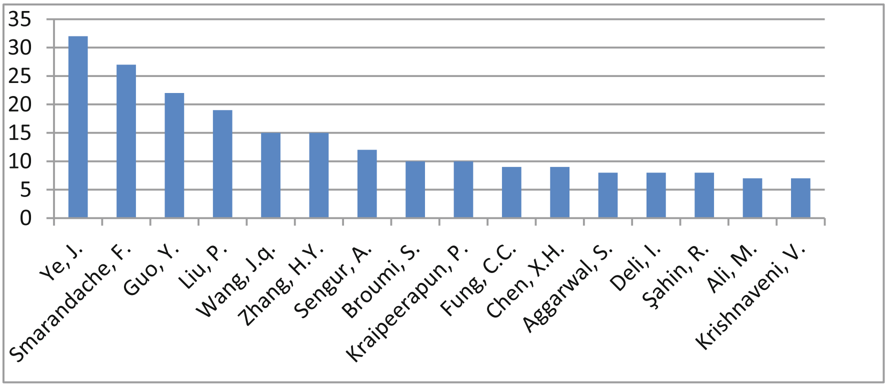


Fig. 2. Researchers publishing Neutrosophic based papers

Figure 3 shows the Universities most publishing neutrosophic based papers. Shaoxing University, University of New Mexico, Central South University China, and Shandong University of Finance are the leading universities in this area.

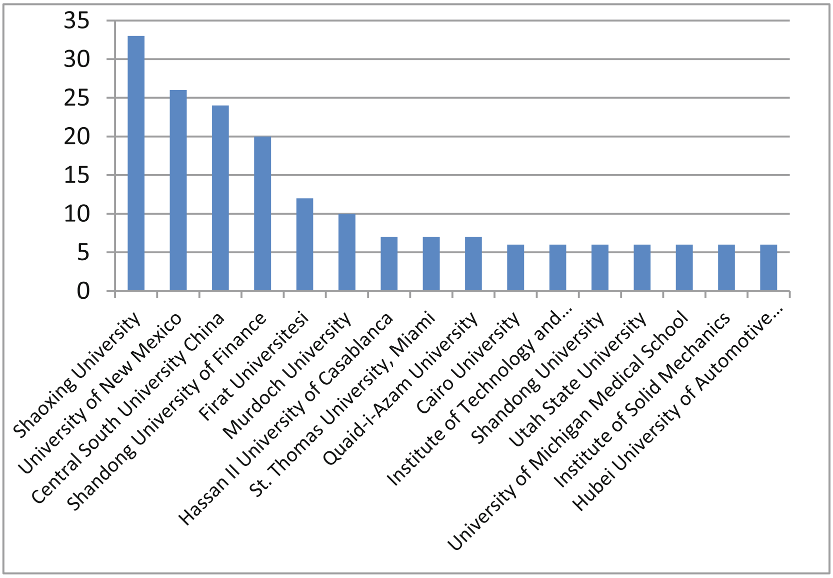


Fig. 3. Affiliations most publishing Neutrosophic based papers

Figure 4 illustrates the countries most publishing neutrosophic based papers. China, United States, India, and Turkey are the leading countries in this area, respectively.

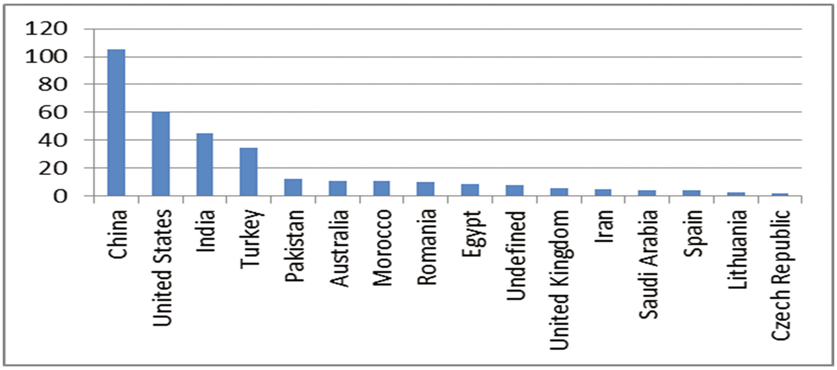


Fig. 4. Countries most publishing Neutrosophic based papers

Figure 5 illustrates the subject areas of the neutrosophic based papers. Computer science, mathematics, and engineering are the first major three areas in which neutrosophic sets are applied.

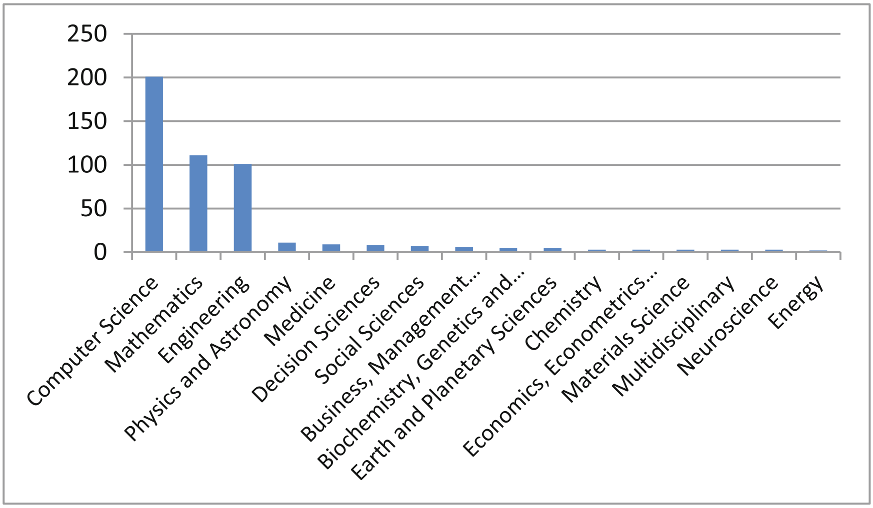


Fig. 5. Subject areas of neutrosophic based publications

In the literature, there are valuable studies on multi-criteria decision making using neutrosophic sets. For instance, Peng et al. [2] proposed a new outranking approach using a simplified neutrosophic sets based on ELECTRE. Peng et al. [3] defined the operations of multi-valued neutrosophic numbers and used Einstein operations and proposed the multi-valued neutrosophic power weighted average operator and geometric operator. Tian et al. [4] applied TOPSIS based on interval neutrosophic sets and proposed a fuzzy cross-entropy approach. Bausys et al. [5] introduced the complex proportional assessment method (COPRAS) method with single value neutrosophic sets. Liu and Zhang [6] extended the VIKOR method using the neutrosophic hesitant fuzzy information and proposed a neutrosophic hesitant fuzzy VIKOR approach. Bausys and Zavadskas [7] developed the extension of VIKOR method with interval-valued neutrosophic sets. Wang and Liu [8] implemented Optimized PROMETHEE method using interval neutrosophic sets to select the best new energy storage alternative. Zavadkas et al. [9] evaluated three different circuit design schemes using neutrosophic WASPAS method with single-valued neutrosophic set. Tian et al. [10] analyzed green product design selection problem integrating power aggregation operators and a TOPSIS-based QUALIFLEX method.

2 Preliminaries of Neutrosophic Sets

2.1 Arithmetic Operations with Neutrosophic Sets

Definition 1. $T_A(x)$, $I_A(x)$, and $F_A(x)$, which are the truth-membership function, the indeterminancy-membership function and the falsity-membership function, respectively are real standard or nonstandard subsets of $]0^-, 1^+[$. Their sum may be at most 3 such that

$$0^- \leq \sup T_A(x) + \sup I_A(x) + \sup F_A(x) \leq 3^+ \quad (1)$$

Definition 2. The union of two neutrosophic sets A and B is a neutrosophic set C whose truth-membership, indeterminancy-membership and false-membership function are $T_C(x) = T_A(x) \oplus T_B(x) \ominus T_A(x) \odot T_B(x)$, $I_C(x) = I_A(x) \oplus I_B(x) \ominus I_A(x) \odot I_B(x)$, and $F_C(x) = F_A(x) \oplus F_B(x) \ominus F_A(x) \odot F_B(x)$ for any x in X, respectively [11].

Definition 3. The intersection of two neutrosophic sets A and B is a neutrosophic set C whose truth-membership, indeterminancy-membership and false-membership functions are $T_C(x) = T_A(x) \odot T_B(x)$, $I_C(x) = I_A(x) \odot I_B(x)$, and $F_C(x) = F_A(x) \odot F_B(x)$ for any x in X, respectively.

2.2 Arithmetic Operations with Interval Neutrosophic Sets

Definition 4. An INS A in X is characterized by $T_A(x) = [\inf T_A(x), \sup T_A(x)]$, $I_A(x) = [\inf I_A(x), \sup I_A(x)]$, $F_A(x) = [\inf F_A(x), \sup F_A(x)] \subseteq [0, 1]$, and $0 \leq \sup T_A(x) + \sup I_A(x) + \sup F_A(x) \leq 3$ $x \in X$ ([11]).

3 Neutrosophic TOPSIS with Group Decision Making

Assume there are m alternatives and n criteria in a decision matrix with the weights of criteria wj where $0 \leq w_j \leq 1, \sum_{j=1}^{n} w_j = 1$

Suppose $NS_{ij}(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ is the neutrosophic score of alternative A_i with respect to criterion C_j. Then, $NS_{ij} = \langle \left[x_{ij}^L, x_{ij}^U\right], \left[T_{ij}^L, T_{ij}^U\right], \left[I_{ij}^L, I_{ij}^U\right], \left[F_{ij}^L, F_{ij}^U\right] \rangle$ where $\left[x_{ij}^L, x_{ij}^U\right]$ is the interval valued score and $T_{ij}^U, I_{ij}^U, F_{ij}^U, T_{ij}^L, I_{ij}^L, F_{ij}^L \in [0, 1]$ and $0 \leq T_{ij}^U + I_{ij}^U + F_{ij}^U \leq 3$ ([12]). For obtaininig the attributes weight vector, the distance between two interval neutrosophic scores is defined as in Eq. (2).

Definition 5. Let $\widetilde{S_1} = <[S_{a_1}, S_{b_1}], \left([T_A^L, T_A^U], [I_A^L, I_A^U], [F_A^L, F_A^U]\right) >$ and $\widetilde{S_2} = <[S_{a_2}, S_{b_2}], \left([T_B^L, T_B^U], [I_B^L, I_B^U], [F_B^L, F_B^U]\right) >$ be any two interval neutrosophic sets. and f: $\widetilde{S} \times \widetilde{S} \rightarrow R$. Then, the Hamming distance between $\widetilde{S_1}$ and $\widetilde{S_2}$ can be defined as follows ([12]):

$$\begin{aligned} d_{IVNS}(\widetilde{S_1}, \widetilde{S_2}) = \frac{1}{12(l-1)} &\left(\left|a_1 \times T_A^L - a_2 \times T_B^L\right| + \left|a_1 \times T_A^U - a_2 \times T_B^U\right|\right) \\ &+ \left|a_1 \times I_A^L - a_2 \times I_B^L\right| + \left|a_1 \times I_A^U - a_2 \times I_B^U\right| + \left|a_1 \times F_A^L - a_2 \times F_B^L\right| \\ &+ \left|a_1 \times F_A^U - a_2 \times F_B^U\right| + \left|b_1 \times T_A^L - b_2 \times T_B^L\right| + \left|b_1 \times T_A^U - b_2 \times T_B^U\right| \\ &+ \left|b_1 \times I_A^L - b_2 \times I_B^L\right| + \left|b_1 \times I_A^U - b_2 \times I_B^U\right| + \left|b_1 \times F_A^L - b_2 \times F_B^L\right| \end{aligned} \tag{2}$$

For a certain criterion $C_j \in C$, the distance $d(z_{ij}, z_{kj})$ to represent the weighted deviation between attribute values z_{ij} and z_{kj} is given by $D_{ij} = \sum_{k=1}^{m} d(z_{ij}, z_{kj}) w_j$. And the normalized attribute weight is given by Eq. (3):

$$w_j = \frac{\sum_{i=1}^{m} \sum_{k=1}^{m} d(z_{ij}, z_{kj})}{\sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} d(z_{ij}, z_{kj})} \tag{3}$$

The steps of the Interval Neutrosophic TOPSIS method are given in the following (Broumi et al. [12]):

Step 1: Normalize the decision matrix

The normalized matrix $R = (r_{ij})$, where $r_{ij} = <[r_{ij}^L, r_{ij}^U], \left([\dot{T}_{ij}^L, \dot{T}_{ij}^U], [\dot{I}_{ij}^L, \dot{I}_{ij}^U], [\dot{F}_{ij}^L, \dot{F}_{ij}^L]\right) >$ is obtained as follows:

For benefit type,

$$\begin{cases} r_{ij}^L = x_{ij}^L, r_{ij}^U = x_{ij}^U \ \text{for} \ (1 \le i \le m, 1 \le j \le n) \\ \dot{T}_{ij}^L = T_{ij}^L, \dot{T}_{ij}^U = T_{ij}^U, \dot{I}_{ij}^L = I_{ij}^L, \dot{I}_{ij}^U = I_{ij}^U, \dot{F}_{ij}^L = F_{ij}^L, \dot{F}_{ij}^U = F_{ij}^U \end{cases} \tag{4}$$

For cost type,

$$\begin{cases} r_{ij}^L = neg(x_{ij}^L), r_{ij}^U = neg(x_{ij}^U) \ \text{for} \ (1 \le i \le m, 1 \le j \le n) \\ \dot{T}_{ij}^L = T_{ij}^L, \dot{T}_{ij}^U = T_{ij}^U, \dot{I}_{ij}^L = I_{ij}^L, \dot{I}_{ij}^U = I_{ij}^U, \dot{F}_{ij}^L = F_{ij}^L, \dot{F}_{ij}^U = F_{ij}^U \end{cases} \tag{5}$$

Step 2: Construct the weighted normalize matrix

$$Y = [y_{ij}]_{m\times n} \begin{bmatrix} <[y_{11}^L,\ y_{11}^U],], \left([\ddot{T}_{11}^L,\ \ddot{T}_{11}^U],\ [\ddot{I}_{11}^L,\ \ddot{I}_{11}^U],\ [\ddot{F}_{11}^L,\ \ddot{F}_{11}^U]\right) > \\ \ldots <[y_{11}^L, y_{11}^U],], \left([\ddot{T}_{1n}^L,\ \ddot{T}_{1n}^U],\ [\ddot{I}_{1n}^L,\ \ddot{I}_{1n}^U],\ [\ddot{F}_{1n}^L,\ \ddot{F}_{1n}^U]\right) > \\ <[y_{21}^L,\ y_{21}^U],], \left([\ddot{T}_{21}^L,\ \ddot{T}_{21}^U],\ [\ddot{I}_{21}^L,\ \ddot{I}],\ [\ddot{F}_{21}^L,\ \ddot{F}_{21}^U]\right) > \\ \ldots <[y_{2n}^L, y_{2n}^U],], \left([\ddot{T}_{2n}^L,\ \ddot{T}_{2n}^U],\ [\ddot{I}_{2n}^L,\ \ddot{I}_{2n}^U],\ [\ddot{F}_{2n}^L,\ \ddot{F}_{2n}^U]\right) > \\ \hline <[y_{mn}^L,\ y_{mn}^U],], \left([\ddot{T}_{mn}^L,\ \ddot{T}_{mn}^U],\ [\ddot{I}_{mn}^L,\ \ddot{I}_{mn}^U],\ [\ddot{F}_{mn}^L,\ \ddot{F}_{mn}^U]\right) > \\ \ldots <[y_{mn}^L,\ y_{mn}^U],], \left([\ddot{T}_{mn}^L,\ \ddot{T}_{mn}^U],\ [\ddot{I}_{mn}^L,\ \ddot{I}_{mn}^U],\ [\ddot{F}_{mn}^L,\ \ddot{F}_{mn}^U]\right) > \end{bmatrix} \tag{6}$$

where

$$\begin{cases} y_{ij}^L = w_j r_{ij}^L,\ y_{ij}^U = w_j r_{ij}^U \\ \ddot{T}_{ij}^L = 1 - (1 - \dot{T}_{ij}^L)^{w_j},\ \ddot{T}_{ij}^U = 1 \quad -(1 - \dot{T}_{ij}^U)^{w_j},\ \ddot{I}_{ij}^L = (\dot{I}_{ij}^L)^{w_j}, \\ \ddot{I}_{ij}^L = (\dot{I}_{ij}^U)^{w_j},\ \ddot{F}_{ij}^L = (\dot{F}_{ij}^L)^{w_j},\ \ddot{F}_{ij}^U = (\dot{F}_{ij}^U)^{w_j} \end{cases}$$

Step 3: Identify, the sets of the positive ideal solution $Y^+ = \left(y_1^+, y_2^+, \ldots, y_m^+\right)$ and the negative ideal solution$Y^- = \left(y_1^-, y_2^-, \ldots, y_m^-\right)$, then we can get

$$\begin{aligned} Y^+ = (y_1^+,\ y_2^+, \ldots, y_m^+) = \quad & \Big(<[y_1^{L+},\ y_1^{U+}],\ ([\ddot{T}_1^{L+},\ \ddot{T}_1^{U+}],\ [\ddot{I}_1^{L+},\ \ddot{I}_1^{U+}],\ [\ddot{F}_1^{L+},\ \ddot{F}_1^{U+}]) >, \\ & <[y_2^{L+},\ y_2^{U+}],\ ([\ddot{T}_2^{L+},\ \ddot{T}_2^{U+}],\ [\ddot{I}_2^{L+},\ \ddot{I}_2^{U+}],\ [\ddot{F}_2^{L+},\ \ddot{F}_2^{U+}]) >, \ldots, \\ & <[y_n^{L+},\ y_n^{U+}],\ ([\ddot{T}_n^{L+},\ \ddot{T}_n^{U+}],\ [\ddot{I}_n^{L+},\ \ddot{I}_n^{U+}],\ [\ddot{F}_n^{L+},\ \ddot{F}_n^{U+}]) > \end{aligned} \tag{7}$$

$$\begin{aligned} Y^- = (y_1^-,\ y_2^-, \ldots, y_m^-) = \quad & \Big(<[y_1^{L-},\ y_1^{U-}],\ ([\ddot{T}_1^{L-},\ \ddot{T}_1^{U-}],\ [\ddot{I}_1^{L-},\ \ddot{I}_1^{U-}],\ [\ddot{F}_1^{L-},\ \ddot{F}_1^{U-}]) >, \\ & <[y_2^{L-},\ y_2^{U-}],\ ([\ddot{T}_2^{L-},\ \ddot{T}_2^{U-}],\ [\ddot{I}_2^{L-},\ \ddot{I}_2^{U-}],\ [\ddot{F}_2^{L-},\ \ddot{F}_2^{U-}]) >, \ldots, \\ & <[y_n^{L-},\ y_n^{U-}],\ ([\ddot{T}_n^{L-},\ \ddot{T}_n^{U-}],\ [\ddot{I}_n^{L-},\ \ddot{I}_n^{U-}],\ [\ddot{F}_n^{L-},\ \ddot{F}_n^{U-}]) > \end{aligned} \tag{8}$$

where

$$\begin{cases} y_j^{L+} = \max_i(y_{ij}^L),\ y_j^{U+} = \max_i(y_{ij}^U), \\ \ddot{T}_j^{L+} = \max_i(\ddot{T}_{ij}^L),\ \ddot{T}_j^{U+} = \max_i(\ddot{T}_{ij}^U),\ \ddot{I}_j^{L+} = \min_i(\ddot{I}_{ij}^L), \\ \ddot{I}_j^{U+} = \min_i(\ddot{I}_{ij}^U),\ \ddot{F}_j^{L+} = \min_i(\ddot{F}_{ij}^L)\ddot{F}_j^{U+} = \min_i(\ddot{F}_{ij}^U) \\ y_j^{L-} = \min_i(y_{ij}^L),\ y_j^{U-} = \min_i(y_{ij}^U), \\ \ddot{T}_j^{L-} = \min_i(\ddot{T}_{ij}^L),\ \ddot{T}_j^{U-} = \min_i(\ddot{T}_{ij}^U),\ \ddot{I}_j^{L-} = \max_i(\ddot{I}_{ij}^L), \\ \ddot{I}_j^{U-} = \max_i(\ddot{I}_{ij}^U),\ \ddot{F}_j^{L-} = \max_i(\ddot{F}_{ij}^L),\ \ddot{F}_j^{U-} = \max_i(\ddot{F}_{ij}^U), \end{cases} \tag{9}$$

Step 4: Obtain the distance between each alternative and the positive ideal solution and between each alternative and the negative ideal solution.

$$D^{+} = \left(d_1^{+}, d_2^{+}, \ldots, d_m^{+}\right);\ D^{-} = \left(d_1^{-}, d_2^{-}, \ldots, d_m^{-}\right) \tag{10}$$

where,

$$\begin{cases} d_i^{+} = \left[\sum_{j=1}^{n}\left(d\left(y_{ij}, y_j^{+}\right)\right)^2\right]^{\frac{1}{2}} \\ d_i^{-} = \left[\sum_{j=1}^{n}\left(d\left(y_{ij}, y_j^{-}\right)\right)^2\right]^{\frac{1}{2}} \end{cases} \tag{11}$$

where, $d\left(y_{ij}, y_j^{+}\right)$. is the distance between the interval valued neutrosophic sets y_{ij}. and y_j^{+}. and $d\left(y_{ij}, y_j^{-}\right)$. is the distance between the interval valued neutrosophic linguistic sets y_{ij}. and y_j^{-}..

Step 5: Obtain the closeness coefficient of each alternative to the ideal solution and then we can get

$$cc_i = \frac{d_i^{+}}{d_i^{+} + d_i^{-}} (i = 1, 2, \ldots, m) \tag{12}$$

Step 6: Rank the alternatives. According to the closeness coefficient above, the alternative with minimum cc_i. is selected.

4 Application

In the paper, we apply one of the multi-criteria decision making tools, e.g. TOPSIS, to solve 6-Sigma Project selection problem by considering a variety of criteria and sub-criteria under uncertain environment. Alternative 6-Sigma Projects are listed in the following Table 1. There are seven criteria used to evaluate alternative Six Sigma projects presented in Table 1 are as follows: C1: Required time to complete projects; C2: The number of employees affected from the project; C3: Tal cost of a project; C4: Customer satisfaction impact; C5: Improvement of product quality; C6: Participation willingness of the employee to the project; C7: The competitive advantage provided by the project.

Table 1. The alternative projects

#	Projects
1	Process capability improvement
2	Process measurement and control
3	Quality improvement of work environment
4	Standardization of tasks and preparation of hand books
5	Work study applications
6	Improving communication channels via internal and external customers

The evaluations are gathered from 3 decision makers from the company. In Tables 2 and 3, evaluation of criteria by decision makers and decision matrix are illustrated.

Table 2. Linguistic evaluations of criteria by decision makers

Decision makers (DM)	C1	C2	C3	C4	C5	C6	C7
DM1	VG	P	VG	G	M	M	G
DM2	G	P	VG	VG	M	P	M
DM3	VG	VP	M	G	M	M	G

Table 3. Decision matrix by decision makers

Alternatives	DM	C1	C2	C3	C4	C5	C6	C7
A1	DM1	VP	VP	P	G	G	P	M
	DM2	P	P	P	VP	P	VP	M
	DM3	P	G	M	VP	P	VP	P
A2	DM1	VG	M	G	G	VG	M	G
	DM2	M	G	M	M	G	G	M
	DM3	M	G	G	VG	G	M	G
A3	DM1	G	G	VG	M	VG	G	G
	DM2	VG	G	G	G	VG	G	M
	DM3	G	G	VG	G	G	VG	M
A4	DM1	VG	G	VG	VG	G	G	G
	DM2	G	G	VG	VG	G	M	VG
	DM3	VG	VG	G	G	VG	VG	G
A5	DM1	M	P	M	P	P	M	G
	DM2	G	M	P	P	M	G	M
	DM3	P	M	P	P	VP	G	VP
A6	DM1	M	G	G	VG	G	M	M
	DM2	P	M	P	G	G	P	P
	DM3	G	P	M	M	G	P	M

Table 4. Evaluation of decision makers

DM1	DM2	DM3
VG	G	M

In Table 4, evaluation of decision makers are presented. By applying aggregation and normalization formulas, we calculate the weights of decision makers as 0.398, 0.36 and 0.242, respectively. In the analysis, the cost attribute criteria are transformed to benefit attribute criteria. First we convert the linguistic evaluations to neutrosophic evaluations. Then, following up the steps of Neutrosophic TOPSIS we obtain aggregated neutrosophic weights of criteria. Once the weights of criteria are obtained, we continue on the steps of the Neutrosophic TOPSIS approach, neutrosophic relative positive and negative ideal solutions are derived. Then, distance measures (d_i^+, d_i^-) and relative closeness coefficient of each alternative (cc_i) are calculated in Table 5.

Table 5. Aggregated neutrosophic weights of criteria

Alternatives	d_i^+	d_i^-	cc_i
A1	0.571	0.117	0.171
A2	0.303	0.444	0.594
A3	0.301	0.282	0.484
A4	0.289	0.394	0.577
A5	0.222	0.189	0.461
A6	0.176	0.369	0.678

5 Conclusion

The study focuses on multi-criteria Six Sigma project selection decision. The study proposes to apply TOPSIS with interval neutrosophic sets. The results highlight that the best project is A6 "Improving communication channels via internal and external customers". It is followed by A2 "Process Measurement and Control". The last ranked alternative is found as A1 "Process Capability Improvement". For the further studies, other MCDM techniques with neutrosophic sets can be applied and the results can be compared with the ones obtained from this study.

References

1. Smarandache, F.: Neutrosophy/Neutrosophic Probability, Set, and Logic. Am. Res., Waltham (1998)
2. Peng, J.-J., Wang, J.-Q., Zhang, H.Y., Chen, X.-H.: An outranking approach for multi-criteria decision making problems with simplified neutrosophic sets. Appl. Soft Comput. **25**, 336–346 (2014)
3. Peng, J.-J., Wang, J.Q., Wu, X.H., Wang, J., Chen, X.H.: Multi-valued neutrosophic sets and power aggregation operators with their applications in multi-criteria group decision-making problems. Int. J. Comput. Intell. Syst. **8**(2), 345–363 (2015)

4. Tian, Z.-P., Zhang, H.-Y., Wang, J., Wang, J.Q., Chen, X.H.: Multi-criteria decision-making method based on a cross-entropy with interval neutrosophic sets. Int. J. Syst. Sci. **47**(15), 3598–3608 (2015)
5. Bausys, R., Zavadskas, E.K., Kaklauskas, A.: Application of neutrosophic set to multicriteria decision making by COPRAS. Econ. Comput. Econ. Cybern. Stud. Res./Acad. Econ. Stud. **49**(2), 1–15 (2015)
6. Liu, P., Zhang, L.: The extended VIKOR method for multiple criteria decision making problem based on neutrosophic hesitant fuzzy set. Gen. Math., 1–13 (2015)
7. Bausys, R., Zavadskas, E.K.: Multicriteria decision making approach by VIKOR under interval neutrosophic set environment. Econ. Comput. Econ. Cybern. Stud. Res./Acad. Econ. Stud. **49**(4), 33–48 (2015)
8. Wang, Z., Liu, L.: Optimized PROMETHEE based on interval neutrosophic sets for new energy storage alternative selection. Rev. Téc. Ing. Univ. Zulia. **39**(9), 69–77 (2016)
9. Zavadkas, E.K., Bausys, R., Stanujkic, D., Magdalinovic-Kalinovic, M.: Selection of lead-zinc flotation circuit design by applying WASPAS method with single-valued neutrosophic set. Acta Montanistica Slovaca **21**(2), 85–92 (2016)
10. Tian, Z.-P., Wang, J., Wang, J.-Q., Zhang, H.-Y.: Simplified neutrosophic linguistic multi-criteria group decision-making approach to green product development. Group Decis. Negot. **26**(3), 597–627 (2017)
11. Ye, J.: A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. J. Intell. Fuzzy Syst. **26**, 2459–2466 (2014)
12. Broumi, S., Ye, J., Smarandache, F.: An extended TOPSIS method for multiple attribute decision making based on interval neutrosophic uncertain linguistic variables. Neutrosophic Sets Syst. **8**, 22–31 (2015)

Integrated Call Center Performance Measurement Using Hierarchical Intuitionistic Fuzzy Axiomatic Design

Basar Oztaysi(✉), Sezi Cevik Onar, and Cengiz Kahraman

Industrial Engineering Department, Faculty of Management,
Istanbul Technical University, 34367 Istanbul, Turkey
oztaysib@itu.edu.tr

Abstract. Measurement of performance is an important management process which deals with assessment and evaluation of a particular process or its' outcomes. Performance measurement is used in different managerial levels for different purposes. While top management, use it to evaluate the results and construct new goals, at the personal level, performance measurement is good for recognising the current weaknesses and motivating for the future accomplishments. For a particular process, team or individual, first critical performance indicators (KPI) are determined, then targets for each KPI is set at the beginning of the period. At end of the assessment period, performance assessment is done for each KPI and the overall performance is calculated. When subjective and qualitative KPIs are used the overall performance measurement has the possibility to be affected by the evaluator. In this study, a performance measurement model for Call Centers are proposed. In the proposed approach hierarchical intuitionistic fuzzy axiomatic de-sign is used to calculate overall performance.

Keywords: Axiomatic design · Intuitionistic fuzzy sets · Call center · Performance measurement

1 Introduction

Performance measurement (PM) focus on analyzing the accomplishment level of an action or its outcome and it involves collecting, analyzing and reporting information regarding the action, person, team or organization (Upadhaya et al. 2014). PM can be used in various levels of a company, at the personal level it helps identifying personal deficiencies and encourage for future actions. At the mid level, PM helps analyzing the outcome of teams or departments, and top management use PM to assess the results of the past activities and set new directions and goals (Meyer 2002).

Because of immense competition, customer service has widely improved. Today companies provide service to their customer via various channels such as, web, mobile, call-center and off-line stores. Among the other channels, call-centers are the most commonly used one. Gartner (2017) define call center as a computer-based organization that delivers call and contact routing for inbound and outbound telephony transactions. Companies from various sectors, either have their own call center or they get it

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_9

as a service from another company. In all cases there is a need to monitor and manage the call center using an objective PM system.

After a literature survey, Oztaysi and Ucal (2009) define six main requirements of a performance measurement technique. These are; capability to reveal significant numerical results to express the global performance, capability to reveal the performance of lower levels, or different viewpoints, capability to trace the changes in performance by time, flexibility modify according to changing situations, and capability to give an understanding about future performance. Multicriteria decision making (MCDM) techniques can be used in performance measurement systems since they can fulfill the above mentioned requirements.

Zadeh (1965) propose fuzzy sets theory to handle uncertainty and imprecision that naturally exist in real world problems. Since then, fuzzy sets are widely used in the literature to provide more realistic and accurate results. Fuzzy sets are capable of representation of linguistic terms better than crisp numbers but due to some critsms and limitations extensions of fuzzy sets are proposed in the literature. Among these extension, type2 fuzzy sets, hesitant fuzzy sets, intuitionistic fuzzy sets, fuzzy multisets are the most commonly used ones. In this study, a performance measurement framework utilizing interval-valued intuitionistic fuzzy sets with axiomatic design is proposed. By integrating interval-valued intuitionistic fuzzy sets decision makers' linguistic evaluations can be better represented so more accurate results are reached.

In the literature there are studies on performance measurement in various industries such as tourism, energy, supply chain, manufacturing (Huang and Coelho 2017; Ke et al. 2017; Cevik Onar et al. 2014; Felício and Rodrigues 2015; Oztaysi and Surer 2014; Oztaysi et al. 2011). There are also studies in the literature that directly focus on measuring overall call-center performance. In one of these studies, Baraka et al. (2015) define *success* and *gap* indices to calculate the performance of the call center. Baraka et al. (2013) introduce a performance evaluation model for call centers based on Information Systems success model using six main criteria namely, system quality, information quality, service quality, usage, user satisfaction, net benefits. In another study, Ma et al. (2011) propose a performance evaluation methodology for call centers at the level of customer-agent interactions. Klement and Snášel (2011) propose a performance measurement approach using Kohonen Self-Organising Map (SOM) algorithm and the Growing Grid algorithm for identifying anomalies.

The originality of the paper comes from two main points. To the best to authors' knowledge this is the first time to apply axiomatic design approach in performance measurement domain, and also this is the initial study to apply a hierarchical decision model into intuitionistic fuzzy axiomatic design approach. The rest of the paper is as follows: Sect. 2 focus on methodology, thus first basics of intuitionistic fuzzy sets are introduced, then information axiom is explained and steps of the methodology is given. Section 3 a case study from call center performance measurement is presented. To this end first the decision model and performance indicators are introduced and then the steps of the methodology is given. In the last section, Sect. 4, the results are discussed and suggestions on future studies are listed.

2 Methodology

2.1 Intuitionistic Fuzzy Sets

Intuitionistic fuzzy sets developed by Atanasov (1986) have been widely used in the literature. Intuitionistic fuzzy sets consider non-membership value in addition to classical membership definition of fuzzy sets. The sum of membership and non-membership cannot exceed the 1.

Let $X \neq \varnothing$ be a given set. An intuitionistic fuzzy set in X is an object A given by

$$\tilde{A} = \left\{ \left\langle x, \mu_{\tilde{A}}(x), v_{\tilde{A}}(x) \right\rangle ; x \in X \right\}, \tag{1}$$

where $\mu_{\underset{A}{\sim}} : X \to [0, 1]$ and $v_{\underset{A}{\sim}} : X \to [0, 1]$ satisfy the condition

$$0 \leq \mu_{\underset{A}{\sim}}(x) + v_{\underset{A}{\sim}}(x) \leq 1, \tag{2}$$

for every $x \in X$. Hesitancy is equal to "$1 - \left(\mu_{\underset{A}{\sim}}(x) + v_{\underset{A}{\sim}}(x) \right)$"

A Triangular Intuitionistic Fuzzy Number, $\widetilde{A}$ is an intuitionistic fuzzy subset in $\mathbb{R}$ with following membership function and non-membership function:

$$\mu_{\underset{A}{\sim}}(x) = \begin{cases} \frac{x-l}{m-l}, & for\ l \leq x \leq m \\ \frac{r-x}{r-m}, & for\ m \leq x \leq r \\ 0, & otherwise \end{cases} \tag{3}$$

and

$$v_{\underset{A}{\sim}}(x) = \begin{cases} \frac{m-x}{m-l'}, & for\ l' \leq x \leq m \\ \frac{x-m}{r'-m}, & for\ m \leq x \leq r' \\ 1, & otherwise \end{cases} \tag{4}$$

where $l' \leq 1 \leq m \leq \mathrm{r} \leq r', 0 \leq \mu_{\underset{A}{\sim}}(x) + v_{\underset{A}{\sim}}(x) \leq 1$ and TIFN is denoted by $\widetilde{A}_{TIFN} = (l, \mathrm{m}, \mathrm{r}; \mathrm{l}', \mathrm{m}', \mathrm{r}')$ (see Fig. 1).

For the defuzzification of TIFN, Kahraman et al. (2017) propose Eq. 5 by considering the heights of the membership functions. Let A be an TIFS, N Considering the intersection function $\tilde{A}_{\mathrm{TIFS}} = (l, (\mathrm{m}, \mu_{\cup}(\mathrm{m})), \mathrm{r}; \mathrm{l}', (\mathrm{m}', \mathrm{v}_{\cup}(\mathrm{m}')), \mathrm{r}')$

$$\overline{x_{\mathrm{t}}} = \frac{(1 - v_{\cup}(\mathrm{m}')) \times (l' + m' + r') + \mu_{\cap}(\mathrm{m}) \times (l + m + r)}{6} \tag{5}$$

2.2 Information Axiom

The Information Axiom is used to select the best design among the alternative that satisfy the Independence Axiom (Suh 2001). It is represent with the information content, I_i, which is a function of the probability of satisfying the given design

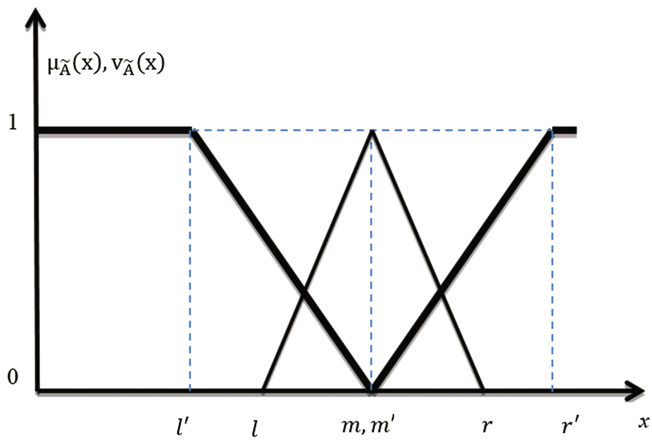


Fig. 1. Membership and non-membership functions of TIFN

requirements. Suh (1990) defines the information content for a given design requirement as follows (Eq. 6):

$$\mathrm{I}_i = \log_2\left(\frac{1}{p_i}\right) \tag{6}$$

where p represents probability of achieving the design requirements. The value $\mathrm{I_i}$ shows that the best alternative is the one with the highest probability of success. When the design fulfills the design requirement for sure, then information content is calculated as zero, on the other hand, when there is no chance to satisfy the design requirements information content becomes infinite.

The probability of fulfilling design requirements can be defined as the overlap between the design range determined by the designer and alternatives' system capability range this overlapped region is called *common range* and the probability to satisfy the requirements can be defined as Eq. (7).

$$\mathrm{p}_i = \left(\frac{Common\ range}{System\ range}\right) \tag{7}$$

Kulak and Kahraman (2005) suggest applying fuzzy information axiom approach into multi criteria decision making in order to use information axiom under certainty or incomplete information. In the fuzzy extension of information axiom, the system and design range for a certain criterion can be expressed by using linguistic terms (Kulak et al. 2010). In general, triangular fuzzy numbers (TFN) or trapezoidal fuzzy numbers are used to mathematically represent these linguistic terms. Graphical representation of *system range* and *common area* when triangular fuzzy numbers are used is given in Fig. 2.

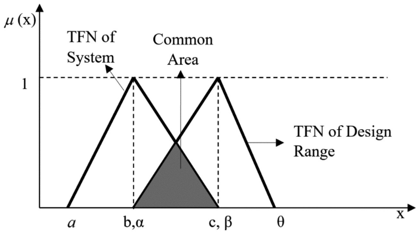


Fig. 2. The common area of system and design ranges (Kulak and Kahraman 2005)

Therefore, when TFNs are used the information content is calculated using Eq. 8.

$$I = \log_2\left(\frac{TFN\ of\ System\ Design}{Common\ Area}\right) \tag{8}$$

While the main formula remains the same, definition of System design and Common area change when TIFNs are used. Figure Y presents intuitionistic fuzzy common area of system range ($\widetilde{SR}$) and design range ($\widetilde{DR}$). It is important to underline that in order to calculate the common area, new functions representing $1 - v$ are formed.

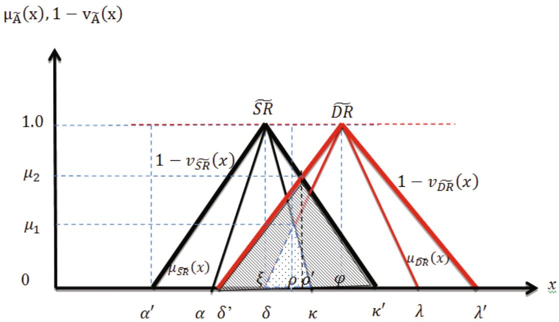


Fig. 3. The common area of system and design ranges (TIFN)

In Fig. 3, $\widetilde{SR}$ represents system range and $\widetilde{DR}$ represent design range. Kahraman et al. (2017) propose using Eq. 9 to find the defuzzified system range

$$DSR = \frac{\alpha + \alpha' + 2\xi + \kappa + \kappa'}{6} \tag{9}$$

On the other hand area covered with dots represent common area of membership values and area covered with lines represent common area of the non-membership values. The defuzzified common area (DCA) can be calculated by using Eq. 10.

$$DCA = \frac{\mu_2 \times (\delta' + \rho' + \kappa') + \mu_1 \times (\delta + \rho + \kappa)}{6} \tag{10}$$

2.3 Hierarchical Intuitionistic Fuzzy Axiomatic Design

In the literature, there are various studies which use information axiom for multicriteria decision making (Cebi et al. 2016; Chen et al. 2015; Kaya et al. 2012; Akay et al. 2011). In this study, Hierarchical Intuitionistic Fuzzy Axiomatic Design is proposed based on the study of Kahraman et al. (2017). The steps of methodology proposed for call center performance measurement is given.

Step 1: Decision model is formed. This step refers to determining performance criteria and performance indicators.
Step 2: Weights of performance criteria and indicators are determined according to the decision model.
Step 3: Values of performance indicators are obtained.
Step 4: Design ranges are defined for each performance indicator by the decision makers.
Step 5: Information content for each indicator is calculated.
Step 6: Weighted information content values are summed up to obtain overall weighted information content (OWIC).
Step 7: The alternative with the lowest (OWIC) is selected as the best alternative.

3 Application

A call center company wants to measure the overall performance of its operations. To this end, a decision-making team is formed to select performance indicators, form decision model and decide the weights of each performance indicator. Different from regular MCDM applications, in this problem we do not have alternatives instead performance measurement values from different time periods are compared.

3.1 Decision Model

One of the most important part of performance measurement is to select performance indicators. To this end meetings with decision making team are made and literature survey is conducted. As a result, for main performance criteria and twelve performance indicators are determined (Fig. 4).

The indicators that takes place in the decision model is briefly explained in the following:

- Service level (C11): a percentage of calls received by the call center that are answered by an agent within a given period.
- Average speed to answer (C12): the average time it takes for calls to be answered in the call center.

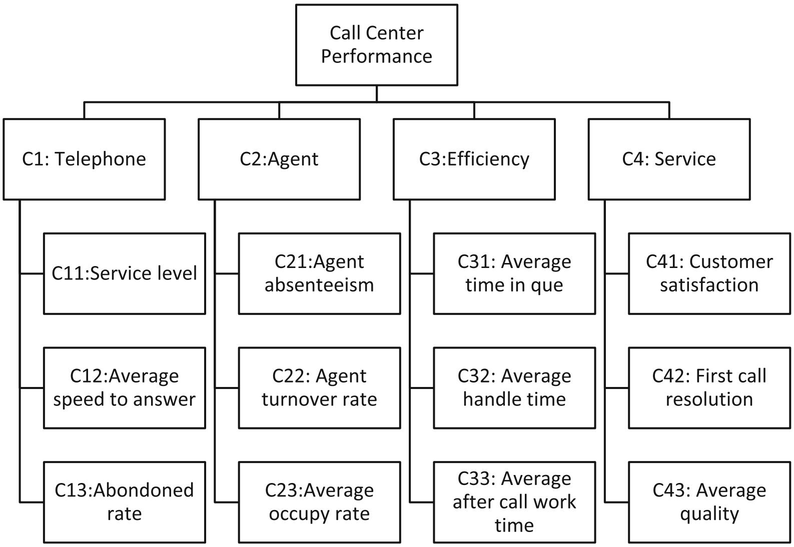


Fig. 4. Performance measurement criteria and indicators.

- Abondoned rate (C13): the percentage of callers who leave the que before reaching an agent.
- Agent absenteeism (C21): the percentage of number of days not served due to agents being absent to the total number of working days.
- Agent turnover rate (C22): the percentage of agents who leave the call center to work elsewhere to the number of all agents.
- Average occupy rate (C23): Average occupancy rate is the amount of time agents spend on live calls and after call work associated with the calls to the total working time.
- Average time in que (C31): the is ratio of the total time customers wait in queues to the total number of calls answered by the call center.
- Average handle time (C32): the average the elapsed time from the moment agent answers a call until the agent disconnects.
- Average after call work time (C33): is the average of time agents spend on after call Works such as updating databases, writing reports etc.
- Customer satisfaction (C41): an average score obtained from customer surveys.
- First call resolution (C42): the percentage of calls that the agent completely addresses the caller's needs without having to transfer, escalate or return the call.
- Average quality (C43): is the average of agents' quality scores. In call centers, quality assurance teams measure the quality of calls based on a set of rules. While this score is important for agents' individual performance, the average quality score is an important indicator for overall performance.

3.2 Call Center Performance Measurement

In this sub section, application of the performance measurement system is explained. The first step of the proposed methodology is to build the decision model and it is already explained in the previous sub-section.

Next step is to determine the weights of the criteria and indicators. In order to calculate the weights Intuitionistic fuzzy AHP method (Cevik Onar et al. 2015) is used. Because of page limitations here we only give the main inputs and outputs of the methodology. According to the methodology the decision makers make pairwise comparison matrics using linguistic scale given in Table 1.

Table 1. Linguistic scale used for Intuitionistic fuzzy AHP

Linguistic variable	Fuzzy representation
Absolutely Low (AL)	([0, 0.2],[0.5, 0.8])
Very Low (VL)	([0.1, 0.3],[0.4, 0.7])
Low (L)	([0.2, 0.4],[0.3, 0.6])
Medium Low (ML)	([0.3, 0.5],[0.2, 0.5])
Equal (E)	([0.4, 0.6],[0.2, 0.4])
Medium High (MH)	([0.5, 0.7],[0.1, 0.3])
High (H)	([0.6,0.8],[0, 0.2])
Very High (VH)	([0.7,0.9],[0,0.1])
Absolutely High (AH)	([0.8,1.0],[0,0])
Exactly Equal	([0.5,0.5],[0.5,0.5])

The pairwise comparison matrices are given in Tables 2 and 3.

Table 2. Pairwise comparison of the criteria

Criteria	C1	C2	C3	C4	Weights
C1	EE	MH	MH	VH	0.345
C2		EE	EH	MH	0.268
C3			EE	MH	0.244
C4				EE	0.143

Table 3. Pairwise comparison of the performance indicators with respect to criteria

w.r.t C1	C11	C12	C13	Weights	w.r.t C3	C31	C32	C33	Weights
C11	EE	H	VH	0.454	C11	EE	MH	MH	0.454
C12		EE	MH	0.334	C12		EE	EE	0.334
C13			EE	0.212	C13			EE	0.212
w.r.t C2	C21	C22	C23	Weights	w.r.t C4	C41	C42	C43	Weights
C21	EE	ML	EE	0.336	C21	EE	MH	E	0.401
C22		EE	MH	0.410	C22		EE	MH	0.358
C23			EE	0.255	C23			EE	0.241

Table 3 shows four different pairwise comparison matrices containing linguistic evaluations of the indicators with respect to the related criteria.

As mentioned before, here we do not give the detailed calculations; instead, local weights are given in both Tables 2 and 3. In order to find the weight of each performance indicator on overall performance, the local weights are multiplied by the weight of the upper level criterion. Table 4 presents the global weights of performance indicators.

Table 4. Global weights of the performance indicators

Indicators	Overall weights	Indicators	Overall weights
C11	0.157	C31	0.11
C12	0.116	C32	0.083
C13	0.074	C33	0.053
C21	0.092	C41	0.057
C22	0.101	C42	0.052
C23	0.068	C43	0.037

After indicator weights are calculated, next step is to determine performance values and design range values. In this study, actual performance values of each indicator is obtained by expert opinions and linguistic terms. To this end, three experts evaluated the performance of the period using linguistic terms, later these terms are transformed into TFIN and aggregated (Kahraman et al. 2017). The aggregated TIFNs are shown in Table 5.

Table 5. Performance values of the indicator

Perf. Indicator	Period 1	Period 2
C11	(0.06,0.41,0.56:0,0.67,0.64)	(0.32,0.52,0.79:0,0.7,1)
C12	(0.49,0.73,0.81:0,0.91,1)	(0.24,0.46,0.78:0,0.95,1)
C13	(0.42,0.67,0.94:0.28,0.41,1)	(0.56,0.89,1:0.44,0.25,1)
C21	(0.35,0.59,0.9:0.21,0.59,1)	(0.46,0.7,0.97:0.33,0.25,1)
C22	(0.12,0.36,0.63:0,0.7,0.7)	(0.21,0.45,0.71:0,0.67,0.79)
C23	(0.38,0.62,0.91:0.23,0.97,1)	(0.4,0.64,0.93:0.26,0.95,1)
C31	(0.42,0.67,0.94:0.28,0.67,1)	(0.68,0.95,1:0.55,0.67,1)
C32	(0.12,0.36,0.63:0,0.89,0.7)	(0.4,0.64,0.93:0.26,0.95,1)
C33	(0.12,0,0.6:0,0.45,0.63)	(0.46,0.7,0.97:0.33,0.3,1)
C41	(0.42,0.67,0.94:0.28,0.67,1)	(0.38,0.62,0.91:0.23,0.59,1)
C42	(0.06,0.3,0.56:0,0.95,0.64)	(0.25,0.5,0.75:0.15,0.67,0.85)
C43	(0.46,0.7,0.97:0.33,0.97,1)	(0.63,0.92,1:0.5,0.89,1)

Since same linguistic scale are used to assess all performance indicators, the design range is defined as (0.7,0.8,0.9;0.6,0.8,1) for all indicators.

For each period and performance indicator, the fuzzy information content is calculated using Eqs. 8, 9 and 10. The results are presented in Table 6.

Table 6. Information content values for each period

Perf. Indicator	Weights	*I* - Period 1	*I* - Period 2
C11	0.157	2.398	2.225
C12	0.116	0.380	1.766
C13	0.074	0.449	0.841
C21	0.092	0.445	3.212
C22	0.101	1.328	2.411
C23	0.068	0.471	1.428
C31	0.11	0.696	0.594
C32	0.083	1.696	1.015
C33	0.053	5.212	1.509
C41	0.057	0.696	1.865
C42	0.052	1.766	2.411
C43	0.037	0.620	0.535

At the very last step, the weighted sum of the performance indicators are calculated. This operation is done for each criteria to observe criteria performance and also for all indicators to obtain overall performance. The results are shown in Table 7.

Table 7. Criteria and overall performance scores

Criteria	Period 1	Period 2	Percentage difference
Telephone	0.454	0.288	+36.6%
Agent	0.207	0.377	−82.3%
Efficiency	0.278	0.119	+57.2%
Service	0.061	0.144	−134.6%
Total	1.001	0.928	+7.2%

When analyzing information content, one must keep in mind that lower *I* values reflect better performance. So, in overall performance Period 2 is better than Period 1. In a similar way, in *Telephone* and *Efficiency* criteria Period 2 outperforms Period 1, however in Agent and Service it is the other way around. According to Table 7, there is nearly 7% increase in overall performance. However, this percentage may be misleading since it is calculated by a logarithm function.

4 Conclusion

In this study, hierarchical intuitionistic fuzzy axiomatic design method is proposed based on Kahraman et al. (2017)'s study. In the proposed method allows a hierarchical decision model to be defined and used in decision-making problems. In the case study, a three level decision model, containing 4 criteria and 12 performance indicators, is defined with and used to measure call center performance.

The results of the study show that, hierarchical intuitionistic fuzzy axiomatic design method can be an effective approach to performance measurement literature since it allows linguistic variables to be represented and used. On the other hand, using information axiom on performance measurement domain has some shortcomings. Although the results can be used to sort or identify best alternative, it is hard to interpret the results to see the degree of improvement. In the future studies, new definitions can be made or formulas can be suggested to maintain more effective results.

References

Akay, D., Kulak, O., Henson, B.: Conceptual design evaluation using interval type-2 fuzzy information axiom. Comput. Ind. **62**(2), 138–146 (2011)

Atanasov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)

Baraka, H.A., Baraka, H.A., El-Gamily, I.H.: Assessing call centers' success: a validation of the DeLone and Mclean model for information system. Egypt. Inform. J. **14**(2), 99–108 (2013)

Baraka, H.A., Baraka, H.A., El-Gamily, I.H.: Information systems performance evaluation, introducing a two-level technique: case study call centers. Egypt. Inform. J. **16**(1), 9–22 (2015)

Cebi, S., Ilbahar, E., Atasoy, A.: A fuzzy information axiom based method to determine the optimal location for a biomass power plant: a case study in Aegean Region of Turkey. Energy **116**(1), 894–907 (2016)

Cevik Onar, S., Oztaysi, S., Otay, I., Kahraman, C.: Multi-expert wind energy technology selection using interval-valued intuitionistic fuzzy sets. Energy **90**(1), 274–285 (2015)

Cevik Onar, S., Oztaysi, B., Kahraman, C.: Dynamic intuitionistic fuzzy multi-attribute aftersales performance evaluation. In: Complex & Intelligent Systems, pp. 1–8 (2017). doi:10.1007/s40747-017-0047-7

Chen, D., Chu, X., Sun, X., Li, Y., Su, Y.: An information axiom based decision making approach under hybrid uncertain environments. Inf. Sci. **312**, 25–39 (2015)

Felício, J.A., Rodrigues, R.: Organizational factors and customers' motivation effect on insurance companies' performance. J. Bus. Res. **68**(7), 1622–1629 (2015)

Gartner: Gartner IT Glossary (2017). http://www.gartner.com/it-glossary/contact-center-systems. Accessed 10 Jan 2017

Huang, Y., Coelho, V.R.: Sustainability performance assessment focusing on coral reef protection by the tourism industry in the Coral Triangle region. Tour. Manag. **59**, 510–527 (2017)

Kahraman, C., Cevik Onar, S., Cebi, S., Oztaysi, B.: Extension of information axiom from ordinary to intuitionistic fuzzy sets: an application to search algorithm selection. Comput. Ind. Eng. **105**, 348–361 (2017)

Kaya, I., Oztaysi, B., Kahraman, C.: A two-phased fuzzy multicriteria selection among public transportation investments for policy-making and risk governance. Int. J. Unc. Fuzz. Knowl. Based. Syst. **20**, 31–48 (2012)

Ke, M.T., Yeh, C.H., Su, C.J.: Cloud computing platform for real-time measurement and verification of energy performance. Appl. Energy **188**, 497–507 (2017)

Klement, P., Snášel, V.: Using SOM in the performance monitoring of the emergency call-taking system. Simul. Model. Pract. Theory **19**(1), 98–109 (2011)

Kulak, O., Kahraman, C.: Fuzzy multi-attribute selection among transportation companies using axiomatic design and analytic hierarchy process. Inf. Sci. **170**, 191–210 (2005)

Kulak, O., Cebi, S., Kahraman, C.: Applications of axiomatic design principles: a literature review. Exp. Syst. Appl. **37**(9), 6705–6717 (2010)

Ma, J., Kim, N., Rothrock, L.: Performance assessment in an interactive call center workforce simulation. Simul. Model. Pract. Theory **19**(1), 227–238 (2011)

Meyer, M.W.: Rethinking Performance Measurement: Beyond the Balanced Scorecard. Cambridge University Press, New York (2002)

Oztaysi, B., Kaya, T., Kahraman, C.: Performance comparison based on customer relationship management using analytic network process. Exp. Syst. Appl. **38**, 9788–9798 (2011)

Oztaysi, B., Ucal, I.: Comparison of MADM techniques usage in performance measurement. In: Proceedings of 10th Annual International Symposium on the Analytical Hierarchy Process (2009)

Oztaysi, B., Surer, O.: Supply chain performance measurement using a SCOR based fuzzy VIKOR approach. In: Kahraman, C., Oztaysi, B. (eds.) Supply Chain Management under Fuzziness, pp. 199–224. Springer, Berlin (2014)

Suh, N.P.: Axiomatic Design: Advances and Applications. Oxford University Press, New York (2001)

Suh, N.P.: The Principles of Design. Oxford University Press, New York (1990)

Upadhaya, B., Munir, R., Blount, Y.: Association between performance measurement systems and organisational effectiveness. Int. J. Oper. Prod. Manag. **34**(7), 2 (2014)

Zadeh, L.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)

Prioritization of Business Analytics Projects Using Interval Type-2 Fuzzy AHP

Basar Oztaysi(✉), Sezi Cevik Onar, and Cengiz Kahraman

Faculty of Management, Industrial Engineering Department,
Istanbul Technical University, 34367 Istanbul, Turkey
oztaysib@itu.edu.tr

Abstract. Because of emerging technologies, a vast amount of data can be stored and processed very easily. These advances also affect companies and many new projects are being proposed. Business analytics is the umbrella term for these projects and it denotes to the skills, technologies, activities aiming at assessment and exploration of past performance to gain an understanding for better decision making. Data and analytical models are the two main pillars of business analytics. Business analytics project can be grouped into three main groups: (i) descriptive analytics, efforts to understand what has happened in the company, (ii) predictive analytics, efforts to figure out the result of an future event, and (iii) prescriptive analytics use mathematical and computational sciences to suggest decision options to take advantage of the results of descriptive and predictive analytics. In this study a prioritization method for possible business analytics projects using Type-2 fuzzy AHP is proposed. Proposed model is composed of six criteria namely, strategic value, competitiveness, customer relations, improved decision-making, improved operations, and data quality.

Keywords: Type-2 fuzzy AHP · Interval type-2 fuzzy sets · Business analytics · Project selection · Multicriteria decision making

1 Introduction

As a result of emerging technologies, both storing and processing large dataset have become cheap and feasible. This trend enabled applications of business analytics to flourish and become important. In its broadest definition, Business analytics (BA) is "evidence-based problem recognition and solving that happen within the context of business situations" (Holsapple et al. 2014). There are various business areas for BA applications, web analytics, marketing analytics, service analytics, talent analytics, process analytics, and risk analytics are among the most popular domains. BA applications are divided into three main groups; descriptive, predictive and prescriptive analytics. (Kiron et al. 2011). Descriptive analytics focus on summarize raw data and make it something that is interpretable by decision makes, thus it helps describing what has happened in the past. Predictive analytics aim at providing companies with actionable insights based on current data and analytical models. The last group, prescriptive analytics aim recommend solutions/actions to the end user.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_10

According to Gartner's 2016 Chief Information Officer (CIO) Agenda Report, Business intelligence and Analytics is the top priority of CIOs, which is followed by cloud technology and mobile (Gartner 2016). This brings the problem of project prioritization and selection since BA projects require considerable financial investment and has potential risks and benefits. BA project selection process can be formulated as a multi criteria decision making (MCDM) which focuses on problems with discrete decision space and predetermined decision alternatives. MCDM approach brings the flexibility to handle various different and potentially conflicting criteria in the decision model. MCDM techniques are applied to various project selection problem in the literature such as; energy (Read et al. 2017; Stojcetovic et al. 2016), urban planning (Wey and Wu 2008; Oztaysi et al. 2016), information systems (Rouhani 2017; Oztaysi 2015), six-sigma (Adebanjo et al. 2016; Ortz et al. 2015), research and development (Oztaysi et al. 2017; Morton et al. 2016), construction (Mousavi et al. 2015).

In the classical MCDM methods, decision makers' evaluations are characterized by numerical numbers. However, in real world applications, using crisp numbers can be impossible. The data may be imprecise by nature, or the decision makers may have problems assigning numerical values to their assessments. Fuzzy set theory developed by Zadeh (1965) can be a good solution to overcome this problem since it provides formalized tools for dealing with the imprecision in decision-making problems. Although fuzzy sets present a better solution than crisp number still some shortcomings are reported. Type-2 fuzzy sets are proposed by Zadeh (1975) in order to better represent imprecision and uncertainty. Type-2 fuzzy sets are extensions of ordinary fuzzy sets, and they model vagueness and linguistic uncertainties since the membership grades, themselves are ordinary fuzzy sets (Mendel 2000; Karnik and Mendel 2001).

The originality of this paper comes from using interval type-2 fuzzy sets into project selection problem for the first time. Besides, to the best of our knowledge this is the first paper which deals with prioritization of business analytics projects and define criteria for this purpose. The rest of the paper is as follows. Section 2 focus on methodology, first basics of interval type-2 fuzzy sets are given and then the steps of interval type-2 fuzzy AHP is explained. Section 3 provides the application information, first the background information about the problem is given, then the decision model is introduced and finally the results of the methodology is presented. In the last section, Sect. 4, the results are discussed and suggestions on future studies are listed.

2 Methodology

In this section, first preliminaries of interval Type-2 fuzzy sets are given with their arithmetic operations. Then the steps of interval Type-2 fuzzy AHP are explained.

2.1 Interval Type-2 Fuzzy Sets

Zadeh (1965) propose fuzzy set theory in order to mathematically represent uncertainty and define mechanisms to handle imprecision and vagueness inherent to real world cases. In the scope of decision process, fuzzy sets allow methodologies to use approximate information instead of crisp values. Using approximate information in problems relaxed the problems and many engineering and decision problems are simplified and enhanced. As a result, classical methods which are originally proposed by using crisp numbers are renewed to enable fuzzy sets to be used. In decision making area, multicriteria decision making (MCDM) techniques are also updated to utilize linguistic terms especially for representing expert evaluations. Using fuzzy sets, linguistic variables could be better represented in the problems. Literature provides studies with different types fuzzy sets such as, triangular fuzzy numbers, trapezoidal fuzzy numbers and Gaussian membership functions (Kahraman and Kaya 2010).

The ordinary fuzzy sets, proposed by Zadeh (1965), suggests that each element in the universe has a degree of membership to a set. This membership is represented by a function called membership function. This is the main property which enable ordinary fuzzy set to handle uncertainties. However, some limitations of ordinary fuzzy sets are reported in the literature such as usage of words, difficulties in aggregation of expert's opinions and working with noisy data (Mendel et al. 2006). In order to overcome this limitation, Zadeh (1975) propose type-2 fuzzy sets. Type-2 fuzzy sets are the fuzzy sets which have membership functions in form of ordinary fuzzy sets. From another perspective, type-2 fuzzy sets brings a new third dimension into membership function. This new dimension brings additional degrees of freedom so that uncertainties can be better modeled. (Mendel et al. 2006).

A type-2 fuzzy sets $\tilde{\tilde{A}}$ in the universe of discourse X can be represented by a type-2 membership function $\mu_{\tilde{\tilde{A}}}(x, \mathrm{u})$, where $x \in \mathrm{X}$ and $u \in \mathrm{J_x} \subseteq [0, 1]$ as follows (Zadeh 1975):

$$\tilde{\tilde{A}} = \left\{ \left((\mathrm{x}, \mathrm{u}), \mu_{\tilde{\tilde{A}}}(\mathrm{x}, \mathrm{u}) \right) | \forall \mathrm{x} \in \mathrm{X}, \forall \mathrm{u} \in \mathrm{J_x} \subseteq [0, 1], 0 \leq \mu_{\tilde{\tilde{A}}}(\mathrm{x}, \mathrm{u}) \leq 1 \right\}, \tag{1}$$

where $\mathrm{J_x}$ denotes an interval [0, 1]. The type-2 fuzzy set $\tilde{\tilde{A}}$ also can be represented as follows (Mendel et al. 2006):

$$\tilde{\tilde{A}} = \int_{\mathrm{x} \in \mathrm{X}} \int_{\mathrm{u} \in \mathrm{J_x}} \mu_{\tilde{\tilde{A}}}(\mathrm{x}, \mathrm{u}) / (\mathrm{x}, \mathrm{u}) \qquad \mathrm{J_x} \subseteq [0, 1] \tag{2}$$

where $\mathrm{J_x} \subseteq [0, 1]$ and $\iint$ denote union over all admissible x and u.

Interval type-2 fuzzy set are a special case of type fuzzy sets where all $\mu_{\tilde{\tilde{A}}}(\mathrm{x}, \mathrm{u}) = 1$, (Buckley 1985).

Mendel et al. (2006) represent interval type-2 fuzzy set $\tilde{\tilde{A}}$ as follows:

$$\tilde{\tilde{A}} = \int_{\mathrm{x} \in \mathrm{X}} \int_{\mathrm{u} \in \mathrm{J_x}} \frac{1}{(\mathrm{x}, \mathrm{u})}, \tag{3}$$

where $\mathrm{J_x} \subseteq [0, 1]$.

In accordance with the given definitions, a trapezoidal interval type-2 fuzzy set can be represented as

$$\tilde{\tilde{A}}_i = \left(\tilde{U}_i;\tilde{L}_i\right) = \left(\left(u_{i1},u_{i2},u_{i3},u_{i4};H_1\left(\tilde{U}_i\right),H_2\left(\tilde{U}_i\right)\right),\left(l_{i1},l_{i2},l_{i3},l_{i4};H_1\left(\tilde{L}_i\right),H_2\left(\tilde{L}_i\right)\right)\right)$$

where $\tilde{U}_i$ and $\tilde{L}_i$ are ordinary fuzzy sets; $u_{i1},u_{i2},u_{i3},u_{i4},l_{i1},l_{i2},l_{i3}$, and l_{i4} are the references points of the interval type-2 fuzzy set $\tilde{\tilde{A}}_i$, $H_1\left(\tilde{U}_i\right)$; shows the membership value of the element $u_{j(j+1)}$ in the upper trapezoidal membership function $\left(\tilde{U}_i\right)$, $1 \le j \le 2$, $H_j\left(\tilde{L}_i\right)$ denotes the membership value of the element $l_{j(j+1)}$ in the lower trapezoidal membership function $\tilde{L}_i$, $1 \le j \le 2$, $H_1\left(\tilde{A}_i^U\right) \in [0,1]$, $H_2\left(\tilde{U}_i\right) \in [0,1]$, $H_1\left(\tilde{A}_i^L\right) \in [0,1]$, $H_2\left(\tilde{L}_i\right) \in [0,1]$ and $1 \le i \le n$ (Chen and Lee 2010). Figure 1 represents a sample trapezoidal interval type-2 fuzzy set.

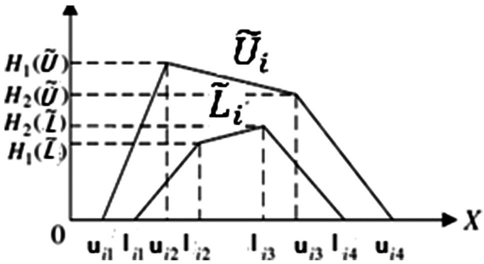


Fig. 1. Interval type-2 fuzzy sets

Let $\tilde{\tilde{A}}_1$ and $\tilde{\tilde{A}}_2$ be interval type-2 fuzzy sets and k be a crisp number,

$$\tilde{\tilde{A}}_1 = \left(\left(u_{11},u_{12},u_{13},u_{14};H_1\left(\tilde{U}_1\right),H_2\left(\tilde{U}_1\right)\right),\left(l_{11},l_{12},l_{13},l_{14};H_1\left(\tilde{L}_1\right),H_2\left(\tilde{L}_1\right)\right)\right)$$

$$\tilde{\tilde{A}}_2 = \left(\left(u_{21},u_{22},u_{23},u_{24};H_1\left(\tilde{U}_2\right),H_2\left(\tilde{U}_2\right)\right),\left(l_{21},l_{22},l_{23},l_{24};H_1\left(\tilde{L}_2\right),H_2\left(\tilde{L}_2\right)\right)\right)$$

the arithmetic operations with these numbers are shown in the following (Chen and Lee 2010).

Addition:

$$\begin{aligned}\tilde{\tilde{A}}_1 \oplus \tilde{\tilde{A}}_2 = &((u_{11}+u_{21},u_{12}+u_{22},u_{13}+u_{23},u_{14}+u_{24};\\ &\min\left(H_1\left(\tilde{U}_1\right);H_1\left(\tilde{U}_2\right)\right),\min\left(H_2\left(\tilde{U}_1\right);H_2\left(\tilde{A}U_2\right)\right)),\\ &(l_{11}+l_{21},l_{12}+l_{22},l_{13}+l_{23},l_{14}+l_{24};\\ &min(H_1(\tilde{L}_1);H_1(\tilde{L}_2)),\ min\,(H_2(\tilde{L}_1);H_2(\tilde{L}_2))))\end{aligned} \tag{4}$$

Subtraction:

$$\begin{aligned}\tilde{\tilde{A}}_1 \Theta \tilde{\tilde{A}}_2 =&((u_{11}-u_{24}, u_{12}-u_{23}, u_{13}-u_{22}, u_{14}-u_{21};\\ &\min(H_1(\tilde{U}_1);H_1(\tilde{U}_2)), \min(H_2(\tilde{U}_1);H_2(\tilde{U}_2))),\\ &(l_{11}-l_{24}, l_{12}-l_{23}, l_{13}-l_{22}, l_{14}-l_{21};\\ &min(H_1(\tilde{L}_1);H_1(\tilde{L}_2)),\ min(H_2(\tilde{L}_1);H_2(\tilde{L}_2))))\end{aligned} \tag{5}$$

Multiplication:

$$\begin{aligned}\tilde{\tilde{A}}_1 \otimes \tilde{\tilde{A}}_2 \cong&((u_{11}\times u_{21}, u_{12}\times u_{22}, u_{13}\times u_{23}, u_{14}\times u_{24};\\ &\min\left(H_1\left(\tilde{U}_1\right);H_1\left(\tilde{U}_2\right)\right), \min\left(H_2\left(\tilde{U}_1\right);H_2\left(\tilde{U}_2\right)\right)),\\ &((l_{11}\times l_{21}, l_{12}\times l_{23}, l_{13}\times l_{23}, l_{14}\times l_{24};\\ &\min\left(H_1\left(\tilde{L}_1\right);H_1\left(\tilde{L}_2\right)\right), \min\left(H_2\left(\tilde{L}_2\right);H_2\left(\tilde{L}_2\right)\right)))\end{aligned} \tag{6}$$

Multiplication with a crisp number:

$$\begin{aligned}k\tilde{\tilde{A}}_1 =&((k\times u_{11}, k\times u_{12}, k\times u_{13}, k\times u_{14}); H_1\left(\tilde{U}_1\right), H_2\left(\tilde{U}_1\right),\\ &\left(k\times l_{11}, k\times l_{12}, k\times l_{13}, k\times l_{14}; H_1\left(\tilde{L}_1\right), H_2\left(\tilde{L}_1\right)\right))\end{aligned} \tag{7}$$

Division by a crisp number:

$$\begin{aligned}\frac{\tilde{\tilde{A}}_1}{k} =&\left(\left(\frac{1}{k}\times u_{11}, \frac{1}{k}\times u_{12}, \frac{1}{k}\times u_{13}, \frac{1}{k}\times u_{14}\right); H_1\left(\tilde{U}_1\right), H_2\left(\tilde{U}_1\right),\\ &\left(\frac{1}{k}\times l_{11}, \frac{1}{k}\times l_{12}, \frac{1}{k}\times l_{13}, \frac{1}{k}\times l_{14}; H_1\left(\tilde{L}_1\right), H_2\left(\tilde{L}_1\right)\right)\right)\end{aligned} \tag{8}$$

where $k > 0$.

Based on these arithmetic operations the interval type-2 fuzzy AHP is introduced in the next section.

2.2 Interval Type-2 Fuzzy AHP

AHP is a method that aim to quantify relative priorities of a given set of alternatives by utilizing pairwise comparisons and decision makers' judgments. Since expert comparisons are the main input of the method, it stresses the consistency of the comparisons to detect inconsistencies. The original scale used for decision maker's evaluations is composed of crisp numbers. However, representing a linguistic term using fuzzy sets provides better results since fuzzy logic provides mathematical tools and operations to handle uncertainties (Kahraman et al. 2010). From this point of view, fuzzy extensions of classical AHP have been proposed.

In the literature there are various studies which use fuzzy AHP with linguistic variables. The first algorithm in fuzzy AHP is form Laarhoven and Pedrycz (1983) which utilizes Lootsma's logarithmic least square method and defines memberships with triangular fuzzy sets. Later, Buckley (1985) uses trapezoidal fuzzy numbers in AHP method and integrates geometric mean method to derive weights. Chang (1996) propose using extent analysis method for the synthetic extent values of the pairwise comparisons and representing pairwise evaluations by triangular fuzzy numbers. In one of the recent studies Zeng et al. (2007), propose using arithmetic averaging method to calculate performance scores with various scales.

In order to handle the uncertainty and vagueness in a better way, type-2 fuzzy sets are integrated into AHP method (Kahraman et al. 2014; Sari et al. 2013). The proposed AHP method integrates trapezoidal interval type-2 fuzzy sets with Buckley's (1985) fuzzy AHP method. The steps of the method are given in the following:

Step 1: The problem is analyzed and the goal is established.

Step 2: The decision model is structured, the top through the intermediate levels by determining the criteria and finally at the lowest level list of alternatives.

Step 3: Pairwise comparison matrices are constructed in accordance with the decision model. Later, the decision makers make comprised evaluations using linguistic variables. In Table 1, the linguistic variables and corresponding trapezoidal interval type-2 fuzzy scales are given (Kahraman et al. 2014).

Table 1. Linguistic variables and fuzzy scales

Linguistic variables	Trapezoidal Internal Type-2 fuzzy scales
Absolutely Strong (AS)	(7, 8, 9, 9; 1, 1) (7.2, 8.2, 8.8, 9; 0.8, 0.8)
Very Strong (VS)	(5, 6, 8, 9;1, 1) (5.2, 6.2, 7.8, 8.8; 0.8, 0.8)
Fairly Strong (FS)	(3, 4, 6, 7; 1, 1) (3.2, 4.2, 5.8, 6.8, 0.8, 0.8)
Slightly Strong (SS)	(1, 2, 4, 5; 1, 1) (1.2, 2.2, 3.8, 4.8; 0.8, 0.8)
Exactly Equal (E)	(1, 1, 1, 1; 1, 1) (1, 1, 1, 1; 1, 1)

The resulting matrice of a pairwise comparison is given in the following;

$$\tilde{\tilde{A}} = \begin{bmatrix} 1 & \tilde{\tilde{a}}_{12} & \cdots & \tilde{\tilde{a}}_{1n} \\ \tilde{\tilde{a}}_{21} & 1 & \cdots & \tilde{\tilde{a}}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\tilde{a}}_{n1} & \tilde{\tilde{a}}_{n2} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & \tilde{\tilde{a}}_{12} & \cdots & \tilde{\tilde{a}}_{1n} \\ 1/\tilde{\tilde{a}}_{12} & 1 & \cdots & \tilde{\tilde{a}}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1\tilde{\tilde{a}}_{\text{ln}} & 1/\tilde{\tilde{a}}_{2n} & \cdots & 1 \end{bmatrix} \tag{9}$$

where

$$1/\tilde{\tilde{a}} = \left(\left(\frac{1}{a_{14}^U}, \frac{1}{a_{13}^U}, \frac{1}{a_{12}^U}, \frac{1}{a_{11}^U}; H_1\left(a_{12}^U\right), H_2\left(a_{13}^U\right) \right), \left(\frac{1}{a_{24}^L}, \frac{1}{a_{23}^L}, \frac{1}{a_{22}^L}, \frac{1}{a_{21}^L}; H_1\left(a_{22}^L\right), H_2\left(a_{23}^L\right) \right) \right)$$

Step 4: The consistency of the pair wise comparisons are surveyed. This is accomplished by defuzzifying the values of matrices and checking the consistencies.

Step 5: Fuzzy weights of each element in the row of the matrix is calculated. In this manner for each row, the geometric mean $\tilde{\tilde{r}}_i$ is calculated using Eq. 10;

$$\tilde{\tilde{r}}_i = \left[\tilde{\tilde{a}}_{i1} \otimes \ldots \otimes \tilde{\tilde{a}}_{ij}\right]^{1/j} \quad (10)$$

Geometric means are normalized using Eq. (11) to obtain fuzzy weights.

$$\tilde{\tilde{p}}_i = \tilde{\tilde{r}}_i \otimes \left[\tilde{\tilde{r}}_1 \oplus \ldots \oplus \tilde{\tilde{r}}_i \oplus \ldots \oplus \tilde{\tilde{r}}_n\right]^{-1} \quad (11)$$

Step 6: Fuzzy performance scores of each alternative is calculated using Eq. (12).

$$\tilde{\tilde{U}}_i = \sum\nolimits_{j=1}^{n} \tilde{\tilde{w}}_j \tilde{\tilde{s}}_j, \forall i. \quad (12)$$

where $\tilde{\tilde{U}}_i$ represents the utility of alternative i, $\tilde{\tilde{w}}_j$ represents the weight of the criterion j, and $\tilde{\tilde{s}}_j$ shows the score of the alternative with respect to criterion j. Recall that $\tilde{\tilde{w}}_j$ and $\tilde{\tilde{s}}_j$ are computed from different pairwise comparison matrices using the same formulas in Step 6. While $\tilde{\tilde{w}}_j$ represents the fuzzy priority of the related pairwise comparison of the criteria, $\tilde{\tilde{s}}_j$ represents the fuzzy priority calculated from the related pairwise comparison of the alternatives with respect to the related criterion.

Step 7: Type-2 interval fuzzy sets are defuzzified in order to determine the importance ranking of the alternatives. The DTtrT method (Kahraman et al. 2014) is used for defuzzification in this step (Eq. 13).

$$DTtrT = \frac{\frac{(u_U - l_U) + (\beta_U . m_{1U} - l_U) + (\alpha_U . m_{2U} - l_U)}{4} + l_U + \left[\frac{(u_L - l_L) + (\beta_L . m_{1L} - l_L) + (\alpha_L . m_{2L} - l_L)}{4} + l_L\right]}{2} \quad (13)$$

Step 8: The best alternative is determined using the defuzzified utility values of the alternatives. The alternative with the highest value is selected.

3 Application

3.1 Background of the Case Study

In this section, a numerical application of interval type-2 fuzzy AHP in a case study is presented. The case study is from a textile manufacturing company, which produces and sells textile products with its own brand. We assume that there are three alternative BA projects for the company. Alternative 1 (Alt. 1) is a project on production visibility. The project promise to capture real time production data, which will support all related planning activities. Using automatic identification technologies managers can trace and analyze all processes and products. Alternative 2 (Alt. 2) is on transportation management. The products are delivered to all retail points and to customers using various transportation routes and carriers. This project aims to optimize routes and shipments to in order to maintain cost reduction. Alternative 3 (Alt. 3) is called customer experience monitoring. The main point of this project is to capture and consolidate customer data through all touch points, including, mobile, web, call-center, phone, fax, and retail store. As a result, the company can form a single view of the customer and build better analytics.

3.2 Decision Model

The criteria used for BA projects prioritization are determined as a result of literature review and comments of decision making team (Laursen and Thorlund 2017, de Araújo et al. 2017). The resulting three level decision making model is constructed as shown in Fig. 2.

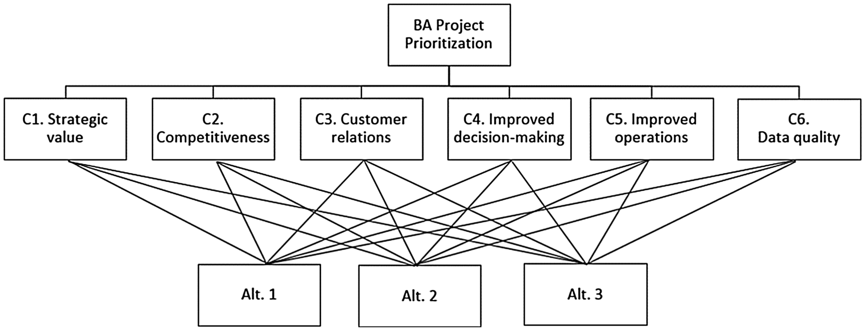


Fig. 2. Hierarchy of BA project prioritization problem

The criteria selected for BA project prioritization are given as follows:

- **Strategic Value (C1)**: A major aspect of BA project evaluation is assessing the relationship with the project and company strategies. The projects, which have a direct link or direct effect on strategic goals, should be prioritized.
- **Competitiveness (C2)**: Competitiveness shows ability and performance of a company, to sell and supply goods and services in a given market, in relation to the performance of other firms. In some cases, projects can directly affect the competitiveness of the company.
- **Customer relations (C3)**: Companies try to manage its interactions with current and potential future customers. Since the number of locations, channels and customers are very high, it is based on the storing and processing customers' data.
- **Improved decision-making (C4)**: The ultimate aim of every BA project is to improve managerial processes by supporting decision-making. However, the contribution of each project may differ. In this criterion, the projects are evaluated with respect to their direct contribution to decision making process.
- **Improved operations (C5):** Companies create value and income by realizing its operations. The performance of operations are measured in terms of efficiency, effectiveness, cost or time. Improving the performance of operations the company gets better results and costs decrease.

- **Data Quality (C6):** Data quality can be defined as fit for its intended uses in operations, decision-making and planning (Redmann 2013). If the data quality is low, the decisions based on this data becomes ineffective and unreliable.

3.3 Prioritization of the Alternatives

This section explains the application steps of interval type-2 fuzzy AHP method defined in Sect. 2. In the first stage of the application the pairwise comparison matrices for criteria and alternatives are constructed. One matrix for comparison of the criteria and six matrices for the comparison of alternatives with respect to each criterion are formed. The decision making team filled the matrices with consensus. Table 2 shows the pairwise comparison of the criteria with respect to the goal. Table 3 shows the pairwise comparison of the alternatives with respect to the criteria

Table 2. The pairwise comparison for the criteria

w.r.t Goal	C1	C2	C3	C4	C5	C6
C1	E	SS	SS	FS	FS	SS
C2	1/SS	E	SS	SS	SS	E
C3	1/SS	1/SS	E	SS	SS	E
C4	1/FS	1/SS	1/SS	E	SS	E
C5	1/FS	1/SS	1/SS	1/SS	E	1/SS
C6	1/SS	E	E	E	SS	E

Table 3. Pairwise comparison matrices with respect to the criteria

w.r.t C1	**ALT.1**	**ALT.2**	**ALT.3**	w.r.t C2	**ALT.1**	**ALT.2**	**ALT.3**
ALT.1	E	SS	E	**AL1/.1**	E	1/SS	E
ALT.2	1/SS	E	E	**AL1/.2**	SS	E	E
ALT.3	E	E	E	**AL1/.3**	E	E	E
w.r.t C3	**ALT.1**	**ALT.2**	**ALT.3**	w.r.t C4	**ALT.1**	**ALT.2**	**ALT.3**
ALT.1	E	E	1/FS	**AL1/.1**	E	E	E
ALT.2	1/E	E	1/W	**AL1/.2**	E	E	SS
ALT.3	FS	W	E	**AL1/.3**	1/E	1/SS	E
w.r.t C5	**ALT.1**	**ALT.2**	**ALT.3**	w.r.t C6	**ALT.1**	**ALT.2**	**ALT.3**
ALT.1	E	1/SS	SS	**AL1/.1**	E	FS	E
ALT.2	SS	E	FS	**AL1/.2**	1/FS	E	1/FS
ALT.3	1/SS	1.FS	E	**AL1/.3**	E	FS	E

Following the steps given in Sect. 3, the fuzzy and defuzzified weights are calculated as given in Table 4. The weights show that the most important criterion is Strategic Value (C1) and the least important criterion is (C5)

Table 4. Interval Type-2 fuzzy weights

Crt.	Fuzzy Weights	Weig.
C1	(0.12, 0.24, 0.61, 1.01; 1, 1) (0.14, 0.27, 0.56, 0.90; 0.8, 0.8)	0.382
C2	(0.067, 0.12, 0.30, 0.52; 1, 1) (0.077, 0.13, 0.27, 0.46; 0.8, 0.8)	0.194
C3	(0.051, 0.087, 0.21, 0.40; 1, 1) (0.058,0.095,0.19,0.34;0.8,0.8)	0.143
C4	(0.037, 0.057, 0.13, 0.25; 1, 1) (0.041,0.062,0.12,0.21;0.8,0.8)	0.092
C5	(0.021, 0.032, 0.085, 0.19; 1, 1) (0.023, 0.035, 0.075, 0.15; 0.8, 0.8)	0.062
C6	(0.066, 0.097, 0.19, 0.30; 1,1) (0.073, 0.10, 0.17, 0.27; 0.8, 0.8)	0.127

Next step is to calculate the weight of the alternatives with respect to each criterion. Later, these weights are multiplied by the weight of the criteria to find the global weights. The global weights are summed up to find the overall priority of the alternative.

Table 5 shows that total defuzzified weights of the weights are 0.529, 0.487, and 0.550 respectively. According to these results, called customer experience monitoring (Alt. 3) has the highest priority and it is followed by production visibility (Alt.1). Transportation management (Alt. 2) has the lowest priority.

Table 5. Global weights of the alternatives

Criteria	Fuzzy Weight of Alt.1	Def. W.
C1	(0.03, 0.092, 0.34, 0.66; 1, 1) (0.043, 0.10, 0.30, 0.57; 08, 08)	0.260
C2	(0.010, 0.023, 0.083, 0.20; 1, 1) (0.012, 0.026, 0.072, 0.16; 0.8, 0.8)	0.072
C3	(0.005, 0.011, 0.042, 0.11; 1, 1) (0.006, 0.012, 0.036, 0.087; 0.8, 0.8)	0.038
C4	(0.009, 0.017, 0.046, 0.099; 1, 1) (0.011, 0.018, 0.041, 0.082; 0.8, 0.8)	0.039
C5	(0.002, 0.005, 0.034, 0.14; 1, 1) (0.002, 0.00, 0.027, 0.10; 0.8, 0.8)	0.040
C6	(0.022, 0.038, 0.099, 0.18; 1, 1) (0.025, 0.042, 0.090, 0.16; 0.8, 0.8)	0.080
Total	**(0.084, 0.18, 0.64, 1.41; 1, 1) (0.10, 0.21, 0.56, 1.17; 0.8, 0.8)**	**0.529**
Criteria	Fuzzy Weight of Alt.2	Def. W.
C1	(0.019, 0.046, 0.17, 0.39; 1, 1) (0.024, 0.052, 0.14, 0.32; 0.8, 0.8)	0.142
C2	(0.017, 0.046, 0.16, 0.34; 1, 1) (0.022, 0.053, 0.14, 0.29; 0.8, 0.8)	0.132
C3	(0.006, 0.01, 0.053, 0.15; 1, 1) (0.007, 0.014, 0.045, 0.12; 0.8, 0.8)	0.051
C4	(0.009, 0.02, 0.074, 0.17; 1, 1) (0.011, 0.024, 0.064, 0.13; 0.8, 0.8)	0.062
C5	(0.005, 0.01, 0.078, 0.27; 1, 1) (0.006, 0.016, 0.064, 0.20; 0.8, 0.8)	0.081
C6	(0.004, 0.007, 0.021, 0.04; 1, 1) (0.004, 0.008, 0.01, 0.039; 0.8, 0.8)	0.018
Total	**(0.063,0.14,0.56,1.39;1,1)(0.077,0.16,0.48,1.11;0.8,0.8)**	**0.487**
Criteria	Fuzzy Weight of Alt.3	Def. W.
C1	(0.03, 0.073, 0.21, 0.39; 1, 1) (0.021, 0.040, 0.094, 0.17; 0.8, 0.8)	0.164
C2	(0.017, 0.03, 0.10, 0.20; 1, 1) (0.021, 0.040, 0.094, 0.17; 0.8, 0.8)	0.084
C3	(0.01, 0.040, 0.19, 0.51; 1, 1) (0.01, 0.047, 0.16, 0.41; 0.8, 0.8)	0.171
C4	(0.005, 0.010, 0.037, 0.099; 1, 1) (0.006, 0.012, 0.031, 0.077; 0.8, 0.8)	0.034
C5	(0.001, 0.002, 0.01, 0.058; 1, 1) (0.001, 0.002, 0.010, 0.040; 0.8, 0.8)	0.016
C6	(0.022, 0.038, 0.099, 0.18; 1, 1) (0.025, 0.042, 0.090, 0.16; 0.8, 0.8)	0.080
Total	**(0.096, 0.20, 0.66, 1.46; 1, 1) (0.11, 0.22, 0.58, 1.20; 0.8, 0.8)**	**0.550**

4 Conclusion

Emerging technologies on data storage and processing has enabled various business analytics projects feasible. As a result, companies need to prioritize and select right BA projects, since the budgets are limited. In this paper BA project selection is modeled as a MCDM problem and interval type-2 fuzzy AHP, is applied in a numerical case study.

In the case study a decision model with six criteria and three alternatives is built. The criteria are determined after a literature survey and with the help of decision making team's comments. As a result of the pairwise comparisons the most important criterion for BA project selection is determined as *strategic value*, which is followed by *competition* and *customer relations*. The results of the study also show that among the three alternative projects, *monitoring customer experience* has the highest priority and it is followed by *production visibility* project.

Both classical AHP and fuzzy AHP methods suggest that they handle linguistic variables into the calculations. In the former method, linguistic variables are handled as crisp numbers and triangular fuzzy numbers are used in the later one. In this study interval type-2 fuzzy sets are used to represent linguistic variables in order to reach more reliable results.

For the future studies, the decision model can be extended to cover the costs as well as benefits. Another way for further studies is to use other MCDM techniques and other fuzzy extensions and compare their results with this study.

References

Gartner: CIO's Agenda (2016) http://www.gartner.com/technology/cio-trends/cio-agenda/?=c. Accessed 16 Apr 2017

Holsapple, C., Lee-Post, A., Pakath, R.: A unified foundation for business analytics. Decis. Support Syst. **64**, 130–141 (2014)

Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., Haydock, M.: Analytics: the widening divide. Sloan Manage. Rev. **53**(3), 1–22 (2011)

Read, L., Madani, K., Mokhtari, S., Hanks, C.: Stakeholder-driven multi-attribute analysis for energy project selection under uncertainty. Energy **119**, 744–753 (2017)

Stojcetovic, B., Nikolic, D., Velinov, V., Bogdanovic, D.: Application of integrated strengths, weaknesses, opportunities, and threats and analytic hierarchy process methodology to renewable energy project selection in Serbia. J. Renew. Sustain. Energ. **8**, 35906 (2016)

Oztaysi, B., Onar, S.C., Goztepe, K.: Kahraman C: Evaluation of research proposals for grant funding using interval-valued intuitionistic fuzzy sets. Soft. Comput. **21**(5), 1203–1218 (2017)

Adebanjo, D., Samaranayake, P., Mafakheri, F., Tritos, L.: Prioritization of Six-Sigma project selection A resource-based view and institutional norms perspective. Benchmarking-An Int. J. **23**(7), 1983–2003 (2016)

Ortz M.A., Felizzola, H.A.; Isaza, S.N.: A contrast between DEMATEL-ANP and ANP methods for six sigma project selection: a case study in healthcare industry, BMC Med. Inform. Decis. Making **15**, Supplement: 3, (2015) Mousavi, S.M., Vahdani, B., Hashemi, H., Ebrahimnejad, S.: An artificial intelligence model-based locally linear neuro-fuzzy for construction project selection. J. Multiple-Valued Logic Soft Comput. **25**(6), 589–604 (2015)

Wey, M.W., Wu, K.Y.: Interdependent Urban renewal project selection under the consideration of resource constraints. Environ. Plann. B: Urban Anal. City Sci. **35**(1), 122–147 (2008)

Oztaysi, B., Cevik Onar, B., Kahraman, C.: Fuzzy multicriteria prioritization of Urban transformation projects for Istanbul. J. Int. Fuzzy Syst. **30**(4), 2459–2474 (2016)

Oztaysi, B.: A group decision making approach using interval type-2 fuzzy AHP for enterprise information systems project selection. J. Multiple-Valued Logic Soft Comput. **24**(5–6), 475–500 (2015)

Rouhani, S.: A fuzzy superiority and inferiority ranking based approach for IT service management software selection. Kybernetes **46**(4), 728–746 (2017)

Morton, A., Keiser, J.M:, Salo, A.: Multicriteria portfolio decision analysis for project selection, multiple criteria decision analysis. In: International Series in Operations Research & Management Science, vol. 233, pp. 1269–1298 (2016)

Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)

Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning – I. Inform. Sci. **8**(3), 199–249 (1975)

Karnik, N.N., Mendel, J.M.: Centroid of a type-2 fuzzy set. Inf. Sci. **132**, 195–220 (2001)

Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**, 233–247 (1985)

Chen, S.M., Lee, L.W.: Fuzzy multiple attributes group decision-making based on the interval type-2 TOPSIS method. Expert Syst. Appl. **37**, 2790–2798 (2010)

Sari, I.U., Oztaysi, B., Kahraman, C.: Fuzzy AHP using type II fuzzy sets: an application to warehouse location selection. In: Doumpos, M., Grigoroudis, E. (eds.) Multicriteria Decision Aid and Artificial Intelligence: Links, Theory and Applications, pp. 258–308. Wiley, Hoboken (2013)

Kahraman, C., Öztayşi, B., Uçal Sarı, I., Turanoğlu, U.: Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. Knowl.-Based Syst. **59**, 48–57 (2014)

Mendel, J.M.: Uncertainty, fuzzy logic, and signal processing. Sign. Process. **80**(6), 913–933 (2000)

Mendel, J.M., John, R.I., Liu, F.L.: Interval type-2 fuzzy logical systems made simple. IEEE Trans. Fuzzy Syst. **14**(6), 808–821 (2006)

Kahraman, C., Kaya, İ.: A fuzzy multicriteria methodology for selection among energy alternatives. Expert Syst. Appl. **37**(9), 6270–6281 (2010)

Chang, D.Y.: Applications of the extent analysis method on fuzzy AHP. Eur. J, Operational Res. **95**, 649–655 (1996)

Zeng, J., Min, A., Smith, N.J.: Application of fuzzy based decision making methodology to construction Project risk assessment. Int. J. Project Manage. **25**, pp. 589–600 (2007)

Laursen, G.H.N., Thorlund, J.: Business Analytics for Managers: Taking Business Intelligence Beyond Reporting. Wiley (2017)

de Araújo, M.C.B., Alencar, L.H., Mota, C.M.M.: Project procurement management: a structured literature review. Int. J. Project Manage. **35**(3), 353–377 (2017)

Redman T.C.: Data quality management past, present, and future: towards a management system for data. In: Sadiq, S. (ed.) Handbook of Data Quality Research and Practice. Springer (2013)

Optimized Fuzzy Transform for Image Compression

Daniel Paternain[1,2(✉)], Aranzazu Jurio[1,2], Javier Ruiz-Aranguren[1], Maria Minárová[3], Zdenko Takáč[3], and Humberto Bustince[1,2]

[1] Departamento de Automatica y Computacion, Universidad Publica de Navarra, Campus Arrosadia sn, 31006 Pamplona, Spain
daniel.paternain@unavarra.es
[2] Institute of Smart Cities, Universidad Publica de Navarra, Campus Arrosadia Sn, 31006 Pamplona, Spain
[3] Slovak Univeristy of Technology in Bratislava, Radlinskeho 9, 813 68 Bratislava, Slovakia

Abstract. In this work we propose an image compression algorithm based on the fuzzy transform. The algorithm tries to find the best fuzzy partition of the functions domain in order to obtain the best compressed image (in terms of quality). To solve the optimization problem we based ourselves in the Gravitational Search Algorithm, in which each agent represents a possible fuzzy partition of a fixed size.

Keywords: Fuzzy transform · Image compression · Gravitational Search Algorithm

1 Introduction

Image compression consists in reducing the amount of data (generally measured as the number of bits) required to represent an image [4]. Generally, image compression algorithms transform or encode the image and this transformation is stored or transmitted. Then, an inverse transformation or decode process is applied and a reconstruction of the original image is obtained.

There exist many image compression algorithms in the literature. In this work, we focus on the fuzzy transform [7] which has been successfully applied in the field of image processing and, more specifically, in image compression (see [1–3,6,8,9]). Broadly speaking, the fuzzy transform is based on a fuzzy partition of the image domain, i.e. $[1, N] \times [1, M]$ where N and M represent, respectively, the number of rows and columns of an image. The fuzzy partition is composed by n and m fuzzy sets defined on the intervals $[1, N]$ and $[1, M]$, respectively, with certain properties. It is known that, given fixed n and m, two different fuzzy partitions of the image domain will yield two different fuzzy transforms, two different compressed images and, accordingly, two different reconstructed images. Since the inverse fuzzy transform is also based on the same fuzzy partition used

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_11

in the compression, it is possible to rank both fuzzy partitions in terms of quality: the better the quality of the reconstructed image, the better the fuzzy partition.

Taking into account this fact, the objective of this work is the following: given an image, to find the best fuzzy partition so that the reconstructed image after applying the fuzzy and the inverse fuzzy transform is as similar as possible to the original image.

In order to find the best fuzzy partition, we propose to use the Gravitational Search Algorithm (GSA), an heuristic optimization algorithm based on the law of gravity and the law of motion [10]. In this algorithm, each agent represents a solution, i.e. a feasible fuzzy partition of the image domain. The quality (fitness) of an specific agent is measured in the following way: taking the fuzzy partition represented by the agent, we apply the fuzzy transform (obtaining the compressed image) and, later, the inverse fuzzy transform. Finally, we measure the quality of the reconstructed image by means of an error measure and we associate this error with the agent. Then, the problem becomes a minimization problem and the GSA tries to find the agent with the minimum error measure.

The structure of this work is as follows. In Sect. 2 we recall the concept of a fuzzy partition and the definition of the fuzzy and inverse fuzzy transform. In Sect. 3 we summarize the steps of the Gravitational Search Algorithm. In Sect. 4 we explain in detail the optimization algorithm we propose in order to find the best fuzzy partition and, in Sect. 5, we show the first preliminary results obtained by our proposal. We finish, in Sect. 6 with some conclusions and future research lines.

2 Fuzzy Transform of Discrete Functions

In this section we recall the concept of the fuzzy transform and the inverse fuzzy transform. For the sake of simplicity, we focus only on the discrete fuzzy transform, that maps a discrete function defined on an interval of real numbers into a real vector.

In order to give the definition of the fuzzy and inverse fuzzy transform, we define the concept of a fuzzy partition of the functions domain.

Definition 1 [7]. *Let $x_1 < \cdots < x_n$ be fixed nodes within $[a, b]$, such that $x_1 = a$, $x_n = b$ and $n \geq 2$. We say that fuzzy sets $A_1, \ldots, A_n$, identified with their membership functions $A_1(x), \ldots, A_n(x)$ defined on $[a, b]$, form a fuzzy partition of $[a, b]$ if they fulfill the following conditions for $k = 1, \ldots, n$:*

(1) $A_k : [a, b] \to [0, 1]$, $A_k(x_k) = 1$;
(2) $A_k(x) = 0$ *if* $x \notin (x_{k-1}, x_{k+1})$ *where* $x_0 = a$ *and* $x_{n+1} = b$;
(3) $A_k(x)$ *is continuous;*
(4) $A_k(x)$, $k = 2, \ldots, n$, *strictly increases on* $[x_{k-1}, x_k]$ *and* $A_k(x)$, $k = 1, \ldots, n-1$, *strictly decreases on* $[x_k, x_{k+1}]$;
(5) *for all* $x \in [a, b]$, $\sum_{k=1}^{n} A_k(x) = 1$.

Definition 2 [7]**.** *Let a fuzzy partition of* $[a, b]$ *be given by fuzzy sets* $A_1, \ldots, A_n$ *in the sense of Definition 1. We say that it is uniform if the nodes* $x_1, \ldots, x_n$, $n \geq 3$, *are equidistant. This means that* $x_k = a + h(k-1)$, $k = 1, \ldots, n$, *where* $h = (b-a)/(n-1)$, *and*

(6) $A_k(x_k - x) = A_k(x_k + x)$, *for all* $x \in [0, h]$, $k = 2, \ldots, n-1$;
(7) $A_k(x) = A_{k-1}(x - h)$, *for all* $k = 2, \ldots, n-1$ *and* $x \in [x_k, x_{k+1}]$, *and* $A_{k+1}(x) = A_k(x - h)$, *for all* $k = 2, \ldots, n-1$ *and* $x \in [x_k, x_{k+1}]$.

Definition 3 [7]**.** *Let* f *be a function given at nodes* $p_1, \ldots, p_l \in [a, b]$ *and* $A_1, \ldots, A_n$, $n < l$, *be basic functions which form a fuzzy partition of* $[a, b]$. *We say that the* n*-tuple of real numbers* $F[f] = (F_1, \ldots, F_n)$ *given by*

$$F_k = \frac{\sum_{j=1}^{l} f(p_j) A_k(p_j)}{\sum_{j=1}^{l} A_k(p_j)}, \quad k = 1, \ldots, n \tag{1}$$

is the (discrete) fuzzy transform of f *with respect to* $A_1, \ldots, A_n$.

Definition 4 [7]**.** *Let* f *be given at nodes* $p_1, \ldots, p_l \in [a, b]$ *and* $F[f]$ *be the fuzzy transform of* f *with respect to the fuzzy partition* $A_1, \ldots, A_n$. *Then, the function*

$$f_F(p_j) = \sum_{k=1}^{n} F_k A_k(p_j)$$

defined at the same nodes is the inverse fuzzy transform.

The definition of the fuzzy and inverse fuzzy transform can be extended to functions of more than one variable. In this work we are interested in working with images, which can be seen as discrete functions of two variables. Therefore, we will use the two-dimensional fuzzy and inverse fuzzy transform.

Definition 5 [7]**.** *Let a function* f *be given at nodes* $(p_i, q_j) \in [a, b] \times [c, d]$, $i = 1, \ldots, N$, $j = 1, \ldots, M$ *and* $A_1, \ldots, A_n$, $B_1, \ldots, B_m$ *where* $n < N$, $m < M$, *be basic functions which form a fuzzy partition of* $[a, b]$ *and* $[c, d]$, *respectively. Suppose that sets* $p_1, \ldots, p_N$ *and* $q_1, \ldots, q_M$ *of these nodes are sufficiently dense with respect to the chosen partition. We say that the* $n \times m$ *matrix of real numbers* $\mathbf{F}[f] = (F_{kl})$ *is the discrete fuzzy transform of* f *with respect to* $A_1, \ldots, A_n$, *and* $B_1, \ldots, B_m$ *if*

$$F_{kl} = \frac{\sum_{j=1}^{M} \sum_{i=1}^{N} f(p_i, q_j) A_k(p_i) B_l(q_j)}{\sum_{j=1}^{M} \sum_{i=1}^{N} A_k(p_i) B_l(q_j)} \tag{2}$$

holds for all $k = 1, \ldots, n$, $l = 1, \ldots, m$.

Definition 6 [7]**.** *Let* $A_1, \ldots, A_n$ *and* $B_1, \ldots, B_m$ *be basic functions which form a fuzzy partition of* $[a, b]$ *and* $[c, d]$ *respectively. Let* f *be given at points* $(p_i, q_j) \in [a, b] \times [c, d]$, $i = 1, \ldots, N$, $j = 1, \ldots, M$ *and* $\mathbf{F}[f]$ *be the fuzzy transform of* f *with respect to* $A_1, \ldots, A_n$ *and* $B_1, \ldots, B_m$. *Then the function*

$$f_{\mathbf{F}}(p_i, q_j) = \sum_{k=1}^{n} \sum_{l=1}^{m} F_{kl} A_k(p_i) B_l(q_j) \tag{3}$$

defined at the same nodes is the inverse fuzzy transform.

3 The Gravitational Search Algorithm

In this section we present the main steps concerning the GSA. Basically, the objective of the GSA is to minimize (or maximize) a fitness function defined on an n-dimensional space (search space). The optimization is performed iteratively by means of a set of agents that represent feasible solutions. Each agent has an acceleration and a velocity that is determined by the effect of the rest of agents: the best agents (the best solutions to the optimization problem) attract each other with a greater force.

Technically, consider a systems of T particles (agents). The position of each agent X_i in a p-dimensional space is given by

$$X_i = (x_i^1, \ldots, x_i^d, \ldots, x_i^p)$$

for each $i \in \{1, \ldots, T\}$. At each specific time t, the mass of a particle X_i represents the adaptation of that specific particle to the problem. This is done by means of a fitness function, that maps the p-dimensional space where the particles are defined into the set of real positive numbers. Each agent X_i has a particular mass, which is determined by the fitness function as follows:

$$m_i(t) = \frac{fitness_i(t) - worst(t)}{best(t) - worst(t)}$$

where

$$best(t) = \min_{j\in\{1,\ldots,T\}} fit_j(t) \text{ and}$$

$$worst(t) = \max_{j\in\{1,\ldots,T\}} fit_j(t)$$

Finally, the masses are normalized by means of:

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^{T} m_j(t)}$$

such that $\sum_{i=1}^{T} M_i(t) = 1$.

Remark 1. Previous formulae are used in minimization problems. In the case of a maximization problem, $best(t)$ and $worst(t)$ are calculated as the maximum and minimum of fitness function, respectively.

As commented before, the GSA is based on the movement of agents, that search along the search space for minima of the fitness function. In order to get this movement, an acceleration is calculated for each agent as the result of the acting forces of the rest of agents. The force acting on agent i by the rest of agents is given as

$$F_i^d(t) = \sum_{j=1 j\neq i}^{T} r_j F_{ij}^d(t)$$

where r_j is a random number in $[0, 1]$ and

$$F_{ij}^d(t) = G(t)\frac{M_i(t)M_j(t)}{R_{ij}(t) + \epsilon}(x_j^d(t) - x_i^d(t)),$$

ϵ is small positive constant and

$$R_{ij}(t) = ||X_i(t) - X_j(t)||_2.$$

Once the force acting over agent i is calculated, the acceleration is given by

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} = \sum_{j=1 j\neq i}^{T} r_j \frac{M_j(t)}{R_{ij}(t)}(x_j^d(t) - x_i^t(d).$$

The next velocity of an agent is given as

$$v_i^d(t+1) = r_i v_i^d(t) + a_i^d(t)$$

where r_i is a random number in $[0, 1]$ and, finally, the next position is given as

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1).$$

Taking into account previous formulae, the most suitable agents will tend to have heavier masses along the iterations and, therefore, tend to attract the other with greater forces. At the end, the agents tend to move toward the best agent. A summarization of the GSA is shown in Algorithm 1.

Algorithm 1. Gravitational Search Algorithm

```
Input: Number of agents T. Fitness function.
Output: Best agent
  for i = 1,...,T do
    Random initialization of X_i
  end for
  while stop criteria not reached do
    for i = 1,...,T do
      Evaluate fitness fitness_i(t) of each agent X_i
    end for
    Update G(t), best(t), worst(t) and M_i(t)
    for i = 1,...,T do
      Calculate the total force F_i(t) acting on agent i
      Calculate the acceleration A_i(t)
      Calculate the velocity V_i(t)
      Update position X_i(t+1)
    end for
  end while
  Return best(t) as the best agent
```

4 GSA for Tuning the Fuzzy Partition of the Fuzzy Transform

As we have stated in the introduction, the fuzzy (and inverse) transform have demonstrated to be a useful tool for image compression. Usually, the use of a uniform fuzzy partition is assumed when compressing images due to its simplicity and the fact that no extra information needs to be stored. However, it is clear that given a fixed number of nodes (fuzzy sets) in the fuzzy partition, different locations of nodes produce different fuzzy transforms. In this section we explain how to optimize the location of the nodes of a fuzzy partition in order to obtain better compressed images.

We recall that, given an image f of $N \times M$ pixels (generally $f : \{1, \dots, N\} \times \{1, \dots, M\} \to \{0, 1, \dots, 255\}$) and fuzzy sets $A_1, \dots, A_n : [1, N] \to [0, 1]$, $B_1, \dots, B_m : [1, M] \to [0, 1]$ forming a fuzzy partition of $[1, N]$ and $[1, M]$, respectively, the fuzzy transform $\mathbf{F}$ of f with respect to $A_1, \dots, A_n$, $B_1, \dots, B_m$ is a new (compressed) matrix of $n \times m$ pixels. The inversion process is performed by the inverse fuzzy transform that, starting from $\mathbf{F}$ and from the same fuzzy partition $A_1, \dots, A_n$, $B_1, \dots, B_m$, obtains a new (uncompressed) matrix $f_{\mathbf{F}}$ of $N \times M$ pixels.

The loss of quality produced by applying the fuzzy and inverse fuzzy transform to image f can be calculated by means of several distance or error measures. One of such measures is the main squared error (MSE), that is calculated as follows:

$$MSE(f, f_{\mathbf{F}}) = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} (f(p_i, q_j) - f_{\mathbf{F}}(p_i, q_j))^2.$$

Since both $\mathbf{F}$ and $f_{\mathbf{F}}$ depends on the fuzzy partition formed by $A_1, \dots, A_n$ and $B_1, \dots, B_m$, one possible way of optimizing the fuzzy transform is by tuning the position of the fuzzy subsets A_i, B_j [11].

Remark 2. In this work we assume that every fuzzy set of the fuzzy partition has triangular membership function. Therefore, a fuzzy partition is totally determined by the position of the fixed nodes (see Definition 1).

In order to do this, we propose to model each possible fuzzy partition by means of a real vector of $n+m$ real numbers. The first n real numbers represent the position of the fixed nodes $x_1, \dots, x_n \in [1, N]$ such that $A_i(x_i) = 1$. The real numbers going from the $n+1$-th to the $n+m$-th position represent the nodes $y_1, \dots, y_m \in [1, M]$ such that $B_j(y_j) = 1$. Therefore, a fuzzy partition is represented by a real vector $X \in [1, N]^n \times [1, M]^m$. However, in order to have a valid representation of a fuzzy partition we need to assure that:

- $x_1^j = 1, x_n^j = N, x_{n+1}^j = 1, x_{n+m}^j = M$, for every $j \in \{1, \dots, p\}$;
- $x_i^j < x_i^{j+1}$ for every $j \in \{1, \dots, n-1\}$ and every $j \in \{n+1, \dots, n+m-1\}$;
- $x_i^{j+1} - x_i^{j+1}$ for every $j \in \{2, \dots, n-1\}$ and every $j \in \{n+2, \dots, n+m-1\}$.

Remark 3. This process is done in our implementation of the GSA and those agents that represent invalid fuzzy partitions automatically transformed in order to assure previous conditions.

Finally, the summarization of our optimization process is explained in Algorithm 2.

Algorithm 2. Optimization of fuzzy transform

Input: Image f of $N \times M$ pixels. Size of compressed image $n < N$ and $m < M$. Number of agents T.

Output: Optimized fuzzy partition $A_1, \ldots, A_n$, $B_1, \ldots, B_m$. Optimized compressed image $\mathbf{F}$. Uncompressed image $f_{\mathbf{F}}$. MSE between f and $f_{\mathbf{F}}$

for $i = 1, \ldots, T$ **do**
  Initialize agent $X_i \in [1, N]^n \times [1, M]^m$ randomly following the restrictions mentioned above.
end for
Execute Algorithm 1 and obtain $Xbest$ as the best particle
Decode $Xbest$ into fuzzy partition $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$
Calculate fuzzy transform $\mathbf{F}$ of f with respect to $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$
Calculate inverse fuzzy transform $f_{\mathbf{F}}$ of F with respect to $A_1, \ldots, A_n$ and $B_1, \ldots, B_m$
Calculate $MSE(f, f_{\mathbf{F}})$

5 Experimental Results

In this section we evaluate our optimization procedure based on the GSA for tuning the fuzzy partition associated with the fuzzy transform. We have first taken an original image of size 321×481 pixels (see Fig. 1) and we have fixed $n = 160, m = 240$. In order to test different settings of the GSA, we evaluate Algorithm 2 with different number of agents, specifically $T = 5, 25, 50, 100$. The MSE of each optimized fuzzy transform is shown in Table 1 (second row). According to the results it is not until we have 100 agents that we outperform the non-optimized uniform fuzzy partition. Besides, the improvement is very small, probably due to the large size of the fuzzy partition (big number of nodes).

With the purpose of testing whether a smaller number of nodes in the fuzzy partition increases the improvement of the optimized fuzzy transform, we have executed Algorithm 2 with

- $n = 107, m = 160$;
- $n = 80, m = 120$ and
- $n = 64, m = 96$.

All of these executions have been also tried with $T = 5, 25, 50, 100$. The MSEs obtained are shown again in Table 1, row number three, four and five, respectively.

Analyzing the results we first realize that having a big number of agents guarantees finding a good enough solution to the optimization problem. If we

Fig. 1. Original image 3096 from [5] (321 × 481 pixels).

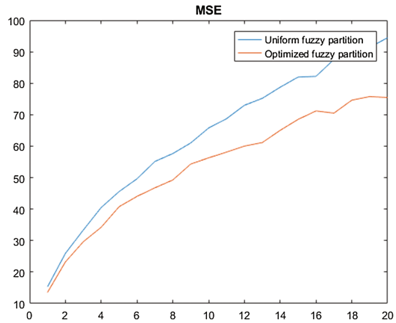


Fig. 2. Comparison of MSE of reconstructed images using a uniform fuzzy partition and Algorithm 2 with $T = 100$.

Table 1. MSE obtained from reconstructed images using uniform and optimized fuzzy partitions.

	Uniform	$T = 5$	$T = 25$	$T = 50$	$T = 100$
$n = 160, m = 240$	12,25	35,40	13,44	13,12	10,62
$n = 107, m = 160$	24,57	68,57	24,09	23,75	21,30
$n = 80, m = 120$	35,56	97,91	37,18	33,09	31,63
$n = 64, m = 96$	46,19	75,74	42,97	41,00	40,86

take $T = 5$ then Algorithm 2 is not able to find a good solution to the problem. However, when $T = 100$ we always outperform the uniform partition.

Now, if we focus on the behavior of the MSE among different partition sizes, we see that the proposed optimization algorithm is able to find better compressed images (if we have enough number of agents). Although the numbers shown in Table 1 do not differ too much depending on the size of the partition (the improvement is around 12% in every case), it seems evident that as long as the size of the compressed image decreases, the improvement should increase. In order to prove this we have executed Algorithm 2 with $T = 100$ and different sizes of the compressed image. The results in terms of MSE are shown in Fig. 2, where the MSE of a uniform and an optimized fuzzy partition are shown. In the horizontal axis we show the different compression rates, i.e. the ratio between the size of the compressed image (size of the fuzzy partition) and the size of the original image. In this sense, a compression rate of 100 means that the original image has 381 × 421 and the compressed image 38 × 42. Observe that the difference between the MSE of the uniform and optimized fuzzy partition increases as long as the compressed ratio increases.

Finally, in Fig. 3 we show the reconstructed images obtained from a uniform (first column) and the reconstructed images obtained from an optimized fuzzy

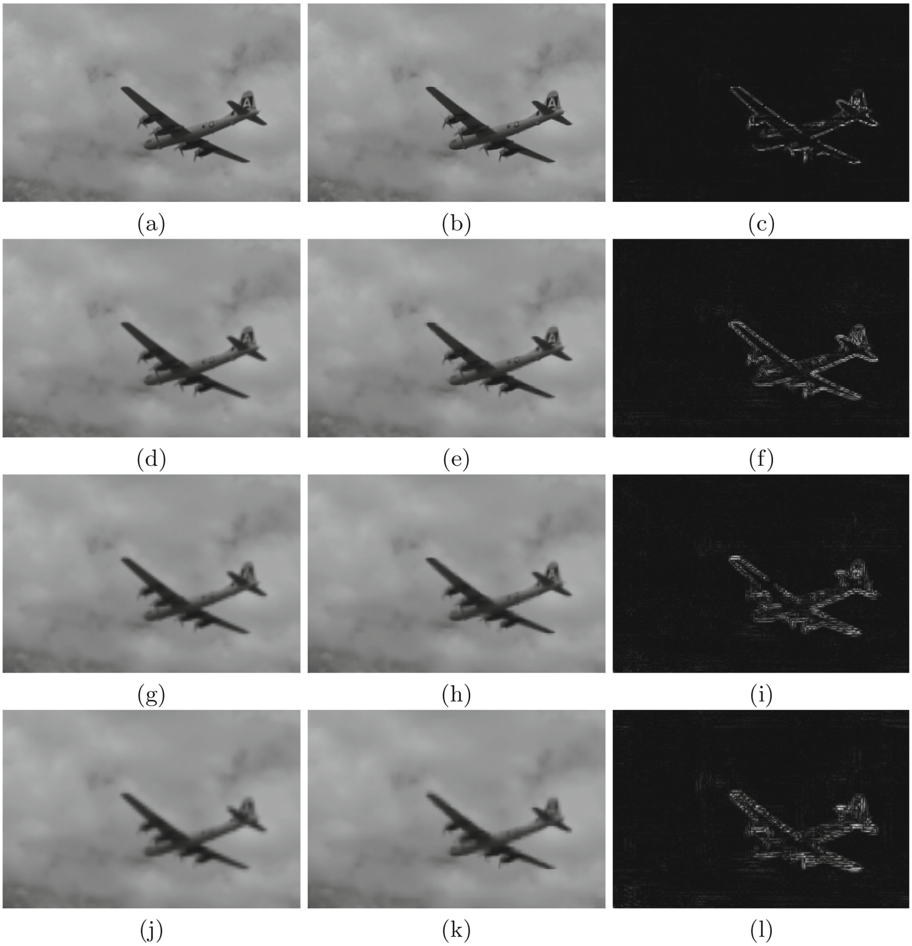


Fig. 3. First column: reconstructed images using a uniform fuzzy partition and sizes $m = 160, n = 240$ (a), $m = 107, n = 160$ (d), $m = 80, n = 120$ (g) and $m = 64, n = 96$ (j). Second column: reconstructed images using an optimized fuzzy partition by Algorithm 2 and sizes $m = 160, n = 240$ (b), $m = 107, n = 160$ (e), $m = 80, n = 120$ (h) and $m = 64, n = 96$ (k). Third column: differences between images a and d (c), d and e (f), g and h (i) and j and k (l).

partition (second column) obtained applying Algorithm 2 and taking $T = 100$ agents and the same sizes used in Table 1. Notice that the differences between the first and second row appears mainly on the edges of objects. This can be better seen in the third column of Fig. 3, where the differences between the images of the first and second column are shown (the differences have been normalized so that the maximum difference appears in white). The conclusion obtained is that while using uniform fuzzy partitions produce blurring in the reconstructed images, the optimization process allows to allocate the fuzzy sets in those areas of interest where there exist large enough changes of intensities. Then, the fuzzy transform is able to capture these changes in a better way and the information is not lost in the process.

6 Conclusions

in this work we have proposed an optimization problem to find the best fuzzy partition associated with an image. The solution of this problem allows to obtain better compressed images by means of the fuzzy transform.

In order to solve the optimization problem we have based on the Gravitational Search Algorithm. The first results obtained show that the GSA is able to obtain optimized fuzzy partitions that minimize the error measure of the image compression procedure.

In the future, we want to extend the experimental study to a wider set of images, analyzing which images are suitable for this algorithm. Moreover, it would be interesting to compare the results when using different optimization algorithms rather than the GSA.

Acknowledgment. This work has been partially supported by MINECO, AEI/FEDER, UE under project TIN2016-77356-P and grant VEGA 1/0420/15.

References

1. Beliakov, G., Bustince, H., Paternain, D.: Image reduction using means on discrete product lattices. IEEE Trans. Image Process. **21**, 1070–1083 (2012)
2. Di Martino, F., Sessa, S.: Compression and decompression of images with discrete fuzzy transforms. Inf. Sci. **177**, 2349–2362 (2007)
3. Di Martino, F., Loia, V., Perfilieva, I., Sessa, S.: An image coding/decoding method based on direct and inverse fuzzy transforms. Int. J. Approx. Reason. **48**, 110–131 (2008)
4. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Pearson International Edition, New Jersey (2008)
5. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th International Conference on Computer Vision, pp. 416–423 (2001)
6. Paternain, D., Fernandez, J., Bustince, H., Mesiar, R., Beliakov, G.: Construction of image reduction operators using averaging aggregation functions. Fuzzy Sets Syst. **261**, 87–111 (2015)

7. Perfilieva, I.: Fuzzy transforms: theory and applications. Fuzzy Sets Syst. **157**, 993–1023 (2006)
8. Perfilieva, I., De Baets, B.: Fuzzy transforms of monotone functions with application to image compression. Inf. Sci. **180**, 3304–3315 (2010)
9. Perfilieva, I., Hurtik, P., Di Martino, F., Sessa, S.: Image reduction method based on the F-transform. Soft Comput. **21**, 1847–1861 (2017)
10. Rashedi, E., Nezamabadi-pour, H., Saryazdi, S.: GSA: a gravitational search algorithm. Inf. Sci. **179**, 2232–2248 (2009)
11. Štěpnička, S., Polakovič, O.: A neural network approach to the fuzzy transform. Fuzzy Sets Syst. **160**, 1037–1047 (2009)

Fuzzy Decision Matrices in Case of a Discrete Underlying Fuzzy Probability Measure

Ondřej Pavlačka(✉) and Pavla Rotterová

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University Olomouc,
17. listopadu 1192/12, 771 46 Olomouc, Czech Republic
{ondrej.pavlacka,pavla.rotterova01}@upol.cz

Abstract. Decision matrices represent a common tool for solving decision-making problems under risk. Elements of the matrix express the outcomes if a decision-maker chooses the particular alternative and the particular state of the world occurs. We deal with the problem of extension of a decision matrix to the case of fuzzy states of the world and fuzzy outcomes of alternatives. We consider the approach based on the idea that a fuzzy decision matrix determines a collection of fuzzy rule-based systems. The aim of the paper is to study extension of this approach to the case where the states of the world are fuzzy sets on the finite universal set and the probabilities of elementary events are determined by a tuple of fuzzy probabilities. We derive the formulas for computations of the fuzzy expected values and fuzzy variances of the outcomes of alternatives, based on which the alternatives can be compared.

Keywords: Decision matrices · Decision-making under risk · Fuzzy probability measure · Fuzzy states of the world · Fuzzy rule-based systems

1 Introduction

In decision making under risk, decision matrices, see Table 1, are often used as a tool of risk analysis, see e.g. [4,6,8,17,18]. They describe how the outcomes of alternatives $x_1, \ldots, x_n$ depend on the fact which of the possible and mutually disjoint states of the world $S_1, \ldots, S_m$ will occur in the future. The probabilities of occurrences of the states of the world are given by $p_1, \ldots, p_m$, where $p_j > 0$, $j = 1, \ldots, m$, and $\sum_{j=1}^{m} p_j = 1$. Thus, the outcome from choosing an alternative x_i, $i \in \{1, \ldots, n\}$, is a discrete random variable H_i that takes on the values $h_{i,1}, \ldots, h_{i,m}$ with the probabilities $p_1, \ldots, p_m$. The alternatives are usually compared on the basis of their expected outcomes $EH_i = \sum_{j=1}^{m} p_j h_{i,j}$, $i = 1, \ldots, n$, and the variances of their outcomes $var\, H_i = \sum_{j=1}^{m} p_j (h_{i,j} - EH_i)^2$, $i = 1, \ldots, n$.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_12

Table 1. Decision matrix

	S_1	S_2	$\cdots$	S_m		
	p_1	p_2	$\cdots$	p_m		
x_1	$h_{1,1}$	$h_{1,2}$	$\cdots$	$h_{1,m}$	EH_1	$var\ H_1$
x_2	$h_{2,1}$	$h_{2,2}$	$\cdots$	$h_{2,m}$	EH_2	$var\ H_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$h_{n,1}$	$h_{n,2}$	$\cdots$	$h_{n,m}$	EH_n	$var\ H_n$

In practical applications, the states of the world as well as the outcomes of the alternatives can be determined vaguely. For instance, the particular state of the world could be defined as "inflation is low". Similarly, the outcome of an alternative under a certain state of the world can be described as "about 5% yield", or as a linguistic evaluation from a given linguistic scale, e.g. "a small profit". The vaguely described pieces of information in decision matrix can be mathematically modelled by means of tools of fuzzy sets theory, see e.g. [5,8,17] and references therein.

In [12–14], the authors considered the following model: A probability space $(\Omega, \mathcal{A}, P)$ is given, where Ω denotes a non-empty set of all elementary events, $\mathcal{A}$ represents the set of all considered random events ($\mathcal{A}$ forms a σ-algebra of subsets of Ω), and $P : \mathcal{A} \rightarrow [0,1]$ is a probability measure that assigns to each random event $A \in \mathcal{A}$ its probability $P(A) \in [0,1]$. The states of the world $S_1, \ldots, S_m$ are fuzzy sets on Ω whose membership functions $\mu_{S_1}, \ldots, \mu_{S_m}$ are $\mathcal{A}$-measurable, i.e. the α-cuts $S_{j\alpha} \in \mathcal{A}$ for any $\alpha \in (0,1]$, $j = 1, \ldots, m$, and that form a fuzzy partition of Ω, i.e. $\sum_{j=1}^{m} \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$. The common approach in such a case, considered in the literature (see e.g. [7,17]), consists in applying the well known crisp probabilities of fuzzy events computed as the expected membership degrees, as was proposed by Zadeh in [19]. However, it was discussed in [13] that it is not appropriate to model the outcomes of alternatives as discrete random variables taking on their values with the Zadeh's probabilities of the fuzzy states of the world, mainly due to a lack of interpretability (for an overview of interpretability of fuzzy systems, see e.g. [1]). Therefore, the authors in [12,14] proposed an alternative way how the information contained in a fuzzy decision matrix can be treated. Their approach is based on the idea that a fuzzy decision matrix does not determine discrete random variables, but a collection of fuzzy rule-based systems (a fuzzy rule-based system was introduced in [20]). In [12], only the crisp (i.e. not fuzzy) outcomes of alternatives were considered which made the problem much simpler. They showed that within this approach, the obtained characteristics of the outcomes of alternatives, like fuzzy expected values or fuzzy variances, are clearly interpretable.

In practice, we can also meet the problem where the underlying probability measure P is ill-known, see e.g. [2,17] and references therein. In such cases, it can be modelled by means of tools of fuzzy sets theory as well, see e.g. [2,3,17]. The aim of the paper is to extend the approach proposed in [12,14] to the case where the states of the world $S_1, \ldots, S_m$ form a fuzzy partition of the finite universal set $\Omega = \{\omega_1, \ldots, \omega_r\}$, $m < r$, and where the probabilities of elementary events $\{\omega_1\}, \ldots, \{\omega_r\}$ are determined expertly, by an r-tuple of fuzzy probabilities $P_1, \ldots, P_r$. The problem will be illustrated by an practical example.

2 Fuzzy Decision Matrices Viewed as a Collection of Fuzzy Rule-Based Systems

First, let us introduce the idea proposed in [12,14] that a fuzzy decision matrix defines a collection of fuzzy rule-based systems.

In concordance with Introduction, let us consider that a probability space $(\Omega, \mathcal{A}, P)$ is given and that the states of the world are described by fuzzy sets $S_1, \ldots, S_m$ forming a fuzzy partition of Ω. Further, let for any $i \in \{1, \ldots, n\}$, the fuzzy outcomes of the alternative x_i under the particular fuzzy states of the world be given by the fuzzy numbers $H_{i,j}$, $j = 1, \ldots, m$. Then, information about the outcome of choosing the alternative x_i can be expressed by the following m-tuple of If-Then rules:

$$\begin{array}{l} \text{If the state of the world is } S_1, \text{ then the outcome of } x_i \text{ is } H_{i,1}. \\ \text{If the state of the world is } S_2, \text{ then the outcome of } x_i \text{ is } H_{i,2}. \\ \ldots \\ \text{If the state of the world is } S_m, \text{ then the outcome of } x_i \text{ is } H_{i,m}. \end{array} \tag{1}$$

In [12], it was shown that in the case of the fuzzy decision matrix with crisp outcomes $H_{i,j} = h_{i,j} \in \mathbb{R}$, $j = 1, \ldots, m$, it is appropriate to use the *Sugeno's method of fuzzy inference*, introduced in [15]. The obtained output from the fuzzy rule-based system (1) is expressed by a real number

$$H_i^S(\omega) = \frac{\sum_{j=1}^m \mu_{S_j}(\omega) h_{i,j}}{\sum_{j=1}^m \mu_{S_j}(\omega)} = \sum_{j=1}^m \mu_{S_j}(\omega) h_{i,j}, \tag{2}$$

where the assumption $\sum_{j=1}^m \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$ was used. In the case of the fuzzy outcomes of alternatives $H_{i,j}$, it was shown in [14] that the so-called *generalised Sugeno's method of fuzzy inference*, introduced in [16], should be applied for obtaining an output from the fuzzy rule-based system (1). According to this method, the fuzzy outcome is for any $\omega \in \Omega$ given by a fuzzy number

$$H_i^S(\omega) = \frac{\sum_{j=1}^m \mu_{S_j}(\omega) H_{i,j}}{\sum_{j=1}^m \mu_{S_j}(\omega)} = \sum_{j=1}^m \mu_{S_j}(\omega) H_{i,j}, \tag{3}$$

where, again, the assumption $\sum_{j=1}^m \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$ was used.

Since we operate within the given probability space $(\Omega, \mathcal{A}, P)$, the mapping $H_i^S : \Omega \to \mathcal{F}_N(\mathbb{R})$, where $\mathcal{F}_N(\mathbb{R})$ denotes the family of all fuzzy numbers, is a fuzzy random variable. It can be easily seen from (2) and (3) that in the case of crisp states of the world $S_1, \ldots, S_m \subset \Omega$ and crisp outcomes $h_{i,j} \in \mathbb{R}$, $i =$

$1, \ldots, n$, $j = 1, \ldots, m$, the random variables $H_1^S, \ldots, H_n^S$ coincide with discrete random variables $H_1, \ldots, H_n$ taking on the values $h_{i,1}, \ldots, h_{i,m}$, $i = 1, \ldots, n$, with the probabilities $p_j = P(S_j)$, $j = 1, \ldots, m$. Hence, this approach represents an extension of a decision matrix to the case of fuzzy states of the world.

Analogously as in the common approach to the fuzzy decision matrix, the ordering of the alternatives $x_1, \ldots, x_n$ can be based on the fuzzy expected values and the fuzzy variances of the random variables $H_1^S, \ldots, H_n^S$. The formulas for computations of EH_i^S and $var H_i^S$, $i = 1, \ldots, n$, derived in [14], are as follows: Let us denote the α-cuts of fuzzy outcomes by $H_{i,j\alpha} = \left[\underline{H_{i,j}}(\alpha), \overline{H_{i,j}}(\alpha)\right]$. For any $\alpha \in (0,1]$, the α-cut of the fuzzy expected outcome EH_i^S, denoted by $EH_{i\alpha}^S = \left[\underline{EH_i^S}(\alpha), \overline{EH_i^S}(\alpha)\right]$, is obtained as follows:

$$\underline{EH_i^S}(\alpha) = \int_{\omega \in \Omega} \sum_{j=1}^{m} \mu_{S_j}(\omega) \underline{H_{i,j}}(\alpha) dP, \tag{4}$$

$$\overline{EH_i^S}(\alpha) = \int_{\omega \in \Omega} \sum_{j=1}^{m} \mu_{S_j}(\omega) \overline{H_{i,j}}(\alpha) dP. \tag{5}$$

Calculation of the α-cuts of the fuzzy variance $var\ H_i^S$, denoted by $var\ H_{i\alpha}^S = \left[\underline{var\ H_i^S}(\alpha), \overline{var\ H_i^S}(\alpha)\right]$, is more complex. Let us denote

$$s(h_{i,1}, \ldots, h_{i,m}) = \int_{\omega \in \Omega} \sum_{j=1}^{m} \left(\mu_{S_j}(\omega) h_{i,j} - \int_{\omega' \in \Omega} \sum_{k=1}^{m} \mu_{S_k}(\omega') h_{i,k} dP \right)^2 dP.$$

Then

$$\underline{var\ H_i^S}(\alpha) = \min \left\{ s(h_{i,1}, \ldots, h_{i,m}) \mid h_{i,j} \in H_{i,j\alpha},\ j = 1, \ldots, m \right\}, \tag{6}$$

$$\overline{var\ H_i^S}(\alpha) = \max \left\{ s(h_{i,1}, \ldots, h_{i,m}) \mid h_{i,j} \in H_{i,j\alpha},\ j = 1, \ldots, m \right\}. \tag{7}$$

The obtained fuzzy decision matrix is given in Table 2. We can see that there are no probabilities of fuzzy states of the world since they are not considered in this approach.

Table 2. Fuzzy decision matrix

	S_1	S_2	$\cdots$	S_m		
x_1	$H_{1,1}$	$H_{1,2}$	$\cdots$	$H_{1,m}$	EH_1^S	$var\ H_1^S$
x_2	$H_{2,1}$	$H_{2,2}$	$\cdots$	$H_{2,m}$	EH_2^S	$var\ H_2^S$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$H_{n,1}$	$H_{n,2}$	$\cdots$	$H_{n,m}$	EH_n^S	$var\ H_n^S$

3 Fuzzy Decision Matrix in the Case of Underlying Discrete Fuzzy Probability Measure

Now, let us extend a fuzzy decision matrix described in Sect. 2 to the case where the universal set Ω is finite and the underlying probability measure P is fuzzy. Let us derive the formulas for computation of the fuzzy expected values and fuzzy variances of the outcomes of alternatives. In next section, the problem will be illustrated by an example.

Let $\Omega = \{\omega_1, \ldots, \omega_r\}$ be a universal set. Let us assume that the probabilities of elementary events $\{\omega_1\}, \ldots, \{\omega_r\}$ are not known precisely and are set, typically on the basis of experts' knowledge and experience, by a special structure of fuzzy numbers $P_1, \ldots, P_r$ defined on $[0,1]$ called an r-tuple of fuzzy probabilities.

Definition 1 (See e.g. [9,17]). *We say that fuzzy numbers $P_1, \ldots, P_r$, defined on $[0,1]$, form an r-tuple of fuzzy probabilities if for all $\alpha \in (0,1]$ and for all $k \in \{1, \ldots, r\}$ the following holds: for any $p_k \in P_{k\alpha}$ there exist $p_l \in P_{l\alpha}$, $l = 1, \ldots, r$, $l \neq k$, such that*

$$p_k + \sum_{l=1, l\neq k}^{r} p_l = 1.$$

Remark 1. Various methods for expert setting of a tuple of fuzzy probabilities are proposed in [9,10]. A more general approach to modelling uncertain values of probabilities, consisting in employing fuzzy vectors, was introduced in [11].

Further, let the fuzzy states of the world $S_1, \ldots, S_m$, $m < r$, be given in the form of fuzzy sets on Ω that form a fuzzy partition of Ω, i.e. $\sum_{j=1}^{m} \mu_{S_j}(\omega_k) = 1$ for any $k \in \{1, \ldots, r\}$. Further, let the information about the outcome of choosing the alternative x_i, $i \in \{1, \ldots, n\}$ be expressed by the m-tuple of If-Then rules 1. According to (3), the fuzzy outcome is for any $\omega_k \in \Omega$ given by a fuzzy number

$$H_i^S(\omega_k) = \sum_{j=1}^{m} \mu_{S_j}(\omega_k) H_{i,j}. \tag{8}$$

Thus, the outcome of the alternative x_i can be seen as a discrete fuzzy random variable H_i^{SF} taking on the fuzzy values $H_i^S(\omega_1), \ldots, H_i^S(\omega_r)$ with the fuzzy probabilities $P_1, \ldots, P_r$. Let us show now how can be the general formulas (4)–(7) for computing the fuzzy expected values and fuzzy variances of the outcomes of alternatives extended to this case.

For any $\alpha \in (0,1]$, the α-cut of the fuzzy expected output EH_i^{SF} from the fuzzy rule-based system (1), denoted by $EH_{i\alpha}^{SF} = \left[\underline{EH_i^{SF}}(\alpha), \overline{EH_i^{SF}}(\alpha)\right]$, is obtained as follows:

$$\underline{EH_i^{SF}}(\alpha) = \min\left\{ \sum_{k=1}^{r} p_k \sum_{j=1}^{m} \mu_{S_j}(\omega_k) \underline{H_{i,j}}(\alpha) \mid p_k \in P_{k\alpha}, k = 1, \ldots, r, \sum_{k=1}^{r} p_k = 1 \right\}, \tag{9}$$

$$\overline{EH_i^{SF}}(\alpha) = \max\left\{\sum_{k=1}^{r} p_k \sum_{j=1}^{m} \mu_{S_j}(\omega_k)\overline{H_{i,j}}(\alpha) \mid p_k \in P_{k\alpha}, k=1,\ldots,r, \sum_{k=1}^{r} p_k = 1\right\}. \tag{10}$$

Remark 2. Let us note that the formulas (9) and (10) correspond to the well known operation called a fuzzy weighted average of fuzzy numbers that is widely studied in the literature (see e.g. [10,11] and the references therein). The algorithm for computing the fuzzy weighted average of fuzzy numbers can be found e.g. in [11].

The α-cut $var\ H_{i\alpha}^{SF} = \left[\underline{var\ H_i^{SF}}(\alpha), \overline{var\ H_i^{SF}}(\alpha)\right]$ of the fuzzy variance of the output from the fuzzy rule-based system (1) is obtained as follows:
Let us denote

$$s(h_{i,1},\ldots,h_{i,m},p_1,\ldots,p_r) = \sum_{k=1}^{r} p_k \left(\sum_{j=1}^{m} \mu_{S_j}(\omega_k)h_{i,j} - \sum_{t=1}^{r} p_t \sum_{u=1}^{m} \mu_{S_u}(\omega_t)h_{i,u}\right)^2.$$

Then,

$$\underline{var\ H_i^{SF}}(\alpha) = \min\left\{s(h_{i,1},\ldots,h_{i,m},p_1,\ldots,p_r) \mid h_{i,j} \in H_{i,j\alpha},\ j=1,\ldots,m,\ p_k \in P_{k\alpha},\ k=1,\ldots,r,\ \sum_{k=1}^{r} p_k = 1\right\}, \tag{11}$$

$$\overline{var\ H_i^{SF}}(\alpha) = \max\left\{s(h_{i,1},\ldots,h_{i,m},p_1,\ldots,p_r) \mid h_{i,j} \in H_{i,j\alpha},\ j=1,\ldots,m,\ p_k \in P_{k\alpha},\ k=1,\ldots,r,\ \sum_{k=1}^{r} p_k = 1\right\}. \tag{12}$$

Let us note that the optimization problems (11) and (12) can be very difficult to solve. The algorithms for solving these problems could be a subject of further research.

4 Illustrative Example

Let us consider the following situation: We can realize one of the two possible projects, denoted by x_1 and x_2. The outcome (future yield) depends solely on the fact what kind of a government coalition will be established after the parliamentary election. Let us assume that only the six possible coalitions, denoted by $\omega_1, \ldots, \omega_6$, can be established after the election. The probabilities of establishing each coalition are not known precisely; they are set expertly the triangular fuzzy numbers $P(\{\omega_k\}) = P_k = \langle 1/12, 1/6, 1/3 \rangle$, $k = 1, \ldots, 6$.

We distinguish three vaguely defined states of the world - a right coalition (S_1), a centre coalition (S_2), and a left coalition (S_3) that are expressed by the following fuzzy sets defined on Ω:

$$S_1 = \left\{ {}^{1}|_{\omega_1}, {}^{0.8}|_{\omega_2}, {}^{0.2}|_{\omega_3}, {}^{0}|_{\omega_4}, {}^{0}|_{\omega_5}, {}^{0}|_{\omega_6} \right\},$$
$$S_2 = \left\{ {}^{0}|_{\omega_1}, {}^{0.2}|_{\omega_2}, {}^{0.8}|_{\omega_3}, {}^{1}|_{\omega_4}, {}^{0.5}|_{\omega_5}, {}^{0}|_{\omega_6} \right\},$$
$$S_3 = \left\{ {}^{0}|_{\omega_1}, {}^{0}|_{\omega_2}, {}^{0}|_{\omega_3}, {}^{0}|_{\omega_4}, {}^{0.5}|_{\omega_5}, {}^{1}|_{\omega_6} \right\},$$

where elements of the sets are in the form ${}^{\mu_{S_j}(\omega_k)}|_{\omega_k}$, $j = 1, 2, 3$, and $k = 1, \ldots, 6$.

The future yields (in %) from realization one of the possible projects x_1 or x_2 in the cases that the particular type of a coalition will be established are also not known precisely. They are expressed by triangular fuzzy numbers that are shown in the fuzzy decision matrix given by Table 3.

Table 3. Fuzzy decision matrix

	S_1	S_2	S_3
x_1	$\langle 10, 15, 20 \rangle$	$\langle -3, 0, 3 \rangle$	$\langle -12, -4, 4 \rangle$
x_2	$\langle 10, 15, 20 \rangle$	$\langle -12, -3, 9 \rangle$	$\langle -4, 0, 4 \rangle$

Now, let us construct the fuzzy random variables H_1^{SF} and H_2^{SF} representing the future fuzzy yields from the projects. The fuzzy outputs of the If-Then rules derived from the fuzzy decision matrix, i.e. the triangular fuzzy numbers $H_1^S(\omega_1), \ldots, H_1^S(\omega_6)$ and $H_2^S(\omega_1), \ldots, H_2^S(\omega_6)$, are given in Table 4.

Table 4. Outputs of If-Then rules for the particular possible coalitions

	x_1	x_2
ω_1	$\langle 10, 15, 20 \rangle$	$\langle 10, 15, 20 \rangle$
ω_2	$\langle 7.4, 12, 16.6 \rangle$	$\langle 5.6, 11.4, 17.8 \rangle$
ω_3	$\langle -0.4, 3, 6.4 \rangle$	$\langle -7.6, 0.6, 11.2 \rangle$
ω_4	$\langle -3, 0, 3 \rangle$	$\langle -12, -3, 9 \rangle$
ω_5	$\langle -7.5, -2, 3.5 \rangle$	$\langle -8, -1.5, 6.5 \rangle$
ω_6	$\langle -12, -4, 4 \rangle$	$\langle -4, 0, 4 \rangle$

The significant values of the fuzzy expected yields EH_1^{SF} and EH_2^{SF} and of the fuzzy variances $var\, H_1^{SF}$ and $var\, H_2^{SF}$, together with their centres of gravity are shown in Table 5.

Table 5. Significant values and centres of gravity

Characteristic	Significant values			Centre of gravity
EH_1^{SF}	−5.33	4.00	13.61	4.17
EH_2^{SF}	−6.33	3.75	15.16	4.33
$var\ H_1^{SF}$	4.13	50.33	196.54	75.61
$var\ H_2^{SF}$	2.20	47.03	200.58	73.77

According to the rule of maximization of the expected value and minimization of the variance of the outcome we can see from Table 5 that both fuzzy expected values as well as fuzzy variances are not comparable. If we employ some defuzzification method, e.g. center of gravity, we would prefer the alternative x_2 over x_1.

5 Conclusion

We have dealt with the problem of extension of a decision matrix to the case of fuzzy states of the world and fuzzy outcomes of alternatives. We have considered the recently developed approach, based on the idea that a fuzzy decision matrix determines a collection of fuzzy rule-based systems. Since in practical applications we can also meet the problem where the underlying probability measure P is not known precisely, the aim of the paper was to study some kind of extension of this approach to the case where the underlying probability measure is fuzzy. We have considered the case where the states of the world are fuzzy sets on the finite universal set and the probabilities of elementary events are determined expertly, by a tuple of fuzzy probabilities. We have derived the formulas for computations of the fuzzy expected values and fuzzy variances of the outcomes of alternatives, based on which the alternatives can be compared.

Next research in this field can be focused on the case of the continuous underlying probability measure where, for instance, the parameters of the probability distribution, like μ and σ in the case of the normal distribution, are fuzzy numbers, set expertly or derived from fuzzy data.

Acknowledgments. The paper is supported by the grant IGA_PrF_2017_019 *Mathematical Models* of the *Internal Grant Agency of Palacký University Olomouc*. The support is greatly acknowledged.

References

1. Alonso, J.M., Castiello, C., Mencar, C.: Interpretability of fuzzy systems: current research trends and prospects. In: Kacprzyk, J., Pedrycz, W. (eds.) Springer Handbook of Computational Intelligence, pp. 219–237. Springer, Heidelberg (2015). doi:10.1007/978-3-662-43505-2_14

2. Baudrit, C., Couso, I., Dubois, D.: Joint propagation of probability and possibility in risk analysis: towards a formal framework. Int. J. Approx. Reason. **45**, 82–105 (2007). doi:10.1016/j.ijar.2006.07.001
3. Dubois, D., Prade, H.: Fuzzy sets, probability and measurement. Eur. J. Oper. Res. **40**(2), 135–154 (1989). doi:10.1016/0377-2217(89)90326-3
4. Ganoulis, J.: Engineering Risk Analysis of Water Pollution: Probabilities and Fuzzy Sets. VCH, Weinheim (2008)
5. Herrera, F., Alonso, S., Chiclana, F., Herrera-Viedma, E.: Computing with words in decision making: foundations, trends and prospects. Fuzzy Optim. Decis. Ma. **8**, 337–364 (2009). doi:10.1007/s10700-009-9065-2
6. Chankong, V., Haimes, Y.Y.: Multiobjective Decision Making: Theory and Methodology. Elsevier Science B.V., New York (1983)
7. Okuda, T., Tanaka, H., Asai, K.: Formulation of fuzzy decision problems with fuzzy information using probability of fuzzy events. Inform. Control **38**(2), 135–147 (1978). doi:10.1016/S0019-9958(78)90151-1
8. Özkan, I., Türken, I.B.: Uncertainty and fuzzy decisions. In: Banerjee, S., et al. (eds.) Chaos Theory in Politics. Understanding Complex Systems, pp. 17–27. Springer, Dordrecht (2014). doi:10.1007/978-94-017-8691-1_2
9. Pan, Y., Yuan, B.: Bayesian inference of fuzzy probabilities. Int. J. Gen. Syst. **26**(1–2), 73–90 (1997). doi:10.1080/03081079708945170
10. Pavlačka, O., Talašová, J.: Application of the fuzzy weighted average of fuzzy numbers in decision making models. In: Bodenhofer, U., Novák, V., Štěpnička, M. (eds.) New Dimensions in Fuzzy Logic and Related Technologies, Vol II. Proceedings of the 5th EUSFLAT Conference, pp. 455–462. Ostravská univerzita, Ostrava (2007)
11. Pavlačka, O.: Modeling uncertain variables of the weighted average operation by fuzzy vectors. Inform. Sci. **181**(22), 4969–4992 (2011). doi:10.1016/j.ins.2011.06.022
12. Pavlačka, O., Rotterová, P.: Fuzzy decision matrices viewed as fuzzy rule-based systems. In: Kocourek, A., Vavroušek, M. (eds.) Proceedings of the 34th International Conference Mathematical Methods in Economics, pp. 641–646. Technical University of Liberec, Liberec (2016)
13. Rotterová, P., Pavlačka, O.: Probabilities of fuzzy events and their use in decision matrices. Int. J. Math. Op. Res. **9**(4), 423–435 (2016). doi:10.1504/IJMOR.2016.10000341
14. Rotterová, P., Pavlačka, O.: New approach to fuzzy decision matrices. Submitted to Acta Polytechnica Hungarica
15. Sugeno, M.: Industrial Applications of Fuzzy Control. Elsevier Science Pub. Co., New York (1985)
16. Talašová, J.: Fuzzy Methods of Multiple Criteria Evaluation and Decision Making (in Czech). Publishing House of Palacký University, Olomouc (2003)
17. Talašová, J., Pavlačka, O.: Fuzzy probability spaces and their applications in decision making. Austrian J. Statist. **35**(2&3), 347–356 (2006). doi:10.17713/ajs.v35i2&3.382
18. Yoon, K.P., Hwang, C.: Multiple Attribute Decision Making: An Introduction. SAGE Publications, California (1995)
19. Zadeh, L.A.: Probability measures of fuzzy events. J. Math. Anal. Appl. **23**(2), 421–427 (1968). doi:10.1016/0022-247X(68)90078-4
20. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning - I. Inform. Sci. **8**, 199–245 (1975). doi:10.1016/0020-0255(75)90036-5

Compositions Consistent with the Modus Ponens Property Used in Approximate Reasoning

Barbara Pękala(✉)

Interdisciplinary Centre for Computational Modelling,
Faculty of Mathematics and Natural Sciences, University of Rzeszów,
Pigonia 1, 35-310 Rzeszów, Poland
bpekala@ur.edu.pl

Abstract. In this paper it is investigated when some kinds of aggregation functions satisfy the Modus Ponens with respect to other aggregation function, or equivalently, when they are $\mathcal{A}$-conditionals. Moreover, some operation connected with $\mathcal{A}$-conditionals is examined and used to algorithm of approximate reasoning.

Keywords: Interval-valued fuzzy relation · Modus Ponens property · Approximate reasoning

1 Introduction

Aggregation functions play important complementary roles in the field of fuzzy logic and its extensions because they are successfully used in many practical applications. Moreover, in this paper we use aggregation functions defined with respect to linear order to create composition using for approximate reasoning is proposed. Approximate reasoning is the process or processes by which a possible imprecise conclusion is deduced from a collection of imprecise premises [7]. We get the Generalized Modus Ponens (GMP):

$$\begin{array}{l} Proposition: \ if\ x\ is\ D\ then\ \ y\ is\ E \\ fact: \ \ x\ is\ D' \\ \hline \qquad\qquad\qquad y\ is\ E', \end{array}$$

where E' is a fuzzy set in the universe Y. The main advantage of the GMP is that we can obtain new information even if D' and D are different. Usually, in the GMP the fuzzy rule: If x is D then y is E is represented by means of a fuzzy relation R on the referential set $X \times Y$. This fuzzy relation R expresses the relationship between the variables x and y involved in the proposition.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_13

In the fuzzy setting, there are situations in which experts have problems building the membership degrees of the elements to the considered fuzzy sets [9]. When this is the case, it is advisable to use extensions of fuzzy sets. One of the most widely used extensions is that interval-valued fuzzy sets (IVFSs) (see for example: [18,20]). In the former sets (IVFSs), the membership degree of each element to the considered IVFS is given by a closed subinterval of the unit interval $[0, 1]$. Taking all these considerations into account, it may happen that experts have problems building the sets D, D' and E appearing in the GMP. For this reason, in this paper we are going to analyze the GMP when the sets D and D' are IVFSs in the referential set X and the sets E and E' are IVFSs in the referential set Y. Clearly, in this setting, the relation R used to represent the rule is an interval-valued fuzzy relation (IVFR) in the referential $X \times Y$ [5,6]. Thus we consider more general method to [11] and we study the composition of IVFRs using interval-valued aggregation functions. We study the conditions under which these new aggregation-based composition preserve some properties, especially the Modus Ponens property. Furthermore, we use generalized concept of aggregation functions for intervals with respect to an admissible order. Moreover, an inference method of the type is developed:

$$E'(y) = \mathcal{B}_{x \in X}(D'(x), R(x, y)),$$

where D' is an IVFS on X, R is the IVFR on $X \times Y$ used to represent the conditional rule and $\mathcal{B}$ is interval-valued aggregation functions defined as

$$\mathcal{B}_{x \in X}(D'(x), R_i(x, y)) = \mathcal{A}_{x \in X}(\mathcal{A}_1(D'(x), R_i(x, y)), \mathcal{A}_2(D'(x), R_i(x, y)))$$

for interval-valued (IV) aggregation functions $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2$ and $i \in \{1, ..., n\}$ for n rules in schema of multiconditional reasoning.

The motivation of this work is the desire to explore the more general algorithm of approximate reasoning by use the composition creating by inspired the General Modus Ponens property with IV aggregations. Moreover, we would like to propose these IV aggregations isotonic with respect to partial or linear order. This work is composed out of the following parts. Firstly, some concepts and results useful in further considerations are recalled (Sect. 2). Next, the general Modus Ponens property is examined (Sect. 3). At the end an algorithm for multiconditional approximate reasoning based on the new aggregation-based composition rules is proposed.

2 Preliminaries

First, we recall the lattice operations and the order for the family of intervals. Let X, Y, Z be finite sets. We denote by L^I the set

$$L^I = \{[x_1, x_2] : \ x_1, x_2 \in [0, 1], \ x_1 \leq x_2\}.$$

Note that if L^I is endowed with the partial order $[x_1, x_2] \leq_P [y_1, y_2]$ if and only if $x_1 \leq y_1$ and $x_2 \leq y_2$, it becomes a complete bounded lattice with the top element given by $\mathbf{1} = [1, 1]$ and the bottom element given by $\mathbf{0} = [0, 0]$. In this lattice, the supremum of any two elements is defined by

$$[x_1, x_2] \vee [y_1, y_2] = [\max(x_1, y_1), \max(x_2, y_2)]$$

and the infimum is defined by

$$[x_1, x_2] \wedge [y_1, y_2] = [\min(x_1, y_1), \min(x_2, y_2)],$$

respectively.

In the next part, we are interested in extending the partial order $\leq_P$ to a linear order (in several ways). We recall the notion of an admissible order, which solve problem of the existence of incomparable elements, a new class of linear orders, called admissible, was introduced in [8] and studied, for example, in [2,16] or [5].

Definition 1 ([8]). *An order $\leq$ in L^I is called admissible if it is linear and satisfies that for all $x, y \in L^I$, such that $x \leq_P y$, then $x \leq y$.*

Simply said, an order $\leq$ on L^I is admissible, if it is linear and refines the order $\leq_P$. In [8] is showed that $\mathbf{1} = [1, 1]$ and $\mathbf{0} = [0, 0]$ are the greatest and the smallest elements of $(L^I, \leq)$, respectively.

Example 1. Admissible (linear) orders on L^I are:
• The Xu and Yager order (see [8]): $[\underline{x}, \overline{x}] \leq_{XY} [\underline{y}, \overline{y}]$ if and only if

$$\underline{x} + \overline{x} < \underline{y} + \overline{y} \text{ or } (\overline{x} + \underline{x} = \overline{y} + \underline{y} \text{ and } \overline{x} - \underline{x} \leq \overline{y} - \underline{y}).$$

• The first lexicographical order (with respect to the first variable), $\leq_{lex1}$ defined as: $[\underline{x}, \overline{x}] \leq_{lex1} [\underline{y}, \overline{y}]$ if and only if

$$\underline{x} < \underline{y} \text{ or } (\underline{x} = \underline{y} \text{ and } \overline{x} \leq \overline{y}).$$

• The second lexicographical order (with respect to the second variable), $\leq_{lex2}$ defined as: $[\underline{x}, \overline{x}] \leq_{lex2} [\underline{y}, \overline{y}]$ if and only if

$$\overline{x} < \overline{y} \text{ or } (\overline{x} = \overline{y} \text{ and } \underline{x} \leq \underline{y}).$$

• Let $K_\alpha : [0, 1]^2 \to [0, 1]$ be the function defined as $K_\alpha(x, y) = \alpha x + (1 - \alpha)y$ for some $\alpha \in [0, 1]$. The order defined as $x \leq_{\alpha\beta} y$ if and only if

$$K_\alpha(\underline{x}, \overline{x}) < K_\alpha(\underline{y}, \overline{y}) \text{ or } (K_\alpha(\underline{x}, \overline{x}) = K_\alpha(\underline{y}, \overline{y}) \text{ and } K_\beta(\underline{x}, \overline{x}) \leq K_\beta(\underline{y}, \overline{y}))$$

is an admissible order for $\alpha \neq \beta$ and $x, y \in L^I$.

We know that orders $\leq_{XY}, \leq_{Lex1}$ and $\leq_{Lex2}$ are special cases of $\leq_{\alpha\beta}$.

In the further part of the work we use the label for partial or linear order $\leq$, for the partial order will be used $\leq_P$ and for the linear order will be used with the appropriate linear order $\leq_L$.

We recall the concept of an aggregation function on L^I, which is a crucial definition for this paper.

Definition 2. (cf. [2,4,12]). *An operation $\mathcal{A} : (L^I)^n \to L^I$ is called an interval-valued (IV) aggregation function if it is increasing with respect to the partial or linear order $\leq$ and*

$$\mathcal{A}(\underbrace{\mathbf{0}, ..., \mathbf{0}}_{n\times}) = \mathbf{0}, \quad \mathcal{A}(\underbrace{\mathbf{1}, ..., \mathbf{1}}_{n\times}) = \mathbf{1}.$$

A relevant class of aggregation functions is that of representable aggregation functions.

So, we may build [10] an IV representable aggregation function $\mathcal{A} : (L^I)^2 \to L^I$ with respect to the partial order $\leq_P$ if there exist two (real) aggregation functions $A_1, A_2 : [0,1]^2 \to [0,1]$ such that, for every $[x_1, x_2], [y_1, y_2] \in L^I$, $A_1 \leq A_2$ it holds that

$$\mathcal{A}([x_1, x_2], [y_1, y_2]) = [A_1(x_1, y_1), A_2(x_2, y_2)].$$

Operations $\wedge$ and $\vee$ on L^I define representable aggregation functions on L^I, with $A_1 = A_2 = \min$ in the first case and $A_1 = A_2 = \max$ in the second. Moreover, many other examples of representable aggregation functions with respect to $\leq_P$ may be considered, such as:

- the representable product $\mathcal{A}_P([x_1, x_2], [y_1, y_2]) = [x_1 y_1, x_2 y_2]$,
- the representable arithmetic mean $\mathcal{A}_{mean}([x_1, x_2], [y_1, y_2]) = [\frac{x_1+y_1}{2}, \frac{x_2+y_2}{2}]$,
- the representable geometric mean $\mathcal{A}_g([x_1, x_2], [y_1, y_2]) = [\sqrt{x_1 y_1}, \sqrt{x_2 y_2}]$,
- the representable product-mean $\mathcal{A}_{P,mean}([x_1, x_2], [y_1, y_2]) = [x_1 y_1, \frac{x_2+y_2}{2}]$.

Whereas, the function $\mathcal{A} : (L^I)^2 \to L^I$,

$$\mathcal{A}(x, y) = \begin{cases} [1, 1], & \textit{if } (x, y) = ([1, 1], [1, 1]) \\ [0, A(x_1, y_2)], & \textit{otherwise} \end{cases}$$

for $x, y \in L^I$ is an IV aggregation function on L^I with respect to $\leq_P$ or $\leq_{Lex1}$, but non-representable, where A is a fuzzy aggregation function. Moreover, $\mathcal{A}_{mean}$ is an aggregation function with respect to $\leq_{\alpha\beta}$ [2].

In the subsequent part of this work we will use following properties of aggregation functions with respect to partial or admissible order, with interval-valued (IV) negation N_{IV} that is decreasing with respect to $\leq$ and $N_{IV}(\mathbf{1}) = \mathbf{0}$, $N_{IV}(\mathbf{0}) = \mathbf{1}$ [3].

Definition 3. (cf. [1,13,17]). *IV aggregation functions $\mathcal{A}, \mathcal{B} : (L^I)^2 \to L^I$ for $x, y, z, t \in L^I$ are said to be:*

(i) commutative, if

$$\mathcal{A}(x, y) = \mathcal{A}(y, x),$$

(ii) associative, if

$$\mathcal{A}(\mathcal{A}(x, y), z) = \mathcal{A}(x, \mathcal{A}(y, z)),$$

(iii) idempotent, if

$$\mathcal{A}(x, x) = x,$$

(iv) $\mathcal{A}$ *is said to have* $e \in L^I$ *an neutral element, if*

$$\mathcal{A}(x, e) = \mathcal{A}(e, x) = x,$$

(v) $\mathcal{A}$ *is said to have* $z \in L^I$ *an absorbing element, if*

$$\mathcal{A}(x, z) = \mathcal{A}(z, x) = z,$$

(vi) $\mathcal{A}$ *distributive with respect to* $\mathcal{B}$*, if*

$$\mathcal{A}(x, \mathcal{B}(y, t)) = \mathcal{B}(\mathcal{A}(x, y), \mathcal{A}(x, t)),$$

(vii) $\mathcal{A}$ *dominates* $\mathcal{B}$ *(*$\mathcal{A} \gg \mathcal{B}$*), if*

$$\mathcal{A}(\mathcal{B}(x, y), \mathcal{B}(z, t)) \geq \mathcal{B}(\mathcal{A}(x, z), \mathcal{A}(y, t)),$$

(viii) $\mathcal{A}$, $\mathcal{B}$ *are modular, if*

$$z \leq x \Rightarrow \mathcal{A}(x, \mathcal{B}(y, z)) = \mathcal{B}(\mathcal{A}(x, y), z),$$

(ix) $\mathcal{A}$ *satisfies the Non-Contradiction principle w.r.t.* N_{IV} *(NC(N)), if*

$$\mathcal{A}(x, N_{IV}(x)) = \boldsymbol{0},$$

(x) $\mathcal{A}$ *satisfies the Excluded-Middle principle w.r.t.* N_{IV} *(EM(N)), if*

$$\mathcal{A}(x, N_{IV}(x)) = \boldsymbol{1},$$

(xi) $\mathcal{A}$ *is a conjunctor (disjunctor) if and only if it satisfies the condition*

$$\mathcal{A}(\boldsymbol{0}, \boldsymbol{1}) = \mathcal{A}(\boldsymbol{1}, \boldsymbol{0}) = \boldsymbol{0} \;\; (\mathcal{A}(\boldsymbol{0}, \boldsymbol{1}) = \mathcal{A}(\boldsymbol{1}, \boldsymbol{0}) = \boldsymbol{1}),$$

(xii) $\mathcal{A}$ *is conjunctive (disjunctive, averaging), if*

$$\mathcal{A} \leq min \;\; (\mathcal{A} \geq max, \; min \leq \mathcal{A} \leq max).$$

3 General Modus Ponens Property

Inference schemes in approximate reasoning are usually based on the Modus Ponens that is carried out through the well known Compositional Rule of Inference (CRI) of Zadeh, based on the sup T-composition, where T is a t-norm. Thus, if I is a fuzzy implication function and T is a t-norm, the Modus Ponens property for I with respect to T (T-conditionality) becomes the functional inequality:

$$T(x, I(x, y)) \leq y, \quad x, y \in [0, 1], \tag{1}$$

where T is continuous fuzzy t-norm and I a fuzzy implication function. Moreover, some generalizations of this T-conditionality have been studied. One of these generalizations is based on uninorms obtaining the so-called RU-implications and (U, N)-implications [15], and in [14] was examined generalization based on continuous uninorm U instead of a t-norm T in study U-conditionality of implication I.

In this section we will examine more general aggregation functions instead of t-norms and we will use the function $I_{A,N}$, such that

$$I_{A,N}(x, y) = A(N(x), y),$$

which is (A, N)-implication if and only if aggregation A has absorbing element 1 and N a negation function (see [17]).

The approximate reasoning is the process or processes by which a possible imprecise conclusion is deduced from a collection of imprecise premises [7], thus instead of (1) we have

$$T_{IV}(x, I_{IV}(x, y)) \leq y, \quad x, y \in L^I, \tag{2}$$

where T_{IV} is continuous interval-valued (IV) t-norm and I_{IV} an interval-valued (IV) implication function.

The previous inequality is known as the Modus Ponens property. We want to generalize of the definition of Modus Ponens using two aggregations instead of t-norm and implication based on IV negation, that is decreasing with respect to $\leq$ and $N_{IV}(\mathbf{1}) = \mathbf{0}$, $N_{IV}(\mathbf{0}) = \mathbf{1}$.

Implication and aggregation functions play important complementary roles in the field of fuzzy logic and its extensions. Both have been intensively investigated since the early 1980s, revealing a tight relationship between them.

Aggregation functions are successfully used in many practical applications, and the interest in them is unceasingly growing. Aggregation and implication functions appear to have a close relation, mainly realized via negation functions, which model the logical negation within the fuzzy framework.

Thus we more generalize the Modus Ponens property further and we will use second interval-valued fuzzy aggregation function instead of t-norm T_{IV}.

Definition 4. *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be IV aggregation functions and N_{IV} be IV negation. It is said that $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$ ($\mathcal{A}_1$-conditionality), if*

$$\mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) \leq y \quad \text{for all} \;\; x, y \in L^I. \tag{3}$$

Proposition 1. *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be IV aggregation functions with respect to the same order $\leq$. Then $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$, if at least one of following conditions is fulfilled:*

1. *$\mathcal{A}_1$, $\mathcal{A}_2$ be IV aggregation functions with neutral element **1**.*
2. *$\mathcal{A}_1$, $\mathcal{A}_2$ are conjunctive.*
3. *$\mathcal{A}_1$ fulfill NC(N) and ($\mathcal{A}_2$ is conjunctive or ($\mathcal{A}_2$ is averaging or has the neutral element **0** by modularity of $\mathcal{A}_1$,$\mathcal{A}_2$)).*

Proof. First and second conditions are obvious. We prove third condition. If $\mathcal{A}_2 \leq \min$ and $\mathcal{A}_1(x, N_{IV}(x)) = \mathbf{0}$, then

$$\mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) \leq \mathcal{A}_1(x, \min(N_{IV}(x), y)) \leq \mathcal{A}_1(x, N_{IV}(x)) = \mathbf{0} \leq y.$$

By modularity of $\mathcal{A}_1$ and $\mathcal{A}_2$ and $\mathcal{A}_2 \leq \max$, $\mathcal{A}_1(x, N_{IV}(x)) = \mathbf{0}$ we conclude

$$\mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) = \mathcal{A}_2(\mathcal{A}_1(x, N_{IV}(x)), y) \leq \max(\mathbf{0}, y) = y.$$

Moreover, by modularity of $\mathcal{A}_1$ and $\mathcal{A}_2$ and by $\mathbf{0}$ as the neutral element of $\mathcal{A}_2$ by condition $\mathcal{A}_1(x, N_{IV}(x)) = \mathbf{0}$ we obtain

$$\mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) = \mathcal{A}_2(\mathcal{A}_1(x, N_{IV}(x)), y) \leq \mathcal{A}_2(\mathbf{0}, y) = y,$$

what finished the proof of the Proposition 1 (for 3. case).

Example 2. IV aggregation functions with respect to the order $\leq_P$:

$$\mathcal{A}_2(x, y) = \begin{cases} [0,0], & if\ [\underline{y}, \overline{y}] \leq [1-\overline{x}, 1-\underline{x}] \\ [1,1], & otherwise \end{cases}$$

and

$$\mathcal{A}_1(x, y) = \begin{cases} \min(x,y), & if\ [\underline{x}+\underline{y}, \overline{x}+\overline{y}] > [1,1] \\ [0,0], & otherwise \end{cases}$$

fulfilling $\mathcal{A}_1(x, N_{IV}(x)) = \mathbf{0}$ by IV negation $N_{IV}(x) = [1-\overline{x}, 1-\underline{x}]$ satisfies condition (3) (Proposition 1, for 3. case).

By monotonicity of aggregation functions we observe following conditions.

Proposition 2. *Let $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ and $\mathcal{A}_4$ be IV aggregation functions with respect to the same order $\leq$.*
- *If $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$ and $\mathcal{A}_3 \leq \mathcal{A}_1$, then $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_3$.*
- *If $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$ and $\mathcal{A}_4 \leq \mathcal{A}_2$, then $\mathcal{A}_4$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$.*

Except that, the Modus Ponens property implies following results.

Proposition 3. *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be IV aggregation functions with respect to the same order $\leq$. If $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$, then must be $\mathcal{A}_1(\mathbf{1}, \mathbf{0}) = \mathbf{0}$.*

Proof. Just taking $x = \mathbf{1}$ in (3) we obtain

$$\mathcal{A}_1(\mathbf{1}, \mathcal{A}_2(\mathbf{0}, \mathbf{0})) \leq \mathbf{0} \Leftrightarrow \mathcal{A}_1(\mathbf{1}, \mathbf{0}) = \mathbf{0}.$$

Proposition 4. *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be interval-valued fuzzy aggregation functions with respect to the same order $\leq$. If $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$, then*

$$\mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) = y$$

for neutral element e_1 of $\mathcal{A}_1$, e_2 of $\mathcal{A}_2$ and $e_1 \leq x$ and $e_2 \leq N_{IV}(x)$.

Proof. Let $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_1$. Then

$$y = \mathcal{A}_1(e_1, \mathcal{A}_2(e_2, y)) \leq \mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) \leq y.$$

As we know, there are different construction methods for building new aggregation functions from data. Some of the most usual ones are those based on either composition or transformation (see [17]).

Proposition 5 (cf. [17]). *If $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2$ are IV aggregation functions with respect to the same order $\leq$, then the function $\mathcal{B} = \mathcal{A}(\mathcal{A}_1, \mathcal{A}_2) : (L^I)^2 \to L^I$, defined as*

$$\mathcal{B}(x, y) = \mathcal{A}(\mathcal{A}_1(x, y), \mathcal{A}_2(x, y)) \tag{4}$$

for any $x, y \in L^I$ is an IV aggregation function.

We will say that $\mathcal{A}$ preserves some property P if $\mathcal{A}(\mathcal{A}_1, \mathcal{A}_2)$ fulfills P whenever $\mathcal{A}_1$ and $\mathcal{A}_2$ fulfill it. We examine a preservation of Modus Ponens property by this composition.

Proposition 6. *Let $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ be IV aggregation functions and N_{IV} be IV negation function with respect to the same order $\leq$. If $\mathcal{A}_1$ or $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_3$ and $\mathcal{A}$ is conjunctive, then $\mathcal{B} = \mathcal{A}(\mathcal{A}_1, \mathcal{A}_2)$ also satisfies the Modus Ponens property with respect to $\mathcal{A}_3$.*

Proof. Let $\mathcal{A}_1$ satisfies the Modus Ponens property with respect to $\mathcal{A}_3$ and $\mathcal{A} \leq \min$. Then

$$\mathcal{A}_3(x, \mathcal{B}(N_{IV}(x), y)) = \mathcal{A}_3(x, \mathcal{A}(\mathcal{A}_1(N_{IV}(x), y), \mathcal{A}_2(N_{IV}(x), y))) \leq$$

$$\mathcal{A}_3(x, \min(\mathcal{A}_1(N_{IV}(x), y), \mathcal{A}_2(N_{IV}(x), y))) \leq \mathcal{A}_3(x, \mathcal{A}_1(N_{IV}(x), y) \leq y.$$

If $\mathcal{A}_2$ satisfies the Modus Ponens property with respect to $\mathcal{A}_3$ proof is similar.

Furthermore, the aggregation $\mathcal{B}$ has following adequate properties:

Proposition 7 (cf. [17]). *Let $\mathcal{B} = \mathcal{A}(\mathcal{A}_1, \mathcal{A}_2)$ be an aggregation function built as in (4). Then:*

1. *$\mathcal{B}$ preserves commutative, the absorbing element, concjunctive, disjunctive, concjunctor, disjunctor, EM(N) or NC(N) property.*
2. *If $\mathcal{A}$ is averaging aggregation function, then $\mathcal{B}$ preserves the neutral element of $\mathcal{A}_1$ and $\mathcal{A}_2$.*

Proposition 8 (cf. [19]). *Let $\mathcal{D}$ be an aggregation function. Then for any $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2$ dominate $\mathcal{D}$, where $\mathcal{A}$ is idempotent, also $\mathcal{B} = \mathcal{A}(\mathcal{A}_1, \mathcal{A}_2) \gg \mathcal{D}$.*

Now we propose the other very interesting composition from the point of view the Modus Ponens property.

Definition 5. *Let N_{IV} be IV negation and $\mathcal{A}_1, \mathcal{A}_2$ be IV aggregation functions with respect to the same order $\leq$. Then the operation (composition) $\mathcal{C} : (L^I)^2 \to L^I$ is defined as follows*

$$\mathcal{C}(x, y) = \mathcal{A}_1(x, \mathcal{A}_2(N_{IV}(x), y)) \tag{5}$$

for $x, y \in L^I$.

Proposition 9. *Let N_{IV} be IV negation, $\mathcal{A}_1, \mathcal{A}_2$ be IV aggregation functions with respect to the same order $\leq$ and $\mathcal{C}$ be the operation defined in Definition 5. Then*
If $\mathcal{A}_1$ and $\mathcal{A}_2$ are $\mathcal{A}_3$-conditionality and $\mathcal{A}_2$ is conjunctive aggregation function, then $\mathcal{C}$ also fulfill $\mathcal{A}_3$-conditionality.

Proof.

$$\mathcal{A}_3(x, \mathcal{C}(N_{IV}(x), y)) = \mathcal{A}_3(x, \mathcal{A}_1(N_{IV}(x), \mathcal{A}_2(N_{IV}(N_{IV}(x)), y))) \leq$$
$$\mathcal{A}_2(N_{IV}(N_{IV}(x)), y) \leq y.$$

Moreover, we consider basic properties of presented composition $\mathcal{C}$.

Proposition 10. *Let $\mathcal{C}$ be the operation defined in Definition 5, N_{IV} be IV negation and $\mathcal{A}_1, \mathcal{A}_2$ be IV aggregation functions with respect to the same order $\leq$. Then*

1. *$\mathcal{C}$ preserve the absorbing element of $\mathcal{A}_1$ and $\mathcal{A}_2$.*
2. *$\mathcal{C}$ preserve the neutral element of $\mathcal{A}_1$ for dual aggregations $\mathcal{A}_1$ and $\mathcal{A}_2$ (i.e. $\mathcal{A}_1(x, y) = N_{IV}\mathcal{A}_2(N_{IV}(x), N_{IV}(y)))$ by strong IV negation N_{IV}.*
3. *If $\mathcal{A}_1, \mathcal{A}_2$ are conjnctive (disjunctive) aggregation functions, then also $\mathcal{C}$ is conjnctive (disjunctive) operation.*
4. *If $\mathcal{A}_1$ fulfill NC(N) (EM(N)) property and $\mathcal{A}_2$ is an averaging aggregation function, then also $\mathcal{C}$ fulfill NC(N) (EM(N)) property.*
5. *$\mathcal{C}(\mathbf{0}, \mathbf{0}) = \mathbf{0}$, if $\mathcal{A}_1$ or $\mathcal{A}_2$ is a conjunctor.*
6. *$\mathcal{C}(\mathbf{1}, \mathbf{1}) = \mathbf{1}$, if $\mathcal{A}_1$ or $\mathcal{A}_2$ is a disjunctor.*
7. *$\mathcal{C}$ satisfy left side isotonicity for both variable and right side isotonicity for the first variable and right side antytonicity for the second variable.*

Proof. The proof we can directly obtain from the Definitions 3 and 5.

Except that, directly by left side isotonicity of $\mathcal{C}$ built as in (5) we obtain for averaging IV aggregation function $\mathcal{A}_2$:

$$(\mathcal{A}_1 \ \ fulfill \ \ NC(N)) \Rightarrow (y \leq N_{IV}(x) \Rightarrow \mathcal{C}(x, y) = \mathbf{0});$$
$$(\mathcal{A}_1 \ \ fulfill \ \ EM(N)) \Rightarrow (y \geq N_{IV}(x) \Rightarrow \mathcal{C}(x, y) = \mathbf{1}).$$

Proposition 11. *Let $\mathcal{A}_1, \mathcal{A}_2, \mathcal{D}$ be IV aggregation functions and $\mathcal{D}$ be N-stable, i.e. $N_{IV}(\mathcal{D}(x, y)) = \mathcal{D}(N_{IV}(x), N_{IV}(y))$ for IV negation function N_{IV}.*
If $\mathcal{A}_1 \gg \mathcal{D}$ and $\mathcal{A}_2 \gg \mathcal{D}$, then $\mathcal{C} \gg \mathcal{D}$, where $\mathcal{C}$ satisfy (5).

Proof. Let $\mathcal{A}_1 \gg \mathcal{D}$ and $\mathcal{A}_2 \gg \mathcal{D}$ and $\mathcal{C}(x_i, y_i) = \mathcal{A}_1(x_i, \mathcal{A}_2(N_{IV}(x_i), y_i))$ for $i \in \{1, 2\}$. Then

$$\mathcal{D}(\mathcal{C}(x_1, y_1), \mathcal{C}(x_2, y_2)) =$$
$$\mathcal{D}(\mathcal{A}_1(x_1, \mathcal{A}_2(N_{IV}(x_1), y_1)), \mathcal{A}_1(x_2, \mathcal{A}_2(N_{IV}(x_2), y_2))) \leq$$
$$\mathcal{A}_1(\mathcal{D}(x_1, x_2), \mathcal{D}(\mathcal{A}_2(N_{IV}(x_1), y_1), \mathcal{A}_2(N_{IV}(x_2), y_2)) \leq$$
$$\mathcal{A}_1(\mathcal{D}(x_1, x_2), \mathcal{A}_2(\mathcal{D}(N_{IV}(x_1), N_{IV}(x_2)), \mathcal{D}(y_1, y_2)) =$$
$$\mathcal{C}(\mathcal{D}(x_1, x_2), \mathcal{D}(y_1, y_2)),$$

what finished the proof.

4 Algorithm for Interval-Valued Multiconditional Approximate Reasoning

The general schema of interval-valued multiconditional reasoning has a form:

$$\begin{array}{l} R_1: \; if \; x \; is \; \mathcal{D}_1 \; then \; y \; \; is \; \mathcal{E}_1 \\ R_2: \; if \; x \; is \; \mathcal{D}_2 \; then \; y \; is \; \mathcal{E}_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ R_n: \; if \; x \; is \; \mathcal{D}_n \; then \; y \; is \; \mathcal{E}_n \\ fact: \quad x \; is \; \mathcal{D}' \\ \hline \qquad\qquad\qquad y \; is \; \mathcal{E}', \end{array}$$

where $\mathcal{D}_1, \cdots, \mathcal{D}_n, \mathcal{D}' \in \mathcal{IVFS}(X)$, $\mathcal{E}_1, \cdots, \mathcal{E}_n, \mathcal{E}' \in \mathcal{IVFS}(Y)$.

The following method to determine $\mathcal{E}'$ is proposed. We omit the aspect of fuzzification and defuzzification in presented algorithm of approximate reasoning.

Algorithm. ApprReasComp:

Inputs: Premises $\mathcal{D}_1, \cdots, \mathcal{D}_n, \mathcal{D}' \in \mathcal{IVFS}(X)$; Conclusions $\mathcal{E}_1, \cdots, \mathcal{E}_n \in \mathcal{IVFS}(Y)$; The composition consistent with the formula (5), i.e. $\mathcal{C} = (\mathcal{A}_1, \mathcal{A}_2)$ by conjunctive interval-valued fuzzy aggregation function $\mathcal{A}_2$ satisfy:

$\mathcal{A}_2(\mathbf{1}, \mathbf{0}) = \mathcal{A}_2(\mathbf{0}, \mathbf{1}) = \mathbf{1}$; The IV aggregation function $\mathcal{B}$ created by (4) with IV aggregation functions $\mathcal{A}, \mathcal{A}_3, \mathcal{A}_4, \mathcal{A}_5$ and $\mathcal{A}$ conjunctive and $\mathcal{A}_5 \geq max$;

Outputs: $\mathcal{E}'$

1. For each rule, the associated interval-valued fuzzy relation R_i is built, where $R_i \in \mathcal{IVFR}(X \times Y)$ for $i = 1, \cdots, n$ and
$$R_i(x, y) = \mathcal{C}(\mathcal{D}_i(x), \mathcal{E}_i(y));$$
2. The interval-valued aggregation function $\mathcal{B} = \mathcal{A}(\mathcal{A}_3, \mathcal{A}_4)$ is taken;
3. For $i = 1, \cdots, n$ is calculated (GMP):
$$\mathcal{E}'_i(y) = \mathcal{B}_{x \in X}(\mathcal{D}'(x), R_i(x, y)),$$
where
$$\mathcal{B}_{x \in X}(\mathcal{D}'(x), R_i(x, y) = \mathcal{A}_{x \in X}(\mathcal{A}_3(\mathcal{D}'(x), R_i(x, y)), \mathcal{A}_4(\mathcal{D}'(x), R_i(x, y)));$$
4. Compute: $\mathcal{E}' = \mathcal{A}_5(\mathcal{E}'_i)$ for each $i = 1, ..., n$.

Preliminary calculations give satisfactory results because, by comparing the algorithm proposed above with algorithms using t-norms and t-conorms, we observe that the received fuzzy set in the new method has higher intervals, That is, it represents a lower degree of uncertainty and, as a result, gives you the possibility of better precision in the application.

5 Conclusions and Future Research

Inference schemes in approximate reasoning are based on the Modus Ponens property, also called $\mathcal{A}$-conditionality. Thus, IV aggregation functions used in the inference process of any IV fuzzy rule based system are required to satisfy this property, which becomes essential in general approximate reasoning. The conditions under which aggregates meet $\mathcal{A}$-conditionality (General Modus Ponens property) are presented. In the future we would like present practical example using presented algorithm and compare it with other known algorithms.

References

1. Aczél, J.: Lectures on Functional Equations and Their Applications. Academic Press, New York (1966)
2. Asiain, M.J., Bustince, H., Bedregal, B., Takáć, Z., Baczyński, M., Paternain, D., Dimuro, G.: About the Use of Admissible Order for Defining Implication Operators, pp. 353–362. Springer International Publishing, Cham (2016)
3. Asiain, M.J., Bustince, H., Mesiar, R., Kolesarova, A., Takac, Z.: Negations with respect to admissible orders in the interval-valued fuzzy set theory. IEEE Trans. Fuzzy Syst. doi:10.1109/TFUZZ.2017.2686372
4. Beliakov, G., Bustince, H., Calvo, T.: A practical Guide to Averaging Functions. In: Studies in Fuzziness and Soft Computing, vol. 329. Springer (2016)
5. Bentkowska, U., Bustince, H., Jurio, A., Pagola, M., Pękala, B.: Decision making with an interval-valued fuzzy preference relation and admissible orders. Appl. Soft Comput. **35**, 792–801 (2015)
6. Bustince, H., Burillo, P.: Interval-valued fuzzy relations in a set structures. J. Fuzzy Math. **4**(4), 765–785 (1996)
7. Bustince, H., Burillo, P.: Mathematical analysis of interval-valued fuzzy relations: application to approximate reasoning. Fuzzy Sets Syst. **113**, 205–219 (2000)
8. Bustince, H., Fernandez, J., Kolesárová, A., Mesiar, R.: Generation of linear orders for intervals by means of aggregation functions. Fuzzy Sets Syst. **220**, 69–77 (2013)
9. Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J., Xu, Z., Bedregal, B., Montero, J., Hagras, H., Herrera, F., De Baets, B.: A historical account of types of fuzzy sets and their relationships. IEEE Trans. Fuzzy Syst. (2015). doi:10.1109/TFUZZ.2015.2451692
10. Deschrijver, G.: Quasi-arithmetic means and OWA functions in interval-valued and Atanassov's intuitionistic fuzzy set theory. In: EUSFLAT-LFA 2011, Aix-les-Bains, France, pp. 506–513 (2011)
11. Elkano, M., Sanz, J.A., Galar, M., Pękala, B., Bentkowska, U., Bustince, H.: Composition of interval-valued fuzzy relations using aggregation functions. Inform. Sci. doi:10.1016/j.ins.2016.07.048
12. Komorníková, M., Mesiar, R.: Aggregation functions on bounded partially ordered sets and their classification. Fuzzy Sets Syst. **175**, 48–56 (2011)
13. Mas, M., Mayor, G., Torrens, J.: The modularity condition for uninorms and t-operators. Fuzzy Sets Syst. **126**, 207–218 (2002)
14. Mas, M., Monserrat, M., Ruiz-Aguilera, D., Torrens, J.: On generalization of the Modus Ponens: U-conditionality. In: Carvalho, J.P., et al. (eds.) IPMU 2016, Part I. CCIS, vol. 610, pp. 387–398 (2016)

15. Mas, M., Monserrat, M., Ruiz-Aguilera, D., Torrens, J.: RU and (U, N)-implications satisfying Modus Ponens. Int. J. Approximate Reasoning **73**, 123–137 (2016)
16. De Miguel, L., Bustince, H., Fernandez, J., Induráin, E., Kolesárová, A., Mesiar, R.: Construction of admissible linear orders for intervalvalued Atanassov intuitionistic fuzzy sets with an application to decision making. Inform. Fus. **27**, 189–197 (2016)
17. Pradera, A., Beliakov, G., Bustince, H., De Baets, B.: A review of the relationships between implication, negation and aggregation functions from the point of view of material implication. Inf. Sci. **329**, 357–380 (2016)
18. Sambuc, R.: Fonctions ϕ-floues: Application á l'aide au diagnostic en pathologie thyroidienne. Ph.D. Thesis, Universit*é* de Marseille, France (1975)
19. Saminger, S., Mesiar, R., Bodenhoffer, U.: Domination of aggregation operators and preservation of transitivity. Int. J. Unc. Fuzz. Knowl. Based Syst. **10**, 11 (2002). http://dx.doi.org/10.1142/S0218488502001806
20. Zadeh, L.A.: A theory of approximate reasoning. In: Hayes, J.E., Michie, D., Mikulich, L.I. (eds.) Machine Intelligence 9, pp. 149–194 (1979)

General Preference Structure with Uncertainty Data Present by Interval-Valued Fuzzy Relation and Used in Decision Making Model

Barbara Pękala(✉)

Faculty of Mathematics and Natural Sciences,
Interdisciplinary Centre for Computational Modelling,
University of Rzeszów, Pigonia 1A, 35–310 Rzeszów, Poland
bpekala@ur.edu.pl

Abstract. Interval-valued fuzzy relations can be interpreted as a tool that may help to model in a better way imperfect information, especially under imperfectly defined facts and imprecise knowledge. Preference structures are of great interest nowadays because of their applications. From a weak preference relation derive the following relations: strict preference, indifference and incomparability, which by aggregations and negations are created and examined in this paper. Moreover, we propose the algorithm of decision making by using new preference structure.

Keywords: Interval-valued fuzzy relations · Preference relations · Reciprocity property

1 Introduction

Interval-Valued Fuzzy Relations (IVFRs) [35] form a generalization of the concept of a fuzzy relation [34] and represent uncertainties, systematic or random uncertainties. Fuzzy sets and relations are applied in diverse areas, e.g. in group decision making [8,21,24,33]. In recent applications to image processing [2] or classification [29,30] it has been proven that, under some circumstances, the use of IVFSs together with the total order provide results that are better than their fuzzy counterparts. Many decision making processes take place in an environment in which the information is not precisely known. As a consequence, experts may feel more comfortable using an interval number rather than an exact crisp numerical value to represent their preference. The concept of a preference relation has been studied by many authors, both in crisp or fuzzy environments [9,27]. But interval-valued fuzzy preference relations (IVFRs) can be considered as an appropriate representation format to capture experts' uncertain preference information. Diverse properties of IVFRs, also in the case of interval-valued fuzzy reciprocal relations, have been studied by a range of authors [17,22]. The assumption of reciprocity is often used for a preference relation both in the interval-valued [2] and classical fuzzy environment [9].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_14

Our main goal this paper is to examine certain aspect of decision making problem based on preference relations built by aggregations and reciprocity property built by negation function, which means that instead of using classical negation in definition of reciprocity, we apply negations. Reciprocity appears in preference relations as a natural assumption. We present generalisation of the concept of defining model of three relations: strict preference, indifference and incomparability, corresponding to preference relation.

This work is composed of the following parts. Firstly, some concepts and results useful in further considerations are reminded (Sect. 2). Next, results connected with preference structure and some properties are presented (Sect. 3). Finally, we present the algorithm of decision making problem with new strict preference, indifference and incomparability relations.

2 Preliminaries

Necessary for our considerations will be the negation function. Thus, we give the definition of negation functions on the unit interval $[0, 1]$ and definition of the dual function which refers also to aggregation functions.

Definition 1 ([19]). *A fuzzy negation function is a decreasing function $N : [0,1] \to [0,1]$ verifying the boundary conditions $N(0) = 1$ and $N(1) = 0$. Strictly decreasing and continuous negation functions are known as strict negations, whereas involutive negation functions (i.e., those for all $x \in [0,1]$ verifying $N(N(x)) = x$) are known as strong negations (and constitute a subclass of strict negations).*

Typical examples of negation functions are:

- $N(x) = 1-x$, which is a strong negation and is called the classical or standard negation;
- $N_S^{\lambda}(x) = \frac{1-x}{1+\lambda x}$, the Sugeno family of fuzzy (strong) negations, where $\lambda \in (-1, \infty)$, and for $\lambda = 0$ we get the classical fuzzy negation.

Definition 2 (cf. [26]). *Let $F : [0,1]^n \to [0,1]$, N be a strong fuzzy negation. The N-dual of F is the function*

$$F^N(x_1, \ldots, x_n) = N(F(N(x_1), \ldots, N(x_n))), \;\; x_1, \ldots, x_n \in [0,1]. \qquad (1)$$

We recall the notion of operations and the order in the family of intervals.

$$L^I = \{[x_1, x_2] : \; x_1, x_2 \in [0,1], \; x_1 \leq x_2\}.$$

Note that L^I endowed with the partial order $[x_1, x_2] \leq_{L^I} [y_1, y_2]$ if and only if $x_1 \leq y_1$ and $x_2 \leq y_2$ is a complete bounded lattice with the top element given by $\mathbf{1} = [1, 1]$ and the bottom element given by $\mathbf{0} = [0, 0]$. In this lattice, the supremum and the infimum of any two elements is defined respectively by
$[x_1, x_2] \vee [y_1, y_2] = [\max(x_1, y_1), \max(x_2, y_2)]$,
$[x_1, x_2] \wedge [y_1, y_2] = [\min(x_1, y_1), \min(x_2, y_2)]$.

Now, we give the definition of interval-valued negation function.

Definition 3 ([2]). *An interval-valued (IV) negation is a function $N_{IV} : L^I \to L^I$ that is decreasing with respect to $\leq_{L^I}$ with $N_{IV}(\boldsymbol{1}) = \boldsymbol{0}$ and $N_{IV}(\boldsymbol{0}) = \boldsymbol{1}$. An IV negation is said to be involutive if it fulfils $N_{IV}(N_{IV}(x)) = x$ for any $x \in L^I$ and is known as strong negation.*

Theorem 1 (cf. [12]). *Let $x = [\underline{x}, \overline{x}] \in L^I$. N_{IV} is an involutive IV negation if and only if there exists a strong fuzzy negation N such that $N_{IV}([\underline{x}, \overline{x}]) = [N(\overline{x}), N(\underline{x})]$.*

The above statement holds true also for fuzzy negations, not only involutive ones.

Moreover, we may define aggregation function on a set of intervals by following way:

Definition 4 (cf. [7], p. 6, [20]). *An operation $\mathcal{A} : (L^I)^n \to L^I$ is called an aggregation function if it is increasing with respect to the order $\leq_{L^I}$ and*

$$\mathcal{A}(\underbrace{\boldsymbol{0}, ..., \boldsymbol{0}}_{n\times}) = \boldsymbol{0}, \quad \mathcal{A}(\underbrace{\boldsymbol{1}, ..., \boldsymbol{1}}_{n\times}) = \boldsymbol{1}.$$

A relevant class of aggregation functions is that of representable aggregation functions.

Definition 5 (cf. [13–15]). *Let $\mathcal{A} : (L^I)^2 \to L^I$ be an aggregation function. $\mathcal{A}$ is said to be a representable aggregation function if there exist two fuzzy aggregation functions $A_1, A_2 : [0,1]^2 \to [0,1]$, $A_1 \leq A_2$ such that, for every $[x_1, x_2], [y_1, y_2] \in L^I$ it holds that*

$$\mathcal{A}([x_1, x_2], [y_1, y_2]) = [A_1(x_1, y_1), A_2(x_2, y_2)].$$

We observe that both $\wedge$ and $\vee$ in L^I define representable aggregation functions in L^I, with $A_1 = A_2 = \min$ in the first case and $A_1 = A_2 = \max$ in the second. Moreover, many other examples may be considered, such as:

- the representable arithmetic mean $\mathcal{A}_{mean}([x_1, x_2], [y_1, y_2]) = [\frac{x_1+y_1}{2}, \frac{x_2+y_2}{2}]$,
- the representable weighted mean with $w_1 + w_2 = 1$, $w_1, w_2 \in [0,1]$ $\mathcal{A}_{wmean}([x_1, x_2], [y_1, y_2]) = [w_1 x_1 + w_2 y_1, w_1 x_2 + w_2 y_2]$,
- the representable geometric mean $\mathcal{A}_g([x_1, x_2], [y_1, y_2]) = [\sqrt{x_1 y_1}, \sqrt{x_2 y_2}]$,
- the representable weighted geometric mean with $w_1 + w_2 = 1$, $w_1, w_2 \in [0,1]$ $\mathcal{A}_{wg}([x_1, x_2], [y_1, y_2]) = [x_1^{w_1} y_1^{w_2}, x_2^{w_1} y_2^{w_2}]$,
- the representable product-mean $\mathcal{A}_{P,mean}([x_1, x_2], [y_1, y_2]) = [x_1 y_1, \frac{x_2+y_2}{2}]$.

In the subsequent part of this work we will use following properties of aggregation functions with respect to the order $\leq_{L^I}$.

Definition 6 (cf. [1,26]). *Let $\mathcal{A}, \mathcal{B} : (L^I)^2 \to L^I$ be aggregation functions, $x, y, z, t \in L^I$:*

- $\mathcal{A}$ *is symmetric, if* $\mathcal{A}(x,y) = \mathcal{A}(y,x)$,
- $\mathcal{A}$ *is idempotent, if* $\mathcal{A}(x,x) = x$,
- $\mathcal{A}$ *is bisymmetric, if* $\mathcal{A}(\mathcal{A}(x,y),\mathcal{A}(z,t)) = \mathcal{A}(\mathcal{A}(x,z),\mathcal{A}(y,t))$,
- $\mathcal{A}$ *is said to have* $e \in L^I$ *a neutral element, if* $\mathcal{A}(x,e) = \mathcal{A}(e,x) = x$,
- $\mathcal{A}$ *is said to have* $z \in L^I$ *an absorbing element, if* $\mathcal{A}(x,z) = \mathcal{A}(z,x) = z$,
- $\mathcal{A}$ *dominates* $\mathcal{B}$ $(\mathcal{A} \gg \mathcal{B})$, *if* $\mathcal{A}(\mathcal{B}(x,y),\mathcal{B}(z,t)) \geq_{L^I} \mathcal{B}(\mathcal{A}(x,z),\mathcal{A}(y,t))$,
- $\mathcal{A}$ *satisfies the Non-Contradiction principle w.r.t.* N_{IV} *(NC(N)), if* $\mathcal{A}(x, N_{IV}(x)) = \mathbf{0}$,
- $\mathcal{A}$ *satisfies the Excluded-Middle principle w.r.t.* N_{IV} *(EM(N)), if* $\mathcal{A}(x, N_{IV}(x)) = \mathbf{1}$,
- $\mathcal{A}$ *is conjunctor (disjunctor) if and only if it satisfies the condition* $\mathcal{A}(\mathbf{0},\mathbf{1}) = \mathcal{A}(\mathbf{1},\mathbf{0}) = \mathbf{0}$ $(\mathcal{A}(\mathbf{0},\mathbf{1}) = \mathcal{A}(\mathbf{1},\mathbf{0}) = \mathbf{1})$,
- $\mathcal{A}$ *is conjunctive (disjunctive, averaging), if* $\mathcal{A} \leq_{L^I} min$ $(\mathcal{A} \geq_{L^I} max,\ min \leq_{L^I} \mathcal{A} \leq_{L^I} max)$.

3 Preference Structure

Firstly, we recall concept of fuzzy preference relation. A fuzzy preference relation R on a set of alternatives $X = \{x_1, ..., x_n\}$ is a fuzzy subset of the Cartesian product $X \times X$, that is $R : X \times X \to [0,1]$ [10,11,16,23] or [27] for each pair of alternatives x_i and x_j, $R_{ij} = R(x_i, xj)$ represents a degree of (weak) preference of x_i over x_j, namely the degree to which x_i is considered as least as good as x_j. The preference relation may be conveniently represented by the $n \times n$ matrix $R = (R_{ij})$ for all $i,j \in \{1,...,n\}$. From a weak preference relation R, Fodor and Roubens [16] or [11] derive the following relations:

1. Strict preference $P_{ij} = P(x_i, x_j)$ is a measure of strict preference of x_i over x_j, indicating that x_i is (weakly) preferred to x_j but x_j is not (weakly) preferred to x_i.
2. Indifference $I_{ij} = I(x_i, x_j)$ is a measure of the simultaneous fulfillment of R_{ij} and R_{ji}. Roughly speaking, x_i and x_j are considered equal in the sense that both x_i is as good as x_j and the other way around.
3. Incomparability $J_{ij} = J(x_i, x_j)$ is a measure of the incomparability of x_i and x_j. More specifically, Fodor and Roubens [16] proposed the following expressions of the above relations in terms of a t-norm T and a strict negation N:

$$P_{ij} = T(R_{ij}, N(R_{ji})),\ I_{ij} = T(R_{ij}, R_{ji}),\ J_{ij} = T(N(R_{ij}), N(R_{ji}))$$

for all $i,j \in \{1,...,n\}$.

Now, we recall the notion of interval-valued fuzzy relation.

Definition 7 (cf. [28,35]). *An interval-valued fuzzy relation (IVFR) R between universes* X, Y *is a mapping* $R : X \times Y \to L^I$ *such that*

$$R(x,y) = [\underline{R}(x,y), \overline{R}(x,y)] \ \ for\ all\ pairs\ \ (x,y) \in X \times Y.$$

The class of all IVFRs between universes X, Y *is denoted by* $\mathcal{IVFR}(X \times Y)$, *or* $\mathcal{IVFR}(X)$ *for* $X = Y$.

Note that, if we consider the order defined in L^I, we see that the family $(\mathcal{IVFR}(X \times Y),\vee,\wedge)$ is a complete and distributive lattice (see [6] for a study on the concept of lattices).

An approach that adds flexibility to represent uncertainty in decision making problems consists of using interval-valued fuzzy relations [18,31,32]. An interval-valued fuzzy preference relation R on X is defined as an interval-valued fuzzy subset of $X \times X$, that is, $R : X \times X \to L^I$, which have been studied deeply (see [2,4,31–33] or [5] created with grouping and overlap functions).

The interval $R(x_i, x_j) = r_{ij} = [\underline{r}_{ij}, \overline{r}_{ij}]$ denotes the degree to which elements x_i and x_j are related (representing the degree of preference of x_i over x_j) in the relation R for all x_i, $x_j \in X$. In [2] for given an IVFR, $R = (r_{ij})$, was examined corresponding Interval-valued strict preference (P), interval-valued indifference (I) and interval-valued incomparability (J) by using IV t-norms and IV negations generated from the standard strict negation. We based on the corresponding ones given by Fodor and Roubens propose generalization of these concept and we use IV aggregations instead of IV t-norms and negations instead of classic negations. According to Independence of Irrelevant Alternatives and Positive Association Principle (see [16]) we propose three functions $p, i, j : (L^I)^2 \to L^I$ such that

$$P(a, b) = p(R(a, b), R(b, a)),$$

$$I(a, b) = i(R(a, b), R(b, a)),$$

$$J(a, b) = j(R(a, b), R(b, a))$$

and functions $p(x, N(y))$, $i(x, y)$ and $j(N(x), N(y))$ are increasing with respect to both arguments. Moreover, according to the considerations in [16] we use two different IV aggregations $\mathcal{A}$ and $\mathcal{B}$:

(1) Interval-valued strict preference

$$\begin{aligned} P_{ij} &= \mathcal{A}(r_{ij}, N_{IV}(r_{ji})) = \mathcal{A}([\underline{R}(i,j), \overline{R}(i,j)], [N(\overline{R}(j,i)), N(\underline{R}(j,i))) \\ &= [A_1(\underline{R}(i,j), N(\overline{R}(j,i))), A_2(\overline{R}(i,j), N(\underline{R}(j,i)))]; \end{aligned} \tag{2}$$

(2) Interval-valued indifference

$$\begin{aligned} I_{ij} &= \mathcal{B}(r_{ij}, r_{ji}) = \mathcal{B}([\underline{R}(i,j), \overline{R}(i,j)], [\underline{R}(j,i), \overline{R}(j,i)]) \\ &= [B_1(\underline{R}(i,j), \underline{R}(j,i)), B_2(\overline{R}(i,j), \overline{R}(j,i))]; \end{aligned} \tag{3}$$

(3) Interval-valued incomparability

$$\begin{aligned} J_{ij} &= \mathcal{B}(N_{IV}(r_{ij}), N_{IV}(r_{ji})) = \mathcal{B}(N_{IV}([\underline{R}(i,j), \overline{R}(i,j)]), N_{IV}([\underline{R}(j,i), \overline{R}(j,i)])) \\ &= [B_1(N(\overline{R}(i,j)), N(\overline{R}(j,i))), B_2(N(\underline{R}(i,j)), N(\underline{R}(j,i)))] \end{aligned} \tag{4}$$

for all $i, j \in \{1, ..., n\}$. It is generalisation by $\mathcal{A}, \mathcal{B}$ IV aggregations and N_{IV} IV negation of preference structure. Now, we recall the crucial definition for this paper of reciprocity property based on negation.

Definition 8 ([25]). *Let $card(X) = n$. An Interval-Valued Fuzzy Reciprocal Relation (IVFRR) R on the set X is a matrix $R = (r_{ij})_{n\times n}$ with $r_{ij} = [\underline{R}(i,j), \overline{R}(i,j)]$, for all $i, j \in \{1, \ldots, n\}$, where $r_{ij} \in L^I$*

$$r_{ii} = [0.5, 0.5], r_{ji} = N_{IV}(r_{ij}) = [N(\overline{R}(i,j)), N(\underline{R}(i,j))]$$

for $i \neq j$, where N is a fuzzy negation and N_{IV} is an IV negation function.

Presented the reciprocity property is based on negation. This notion is a generalization of the reciprocity property introduced in [32], where N was a standard negation. However, the assumption $r_{ji} = 1 - r_{ij}$ for $i, j \in \{1, \ldots, n\}$, which stems from the reciprocity property, is rather strong and frequently violated by decision makers in real-life situations. This is why we use a fuzzy negation instead of the classical negation $N(x) = 1 - x$. Especially, if $\overline{R}(i,j) = \underline{R}(i,j) = r_{ij}$ for $i, j \in \{1, \ldots, n\}$, then an IVFRR reduces to a reciprocal fuzzy relation (it is also worth mentioning that IVFRRs may be built from the fuzzy ones using the concept of ignorance function [2]). The interval r_{ij} indicates the interval-valued reciprocal degree or intensity of the alternative x_i over alternative x_j and $\underline{R}(i,j)$, $\overline{R}(i,j)$ are the lower and upper limits of r_{ij}, respectively.
The following results extends Theorem from [2].

Proposition 1 ([4]). *Let $R \in \mathcal{IVFR}(X \times X)$ be reciprocal and let P_{ij} be its associated interval-valued strict fuzzy preference relation given. The following equivalence holds:*
$P_{ij} = r_{ij}$ for all $i, j \in \{1, ..., n\}$ if and only if IV aggregation $\mathcal{A}$ is idempotent.

Corollary 1. *$P_{ij} = r_{ij} \Leftrightarrow \mathcal{A}$ is averaging aggregation function.*

Proposition 2. *Let $R \in \mathcal{IVFR}(X)$ be reciprocal and I, J be associated interval-valued indifference and incomparability fuzzy preference relations. If $\mathcal{B}$ is symmetric, then $I = J$.*

Now, we consider the most important property in point of view of consistency of the group decision making, i.e. transitivity. We will consider $\mathcal{A}$-transitivity of the relation $R \in IVFR(X)$, i.e. $\mathcal{A}(R(x,y), R(y,z)) \leq_{L^I} R(x,z)$ for $x, y, z \in X$.

Proposition 3. *Let $R \in \mathcal{IVFR}(X)$ be reciprocal and P be associated interval-valued strict fuzzy preference relation and $\mathcal{A}$ be bisymmetric IV aggregation function. If R is $\mathcal{A}$-transitive, then P is also $\mathcal{A}$-transitive.*

Proof. By reciprocity of R and $\mathcal{A}$-transitivity by bisymmetry of $\mathcal{A}$ we have:

$$\mathcal{A}(P(x,y), P(y,z)) = \mathcal{A}(\mathcal{A}(R(x,y), N_{IV}(R(y,x))), \mathcal{A}(R(y,z), N_{IV}(R(z,y))) =$$

$$\mathcal{A}(\mathcal{A}(R(x,y), R(x,y)), \mathcal{A}(R(y,z), R(y,z)) =$$

$$\mathcal{A}(\mathcal{A}(R(x,y), R(y,z)), \mathcal{A}(R(x,y), R(y,z)) \leq \mathcal{A}(R(x,z), R(x,z)) =$$

$$\mathcal{A}(R(x,z), N_{IV}R(z,x)) = P(x,z).$$

So P is $\mathcal{A}$-transitive.

Now, we consider asymmetry property and more practical property, i.e. weak asymmetry property of relation $R \in \mathcal{IVFR}(X)$, $card(X) = n$:

- R is $\mathcal{A}$-asymmetric, if $\mathcal{A}(R_{ij}, R_{ji}) = \mathbf{0}$,
- R is weakly $\mathcal{A}$-asymmetric, if $\mathcal{A}(R_{ij}, R_{ji}) \leq_{L^I} [\frac{1}{2}, \frac{1}{2}]$ for all $i, j = \{1, ..., n\}$, $n \in \mathbb{N}$.

Proposition 4. *If R is $\mathcal{A}$-asymmetric, $\mathcal{A}$ is N-stable (i.e. $\mathcal{A}(N_{IV}(R_{ij}), N_{IV}(R_{ji})) = N_{IV}(\mathcal{A})(R_{ij}, R_{ji})$) and bisymmetric aggregation function such that $\mathcal{A}(\mathbf{0}, \mathbf{1}) = \mathbf{0}$, then P is also $\mathcal{A}$-asymmetric.*

Proof. Let $\mathcal{A}$ be bisymmetric aggregation function, i.e.

$$\mathcal{A}(\mathcal{A}(x_{11}, ..., x_{1m}), ..., \mathcal{A}(x_{m1}, ..., x_{mm})) = \mathcal{A}(\mathcal{A}(x_{11}, ..., x_{m1}), ..., \mathcal{A}(x_{1m}, ..., x_{mm})).$$

Then for $m = 2$ we have

$$\mathcal{A}(P_{ij}, P_{ji}) = \mathcal{A}(\mathcal{A}(R_{ij}, N_{IV}(R_{ji})), \mathcal{A}(R_{ji}, N_{IV}(R_{ij})) =$$

$$\mathcal{A}(\mathcal{A}(R_{ij}, R_{ji}), \mathcal{A}(N_{IV}(R_{ji}, N_{IV}(R_{ij})) = \mathcal{A}(\mathcal{A}(R_{ij}, R_{ji}), N_{IV}\mathcal{A}(R_{ji}, R_{ij}) =$$

$$\mathcal{A}(\mathbf{0}, \mathbf{1}) = \mathbf{0}.$$

Thus P is $\mathcal{A}$-asymmetric.

Proposition 5. *Let N_{IV} be IV negation function fulfil $N_{IV}(x) \leq_{L^I} x$. If R is weakly $\mathcal{A}$-asymmetric and $\mathcal{A}$ is N-stable and bisymmetric, subidempotent aggregation function, then P is also weakly $\mathcal{A}$-asymmetric.*

Proposition 6. *For an interval-valued fuzzy preference structure (P, I, J) the following equalities are fulfilled:*
$p(\mathbf{0}, \mathbf{1}) = i(\mathbf{0}, \mathbf{0}) = j(\mathbf{1}, \mathbf{1}) = \mathbf{0}$ and $p(\mathbf{1}, \mathbf{0}) = i(\mathbf{1}, \mathbf{1}) = j(\mathbf{0}, \mathbf{0}) = \mathbf{1}$.

We would like to consider the system of the functional equations:

$$\mathcal{C}(P(a, b), I(a, b)) = R(a, b),$$

$$\mathcal{C}(P(a, b), J(a, b)) = N(R(b, a)),$$

where IV aggregation function $\mathcal{C}$ fulfils $\mathcal{C}(s, t) = \mathbf{0} \Leftrightarrow s = \mathbf{0}$ *and* $t = \mathbf{0}$ and $\mathcal{C}(s, t) = \mathbf{1} \Leftrightarrow s = \mathbf{1}$ *and* $t = \mathbf{1}$.
If we denote $R(a, b) = x$ and $R(b, a) = y$ we can write our system in the following way:

$$\mathcal{C}(p(x, y), i(x, y)) = x, \tag{5}$$

$$\mathcal{C}(p(x, y), j(x, y)) = N(y). \tag{6}$$

Proposition 7. *If (5) and (6) are fulfilled, then*

$$i(\mathbf{0}, y) = p(\mathbf{0}, y) = j(x, \mathbf{1}) = p(x, \mathbf{1}) = \mathbf{0}.$$

Directly from the Eqs. (5) and (2–3) appears the corresponding following composition:

Definition 9. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be IV aggregation functions and N_{IV} be IV negation function. The operation $\mathcal{D} : (L^I)^2 \to L^I$ associated with aggregation functions $\mathcal{A}, \mathcal{B}, \mathcal{C}$ we define as follows*

$$\mathcal{D}(x, y) = \mathcal{C}(\mathcal{A}(x, N_{IV}(y)), \mathcal{B}(x, y)). \tag{7}$$

We can also write this operation as follows:

$$\mathcal{D}(x, y) = \mathcal{C}(P(x, y), I(x, y)).$$

Example 1. Let $x, y \in L^I$.

$$\mathcal{A}(x, y) = \begin{cases} \mathbf{1}, & if \ \ y = \mathbf{1} \\ \mathcal{A}_{mean}, & otherwise \end{cases}$$

for N_{IV} IV negation function,

$$\mathcal{B}(x, y) = \begin{cases} \mathbf{1}, & if \ \ x \geq_{L^I} [\frac{1}{2}, \frac{1}{2}] \ or \ y \geq_{L^I} [\frac{1}{2}, \frac{1}{2}] \\ \mathbf{0}, & if \ \ x = y = \mathbf{0} \\ \alpha, & otherwise \end{cases}$$

and

$$\mathcal{C}(x, y) = \begin{cases} \mathbf{1}, & if \ \ x = \mathbf{1} \ or \ y = \mathbf{1} \\ \mathcal{A}_{mean}, & otherwise. \end{cases}$$

Then

$$\mathcal{D}(x, y) = \begin{cases} \frac{1}{4}x + \frac{1}{4}N_{IV}(y) + \frac{1}{2}\alpha, & if \ \ x, y <_{L^I} [\frac{1}{2}, \frac{1}{2}], y \neq \mathbf{0} \\ \mathbf{1}, & otherwise \end{cases}$$

for $\alpha \in (0, 1)$.

The operation $\mathcal{D}$ has following properties.

Proposition 8. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be IV aggregation functions and $\mathcal{D}$ be created by (7).*

1. *$\mathcal{D}$ has the absorbing element $\mathbf{0}$ ($\mathbf{1}$), if $\mathcal{A}, \mathcal{B}, \mathcal{C}$ have the absorbing element $\mathbf{0}$ ($\mathbf{1}$).*
2. *$\mathcal{D}(\mathbf{0}, \mathbf{0}) = \mathbf{0}$ ($\mathcal{D}(\mathbf{1}, \mathbf{1}) = \mathbf{1}$), if $\mathcal{A}(\mathbf{0}, \mathbf{1}) = \mathbf{0}$ ($\mathcal{A}(\mathbf{1}, \mathbf{0}) = \mathbf{1}$).*
3. *$\mathcal{D}$ has NC(N) (EM(N)) property, if $\mathcal{B}$ has this property and $\mathcal{C}$ has the absorbing element $\mathbf{0}$ ($\mathbf{1}$).*
4. *$\mathcal{D}$ is isotonic with respect to the first variable.*
5. *$\mathcal{D} \gg \mathcal{E}$, if $\mathcal{A} \gg \mathcal{E}$, $\mathcal{B} \gg \mathcal{E}$ and $\mathcal{C} \gg \mathcal{E}$.*
6. *$\mathcal{D}$ is conjunctive (disjunctive), if $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are conjunctive (disjunctive).*

In [4] we check preservation reciprocity property by aggregation function, now we study this property by operation $\mathcal{D}$.

Proposition 9. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be IV representable aggregation functions, such that $A_1 = A_2^N$, $B_1 = B_2^N$, $C_1 = C_2^N$ and N be IV strong negation function. Then operation $\mathcal{D}$ preserves reciprocity property of R.*

4 Application

Our above results allow to perform the following applications. We consider an interval-valued fuzzy relation on $X = \{x_1, \ldots, x_n\}$ (set of alternatives) which represents the expert's opinion of each alternative over another one, i.e. preferences. The preferences will be represented with respect to a finite number of criteria, mathematically these are relations $R_1, \ldots, R_n \in \mathcal{IVFR}(X)$. To find the solution alternative we apply modified voting method by generalized preference structure (P, I, J) and a linear order generated by aggregation functions $\leq_{K_{1,2}}$ [3] defined in the following way: $x \leq_{K_{1,2}} y$ if and only if

$$K_1(\underline{x}, \overline{x}) < K_1(\underline{y}, \overline{y})$$

$$or\ (K_1(\underline{x}, \overline{x}) = K_1(\underline{y}, \overline{y})\ and\ K_2(\underline{x}, \overline{x}) \leq K_2(\underline{y}, \overline{y}))$$

for two continuous aggregation functions, such that, for all $x, y \in L^I$, the equalities $K_1(\underline{x}, \overline{x}) = K_1(\underline{y}, \overline{y})$ and $K_2(\underline{x}, \overline{x}) = K_2(\underline{y}, \overline{y})$ hold if and only if $x = y$.

It is worth to mention that at the beginning of algorithm it may be checked if $R_1, \ldots, R_n \in \mathcal{IVFR}(X)$ are reciprocal with respect to some fuzzy negation N. If the answer is positive we may apply the presented in this paper results in order to consider the appropriate aggregation function to aggregate these relations and obtain reciprocal aggregated result. We will present the algorithm in the case when we do not check reciprocity of input relations. In such situation the aggregated IVFR is normalized to obtain the given reciprocity with the use of the following formula

$$R_{ij} = \begin{cases} R_{ij} & if\ R_{ij} \geq_{Lex1} R_{ji}, \\ N_{IV}(R_{ji}) & else, \end{cases} \tag{8}$$

where $[\underline{x}, \overline{x}] \leq_{lex1} [\underline{y}, \overline{y}] \Leftrightarrow \underline{x} < \underline{y}\ \ or\ \ (\underline{x} = \underline{y}\ \ and\ \ \overline{x} \leq \overline{y})$.

The following algorithm gives an alternative who has the worst/best relationships in a considered X.

Algorithm. $\mathcal{D}$ – **composition**.

Inputs : $X = \{x_1, \ldots, x_n\}$ set of alternatives; $\mathcal{A}$, $\mathcal{B}$ IV aggregation functions; $R_1, \ldots, R_n \in \mathcal{IVFR}(X)$; $\mathcal{D}$ created according to (7); Interval-valued fuzzy preference (reciprocal and transitive or not) relations; The linear order $\leq_{K_{1,2}}$.

Output : The best alternative: $x_{selection}$.

(**Step 1**) Aggregation of given relations $R_1, \ldots, R_n \in \mathcal{IVFR}(X)$ to obtain $R \in \mathcal{IVFR}(X)$;

(**Step 2**) Normalization of relation R with the use of Eq. (8);

(**Step 3**) Build J interval-valued fuzzy relation according to (4);

(**Step 4**) Calculate

$$M_{ij} = \mathcal{A}(\mathcal{D}(R_{ij}, R_{ji}), J_{ij});$$

(**Step 5**) For $m = 1$ to $n + 1$
Find

$$x_{selection} = arg \max_i (\mathcal{B}(M_{ij})),$$

where $1 \leq j \neq i \leq n$, $\mathcal{B} \geq_{L^I} \max$, using a linear order $\leq_{K^m_{1,2}}$;

If only one alternative is the "best" solution, then this alternative is the final solution of the decision making problem;

Else we chose the one with the smallest interval length as the final solution of the decision making problem (if they have the same lengths, then we change aggregation $\mathcal{B}$ and we repeat Step 5)

End

5 Conclusion

We present certain aspect of decision making problem based on preference relations built by IV aggregations and reciprocity property built by negation functions. We propose new idea of reciprocity property and new preference structure. In future we would like to study more properties and classification of these reciprocity and structure. Moreover, we would like to consider solutions of the presented system of the functional Eqs. (5–6). Especially, assumptions about $\mathcal{C}$, $\mathcal{A}$ and $\mathcal{B}$ by which mentioned solutions exist. We would like to propose the practical example of the presented algorithm and compare obtained results with similar algorithms, e.g. from [2,5] or [4].

Acknowledgment. This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge of University of Rzeszów, Poland, project RPPK.01.03.00-18-001/10.

References

1. Aczél, J.: Lectures on Functional Equations and their Applications. Academic Press, New York, London (1966)
2. Barrenechea, E., Fernandez, J., Pagola, M., Chiclana, F., Bustince, H.: Construction of interval-valued fuzzy preference relations from ignorance functions and fuzzy preference relations: application to decision making. Knowl. Based Syst. **58**, 33–44 (2014)
3. Bustince, H., Fernandez, J., Kolesárová, A., Mesiar, R.: Generation of linear orders for intervals by means of aggregation functions. Fuzzy Sets Syst. **220**, 69–77 (2013)

4. Bentkowska, U., Pękala, B., Bustince, H., Fernandez, J., Jurio, A., Balicki, K.: N-reciprocity property for interval-valued fuzzy relations with an application to group decision making problems in social networks. Inte. J. Uncertain. Fuzziness Knowl. Based Syst. (submitted)
5. Bustince, H., Pagola, M., Mesiar, R., Hüllermeier, E., Herrera, F.: Grouping, overlap, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. IEEE Trans. Fuzzy Syst. **20**(3), 405–415 (2012)
6. Birkhoff, G.: Lattice Theory. AMS Coll. Publ. 25, Providence (1967)
7. Calvo, T., Kolesárová, A., Komorniková, M., Mesiar, R.: Aggregation operators: properties, classes and construction methods. In: Calvo, T., et al. (eds.) Aggregation Operators, pp. 3–104. Physica-Verlag, Heidelberg (2002)
8. Chen, H., Zhou, L.: An approach to group decision making with interval fuzzy preference relations based on induced generalized continuous ordered weighted averaging operator. Expert Syst. Appl. **38**, 13432–13440 (2011)
9. Chiclana, F., Herrera-Viedma, E., Alonso, S., Pereira, R.A.M.: Preferences and consistency issues in group decision making. In: Bustince, H., et al. (eds.) Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, pp. 219–237. Springer, Berlin (2008)
10. Chiclana, F., Herrera, F., Herrera-Viedma, E.: Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. Fuzzy Sets Syst. **97**(1), 33–48 (1998)
11. De Baets, B., Van de Walle, B., Kerre, E.: Fuzzy preference structures without incomparability. Fuzzy Sets Syst. **76**(3), 333–348 (1995)
12. Deschrijver, G., Cornelis, C., Kerre, E.E.: On the representation of intuitonistic fuzzy t-Norms and t-Conorms. IEEE Trans. Fuzzy Syst. **12**, 45–61 (2004)
13. Deschrijver, G., Kerre, E.: Aggregation operators in interval-valued fuzzy and atanassov intuitionistic fuzzy set theory. In: Bustince, H., et al. (eds.) Fuzzy Sets and their Extensions: Representation, Aggregation and Models, pp. 183–203. Springer (2008)
14. Deschrijver, G.: Quasi-arithmetic means and OWA functions in interval-valued and Atanassovs intuitionistic fuzzy set theory. In: Galichet, S., et al. (eds.) Proceedings of EUSFLAT-LFA 2011, July 18–22, 2011, pp. 506–513. Aix-les-Bains, France (2011)
15. Drygaś, P., Pękala, B.: Properties of decomposable operations on some extension of the fuzzy set theory. In: Atanassov, K.T., Hryniewicz, O., Kacprzyk, J., et al. (eds.) Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics, pp. 105–118. EXIT, Warsaw (2008)
16. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support in Theory and Decision Library. Kluwer Academic Publishers, Dordrecht (1994)
17. Freson, S., De Baets, B., De Meyer, H.: Closing reciprocal relations w.r.t. stochastic transitivity. Fuzzy Sets Syst. **241**, 2–26 (2014)
18. Gorzałczany, M.B.: A method of inference in approximate reasoning based on interval-valued fuzzy sets. Fuzzy Sets Syst. **21**(1), 1–17 (1987)
19. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
20. Komorníková, M., Mesiar, R.: Aggregation functions on bounded partially ordered sets and their classification. Fuzzy Sets Syst. **175**, 48–56 (2011)
21. Liu, F., Zhang, W.-G., Zhang, L.-H.: A group decision making model based on a generalized ordered weighted geometric average operator with interval preference matrices. Fuzzy Sets Syst. **246**, 1–18 (2014)

22. Llamazares, B.D., Baets, B.: Fuzzy strict preference relations compatible with fuzzy orderings. Int. Uncertain. Fuzziness Knowl. Based Syst. **18**(1), 13–24 (2010)
23. Ovchinnikov, S.V., Roubens, M.: On strict preference relations. Fuzzy Sets Syst. **43**, 319–326 (1991)
24. Paternain, D., Jurio, A., Barrenechea, E., Bustince, H., Bedregal, B., Szmidt, E.: An alternative to fuzzy methods in decision-making problems. Expert Syst. Appl. **39**(9), 7729–7735 (2012)
25. Pękala, B., Bentkowska, U.: Generalized reciprocity property for interval-valued fuzzy setting in some aspect of social network. In: IWIFSGN 2016 Fifteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warsaw, Poland, October 12–14 (2016)
26. Pradera, A., Beliakov, G., Bustince, H., De Baets, B.: A review of the relationships between implication, negation and aggregation functions from the point of view of material implication. Inf. Sci. **329**, 357–380 (2016)
27. Roubens, M., Vincke, P.: Preference Modelling. Springer, Berlin (1985)
28. Sambuc, R.: Fonctions ϕ-floues: Application á l'aide au diagnostic en pathologie thyroidienne. Ph.D. Thesis, Universit$\acute{e}$ de Marseille, France (1975) (in French)
29. Sanz, J., Fernandez, A., Bustince, H., Herrera, F.: Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning. Inf. Sci. **180**(19), 3674–3685 (2010)
30. Sanz, J., Fernandez, A., Bustince, H., Herrera, F.: A genetic tuning to improve the performance of fuzzy rule-based classification systems with intervalvalued fuzzy sets: degree of ignorance and lateral position. Int. J. Approx. Reason. **52**(6), 751–766 (2011)
31. Türksen, I.B., Bilgiç, T.: Interval-valued strict preference relations with Zadeh triples. Fuzzy Sets. Syst. **78**, 183–195 (1996)
32. Xu, Z.: On compatibility of interval fuzzy preference relations. Fuzzy Optim. Decis. Mak. **3**, 217–225 (2004)
33. Xu, G.-L., Liu, F.: An approach to group decision making based on interval multiplicative and fuzzy preference relations by using projection. Appl. Math. Model. **37**, 3929–3943 (2013)
34. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
35. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning-I. Inf. Sci. **8**, 199–249 (1975)

Comparative Study of Type-1 and Interval Type-2 Fuzzy Systems in the Fuzzy Harmony Search Algorithm Applied to Benchmark Functions

Cinthia Peraza, Fevrier Valdez(✉), and Oscar Castillo

Tijuana Institute of Technology, Tijuana, BC, Mexico
fevrier@tectijuana.mx

Abstract. At present the use of fuzzy systems applied to problem solving is very common, since the use of linguistic variables is less complex when solving a problem. This article presents a study of the use of type-1 and interval type-2 fuzzy system applied to the solution of problems of optimization using meta-heuristic algorithms. There are many types of algorithms that mimic social, biological, etc. behaviors. In this case the work focuses on the metaheuristic algorithms in specific the fuzzy harmony search algorithm (FHS), the meta-heuristic algorithms use a technique to obtain a suitable exploration in a definite space to finish with an exploitation around the best position found, with this it is possible to obtain a good solution of the problem. In particular, it was applied to 11 mathematical reference functions using different numbers of dimensions.

Keywords: Metaheuristic algorithms · Harmony search · Type-1 fuzzy logic · Type-2 fuzzy logic · Dynamic parameter adaptation

1 Introduction

The use of fuzzy systems at present is increasing as they take advantage of the concepts of fuzzy sets, these sets use terms and concepts that are easily understood by people and in turn these apply them to solve all kinds of problems of life real. According to [16, 19, 20], fuzzy logic was conceived by Lotfi. A. Zadeh in 1965, on the basis of a theory of fuzzy sets, which differ from traditional ones, because they considered the degree of membership. The degree of membership is represented by a membership function, or membership, which evaluates the input, and certain predefined rules, assigns the degree of membership to a fuzzy set. These values range from 0 to 1, with 0 none and 1 total membership. There is another classification called type-2 fuzzy systems, which were theoretically proposed by Lotfi A. Zadeh in 1975 [7–10]. The reason for the original fuzzy systems to evolve is to consider levels of uncertainty, expanding its scope. In type-2 fuzzy systems the membership functions can now return a range of values that varies depending on the uncertainty involved, not only in the input but also based on the same membership. Type-2 fuzzy systems use a footprint of uncertainty and it is the value of the function at each point in the two-dimensional space. In type-1 we have

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_15

uncertainty only in the antecedent of the rule, whereas in type-2 we have uncertainty both in the antecedent and in the consequent of the sentence.

This work is based on metaheuristic algorithms which are used a lot to solve real life problems using evolutionary computation techniques, fuzzy systems, neural networks, data mining, etc. as can be observed in [1, 2, 15, 17, 18]. In this case we used the metaheuristic called the harmony search algorithm [4], which is inspired by the music and its aim is to imitate jazz improvisation, some of the most relevant works of the present time with this method are the following [3, 5, 6, 11, 12].

In previous works [13, 14] a fuzzy harmony search algorithm was developed applied to benchmark mathematical functions, achieving with this the control of internal parameters of the algorithm by type-1 and interval type-2 fuzzy systems, removing the update of these parameters manually. It is worth mentioning that in these works only the parameters are updated as the number of iterations advance.

The objective of this research is to analyze changes in the fuzzy harmony search algorithm to improve it, mainly with new input parameters and with techniques that allow an improvement to the method to obtain better solutions.

The document is structured as follows: Sect. 2 describes the problem description and the proposed method, Sect. 3 presents the benchmark functions and the results of the simulation and finally in Sect. 4 the conclusions are presented.

2 Proposed FHS Algorithm

This section describes the main contribution of this work, as mentioned above this paper focuses on a metaheuristic based on music, in specific we refer to the fuzzy harmony search algorithm (FHS), which is based on the original algorithm. FHS dynamically adjusts internal parameters of the previous algorithm to a detailed study of the original method using a type-1 fuzzy system as the number of iterations progresses. The difference with previous work is to incorporate a second input to the type-1 and interval type-2 fuzzy system and combine the two parameters in the outputs to achieve a more complex method, with which problems are solved more effectively. In this case the proposed method focused on the minimization of benchmark mathematical functions. In the Fig. 1 the diagram of the proposed method can be observed, in the part of the process of improvisation is executed the adjustment of dynamic parameters.

Figure 1 described the proposal, in the improvisation step the dynamically adjusted parameters are the harmony memory accepting (*HMR*) and pitch adjustment (*PArate*) parameters, are responsible for achieving a control of exploitation and exploration within a specified range.

To achieve the control of exploration and exploitation within a specified range the proposed method uses two measures in the inputs of the fuzzy system, the first are the iterations shown in Eq. 1 and the second is the diversity shown in the Eq. 2, with the purpose of achieving the overall optimum.

$$\text{Iteration} = \frac{\text{Initial Iteration}}{\text{Final Iterations}} \tag{1}$$

$$\text{Diversity}(S(t)) = \frac{1}{n_S}\sum_{i=1}^{n_S}\sqrt{\sum_{j=1}^{n_x}\left(x_{ij}(t) - \bar{x}_j(t)\right)^2} \quad (2)$$

Where Eq. 1, the initial iteration is the current iteration and final iterations are the maximum iterations. In Eq. 2, S is the harmonies or the population of HS; t is the current improvisation or time, n_s is the size of the harmonies, i is the number of the harmony, n_x is the total number of dimensions, j is the number of the dimension, x_{ij} is the j dimension of the harmony i, $\bar{x}_j$ is the j dimension of the current best harmony of the harmonies.

The fuzzy systems that one used are illustrated in Fig. 2 (type-1 fuzzy system) and Fig. 3 (Interval type-2 fuzzy system). In the two proposed fuzzy systems we use as input the iterations and the diversity and as output the *HMR* and *PArate* parameters. In this case used triangular membership functions were used in all fuzzy systems and all are granulated in three membership functions.

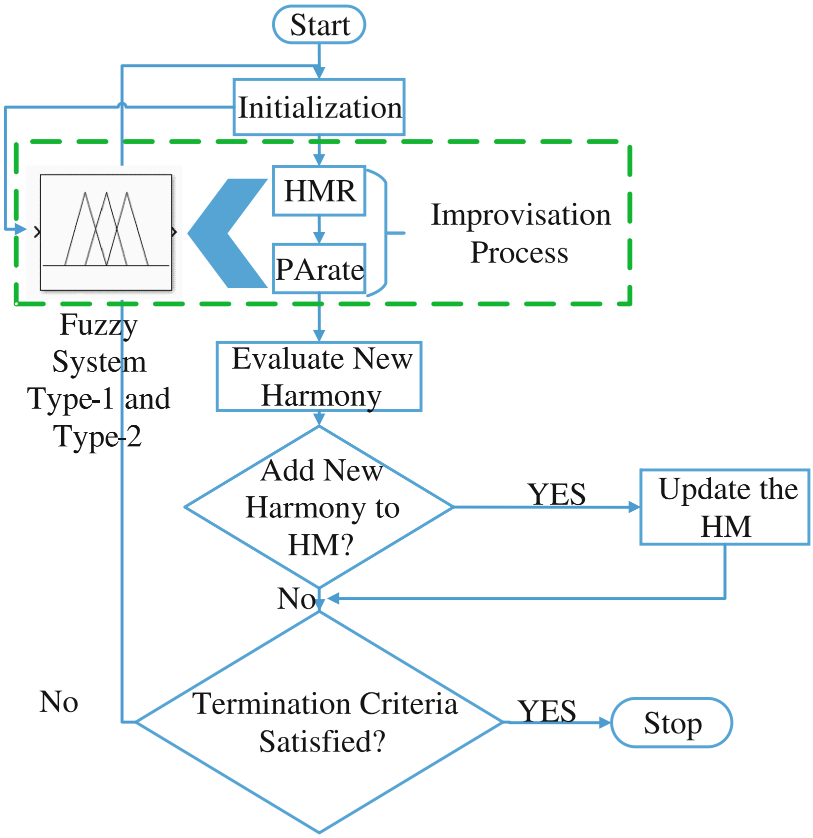


Fig. 1. Schema of the proposed method

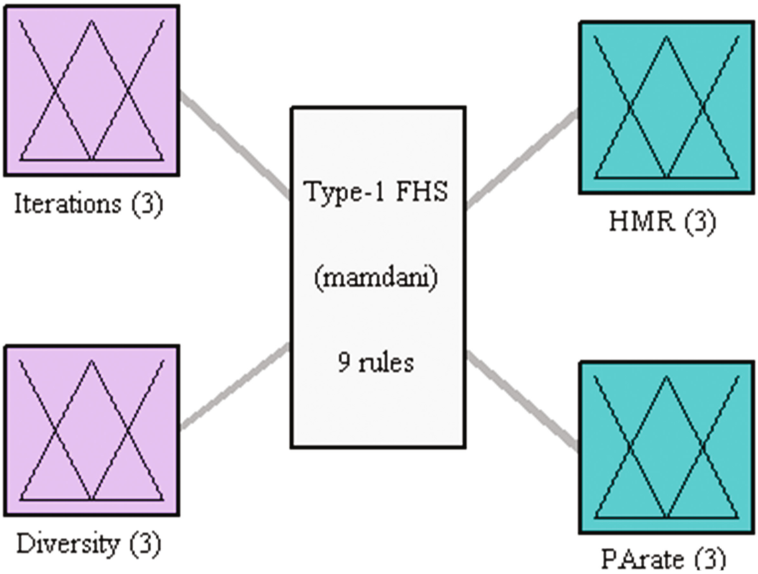


Fig. 2. Type-1 fuzzy system (FHS1).

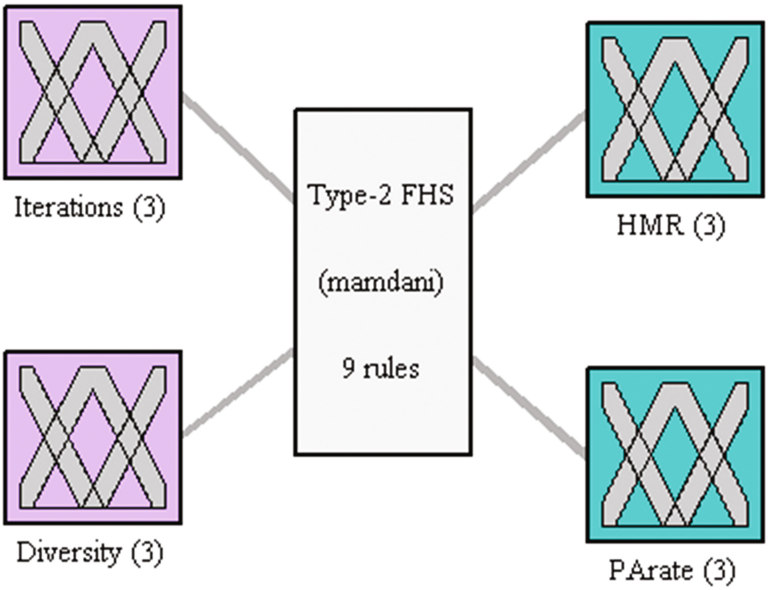


Fig. 3. Interval type-2 fuzzy system (FHS2).

In Fig. 4 there are inputs that are used by each fuzzy system, Fig. 4a and b show the inputs of type-1 fuzzy system, Fig. 4c and d show the inputs of the interval type-2 fuzzy system.

In Fig. 5 shows the outputs used in each fuzzy system, Fig. 5a and b show the outputs of the type-1 fuzzy system, it can be observed that granulated three triangular type membership functions. Figure 5c and d show the outputs of the interval type-2 fuzzy system.

The rules used in fuzzy systems proposed, were created in base knowledge about the behavior of the algorithm and its parameters, thus achieving explore in low iterations and exploiting in high iterations (Fig. 6).

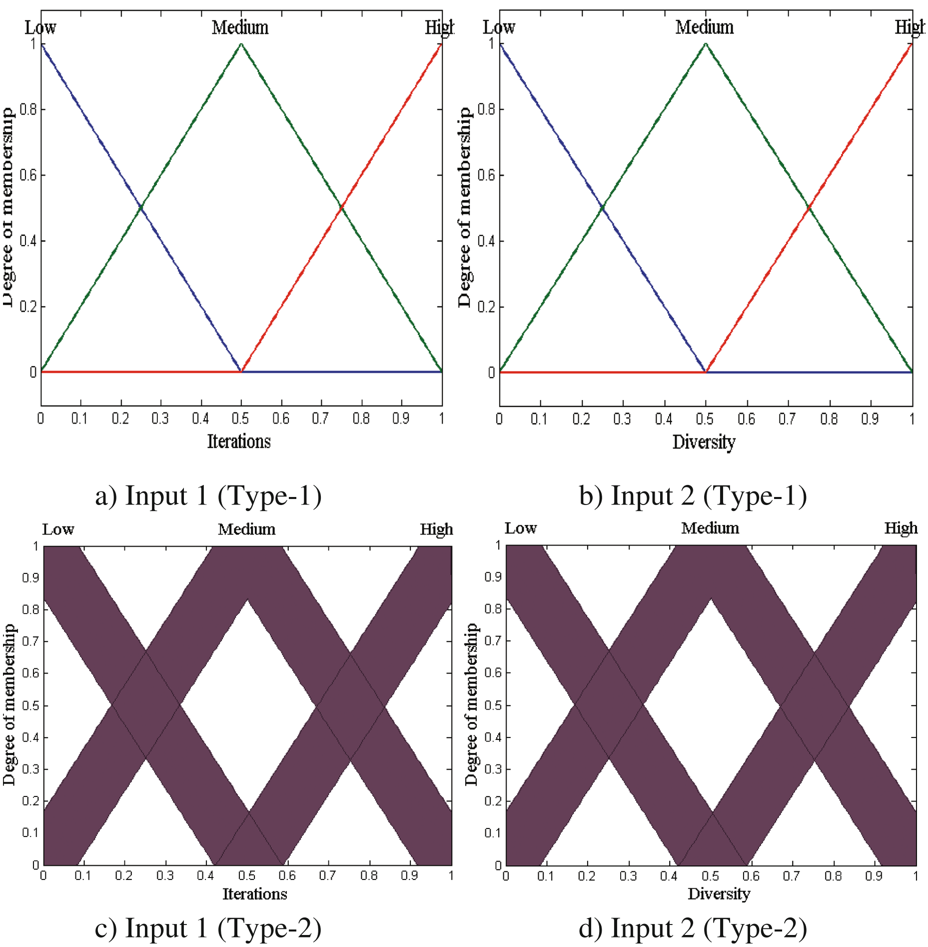
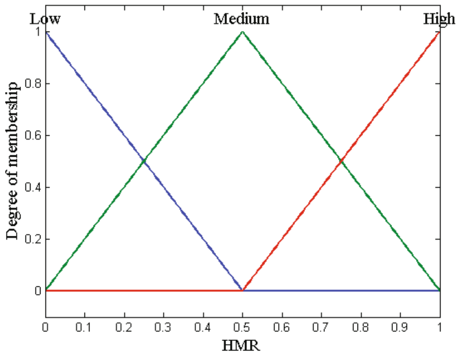
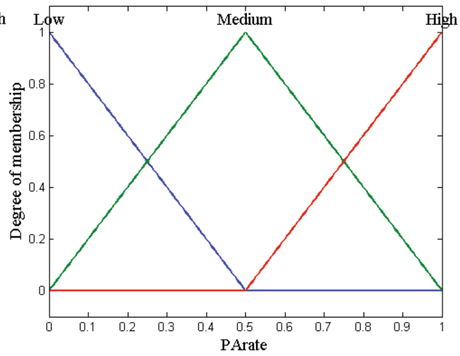


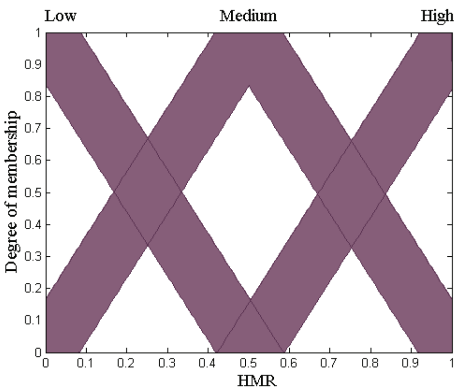
Fig. 4. Inputs of type-1 and interval type-2 fuzzy systems. (a) Input 1 (Type-1). (b) Input 2 (Type-1). (c) Input 1 (Type-2). (d) Input 2 (Type-2)



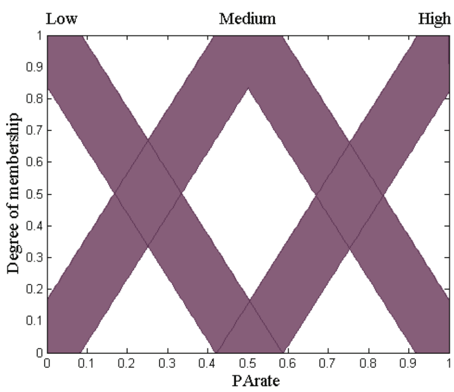
a) Output 1 (Type-1)



b) Output 2 (Type-1)



c) Output 1 (Type-2)



d) Output 2 (Type-2)

Fig. 5. Outputs of type-1 and interval type-2 fuzzy systems. (a) Output 1 (Type-1) (b) Output 2 (Type-1) (c) Output 1 (Type-2) (d) Output 2 (Type-2)

1. If (Improvisation is Low) and (D is Low) then (HMR is High) (PArate is Low) (1)
2. If (Improvisation is Low) and (D is Medium) then (HMR is Medium) (PArate is Medium) (1)
3. If (Improvisation is Low) and (D is High) then (HMR is Medium) (PArate is Medium) (1)
4. If (Improvisation is Medium) and (D is Low) then (HMR is Medium) (PArate is Medium) (1)
5. If (Improvisation is Medium) and (D is Medium) then (HMR is Medium) (PArate is Medium) (1)
6. If (Improvisation is Medium) and (D is High) then (HMR is Medium) (PArate is Medium) (1)
7. If (Improvisation is High) and (D is Low) then (HMR is Medium) (PArate is High) (1)
8. If (Improvisation is High) and (D is Medium) then (HMR is Medium) (PArate is Medium) (1)
9. If (Improvisation is High) and (D is High) then (HMR is Low) (PArate is High) (1)

Fig. 6. Rules for the type-1 and interval type-2 fuzzy systems

3 Simulation Results

The simulations obtained are shown in this section; the proposed method was tested using the mathematical functions shown in Table 1. For the experiments used the dimensions between 2, 6 and 1. In maximum of the functions their global optimum is zero, except for the Shubert and Trid function. 1000 iterations and 50 runs were used for each type-1 and interval type-2 method.

Table 1. Benchmark functions and parameters

Function	Dimension	Search Domain	Global minimum
Rosenbrock	10	[−5, 10]	0
Sphere	10	[−5.12, 5.12]	0
Hump	10	[−5, 5]	0
Rastrigin	10	[−5.12, 5.12]	0
Schwefel	10	[−500, 500]	0
Shubert	2	[−10, 10]	−186.7309
Sum Square	10	[−10, 10]	0
Zakharov	10	[−5, 10]	0
Griewank	10	[−600, 600]	0
Powel	10	[−4, 5]	0
Trid	6	[−36, 36]	−50
Trid	10	[−100, 100]	−200

In Table 1, the benchmark functions with which the proposed method was tested this article, also the range used, dimensions, and its global minimum for each function is shown.

50 runs were performed for each mathematical function using the original HS, type-1 FHS and interval type-2 FHS methods. The average was obtained for each function as shown in Table 2.

Table 2. Values obtained in each function

Function	HS	Type-1 FHS	Type-2 FHS
Rosenbrock	2.57E-02	9.16E-03	5.67E-08
Sphere	1.00E+01	1.07E-02	0.00E+00
Hump	-1.02E+00	6.49E-01	0.00E+00
Rastrigin	1.07E+00	1.59E-02	6.44E-08
Schwefel	1.87E+01	4.82E + 00	1.27E-07
Shubert	-1.85E+02	-1.86E + 02	-1.86E+02
Sum Square	1.47E-01	8.35E-03	3.86E-10
Zakharov	1.65E-01	2.38E-03	8.64E-10
Griewank	3.90E-01	2.11E-01	1.05E-10
Powel	2.66E+00	0.00E+00	0.00E+00
Trid	-1.51E+01	-1.98E+00	-3.31E+01
Trid	6.83E-01	-7.10E-03	-6.03E+00

In Table 2 the averages of the 50 experiments applied to each function are shown, note the improvement in the use of the interval type-2 FHS compared to the original and the type-1 FHS algorithms, in most cases better results are achieved.

4 Conclusions

In this paper the type-1 and interval type-2 FHS algorithm is proposed. This method applies to 11 mathematical reference functions for validation, achieving in most cases to obtain better results when using interval type-2, it can be verified that the greater complexity interval type-2 manages to maintain better results. Unlike the previously created methods based on this same algorithm, this proposal uses two inputs the “iterations” and “diversity” and two outputs the *HMR* and *PArate* to achieve total control over the exploration and exploitation of the algorithm.

Type 2 is more complex and time consuming but it obtains better results in the majority of the 50 experiments by each mathematical function, Type 1 in 1000 iterations and with the same execution time that type 2 does not get better results than type 2. In future works this methodology will be tested to different functions and problems with greater complexity and number of dimensions.

References

1. Arias, N.B., et al.: Metaheuristic optimization algorithms for the optimal coordination of plug-in electric vehicle charging in distribution systems with distributed generation. Electr. Power Syst. Res. **142**, 351–361 (2017)
2. Askarzadeh, A.: A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm. Comput. Struct. **169**, 1–12 (2016)
3. Assad, A., Deep, K.: Applications of harmony search algorithm in data mining: a survey. In: Proceedings of Fifth International Conference on Soft Computing for Problem Solving. Springer Singapore (2016)
4. Geem, Z.W., Kim, J.H., Loganathan, G.V.: A new heuristic optimization algorithm: harmony search. Simulation **76**(2), 60–68 (2001)
5. Kar, P., Swain, S.C.: A harmony search-firefly algorithm based controller for damping power oscillations. In: 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT). IEEE (2016)
6. Lee, A., Geem, Z.W., Suh, K.-D.: Determination of optimal initial weights of an artificial neural network by using the harmony search algorithm: application to breakwater armor stones. Appl. Sci. **6**(6), 164 (2016)
7. Mendel, J.: Type-2 fuzzy sets and systems: an overview [corrected reprint]. IEEE Comput. Intell. Mag. **2**(2), 20–29 (2007)
8. Mendel, J.: Type-2 Fuzzy sets and systems: How to learn about them. IEEE SMC eNewsletter **27** (2009)
9. Mendel, J., John, R.I.B.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**(2), 117–127 (2002)
10. Mendel, J., John, R.I., Liu, F.: Interval type-2 fuzzy logic systems made simple. IEEE Trans. Fuzzy Syst. **14**(6), 808–821 (2006)

11. Molina-Moreno, F., et al.: Optimization of buttressed earth-retaining walls using hybrid harmony search algorithms. Eng. Struct. **134**, 205–216 (2017)
12. Nigdeli, S.M., Bekdaş, G., Yang, X.-S.: Optimum tuning of mass dampers by using a hybrid method using harmony search and flower pollination algorithm. In: International Conference on Harmony Search Algorithm. Springer, Singapore (2017)
13. Peraza, C., et al.: A new fuzzy harmony search algorithm using fuzzy logic for dynamic parameter adaptation. Algorithms **9**(4), 69 (2016)
14. Peraza, C., Valdez, F., Castillo, O.: Interval type-2 fuzzy logic for dynamic parameter adaptation in the Harmony search algorithm. In: 2016 IEEE 8th International Conference on Intelligent Systems (IS). IEEE (2016)
15. Shaddiq, S., et al.: Optimal capacity and placement of distributed generation using metaheuristic optimization algorithm to reduce power losses in Bantul distribution system, Yogyakarta. In: 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE (2016)
16. Terano, T., Asai, K., Sugeno, M.: Fuzzy systems theory and its applications. Academic Press Professional Inc, New York (1992)
17. Thanh, L.T., et al.: A computational study of hybrid approaches of metaheuristic algorithms for the cell formation problem. J. Oper. Res. Soc. **67**(1), 20–36 (2016)
18. Wang, G.-G., et al.: A new metaheuristic optimisation algorithm motivated by elephant herding behaviour. Int. J. Bio-Inspir. Comput. **8**(6), 394–409 (2016)
19. Zadeh, L.A.: Fuzzy sets. Inform. Control **8**(3), 338–353 (1965)
20. Zadeh, L.A.: Fuzzy sets and applications: selected papers (1987)

Penalty-Based Aggregation Beyond the Current Confinement to Real Numbers: The Method of Kemeny Revisited

Raúl Pérez-Fernández(✉) and Bernard De Baets

KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium
{raul.perezfernandez,bernard.debaets}@ugent.be

Abstract. The field of aggregation theory addresses the mathematical formalization of aggregation processes. Historically, the developed mathematical framework has been largely confined to the aggregation of real numbers, while the aggregation of other types of structures, such as rankings, has been independently considered in different fields of application. However, one could lately perceive an increasing interest in the study and formalization of aggregation processes on new types of data. Mostly, this aggregation outside the framework of real numbers is based on the use of a penalty function measuring the disagreement with a consensus element. Unfortunately, there does not exist a comprehensive theoretical framework yet. In this paper, we propose a natural extension of the definition of a penalty function to a more general setting based on the compatibility with a given betweenness relation. In particular, we revisit one of the most common methods for the aggregation of rankings – the method of Kemeny – which will be positioned in the penalty-based aggregation framework.

Keywords: Penalty function · Aggregation of rankings · Kemeny · Monometric

1 Introduction

The use of penalty[1] functions measuring the disagreement of a list of values with a consensus value is a common approach to data aggregation [1,3,4,16,17]. Mostly, the theoretical framework of penalty-based data aggregation has been developed for real numbers [2]. This confinement contrasts with the increasing interest of the research community in the aggregation of new types of data. For instance, we refer to the problems of the aggregation of rankings [13], the aggregation of multidimensional data [7] and the aggregation of mappings [6].

[1] Some other terms, such as 'cost', 'disagreement', 'discrepancy', 'divergence' or 'error', have been occasionally used for replacing the term 'penalty'. However, in the field of aggregation theory, the term 'penalty' is nowadays considered the standard.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_16

In this contribution, we take a step towards a comprehensive theoretical framework for penalty-based aggregation outside this current confinement to real numbers. In particular, we restrict our attention to the aggregation of rankings, which has been addressed independently by the field of social choice theory since the eighteenth century, and we propose to revisit one of its best-known methods: the method of Kemeny [8]. This method will be proved to be a prominent example of a penalty-based aggregation process.

The rest of the paper is structured as follows. First, we recall the definition of a penalty function in Sect. 2. We expand the definition of a penalty function beyond the current restriction to real numbers in Sect. 3. We discuss the particular setting of the aggregation of rankings in Sect. 4 and prove that the method of Kemeny is a prominent example of penalty-based aggregation of rankings in Sect. 5. We end with some conclusions in Sect. 6.

2 Penalty Functions in the Setting of Real Numbers

Back in 1993, Yager [16] proposed for the first time the use of penalty functions in data aggregation. Given a penalty function measuring the degree of disagreement of a list of values with a consensus value, the value(s) that minimizes the penalty is considered the result of the aggregation. Usually, the considered penalty function has been provided with a well-founded semantic basis ($L(y,x) \leq L(z,x)$,[2] if $|y-x| < |z-x|$, in [17]; $L(y,x) \leq L(z,x)$, if $x \leq y \leq z$ or $z \leq y \leq x$, in [4]; quasi-convexity in the second argument in [2,15]). After three decades of studies of the notion of a penalty function, the current understanding of a penalty function is the following [2].

Definition 1. *Consider $n \in \mathbb{N}$ and a closed interval $I \subseteq \mathbb{R}$. A function $P : I^{n+1} \to \mathbb{R}^+$ is called a penalty function if there exists $c \in \mathbb{R}^+$ such that:*

(i) $P(\mathbf{x}, y) \geq c$, for any $\mathbf{x} \in I^n$ and any $y \in I$;
(ii) $P(\mathbf{x}, y) = c$ if and only if $\mathbf{x} = (y, \ldots, y)$;
(ii) $P(\mathbf{x}, \cdot)$ is quasi-convex and lower semi-continuous for any $\mathbf{x} \in I^n$.

Remark 1. In most cases, the value c is set to zero.

The properties of quasi-convexity and lower semi-continuity of a penalty function P imply that the set of minimizers of $P(\mathbf{x}, \cdot)$ is either a singleton or an interval. This minimizer (or middle point of the interval) is usually considered as the result of the 'aggregation',[3] which is given by the so-called penalty-based function associated with P.

---

[2] We denote (local) penalty functions defined on $\mathbb{R} \times \mathbb{R}$ by L and penalty functions defined on $\mathbb{R}^n \times \mathbb{R}$ by P. For more details, we refer to [2].

[3] Note that we write the term 'aggregation' between quotation marks due to the fact that the penalty-based function associated with P is only assured to satisfy the boundary conditions (it is actually idempotent), while the property of increasingness might not hold. For more details, we refer to [2,3].

Definition 2. *Consider $n \in \mathbb{N}$, a closed interval $I \subseteq \mathbb{R}$ and a penalty function $P : I^{n+1} \to \mathbb{R}^+$. A function $f : I^n \to I$ is called the penalty-based function associated with P if, for any $\mathbf{x} \in I^n$, it holds that*

$$f(\mathbf{x}) = \frac{a+b}{2}\,,$$

where $[a, b]$ is the interval closure of the set of minimizers of $P(\mathbf{x}, \cdot)$.

3 The Extension Beyond the Setting of Real Numbers

Condition (iii) of a penalty function has a twofold goal: first, provide the penalty with a well-founded semantic basis, and, second, assure that the set of minimizers of $P(\mathbf{x}, \cdot)$ is either a singleton or an interval. In case there exists at least one minimizer, the quasi-convexity property is equivalent to the fact that $P(\mathbf{x}, \cdot)$ decreases up to the set of minimizers and increases from the set of minimizers on. Unfortunately, this property can no longer be defined in case we are dealing with other types of data. We propose to introduce the notion of a betweenness relation, which is a ternary relation [10] that describes when an element is in between two other ones, in order to provide the penalty with a well-founded semantic basis. In what follows, we adhere to the formal relaxed definition given by Pitcher and Smiley [14], requiring a minimal set of axioms.

Definition 3. *A ternary relation B on a set X is called a betweenness relation if it satisfies the following two properties:*

(i) Symmetry in the end points: for any $x, y, z \in X$, it holds that

$$(x, y, z) \in B \quad \Leftrightarrow \quad (z, y, x) \in B\,.$$

(ii) Closure: for any $a, b, c \in X$, it holds that

$$\big((x, y, z) \in B \wedge (x, z, y) \in B\big) \quad \Leftrightarrow \quad y = z\,.$$

The fact that $(x, y, z) \in B$ is referred to as 'y is in between x and z'.

As shown in [12], any betweenness relation on a set X can be easily extended to X^n.

Proposition 1. *Consider $n \in \mathbb{N}$ and a betweenness relation B on a set X. The ternary relation $B^{(n)}$ on X^n defined as*

$$B^{(n)} = \{(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in (X^n)^3 \mid (\forall i \in \{1, \ldots, n\})((x_i, y_i, z_i) \in B)\}\,,$$

is a betweenness relation on X^n, called the product betweenness relation.

The definition of a penalty function can then be adapted by requiring the preservation of a given betweenness relation.

Definition 4. *Consider* $n \in \mathbb{N}$*, a set* X *and a betweenness relation* B_n *on* X^n*. A function* $P : X^{n+1} \to \mathbb{R}^+$ *is called a penalty function (compatible with* B_n*) if there exists* $c \in \mathbb{R}^+$ *such that:*

(i) $P(\mathbf{x}, y) \geq c$*, for any* $\mathbf{x} \in X^n$ *and any* $y \in X$*;*
(ii) $P(\mathbf{x}, y) = c$ *if and only if* $\mathbf{x} = (y, \ldots, y)$*;*
(iii) $P(\mathbf{x}, y) \leq P(\mathbf{x}', y)$*, for any* $\mathbf{x}, \mathbf{x}' \in X^n$ *and any* $y \in X$ *such that* $((y, \ldots, y), \mathbf{x}, \mathbf{x}') \in B_n$*.*

Remark 2. An alternative condition requiring the penalty to increase when moving away from the minimizer(s) may be considered: (iii') $P(\mathbf{x}, y) \leq P(\mathbf{x}, y')$, for any $\mathbf{x} \in X^n$, any $y, y' \in X$ and any minimizer $z \in X$ of $P(\mathbf{x}, \cdot)$ such that $((z, \ldots, z), (y, \ldots, y), (y', \ldots, y')) \in B_n$. Note that, while condition (iii) is the natural extension of the condition required in [4] ($L(y, x) \leq L(z, x)$, for any $x \leq y \leq z$ or $z \leq y \leq x$), condition (iii') is the natural extension of quasi-convexity required in [2].

Remark 3. The set of minimizers might be empty, unless, of course, the set X is finite. This problem could be avoided by adding a fourth condition: (iv) The set of minimizers of $P(\mathbf{x}, \cdot)$ is non-empty, for any $\mathbf{x} \in X^n$.

As the set of minimizers is not assured to be a singleton, nor non-empty, the codomain of the corresponding penalty-based function associated with P needs to be the power set $\mathcal{P}(X)$ of X rather than X.

Definition 5. *Consider* $n \in \mathbb{N}$*, a set* X*, a betweenness relation* B_n *on* X^n *and a penalty function* $P : X^{n+1} \to \mathbb{R}^+$ *compatible with* B_n*. A function* $f : X^n \to \mathcal{P}(X)$ *is called the penalty-based function associated with* P *if, for any* $\mathbf{x} \in X^n$*, it holds that* $f(\mathbf{x})$ *equals the set of minimizers of* $P(\mathbf{x}, \cdot)$*.*

Remark 4. Any penalty-based function f associated with P is idempotent, i.e., it holds that $f(x, \ldots, x) = \{x\}$, for any $x \in X$.

4 The Framework of the Aggregation of Rankings

Here, we consider the particular case of the aggregation of rankings. A ranking is a strict total order relation $\succ$ on a set $\mathscr{C} = \{a_1, \ldots, a_k\}$ of k elements, i.e., the asymmetric part of a total order relation $\succeq$ on $\mathscr{C}$. The set of all rankings on $\mathscr{C}$ is denoted by $\mathcal{L}(\mathscr{C})$. Each ranking $\succ$ on $\mathscr{C}$ defines an order relation $\supseteqq_\succ$ on $\mathcal{L}(\mathscr{C})$ according to how far two rankings in $\mathcal{L}(\mathscr{C})$ are from $\succ$ in terms of reversals [11].

Definition 6. *Let* $\mathscr{C}$ *be a set of* k *elements. A ranking* $\succ$ *on* $\mathscr{C}$ *induces the following partial order relation* $\supseteqq_\succ$ *on* $\mathcal{L}(\mathscr{C})$*:*

$$\supseteqq_\succ = \left\{(\succ', \succ'') \in \mathcal{L}(\mathscr{C})^2 \mid \left(\forall a_{i_1}, a_{i_2} \in \mathscr{C}\right) \left(\left((a_{i_1} \succ a_{i_2}) \wedge (a_{i_1} \succ'' a_{i_2})\right) \Rightarrow (a_{i_1} \succ' a_{i_2})\right)\right\} .$$

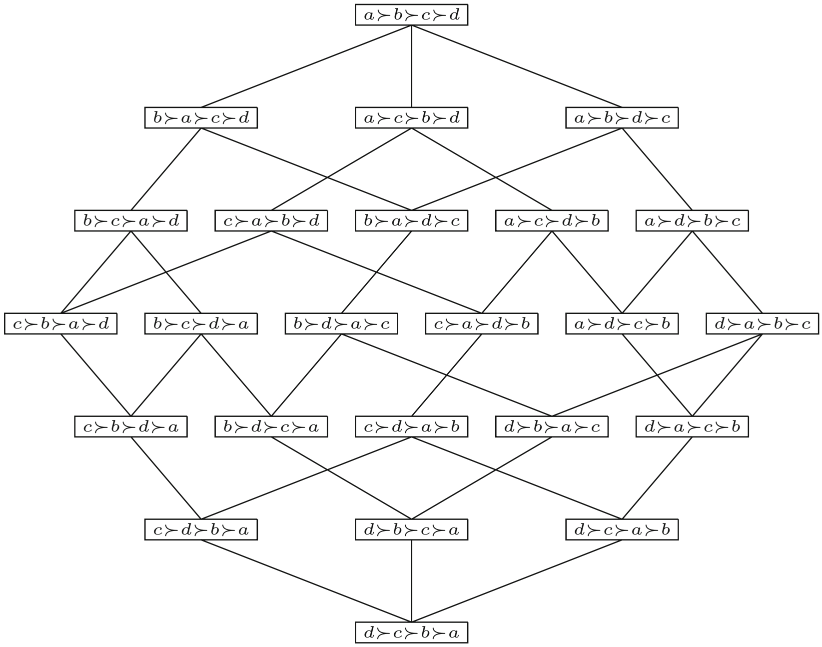


Fig. 1. Hasse diagram of the order relation $\trianglerighteq_{\succ}$ for the ranking $a \succ b \succ c \succ d$.

Figure 1 displays the Hasse diagram of the order relation $\trianglerighteq_{\succ}$ for the ranking $a \succ b \succ c \succ d$ on the set of four elements $\mathscr{C} = \{a, b, c, d\}$. Clearly, every ranking $\succ'$ is closer (in terms of reversals) to $\succ$ than $\succ''$ if it holds that $\succ' \trianglerighteq_{\succ} \succ''$.

The most common notion of distance on rankings is measured by means of the Kendall distance function K between rankings [9]. This distance function assigns to each couple of rankings the number of pairwise disagreements between them. Formally, for any two rankings $\succ_1$ and $\succ_2$, the Kendall distance is defined as

$$K(\succ_1, \succ_2) = \#\{(a_{i_1}, a_{i_2}) \in \mathscr{C}^2_{\neq} \mid a_{i_1} \succ_1 a_{i_2} \wedge a_{i_2} \succ_2 a_{i_1}\}\,.$$

The Kendall distance function induces a natural betweenness relation $B_{\mathcal{L}(\mathscr{C})}$ on $\mathcal{L}(\mathscr{C})$, which is defined as follows:

$$B_{\mathcal{L}(\mathscr{C})} = \{(\succ_1, \succ_2, \succ_3) \in \mathcal{L}(\mathscr{C})^3 \mid K(\succ_1, \succ_3) = K(\succ_1, \succ_2) + K(\succ_2, \succ_3)\}\,.$$

Figure 2 illustrates the rankings that are in between the rankings $a \succ b \succ c \succ d$ and $d \succ b \succ a \succ c$ according to the betweenness relation $B_{\mathcal{L}(\mathscr{C})}$ for the set $\mathscr{C} = \{a, b, c, d\}$. Intuitively, the fact that $(\succ_1, \succ_2, \succ_3) \in B_{\mathcal{L}(\mathscr{C})}$ is equivalent to the fact that $\succ_2 \trianglerighteq_{\succ_1} \succ_3$ (or to the analogous $\succ_2 \trianglerighteq_{\succ_3} \succ_1$).

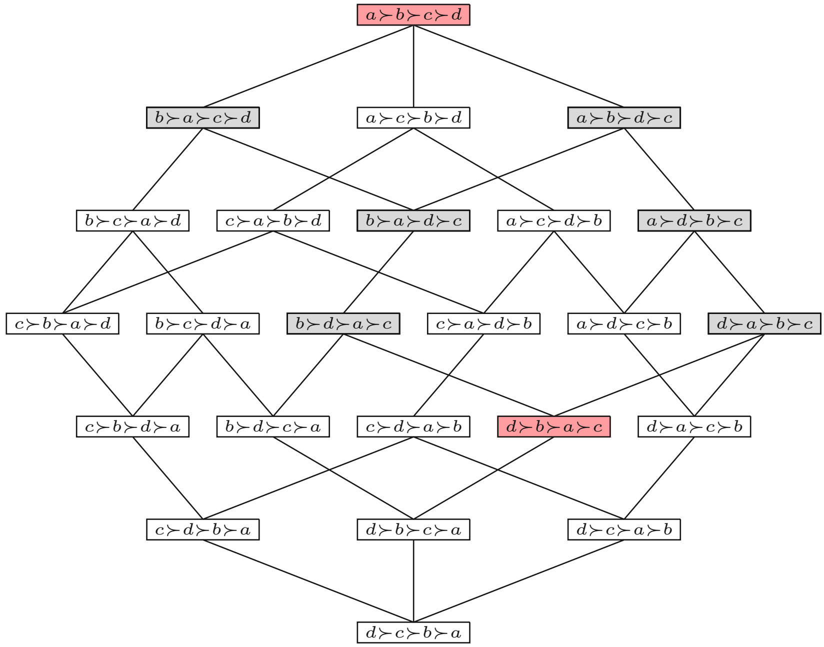


Fig. 2. Graphical representation of $B_{\mathcal{L}}(\mathscr{C})$

5 A Prominent Example of Penalty-Based Aggregation Outside the Setting of Real Numbers: The Method of Kemeny

The method of Kemeny [8] is one of the most (if not the most) common methods for the aggregation of rankings. The winning ranking according to this method is the ranking that minimizes the sum of the Kendall distances to the given list of rankings. Considering the terminology of the penalty-based aggregation framework, we have that the penalty function $P : \mathcal{L}(\mathscr{C})^{n+1} \to \mathbb{R}^+$ associated with the method of Kemeny is defined as

$$P(\mathscr{R}, \succ) = \sum_{j=1}^{n} K(\succ_j, \succ) = \sum_{a_{i_1} \succ a_{i_2}} \#\{\succ_j \in \mathscr{R} \mid a_{i_2} \succ_j a_{i_1}\},$$

for any list of rankings $\mathscr{R} = (\succ_j)_{j=1}^n$ and any ranking $\succ$.

Note that this function is indeed a penalty function (compatible with $B^{(n)}_{\mathcal{L}(\mathscr{C})}$) in the sense of Definition 4, as we will prove in the remainder of this section.

First, as the Kendall distance function is a metric, it holds that condition (i) is trivially satisfied. Second, due to the coincidence axiom of a metric, it holds that condition (ii) is satisfied. Third, consider the betweenness relation $B^{(n)}_{\mathcal{L}(\mathscr{C})}$.

For any $\mathscr{R}, \mathscr{R}' \in \mathcal{L}(\mathscr{C})^n$ and any $\succ \in \mathcal{L}(\mathscr{C})$ such that $((\succ, \ldots, \succ), \mathscr{R}, \mathscr{R}') \in B^{(n)}_{\mathcal{L}(\mathscr{C})}$, it holds that

$$(\forall j \in \{1, \ldots, n\})(K(\succ'_j, \succ) = K(\succ'_j, \succ_j) + K(\succ_j, \succ)) .$$

In particular, as the Kendall distance function is a metric, it holds that

$$(\forall j \in \{1, \ldots, n\})(K(\succ'_j, \succ) \geq K(\succ_j, \succ)) .$$

Therefore, we conclude that condition (iii) is satisfied, i.e.,

$$P(\mathscr{R}, \succ) = \sum_{j=1}^{n} K(\succ_j, \succ) \leq \sum_{j=1}^{n} K(\succ'_j, \succ) = P(\mathscr{R}', \succ) .$$

Note that, in case the winning ranking $\succ$ according to the method of Kemeny is unique, condition (iii') is equivalent to the penalty $P(\mathscr{R}, \cdot)$ being decreasing when going from top to bottom in the Hasse diagram of the order relation $\trianglerighteq_{\succ}$. In general, this condition is not satisfied by the penalty function associated with the method of Kemeny. For instance, in case we consider the list of rankings on $\mathscr{C} = \{a, b, c, d\}$ given in Table 1.

Table 1. List of rankings on $\mathscr{C} = \{a, b, c, d\}$ and their respective frequencies.

Freq.	Ranking
9	$a \succ b \succ c \succ d$
5	$a \succ c \succ b \succ d$
5	$b \succ d \succ c \succ a$
5	$c \succ d \succ b \succ a$
5	$d \succ b \succ a \succ c$
5	$d \succ c \succ a \succ b$

The penalty $P(\mathscr{R}, \succ)$ associated with each ranking $\succ$ on $\mathscr{C} = \{a, b, c, d\}$ for the list $\mathscr{R}$ of rankings given in Table 1 is shown in Table 2.

We see that the unique minimizer of $P(\mathscr{R}, \cdot)$ is the ranking $a \succ b \succ c \succ d$. Due to the fact that $(a \succ b \succ c \succ d, a \succ b \succ d \succ c, d \succ a \succ b \succ c) \in B_{\mathcal{L}(\mathscr{C})}$, it holds that

$$\left(\mathscr{R}_{a \succ b \succ c \succ d}, \mathscr{R}_{a \succ b \succ d \succ c}, \mathscr{R}_{d \succ a \succ b \succ c}\right) \in B^{(n)}_{\mathcal{L}(\mathscr{C})} ,$$

where $\mathscr{R}_{a \succ b \succ c \succ d}$, $\mathscr{R}_{a \succ b \succ d \succ c}$ and $\mathscr{R}_{d \succ a \succ b \succ c}$ denote the unanimous lists of rankings where all the rankings are $a \succ b \succ c \succ d$, $a \succ b \succ d \succ c$ and $d \succ a \succ b \succ c$, respectively. Due to the fact that $P(\mathscr{R}, a \succ b \succ d \succ c) = 99 > 97 = P(\mathscr{R}, d \succ a \succ b \succ c)$, we conclude that condition (iii') is not satisfied.

A common notion in the field of the aggregation of rankings is that of the existence of the Condorcet ranking, which is named after one of the forefathers

Table 2. Penalty $P(\mathscr{R}, \succ)$ associated with each ranking $\succ$ on $\mathscr{C} = \{a, b, c, d\}$ for the list $\mathscr{R}$ of rankings given in Table 1

Ranking	Penalty	Ranking	Penalty	Ranking	Penalty	Ranking	Penalty
$a \succ b \succ c \succ d$	95	$b \succ a \succ c \succ d$	99	$c \succ a \succ b \succ d$	103	$d \succ a \succ b \succ c$	97
$a \succ b \succ d \succ c$	99	$b \succ a \succ d \succ c$	103	$c \succ a \succ d \succ b$	107	$d \succ a \succ c \succ b$	101
$a \succ c \succ b \succ d$	99	$b \succ c \succ a \succ d$	103	$c \succ b \succ a \succ d$	107	$d \succ b \succ a \succ c$	101
$a \succ c \succ d \succ b$	103	$b \succ c \succ d \succ a$	97	$c \succ b \succ d \succ a$	101	$d \succ b \succ c \succ a$	105
$a \succ d \succ b \succ c$	103	$b \succ d \succ a \succ c$	97	$c \succ d \succ a \succ b$	101	$d \succ c \succ a \succ b$	105
$a \succ d \succ c \succ b$	107	$b \succ d \succ c \succ a$	101	$c \succ d \succ b \succ a$	105	$d \succ c \succ b \succ a$	109

of social choice theory: Marquis de Condorcet [5]. For a given list $\mathscr{R}$ of rankings, we say that a ranking $\succ$ is the Condorcet ranking (associated with $\mathscr{R}$) if, for any $a_{i_1}, a_{i_2} \in \mathscr{C}$ such that $a_{i_1} \succ a_{i_2}$, it holds that a_{i_1} is ranked at a better position than a_{i_2} in a greater number of rankings in $\mathscr{R}$ than the number of rankings in $\mathscr{R}$ in which a_{i_2} is ranked at a better position than a_{i_1}.

We now prove that the existence of a Condorcet ranking associated with $\mathscr{R}$ implies the compliance of condition (iii'). It is known that, in case the Condorcet ranking $\succ$ exists, it is the unique winning ranking according to the method of Kemeny. Therefore, it suffices to prove that the penalty $P(\mathscr{R}, \cdot)$ decreases when going from top to bottom in the Hasse diagram of the order relation $\trianglerighteq_{\succ}$. For any two consecutive rankings $\succ^1, \succ^2 \in \mathcal{L}(\mathscr{C})$ in the order relation $\trianglerighteq_{\succ}$ (i.e. $\succ^1 \trianglerighteq_{\succ} \succ^2$ and there does not exist $\succ^3 \in \mathcal{L}(\mathscr{C})$ such that $\succ^1 \trianglerighteq_{\succ} \succ^3 \trianglerighteq_{\succ} \succ^2$), it holds that there exists a unique pair of elements $a_{i_1}, a_{i_2} \in \mathscr{C}$ such that $a_{i_1} \succ^1 a_{i_2}$ while $a_{i_2} \succ^2 a_{i_1}$ (it obviously holds that $a_{i_1} \succ a_{i_2}$). By definition of the penalty associated with the method of Kemeny, it follows that

$$P(\mathscr{R}, \succ^1) = P(\mathscr{R}, \succ^2) - \#\{\succ_j \in \mathscr{R} \mid a_{i_1} \succ_j a_{i_2}\} + \#\{\succ_j \in \mathscr{R} \mid a_{i_2} \succ_j a_{i_1}\}.$$

As $\succ$ is the Condorcet ranking associated with $\mathscr{R}$ and $a_{i_1} \succ a_{i_2}$, it holds that

$$\#\{\succ_j \in \mathscr{R} \mid a_{i_1} \succ_j a_{i_2}\} > \#\{\succ_j \in \mathscr{R} \mid a_{i_2} \succ_j a_{i_1}\},$$

and, therefore, $P(\mathscr{R}, \succ^1) \leq P(\mathscr{R}, \succ^2)$. We conclude that, in case a Condorcet ranking exists, condition (iii') is satisfied.

As the set of rankings on $\mathscr{C}$ is finite, condition (iv) is trivially satisfied. Note that, like in most methods for the aggregation of rankings, the set of minimizers is not assured to be a singleton. For instance, it is not possible to decide which ranking on $\mathscr{C} = \{a, b, c\}$ should be the result of aggregating the list of rankings on $\mathscr{C}$ given in Table 3.

We conclude that the associated penalty function satisfies conditions (i), (ii), (iii) and (iv), while the additional condition (iii') is assured in case of existence of a Condorcet ranking.

Table 3. List of rankings on $\mathscr{C} = \{a, b, c\}$.

$a \succ b \succ c$
$a \succ c \succ b$
$b \succ a \succ c$
$b \succ c \succ a$
$c \succ a \succ b$
$c \succ b \succ a$

6 Conclusions

In this contribution, we have proposed a natural extension of the definition of a penalty function beyond its current confinement to real numbers based on the compatibility with a given betweenness relation. In particular, we have paid special attention to one of the best-known methods for the aggregation of rankings: the method of Kemeny. We have proved that the method of Kemeny can be understood as a prominent example of penalty-based aggregation outside the setting of real numbers satisfying all the proposed conditions.

References

1. Beliakov, G., Bustince, H., Fernandez, J.: The median and its extensions. Fuzzy Sets Syst. **175**, 36–47 (2011)
2. Bustince, H., Beliakov, G., Dimuro, G.P., Bedregal, B., Mesiar, R.: On the definition of penalty functions in data aggregation. In: Fuzzy Sets and Systems (2016). http://dx.doi.org/10.1016/j.fss.2016.09.011
3. Calvo, T., Beliakov, G.: Aggregation functions based on penalties. Fuzzy Sets Syst. **161**, 1420–1436 (2010)
4. Calvo, T., Mesiar, R., Yager, R.R.: Quantitative weights and aggregation. IEEE Trans. Fuzzy Syst. **12**, 62–69 (2004)
5. Condorcet, M.: Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix, De l'Imprimerie Royale, Paris (1785)
6. De Miguel, L., Campión, M.J., Candeal, J.C., Induráin, E., Paternain, D.: Pointwise aggregation of maps: its structural functional equation and some applications to social choice theory. In: Fuzzy Sets and Systems (in press). http://dx.doi.org/10.1016/j.fss.2016.05.010
7. Gagolewski, M.: Penalty-based aggregation of multidimensional data. In: Fuzzy Sets and Systems (in press). http://dx.doi.org/10.1016/j.fss.2016.12.009
8. Kemeny, J.G.: Mathematics without numbers. Daedalus **88**(4), 577–591 (1959)
9. Kendall, M.G.: A new measure of rank correlation. Biometrika **30**, 81–93 (1938)
10. Pasch, M.: Vorlesungen über neuere Geometrie, vol. 23. Teubner, Leipzig, Berlin (1882)
11. Pérez-Fernández, R., Alonso, P., Díaz, I., Montes, S., De Baets, B.: Monotonicity as a tool for differentiating between truth and optimality in the aggregation of rankings. J. Math. Psychol. **77**, 1–9 (2017)

12. Pérez-Fernández, R., Rademaker, M., Alonso, P., Díaz, I., Montes, S., De Baets, B.: Monotonicity-based ranking on the basis of multiple partially specified reciprocal relations. In: Fuzzy Sets and Systems (in press). http://dx.doi.org/10.1016/j.fss.2016.12.008
13. Pérez-Fernández, R., Rademaker, M., De Baets, B.: Monometrics and their role in the rationalisation of ranking rules. Inf. Fusion **34**, 16–27 (2017)
14. Pitcher, E., Smiley, M.F.: Transitivities of betweenness. Trans. Am. Math. Soc. **52**(1), 95–114 (1942)
15. Wilkin, T., Beliakov, G.: Weakly monotonic averaging functions. Int. J. Intell. Syst. **30**(2), 144–169 (2015)
16. Yager, R.R.: Toward a general theory of information aggregation. Inf. Sci. **68**, 191–206 (1993)
17. Yager, R.R., Rybalov, A.: Understanding the median as a fusion operator. Int. J. Gen. Syst. **26**, 239–263 (1997)

Is Fuzzy Number the Right Result of Arithmetic Operations on Fuzzy Numbers?

Andrzej Piegat[1] and Marek Landowski[2](✉)

[1] Faculty of Computer Science and Information Systems, West Pomeranian University of Technology, Zolnierska Street 49, 71-210 Szczecin, Poland
apiegat@wi.zut.edu.pl

[2] Department of Mathematical Methods, Maritime University of Szczecin, Waly Chrobrego 1-2, 70-500 Szczecin, Poland
m.landowski@am.szczecin.pl

Abstract. Present versions of fuzzy arithmetic (FA) are not ideal. For some computational problems they deliver credible results. However for many other problems the results are less credible or sometimes clearly incredible. Reason of this state of matter is the fact that present FA-versions partially or fully (depending on a method) do not possess mathematical properties that are necessary for achieving correct calculation results as: distributivity law, cancellation law, neutral elements of addition and multiplication, property of restoration, possibility of decomposition of calculation in parts, ability of credible equations' solving, property of delivering universal algebraic solutions, possibility of formula transformation, and other. Lack of above properties is, in the authors' opinion, caused by incorrect assumption of all existing FA-versions that result of arithmetic operations on unidimensional fuzzy intervals is also a unidimensional fuzzy interval. In the paper authors show that the correct result is a multidimensional fuzzy set and present a fuzzy arithmetic based on this proposition, which possess all necessary mathematical properties and delivers credible results.

Keywords: Fuzzy arithmetic · Fuzzy computations · Uncertainty theory · Granular computing · Soft computing · Artificial intelligence

1 Introduction

Fuzzy arithmetic is very important for uncertainty theory [2], granular computing [18], soft computing [16] and generally for artificial intelligence [11] because human thinking uses information granules that can be modeled as fuzzy sets. Motivation of this paper is explanation that fuzzy interval is not solution of arithmetic operations on fuzzy intervals and also explanation of other basic problems of fuzzy arithmetic [6,9,18,19] and of the sense of the calculation result because incorrect understanding of it leads to many misunderstandings

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_17

and computational paradoxes in fuzzy arithmetic (shortly: FA, F-arithmetic). There exist few versions of FA mentioned below.

FA1. Fuzzy arithmetic based on Zadeh's extension principle [6,9,19].
FA2. Fuzzy arithmetic of decomposed fuzzy numbers based on α-cuts of fuzzy sets and on interval arithmetic [6,18]. Because there exist few versions of interval arithmetic also there exist few versions of FA2. However, the mostly used is the standard interval arithmetic of Moore [8,17,18].
FA3. Left-Right (L-R) version of fuzzy arithmetic [4].
FA4. Advanced fuzzy arithmetic based on transformation method or on extended transformation method [6].
FA5. Constrained fuzzy arithmetic [10].
FA6. Ordered fuzzy numbers arithmetic [12].

Each of the cited F-arithmetic versions assumes that result of arithmetic operations $\{+, -, \times, /\}$ on unidimensional fuzzy numbers (shortly FNs, F-numbers) also is a unidimensional FN. E.g. result of addition of two triangular FNs $A = (1, 2, 3)$ and $B = (3, 4, 5)$ is the number $X = (4, 6, 8)$ shown in Fig. 1. when FA2-version is used.

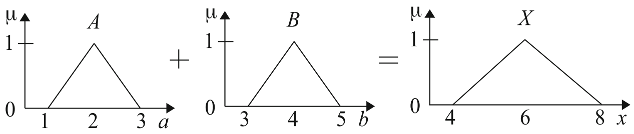


Fig. 1. Example of addition of triangular fuzzy numbers A and B with use of existing versions of FA: the result is also a triangular fuzzy number.

The addition result in Fig. 1 is also a triangular FN. This result seems intuitively correct. It is also consistent with axiom of closure [1,3] that for triangular FNs can be formulated as follows: "the set of triangular FNs has closure under an operation (e.g. of addition) if performance of that operation on numbers of this set always produces a member of the same set, i.e. a triangular FN". Because the closure axiom (property) seems intuitively obvious many authors of existing FA-versions have assumed that result of operations on FN-s also is a FN. However, using the axiom of closure results in many calculation paradoxes and weaknesses of FA. E.g. in [5,28] Dymova and Sevastjanov describe phenomenon of "multiple results". It consists in achieving different calculation results for one and the same system of data processing if contemporary FA or interval arithmetic is used. Let us assume that a system realizes addition $a + b = x$, where a and b are input values and x is the output value. Mathematical model (M-model) of the system can be expressed not only as (1) $a + b = x$ but also in form of equivalent M-models (2) $a = x - b$, (3) $b = x - a$, and (4) $a + b - x = 0$. If crisp input values are known, e.g. $a = 2$ and $b = 4$ then on the basis of all 4 models we achieve the same result $x = 6$. This result satisfies principle of solution universality. The universal solution does not depend on the form of M-model used in calculations. In conventional arithmetic of crisp numbers this principle is satisfied. However, let us assume that not crisp but only approximate input values are known in form of triangular FNs: $A = (1, 2, 3)$ and $B = (3, 4, 5)$ and that we want to calculate their addition result X. With use of FA2 for fuzzy extension (F-extension) $A + B = X$ of the system dependence $a + b = x$ we achieve the result $X = (4, 6, 8)$ shown in Fig. 1. For second F-extension $A = X - B$ the result $X = (0, 3, 6)$ is achieved. For third F-extension $B = X - A$ the result is $X = (0, 3, 6)$ and for forth extension $A + B - X = 0$ we achieve the result $X = (8, 6, 4)$ being improper fuzzy number [8] in which the lower limit 8 is greater than the upper limit 4. Which of the achieved results is correct? It can easily be checked that no one of them is universal result which satisfies all four F-extensions $A + B = X$, $A = X - B$, $B = X - A$, $A + B - X = 0$. Then, one can ask: does there exists a universal result of the addition operation of two FNs in general? Yes, it exists, but, as it will be shown later, this result is not an ordinary F-number. Existing FA-versions with FN-results have also many further weaknesses which hamper effective calculations. E.g., in the present FA and IA the interval difference $X = A - B$ can be calculated in two ways: as ordinary difference $X = A - B$ and as Hukuhara difference X^H [7,18,29]. Hukuhara difference is calculated from equation $A = X^H + B$ and has smaller uncertainty than ordinary difference X. Let us use two triangle FNs : $A = (0, 3, 5)$ and $B = (2, 3, 6)$. Then the ordinary difference $X = (-6, 0, 1)$ and Hukuhara difference $X^H = (-2, 0, 1)$. One can ask: which difference is true and correct? This situation is strange and seems illogical. Next disadvantage of FA consists in that. the version FA2 and other versions have not inverse elements of addition $-X$ and of multiplication $1/X$ that satisfy conditions (1) and (2).

$$X - X = 0 \tag{1}$$

$$X/X = 1 \tag{2}$$

In FA2-arithmetic a very important distributivity law (3) does not hold.

$$X(Y + Z) = XY + XZ \tag{3}$$

Without this law formula transformations can not be made because it can change calculation result. If X, Y, Z are triangle F-sets and $X = (1, 2, 3)$, $Y = (2, 4, 5)$, $Z = (-2, 0, 1)$ then with FA2 we achieve following results: $X(Y + Z) = (0, 8, 18)$ and $XY + XZ = (-6, 8, 18)$. As can be seen $X(Y + Z) \neq XY + XZ$ in this case. Similarly in FA2 cancellation law (4) does not hold, what also can cause calculative paradoxes.

$$\text{IF } (XZ = YZ) \text{ THEN } (X = Y) \tag{4}$$

Why FA2 and other FA-versions do not have such important mathematical properties? As it will be shown later, the reason is the assumption that result of arithmetic operations on F-numbers is also a F-number. An important feature required from any FA is not only correct realization of basic arithmetic operations $\{+, -, \times, /\}$ but also ability of solving arithmetic equations as e.g.

$A + B = X$, $A + BX + CX^2 = 0$, etc. Solving of even the simplest fuzzy equations as $A + X = B$ at present is difficult and solutions delivered by particular existing methods [16,23] are different. To verify whether a particular FA delivers correct solution X of an equation conditions of the correct solution should be determined. One of important conditions is that solution X should have property of algebraic solution. E.D. Popova in [27] gives definition of algebraic solution: "interval algebraic solution of a linear equation is an interval (interval vector) that substituting it into the equation and performing all interval operations results in valid equality". It should be noted that Popova also assumes that result of interval equation should be an interval. Further on a little corrected definition formulated by the paper authors will be given: "Universal algebraic solution of a fuzzy equation is a solution that after substituting it into the equation and performing fuzzy arithmetic operations according to rules of this arithmetic gives equality of both sides of the equation, independently of mathematical equation form used in calculations". Universality of solution X means e.g. in the case of equation $A + X = B$, that the solution should satisfy also other alternative forms of the equation $A = X - B$, $B = X - A$, $A + B - X = 0$. It should be noted that the proposed definition of algebraic solution does not assume that solution of a fuzzy equation is a fuzzy number.

2 Relationship Between Fuzzy and Interval Arithmetic

One of realization methods of FA-operations is decomposition of FNs on α-cuts [6,18]. Further on, instead of the name α-cuts the name μ-cut will be used. Let D be a set of domain and x be an element in D. Then a fuzzy set A in D is characterized by (5),

$$A = \{x, \mu(x); x \in D\} \tag{5}$$

where $\mu(x)$ is the grade of membership of x in A. For an ordinary set $\mu(x)$ is either 0 or 1 while for a fuzzy set $\mu(x) \in [0, 1]$. A μ^*-cut denoted by A_{μ^*} of a fuzzy set A is an ordinary set of elements with membership not less than μ^* for $0 \leq \mu \leq 1$, (6).

$$A_{\mu^*} = \{x \in D; \mu(x) \geq \mu^*\} \tag{6}$$

Let A and B be two fuzzy numbers and A_{μ^*} and B_{μ^*} be their intervals of confidence for the level μ^*, $0 \leq \mu \leq 1$. Then, we can write (7),

$$A_{\mu^*} * B_{\mu^*} = X_{\mu^*} \tag{7}$$

where $*$ is one of arithmetic operations $\{+, -, \times, /\}$ and X_{μ^*} is the operation result achieved with I-arithmetic. As can be seen from above, arithmetic operation on F-numbers can be decomposed on set of operations on intervals representing μ-cuts made on various levels. It means that F-arithmetic is strongly connected with I-arithmetic which is its basis. If we possess a correct I-arithmetic then we can create correct F-arithmetic. According to all versions of present I-arithmetic [8,17,18] result X_μ of arithmetic operations on μ-cuts A_μ and B_μ being intervals is also an interval. As it will be shown this interpretation is incorrect.

3 What Is the Result of Arithmetic Operations on Ordinary Intervals?

Let us begin the problem analysis from discrete intervals. Let us consider a simple example.

Example 1. On Sunday in village C a dance party is organized. It is known that from village A maximally 2 persons will come on motor bike. However, it also is possible that nobody will come. Similar situation is with village B. How many persons will come from both villages? Let us denote the real number of persons from village A by a and from village B by b. We know that $a \in A = \{0, 1, 2\}$ and $b \in B = \{0, 1, 2\}$. A and B are sets of persons that can come from village A and B. Number a of persons from village A does not depend on number b of persons from village B. Let us notice that true values of a and b are not known for us. Thus, they are unknown variables. Hence, it is not possible to predict with full certainty the true sum $x = a + b$ of person that will come from both villages. All what we can do is determining all possibilities. Set$X$ of possible point-results is shown in Fig. 2.

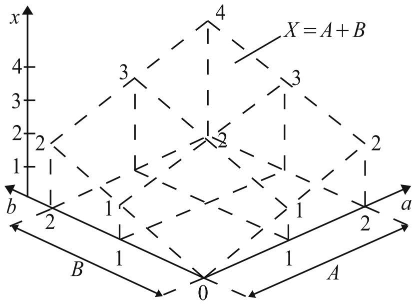


Fig. 2. Visualization of addition of discrete intervals $A = \{0, 1, 2\}$ and $B = \{0, 1, 2\}$.

The result set $X = \{X_1, \ldots, X_9\}$ consists of 9 possible single results X_i, $i = 1, \ldots, 9$, of conditional character $(x|a, b)$ which linguistically can be interpreted as rules (8).

$$(x|a, b): \text{ IF } (a = a^*) \text{ AND } (b = b^*) \text{ THEN } (x = a^* + b^*) \tag{8}$$

E.g. for $a = 1$ and $b = 2$ the single conditional addition is expressed by (9).

$$(3|1, 2): \text{ IF } (a = 1) \text{ AND } (b = 2) \text{ THEN } (x = 3) \tag{9}$$

Full set of single conditional results for the considered addition has form of (10).

$$\begin{aligned} X = \{&(0|0, 0), (1|1, 0), (2|2, 0), (1|0, 1), (2|1, 1,), (3|2, 1), \\ &(2|0, 2), (3|1, 2), (2|0, 2), (3|1, 2), (4|2, 2)\} \end{aligned} \tag{10}$$

The above example shows that in the case of addition of discrete intervals the result is not a discrete interval but set X of addressed, possible, conditional results existing not in 1D-space but in 3D-space. Because we do not know real values of variables a and b therefore we can only find conditional point solutions as (9). Full addition of discrete intervals A+B is given by (11).

$$\begin{aligned} A+B &= X \\ \{0,1,2\}+\{0,1,2\} &= \{(0|0,0),(1|1,0),(2|2,0),(1|0,1),(2|1,1,),(3|2,1), \\ & \quad (2|0,2),(3|1,2),(2|0,2),(3|1,2),(4|2,2)\} \end{aligned} \tag{11}$$

Analysis of the result set X, formula (11) or Fig. 2, shows that this set contains one solution for $x = 0$, 2 solutions for $x = 1$, 3 solutions for $x = 2$, 2 solutions for $x = 3$, and 1 solution for $x = 4$. The solution number in particular subsets of X can be called *cardinality* ($card.x$) or *frequency* ($freq.x$) [13]. Cardinality is a valuable representative or indicator of result sets. Distribution of cardinality for the considered example is shown in Fig. 3.

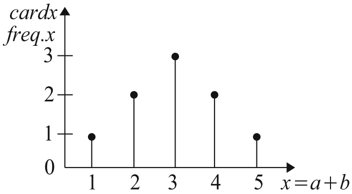


Fig. 3. Not normalized distribution of cardinality (frequency) of resul subsets with identical result value $x = a + b = const$.

Cardinality distribution from Fig. 3 can be interpreted as a priori information (not verified experimentally) which value x of the result set has smaller or greater chance to occur. $Card.x$ is not the direct result of interval addition. It is only an indicator (simplified information piece) about the result set X. The next indicator (representative) of this set can be its span $s(x)$, see Fig. 3 and formula (12).

$$s(X) = \{(\min x), (\min x)+1, (\min x)+2, \ldots, (\max x)\} = \{0,1,2,3,4\} \tag{12}$$

It should be noticed that according to present interval arithmetic versions just this indicator is understood as direct result of I-arithmetic operations. It is a mistake. Span is only one of many possible indicators of the result set similarly as cardinality distribution or center of gravity of the set. At the very most it can be interpreted as "secondary" result. Treating it as direct result is reason of calculation paradoxes observed in I-arithmetic and F-arithmetic [5,20,28].

Third valuable indicator (representative) of the result set X is its *center of gravity* ($COG(X)$). If the set X contains n single conditional result x_i then $COG(X)$ can be calculated from (13).

$$COG(X) = \sum_{i=1}^{n} x_i \cdot cardx_i \Bigg/ \sum_{i=1}^{n} cardx_i \tag{13}$$

In the considered discrete example position of $COG(X)$ is given by (14).

$$COG(X) = \frac{1 \cdot 0 + 2 \cdot 1 + 3 \cdot 2 + 2 \cdot 3 + 1 \cdot 4}{1 + 2 + 3 + 2 + 1} = 2 \tag{14}$$

The center of gravity informs us about the average weighted value of the result x we can hope on the basis of possessed knowledge.

Now, let us consider continues intervals and realization of arithmetic operations leading to achievement of the complete result set. If set A is set of variable a values such that $a \in [\underline{a}, \overline{a}] = A$ and set B is interval of variable b values such that $b \in [\underline{b}, \overline{b}] = B$ then all present I-arithmetic versions assume that result X of arithmetic operation $A * B = X$ is an interval, where $* = \{+, -, \times, /\}$. E.g., with such assumptions for addition of $A = [0, 2]$ and $B = [0, 2]$ we achieve $[0, 2] + [0, 2] = [0, 4]$. This notation of the arithmetic operation is incorrect and imprecise. First of all it is not addressed. This notation suggests that we can take any number from interval $A = [0, 2]$ and any number from $B = [0, 2]$ and can get any number from $X = [0, 4]$. E.g., a can be equal 0.1, b can be equal to 0.2 and the addition result can be equal to e.g. 3.5 ($0.1 + 0.2 = 3.5$). The equality sign in equation $A + B = X = [\underline{x}, \overline{x}]$ is incorrect. More sensible though also incorrect could seem the notation $(a|a \in A) + (b|b \in B) \in X$ e.g.: $(a|a \in [0, 2]) + (b|b \in [0, 2]) \in [0, 4]$. About incorrectness of the notation used in present I-arithmetic have also written W. Lodwick and D. Dubois in their comprehensive paper [14], where they propose the sign $\approx$ instead of equation sign $=$. To elaborate correct I -arithmetic, addressing all values inside an interval should be introduced. Such addressing can be realized with RDM (Relative-Distance-Measure) variable α, $\alpha \in [0, 1]$ used in RDM I-arithmetic [20–22]. M-model of a single variable value contained in interval $X = [\underline{x}, \overline{x}]$ has form of (15).

$$x = \underline{x} + \alpha_x(\overline{x} - \underline{x}),\ \alpha_x \in [0, 1] \tag{15}$$

E.g. M-model of a single variable-value $x \in [2, 5]$ has form $x = 2 + 3\alpha_x$, $\alpha_x \in [0, 1]$. In (15) each value contained in the interval is not anonymous because it can be distinguished by a value of RDM-variable α assigned to it. M-model (15) is not only a model of a single variable-value, it is can be interpreted a model of the true, though precisely unknown value of the variable that occurred in a real system. If we know that an uncertain variable value $a \in [\underline{a}, \overline{a}]$ and $b \in [\underline{b}, \overline{b}]$ and variables a and b are independent then any operation $*$ from the set $\{+, -, \times, /\}$ can be realized according to (16).

$$\begin{aligned} &A * B = X \\ &A * B:\ a * b = [\underline{a} + \alpha_a(\overline{a} - \underline{a})] * [\underline{b} + \alpha_b(\overline{b} - \underline{b})] = X(\alpha_a, \alpha_b), \\ &\alpha_a, \alpha_b \in [0, 1] \end{aligned} \tag{16}$$

E.g., if $a \in [0, 2]$ and $b \in [0, 2]$, then their addressed values have form: $a = 0 + 2\alpha_a$, $b = 0 + 2\alpha_b$, $\alpha_a, \alpha_b \in [0, 1]$. It can easily be noticed that the sum X

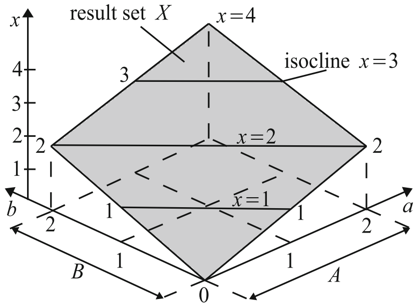


Fig. 4. 3-dimensional result set of interval addition $A + B = [0, 2] + [0, 2] = X$ with isoclines of constant result x-values.

does not exists in 1D-space but in multidimensional 3D-space $\alpha_a \times \alpha_b \times X$. The set of possible results X of interval addition $A = [0, 2]$ and $B = [0, 2]$ is shown in Fig. 4.

Result set $X(\alpha_a, \alpha_b)$ is the complete set of infinite number of possible single results x. Formula (17) can be used for generating any possible conditional result $(x|a, b) = (x|\alpha_a, \alpha_b)$. E.g. for $\alpha_a = 0.3$ and $\alpha_b = 0.7$ we get $a = 0.6$ and $b = 1.4$, $x = 2$, i.e. we get the triple $(x|a, b) = (2|0.6, 1.4)$ or $((x|\alpha_a, \alpha_b) = (2|0.3, 0.7)$. On the basis of RDM I-arithmetic we can correctly realize arithmetic operations for particular μ-cuts of fuzzy numbers. This arithmetic possess similar mathematic properties as arithmetic of crisp numbers. It possess property of *commutativity* $(X+Y = Y+X, XY+YX)$, property of *associativity* $[X+(Y+Z) = (X+Y)+Z]$, $[X(YZ) = (XY)Z)]$, it possess *neutral elements of addition and multiplication* $(X + 0 = 0 + X = X)$,$(X \cdot 1 = 1 \cdot X)$, *distributive law* $[X(Y + Z) = XY + XZ]$, *cancellation law for addition and multiplication* $(X + Z = Y + Z \Rightarrow X = Y)$, $(XZ = YZ \Rightarrow X = Y)$. RDM I-arithmetic possess also a very important *property of restoration for addition*: if after adding intervals $A+B$ the result X has been achieved then knowing X and B restoration of A should be possible $(A + B = X \Rightarrow X - B = A$ and $X - A = B)$. Similarly *restoration property for multiplication* has form:$(AB = X \Rightarrow X/B = A$ and $X/A = B)$. Both restoration properties possess crisp-number arithmetic. E.g. $(2 + 3 = 5 \Rightarrow 5 - 3 = 2$ and $5 - 2 = 3)$. However, no I-arithmetic assuming that arithmetic operation result is an interval possess this property. E.g., let $A = [0, 2]$ and $B = [0, 2]$. Then, according to SI-arithmetic $A + B = [0, 4]$. However, $X - B = [0, 4] - [0, 2] = [-2, 4] \neq A = [0, 2]$ and $X - A = [0, 4] - [0, 2] = [-2, 4] \neq B = [0, 2]$. In the case of RDM I-arithmetic we have: $a = 2\alpha_a$ and $b = 2\alpha_b$, $x - b = (2\alpha_a + 2\alpha_b) - 2\alpha_b = 2\alpha_a = a$. Similar result is achieved for multiplication. Thus, RDM I-arithmetic possess property of restoration thanks to addressing variable values taking part in operations. Figure 4 presents a 3D-result of addition operation realized with use of RDM I-arithmetic. For this result 3 basic indicators (representatives) can be determined: *cardinality distribution* $card(x)$ of possible single results $x = const$, *span* $s(x)$ of possible x results, and *center of gravity* $COG(x)$ of possible x-results (average weighted value of the results). Cardinality is a measure of frequency of occurrence [13] of triples $(x|a, b)$ with the same x-value. This frequency measure is length of particular isoclines shown in Fig. 4. In the case of addition the length can be calculated with formula (17).

$$L(x = const) = \sqrt{2}(a_{\max}(x = const) - a_{\min}(x = const)) \tag{17}$$

The cardinality distribution is shown in Fig. 5.

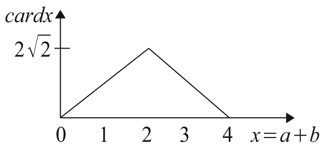


Fig. 5. Distribution of cardinality (frequency measure) of possible addition results $a + b = x$ of elements of intervals $A = [0, 2]$ and $B = [0, 2]$ as indicator (representative) of the 3D set $X(\alpha_a, \alpha_b)$ of possible results x.

Second indicator of the 3D-result set $X(\alpha_a, \alpha_b)$ is span $s(X)$ which can be calculated with (18).

$$s(X(\alpha_a, \alpha_b)) = [\min x(\alpha_a, \alpha_b), \max x(\alpha_a, \alpha_b)] \tag{18}$$

In the considered addition of intervals $A+B$ span of the result is $s(X) = [0, 4]$.

Third indicator (representative) of the set X of possible results is center of gravity $COG(X)$ which can be calculated from (19).

$$COG(X) = \int_{x_{\min}}^{x_{\max}} (x \cdot cardx)dx \Bigg/ \int_{x_{\min}}^{x_{\max}} (cardx)dx \tag{19}$$

In the considered example of interval addition $A + B$ position of the center of gravity of the result set X is equal to $COG(X) = 2$. The center and the span of the set X are shown in Fig. 6.

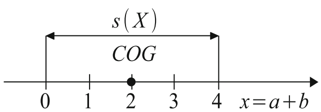


Fig. 6. Visualization of two indicators of the result set X of possible results of addition of two intervals $A+B = [0, 2]+[0, 2]$: span $s(X)$ and center of gravity of results $COG(x)$.

4 Arithmetic Operations on Fuzzy Intervals

Arithmetic operations on fuzzy intervals can separately be performed on chosen μ-cuts, e.g. for levels $\mu = 0, 1, 2/4, 3/4, 1$. Then, for the achieved result sets of particular μ-cuts their indicators can be determined and presented together. However, arithmetic operations can also be performed not separately but for the full range of membership $\mu \in [0, 1]$ if horizontal membership functions (HMFs, HM-functions) are used [23–26]. Horizontal membership functions were used for solving fuzzy differential equations of motion for the Boeing 747 [15]. Figure 7 shows example of trapezoidal and triangle membership function.

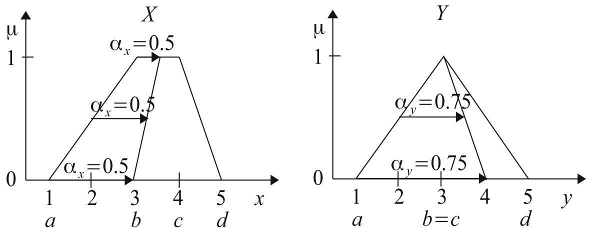


Fig. 7. Trapezoidal and triangle membership function with denotations.

If usual models of MFs have form $\mu = f_\mu(x)$ then horizontal MFs have the inverse form $x = f_x(\mu, \alpha_x)$. And so horizontal MF of the trapezoidal fuzzy set from Fig. 7 is determined by (20), where $\alpha_x \in [0, 1]$ is RDM variable.

$$x = [a + (b - a)\mu] + [(d - a) - \mu(d - a + b - c)]\alpha_x,\ \alpha_x \in [0, 1] \tag{20}$$

For parameter values as in Fig. 7 the horizontal MF is given by (21).

$$x = (1 + 2\mu) + (4 - 3\mu)\alpha_x,\ \alpha_x \in [0, 1] \tag{21}$$

Assuming any cut level μ a corresponding RDM interval model can be determined from (21). E.g. for $\mu = 0.5$ the cut of trapezoidal MF from Fig. 7 is described by formula $x = 2 + 2.5\alpha_x$, $\alpha_x \in [0, 1]$. The result can easily be checked on this figure. In the case of triangle MFs formula (22) concerning trapezoidal MFs also can be used after substituting $b = c$ because triangle is a special case of trapezoid, see (22).

$$y = [a + (b - a)\mu] + [(d - a) - \mu(d - a)]\alpha_y,\ \alpha_y \in [0, 1] \tag{22}$$

For numerical values as in Fig. 7 M-model of horizontal MF is given by (23).

$$y = (1 + 2\mu) + 4(1 - \mu)\alpha_y,\ \alpha_y \in [0, 1] \tag{23}$$

For the cut level $\mu = 0.5$ the interval horizontal model has form $x = 2 + 2\alpha_x$, $\alpha_x \in [0, 1]$. One can check in Fig. 7 that this model is correct. Operations $\{+, -, \times, /\}$ of fuzzy arithmetic can with use of horizontal MFs be performed very simply. If $x = f_x(\mu, \alpha_x)$ is the horizontal MF representing one uncertain variable value and $y = f_y(\mu, \alpha_y)$ represents second uncertain variable value and both variables are independent then operations of fuzzy arithmetic can be realized with formula (24).

$$X * Y = Z : x(\mu, \alpha_x) * y(\mu, \alpha_y) = z(\mu, \alpha_x, \alpha_y),\ \mu, \alpha_x, \alpha_y \in [0, 1] \tag{24}$$

E.g., if trapezoidal fuzzy set $X = (1, 3, 4, 5)$ from Fig. 7 is described by horizontal MF $x = (1 + 2\mu) + (4 - 3\mu)\alpha_x,\ \alpha_x \in [0, 1]$ and the triangle fuzzy set $Y = (1, 3, 5)$ is described by the function $y = (1 + 2\mu) + 4(1 - \mu)\alpha_y,\ \alpha_y \in [0, 1]$ then their sum $X + Y = Z$ is expressed by (25).

$$X + Y = Z : z = [(1 + 2\mu) + (4 - 3\mu)\alpha_x] + [(1 + 2\mu) + 4(1 - \mu)\alpha_y] = z(\mu, \alpha_x, \alpha_y) \tag{25}$$

As can be seen from (25) the result set Z exists in 4D-space $Z \times_m u \times \alpha_x \times \alpha_y$ and hence it can not be visualized directly. However, its particular μ-cuts can. E.g., for cut at level $\mu = 0$ function (25) takes form of (26).

$$Z(\mu = 0) = (1 + 4\alpha_x) + (1 + 4\alpha_y),\ \alpha_x, \alpha_y \in [0, 1] \tag{26}$$

For particular μ-cuts one can determine indicators such as cardinality distribution of the result set, span of the set, and center of gravity of the set. Relatively easy is determining of the span $s(\mu)$ of the result set Z, formula (27).

$$s(\mu) = [\min_{\alpha_x, \alpha_y} z, \max_{\alpha_x, \alpha_y} z],\ \mu, \alpha_x, \alpha_y \in [0, 1] \tag{27}$$

In the addition example of two fuzzy numbers $X + Y = Z$ determined by formula (25) $\min z$ is achieved for $\alpha_x = \alpha_y = 0$ and $\max z$ for $\alpha_x = \alpha_y = 1$. Hence, the span indicator of Z is determined by formula (28). It is also shown in Fig. 8.

$$s(\mu) = [2 + 4\mu, 10 - 3\mu] \tag{28}$$

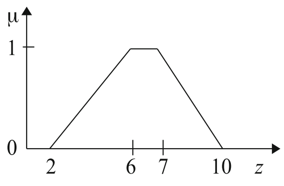


Fig. 8. Distribution of the span indicator $s(\mu)$ of the addition result set $Z = X + Y$ for various levels of membership μ.

Finishing the paper authors would like to give a convincing example proving that solution of arithmetic operations with fuzzy intervals cannot be a fuzzy interval. Let the example be the fuzzy interval equation $A + X = B$ with triangle fuzzy sets $A = (1, 3, 5)$ and $B = (7, 7.5, 8)$. One can try to solve this equation with arithmetic FA2 in which fuzzy sets are decomposed in ordinary intervals for particular μ-levels and one tries to determine interval solutions for each μ-level. For checking whether solution can be an interval it is sufficient to analyze the problem for the support level $\mu = 0$. For this level the fuzzy equation $A + X = B$ takes the form $A_0 + X_0 = B_0$ ($[1, 5] + X_0 = [7, 8]$). Let us assume that the solution X_0 is an interval $[\underline{x_0}, \overline{x_0}]$. Then the equation takes the form $[1, 5] + [\underline{x_0}, \overline{x_0}] = [7, 8]$. After solving this equation we achieve the "solution" $X_0 = [\underline{x_0}, \overline{x_0}] = [6, 3]$, which is improper interval in which lower limit is greater than the upper one: $\underline{x_0} > \overline{x_0}$. Such intervals cannot be realized in the practice as solution [14] because each value x_0 inside the considered interval should satisfy the condition $\underline{x_0} \le x_0 \le \overline{x_0}$ which in this case means $x_0 \ge 6$ and $x_0 \le 3$ that is impossible. There exists no proper interval which could be solution of the considered equation which means that in terms of any existing arithmetic versions the considered equation has no solution. However, with use of the "common sense" one can easily find many possible point solutions. E.g. for $a = 4$ and $b = 7.5$ the solution is $x_0 = 3.5$. The real solution set X_0 can be found if one accepts the truth that the solution X_0 is not a 1D-interval but a 3-dimensional information granule consisting of triples $(x_0|a, b)$. This solution can be found with use of RDM I-arithmetic. In terms of this arithmetic model of interval $A = [1, 5]$ has form $a = 1 + 4\alpha_a$, $\alpha_a \in [0, 1]$, and model of interval $B = [7, 8]$ has form $b = 7 + \alpha_b$, $\alpha_b \in [0, 1]$. The RDM IA-solution of the equation has form of a 3D granule $x = b - a = (7 + \alpha_b) - (1 + 4\alpha_a)$, with $\alpha_a, \alpha_b \in [0, 1]$. Choosing any allowable values of α_a, α_b we achieve corresponding values of $a = 1 + 4\alpha_a$, $b = 7 + \alpha_b$, and corresponding possible value of $x = b - a$. E.g. for $\alpha_a = 0.75$, $\alpha_b = 0.5$ the result is $x = 3.5$ or more formally $(x|a, b) = (3.5|4, 7.5)$. Concluding: there exists RDM IA-solution X_0 of the equation $[1, 5] + X_0 = [7, 8]$ for which no interval solution exists. Figure 9 shows the set X_0 of possible point-solutions x_0 in the space $A \times B$ and indicators of this solutions cardinality distribution, span and center of gravity of the solutions' set.

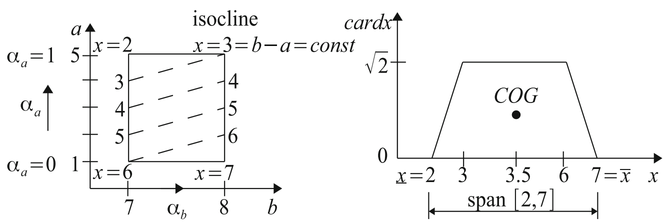


Fig. 9. Visualization of the solution set X_0 of interval equation $[1, 5] + X_0 = [7, 8]$ in the space $A \times B$ and indicators of this solution: cardinality (frequency) distribution, span and center of gravity.

5 Conclusions

It was shown in the paper that the assumption speaking that results of F-arithmetic operations on F-intervals is a F-interval is incorrect. This assumption is typical for all existing FA-versions and frequently leads to incredible problems' solutions and calculative paradoxes described in scientific papers. It was explained in the paper that the correct result is a multidimensional fuzzy set and its fuzzy span is only one of its possible indicators. The paper presented also a fuzzy RDM arithmetic based on this proposition which applies horizontal membership functions. This multidimensional arithmetic possess all mathematical properties necessary for delivering credible calculation results.

References

1. Bader, F., Nipkov, T.: Term Rewriting and All That. Cambridge University Press, Cambridge (1998)
2. Baoding, L.: Uncertainty Theory, 4th edn. Springer, Berlin (2015)
3. Birkhoff, G.: Lattice Theory. Colloquium Publications, vol. 25. American Mathematical Society (1967)
4. Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. J. Syst. Sci. **9**(6), 613–626 (1978)
5. Dymova, L.: Soft Computing in Economics and Finance. Springer, Berlin (2011)
6. Hanss, M.: Applied Fuzzy Arithmetic. Springer, Berlin (2005)
7. Hukuhara, M.: Integration des applications measurables don't la valeur est un compact convexe. Funkcialaj Ekvacioj **10**, 205–223 (1967)
8. Kaucher, E.: Interval analysis in extended interval space IR. Computing **2**(Supplement), 33–49 (1980)
9. Kaufmann, A., Gupta, M.M.: Introduction to Fuzzy Arithmetic. Van Nostrand Reinhold, New York (1991)
10. Klir, G.J.: Fuzzy arithmetic with requisite constraints. Fuzzy Sets Syst. **91**, 165–175 (1997)
11. Konar, A.: Artificial Intelligence and Soft Computing. CRC Press, London, New York (2000)
12. Kosinski, W., Prokopowicz, P., Slezak, D.: Ordered fuzzy numbers. Bull. Pol. Acad. Sci. Ser. Sci. Math. **51**(3), 327–338 (2003)
13. Kovalerchuk, B., Kreinovich, V.: Comparison of applied tasks with intervals, fuzzy sets and probability approaches. In: Proceedings of 2016 IEEE International Conference on Fuzzy Systems (FUZZ), 24–29 July 2016, Vancouver, Canada, pp. 1478–1483 (2016)
14. Lodwick, W.A., Dubois, D.: Interval linear systems as necessary step in fuzzy linear systems. Fuzzy Sets Syst. **281**, 227–251 (2015)
15. Mazandarani, M., Pariz, N., Kamyad, A.V.: Granular differentiability of fuzzy-number-valued functions. IEEE Trans. Fuzzy Syst. (2017). doi:10.1109/TFUZZ.2017.2659731
16. Mazarbhuiya, F.A., Mahanta, A.K.,Baruah, H.K.: Solution of the fuzzy equation $A + X = B$ using the method of superimposition. Appl. Math. **2**, 1039–1045. http://www.SciRP.org/journal/am
17. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. Society for Industrial and Applied Mathematics, Philadelphia (2009)

18. Pedrycz, W., Skowron, A., Kreinovich, V.: Handbook of Granular Computing. Wiley, Chichester (2008)
19. Piegat, A.: Fuzzy modeling and control. Physica-Verlag, A Springer-Verlag Company, Heidelberg, New York (2001)
20. Piegat, A., Landowski, M.: Is the conventional interval-arithmetic correct? J. Theor. Appl. Comput. Sci. **6**(2), 27–44 (2012)
21. Piegat, A., Landowski, M.: Two interpretations of multidimensional RDM interval arithmetic - multiplication and division. Int. J. Fuzzy Syst. **15**(4), 486–496 (2013)
22. Piegat, A., Landowski, M.: Correctness checking of uncertain-equation solutions on example of the interval-modal method. In: Atanasov, T., et al. (ed.) Modern approaches in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Foundations, SRI PAS-IBS PAN, Warsaw, Poland, vol.1, pp. 159–170 (2014)
23. Piegat, A., Plucinski, M.: Fuzzy number addition with the application of horizontal membership functions. Sci. World J., 1–16, Article ID: 367214 (2015). (Hindawi Publishing Corporation)
24. Piegat, A., Landowski, M.: Horizontal membership function and examples of its applications. Int. J. Fuzzy Syst. **17**(1), 22–30 (2015)
25. Piegat, A., Plucinski, M.: Computing with Words with the use of inverse RDM models of membership functions. Int. J. Appl. Math. Comput. Sci. **25**(3), 675–688 (2015)
26. Piegat, A., Landowski, M.: Fuzzy arithmetic type 1 with horizontal membership functions. In: Kacprzyk, J., Kreinovich, V. (eds.) Uncertainty Modeling. Springer (2017)
27. Popova, E.D.: Algebraic solutions to a class of interval equations. J. Univ. Comput. Sci. **4**(1), 48–67 (1998)
28. Sevastjanov, F.A., Dymova, L.: A new method for solving interval and fuzzy equations: linear case. Inf. Sci. **179**, 925–937 (2009)
29. Stefanini, L.: A generalization of Hukuhara difference and division for interval and fuzzy arithmetic. Fuzzy Sets Syst. **161**, 1564–1584 (2010)

Analysis of Different Proposals to Improve the Dissemination of Information in University Digital Libraries

Carlos Porcel[1(✉)], Alberto Ching-López[2], Alvaro Tejeda-Lorente[2], Juan Bernabé-Moreno[2], and Enrique Herrera-Viedma[2]

[1] Department of Computer Science, University of Jaén, 23071 Jaén, Spain
cporcel@ujaen.es

[2] Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain

Abstract. Currently the great advances in Web technologies are changing the process of access to information and the Web is one of the most important source of information. Furthermore, the Web influences the development of others media, for example, newspapers, journals, books, libraries, etc. In this paper we analyze its impact in the development of the university digital libraries. As well as on the Web, the information growth is a big problem for academic digital libraries, and similar tools can be applied in university digital libraries to provide users with access to the information. Given the importance of this aspect, in this paper we analyze and review different proposals that improve the processes of dissemination of information in these university digital libraries, promoting access to information of interest. These proposals manage to adapt access to information according to the needs and preferences of each user. As we can see in the literature, one of the techniques with the best results, is the application of recommender systems. Recommender systems are tools whose objective is to evaluate and filter the large amount of information available on the Web to assist users in their process of access to information. Thus, in this paper we analyze some proposals based on recommender system to help students, teachers and researchers to find research resources that can improve the services provided by the university digital libraries.

Keywords: Digital libraries · Dissemination of information · Recommender systems · Fuzzy linguistic modelling

1 Introduction

Digital libraries are collections of information that have associated services offered to the users community using a variety of technologies. The information collections can be scientific, business or personal data, and can be represented in

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_18

different formats such as digital text, images, audio, video, or any other digital media. This information can be digitalized or digitally born information and the services offered with this content can be varied, and can be offered to individuals or user communities. Internet access has made digital libraries more and more used by diverse communities for various purposes, in which sharing and collaboration have become important social elements. As digital libraries have become common spaces, and their contents and services are more varied, people expect more sophisticated services from their digital libraries [6,9,23]. The digital libraries are composed by human resources (staff) that are responsible for managing and enabling access to the most interesting documents for users, taking into account both their areas of interest and their needs [18]. The library staff search, evaluate, select, catalogue, classify, preserve and schedule the access to the digital documents [14]. Digital libraries have been incorporated into many environments, but we will focus on the academic context. specifically, we talk about *University Digital Libraries* (UDL), which provide information resources and services to students, faculty and staff in an environment that supports learning, teaching and research [7,25].

The exponential growth of Web sites and documents contributes to users not being able to find the information they are looking for in a simple and time-effective way. Users need tools to help them deal with the large amount of information available to them on the Web [12]. Therefore, search and mining techniques of the Web are becoming vital. Furthermore, the Web influences the development of others information media, for example, newspapers, journals, books, etc. and specifically the development of academic digital libraries [25]. As on the Web, the exponential growth of information is the mayor problem of these libraries because the employees have problems carrying out the tasks of delivering the information to the users. For this we can use Web context tools in UDL, to facilitate the tasks of employees and therefore improve access to information for students, teachers and researchers.

A traditional search function is an essential part of any digital library, but the frustration of users increase as their needs are more complex and the volume of information handled by the library grows. Digital libraries should move from being passive to being more proactive in offering and tailoring information for individuals and communities, and in supporting community efforts to capture, structure and share knowledge [6]. So, the digital libraries can anticipate the users' needs and recommend resources that could be of their interest. Given these features, in a UDL a service that is particularly important is the selective dissemination of information or filtering. Users develope their interests profile, so when new materials are added to the collection of information, the UDL can notify the users with relevant items [14]. Due to the problem of information overload, although there is a great abundance of information available, sometimes it is difficult to obtain useful or relevant information when necessary. When the users of a UDL try to access to useful information, they often obtain irrelevant information or information which does not meet their needs. So, users need easier access to the thousands of resources that are available but yet hard to find [17,24].

As on the Web, we can use recommender systems to facilitate the access to information. A *recommender system* attempts to discover information items that are likely of interest to a user. Recommender systems are especially useful when they identify information that a person was previously unaware of. Furthermore, recommender systems are personalized services because they may treat each user in a different way. These recommender systems play an important role in highly rated Web sites, such as Amazon,[1] YouTube,[2] Netflix,[3] Tripadvisor[4] or IMDb[5] [8].

The provision of personalized recommendations, requires that the system knows something about every user, such as the ratings provided by the users about the explored items [4,28]. This knowledge implies that the system must maintain users' profile containing the users' preferences or needs. But the way in which this information is acquired and exploited depends on the particular recommendation approach. The system could acquired implicit information about the users analyzing the users behavior, or the system might request the users insert explicitly their preferences. Another question to consider is what additional information is required by the system, and how this information is processed and managed to generate a list of personalized recommendations.

Following these ideas, in this paper we review and analyze different proposals, which favor the dissemination of information in UDL. Based on the success shown by the application of recommender system we focus on proposals based on these recommendation techniques [24]. Besides, these proposals also face the problem of the wide variety of representations and evaluations of information, which is more pronounced when users are part of the process, as is the case of UDL. Therefore, we also expose the fuzzy linguistic modelling that will help us to represent and efficiently manage the qualitative information present in the communication processes, as in previous proposals in which fuzzy approaches were applied [25,26]. Specifically, we analyze the multi-granular approach that gives us greater flexibility in the system-user interaction [16,19].

We analyzed four proposals, each of them improving the performance of the previous one. The first one proposes a fuzzy linguistic recommender system that recommends both specialized resources of the user interest area, and complementary resources that could be interesting to form multi-disciplinar groups [22]. The second one proposes a new method for acquiring the user profiles reducing the great effort of previous proposals; users provide their preferences on some research resources (by means of incomplete fuzzy linguistic preference relations) and from this information the system obtain their respective preference vectors on topics of interest [21]. The third one improves the previous proposals with

[1] http://www.amazon.es/.
[2] www.youtube.com/.
[3] www.netflix.com/.
[4] www.tripadvisor.es/.
[5] www.imdb.com/.

a recommender system which uses a memory to avoid the information overload problem still persistent in UDL; the main idea is to use previous selected items to make a new selection in a new recommendation round [20]. Finally, the last proposal faces the recommendations generation process about research resources as a task with two distinct elements: On one hand, finding research resources that are relevant to the users, and on the other hand, finding valid research resources from the standpoint of the quality of items [28].

The paper is structured as follows. Section 2 revises the preliminaries needed to understand the analyzed proposals. In Sect. 3 we analyze several proposals to improve the dissemination of information in digital libraries. Finally, some conclusions and future research are pointed out.

2 Preliminaries

2.1 Basis of Recommender Systems

Recommender systems try to guide the users in a personalized way towards suitable tasks among a wide range of possible options [4,28]. Personalized recommendations rely on some knowledge about the users, which might be tastes, preferences as well as the ratings of previously explored items. The way of acquiring this information may vary from implicit information, obtained analyzing users behavior, or explicit information, where users directly provide their preferences.

Other aspect to take care of is the way of generating recommendations. In the literature we can find them mainly pooled in two categories [4,27]. In the first one authors consider two different approaches: On one side, the *content-based approaches* generate the recommendations taking into account the characteristics used to represent the items and the ratings that a user has given to them. On the other side, the *collaborative approaches* generate recommendations using explicit or implicit preferences from many users, ignoring the items representation. The second one extends the categorization with another three approaches: *Demographic systems*, *Knowledge-based systems* and *Utility-based systems* [4].

Since each approach has certain advantages and disadvantages, depending on the scope settings. In order to combine different approaches to reduce the disadvantages of each one and to exploit their benefits, a widespread solution is the combination of approaches, known as *hybrid approach* [4].

2.2 Fuzzy Linguistic Approach

The fuzzy linguistic approach is a tool based on the concept of linguistic variable proposed by Zadeh [29]. This theory has given very good results to model qualitative information and it has been proven to be useful in many problems. We briefly describe the approaches used in the reviewed proposals.

The 2-Tuple Fuzzy Linguistic Approach. In order to reduce the loss of information of other methods such as classical or ordinal, in [10] was proposed a continuous model of information representation based on 2-tuple fuzzy linguistic modelling. To define it both the 2-tuple representation model and the 2-tuple computational model to represent and aggregate the linguistic information have to be established.

Let $S = \{s_0, ..., s_g\}$ be a linguistic term set with odd cardinality. We assume that the semantics of labels is given by means of triangular membership functions and consider all terms distributed on a scale on which a total order is defined. In this fuzzy linguistic context, if a symbolic method aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, ..., g\}$, we can represent β as a 2-tuple (s_i, α_i), where s_i represents the linguistic label, and α_i is a numerical value expressing the value of the translation between numerical values and 2-tuple: $\Delta(\beta) = (s_i, \alpha)$ and $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ [10].

In order to establish the computational model negation, comparison and aggregation operators are defined. Using functions Δ and Δ^{-1}, any of the existing aggregation operators can be easily be extended for dealing with linguistic 2-tuples without loss of information [10]. For instance arithmetic mean, weighted average operator or linguistic weighted average operator could be used.

Multi-granular Linguistic Information Approach. A problem modelling the information arises when different experts have different uncertainty degrees on the same phenomenon or when an expert has to evaluate different concepts. Then, several linguistic term sets with a different granularity of uncertainty are necessary. In such situations, we need tools to manage multi-granular linguistic information [11]. In [11] a multi-granular 2-tuple fuzzy linguistic modelling based on the concept of linguistic hierarchy is proposed. A *Linguistic Hierarchy LH*, is a set of levels $l(t, n(t))$, where each level t is a linguistic term set with different granularity $n(t)$. In [11] a family of transformation functions between labels from different levels was introduced. To establish the computational model we select a level that we use to make the information uniform and thereby we can use the defined operator in the 2-tuple model. This result guarantees that the transformations between levels of a linguistic hierarchy are carried out without any loss of information.

Incomplete Fuzzy Preference Relations. A fuzzy preference relation P on a set of alternatives $X = \{x_1, .., x_n\}$ is a fuzzy set on the product set $X \times X$, i.e., it is characterized by a membership function $\mu_P : X \times X \longrightarrow [0, 1]$. When cardinality of X is small, the preference relation may be conveniently represented by the $n \times n$ matrix $P = (p_{ij})$, being $p_{ij} = \mu_P(x_i, x_j)$ $(\forall i, j \in \{1, \ldots, n\})$ interpreted as the preference degree of the alternative x_i over x_j, where $p_{ij} = 1/2$ indicates indifference between x_i and x_j, $p_{ij} = 1$ indicates that x_i is absolutely preferred to x_j, and $p_{ij} > 1/2$ indicates that x_i is preferred to x_j.

As the proposals analyzed integrate the multi-granular fuzzy linguistic modeling based on 2-tuples, a linguistic preference relation must be defined.

Let $X = \{x_1, .., x_n\}$ a set of alternatives and S a linguistic term set. A linguistic preference relation $P = p_{ij}(\forall i, j \in \{1, \ldots, n\})$ on X is:

$$\mu_P : X \times X \longrightarrow S \times [0.5, 0.5) \tag{1}$$

where $p_{ij} = \mu_P(x_i, x_j)$ is a 2-tuple which denotes the preference degree of alternative x_i regarding to x_j.

However, in many problems the experts are often not able to provide all the preference values that are required. In order to model these situations, incomplete fuzzy preference relations are used [1,2,15]. A function $f\colon X \longrightarrow Y$ is *partial* when not every element in the set X necessarily maps onto an element in the set Y. When every element from the set X maps onto one element of the set Y, then we have a *total* function. A two-tuple fuzzy linguistic preference relation P on a set of alternatives X with a partial membership function is an *incomplete two-tuple fuzzy linguistic preference relation*.

3 Proposals to Improve the Dissemination of Information in Digital Libraries

3.1 A Multi-disciplinar Recommender System to Advice Research Resources in University Digital Libraries

The first proposal was presented in [22]. This paper presents a fuzzy linguistic recommender system that recommends two types of resources: specialized resources of the user research area, and complementary resources in order to include resources from related areas that could lead to interesting collaboration possibilities with other researchers and form multi-disciplinar groups. The *vector model* [13] is used to represent both the resource scope and the topics of interest that characterize the users profiles. A classification composed by 25 disciplines is used, and in each position of the resource or user vector, a linguistic 2-tuple value represents the importance degree of the discipline regarding to the resource or the user topics of interest is stored. The recommendation approach is based in a matching process among the terms used in the users and resources representations [13]. The vector model is used to represent both the resource scope and the users topics of interest. Since the system works with linguistic values, a linguistic similarity measure $\sigma_l(V_1, V_2)$ is defined, based on cosine measure but defined in a linguistic context. The recommendation strategy has two phases:

– To generate recommendations for a resource i, $\sigma_l(V_i, V_j)$ is computed among the resource scope vector (V_i) against all the stored resources in the system (V_j, $j = 1 \ldots m$ where m is the number of resources). If $\sigma_l(V_i, V_j) \geq \alpha$ (linguistic threshold value to filter out the information), the resource j is chosen. Next, the system searches for the users which were satisfied with these chosen resources. To obtain the relevance of the resource i for a selected user x, the system aggregates (using the arithmetic mean) $\sigma_l(V_i, V_j)$ with the assessments previously provided by x about the similar resources and with the assessments provided by others users. If the calculated relevance degree is greater than a linguistic threshold μ, then, the system sends the resource information and its calculated linguistic relevance degree to the selected users. If not, the system proceeds to estimate if the resource could be interesting as a complementary recommendation.

To obtain the complementary recommendations, the system computes $\sigma_l(V_i, V_x)$ among the resource i and the user x (for all users). Then, it applies a multi-disciplinar function to the value $\sigma_l(V_i, V_x)$. This function must give greatest weights to similarity middle values (near 0.5), because values of total similarity contribute with efficient recommendations but are probably known for the users. Like null values of similarity show a null relationship between areas. In the proposed system a triangular function, $g(x)$ is used. Next, if the obtained multi-disciplinar value is greatest than a previously defined linguistic threshold γ, the system recommends the complementary resource.

- The proccess of generating recommendations for a user x, is similar, but computing $\sigma_l(V_x, V_y)$ between the topics of interest vectors of the new user (V_x) against all users in the system (V_y, $y = 1..n$ where n is the number of users). If $\sigma_l(V_x, V_y) \geq \delta$ (linguistic threshold value), the user y is chosen as near neighbor of x. Next, the system searches for the resources that satisfied these users. To obtain the relevance of a resource i for the user x, the system aggregates $\sigma_l(V_x, V_y)$ with the assessments previously provided about i by the nearest neighbors of x. If the calculated relevance degree is greater than the linguistic threshold μ, then, the system recommends to the new user the resource information and its calculated linguistic relevance degree. If not, the system proceeds to estimate if the resource could be interesting as a complementary recommendation for the user. The system computes $\sigma_l(V_x, V_i)$ among the user x and the resource i (for all resources). Then, it applies the multi-disciplinar function $g(x)$ to the value $\sigma_l(V_x, V_i)$. If the obtained multi-disciplinar value is greatest than the linguistic threshold γ, the system recommends the resource as complementary.

3.2 Dealing with Incomplete Information in a Fuzzy Linguistic Recommender System to Disseminate Information in a University Digital Library

The second proposal is presented in [21]. The problem of the previous proposal is that users must directly specify their user profiles by providing their preferences on all topics of interest and it requires too much effort by the user. The system presented in [21] allows users to provide their preferences by means of incomplete fuzzy linguistic preference relations [1], and this facilitate the determination of user profiles. To reduce that effort and make the process of acquiring user preferences easier, an alternative method to obtain the user preferences on topics of interest is proposed. The system shows to the users only a selection of the most representative resources, and the users stablish their preferences about these resources by means of an incomplete fuzzy preference relation. Furthermore, according to results presented in [2], it is enough that the users provide only a row of the preference relation. Then the method proposed in [2] is used to complete the relations. Once the system completes the fuzzy linguistic preference relation provided by the user, it is possible to obtain a vector representing the user preferences on the topics of interest.

The recommendation strategy is based on a matching process developed between user profiles and resource representations, using a linguistic similarity measure based on cosine measure, $\sigma_l(V_1, V_2)$. To generate the recommendations for a resource i, $\sigma_l(VR_i, VU_j)$ is computed, between the representation vector of the resource (VR_i) and all the user preference vectors, $\{VU_1, \ldots, VU_m\}$, where m is the number of users in the system. If $\sigma_l(VR_i, VU_j) \geq \psi$ (linguistic threshold previously defined), the user j is selected to receive recommendations about resource i. For users who want it, the system also recommends collaboration possibilities. The linguistic compatibility degree is obtained computing $\sigma_l(VU_x, VU_y)$ between each two users x and y who want to collaborate.

3.3 An Improved Recommender System to Avoid the Persistent Information Overload in a University Digital Library

In third place, we analyze the proposal presented in [20]. Despite that the use of the two previous techniques to avoid the information overload problem was successful, the number of electronic resources daily generated keeps growing continuously and the problem rises again. Therefore, a persistent problem of information overload was found. The idea is to use a memory to remember selected items but not recommended previously, and in such a way, the system could incorporate them in future recommendations to complete the set of recommendations. For example, if there are a few items to be recommended or if the user wishes outputs obtained by combination of items selected in different recommendation rounds. Users are asked to express restrictions on the quantity of items to receive in each recommendation round and about the novelty of such items.

This system works in two phases:

1. To generate the recommendations using the recommendation approach of the previous proposal [21].
2. To apply a second filter or selection process according to the user's restrictions. Taking into account the number of recommendations that the user would like to receive:
   (a) If there are not enough resources to satisfy the amount of recommended resources specified by the user, the system remembers the items previously selected but not recommended and now could be recommended. The system then repeats the recommendation process detailed in phase 1, but now incorporating these remembered resources.
   (b) If the amount of selected resources is enough, the system checks the restrictions talking about the novelty of the resources or if the user is also interested in previous resources but still with validity, which could

be most interesting than a new resource. If the user wants both kinds of resources, the system repeats the recommendation process of the first phase, but now incorporating these remembered resources.

Finally, the system shows to the users the resource information and its calculated linguistic relevance degree, and for the users who want to collaborate, the system sends the resource information, its calculated linguistic relevance degree and the collaboration possibilities characterized by its linguistic compatibility degrees.

3.4 A Quality Based Recommender System to Disseminate Information in a University Digital Library

Finally, we analyze the proposal presented in [28]. Analyzing the previous proposals, different aspects that may limit their performance were found. Really they worked as an information retrieval system based on matching functions which acted among the resources representation and user profiles, and this limited their performance. Furthermore, the number of electronic resources daily generated grows continuously, so the problem appears again and the system performance decreases.

In this proposal the system implements a hybrid recommendation strategy based on a switching hybrid approach [3], which switches between a content-based recommendation approach and a collaborative one to share the user individual experience and social wisdom. With this dual perspective, the cold-start problem is minimized because the system switch from one approach to another, depending on the circumstances.

Besides, now the recommendations generation process is a task with two distinct elements: On one hand, finding resources that are relevant to the users and on the other hand, finding valid resources from the standpoint of the *quality of the items* [5]. The system incorporates a new module which performs a *re-ranking* process which takes into account the estimated relevance of an item along with the item quality. But the problem is how to obtain the resource quality without much interaction from users. So, a new way to evaluate the quality of resources is proposed. Based on the idea of whether one resource is usually preferred than others, indicates that the resource has a certain quality. To do that, the system incorporates the method presented in [21] in which the users are asked to provide their preferences on five research resources, by means of an incomplete fuzzy preference relation. Then, the system completes this preference relation. This method is used to obtain the user profiles, but it is also used to estimate the quality of these resources. It is assumed that resources usually preferred over others have a higher quality. So, you can count the times that each resource has been selected to be shown as well as the times that each resource has been preferred over other. The displayed resources will vary over time, so the system must record each time a resource is selected and each time a resource is preferred to other. The quality of a resource is estimated as the probability that the resource is chosen against another.

Once a research resource is considered relevant for a user, and both the estimated relevance degree and the resource quality score, have been computed, the last step is to aggregate both in a single score. To do this, the system uses a multiplicative aggregation in which the estimated relevance is multiplied by the translated quality score (with the corresponding linguistic transformations). Then, the systems recommends the user these resources along with these final estimated scores to justify the recommendations.

3.5 Comparative Analysis

Now, we include some brief comments about the capacity of ratings predictions. In order to obtain data to compare the Mean Absolute Error (MAE) was computed for the different propsosals, i.e. the average absolute deviation between a predicted rating and the user's true rating. The first two approaches get similar values of yield, but the advantage of the second is the lower participation of users, thus improving the satisfaction of users. The third approach presents a small performance improvement, with greater precision. But it is the fourth approach that performs best. Therefore, the predictions obtained by using the quality of resources are better than the predictions obtained only with the relevance or memory. Specifically an improvement of 4.80% is obtained. That is, the predictions generated with the new system are closer to the users' preferences.

4 Conclusions and Future Work

In a UDL the selective dissemination of information about research resources is a particularly important service. The UDL staff and users need tools to help them in their processes of information discovering because of the large amount of information available on these systems. Recommender systems have been successfully applied in academic environments to assist users in their access to relevant information. For this reason, we found it really interesting and in this paper we have reviewed and analyzed several proposals based on recommender system that help students, teachers and researchers to find information. These proposals can improve the services provided by the UDL to their users. The four different proposals reviewed follow an evolution in the time. All of them are based on the application of recommender system and used the fuzzy linguistic modeling, besides each one improves the performance of its predecessors.

Analyzing these proposals, we could conclude and point out that although some progress has been made, it is fundamental to continue working to solve the information overload problem, even more pressing with the continuous advances in technology and especially social networks. In this sense, and focusing on future research, we believe that a promising direction is to study automatic techniques to establish the representation of resources. Moreover, given the current situation of intensive use of social networks, other idea is to explore new improvements in the recommendation approach, exploring new methodologies for the generation of recommendations, for example, extracting knowledge from the information we share in social networks.

Acknowledgments. This paper has been developed with the financing of Projects TIN2013-40658-P and TIN2016-75850-R.

References

1. Alonso, S., Cabrerizo, F., Chiclana, F., Herrera, F., Herrera-Viedma, E.: Group decision-making with incomplete fuzzy linguistic preference relations. Int. J. Intell. Syst. **24**(2), 201–222 (2009)
2. Alonso, S., Chiclana, F., Herrera, F., Herrera-Viedma, E., Alcalá-Fdez, J., Porcel, C.: A consistency-based procedure to estimating missing pairwise preference values. Int. J. Intell. Syst. **23**, 155–175 (2008)
3. Burke, R.: Hybrid web recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, LNCS, vol. 4321, pp. 377–408 (2007)
4. Burke, R., Felfernig, A., Göker, M.: Recommender systems: an overview. AI Mag. **32**, 13–18 (2011)
5. Cabrerizo, F.J., Morente-Molinera, J.A., Pérez, I.J., López-Gijón, J., Herrera-Viedma, E.: A decision support system to develop a quality management in academic digital libraries. Inf. Sci. **323**, 48–58 (2015)
6. Callan, J., Smeaton, A.: Personalization and recommender systems in digital libraries. In: Joint NSF-EU DELOS Working Group Report (2003)
7. Chao, H.: Assessing the quality of academic libraries on the web: the development and testing of criteria. Libr. Inf. Sci. Res. **24**, 169–194 (2002)
8. Charlotte-Ahrens, S.: Recommender Systems: Relevance in the Consumer Purchasing Process. Epubli, Berlin (2011)
9. Gonçalves, M.A., Fox, E., Watson, L., Kipp, N.: Streams, structures, spaces, scenarios, societies (5s): a formal model for digital libraries. ACM Trans. Inf. Syst. **22**(2), 270–312 (2004)
10. Herrera, F., Martínez, L.: A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans. Fuzzy Syst. **8**(6), 746–752 (2000)
11. Herrera, F., Martínez, L.: A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making. IEEE Trans. Syst. Man Cybern. Part B Cybern. **31**(2), 227–234 (2001)
12. Kobayashi, M., Takeda, K.: Information retrieval on the web. ACM Comput. Surv. **32**(2), 148–173 (2000)
13. Korfhage, R.: Information Storage and Retrieval. Wiley Computer Publishing, New York (1997)
14. Marchionini, G.: Research and development in digital libraries (2000). http://ils.unc.edu/~march/digital_library_R_and_D.html
15. Martínez, L., Pérez, L., Barranco, M., Espinilla, M.: Improving the effectiveness of knowledge based recommender systems using incomplete linguistic preference relations. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **16**(2), 33–56 (2008)
16. Mata, F., Martínez, L., Herrera-Viedma, E.: An adaptive consensus support model for group decision making problems in a multi-granular fuzzy linguistic context. IEEE Trans. Fuzzy Syst. **17**(2), 279–290 (2009)
17. Meghabghab, G., Kandel, A.: Search Engines, Link Analysis, and User's Web Behavior. Springer, Heidelberg (2008)
18. Montoya, R.: Boundary objects/boundary staff: Supporting digital scholarship in academic libraries. J. Acad. Librarianship (2017, in press)

19. Morente-Molinera, J., Pérez, I., Ureña, R., Herrera-Viedma, E.: On multi-granular fuzzy linguistic modelling in group decision making problems: a systematic review and future trends. Knowl. Based Syst. **74**, 49–60 (2015)
20. Porcel, C., Morales-del Castillo, J., Cobo, M., Ruíz, A., Herrera-Viedma, E.: An improved recommender system to avoid the persistent information overload in a university digital library. Control Cybern.-ICAE J. Spec. Issue **39**(4), 899–924 (2010)
21. Porcel, C., Herrera-Viedma, E.: Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. Knowl.-Based Syst. **23**, 32–39 (2010)
22. Porcel, C., Moreno, J., Herrera-Viedma, E.: A multi-disciplinar recommender system to advice research resources in university digital libraries. Expert Syst. Appl. **36**(10), 12520–12528 (2009)
23. Ross, L., Sennyey, P.: The library is dead, long live the library! The practice of academic librarianship and the digital revolution. J. Acad. Librarianship **34**(2), 145–152 (2008)
24. Serrano-Guerrero, J., Herrera-Viedma, E., Olivas, J., Cerezo, A., Romero, F.: A google wave-based fuzzy recommender system to disseminate information in university digital libraries 2.0. Inf. Sci. **181**, 1503–1516 (2011)
25. Shannon, A., Orozova, D., Sotirova, E., Atanassov, K., Melo-Pinto, P., Kim, T.: Generalized net model of a university electronic library, using intuitionistic fuzzy estimations. In: Proceedings of the Eight International Conference on Intuitionistic Fuzzy Sets, vol. 10, pp. 91–96 (2004)
26. Shannon, A., Riečan, B., Sotirova, E., Inovska, G., Atanassov, K., Krawczak, M., Melo-Pinto, P., Kim, T.: A generalized net model of university subjects rating with intuitionistic fuzzy estimations. Notes on Intuitionistic Fuzzy Sets **18**, 61–67 (2012)
27. Tejeda-Lorente, A., Porcel, C., Bernabé-Moreno, J., Herrera-Viedma, E.: Refore: a recommender system for researchers based on bibliometrics. Appl. Soft Comput. **30**, 778–791 (2015)
28. Tejeda-Lorente, A., Porcel, C., Peis, E., Sanz, R., Herrera-Viedma, E.: A quality based recommender system to disseminate information in a university digital library. Inf. Sci. **261**, 52–69 (2014)
29. Zadeh, L.: The concept of a linguistic variable and its applications to approximate reasoning. Part I, Inf. Sci. **8**, 199–249 (1975), Part II, Inf. Sci. **8**, 301–357 (1975), Part III. Inf. Sci. **9**, 43–80 (1975)

Modeling Trends in the Hierarchical Fuzzy System for Multi-criteria Evaluation of Medical Data

Piotr Prokopowicz[1(✉)], Dariusz Mikołajewski[1], Emilia Mikołajewska[2], and Krzysztof Tyburek[1]

[1] Institute of Mechanics and Applied Computer Science, Kazimierz Wielki University, ul. Kopernika 1, 85-074 Bydgoszcz, Poland
{piotrekp,dmikolaj,krzysiekkt}@ukw.edu.pl

[2] Department of Physiotherapy, Ludwik Rydygier Collegium Medicum, Nicolaus Copernicus University, ul. Jagiellonska 13-15, 85-094 Bydgoszcz, Poland
emiliam@cm.umk.pl

Abstract. The paper presents the analysis and application of hierarchical fuzzy system to the problem of evaluation/measurement of the rehabilitation effects in post-stroke patients. Healthy people constitute reference group. Prevalence and impact of the stroke-related disorders on Health-Related Quality of Life (HRQoL) as a recognized and important outcome after stroke is huge. Quick, valid and reliable assessment of HRQoL in people after stroke constitutes a worldwide significant problem for scientists and clinicians - there are many tools, but no one fulfills all requirements or has prevailing advantages. Evaluation model presented here is improved version of earlier attempts and applies the potential of fuzzy systems for linguistic modeling of rules. It provides a great advantage as there are experienced clinicians working on the improvement of the rehabilitation methods but there is no intuitive formal model to measure their effects. The innovative element here is the use of Ordered Fuzzy Number model. It is a good tool for modeling the trends in information used to create the fuzzy rules of small fuzzy systems which together form a hierarchical fuzzy evaluation model.

Keywords: Ordered Fuzzy Number · Kosinski's Fuzzy Number · Fuzzy system · Hierarchical fuzzy system · Linguistic modeling · Stroke rehabilitation · Health-related quality of life

1 Introduction

There are many medical practitioners experienced in post-stroke rehabilitation, but often there is also a lack of formal models for important data processing. In general, we deal here with a problem how to transform rich practical experience into a formal tool, which can be easily used and help in work for the less experienced medics.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_19

In such situation, we need the tool which will transform the linguistic model into the formal algorithm. The fuzzy system Mamdani's type is a good proposition.

Measurement of the patient-centered outcomes such as functional status and health-related quality of life (HRQoL) is very important within current health care, especially in rehabilitation after severe diseases, injuries and cerebrovascular accidents such as stroke [17]. Conceptual and methodological issues of patients quality of life (QoL) measurement are not easy and are still under debate. Many developed tools such as Ferrans and Powers QOL Index-Stroke Version, Niemi QOL scale, SA-SIP30, Sickness Impact Profile, etc. have advantages and disadvantages (or even some unresolved issues), thus selection of the proper tool need for particular caution, and taking into consideration goals, context, and limitations of the particular application.

The HRQoL is a very general concept. There is many elements which are difficult to calculate. Therefore if a precise model is out of reach, we can use the tools for the imprecise information processing - fuzzy systems. Key advantage here is the flexibility, intuitiveness and clarity of rules that are easy to describe linguistically.

In this paper, an algorithm for evaluation of a general quality of life of people after stroke is presented. A hierarchical fuzzy system [12,27,28] is used here as the main evaluating mechanism. Presented tool has work-name Multicriterial Fuzzy Evaluator of Health-Related Quality of Life (abbr. MuFE-HRQoL) and it is an extended variant of the proposition from [24].

The new feature of the present proposition is considering a kind of trend (or tendency) in the model of data processing. It is carried by using in low-level parts of fuzzy system hierarchy new kind of methods based on the special model - the Ordered Fuzzy Number.

That model was started by three authors Witold Kosiński, Piotr Prokopowicz and Dominik Ślęzak in [10,11]. Since the paper [26], the alternative name the *Kosiński's Fuzzy Numbers* (KFN) is used to honor the contribution of late Witold Kosiński in development of the considered model. This name also will be used here in following parts of this paper.

Using KFN concept we can model with one object a situation like: 'gait velocity is high and growing'. There are indications about some features of HRQoL, which can be better formulated using a trend in the data.

This paper also presents another innovative element which is the combination of both ideas together in one application: the classical fuzzy system and concepts based on the KFN model.

The order of this paper is as follows: a short description of medical scores used in the estimation of HRQoL, presentation of main ideas of Kosinski's Fuzzy Numbers and key-method for the processing directed data with them. Next description of the model of evaluation, then results for processing of data gathered in post-stroke patients and the healthy reference group. Finally, an analysis of results and summary will be provided. Thus main aim of this paper is to present a novel approach to HRQoL assessment based on KFN calculations.

2 Clinical Scores to Evaluate

Bobath Scale (to assess hand functions), Barthel Index (to assess activities of daily living - ADL), and normalized values of the gait parameters (normalized gait velocity, normalized cadence, and normalized stride length) were applied to asses functional status and independence of the subjects. Measures above are often used in everyday clinical practice, assessed as valid and reliable. Measurements were performed in every post-stroke patient (i.e. belonging to the study group) twice:

– before the therapy (before the first session of the therapy),
– after the therapy (after the last session of the therapy) - to compare results and assess rehabilitation effects.

Ten sessions of the NDT-Bobath therapy were provided during the course of 2 weeks (10 days of the therapy rehabilitation was performed every day for 5 days a week). Each session lasted 30 min. NDT-Bobath Concept (NDT stands for neurodevelopmental treatment) constitutes the most popular treatment approach applied in stroke rehabilitation, despite the superiority of the one particular approach has not been established yet due to methodological limitations and scarce compartmental studies. Current evidence syntheses are weak, pose too many methodological shortcomings, and lack of the high-quality trials. Lack of detailed clinical guidelines in the area of post-stroke physiotherapy cause that preferences and experience of the therapist constitute framework of the most effective treatment (so-called mixed/eclectic approach) [6,7,13,14]. Patients were treated according to the rules of the method by experienced (>15 years of experience) therapists of NDT-Bobath method for adults with international certificates: by IBITA (basic and advanced course) and EBTA (basic and advanced course). Measurements were performed in every member of the reference group (healthy people) once. The study was accepted by the appropriate Bioethical Committee. The subjects gave written informed consent before entering the study, in accordance with the recommendations of the Bioethical Committee, acting on the rules of Good Clinical Practice and the Helsinki Declaration.

3 Kosinski's Fuzzy Number Model

The concept of the model of Kosinski's Fuzzy Number was defined as a result of searching for simple and flexible algorithms performing calculations on fuzzy numbers [9–11]. It is alternative for Zadeh's fuzzy numbers approach for managing a vagueness in the quantitative data. There is the latest monograph [22] which gathers in one place most of the present knowledge about this model. Here are introduced only basic concepts to understand an idea of linguistic modeling of direction in presented in this paper evaluation process.

Each convex membership function of fuzzy number can be split into two parts: first is non-decreasing and second is non-increasing. It is a base of widely known classical fuzzy numbers model called the *L-R fuzzy numbers* [3]. The

KFNs are also based on the idea of membership function splitting. However, the order of parts independent of the values from the function domain is introduced. Figure 1 shows the general idea and basic labels used in the further part of this paper. We denote direction by the arrow. It points an order from the *up part* to the *down part*. For purposes of this paper the *up part* and *down part* are linear functions (we use the trapezoidal shapes), so we use a simplified representation of the any KFN by four (see Fig. 1):

$$A = (s, 1^-, 1^+, e). \tag{1}$$

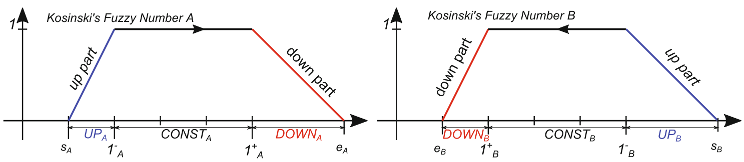


Fig. 1. Two KFNs with the same shape, but different direction/orientation.

4 Processing the Direction

The Kosinski's Fuzzy Number model introduces new feature - the direction. Its useful interpretation was already proposed (see [9,19]. We got a possibility to describe the situation with the trend/tendency in data i.e. *'a temperature is about* 15 °C *and it is rising'*.

It should be noted that some similar idea was proposed for the classical fuzzy model - the gradual fuzzy system (see [4]). It was also extended and analyzed in an interesting approach to trend modeling in [5], where the trend is understood as a gradual dependence between attributes.

However, the gradual fuzzy rules have a form *'The more X is F, the more Y is G'*, but for the KFN model more appropriate is *IF X is in F which is growing/lowering, THEN Y is in G which is growing/lowering*. Here, the KFNs in a natural way have advance - the trend is represented by their direction. Various methods presented in the publications [18,20,22,23,26] are designed to process data with the trend.

4.1 Inference Mechanism Based on the KFN Model

The basis for the processing of fuzzy rules are the operators of inference (i.e. see [1,16]). Generally they are based on logical implications. However, there are also popular solutions like the MIN or PROD, which formally are not the implications, but their practical usefulness is proved.

The method used in the evaluation model in the further part of this paper is named the **'Directed Inference by the Multiplication with a Shift'** and was presented in the [20,21,26]. We emphasize here only its directed character, which let to model the trend in the linguistic rules. We consider the rule as follows:

$$\textbf{IF } X \textit{ is } A \textbf{ THEN } Y \textit{ is } B \tag{2}$$

where A, B - are fuzzy values modeling the rule, X, Y - input and output variable.

Figures 2 and 3 explain the main mechanism of the directed inference. On the Fig. 2(a) we see that the level of activation of the rule is 0.66 and it is on the *up part* side of KFN from premise part of the rule.

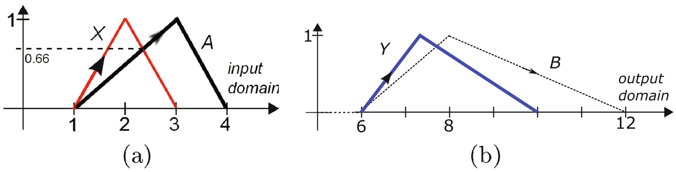


Fig. 2. (a) KFNs X - input, A - model of rule premise (b) Y - result, B - model of conclusion

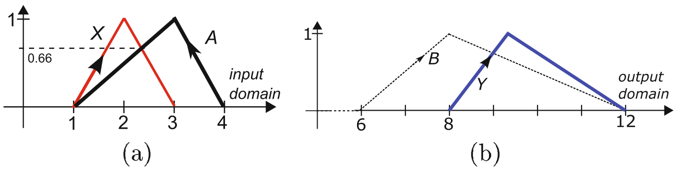


Fig. 3. (a) KFN A with opposite direction while X is the same like in Fig. 2(a), (b) New result of inference operation.

As we see on Fig. 2(b) the answer to the rule is in bounds of the KFN from the consequent part, but it is shifted in its *up part*. Figure 3(a) presents the same shape with only the direction changed. Thus activation is the same (0.66), but this time it is on the *down part* part of premise part KFN. We see (Fig. 3(b)) that the result is also in the bounds of KFN from the consequent part of rule but now it is shifted in the down part side.

5 Hierarchical Fuzzy Evaluator of HRQoL

Fuzzy system used in the MuFE-HRQoL works in two steps. First one uses a group of small fuzzy systems to evaluate separately every singular feature describing HRQoL. Next step uses outputs of the first one as inputs where the

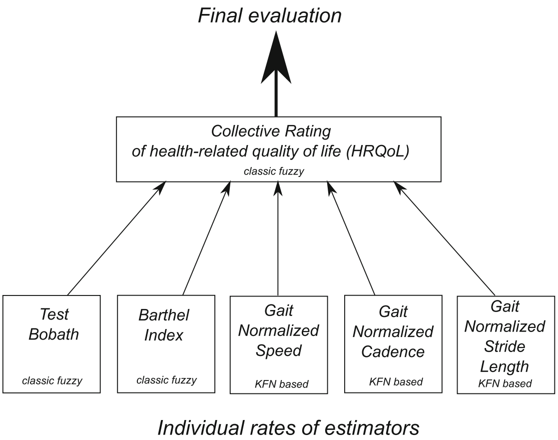


Fig. 4. Hierarchy of fuzzy model of evaluation.

final result - the general quality of HRQoL is calculated. Ideas presented in this paper are strictly connected with [15,25]. The model presented in further sections is modified variant of that proposed in the [24]. Main changes are: using the KFNs, and trapezoidal shapes for modeling the reference group fuzzy values.

The proposed structure of evaluation is a kind of hierarchical fuzzy systems [12,27,28]. Generally, the hierarchical organization of fuzzy systems is used to decrease the total number of rules. However additionally, in our proposition the hierarchy let us to separate the context of the medical properties what makes easier to formulate the model of evaluation linguistically.

Technically MuFE-HRQoL is a hierarchical fuzzy system with five inputs and one output value (see Fig. 4). Input values are the medical scores for Bobath Scale, Barthel Index and three descriptors of gait: velocity, cadence and stride length. The output - general quality - range is $[0;1]$ interval which can be easily transformed to any other. The fuzzification of all inputs is singleton type.

The calculations of evaluation were done with the help of a dedicated tool (implemented by first author Piotr Prokopowicz). However, all KFN calculations can also be done with the use of spreadsheet tools like Microsoft Excel or LibreOffice Calc.

5.1 Evaluator - Low-Level Fuzzy Systems

First two systems evaluate Bobath Scale and Barthel Index. These are strictly defined tests and their quality results represent a typical monotonic tendency, ‘low input, so low output’, where maximum means a normal/healthy condition, and minimum very bad condition. Thus it is modeled by classic fuzzy systems with MIN as inference operator, MAX as conclusions aggregation and COG (center of gravity) defuzzification.

Technically input linguistic variables are divided into two triangular fuzzy sets each. They represent ‘bad/low’ and ‘good/high’ opinion on values - see

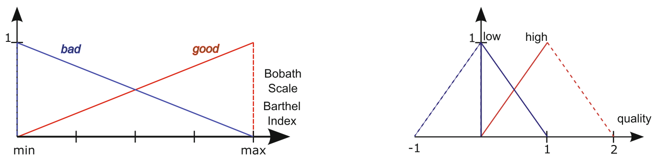


Fig. 5. Left - two input linguistic variables 'Bobath Scale' and 'Barthel Index', right - output linguistic variables

Fig. 5 left side. The assumption is that the range of output values is interval $[0; 1]$, however we observe on the right side graph that we have symmetrical triangular fuzzy sets. Thanks to that using defuzzification COG we really gain range from exactly 0 to exactly 1 without additional normalization (compare with [24]). The rules for the Bobath Scale and Barthel Index are simply illustration of the idea 'high input, then high output' and 'low input, then low output'.

Describing the features of gait is more complex. Too low and too high vales are not 'wanted'. So, each gait parameter is represented by three fuzzy values - two for 'bad' and one for 'good' quality. Additionally, we want to express a trend for 'good' value. Although, too high values are not wanted, in the range of 'good' we prefer the higher values rather than the lower. The same is with gait velocity, cadence and stride length.

Technically each of gait parameters is represented by a separate linguistic variable. Unlike in the [24], here we using KFNs to model these values.

The values representing 'ideal' gait are based on the data given for the reference group - people without stroke. They have a trapezoidal shape and are determined on the all available data. For the inference, operation presented earlier in Sect. 4.1 was used and for defuzzification - the Mean of Maxima. In the papers [2,8] some special KFN defuzzification operators were presented.

All three gait linguistic variables are proposed in the same pattern. Lets look at set *Good Gait Velocity* - ($good^{GV}$). We base on three characteristic values: $x_{min}/x_{max}/x_{mean}$ – the minimum/maximum/mean value of the Gait Velocity parameters for data about healthy (non post-stroke) people. Next we count two deltas: $\Delta_- = x_{mean} - x_{min}$, $\Delta_+ = x_{max} - x_{mean}$. We using the arithmetical mean. Next we describe the KFNs by fourths (see formula 1):

$$good^{GV} = (G_S, N^S, N^E, G_E) \tag{3}$$

where: $G_S = x_{mean} - 1.5 \cdot \Delta_-$, $N^S = x_{mean} - 0.5 \cdot \Delta_-$, $N^E = x_{mean} + 0.5 \cdot \Delta_+$, $G_E = x_{mean} + 1.5 \cdot \Delta_+$.

The KFNs representing 'bad' quality of gait:

$$\begin{aligned} bad^{GV}_{too-low} &= (N^S, G_S, x_0, x_0), \\ bad^{GV}_{too-high} &= (N^E, G_E, 2 \cdot x_{mean}, 2 \cdot x_{mean}), \end{aligned} \tag{4}$$

where $x_0 = MIN(0, x_{mean} - 2 \cdot \Delta_-)$.

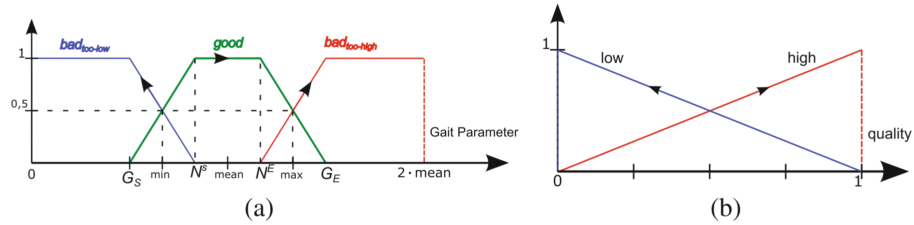


Fig. 6. (a) A pattern of the KFNs describing gait parameters, (b) KFNs for the consequent parts of the rules

Figure 6 shows the general idea such assumptions. It is worth to notice the directions of the KFNs, because it will be important at the determination of rules. The range of linguistic variable is defined as interval $[x_0; 2 \cdot x_{mean}]$. It is enough to cover the all available data for healthy and post-stroke people.

The output KFNs are presented on the Fig. 6(b). The result will be a number from the continuous interval $[0, 1]$ where the higher value means better quality. It can be presented as convenient and intuitive percentage scale. However, it should be stressed that the upper bound stands for the ideal gait parameters, but in the real life there are natural individual differences between healthy people. Thus, the model is formulated to point evaluation each of this persons at least like 50% result value.

Rules for the gait features use the trend/tendency describing linguistic expressions:

– if parameter is 'too low and decreasing' the quality is 'bad and getting worse',
– if parameter is 'good and increasing' the quality is 'good and getting better',
– if parameter is 'too high and increasing', the quality is 'bad and getting worse'.

Considering directions of the KFNs (see Fig. 6) with these rules we are intuitively expressing an expected dependency between input and outputs. The third rule i.e. means 'parameter too high, then quality bad', and at the same time express 'parameter is increasing then quality is getting worse'.

5.2 Evaluator - Fuzzy Sets for Output Fuzzy System

The fuzzy system from the final/second level in hierarchy aggregates the partial evaluations. The input variables are simply two triangular values 'low' and 'high' defined on interval $[0; 1]$. However the output 'Quality' is divided into six fuzzy values (see Fig. 7). They represents terms from an 'extremal low' to a 'normal'. It situation like for Bobath Scale and Barthel Index outputs. We define here fuzzy sets 'extremal low' and 'normal' as symmetrical to get for COG defuzzification in real $[0; 1]$ interval. Aggregation of the premise parts is MIN, the implication operator - MIN, the aggregation of fuzzy outputs (accumulation) – MAX.

As for the rules, there are five input variables two fuzzy sets ('low' and 'high') each. Therefore for complete rule base we have $2^5 = 32$ rules here. To shorten

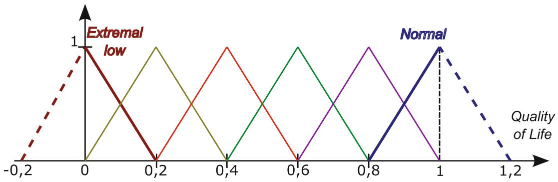


Fig. 7. A pattern of fuzzy sets for the gait describing linguistic variables.

the description we will use the abbreviations of qualities: qV - quality of gait velocity, qC - of gait cadence, qS - of gait stride length, qBT - of Bobath Test, qBI - of Barthel Index.

The output variable (see Fig. 7) is represented by values denoted in the further description as Out_i where $i = 0..5$. The value Out_0 is interpreted as worst and Out_5 is desirable health condition understand as the 'normal quality of life'. Rules follow the pattern:

$$\text{IF } qV \text{ and } qC \text{ and } qS \text{ and } qBT \text{ and } qBI \text{ THEN } Out_s \tag{5}$$

where s is a number of times the term 'high' was used in the premise part of rule. For the example the highest Out_5 is used only in one rule:

$$\text{IF } qV = h \text{ and } qC = h \text{ and } qS = h \text{ and } qBT = h \text{ and } qBI = h \text{ THEN } Out = Out_5 \tag{6}$$

the letter 'h' stands for 'high'. Output is Out_5 because the 'high' was used five times in premise of the rule. Adequately the Out_0 also is used only in one rule ('l' stands for 'low'):

$$\text{IF } qV = l \text{ and } qC = l \text{ and } qS = l \text{ and } qBT = l \text{ and } qBI = l \text{ THEN } Out = Out_0 \tag{7}$$

6 Practical Results

The evaluator was used to measure a condition of 40 patients twice. First - before beginning a cycle of rehabilitation, and second - after it was finished. For comparing/validation purposes we also evaluated a reference group (20 persons), which was the source of a pattern of the gait quality. See Table 1 for collective data about the results. Comparing the averages we see that the rehabilitation in general improves the patient's life quality. As we can see for the reference group, their results are much higher than patients.

With the Table 2 we want to do some validation of the presented evaluator. As the trend expressed in the model is important for this paper, the Table 1 shows evaluation for a special artificially generated set of data. We do not analyze here the inputs for other used HRQoL scores. They are modeled by the classical fuzzy systems and there is no trend expression in the rules. However, it should

Table 1. Collective data about results of evaluation with the MuFE-HRQoL

	Before rehab.	After rehab.	Reference group
Min	0.1178	0.1778	0.7907
Max	0.5782	0.7325	1
Average	0.37352	0.48477	0.90584

Table 2. Trend sensitive results

S.n.	Gait velocity	Gait cadence	Stride length	HRQoL evaluation
1	Average	Average	Average	1
2	Min	Average	Average	0.8999
3	Max	Average	Average	0.942
4	Average	Min	Min	0.7999
5	Average	Max	Max	0.8695
6	Min	Min	Max	0.7406
7	Max	Max	Min	0.7642

be clarified that for those inputs were used ideal values generating full activation of rules pointing ‘high quality’.

For better analysis the Table 2 it is good to look for construction of the gait input linguistic values presented on the Fig. 6(a). The KFN representing ‘good-quality’ is based on the numeric interval generated for the reference group. In the mechanism of evaluation, we model the idea that non post-stroke people’s gait quality is at least 0.5. Thus the borders of that interval - min and max values - basically gives 0.5 evaluation as partial evaluation of whole MuFEG-HRQoL (if it would be a classic fuzzy system). However, modeling the trend, we are changing it and the min value still gives 0.5 result but the max value should give more. In fact, presented case max value gives the partial result equal to 0.75.

As for the average value on input, it represents the ideal expected situation so it always gives the quality 1.

The evaluation results from the Table 2, confirms modeled trend. The first line produces quality 1, because the inputs are averages ideal values for gait parameters. If we change only one input on min we get 0.8999 result, however if this input will be max value it will generate 0.942 quality - table rows 2, 3. The difference between two system with minimal values an two with maximal is even greater - see rows 4, 5. Table rows 6, 7 shows the result where there is no ideal value on input, but still one more maximal input than minimal generates slightly higher evaluation result.

That validation results also shows the expected general pattern. The more ideal are the inputs, the evaluations are higher.

7 Summary and Future Actions

In this paper, we presented the tool MuFE-HRQoL for evaluation of the quality of life for the people after stroke. Thanks to the use of the KFN model we get the possibility of managing vagueness with the trend which is described linguistically. The results for real life data show the potential for application in the practical problems.

The evaluation tool has also a potential to compare the different rehabilitation methods. Therefore, it could also be a tool for evaluation of them. Application of KFNs in the modeling gives the additional potential for involving information about trend/tendency of data used to build fuzzy evaluation system.

In addition, this publication shows that a hierarchical fuzzy system is the right construction to successfully combine classic fuzzy systems with KFN-based solutions. The limitation of the proposed approach it a fact, that it is still in the beginning of its development. Such complex computational approach need for further studies, including careful validation procedures in clinical setting. But our current results strengthen values of the proposed approach as an useful tool in everyday clinical practice.

As for the future actions a gait is only one aspect of life it could be modeled by the independent level of hierarchy which gives as result one value representing the gait quality. Then, it could be used as one independent value among other features affects the final result.

Finally, an improvement of evaluator would be using more medical scores in evaluation of quality of life. However, it is restricted by availability of the another tests results, which should be performed on the same group of patients. We continue our efforts toward web-based version of our tool, which may increase the number of assessed patients and compared scores.

References

1. Buckley, J.J., Eslami, E.: Advances in Soft Computing: An Introduction to Fuzzy Logic and Fuzzy Sets. Physica-Verlag GmbH, Heidelberg (2002)
2. Czerniak, J.M., Dobrosielski, W.T., Iwona, F.: Some cases of comparing fuzzy numbers using defuzzificators on the catalog of ofn shapes. In: Prokopowicz, P., et al. (eds.) Theory and Applications of Ordered Fuzzy Numbers: A Tribute to Professor Witold Kosński. Studies in Fuzziness and Soft Computing, vol. 356. Springer (2017, in print)
3. Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. J. Syst. Sci. **9**(6), 613–626 (1978)
4. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Inf. Sci. **61**(1), 103–122 (1992)
5. Hüllermeier, E.: Association rules for expressing gradual dependencies. In: Elomaa, T., Mannila, H., Toivonen, H. (eds.) Principles of Data Mining and Knowledge Discovery. PKDD 2002, pp. 200–211. Springer, Heidelberg (2002)
6. Klimkiewicz, P., Kubsik, A., Woldańska-Okońska, M.: NDT-Bobath method used in the rehabilitation of patients with a history of ischemic stroke. Wiad. Lek. **65**(2), 102–107 (2012)

7. Kollen, B.J., Lennon, S., Lyons, B., Wheatley-Smith, L., Scheper, M., Buurke, J.H., Halfens, J., Geurts, A.C., Kwakkel, G.: The effectiveness of the Bobath concept in stroke rehabilitation: what is the evidence? Stroke **40**(4), 89–97 (2009)
8. Kosiński, W., Prokopowicz, P., Rosa, A.: Defuzzification functionals of ordered fuzzy numbers. IEEE Trans. Fuzzy Syst. **21**(6), 1163–1169 (2013)
9. Kosiński, W., Prokopowicz, P., Kacprzak, D.: Fuzziness - representation of dynamic changes by ordered fuzzy numbers. In: Seising, R. (ed.) Views on Fuzzy Sets and Systems from Different Perspectives: Philosophy and Logic, Criticisms and Applications, pp. 485–508. Springer, Heidelberg (2009)
10. Kosiński, W., Prokopowicz, P., Ślęzak, D.: Ordered fuzzy numbers. Bull. Polish Acad. Sci. Math. **51**(3), 327–338 (2003)
11. Kosiński, W., Prokopowicz, P., Ślęzak, D.: On algebraic operations on fuzzy numbers. In: Kłopotek, M.A., et al. (eds.) Intelligent Information Processing and Web Mining: Proceedings of the International IIS: IIPWM 2003 Conference Held in Zakopane, Poland, 2–5 June 2003, pp. 353–362. Springer, Heidelberg (2003)
12. Lee, M.L., Chung, H.Y., Yu, F.M.: Modeling of hierarchical fuzzy systems. Fuzzy Sets Syst. **138**(2), 343–361 (2003)
13. Mikołajewska, E.: NDT-Bobath method in normalization of muscle tone in post-stroke patients. Adv. Clin. Exp. Med. **21**(4), 513–517 (2012)
14. Mikołajewska, E.: Associations between results of post-stroke NDT-Bobath rehabilitation in gait parameters, ADL and hand functions. Adv. Clin. Exp. Med. **22**(5), 731–738 (2013)
15. Mikołajewska, E., Prokopowicz, P., Mikołajewski, D.: Computational gait analysis using fuzzy logic for everyday clinical purposes - preliminary findings. Bio-Algorithms Med-Syst. **13**(1), 37–42 (2017)
16. Pedrycz, W., Gomide, F.: An Introduction to Fuzzy Sets: Analysis and Design. With a Foreword by Lotfi A. Zadeh. MIT Press, Cambridge (1998)
17. Pickard, A.S., Johnson, J.A., Feeny, D.H.: Responsiveness of generic health-related quality of life measures in stroke. Qual. Life Res. **14**(1), 207–219 (2005)
18. Prokopowicz, P.: Adaptation of rules in the fuzzy control system using the arithmetic of ordered fuzzy numbers. In: Rutkowski, L., et al. (eds.) Artificial Intelligence and Soft Computing - ICAISC 2008. LNCS, vol. 5097, pp. 306–316. Springer, Heidelberg (2008)
19. Prokopowicz, P.: Analysis of the changes in processes using the Kosinski's fuzzy numbers. Ann. Comput. Sci. Inf. Syst. **8**, 121–128 (2016). IEEE
20. Prokopowicz, P.: The directed inference for the Kosinski's fuzzy number model. In: Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015, pp. 493–503. Springer International (2016)
21. Prokopowicz, P.: Processing the direction with ordered fuzzy numbers. In: Prokopowicz, P., et al. (eds.) Theory and Applications of Ordered Fuzzy Numbers: A Tribute to Professor Witold Kosiński. Studies in Fuzziness and Soft Computing, vol. 356. Springer (2017, in print)
22. Prokopowicz, P., Czerniak, J., Mikołajewski, D., Apiecionek, Ł., Ślęzak, D.: Theory and Applications of Ordered Fuzzy Numbers: A Tribute to Professor Witold Kosiński. Studies in Fuzziness and Soft Computing, vol. 356. Springer (2017) (in print)
23. Prokopowicz, P., Golsefid, S.: Aggregation operator for ordered fuzzy numbers concerning the direction. In: Rutkowski, L., et al. (eds.) Artificial Intelligence and Soft Computing. LNCS, vol. 8467, pp. 267–278. Springer International, Cham (2014)

24. Prokopowicz, P., Mikolajewska, E., Mikolajewski, D., Kotlarz, P.: Fuzzy system as an assessment tool for analysis of the health-related quality of life for the people after stroke. In: Rutkowski, L., et al. (eds.) Artificial Intelligence and Soft Computing. LNAI, vol. 10245, pp. 710–721. Springer, Cham (2017)
25. Prokopowicz, P., Mikolajewska, E., Mikolajewski, D., Kotlarz, P.: Traditional vs ofn-based analysis of temporo-spatial gait parameters. In: Prokopowicz, P., et al. (eds.) Theory and Applications of Ordered Fuzzy Numbers: A Tribute to Professor Witold Kosński. Studies in Fuzziness and Soft Computing, vol. 356. Springer (2017) (in print)
26. Prokopowicz, P., Pedrycz, W.: The directed compatibility between ordered fuzzy numbers - a base tool for a direction sensitive fuzzy information processing. In: Rutkowski, L., et al. (eds.) Artificial Intelligence and Soft Computing. LNCS, vol. 9119, pp. 249–259. Springer International, Cham (2015)
27. Raju, G.V.S., Zhou, J., Kisner, R.A.: Hierarchical fuzzy control. Int. J. Control **54**(5), 1201–1216 (1991)
28. Torra, V.: A review of the construction of hierarchical fuzzy systems. Int. J. Intell. Syst. **17**(5), 531–543 (2002)

Using Fuzzy Sets in a Data-to-Text System for Business Service Intelligence

A. Ramos-Soto[1(✉)], J. Janeiro[1], J.M. Alonso[1], A. Bugarin[1], and D. Berea-Cabaleiro[2]

[1] Centro Singular En Investigación En Tecnoloxías da Información (CiTIUS), Universidade de Santiago de Compostela, Rúa de Jenaro de la Fuente Domínguez, Santiago de Compostela, Spain
{alejandro.ramos,julio.janeiro,josemaria.alonso.moral, alberto.bugarin.diz}@usc.es

[2] Ozona Consulting S.L., Avda San Marcos 31B, 15890 Santiago de Compostela, Spain
diego.berea@ozonaconsulting.com

Abstract. We describe the use of fuzzy sets within MonitorSI-Text. It is a real and operative data-to-text system that generates textual information about the operational state of Information Technology services, monitored by the commercial software platform Obsidian. Until now, Obsidian provided several dashboards that allowed to monitor in real time the state of the service infrastructure of the clients. MonitorSI-Text extends the capabilities of Obsidian with the automatic generation of textual reports, live descriptions and notifications that complement the visualization dashboards with enhanced textual information. Moreover, our system performs an analysis of time series data based on a fuzzy filtering approach as part of its content determination process. MonitorSI-Text has been tested, commercialized and deployed as part of the Obsidian Business Service Intelligence platform, which is currently in use by several customer companies, such as Camper and PwC.

Keywords: Data-to-Text · Fuzzy sets · Time series data · Business service intelligence · Real application

1 Introduction

In recent times, concepts such as Business Intelligence (BI) and Business Analytics (BA) [2] are increasingly gaining attention from the business scene. In general, these terms refer to the ability of a company to manage and analyze effectively its data to improve its decision-making processes. For instance, companies such as Microsoft, Qlik or Tableau offer BI solutions for third-parties and are being extensively used nowadays. In this context, the application of Artificial Intelligence techniques to improve the analysis and interpretation of data in BI is currently a hot topic in this field.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_20

In the realm of Computational Intelligence, fuzzy sets theory has contributed to the emergence of novel proposals in the context of BI. For example, we can highlight the use of fuzzy inference mechanisms and fuzzy clustering in a fraud prevention and risk assessment solution for banking [12], fuzzy cognitive maps in Business Analysis [21], and fuzzy partitions to model an expert's vocabulary [19].

However, one of the disciplines that is currently receiving increased attention from the BI panorama is natural language generation (NLG) [7]. With several decades of research, but a relatively small community when compared to other disciplines such as fuzzy logic, NLG actually encompasses different types of text generation that characterize well-differentiated subfields. For instance, NLG in dialogue systems, narratives and summarization (of other texts) are well-recognized categories. But perhaps the subfield that is currently enjoying the most attention is data-to-text (D2T), defined by Reiter in [18] as the branch of NLG that *deals with the generation of texts from non-linguistic data*.

D2T and NLG have been a source of successfully deployed systems in many application domains for quite a long time (e.g., see the NLG systems timeline in [13]), with early companies like CoGenTex [3] emerging in the 1990s. However, it was not until the late 2000s that commercial NLG started to experience an important growth, with the apparition of bigger companies in recent years, such as NarrativeScience, AutomatedInsights, YSEOP and ARRIA (these are briefly described in [15]). In this context, NLG is a hot topic in BI and, according to the IT consultant company Gartner, *by 2019, natural-language generation will be a standard feature of 90% of modern BI and analytics platforms* [6].

From a fuzzy sets research perspective, the interest in D2T is also important. Paradigms such as computing with words and perceptions, and mainly related research topics as fuzzy linguistic summarization and description of data, have become increasingly aware of the usefulness and complementarity of D2T. This paradigm allows to provide a textual interface to the already human-like linguistic concepts that fuzzy sets theory and fuzzy logic usually model upon numeric data. In fact, much has happened since Kacprzyk and Zadrȯzny identified in [10] NLG as a powerful tool to convey the information computed by fuzzy protoforms [22].

On the one hand, the original research on the theoretical side of fuzzy linguistic summarization, which is also plentiful in illustrative use cases [1], is still continuing in search of new types of protoforms (e.g. [20]). However, on the other hand, complementary research lines have emerged that explore how D2T and NLG in general may benefit from the imprecision modeling capabilities of words and expressions that fuzzy sets can provide [4,5,8,11,15–17].

This paper brings together the three main elements described in this introduction, namely BI, fuzzy sets and D2T. For this, we present an actual D2T system, MonitorSI-Text, that has been developed in collaboration with Ozona Consulting S.L., an internationally established company specialized in consultancy about Information Technology (IT) services under the ISO/IEC 20000 standard [9]. Our system has already been commercialized and deployed as part of a Business Service Intelligence platform, Obsidian [14], which is currently in use by several customer companies, such as Camper and PwC. MonitorSI-Text allows the generation of monthly reports and real time notifications about the state of any kind of IT services within a company's infrastructure. The generation of monthly reports is based on time series data and statistical analysis, but also includes fuzzy set techniques to address the imprecision of some of the terms and expressions that are conveyed in the generated texts.

The rest of this paper is structured in three sections. Section 2 describes in more depth our specific application domain, Business Service Intelligence, and the Obsidian platform. Section 3 depicts MonitorSI-Text from a general perspective, and describes the usual NLG content determination process, where we use of fuzzy sets to model imprecise terms and analyze historic time series for several metrics about the state of IT services. Finally, Sect. 4 summarizes the main contributions described in this paper.

2 Business Service Intelligence: The Obsidian Platform

Nowadays many companies and organizations rely on complex IT infrastructures that are essential for a correct functioning of internal and external services, such as mail services, web platforms, storage, virtualization, firewalling, etc. Although data are usually available to managing these IT services, they are usually found in unstructured formats, distributed into multiple tools and under the responsibility of different units or departments.

In this context, monitoring and managing tools for technical platforms, service desk tools, or business applications like enterprise resource planners (ERPs) are, among others, valid sources of information to calculate in real time the performance of IT services and align them with the business needs. However, there is a lack of tools that allow to integrate and analyze this kind of information. This problem is addressed by Obsidian, a platform that allows to model IT services from data produced by heterogeneous sources. It also calculates indicators at service and business process levels.

Obsidian is a Business Service Intelligence and Analytics platform developed by Ozona Consulting S.L., and designed with the objective of helping align IT services with business objectives. It provides real time metrics and indicators, historical data and predictors through dashboards and reports. This allows Obsidian to provide a holistic view of the IT services. Thus, Obsidian monitors and contributes to enhance the service alignment with the business process (Fig. 1). In other words, Obsidian can be deemed as a BI solution based on IT Service Management.

Obsidian 's main functionalities can be classified into three different categories:

- Service modeling. Obsidian allows to model the architecture of IT services, their internal and external dependencies, and their impact on business processes.

Fig. 1. Snapshots of different visualization dashboards provided by Obsidian.

– Real time computation. A multi-threaded computing engine which allows to calculate in parallel real time metrics about multiple services and their impact on business processes.
– Service monitoring through dashboards and reports. Obsidian provides several out-of-the-box reports and web-based dashboards, as well as tools to design any kind of dashboards and specific reports.

The dashboards and the report generator in Obsidian include D2T capabilities, which are provided by MonitorSI-Text. This D2T system makes use of different data and information about service state metrics and service dependencies to generate textual information, as we will describe in the next section.

3 MonitorSI-Text: A D2T System Using Fuzzy Sets

3.1 General Description

MonitorSI-Text is a D2T system developed to provide the Obsidian platform with enriched information, which is expressed in natural language as a complement to the graphical information already given in dashboards and reports. From a high level view, it has been designed as a set of RESTful web services and an associated client which facilitates the integration with the rest of the platform.

Currently, the system provides D2T services for three different purposes: generation of notifications, generation of monthly reports, and generation of real time descriptions for dashboards. Although each set of services is meant to address different text generation needs within Obsidian (see Fig. 2), all of them have been implemented using a similar D2T approach.

From a D2T perspective, our system follows a simplified version of the D2T architecture proposed by Reiter in [18], consisting of two main processing stages: a content determination stage supported by fuzzy sets and a realization stage based on templates. In this regard, the general architecture of MonitorSI-Text is similar to the already operational D2T system GALiWeather [16].

3.2 Content Determination

Most of the data and information the system needs to generate the texts is already provided by the Obsidian platform, including numeric measurements and statistical values. However, more sophisticated data processing is also present, including the analysis of time series data using fuzzy sets (which is limited to the monthly report generation), as well as the analysis of service dependencies to determine the root cause of specific problems (common for all D2T services)

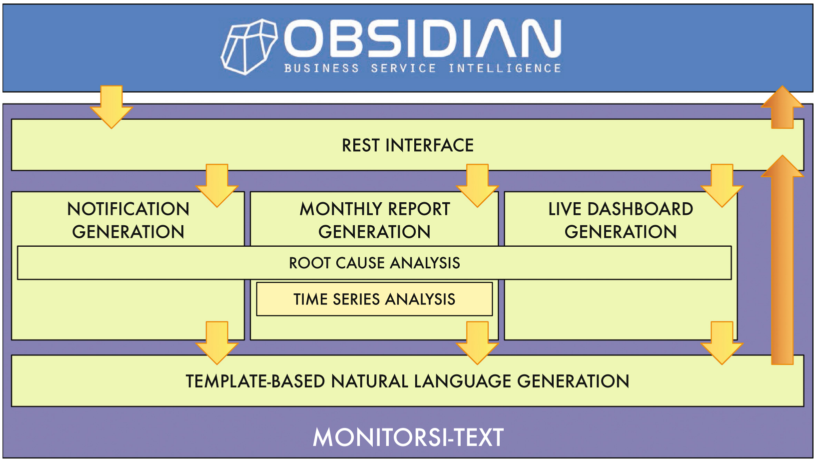


Fig. 2. General architecture of MonitorSI-Text.

Input data. The textual information generated by MonitorSI-Text focuses on four main numeric indicators in the range [0,100] about the state of IT services:

- **Availability**, which measures the aggregated operational state of a given service based on the availability monitor measurements of the service itself and its dependencies (for instance, a web server may depend partly on other services such as a mail server, and depending on the dependency strength, an inoperative dependency might propagate to the main service).
- **Capacity**, which measures performance and available resources at the service level, once more based on the monitor measurements of the service itself and

its dependencies. For instance, a level of 100 in capacity means that the service has enough resources.
- **Service desk**, a synthetic indicator which measures the quality of service support, according to the service desk provider tool. For instance, this indicator can depend on the current number of unresolved issues.
- **Service level**, an indicator which aggregates the values of the three previous metrics.

The numeric values associated with these indicators can be categorized into different labels that represent the possible states of the services according to the service level agreement (SLA) contracted by the company. Particularly, three different states are considered: admissible (green), partially admissible (yellow), and inadmissible (red). Each state corresponds to a closed range within [0,100], whose exact definition depends on the SLA. This means that, for instance, for one company values in the admissible range may correspond to [95,100], but for others this range can be wider or even narrower.

All these indicators are measured and/or calculated regularly on a short temporal basis, providing lengthy sets of time series data. MonitorSI-Text uses these data to generate different kinds of texts. For instance, for real time descriptions our system takes into account only the most recent values, but for the generation of monthly reports it considers aggregated data which may encompass several months backwards for comparison purposes.

Using fuzzy sets in MonitorSI-Text. We will focus now on the particular use that MonitorSI-Text makes of fuzzy sets within its content determination task. Namely, we use fuzzy sets to analyze time series data in the task of generating the more sophisticated monthly reports about the state of IT services. This allows us to fulfill three similar content determination tasks: determining intervals of time that correspond to episodes with values in specific SLA states according to their severity, and evaluating the stability and the variation trend of the indicators at a monthly level.

This analysis is performed regarding all four indicators (availability, capacity, service desk and service level) at different time granularities. Regarding the severity values the analysis is made using aggregated daily data, in the case of the stability we analyze the difference between monthly standard deviations of the time series data for several months, and in the case of the variation trend we perform the same analysis on the average monthly values. This allows the system to extract descriptions such as *"last month there was a partial unavailability from the 15 to the 19"* in the case of the severity analysis, *"the service is more stable compared to the previous month"* in the case of the stability analysis, and *"the service has increased its average capacity value during the previous six months"* in the case of the variation trend.

Our approach is based on a fuzzy filtering of the time series data, which is made for each task according to three different linguistic variables that model the severity of the daily indicator values, the stability over the difference between monthly standard deviation values, and the variation trend over the difference between monthly average values, respectively. Particularly, the severity linguistic

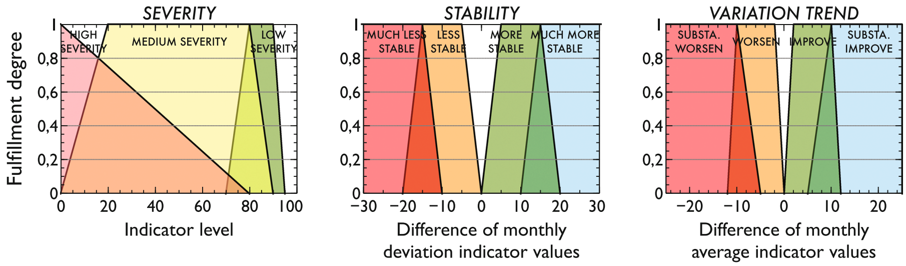


Fig. 3. Linguistic variables used in our approach, which were refined experimentally. From left to right, the severity linguistic variable SEV, the stability linguistic variable STV, and the trend variation linguistic variable VAV.

Algorithm 1. General fuzzy-based algorithm for analyzing time series data in MonitorSI-Text.

Input: TSD, VAV or SEV or STV
Output: EL {A list of relevant episodes}

```
EL ← {}
for all l ∈ VAV|SEV|STV do
    FSD ← μ_l(TSD) {Fuzzify the time series data}
    episode ← {}
    startindex ← 0
    for i=startindex, fv_i ∈ FSD do
        episode ← episode ∪ fv_i {Add fuzzy values until mean(episode) < ε}
        if mean(episode) < ε then
            startindex ← i
            break
        end if
    end for
    EL ← EL ∪ episode
end for
return EL
```

---

variable is dynamically defined by fuzzifying the crisp intervals that define the indicator levels in accordance to the company's SLA. The variation and stability linguistic variables, on the contrary, are statically defined, independently from the SLA.

Regarding the specific algorithms that perform the analysis of the severity, the stability and the variation trend, they all share the same underlying idea. The time series data are filtered for each label in the linguistic variables, providing a linguistic signal which is then analyzed to search for periods with an average membership degree that exceeds an experimental threshold. Consequently, this method provides a more flexible linguistic layer which is tolerant to values that do not fulfill specific labels entirely, but are still relevant enough to be considered during the analysis.

Formally, all algorithms receive as input the following elements:

- The time series data of an indicator, $TSD = \{v_0, ..., v_i, ...v_n\}$, where n depends on the type of task and the current month (only a few values in the variation analysis, but ≈ 30 in the severity analysis), and $v_i = \{i, val\}, i \in \mathbb{N}, val \in [-100, 100]$.
- A linguistic variable which can be either the trend variation (VAV), the severity variable (SEV), the stability variable (STV). Namely, VAV = {substantially worsen, worsen, improve, substantially improve}, SEV = {high severity, medium severity, low severity}, and STV = {much less stable, less stable, more stable, much more stable}.
- A threshold value, ϵ, that determines the minimum average membership degree that the extracted subperiods must fulfill.

The membership functions (modeled as trapezoidal functions in our case, as shown in Fig. 3) associated to each label in the linguistic variables are defined as $\mu : [-100, 100] \rightarrow [0, 1]$. Based on these definitions, Algorithm 1 describes in more detail the general procedure which is shared for the stability, variation and severity content determination tasks.

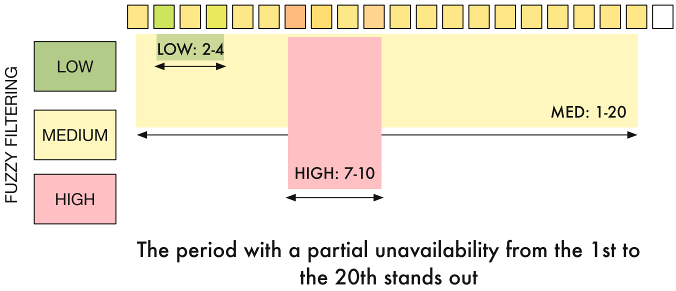


Fig. 4. The algorithm for severity detection selects the longest periods with higher fulfillment degrees.

Algorithm 1 produces EL as output, a list of temporal intervals that are deemed relevant to be included in the generated monthly reports. These intervals are ranked according to their average membership degree, and their associated information (linguistic label, start and end temporal references) is further processed in the natural language generation stage to be properly verbalized.

In the case of the severity analysis, the corresponding version of Algorithm 1 includes several restrictions. For instance, ϵ is set to 0.8, each temporal episode also requires that its first and last values fulfill that $\mu_l(v_i) > 0.8$, and it also adds an overlapping detection feature that controls situations where two episodes identified for different labels may overlap in time (see Fig. 4). Given that there may appear several episodes throughout a month where the indicators drop to specific severity states, we are more restrictive about the conditions that the detected episodes need to meet in order to be deemed as relevant for being conveyed. Figure 4 illustrates how the severity analysis is performed in more detail.

The severity analysis is complemented by the root cause analysis. This allows the D2T system to communicate not only time periods on specific severity levels, but also the actual cause of that event, as in the following description generated by MonitorSI-Text: *"The period with a partial unavailability from the 1st to the 20th stands out, caused by an unavailability in the Directory Service, on which this service depends."*

In the case of the stability and trend variation analysis, the algorithm is simplified to search backwards only a single monthly interval that starts from the last available month, with $\epsilon \geq 0.5$. This allows the system to determine, for a given indicator, if the current month is more or less stable than previous ones, and whether the indicator values have been improving or decreasing. In this case, as the length of the input time series is greatly reduced and we are interested in obtaining a single interval, the search thresholds are relaxed.

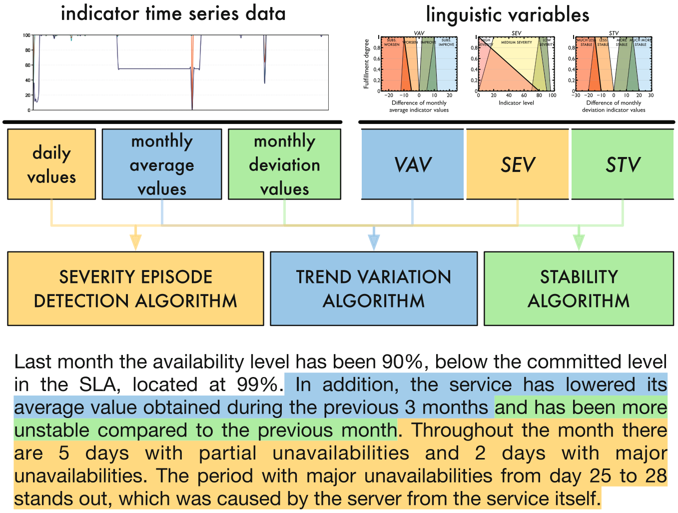


Fig. 5. Illustration of the content determination task using fuzzy sets and corresponding example texts

For instance, the trend variation analysis allows to generate textual descriptions like *"the average value of the service has been decreasing significantly during the previous two months."* In the case of the stability, MonitorSI-Text generates sentences like *"the service has been more unstable compared to the previous month."* The whole content determination based on fuzzy sets is illustrated in Fig. 5, which also includes a text example generated by the service that conveys information obtained by the three content determination tasks.

3.3 Text Realization

We follow a template-based approach to generate the texts requested by the Obsidian platform. Real time descriptions and notifications require simple and mostly static templates. Monthly reports, on the contrary, are more sophisticated, as they may include dynamic information and linguistic resources that depend on the content which is to be communicated. For this, we have developed an additional module that addresses tasks like the aggregation of individual indicator reports. Figure 6 shows several actual examples that MonitorSI-Text is currently able to generate. This includes a live description and a notification, that communicate quantitative and qualitative data about several features, such as the remaining time to breach the committed SLA and changes on the state of a specific indicator with detailed explanations. Figure 6 also provides two different examples of monthly reports. The first one aggregates the information about all indicators as they present similar states, while the second one is a description of one indicator that presents several issues.

MonitorSI-Text currently supports generation for several languages, based on the needs of Obsidian's current customers. These include Spanish, English, Portuguese, and French.

Live description	The current availability level is within the committed levels in the SLA. The average monthly value to date is 99.83%. Currently there are 5.58 hours left to breach the committed monthly availability level.
Notification	The capacity level has improved and stays within the committed levels in the SLA, caused by a rise on the availability of the application SAP AFS 6.0 from the service itself. The capacity fall has lasted for 11.02 minutes. During the rest of the month we could only have 7.44 hours of falls to keep the committed target level.
Monthly report (integrated)	In February all indicators have remained within the committed levels in the SLA. Nor have there been any remarkable falls.
Monthly report (individual)	In the current month the availability level has been 94.69%, below the committed level in the SLA, located at 99%. However, the service has been much more stable compared to the previous month. If we compare it with the other services, this one is much more unstable. Throughout the month there have been 12 days with major unavailabilities. We can highlight the following periods with major unavailabilities: - The day 1 caused by the server CGZTNP097-1 from the service itself. - The day 5 caused by the server CGZTNP097-1 from the service itself. - The day 12 caused by the server CGZTNP097-1 from the service itself.

Fig. 6. Real examples of the different kinds of texts that MonitorSI-Text generates.

4 Conclusions

We have described the use of fuzzy sets to provide flexibility in the analysis of time series data within a real data-to-text system, MonitorSI-Text. This system generates reports about the monthly state of different metrics (capacity, availability, service desk and service level) that measure the state of the services within a company's IT infrastructure through the Obsidian platform.

Although our D2T system is still evolving with the addition of new features, thanks to its previous testing and subsequent commercialization and deployment as part of Obsidian, MonitorSI-Text is already facilitating the work of IT technicians in several companies by automatically analyzing and interpreting the monthly evolution of the services. Likewise, thanks to the automatic generation of real time notifications, MonitorSI-Text allows both technicians and managers to be aware of the state of their IT services without the need of having them constantly monitored.

As future work, we intend to improve the natural language generation capabilities of our system by adapting the generated texts to different job profiles (such as consultants, technicians or managers). We also intend to perform an extrinsic evaluation of the system, to study the actual impact of the texts generated by MonitorSI-Text on final users. From a fuzzy sets perspective, we are interested in exploring new ways of integrating fuzzy set-based techniques for this system that allow to determine the content of new pieces of textual information as new requirements arise.

Acknowledgments. This work has been funded by TIN2014-56633-C3-1-R and TIN2014-56633-C3-3-R projects from the Spanish "Ministerio de Economía y Competitividad" and by the "Consellería de Cultura, Educación e Ordenación Universitaria" (accreditation 2016-2019, ED431G/08) and the European Regional Development Fund (ERDF).

References

1. Boran, F.E., Akay, D., Yager, R.R.: An overview of methods for linguistic summarization with fuzzy sets. Expert Syst. Appl. **61**, 356–377 (2016)
2. Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics: from big data to big impact. MIS Q. **36**(4), 1165–1188 (2012). http://dl.acm.org/citation.cfm?id=2481674.2481683
3. CoGenTex, Inc. http://www.cogentex.com. Accessed 19 March 2017
4. Conde-Clemente, P., Alonso, J.M., Trivino, G.: Towards automatic generation of linguistic advice for saving energy at home. In: Soft Computing, pp. 1–15 (2016)
5. Conde-Clemente, P., Trivino, G., Alonso, J.M.: Generating automatic linguistic descriptions with big data. Inf. Sci. **380**, 12–30 (2017)
6. Gartner: Neural Networks and Modern BI Platforms Will Evolve Data and Analytics. http://www.gartner.com/smarterwithgartner/nueral-networks-and-modern-bi-platforms-will-evolve-data-and-analytics/. Accessed 14 March 2017
7. Gatt, A., Krahmer, E.: Survey of the state of the art in natural language generation: core tasks, applications and evaluation. ArXiv e-prints. https://arxiv.org/abs/1703.09902, March 2017
8. Gatt, A., Portet, F.: Multilingual generation of uncertain temporal expressions from data: a study of a possibilistic formalism and its consistency with human subjective evaluations. Fuzzy Sets Syst. **285**, 73–93 (2016). Special Issue on Linguistic Description of Time Series
9. International Organization for Standardization: ISO/IEC 20000-1:2011. https://www.iso.org/standard/51986.html. Accessed 21 March 2017

10. Kacprzyk, J.: Computing with words is an implementable paradigm: fuzzy queries, linguistic data summaries, and natural-language generation. IEEE Trans. Fuzzy Syst. **18**, 451–472 (2010)
11. Marín, N., Sánchez, D.: On generating linguistic descriptions of time series. Fuzzy Sets Syst. **285**, 6–30 (2016). Special Issue on Linguistic Description of Time Series
12. Meyer, A., Zimmermann, H.J.: Applications of fuzzy technology in business intelligence. Int. J. Comput. Commun. Control V **I**(3), 428–441 (2011)
13. NLG Systems Wiki: Timeline of NLG systems. http://www.nlg-wiki.org/systems/Timeline_of_NLG_systems. Accessed 19 March 2017
14. Obsidian Soft: Obsidian Platform. https://www.obsidiansoft.com/en. Accessed 27 March 2017
15. Ramos-Soto, A., Bugarín, A., Barro, S.: On the role of linguistic descriptions of data in the building of natural language generation systems. Fuzzy Sets Syst. **285**, 31–51 (2016)
16. Ramos-Soto, A., Bugarín, A., Barro, S., Taboada, J.: Linguistic descriptions for automatic generation of textual short-term weather forecasts on real prediction data. IEEE Trans. Fuzzy Syst. **23**(1), 44–57 (2015)
17. Ramos-Soto, A., Tintarev, N., de Oliveira, R., Reiter, E., van Deemter, K.: Natural language generation and fuzzy sets: an exploratory study on geographical referring expression generation. In: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 587–594, July 2016
18. Reiter, E.: An architecture for data-to-text systems. In: Busemann, S. (ed.) Proceedings of the 11th European Workshop on Natural Language Generation, pp. 97–104 (2007). http://www.csd.abdn.ac.uk/~ereiter/papers/enlg07.pdf
19. Smits, G., Pivert, O., Yager, R.R.: A soft computing approach to agile business intelligence. In: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1850–1857, July 2016
20. Wilbik, A., Dijkman, R.M.: Linguistic summaries of process data. In: 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–7, August 2015
21. Xirogiannis, G., Glykas, M.: Fuzzy cognitive maps in business analysis and performance-driven change. IEEE Trans. Eng. Manage. **51**(3), 334–351 (2004)
22. Yager, R.R.: A new approach to the summarization of data. Inf. Sci. **28**(1), 69–86 (1982)

An Approach to Fault Diagnosis Using Fuzzy Clustering Techniques

Adrián Rodríguez Ramos[1], José Manuel Bernal de Lázaro[1], Antônio J. da Silva Neto[2], CarlosCruz Corona[3], José Luís Verdegay[3], and Orestes Llanes-Santiago[1(✉)]

[1] Automation and Computing Department, Universidad Tecnológica de La Habana, CUJAE, Marianao, CP 19390 Havana, Cuba
{adrian.rr,jbernal}@automatica.cujae.edu.cu
[2] Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, Brazil
ajs_neto@uol.com.br
[3] DECSAI, Universidad de Granada, Granada, Spain
carloscruz@decsai.ugr.es, verdegay@ugr.es, orestes@tesla.cujae.edu.cu

Abstract. In this paper a novel approach to design data driven based fault diagnosis systems using fuzzy clustering techniques is presented. In the proposal, the data was first pre-processed using the Noise Clustering algorithm. This permits to eliminate outliers and reduce the confusion as a first part of the classification process. Secondly, the Kernel Fuzzy C-means algorithm was used to achieve greater separability among the classes, and reduce the classification errors. Finally, it can be implemented a step for optimizing the parameters of the NC and KFCM algorithms. The proposed approach was validated using the iris benchmark data sets. The obtained results indicate the feasibility of the proposal.

Keywords: Fault diagnosis · Fuzzy clustering · FCM algorithm · NC algorithm · KFCM algorithm · Metaheuristics

1 Introduction

In current industries, there is a marked necessity to improve the processes efficiency in order to produce with higher quality besides attending the environmental and industrial safety regulations [9]. In industries the faults in equipments can have an unfavorable impact in the availability of the systems, the environment and the safety of operators. For such reason, the faults need to be detected and isolated, being these tasks associated to the fault diagnosis systems [10].

Within the fault diagnosis methods there are those based on the process historical data [7,15]. These approaches do not need a mathematical model, and they do not require much prior knowledge of the process parameters [16]. These characteristics constitute an advantage for complex systems, where relationships

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_21

among variables are nonlinear, not totally known, and it is very difficult to obtain an analytical model that describes efficiently the dynamics of the process.

By performing an analysis of the different techniques developed in the recent years for control and fault diagnosis tasks, it is significative the increment in the use of the fuzzy clustering methods [2,14]. Fuzzy clustering techniques are very important unsupervised tools of data classification [8], that can be used to organize data into groups based on similarities among the individual data. Fuzzy clustering deals with uncertainty and vagueness that can be found in a wide variety of applications. The main focus of all fuzzy clustering techniques is to improve the clustering by avoiding the influence of the noise and outlier data.

The Fuzzy C-Means (FCM) algorithm [3], is one of the most widely used algorithm for clustering due to its robust results for overlapped data. Unlike k-means algorithm, data points in the FCM may belong to more than one cluster center. FCM obtains very good results with noise free data but are highly sensitive to noisy data and outliers [8]. Other similar techniques as, Possibilistic C-Means (PCM) [12] and Possibilistic Fuzzy C-Means (PFCM) [13] interprets clustering as a possibilistic partition and work better in presence of noise in comparison to FCM. However, PCM fails to find optimal clusters in the presence of noise [8] and PFCM does not yield satisfactory results when dataset consists of two clusters which are highly unlike in size and outliers are present. Noise Clustering (NC) [6], Credibility Fuzzy C-Means (CFCM) [4], and Density Oriented Fuzzy C-Means (DOFCM) [11] algorithms were proposed specifically to work efficiently with noisy data.

The clustering output depends upon various parameters such as distribution of data points inside and outside the cluster, shape of the cluster and linear or non-linear separability. The effectiveness of the clustering method highly relies on the choice of the distance metric adopted. FCM uses Euclidean distance as the distance measure, and therefore, it can only be able to detect hyper spherical clusters. Researchers have proposed various other distance measures like Mahalanobis and kernel based distance has been used as measures to detect non-hyper spherical/non-linear clusters in the data and high dimensional feature spaces respectively [17].

In this paper a new fault diagnosis methodology using fuzzy clustering techniques is proposed. The methodology consists of two basic steps. First the preprocessing of data to remove outliers is performed. To achieve this goal the NC algorithm is used. Second, the classification process is developed. For this, the Kernel Fuzzy C-Means (KFCM) algorithm is used to obtain a better separability among classes and therefore the classification results are improved. Finally, an optional step is used to optimize the parameters of the algorithms used in the previous stages.

The main contribution of this paper is the obtaining of a methodology that adequately combines fuzzy clustering algorithms to solve the drawbacks of this type of technique when the data is affected by noise and outliers, and improving the classification by using kernel tools.

The organization of the paper is as follows: in Sect. 2 a description of the new classification methodology using fuzzy clustering techniques is presented. The Sect. 3 presents the experiment design and the data set used to validate the proposed methodology. In Sect. 4 an analysis of the obtained results is presented. Finally, the conclusions are posed.

2 Proposed Classification Methodology Using Fuzzy Clustering

The classification scheme proposed in this work is shown in Fig. 1. In the first step, a set of N observations (data points) $X = [x_1, x_2, ..., x_N]$ are classified into $c+1$ groups or classes using the NC algorithm. The first c classes represent the faults to be diagnosed, as well as the normal operation conditions of the process, and they contain the data points to be used in the next step of the classification methodology. The other remaining class contains the data points identified as outliers by the NC algorithm, and they are not used in the next step.

In the second step, the Kernel Fuzzy C-Means (KFCM) algorithm is applied. This algorithm receives the set of observations classified by the NC algorithm in the c classes as a set of observations to be classified. The KFCM algorithm maps these observations into a higher dimensional space in which the classification process obtains better results. Finally, it can be implemented a final step for optimizing the parameters of the NC and KFCM algorithms.

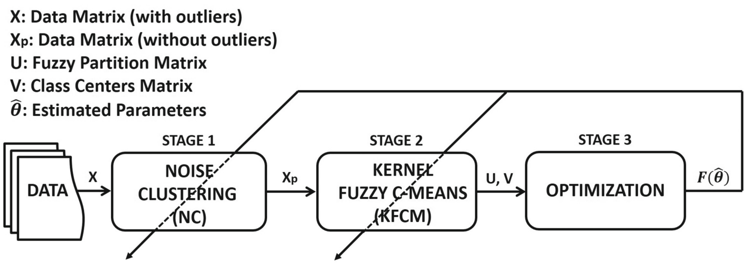


Fig. 1. Classification scheme using fuzzy clustering.

2.1 Fuzzy C-Means (FCM)

Different methods have been proposed for the fuzzy clustering. Among them, the most common are the ones based on distance. One of these methods is the Fuzzy C-Means (FCM) algorithm which uses the optimization criterion (1) to group the data according to the similitude among themselves.

$$J(X; U, \mathbf{v}) = \sum_{i=1}^{c} \sum_{k=1}^{N} (\mu_{ik})^m (d_{ik})^2 \tag{1}$$

The fuzzy clustering allows to obtain the membership degrees matrix $U = [\mu_{ik}]_{c\times N}$ where μ_{ik} represents the degree of fuzzy membership of the sample k to the i^{-th} class, which satisfies the following relationship:

$$\sum_{i=1}^{c} \mu_{ik} = 1, \forall k, i = 1, 2, ..., N \tag{2}$$

In this algorithm, the similitude is evaluated by means of the distance function d_{ik}, represented by the Eq. (3). This function provides a measure of the distance between the data and the classes centers $v = v_1, v_2, ..., v_c$, being $A \in \Re^{n\times n}$ the norm induction matrix, where n is the quantity of measured variables.

$$d_{ik}^2 = (\mathbf{x}_k - \mathbf{v}_i)^T \mathbf{A} (\mathbf{x}_k - \mathbf{v}_i) \tag{3}$$

The exponent $m > 1$ in (1), is an important factor that regulates the fuzziness of the resulting partition. The measure of dissimilarity is the square distance between each data point and the clustering center v_i. This distance is weighted by a power of the membership degree $(\mu_{ik})^m$. The value of the cost function J is a measure of the weighted total quadratic error and statistically it can be seen as a measure of the total variance of $\mathbf{x}_k$ regarding $\mathbf{v}_i$.

The conditions for local extreme for the Eqs. (1) and (2) are derived using Lagrangian multipliers [3]:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} (d_{ik,\mathbf{A}}/d_{jk,\mathbf{A}})^{2/(m-1)}} \tag{4}$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^m \mathbf{x}_k}{\sum_{k=1}^{N} (\mu_{ik})^m} \tag{5}$$

In Eq. (5) should be noted that $\mathbf{v}_i$ is the weighted average of the data elements that belong to a cluster, i.e., it is the center of the cluster i. The FCM algorithm is an iterative procedures where N data are grouped in c classes. Initially, the user should establish the number of classes (c). The centers of the c classes are initialized in a random form, and they are modified during the iterative process. In a similar way the membership degrees matrix U is modified until it is stabilized, i.e. $\|U_t - U_{t-1}\| < \epsilon$, where ϵ is a tolerance limit prescribed a priori, and t is an iteration counter.

2.2 Noise Clustering (NC)

The main idea in the noise clustering algorithm is the concept of a "*noise-prototype*". A noise prototype is an universal entity which always will be at the same distance from every point in the dataset.

Let $\mathbf{v}_n$ be the noise prototype, and $\mathbf{x}_k$ be the point in the feature space, $\mathbf{v}_n$, $\mathbf{x}_k \in \Re^p$. Then, the noise prototype is such that the distance d_{nk}, i.e. the distance of point $\mathbf{x}_k$ from $\mathbf{v}_n$, is:

$$d_{nk} = \delta, k = 1, 2, ..., N \tag{6}$$

Although the definition mentioned above does not specify what the distance δ is, it implies that all the points in the dataset are at the same distance from the noise cluster, and thus defines the noise prototype.

The conventional FCM algorithm was re-formulated using this concept. Let there be c *good* clusters in the dataset, and one noise cluster is added. Then, NC reformulates FCM objective function as:

$$J_{NC}\left(X;U,\mathbf{v}\right)=\sum_{i=1}^{c+1}\sum_{k=1}^{N}\left(\mu_{ik}\right)^{m}\left(d_{ik}\right)^{2} \tag{7}$$

where, the distances are defined by,

$$d_{ik}^{2}=\left(\mathbf{x}_{k}-\mathbf{v}_{i}\right)^{T}\mathbf{A}_{i}\left(\mathbf{x}_{k}-\mathbf{v}_{i}\right),\forall k,i=1\,....\,c \tag{8}$$

$$d_{ik}^{2}=\delta^{2},\ \ for\, i=n=c+1 \tag{9}$$

By assuming that the distance δ is specified, after a minimization process μ_{ik} and $\mathbf{v}_i$ given by Eqs. (4) and (5) are obtained again. In Eq. (5) must be observed that: $i=1\,....\,c$. The main difference in this formulation is that the constraint on the membership values is now equivalent to:

$$0\leq\sum_{i=1}^{c}\mu_{ik}\leq 1,k=1,2,...,N \tag{10}$$

The noise distance δ is a critical parameter in this algorithm, and would be different for different problems. The noise distance proposed by [5] is a simplified statistical average over the non-weighted distances of all feature vectors to all prototype vectors.

$$\delta^{2}=\lambda\left[\frac{\sum_{i=1}^{c}\sum_{k=1}^{N}\left(d_{ik}\right)^{2}}{Nc}\right] \tag{11}$$

where λ is the value of the multiplier used to obtain δ from the average of distances. Based on Eq. (11), δ can be calculated at each iteration of the algorithm.

2.3 Kernel Fuzzy C-Means (KFCM)

KFCM represents the kernel version of FCM. This algorithm uses a kernel function for mapping the data points from the input space to a high dimensional space.

KFCM algorithm modifies the objective function of FCM using the mapping $\mathbf{\Phi}$ as:

$$J_{KFCM}=\sum_{i=1}^{c}\sum_{k=1}^{N}\left(\mu_{ik}\right)^{m}\left\|\mathbf{\Phi}(\mathbf{x_k})-\mathbf{\Phi}(\mathbf{v_i})\right\|^{2} \tag{12}$$

where $\left\|\mathbf{\Phi}(\mathbf{x_k})-\mathbf{\Phi}(\mathbf{v_i})\right\|^{2}$ is the square of the distance between $\mathbf{\Phi}(\mathbf{x_k})$ and $\mathbf{\Phi}(\mathbf{v_i})$. The distance in the feature space is calculated through the kernel in the input space as follows:

$$\|\mathbf{\Phi}(\mathbf{x_k}) - \mathbf{\Phi}(\mathbf{v_i})\|^2 = \mathbf{K}(\mathbf{x_k}, \mathbf{x_k}) - \mathbf{2K}(\mathbf{x_k}, \mathbf{v_i}) + \mathbf{K}(\mathbf{v_i}, \mathbf{v_i}) \tag{13}$$

If the Gaussian kernel is used, then $\mathbf{K}(\mathbf{x}, \mathbf{x}) = \mathbf{1}$ and $\|\mathbf{\Phi}(\mathbf{x_k}) - \mathbf{\Phi}(\mathbf{v_i})\|^2 = \mathbf{2}\,(\mathbf{1} - \mathbf{K}(\mathbf{x_k}, \mathbf{v_i}))$. Thus the Eq. (12) can be written as:

$$J_{KFCM} = 2\sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m \,\|1 - \mathbf{K}(\mathbf{x_k}, \mathbf{v_i})\|^2 \tag{14}$$

where $\mathbf{K}(\mathbf{x_k}, \mathbf{v_i}) = e^{-\|x_k - v_i\|^2/\sigma^2}$.

Minimizing the Eq. (14) under the constraint shown in Eq. (2), yields:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{1-\mathbf{K}(\mathbf{x_k},\mathbf{v_i})}{1-\mathbf{K}(\mathbf{x_k},\mathbf{v_j})}\right)^{1/(m-1)}} \tag{15}$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^m \,\mathbf{K}(\mathbf{x_k}, \mathbf{v_i})\mathbf{x_k}}{\sum_{k=1}^{N} (\mu_{ik})^m \,\mathbf{K}(\mathbf{x_k}, \mathbf{v_i})} \tag{16}$$

3 Study Case and Experimental Design

The iris benchmark dataset from UCI Machine Learning Repository [1] is used for the performance validation of the new classification proposal for fault diagnosis. This dataset presents three classes (*setosa, versicolor, virginica*) with 50 observations each one, and each class has four variables: *sepal length, sepal width, petal length, petal width.* To the original dataset, 48 new observations were added and evenly distributed among the classes in order to represent the possible outliers for each class. Figure 2 shows the iris dataset modified. The *setosa* class (in blue circle) will be considered the normal operation state, while the *versicolor*

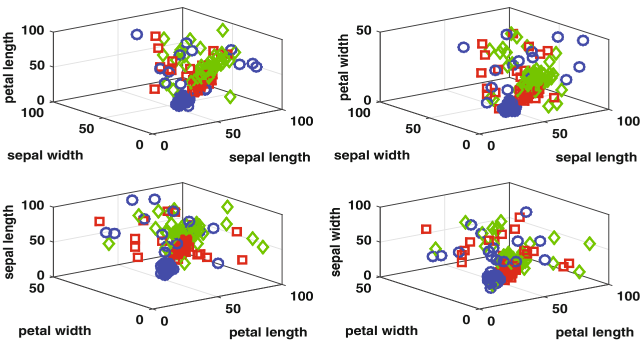


Fig. 2. Modified iris data set

Table 1. Experiments performed

Experiment	Stage 1	Stage 2
1	-	FCM
2	-	KFCM
3	NC	FCM
4	NC	KFCM

(in red square) and *virginica* (in green diamond) classes will represent faults 1 and 2 respectively.

Table 1 presents the four experiments performed. In the first and the second experiments, the step 1 of the proposed classification scheme was not applied. In the first experiment the FCM algorithm was applied in the step 2, and in the second experiment the KFCM algorithm was used. For the experiments 3 and 4 the NC algorithm was selected to be applied in the step 1, and the FCM and KFCM algorithms are applied in the second step, respectively. The values of the parameters used for the applied algorithms are: Number of iterations = 100, $\epsilon = 10^{-5}$, $m = 2$, $\sigma = 10$ (only used for the KFCM algorithm) and $\lambda = 0.01$ (only used for the NC algorithm), these values in the parameters are used in [11].

4 Analysis of the Results

A very important step in the design of the fault diagnosis systems consists on verifying the quality of the performed task. The most used criterion for this analysis is the confusion matrix (CM). The confusion matrix is an indicator that allows to visualize the performance of the classifier in the classification process. Each CM_{rs} element of a confusion matrix for $r \neq s$, indicates the number of times that the classifier confuses a state r with a state s in a set of L experiments. The results obtained from the application of the proposed methodology to fault diagnosis in the modified iris data set are presented next.

4.1 Experiment 1

Figure 3 shows the classification results performed by the FCM algorithm for the modified iris dataset.

Table 2 shows the confusion matrix for experiment 1 where NOC: Normal Operation Condition, F1: Fault 1 and F2: Fault 2. The main diagonal is associated with the number of observations successfully classified. Since the total number of observations per class is known, the accuracy (TA), and the overall error (E) can also be computed. The last row shows the average (AVE) of TA and E. Note that the classification errors have significant values. These results indicate the difficulty of the FCM algorithm to obtain satisfactory results in the classification in the presence of outliers.

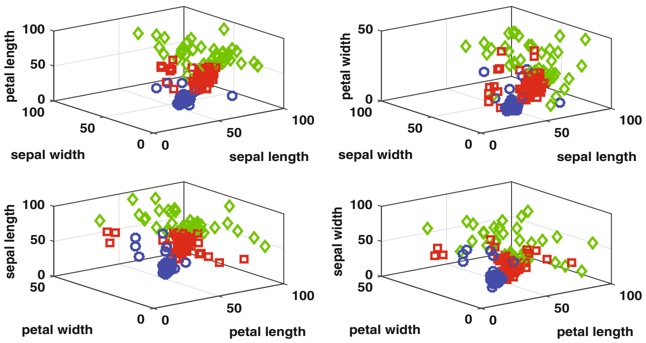


Fig. 3. Classification results for experiment 1 using the FCM algorithm

Table 2. CM for experiment 1. (NOC: 66, F1: 66, F2: 66)

	NOC	F1	F2	TA (%)	E (%)
NOC	**52**	6	8	78.79	21.21
F1	4	**54**	8	81.82	18.18
F2	2	35	**29**	43.94	56.06
AVE				**68.18**	**31.82**

4.2 Experiment 2

Figure 4 shows the classification results of the KFCM algorithm for the modified iris dataset. As shown in Table 3, KFCM algorithm has similar difficulties to FCM in the classification process in presence of outliers.

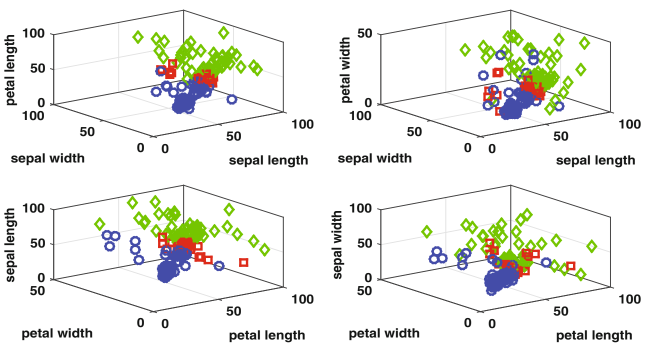


Fig. 4. Classification results for experiment 2 using the KFCM algorithm

Table 3. CM for experiment 2. (NOC: 66, F1: 66, F2: 66)

	NOC	F1	F2	TA (%)	E (%)
NOC	**52**	5	9	78.79	21.21
F1	5	**46**	15	69.70	30.30
F2	2	6	**58**	87.88	12.12
AVE				**78.79**	**21.21**

4.3 Experiment 3

Stage 1

As shown in Fig. 5, the NC algorithm is able to classify the outliers (shown in black color) in the first stage of the classification process. Table 4 shows that the NC algorithm classifies as outliers 48 observations (O class). In addition, NC algorithm obtains good results in the classification of the states NOC, F1 and F2, although these results will not be used in the next step.

Stage 2

Figure 6 shows the classification results of the FCM algorithm after the outlier data were removed in the first step.

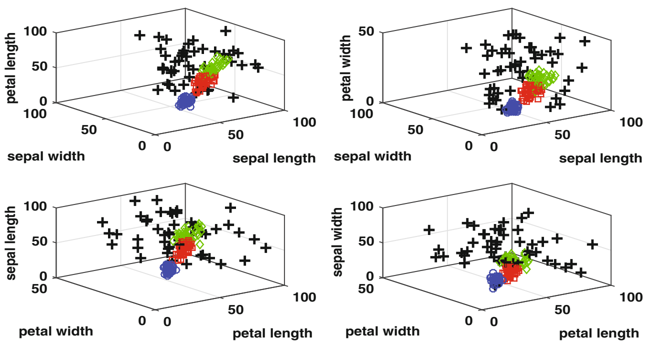


Fig. 5. Stage 1: Results of the outlier classification with NC algorithm.

Table 4. CM for experiment 3: Step 1. (NOC: 50, F1: 50, F2: 50, O: 48)

	NOC	F1	F2	O	TA (%)	E (%)
NOC	**50**	0	0	0	100	0
F1	0	**44**	6	0	88	12
F2	0	15	**35**	0	70	30
O	0	0	0	**48**	100	0
AVE					**89.50**	**10.50**

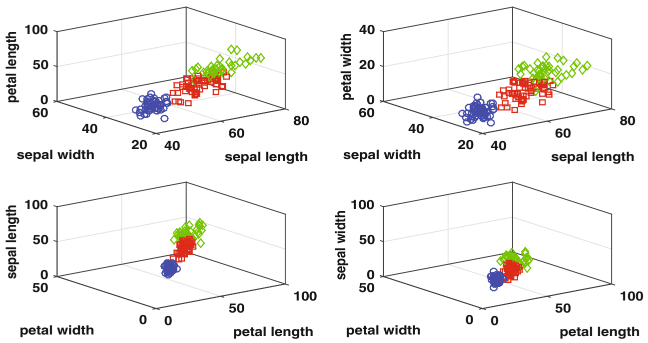


Fig. 6. Stage 2: Classification results with the FCM algorithm.

After applying the algorithm FCM in the stage 2 to classify the observations from the stage 1 the results are the same to those shown in Table 4. This is because the main difference between NC and FCM algorithms is the capacity of the former to classify outliers, therefore, when the data to classify are clean of outliers the results are similar.

4.4 Experiment 4

Stage 1

The classification results of the stage 1 in this experiment are the same of the experiment 3 (Fig. 5 and Table 4).

Stage 2

Figure 7 shows how the KFCM algorithm classifies the observations after the outliers were eliminated in stage 1. Table 5 shows the confusion matrix, and in this case better classification results are achieved compared with the FCM algorithm, as result of the better separability of the classes due to the application

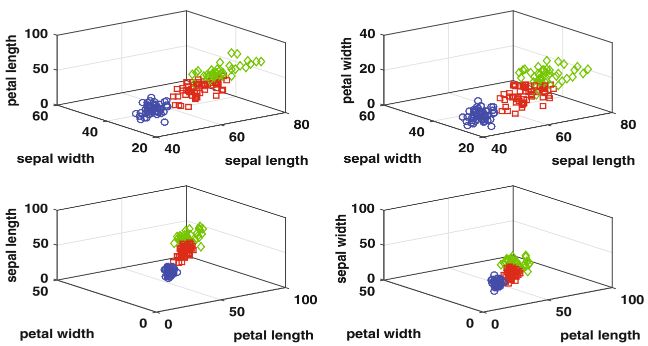


Fig. 7. Stage 2: Classification with KFCM algorithm.

Table 5. CM for experiment 4. (NOC: 50, F1: 50, F2: 50)

	NOC	F1	F2	TA (%)	E (%)
NOC	**50**	0	0	100	0
F1	0	**46**	4	92	8
F2	0	6	**44**	88	12
AVE				**93.33**	**6.67**

of the kernel function. This experiment validates the main propose of this article, i.e. the obtaining of a new classification approach to be applied in fault diagnosis of industrial systems that adequately combines fuzzy clustering algorithms to solve the drawbacks of this type of technique when the data is affected by noise and outliers, and improving the classification process using kernel tools.

5 Conclusions

In the present paper a new classification scheme to fault diagnosis using fuzzy clustering techniques is proposed. In the proposal, the NC algorithm is used in a first stage of preprocessing data to remove the outliers, and the KFCM algorithm is used in a second stage of data classification to make use of the advantages introduced by the kernel function in the separability of the classes, in order to obtain better classification results. Some experiments were performed and their results show the feasibility of the proposal. A possible third stage could be used to optimize the parameters of the used algorithms.

Acknowledgements. The authors acknowledge the financial support provided by the project TIN2014-55024-P from the Spanish Ministry of Economy and Competitiveness and P11-TIC-8001 from the Andalusian Government, both with FEDER funds; FAPERJ, Fundacão Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro; CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico; CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, research supporting agencies from Brazil; and CUJAE, Universidad Tecnológica de La Habana José Antonio Echeverría.

References

1. Asuncion, A., Newman, D.: UCI machine learning repository. University of California, School of Information and Computer Science, Irvine, CA (2007). http://archive.ics.uci.edu/beta
2. Bedoya, C., Uribe, C., Isaza, C.: Unsupervised feature selection based on fuzzy clustering for fault detection of the tennessee eastman process. Advances in Artificial Intelligence. LNAI, vol. 7637, pp. 350–360. Springer (2012)
3. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York (1981)

4. Chintalapudi, K.K., Kam, M.: A noise resistant fuzzy C-means algorithm for clustering. In: IEEE Conference on Fuzzy Systems Proceedings, vol. 2, pp. 1458–1463 (1998)
5. Dave, R.N.: Characterization and detection of noise in clustering. Pattern Recogn. Lett. **12**, 657–664 (1991)
6. Dave, R.N., Krishnapuram, R.: Robust clustering methods: a unified view. IEEE Trans. Fuzzy Syst. **5**, 270–293 (1997)
7. Fan, J., Wang, Y.: Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis. Inf. Sci. **259**, 369–379 (2014)
8. Gosain, A., Dahika, S.: Performance analysis of various fuzzy clustering algorithms: a review. In: 2016 Proceedings of the 7th International Conference on Communication, Computing and Virtualization, vol. 79, pp. 100–111 (2016)
9. Hwang, I., Kim, S., Kim, Y., Seah, C.: A survey of fault detection, isolation, and reconfiguration methods. IEEE Trans. Control Syst. Technol. **18**, 636–656 (2010)
10. Isermann, R.: Fault-Diagnosis Applications: Model-based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors and Fault-tolerant Systems. Springer, Heidelberg (2011)
11. Kaur, P.: A density oriented fuzzy C-means clustering algorithm for recognising original cluster shapes from noisy data. Int. J. Innov. Comput. Appl. **3**, 77–87 (2011)
12. Krishnapuram, R., Keller, J.M.: A possibilistic approach to clustering. IEEE Trans. Fuzzy Syst. **1**, 98–110 (1993)
13. Pal, N.R., Pal, K., Keller, J.M., Bezdek, J.C.: A possibilistic fuzzy C-means clustering algorithm. IEEE Trans. Fuzzy Syst. **13**, 517–530 (2005)
14. Qin, S.J.: Survey on data-driven industrial process monitoring and diagnosis. Annu. Rev. Control **36**(2), 220–234 (2012)
15. Sina, S., Sadough, Z.N., Khorasani, K.: Dynamic neural network-based fault diagnosis of gas turbine engines. Neurocomputing **125**, 153–165 (2014)
16. Wang, J., Hu, H.: Vibration-based fault diagnosis of pump using fuzzy technique. Measurement **39**, 176–185 (2009)
17. Zhang, D.Q., Chen, S.C.: A novel kernelized fuzzy C-means algorithm with application in medical image segmentation. Artif. Intell. Med. **32**, 37–50 (2004)

Universal Generalized Net Model for Description of Metaheuristic Algorithms: Verification with the Bat Algorithm

Olympia Roeva(✉) and Vassia Atanassova

Bioinformatics and Mathematical Modelling Department,
Institute of Biophysics and Biomedical Engineering,
Bulgarian Academy of Sciences, 105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
olympia@biomed.bas.bg, vassia.atanassova@gmail.com

Abstract. In the present paper, the apparatus of generalized nets is used to describe the metaheuristic technique Bat algorithm. Generalized nets are considered an effective and appropriate tool for description of the logics of different optimization techniques. As a result, the developed generalized net model executes the Bat algorithm procedures, conducting basic steps and performing optimal search. The paper elaborates on the already proposed Universal generalized net model for description of the population-based metaheuristic algorithms, which was used so far to model the Cuckoo search, Firefly algorithm and Artificial bee colony optimization, and is used here for modelling of Bat algorithm. It is shown that the Bat algorithm can be described in terms of Universal generalized net model by only varying the characteristic functions of the tokens. Thus, verification of the Universal generalized net model is performed.

Keywords: Generalized nets · Modelling · Metaheuristic · Bat algorithm

1 Introduction

Metaheuristic is a top-level strategy that guides an underlying heuristic solving of a given problem. Following Glover [1], "metaheuristics in their modern forms are based on a variety of interpretations of what constitutes intelligent search".

Evolutionary algorithms like Genetic Algorithms (GA) [2] and Evolution Strategies [3], Ant Colony Optimization (ACO) [4], Particle Swarm Optimization [5], Tabu Search [6], Dynamic Virtual Bats Algorithm [7], Gradient Evolution Algorithm [8], Multi-objective Vortex Search, [9], Simulated Annealing [10], Cuckoo search (CS) [11], Firefly algorithm (FA) [12], Artificial Bee Colony (ABC) optimization [13], Estimation of Distribution Algorithms [14], Scatter Search and Path Relinking [15], Greedy Randomized Adaptive Search Procedure [16,17], Multi-start and Iterated Local Search [18], Guided Local Search and Variable Neighborhood Search [19] are – among others – often listed as

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_22

examples of classical metaheuristics, and they have individual historical backgrounds and follow different paradigms and philosophies [20].

In the present paper, the authors were motivated to use the paradigm of Generalized nets (GNs) [21–24] for description of the mentioned above metaheuristic algorithms. GNs can open opportunities for different online applications; search for optimal conditions; learning on the basis of experimental data; control on the basis of expert systems, etc. Until now, the apparatus of GNs has been used as a tool for a description of parallel processes in several areas – economics, transport, medicine, bioprocess, computer technologies, etc. [25]. The facility of obtaining GN-models demonstrates the flexibility and the efficiency of generalized nets as modelling tools in different fields – biology and biotechnology [26–29], medicine [30,31], optimization [32,33], neural networks [34,35], expert systems [36–38], e-learning [39], intercriteria analysis [40], pattern recognition [41].

So far, GNs have been used as a tool for modelling of various metaheuristics: GAs [42], ACO [44], ABC optimization [46], etc. In [45], six modifications of simple GA have been proposed with different user defined order of implementation of selection, crossover and mutation operators. The monograph [44] contains detailed description of the process of ACO with thorough tests performed for the multiple knapsack problem. In [43], the theory of GNs has been used to describe the FA. The developed model executes the algorithm procedure performing basic steps and realizes an optimal search.

In [46], a universal GN-model that describes the metaheuristics CS, FA and ABC algorithm is proposed. This research contributes to the open problem, defined in [22], namely "to present each of the artificial intelligence areas by GNs". Therefore, a universal GN-model that describes any nature-inspired metaheuristic algorithm is searched. The aim of this study is to verify the so-proposed universal GN-model with the metaheuristic Bat algorithm (BA).

BA is a relative new metaheuristics based on the echolocation behaviour of bats [47]. This algorithm was proposed by Yang [48,49]. The capability of echolocation of microbats is fascinating as these bats can find their prey and discriminate between different types of insects even in complete darkness. Yang [49] formulates the BA by idealizing the echolocation behavior of bats. The resulting algorithm is simple in concept and simultaneously powerful in implementation [50].

The paper is organized as follows. In The Sect. 2 the Bat algorithm is described. In Sect. 3 a GN-model of Bat algorithm is presented. In Sect. 4 the considered Bat algorithm is described by the universal GN-model proposed in [46]. Some conclusions are discussed in Sect. 5.

2 Bat Algorithm

BA, as proposed by [48,49], is based on the following idealized rules:

- All bats use echolocation to sense distance, and they also "know" the difference between food/prey and background barriers.

– Bats fly randomly with velocity v_i at position x_i with a fixed frequency f_{min}, varying wavelength λ and loudness L_0 to search for prey. They can automatically adjust the wavelength (or frequency) of their emitted pulses and adjust the rate of pulse emission r in the range of $[0, 1]$, depending on the proximity of their target.
– Although the loudness can vary in many ways, for example the loudness could vary from a positive L_0 to a minimum constant value L_{min}.

The new solutions $x_i(t)$ and velocities $v_i(t)$ at time step t are given by [48,49]:

$$v_i(t) = v_i(t-1) + (x_i(t) - x_*)f_i, \tag{1}$$

$$x_i(t) = x_i(t-1) + v_i(t), \tag{2}$$

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta, \tag{3}$$

where $\beta \in [0, 1]$ is a random vector drawn from a uniform distribution, x_* is the current global best solution that is located after comparing all the solutions among all the n bats, f_i is used to adjust the velocity change.

For the local search part, once a solution is selected among the current best solutions, a new solution for each bat is generated locally using random walk [48,49]:

$$x_{new} = x_{old} + \eta L_i(t), \tag{4}$$

```
Objective function f(x), x = (x_1, ..., x_d)^T
Initialize the bat population x_i (i = 1, 2, ..., n) and v_i
Define pulse frequency f_i at x_i
Initialize pulse rates r_i and the loudness A_i
   while (t < Max number of iterations)
      Generate new solutions by adjusting frequency, and update
velocities
      and locations/solutions [equations (2) to (4)]
         if (rand > r_i)
            Select a solution among the best solutions
            Generate a local solution around the selected best solution
         end if
      Generate a new solution by flying randomly
         if (rand < A_i & f(x_i) < f(x_*))
            Accept the new solutions
            Increase r_i and reduce A_i
         end if
      Rank the bats and find the current best x_*
   end while
Post-process results and visualization
```

Fig. 1. Pseudo-code of the Bat algorithm.

where $\eta \in [-1,\ 1]$ is a random number, $L_i(t)$ is the average loudness of all the bats at this time step.

The loudness $L_i(t)$ and the rate $r_i(t)$ of pulse emission have to be updated accord-ingly as the iterations proceed:

$$\begin{aligned} L_i(t+1) &= \alpha L_i(t), \\ r_i(t+1) &= r_i(0)[1 - \exp(-\gamma t)], \end{aligned} \tag{5}$$

where α and γ are constants, whose choice requires some experimenting.

For any $0 < \alpha < 1, 0 < \gamma$, we have [48,49]:

$$L_i(t) \to 0, r_i(t) \to r_i(0), \text{ as } t \to \infty. \tag{6}$$

The loudness $L_i(t)$ and emission rates $r_i(t)$ will be updated only if the new solutions are improved, which means that these bats are moving towards the optimal solution.

The BA can be presented as pseudo-code, shown in Fig. 1.

3 Generalized Net Model of the Bat Algorithm

The GN-model, describing the Bat algorithm, is presented in Fig. 2. The token α enters GN through place l_1 with an initial characteristic:

"BA parameters: n, N_{gen}, A, r, Q_{min}, Q_{max}, d, Lb, Ub",

where n is the population size; N_{gen} is the number of generations; A is the loudness; r is the pulse rate; Q_{min} and Q_{max} are frequency minimum and maximum; d is the number of dimensions; Lb and Ub are the lower and upper limit/bounds of the search parameters.

The form of the first transition of the GN-model is

$$Z_1 = \langle \{l_1\}, \{l_2, l_3\}, r_1, \vee(l_1) \rangle,$$

$$r_1 = \begin{array}{c|cc} & l_2 & l_3 \\ \hline l_1 & true & true \end{array}.$$

The token α is splited in two new tokens ε and χ. In place l_2 the token ε obtains the characteristic:

"Q (velocities), v (frequency), Q_{min}, Q_{max}, A, r, Lb, Ub",

where $Q = \text{zeros}(n, 1)$ and $v = \text{zeros}(n, d)$.

In place l_3 the token χ obtains the characteristic:

"Sol (solution initialization)",

where $Sol = Lb + (Ub - Lb)\text{rand}(1, d)$.

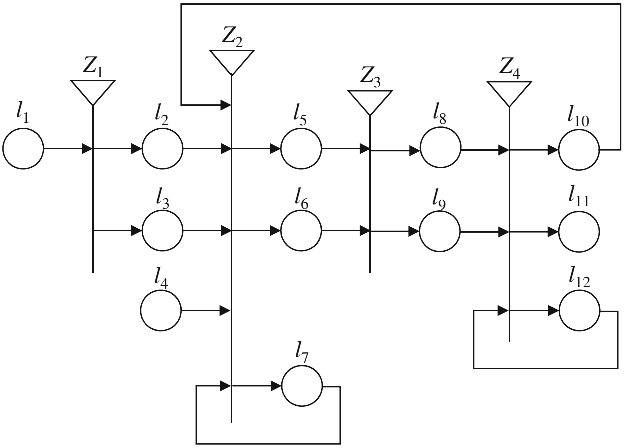


Fig. 2. Generalized net model of Bat algorithm

The token δ enters GN through place l_4 with an initial characteristic:

$$\text{“}Fun \text{ (objective function } f(x)\text{)”}.$$

The form of the second transition of the GN-model is:

$$Z_2 = \langle \{l_2, l_3, l_4, l_7, l_9\}, \{l_5, l_6, l_7\}, r_2, \vee(l_2, l_3, l_4)\rangle,$$

$$r_2 = \begin{array}{c|ccc} & l_5 & l_6 & l_7 \\ \hline l_2 & false & true & false \\ l_3 & true & false & false \\ l_4 & true & false & true \\ l_7 & true & false & false \\ l_9 & false & false & true \end{array}.$$

In place l_5 the tokens χ and δ are combined in a new token γ with the characteristic:

$$\text{“}Fitness \text{ (fitness function), } Fun, \ Sol\text{”},$$

according to $Fitness(i) = Fun(Sol(i, :))$.

In place l_6 the token ε keeps the same characteristic:

$$\text{“}Q, v, Q_{min}, Q_{max}, A, r, Lb, Ub\text{”}.$$

In place l_7 the token γ keeps the same characteristic:

$$\text{“}Fitness, Fun, Sol\text{”}.$$

The form of the third transition of the GN-model is

$$Z_3 = \langle \{l_5, l_6\}, \{l_8, l_9\}, r_3, \vee(l_5, l_6)\rangle,$$

$$r_3 = \begin{array}{c|cc} & l_8 & l_9 \\ \hline l_5 & true & false \\ l_6 & false & true \end{array}.$$

In place l_8 the token γ obtains a characteristic:

$$\text{“}Fitness, Fun, Sol_{best} \text{ (current best solution)”}$$

according to

$$[f_{min}, I] = min(Fitness);$$
$$Sol_{best} = Sol(I, :).$$

The token ε keeps the same characteristic:

$$\text{“}Q, v, Q_{min}, Q_{max}, A, r, Lb, Ub\text{”}$$

in place l_9.

The form of the fourth transition of the GN-model is

$$Z_4 = \langle \{l_8, l_9, l_{12}\}, \{l_{10}, l_{11}, l_{12}\}, r_4, \vee(l_8, l_9)\rangle,$$

$$r_4 = \begin{array}{c|ccc} & l_{10} & l_{11} & l_{12} \\ \hline l_8 & false & false & true \\ l_9 & false & false & true \\ l_{12} & W_{12,10} & W_{12,11} & true \end{array}.$$

where:

- $W_{12,11} =$“End of the BA is reached”;
- $W_{12,10} = \neg W_{12,11}$.

```
for i = 1:n
    Q_new(i) = Q_min + (Q_max − Q_min)rand;
    v_new(i,:) = v(i,:) + (Sol(i,:) − Sol_best)Q(i);
    Sol_new(i,:) = Sol(i,:) + v(i,:);
    Sol_new(i,:) = check_bounds(Sol_new(i,:), Lb, Ub);
    if rand > r
        Sol_new(i,:) = Sol_best + 0.001rand(1,d);
    end if
        Fitness_new = Fun(Sol_new(i,:));
        if (Fitness_new <= Fitness(i))&(rand < A),
            Sol(i,:) = Sol_new(i,:);
            Fitness(i) = Fitness_new;
        end if
        if Fitness_new <= f_min,
            Sol_best = Sol_new(i,:)
            f_min = Fitness_new
        end if
    end if
```

Fig. 3. Pseudo-code of the Bat algorithm.

The token γ obtains the following characteristic:s:

– in place l_{10} – "$Fitness_{new}, Sol_{new}$";
– in place l_{11} – "$Fitness_{final}, Sol_{final}$";
– in place l_{12} – "$Q_{new}, v_{new}, Fitness_{new}, Sol_{new}$",

according to the source-code below (Fig. 3).

In this transition, a new solution (Sol_{new}) is generated by adjusting frequency and velocities and by flying randomly. If termination condition of the BA is reached the final best solution is obtained (place l_{11}), otherwise the next iteration is performed (place l_{10} and second transition).

4 Bat Algorithm Described by the Universal Generalized Net

Here, the universal GN-model, proposed in [46] is used for description of a Bat algorithm. The GN-model is presented in Fig. 4.

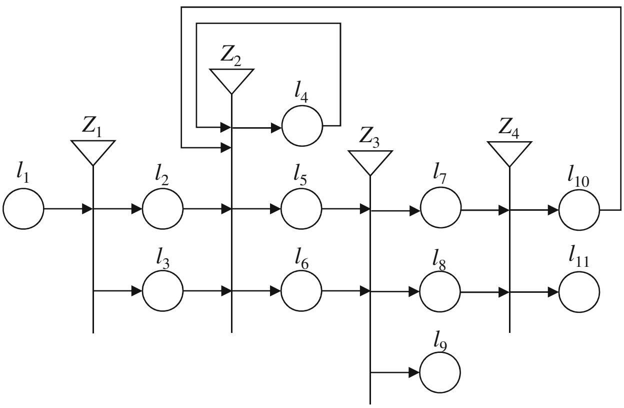


Fig. 4. Universal generalized net model of population-based metaheuristics

Concidering a Bat algorithm the token χ enters GN through place l_1 with an initial characteristic:

"Bat algorithm parameters – $n, N_{gen}, A, r, Q_{min}, Q_{max}$, and problem parameters – objective function, d, Lb, Ub)".

The form of the first transition of the GN-model is

$$Z_1 = \langle \{l_1\}, \{l_2, l_3\}, r_1, \vee(l_1) \rangle,$$

$$r_1 = \begin{array}{c|cc} & l_2 & l_3 \\ \hline l_1 & true & true \end{array}.$$

The token χ is split in two new tokens δ and τ. In place l_2 the token δ obtains the characteristic:

"Initial population, $Sol = Lb + (Ub - Lb)rand(1,d)$; objective function value".

The token τ obtains in place l_3 the characteristic:

"Algorithm parameters and problem parameters".

The form of the second transition of the GN-model is

$$Z_2 = \langle \{l_2, l_3, l_4, l_{10}\}, \{l_4, l_5, l_6\}, r_2, \vee(l_2, l_3, l_4, l_{10})\rangle,$$

$$r_2 = \begin{array}{c|ccc} & l_4 & l_5 & l_6 \\ \hline l_2 & true & false & false \\ l_3 & true & false & true \\ l_4 & true & true & false \\ l_{10} & true & true & false \end{array}.$$

In place l_4, the tokens τ and δ are combined in a new token ν. The token ν obtains the characteristic:

"Evaluated modified population; objective function value".

Here, based on the Bat algorithm, modifications of the initially generated solution are made (by adjusting frequency and velocities and by flying randomly) in order to find a better solution in the search space.

In place l_5, the token ν obtains the characteristic:

"New population (solution)".

The token τ (from place l_3) keeps the same characteristic in place l_6.

The form of the third transition of the GN-model is

$$Z_3 = \langle \{l_5, l_6\}, \{l_7, l_8, l_9\}, r_3, \vee(l_5, l_6)\rangle,$$

$$r_3 = \begin{array}{c|ccc} & l_7 & l_8 & l_9 \\ \hline l_5 & true & false & true \\ l_6 & true & true & true \end{array}.$$

The token ν is split in two new tokens ε and β. In place l_7, token ε obtains the characteristic:

"Ranked solutions with corresponded objective function values".

In place l_9, token β obtains the characteristic:

"Worst solutions".

The token τ (from place l_6) keeps the same characteristic in place l_8.

Here, based on each solution performance and defined objective function, all solutions are ranked.

The form of the fourth transition of the GN-model is:

$$Z_4 = \langle \{l_7, l_8\}, \{l_{10}, l_{11}\}, r_4, \vee(l_7, l_8)\rangle,$$

$$r_4 = \begin{array}{c|cc} & l_{10} & l_{11} \\ \hline l_7 & W_{7,10} & W_{7,11} \\ l_8 & W_{8,10} & W_{8,11} \end{array}.$$

where:

- $W_{7,10} = W_{8,10} =$ "End of the Bat algorithm is not reached";
- $W_{7,11} = W_{8,11} = \neg W_{7,10}$.

The token ε keeps the characteristic:

"Ranked solutions (population) with corresponded objective function values"

in place l_{10}. In the end of algorithm, in place l_{11}, the token ε obtains a new characteristic:

"Best solution and objective function".

It is shown that the proposed universal GN-model (Fig. 4) could be used for description of the considered here Bat algorithm. The Bat algorithm is modelled only through variation of the characteristic functions of the tokens in the GN-model.

5 Conclusion

In this paper, the generalized nets theory is used to describe the Bat algorithm. A GN-model of the basic steps of Bat algorithm is proposed. Moreover, a verification of the universal GN-model for description of the population-based metaheuristic algorithms is presented. Up to now some of the population-based algorithms, namely Cuckoo search, Firefly algorithm and Artificial bee colony optimization, have been described by universal GN-model and shown to do so only by a variation of the characteristic functions of the tokens in their GN-models. As a further step, the universal GN-model is verified here with the metaheuristic Bat algorithm for global optimization. It is shown that the Bat algorithm can be modelled by the universal GN-model based on a slight variation of the characteristic functions. This is achieved because of the specific peculiarities, and especially the universality of the GN theory. Such characteristics allow already developed models to be expanded and/or easily modified resulting in a new model, as well as ensures the compatible and integral representation of various heterogenous paradigms in uniform mathematical modelling terms.

Future work will include development of GN-models of other metaheuristic algorithms as Evolutionary computation, Particle swarm optimization, Glowworm swarm optimization, Bacterial foraging, etc. In addition to the further verification of the universal GN-model for description of the population-based metaheuristic algorithms, research will focus on the possibility for the universal GN-model to be transformed to any of the developed GN-models of different metaheuristic algorithms using the hierarchical operator H_5 [24].

Acknowledgements. This work has been partially supported by Grant DFNP-124, Programme for career development of young scientists of the Bulgarian Academy of Sciences, and by Grant DN02/10 "New Instruments for Knowledge Discovery from Data, and their Modelling", Bulgarian National Science Fund.

References

1. Glover, F.: Tabu search - wellsprings and challenges. Eur. J. Oper. Res. **106**, 221–225 (1998)
2. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley Longman, London (2006)
3. Beyer, H.-G., Schwefel, H.-P.: Evolution strategies: a comprehensive introduction. J. Natural Comput. **1**(1), 3–52 (2002)
4. Dorigo, M., Stützle, T.: Ant Colony Optimization. MIT Press, Cambridge (2004)
5. Umarani, R., Selvi, V.: Particle swarm optimization: evolution, overview and applications. Int. J. Eng. Sci. Technol. **2**(7), 2802–2806 (2010)
6. Yusof, M.K., Stapa, M.A.: Achieving of tabu search algorithm for scheduling technique in grid computing using gridsim simulation tool: multiple jobs on limited resource. Int. J. Grid Distrib. Comput. **3**(4), 19–31 (2010)
7. Topal, A.O., Altun, O.: A novel meta-heuristic algorithm: dynamic virtual bats algorithm. Inf. Sci. **354**, 222–235 (2016)
8. Kuo, R.J., Zulvia, F.E.: The gradient evolution algorithm: a new metaheuristic. Inf. Sci. **316**, 246–265 (2015)
9. Özkış, A., Babalık, A.: A novel metaheuristic for multi-objective optimization problems: the multi-objective vortex search algorithm. Inf. Sci. **402**, 124–148 (2017)
10. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science **220**(4598), 671–680 (1983). New Series
11. Yang, X.-S., Deb, S.: Cuckoo search via Lévy flights. In: Proceedings of World Congress on Nature, Biologically Inspired Computing (NaBIC 2009), USA, pp. 210–214. IEEE Publications (2009)
12. Yang, X.-S.: Firefly algorithm for multimodal optimization. In: Lecture Notes in Computing Sciences, vol. 5792, pp. 169–178 (2009)
13. Karaboga, D.: An idea based on honeybee swarm for numerical optimization, Technical report TR06. Erciyes University, Engineering Faculty, Computer Engineering Department (2005)
14. Ceberio, J., Irurozki, E., Mendiburu, A., Lozano, J.A.: A review on estimation of distribution algorithms in permutation-based combinatorial optimization problems. Prog. Artif. Intell. **1**, 103–117 (2012)
15. Glover, F.: A template for scatter search and path relinking. In: Lecture Notes in Computing Sciences, vol. 1363, pp. 13–54 (1997)
16. Feo, T.A., Resende, M.G.C.: Greedy randomized adaptive search procedures. J. Global Optim. **6**, 109–133 (1995)
17. Colmenar, J.M., Greistorfer, P., Martí, R., Duarte, A.: Advanced greedy randomized adaptive search procedure for the obnoxious p-median problem. Eur. J. Oper. Res. **252**(2), 432–442 (2016)
18. Mladenovic, N., Hansen, P.: Variable neighborhood search. Comput. Oper. Res. **24**, 1097–1100 (1997)
19. Stützle, T.: Local search algorithms for combinatorial problems: analysis, improvements, and new applications, Ph.D. thesis, Darmstadt University of Technology (1998)

20. Brownlee, J.: Clever Algorithms, Nature-Inspired Programming Recipes. LuLu, Raleigh (2011)
21. Atanassov, K.T.: Generalized Nets and Systems Theory. Academic Publishing House "Prof. M. Drinov", Sofia (1997)
22. Atanassov, K.T.: Generalized Nets in Artificial Intelligence. Generalized Nets and Expert Systems, vol. 1. Academic Publishing House "Prof. M. Drinov", Sofia (1998)
23. Atanassov, K.T.: Generalized Nets. World Scientific, Singapore (1991)
24. Atanassov, K.T.: On Generalized Nets Theory. Academic Publishing House "Prof. M. Drinov", Sofia (2007)
25. Choy, E., Krawczak, M., Shannon, A., Szmidt, E. (eds.) A Survey of Generalized Nets, Raffles KvB Monograph No. 10, Australia (2007)
26. Georgieva, V., Sotirova, E.: Generalized net model of biological treatment of wastewater. In: Issues in Intuitionistic Fuzzy Sets and Generalized Nets, vol. 11, pp. 63–72 (2014)
27. Georgieva, V., Angelova, N., Roeva, O., Pencheva, T.: Simulation of parallel processes in wastewater treatment plant using generalized net integrated development environment. Comptes rendus de l'Academie bulgare des Sci. **69**(11), 1493–1502 (2016)
28. Kosev, K., Melo-Pinto, P., Roeva, O.: Generalized net model of the lac operon in bacterium *E. coli*. In: Proceedings of the 6th IEEE International Conference on Intelligent Systems, pp. 237–241 (2012)
29. Pencheva, T., Roeva, O., Bentes, I., Barroso, J.: Generalized nets model for fixed-bed bioreactors. In: Proceedings of the 10th ISPE International Conference on Concurrent Engineering – Advanced Design, Production and Management Systems, Madeira, Portugal, 26–30 July 2003, pp. 1025–1028 (2013)
30. Ribagin, S., Roeva, O., Pencheva, T.: Generalized net model of asymptomatic osteoporosis diagnosing. In: Proceedings of the IEEE 8th International Conference on Intelligent Systems, pp. 604–608 (2016)
31. Ribagin, S., Chakarov, V., Atanassov, K.: Generalized net model of the scapulohumeral rhythm. In: Studies in Computational Intelligence, vol. 657, pp. 229–247 (2017)
32. Dimitrov, D., Roeva, O.: Development of generalized net for testing of different mathematical models of *E. coli* cultivation process. In: Advances in Intelligent Systems and Computing, vol. 322, pp. 657–668 (2015)
33. Roeva, O., Atanassova, V.: Generalized net model of cuckoo search algorithm. In: Proceedings of the IEEE 8th International Conference on Intelligent Systems, pp. 589–592 (2016)
34. Krawczak, M., Sotirov, S., Sotirova, E.: Generalized net model for parallel optimization of multilayer neural network with time limit. In: Proceedings of the 6th IEEE International Conference on Intelligent Systems, pp. 173–177 (2012)
35. Sotirov, S., Orozova, D., Sotirova, E.: Generalized net model of the process of the prognosis with feedforward neural network. In: Proceedings of the 16th International Symposium on Electrical Apparatus and Technologies, vol. 1, pp. 272–278 (2009)
36. Chountas, P., Atanassov, K., Sotirova, E., Bureva, V.: Generalized net model of an expert system dealing with temporal hypothesis. In: Advances in Intelligent Systems and Computing, vol. 400, pp. 473–481 (2016)
37. Peneva, D., Tasseva, V., Kodogiannis, V., Sotirova, E., Atanassov, K.: Generalized nets as an instrument for description of the process of expert system construction. In: Proceedings of the 3rd IEEE International Conference on Intelligent Systems, pp. 755–759 (2006)

38. Shannon, A.G., Riečan, B., Sotirova, E., Atanassov, K., Krawczak, M., Melo-Pinto, P., Parvathi, R., Kim, T.: Generalized net models of academic promotion and doctoral candidature. In: Studies in Computational Intelligence, vol. 657, pp. 263–277 (2017)
39. Shannon, A.G., Riečan, B., Orozova, D., Sotirova, E., Atanassov, K., Krawczak, M., Melo-Pinto, P., Parvathi, R., Kim, T.: Generalized net model of the process of selection and usage of an intelligent e-learning system. In: Proceedings of the 6th IEEE International Conference on Intelligent Systems, pp. 233–236 (2012)
40. Sotirova, E., Bureva, V., Sotirov, S.: A generalized net model for evaluation process using intercriteria analysis method in the university. In: Studies in Fuzziness and Soft Computing, vol. 332, pp. 389–399 (2016)
41. Atanassov, K., Gluhchev, G., Hadjitodorov, S., Kacprzyk, J., Shannon, A., Szmidt, E., Vassilev, V.: Generalized Nets Decision Making and Pattern Recognition. Warsaw School of Information Technology, Warszawa (2006)
42. Pencheva, T., Roeva, O., Shannon, A.: Generalized net models of basic genetic algorithm operators. In: Studies in Fuzziness and Soft Computing, vol. 332, pp. 305–325 (2016)
43. Roeva, O., Melo-Pinto, P.: Generalized net model of firefly algorithm. In: Proceedings of the 14th International Workshop on Generalized Nets, Burgas, 29 November 2013, pp. 22–27 (2013)
44. Fidanova, S., Atanassov, K., Marinov, P.: Generalized Nets in Artificial Intelligence. Generalized Nets and Ant Colony Optimization, vol. 5. Academic Publishing House "Prof. Marin Drinov", Sofia (2011)
45. Roeva, O., Shannon, A., Pencheva, T.: Description of simple genetic algorithm modifications using generalized nets. In: Proceedings of 6th IEEE International Conference on Intelligent Systems, Sofia, Bulgaria, Vol. 2, pp. 178–183 (2012)
46. Roeva, O., Atanassova, V.: Cuckoo search algorithm, firefly algorithm and artificial bee colony optimization in terms of generalized net theory. In: Computational Intelligence series. Springer (in press)
47. Altringham, J.D.: Bats: Biology and Behaviour. Oxford Univesity Press, Oxford (1996)
48. Yang, X.-S., Gandomi, A.H.: Bat algorithm: a novel approach for global engineering optimization. Eng. Comput. **29**(5), 464–483 (2012)
49. Yang, X.-S.: A new metaheuristic bat-inspired algorithm. In: Nature-Inspired Cooperative Strategies for Optimization. In: Studies in Computational Intelligence, vol. 284, pp. 65–74 (2010)
50. Gandomi, A.H., Yang, X.-S., Alavi, A.H., Talatahari, S.: Bat algorithm for constrained optimization tasks. Neural Comput. Appl. **22**(6), 1239–1255 (2012)

Insurance Portfolio Containing a Catastrophe Bond and an External Help with Imprecise Level—A Numerical Analysis

Maciej Romaniuk(✉)

Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, 01-447 Warszawa, Poland
mroman@ibspan.waw.pl

Abstract. In this paper, an integrated insurer's portfolio, which consists of a few layers of insurance and financial instruments, is numerically analysed. A future behaviour of such a portfolio is related to stochastic processes (like a random interest rate yield and uncertain catastrophic losses), therefore the Monte Carlo (MC) approach is applied. A special attention is paid to a problem of a share of catastrophe bonds in such a portfolio and to an analysis of an influence of an additional layer—an external (e.g. governmental) help. Some important measures of an insurer's risk (like a probability of his bankruptcy) are then numerically analysed. In considered examples, apart from strictly crisp sets of parameters, also fuzzy numbers are used to model an imprecise information concerning the possible external help.

Keywords: Risk process · Insurance portfolio · Catastrophe bond · Monte Carlo simulations · Probability of ruin · Governmental help · Fuzzy numbers

1 Introduction

Nowadays, the insurers face the problem of catastrophic losses, which are caused by earthquakes, tsunamis and other natural catastrophes. Therefore, a problem of an estimation and an analysis of a probability of an insurer's ruin is even more significant and urgent. Moreover, the insurers apply new, financial (or, simultaneously, financial and insurance) instruments, which are intended to lower this probability. A catastrophe bond (or a cat bond in short) is an example of such an instrument (see, e.g., [8,10,11]). However, an issuance of additional instruments changes a whole structure of an insurer's portfolio. Then, a classical risk process, which describes the cash flows of an insurer, should be also generalized to take into account these additional layers of the portfolio. This new formula of the risk process requires more complex approaches and supplementary numerical simulations in order to estimate the probability of an insurer's ruin and other statistics, which are important for an insurer.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_23

In this paper, we continue a work, which was started in [13,14]. Then, the generalized form of the classical risk process for the insurer's portfolio is considered. A cat bond, which is issued by an insurer, and an external help are examples of the layers in such a portfolio. Contrary to the classical approach, we also assume that there is dependency between time and money, i.e., one unit of money, which is paid now, has other value than the same unit, which will be paid in the future. In the following, cash flows for the insurer's portfolio are analysed using Monte Carlo (MC) simulations.

A contribution of this paper is fourfold. Firstly, a special attention is paid to a problem of a share of catastrophe bonds in the portfolio. An optimum level of the issued bonds is an important factor for the insurer. A larger share minimizes a probability of his bankruptcy, but it also minimizes expected profits of the insurer. Therefore a relevant numerical analysis is conducted. Secondly, a structure of the portfolio is further developed and an additional layer—an external (e.g., governmental or foreign) help—is incorporated. This next layer changes the mentioned generalized form of the classical risk process in a new way. Thirdly, we consider both a probabilistic and an imprecise approach to a value of such a help. In this second case, fuzzy triangular numbers are used to model this external help. Fourthly, in order to directly compare some important risk factors for the insurer, a method of a reduction of an estimation error is applied.

This paper is organized as follows. In Sect. 2, the generalized version of the classical risk process is introduced. Also an applied model of an interest rate (the one–factor Vasicek model) is recalled there. Some notes about a possible optimization procedure, which maximizes the cash flow for an insurer and minimizes his probability of a ruin, are included in Sect. 3. Section 4 is devoted to a numerical analysis of some examples, which are close to practical situations. Section 5 concludes the paper with some final remarks.

2 Risk Reserve Process and Its Generalization

Traditionally, in the insurance industry, a risk reserve process R_t is defined as a model of the financial reserves of an insurer depending on time t, i.e.

$$R_t = u + pt - C_t^* \tag{1}$$

where u is an initial reserve of the insurer, p is a rate of premiums paid by the insureds per unit time and C_t^* is a claim process, which is given by

$$C_t^* = \sum_{i=1}^{N_t} C_i \tag{2}$$

where $C_1, C_2, \ldots$ are iid random values of the claims. These claims are traditionally identified with the losses U_i, which are caused by the natural catastrophes, so we have $C_i = U_i$. There are also models, where the claims are only some part of the losses, e.g.,

$$C_i = \alpha_{\text{claim}} Z_i U_i \ , \tag{3}$$

so $\alpha_{\text{claim}} \in [0,1]$, $Z_i \sim U[c_{\min}, c_{\max}]$, and Z_i, U_i are mutually independent variables. The parameter α_{claim} describes a deterministic share of the considered insurer in the whole insurance market (for the given region) and the random variable Z_i models a random part of the claim C_i in the loss U_i. In this case, a non-informative random distribution, i.e. a uniform distribution $U[c_{\min}, c_{\max}]$ (for $0 \leq c_{\min} \leq c_{\max} \leq 1$), is used. Then, we have a process of the losses, which is given by

$$N_t^* = \sum_{i=1}^{N_t} U_i \tag{4}$$

If the assumption (3) is applied, it can lead to a hedging problem (see, e.g., [13]).

A process of a number of the claims $N_t \geq 0$ is usually driven by a homogeneous Poisson process (HPP), or a non-homogeneous Poisson process (NHPP). In this paper, we assume that a cyclic intensity function

$$\lambda_{\text{NHPP}}(t) = a + b2\pi \sin(2\pi(t-c)) \tag{5}$$

is used to model NHPP of the number of the claims N_t. The parameters of (5), which are applied in the following part of the paper, were estimated in [2], based on the data from the United States, provided by the Property Claim Services (PCS) of the ISO (Insurance Service Office Inc.). Then we have $a = 30.875, b = 1.684, c = 0.3396$. Also, using a method described in [2], the value of the single loss U_i is further modelled by a lognormal distribution with parameters $\mu_{\text{LN}} = 17.357, \sigma_{\text{LN}} = 1.7643$.

Because a non-constant intensity function (5) is applied, then the premium in (1) is fixed as a constant function for some deterministic moment T (see also [14] for further details), so

$$p(T) = (1+\nu_p) E C_i \int_0^T \lambda_{\text{NHPP}}(s) ds \tag{6}$$

where ν_p is a safety loading (or security loading) of the insurer, which is usually, in practical situations, about 10%–20%.

In the following, we consider a more complex insurer's portfolio, which consists of an additional layer—a special financial instrument, which is known as a catastrophe bond (or a cat bond, see, e.g., [8,10,11,14]) Therefore, the classical risk process (1) has to be generalized into a more suitable form, so that the cash flows related to the cat bond can be taken into account.

In general, when a catastrophe bond is issued, the insurer pays an insurance premium p_{cb} in exchange for a coverage, when a triggering point (usually some catastrophic event, like an earthquake) occurs. The investors purchase an insurance–linked security for cash. The above mentioned premium and cash flows are usually managed by a SPV (Special Purpose Vehicle), which also issues the catastrophe bonds. The investors hold the issued assets, whose coupons and/or principal depend on the occurrence of the mentioned triggering point. If such a catastrophic event occurs during the specified period, then the SPV compensates the insurer and the cash flows for the investors are changed. Usually, these flows are lowered, i.e. there is full or partial forgiveness of the repayment of principal and/or interest. However, if the triggering point does not occur, the investors usually receive the full payment from a cat bond (see, e.g., [11,13,14]).

Taking into account the described cash flows of a catastrophe bond, the classical risk process (1) should be written as

$$R_T = \mathrm{FV}_T(u - p_{\mathrm{cb}}) + \mathrm{FV}_T(p(T)) - \mathrm{FV}_T(C_T^*) + n_{\mathrm{cb}} f_{\mathrm{cb}}^i(N_T^*) \; , \tag{7}$$

where $f_{\mathrm{cb}}^i(N_T^*)$ is a payment function of the single cat bond for the insurer and p_{cb} is an insurance premium. We assume, that p_{cb} is proportional to both a part α_{cb} of a whole price of the single catastrophic bond I_{cb}, and to a number of the issued bonds n_{cb}, so that $p_{\mathrm{cb}} = \alpha_{\mathrm{cb}} n_{\mathrm{cb}} I_{\mathrm{cb}}$.

Moreover, in our setting (which is contrary to the classical approach, see also [13,14]), a value of money depending on time is taken into account. Therefore, $\mathrm{FV}_T(.)$ denotes a future value of the cash flow in (7). In the following, to calculate this future value, the one–factor Vasicek model

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t \tag{8}$$

is applied. The parameters for (8) are fitted in [1], so we get $\kappa = 0.1179, \theta = 0.086565, \sigma^2 = 0.0004$.

We can also enrich the considered portfolio and add some other layers (i.e. financial or insurance instruments), e.g. a reinsurance contract (see [14] for a more detailed discussion). But, in this paper, we focus only on a governmental (or, e.g., foreign), external help. We assume, that this help is supplied only if the losses surpass some given minimal limit A_{hlp}, and only with a fixed probability p_{hlp} (i.e. $\Pr(H = 1) = p_{\mathrm{hlp}}$ and $\Pr(H = 0) = 1 - p_{\mathrm{hlp}}$, where H is a binomial variable, which indicates, if this external fund is used or not). Then, a value of this help can be modelled by some function $f_{\mathrm{cb}}^i(N_T^*)$, e.g. by a constant value. If this external fund is incorporated into the generalized risk process (7), we get a new formula

$$\begin{aligned} R_T = \mathrm{FV}_T(u - p_{\mathrm{cb}}) + \mathrm{FV}_T(p(T)) - \mathrm{FV}_T(C_T^*) + n_{\mathrm{cb}} f_{\mathrm{cb}}^i(N_T^*) \\ + I(H = 1, N_T^* \geq A_{\mathrm{hlp}}) f_{\mathrm{hlp}}(N_T^*) \; , \end{aligned} \tag{9}$$

where $I(.)$ is an indicator function. Easily seen, such a help is treated as an additional source of funds by the insurer, because it lowers the overall losses and mitigates his expanses.

3 Optimization Goals

In a classical problem statement, an insurer is interested in a minimization of a probability of his ruin. For the given moment T, a probability of a ruin at the end of time interval T is given as

$$\phi(T) = \Pr(R_T < 0) \; . \tag{10}$$

Moreover, an insurer wants to maximize an overall cash flow for his portfolio, which is described by the generalized risk process (7) or (9). Therefore, in the following, we focus on an analysis of these two characteristics. Because of a stochastic and uncertain nature of (7) and (9), the MC approach is used to estimate an expected value of (7) or (9), namely ER_T. In practical situations, an insurer can be also interested in an overall optimization of his portfolio. Then, both the probability of the ruin and the expected value of the future cash flows can be combined in one optimization goal, e.g.,

$$\max(ER_T - \alpha_{\text{pen}} \Pr(R_T < 0)) , \tag{11}$$

where α_{pen} is some penalty factor, which is related to an occurrence of a ruin, and the maximum is taken for selected parameters of the portfolio (see, e.g., [3] for other approaches). In order to solve the problem (11), a stochastic optimization procedure can be necessary (see, e.g., [5]).

4 Numerical Analysis

As it was mentioned in Sect. 2, to model the trajectory of the process R_T, we apply NHPP with the intensity function (5) for the lognormal catastrophic losses. As for a payment function $f(N_T^*)$ for a holder of the considered cat bond, a piecewise linear function is applied (see [8,10,11,14] for a necessary introduction and an additional discussion), so

$$f(C_T^*) = \text{Fv}\left(1 - \sum_{i=1}^{n} \frac{\min(N_T^*, K_i) - \min(N_T^*, K_{i-1})}{K_i - K_{i-1}} w_i\right) \tag{12}$$

where Fv is a face value of the cat bond, $w_1, \ldots, w_n > 0$ are payoff decreases, and $0 \leq K_0 \leq K_1 \leq \ldots \leq K_n$ are the triggering points. We set $\text{Fv} = 1$ (i.e. one monetary unit assumption is used), and

$$K_0 = Q^{\text{loss}}_{\text{NHPP-LN}}(0.75), K_1 = Q^{\text{loss}}_{\text{NHPP-LN}}(0.9) , \tag{13}$$

where $Q^{\text{loss}}_{\text{NHPP-LN}}$ is x-th quantile of the cumulated value of the losses (for the considered NHPP and the lognormal distribution of the single loss). The payoff decrease is equal to $w_1 = 1$ and one year time horizon is applied, so $T = 1$. Then, if after one year, the cumulated value of losses surpasses K_1, the bond holder receives nothing. To find the price of such a catastrophe bond, we apply the method introduced in [8,10,11,14]. It requires analytical formulas and additional Monte Carlo simulations. Then, the mentioned price is estimated as $I_{\text{cb}} = 0.809896$ (see also [8,10,11,14] for a more detailed discussion), so such a value will be used further in this paper. We also assume, that $u = Q^{\text{loss}}_{\text{NHPP-LN}}(0.25)$, i.e. the initial reserve of the insurer is equal to 0.25-th quantile of the cumulated value of the claims, and that $\alpha_{\text{cb}} = 0.3$ (so 30% of the cat bond price is covered by the insurer) and $\nu_p = 0.1$ (i.e. the safety loading for the premiums is equal to 10%). For a better readability of results, the losses (hence, the claims also) are scaled in millions of money units.

4.1 Number of the Issued Cat Bonds

We start from an analysis of an influence of the number of the issued bonds n_{cb} on some key factors for the insurer, like a probability of his ruin. For a larger share of the cat bonds in the portfolio, the potentially catastrophic losses have lower impact on the insurer. This directly leads to the lower probability of his bankruptcy. On the other hand, the larger share also reduces the overall cash flows in the portfolio, even if the issued catastrophe bond will not be used afterwards (because a fixed triggering point of this cat bond is not even achieved). Hence, an issuance of the cat bonds works as an alternative way of a reinsurance (see also [14] for a comparison of these two approaches). Therefore, the insurer should choose an optimal level of the share of the catastrophe bonds in his portfolio. It should be not too large (because it does not maximize the expected insurer's profits) and not too low (because it leads to the higher probability of the bankruptcy at time T). In [13,14] there is no such an exact analysis.

In order to compare simulated outcomes for different values of n_{cb}, it is necessary to minimize other possible sources of variability. Therefore, to reduce a variance (and, furthermore, an estimation error), for an each value of n_{cb} the same set of $n = 1000000$ simulated trajectories is used. We also analyse three possible kinds of dependencies between the claims and the losses, namely $C_i = U_i$ (which is denoted further as Example I), $C_i = 0.5U_i$ (Example II, in this case each claim is always equal to 50% of the loss) and $C_i = Z_i U_i$, where $Z_i \sim U[0, 1]$ (Example III, the loss is transformed to the claim using a standard uniform distribution). Then, Example II reflects a situation, when the insurer has 50% of a whole insurance market, and Example III means, that there is no strict information about a level of such a share. Then, only a very general, non-informative statistical approach can be used (see [13] for a different approach to this problem).

Our analysis is done for a wide range of possible values of n_{cb}, which allows the insurer to directly compare his different possibilities in a construction of the portfolio. The estimated averages of the final value of the portfolio $\bar{R}_T$, as a function of n_{cb}, are plotted in Fig. 1 (outcomes for Example I are denoted by circles, for Example II—by squares, and for Example III—by rhombuses). They are almost linearly decreasing functions, which behave in a very similar way. However, an observed reduction of the estimated expected value is not very fast, e.g. in Example I for $n_{\text{cb}} = 50$ we have $\bar{R}_T = 3487.55$, and for $n_{\text{cb}} = 1600$ (i.e., the share of the cat bonds in such a portfolio is 32 times higher than in the previous case) we get $\bar{R}_T = 3349.06$ (only about 4% reduction).

It should be noted, that the averages $\bar{R}_T$ in Example II and Example III are very similar, but still they are not completely equal. It means, that even if an expected value of the loss in Example III is the same as a deterministic part of U_i in Example II, the outcomes are significantly different, which is rather in contrary to an "intuitive thinking".

Also the estimators of the ruin probabilities $\hat{\phi}(T)$ can be found in the similar way, using numerical simulations (see Fig. 2, the relevant plots are labelled in the same way, as previously). These probabilities are non-linearly decreasing

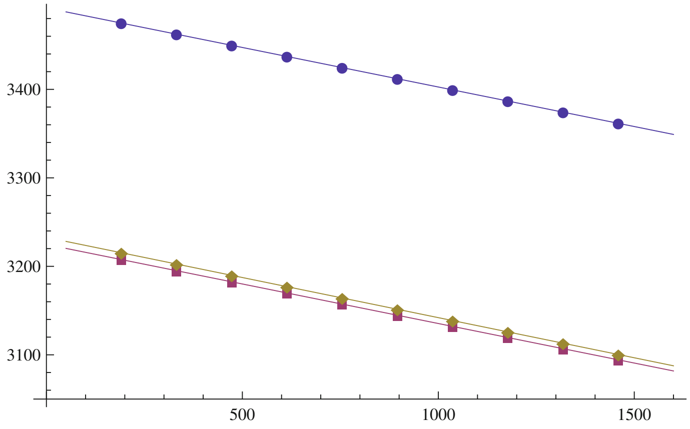


Fig. 1. Estimated averages of the final value of the portfolio (in Example I–Example III)

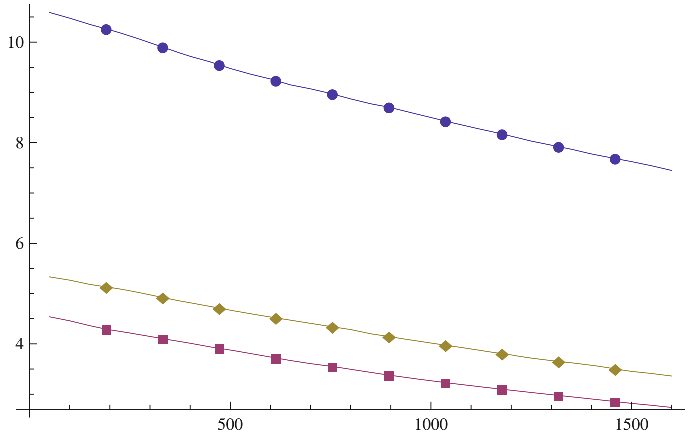


Fig. 2. Estimated probabilities of the final ruin (in Example I–Example III)

functions of n_{cb}. But now, the observed reduction for the increasing values of n_{cb} is more significant. In Example I, for $n_{\mathrm{cb}} = 50$ we have $\hat{\phi}(T) = 10.589\%$, and for $n_{\mathrm{cb}} = 1600$ we get $\hat{\phi}(T) = 7.448\%$ (almost 30% reduction of the ruin probability).

Because $\bar{R}_T$ and $\hat{\phi}(T)$ behave in a different way (linear vs. non-linear) as the functions of n_{cb}, then the outcomes, which are summarized in Figs. 1 and 2, can be directly merged using the optimization function (11) (or other one). Then, the optimal level of the issued cat bonds for the insurer can be directly found.

4.2 Influence of the External Help

We further develop our analysis of the insurer's portfolio and incorporate an additional layer—the external help (see (9)). As it was mentioned, the payment function for this help $f_{\text{hlp}}(N_T^*)$ can be modelled in various ways. In this paper, a function similar to a classical excess-of-loss policy is adopted. Then, we have

$$f_{\text{hlp}}(N_T^*) = \begin{cases} B_{\text{hlp}} - A_{\text{hlp}} & \text{if } \quad N_T^* \geq B_{\text{hlp}} \\ N_T^* - A_{\text{hlp}} & \text{if } B_{\text{hlp}} \geq N_T^* \geq A_{\text{hlp}} \end{cases} , \tag{14}$$

where B_{hlp} is a maximum limit for this help. Formulae (14) is, in some way, similar to commonly used reinsurance contracts (see, e.g., [14]), but without additional costs incurred by an insurer.

Let us suppose, that $n_{\text{cb}} = 1000$, the claims are equal to the losses, $p_{\text{hlp}} = 1$ (i.e., the help is always available, if the minimum limit of the losses is surpassed), and that $A_{\text{hlp}} = Q^{\text{loss}}_{\text{NHPP-LN}}(0.95), B_{\text{hlp}} = Q^{\text{loss}}_{\text{NHPP-LN}}(0.99)$, so the minimal limit for the external help is equal to 0.95-th quantile of the cumulated value of the losses and the maximal limit is given by 0.99-th quantile. Such a set of the parameters constitutes Example IV. Then, using simulations for the same set of trajectories as in Example I, the relevant outcomes can be easily compared. The average for the final value of the portfolio in Example IV is equal to 3598.21, comparing to 3402.67 in Example I (about 5.75% more in Example IV). However, a difference in the ruin probability is less visible—only about 0.01% (8.488% in Example IV vs. 8.498% in Example I).

Of course, in practical situations, A_{hlp} and B_{hlp} can be given as imprecise values, not as strictly precise information. For example, the minimum limit can be stated as "about $Q^{\text{loss}}_{\text{NHPP-LN}}(0.95)$". Such inexact data can be modelled with, e.g., fuzzy sets, in contrary to an application of real numbers (i.e., "exact" information, see [6,7,9,11,13] for examples of applications of the fuzzy numbers in some areas). Fuzzy sets can be also combined with a probabilistic approach, and this leads to random fuzzy variables (see, e.g., [4] for a more detailed review). Therefore, in the next case—Example V—we use triangular fuzzy numbers to describe A_{hlp} and B_{hlp}, and analyse influence of such an assumption on the simulated output. We restrict ourselves to the triangular fuzzy numbers, but the presented further approach can be also used for other kinds of L–R fuzzy numbers.

Let $\tilde{a} = [a_L, a_C, a_R]$ denote a triangular fuzzy number, where a_L is its left end of a support, a_R—its right end of a support, and a_C—a core. Then, $\tilde{a}[\alpha] = [a_L[\alpha], a_R[\alpha]]$ is an α-cut of $\tilde{a}$, if $\alpha \in [0, 1]$.

We assume that

$$\tilde{A}_{\text{hlp}} = [Q^{\text{loss}}_{\text{NHPP-LN}}(0.95) - 200, Q^{\text{loss}}_{\text{NHPP-LN}}(0.95), Q^{\text{loss}}_{\text{NHPP-LN}}(0.95) + 200] \tag{15}$$

(so, the minimum limit of the external help is 0.95-th quantile ± 200), and, in the same way,

$$\tilde{B}_{\text{hlp}} = [Q^{\text{loss}}_{\text{NHPP-LN}}(0.99) - 200, Q^{\text{loss}}_{\text{NHPP-LN}}(0.99), Q^{\text{loss}}_{\text{NHPP-LN}}(0.99) + 200] . \tag{16}$$

Then, simulations for consecutive α-cuts of $\tilde{A}_{\text{hlp}}$ and $\tilde{B}_{\text{hlp}}$ can be performed to obtain an outcome (i.e. an approximation of a fuzzy number) for a desired function $f(x)$. During the MC simulations, α is changed from some starting value $\alpha_0 \geq 0$ up to an upper bound $\alpha_1 \in (\alpha_0, 1]$ with an increment $\Delta\alpha > 0$. After an evaluation of the left and right end points of the different α-level sets of the considered function of the output, i.e. $[\tilde{f}_L[\alpha](x), \tilde{f}_R[\alpha](x)]$, the obtained intervals are put on one another, so they form an approximation of a final fuzzy outcome $\tilde{f}(x)$. During this procedure, we should keep in mind, if $f(x)$ is an increasing or decreasing function of the fixed x, in order to select relevant left or right ends of the α-cuts for $\tilde{A}_{\text{hlp}}$ and $\tilde{B}_{\text{hlp}}$ (see [11–13] for further details of this approach).

The estimated average of the final value of the portfolio forms a L–R fuzzy number, which is almost a triangular fuzzy number (see Fig. 3, a plot labelled with circles). Its support is equal to $[3586.59, 3610.27]$ (respectively, 5.5% and 6.1% more than in Example I) and its core is given by the relevant value from Example IV. A supplier of the external help can be also interested in an evaluation of a probability of using such a help. This value can be directly estimated, if the introduced approach is applied (see Fig. 4, a plot labelled with circles), and it is also a L–R fuzzy number. Its support is equal to $[4.65\%, 5.141\%]$ and its core is given by 4.88%.

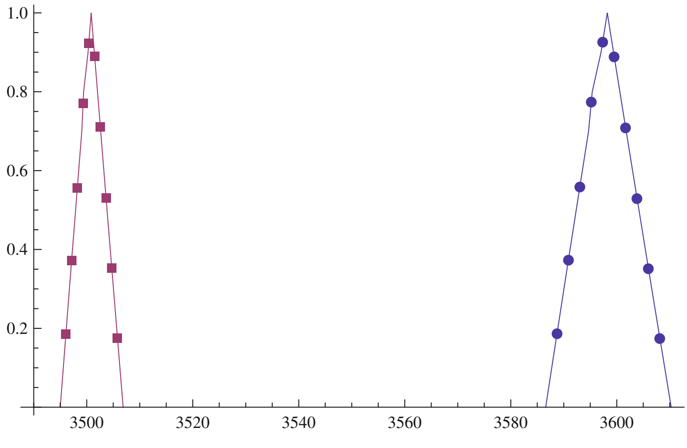


Fig. 3. Estimated averages of the final value of the portfolio (in Example V and Example VI)

An average is an important measure, however, a practitioner can be also interested in a more detailed analysis of other characteristics of the portfolio, e.g., a statistical behaviour of its final value. An example of such a study can be seen in Fig. 5, where a quantile plot for the final value of the insurer's portfolio is plotted. In this case, the quantiles for $\alpha = 0$ of $\tilde{A}_{\text{hlp}}$ and $\tilde{B}_{\text{hlp}}$ are calculated, using the approach described previously. Main differences between the portfolios are seen in Fig. 5 only for lower ranks of the quantiles, e.g., for 0.01-th quantile

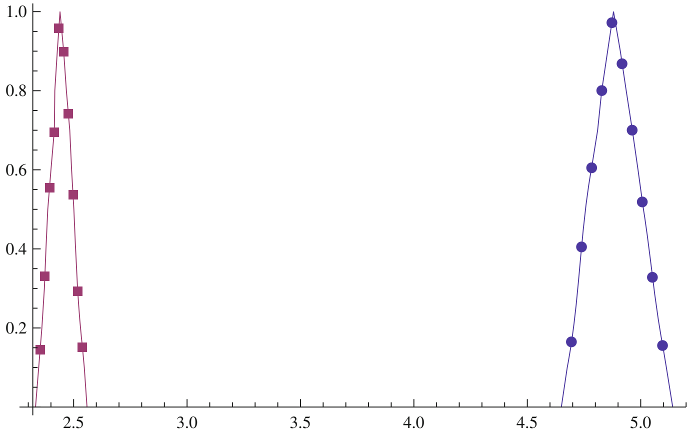


Fig. 4. Estimated probabilities of using the external help (in Example V and Example VI)

we have the final value of the portfolio −2722.21 versus −2247.15 (the difference is equal to 475.06), and for 0.99-th quantile we have the final value 7224.41 versus 7224.56 (so the difference is only 0.15). Then, a major effect of the external help is related rather to the "really catastrophic" events, which are statistically rare (only about 5% cases).

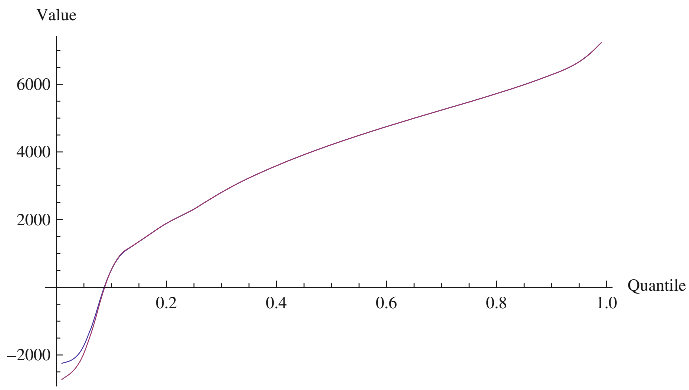


Fig. 5. Quantile plot of the final value of the portfolio (Example V)

In practical situations, we are not completely sure, if the external help will be supplied, i.e., we have $p_{\text{hlp}} \leq 1$. Therefore, we analyse a case (which is labelled further as Example VI), when $p_{\text{hlp}} = 0.5$ (so, there is 50% chance, that the external help can be used in a relevant situation) and all of the other parameters are the same as in Example V. Then, the estimated average of the final value of the portfolio and the probability, if the external help is used, can be seen in Figs. 3 and in 4 (plots labelled with squares). Easily seen, both of these fuzzy numbers are shifted to a left hand side, and their supports are narrower than in Example V.

5 Conclusions

In this paper, we focus on the analysis of the influence of the catastrophe bonds and the external help on the behaviour of the integrated insurer's portfolio. In order to evaluate the probability of the ruin and other important factors for the insurer, Monte Carlo simulations, together with the reduction of the estimation error, are applied. Then, various scenarios for the insurer's portfolio with different parameters are analysed. The outcomes from these examples are compared, using statistical measures. Apart from the crisp approach, the fuzzy numbers are also used to model an imprecise information, like the borderline limits of the external help.

References

1. Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B.: An empirical comparison of alternative models of the short-term interest rate. J. Finan. **47**(3), 1209–1227 (1992)
2. Chernobai, A., Burnecki, K., Rachev, S., Trück, S., Weron, R.: Modeling catastrophe claims with left-truncated severity distributions. Comput. Stat. **21**(3), 537–555 (2006). doi:10.1007/s00180-006-0011-2
3. Ermoliev, Y.M., Ermolieva, T.Y., MacDonald, G.J., Norkin, V.I.: Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks. Ann. Oper. Res. **99**(1), 207–225 (2000). doi:10.1023/A:1019244405392
4. Gil, M.A., Hryniewicz, O.: Statistics with imprecise data. In: Meyers, R.A. (ed.) Encyclopedia of Complexity and Systems Science, pp. 8679–8690. Springer, Heidelberg (2009)
5. Homem-de-Mello, T., Bayraksan, G.: Monte Carlo sampling-based methods for stochastic optimization. Surv. Oper. Res. Manag. Sci. **19**(1), 56–85 (2014). doi:10.1016/j.sorms.2014.05.001
6. Hryniewicz, O., Kaczmarek, K., Nowak, P.: Bayes statistical decisions with random fuzzy data–an application for the Weibull distribution. Eksploatacja i Niezawodnosc (Maintenance and Reliability) **17**(4), 610–616 (2015). doi:10.17531/ein.2015.4.18
7. Nowak, P., Pawłowski, M.: Option pricing with application of Levy processes and the minimal variance equivalent martingale measure under uncertainty. IEEE Trans. Fuzzy Syst. **25**(2), 402–416 (2017). doi:10.1109/TFUZZ.2016.2637372
8. Nowak, P., Romaniuk, M.: Pricing and simulations of catastrophe bonds. Insur. Math. Econ. **52**(1), 18–28 (2013). doi:10.1016/j.insmatheco.2012.10.006
9. Nowak, P., Romaniuk, M.: Application of Levy processes and Esscher transformed martingale measures for option pricing in fuzzy framework. J. Comput. Appl. Math. **263**, 129–151 (2014). doi:10.1016/j.cam.2013.11.031

10. Nowak P., Romaniuk, M.: Valuing catastrophe bond involving correlation and CIR interest rate model. Computational and Applied Mathematics (2016). doi:10.1007/s40314-016-0348-2
11. Nowak, P., Romaniuk, M.: Catastrophe bond pricing for the two-factor Vasicek interest rate model with automatized fuzzy decision making. Soft Comput. **21**(10), 2575–2597 (2017). doi:10.1007/s00500-015-1957-1
12. Romaniuk, M.: On simulation of maintenance costs for water distribution system with fuzzy parameters. Eksploatacja i Niezawodnosc (Maintenance and Reliability) **18**(4), 514–527 (2016). doi:10.17531/ein.2016.4.6
13. Romaniuk, M.: Analysis of the insurance portfolio with an embedded catastrophe bond in a case of uncertain parameter of the insurer's share. In: Wilimowska, Z., Borzemski, L., Grzech, A., Świątek, J. (eds.) Information Systems Architecture and Technology: Proceedings of 37th International Conference on Information Systems Architecture and Technology—ISAT 2016—Part IV. Advances in Intelligent Systems and Computing, vol. 524, pp. 33–43. Springer International Publishing (2017). doi:10.1007/978-3-319-46592-0_3
14. Romaniuk, M., Nowak, P.: Monte Carlo Methods: Theory, Algorithms and Applications to Selected Financial Problems. ICS PAS, Warszawa (2015)

Global Quality Measures for Fuzzy Association Rule Bases

Pavel Rusnok(✉) and Michal Burda

Centre of Excellence IT4Innovations,
Institute for Research and Applications of Fuzzy Modeling,
University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic
pavel.rusnok@osu.cz, michal.burda@osu.cz

Abstract. Association rules and fuzzy association rules are vastly studied topics. Various measures for quantifying a quality of a (fuzzy) association rule were proposed in the past. In this article, we survey existing and propose some new quality measures for the whole rule bases of fuzzy association rules.

Keywords: Global measures · Fuzzy rules · Fuzzy associations · Data mining · Rule bases

1 Introduction

Association rules belong to an exploratory part of data mining and pattern recognition. Association rule is a formula of the form $A \rightharpoonup C$, where A is called antecedent and C is a consequent, and which denotes some implicative relationship between A and C. Where the relationship is not directional, the rules have the form $A \sim B$. Also, other rule types were defined in the literature. Association analysis is a process for automatic search of such rules in data. In this paper, we focus on implicative association rules.

The first who come with the idea of association analysis was Hájek et al. in the late 1960s. They formulated GUHA (General Unary Hypotheses Automaton) method [1], a very general theory for pattern recognition and automatic hypotheses formation. For a more recent survey, see [2] or a book treating its part called Observational Calculi [3].

A very similar approach to the original GUHA method was independently reinvented by Agrawal [4] in 1993. Fuzzification of association rules was proposed later by many authors – a very good survey might be found in [5].

The so-called linguistic summaries, which were independently proposed by Yager in [6] (even before Agrawal) and further developed by Kacprzyk [7], are very closely related to fuzzy association rules. There is a significant overlap between these two research directions. For example, *support* in association rules is in fact *degree of focus* in linguistic summaries [8].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_24

In all of the mentioned research, a lot of attention was given to measuring and evaluating the quality of a single association rule $A \rightharpoonup C$. However, there is not much developed to measure the quality of a set of fuzzy association rules (linguistic summaries) as a whole. There are some generalizations made for classification rules to fuzzy classification rules in [9], but the case with fuzzy consequent is not treated. Furthermore, some measures for sets of IF-THEN rules of fuzzy inference systems are proposed in [10], but only with stress on interpretability. Otherwise, there is not much done for measuring a quality of the sets of fuzzy rules (fuzzy associations).

In this paper, we survey the work that was done so far, both for crisp and fuzzy rules, and propose some generalizations. We believe that it is very useful to measure the quality of a rule base as a whole. In certain contexts, the quality may be understood as interpretability of the rules, whereas in other contexts, it might be the coverage of data, i.e. how big part of data is summarized in the rule base or what fraction of data is not described by the rules. There are many different approaches for capturing the rule base quality, and we provide here a varied set of measures that might be used and successfully implemented.

In this text, we concentrate on rule bases of fuzzy associations. We use language that is usual for the audience focusing on association analysis. However, we are convinced that a researcher in linguistic summaries might also benefit from this survey.

Please note also that the word *measure* will be used in this study rather loosely and not all measures presented and proposed here are fulfilling axioms of measure in the mathematical sense.

Our contribution is organized in the following way. In Sect. 2 we provide and review definitions of basic notions. Then in Sects. 3 through 6 we define various measures of rule bases divided to different classes and finally in Sect. 7 we conclude.

2 Preliminaries

Let $\mathcal{O} = \{o_1, o_2, \ldots, o_N\}$, $N > 0$, be a finite set of objects and $\mathcal{A} = \{a_1, a_2, \ldots, a_M\}$, $M > 0$, be a finite set of attributes. Dataset D is a mapping that assigns to each object $o \in \mathcal{O}$ and attribute $a \in \mathcal{A}$ a degree $D(a, o) \in [0, 1]$, which represents the intensity of assignment of attribute a to object o. For fixed D, we can treat the attribute a as a predicate, which assigns a truth value $a(o) \in [0, 1]$ to each $o \in \mathcal{O}$. Similarly for each subset $X \subseteq \mathcal{A}$ of attributes, we define a predicate $X(o)$ by using a t-norm $\otimes$ as follows:

$$X(o) = \bigotimes_{a \in X} a(o). \tag{1}$$

T-norm $\otimes$ is a generalized logical conjunction, i.e. a function $[0, 1] \times [0, 1] \to [0, 1]$ which is associative, commutative, monotone increasing (in both places) and which satisfies the boundary conditions $\alpha \otimes 0 = 0$ and $\alpha \otimes 1 = \alpha$ for each $\alpha \in [0, 1]$. Some well-known examples of t-norms are for $\alpha, \beta \in [0, 1]$ as follows:

- product t-norm: $\alpha\beta$;
- minimum t-norm: $\min(\alpha, \beta)$;
- Łukasiewicz t-norm: $\max(0, \alpha + \beta - 1)$.

Since we use X to represent a crisp set of attributes as well as a predicate with fuzzy truth value, we need to define two types of cardinalities: $|X|$ will represent a cardinality of the crisp set X and $|X|_\Sigma$ will represent the sigma-count derived from (1) as follows (Table 1):

$$|X|_\Sigma = \sum_{o \in \mathcal{O}} X(o).$$

Table 1. A table representing an initial dataset D where $e_{ij} = a_j(o_i)$.

	a_1	a_2	$\dots$	a_m
o_1	e_{11}	e_{21}	$\dots$	e_{m1}
o_2	e_{12}	e_{22}	$\dots$	e_{m2}
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
o_N	e_{1N}	e_{2N}	$\dots$	e_{mN}

Association rule is a formula $A \rightharpoonup C$, where $A \subseteq \mathcal{A}$ is the *antecedent* and $C \subseteq \mathcal{A}$ is the *consequent*. It is natural to assume $A \cap C = \emptyset$. We have often also $|C| = 1$. A finite set of association rules is called as *rule base*.

So far, a lot has been written about association rules and their properties [1,4,5,11]. Both crisp and fuzzy association rules were considered. Especially, the algorithms for automated extraction of interesting rules from data have been developed in past research. These algorithms are mainly based on the traversal through the search space where various combinations of antecedents and consequents are examined. However, a fundamental problem of the search is how to define and measure the interestingness of a rule. Perhaps the most commonly known indicators of rule quality are the *support* and *confidence*, which are defined for a fuzzy association rule $A \rightharpoonup C$ as follows:

$$\operatorname{supp}(A \rightharpoonup C) = \frac{|A \cup C|_\Sigma}{|\mathcal{O}|} = \frac{1}{N} \sum_{o \in \mathcal{O}} A(o) \otimes C(o), \tag{2}$$

$$\operatorname{conf}(A \rightharpoonup C) = \frac{|A \cup C|_\Sigma}{|A|_\Sigma} = \frac{\sum_{o \in \mathcal{O}} A(o) \otimes C(o)}{\sum_{o \in \mathcal{O}} A(o)}. \tag{3}$$

The search for fuzzy association rules is then driven by the user defined thresholds $\operatorname{supp}_t, \operatorname{conf}_t \in [0, 1]$ and only rules with support and confidence above the given thresholds are included into the resulting rule base.

There exist many other rule quality measures. An overview of them can be found in [12]. Fuzzy generalization of other measures besides support and confidence (e.g. lift, leverage or conviction) is discussed in [13,14].

Although we set considerable constraints on the quality of obtained rules, it is quite usual that the association rules extracting process ends up with a rule base containing hundreds or thousands of rules. A natural question is then how to measure the quality of the obtained rule base as a whole. Surprisingly, not so much work has been done so far on that topic and this is where our paper tries to fill the gap: in the subsequent sections, we try to outline several definitions of rule base quality measures.

3 Statistical Characteristics of Rule Quality Measures

An obvious way to measure the differences between sets of rules is via various statistics of known quality measures for particular rules. Let $R = \{r_1, r_2, \ldots, r_K\}$ be a rule base consisting of $K > 0$ association rules. Let q be a rule quality measure. (For instance, q could be support, confidence, lift or any other measure.) As long as the domain of q is a subset of real numbers, we can define, for any rule base R, a vector $Q_R = (q(r_1), q(r_2), \ldots, q(r_K))$ and characterize the quality of the rule base R by describing the distribution of values in Q_R, e.g. by:

- arithmetic mean and standard deviation;
- median and quartiles;
- minimum and maximum;
- graphically, by depicting a histogram of values in Q_R, or a box-plot.

Moreover, two rule bases R_1 and R_2 may be compared using the appropriate two-sample statistical test to evaluate the significance of their difference. (Perhaps the Wilcoxon rank sum test would be the most appropriate as it does not require the values to be normal.)

4 Inference Driven Rule Base Measures

The utility from the base of association rules may be two-fold. The association rules may be analyzed directly by the user, who may transform them into some useful knowledge. Alternatively, the rule base may serve as an input to some inference mechanism in order to perform automatic classifications or predictions on unseen future data.

A set $\{y_1, \ldots, y_K\}$ of target attributes usually forms a (fuzzy) partition on $\mathcal{O}$. The rest of attributes ($\mathcal{A} \setminus \{y_1, \ldots, y_K\}$) are considered as inputs. The process of creation of rule base R usually assumes the set $\mathcal{O}$ of objects to be split into two disjoint parts: training $\mathcal{O}_{\text{train}}$ and testing $\mathcal{O}_{\text{test}}$. The training set is used to create the rule base, i.e. a set of rules $A \rightharpoonup \{y_k\}$ for $k \in \{1, \ldots, K\}$. (There exist strategies of how to split the source dataset into training and testing parts, e.g. cross-validation or leave-one-out strategy, see [15, Chap. 6.13] for more information.) The testing part of objects can then be used to evaluate the quality of predictions.

Accordingly, to the target, we distinguish at least between two types of predictions: *classification* (a prediction of a categorical variable) and *regression* (a prediction of a numerical variable).

In regression, the inference mechanism I employs some defuzzification method to obtain a numeric value $\hat{y}(o)$ as a result. The inferred $\hat{y}(o)$ can be compared with known $y(o)$ (for each $o \in \mathcal{O}_{\text{test}}$) by computing some error measure, e.g. root mean squared error,

$$RMSE = \sqrt{\frac{1}{|\mathcal{O}_{\text{test}}|} \sum_{o \in \mathcal{O}_{\text{test}}} (y(o) - \hat{y}(o))^2},$$

or other such as absolute error, mean absolute error, or precision in the case of classification [15, Chap. 6.12].

It might be computationally expensive to run the inference for all objects in $\mathcal{O}_{\text{test}}$, as pointed out in [16]. For that purpose, a similarity measure of rule bases was proposed in [16]. It was shown there that the proposed similarity was consistent with the difference in $RMSE$.

A similar investigation was done in [8], where the similarity between linguistic summaries was proposed.

This leads us to a task how to evaluate a rule base without using it in the inference mechanism I at all.

5 Rule Base Measures Based on Singe Rule Validity

In classification, we may use error measures designed for crisp classifiers and generalize them for fuzzy rules as proposed by Holeňa in [9] who defines the *inaccuracy* (also called Brier score) and *imprecision* based on cardinalities of fuzzy subsets of $\mathcal{O}$ as follows:

$$\text{Inacc} = 1 - \frac{|\mathcal{O}^+|_\Sigma - |\mathcal{O}^-|_\Sigma}{|\mathcal{O}|}, \quad \text{Impr}_1 = 1 - \frac{|\mathcal{O}^+|_\Sigma}{|\mathcal{O}|}, \quad \text{Impr}_2 = 1 - \frac{|\mathcal{O}^+|_\Sigma}{|\mathcal{O}_R|_\Sigma}, \quad (4)$$

where $\mathcal{O}^+$, $\mathcal{O}^-$ and $\mathcal{O}_R$ are fuzzy subsets of $\mathcal{O}$ such that:

- $\mathcal{O}^+(o)$ is a degree in which a *rule base R is valid for o*,
- $\mathcal{O}^-(o)$ is a degree in which a *rule base R is not valid for o*,
- $\mathcal{O}_R(o) = \max\{A(o) : (A \rightharpoonup C) \in R\}$.

Holeňa's approach [9] requires a definition of arbitrary *validity of rule $(A \rightharpoonup C) \in R$ for object $o \in \mathcal{O}$*, denoted with $(A \rightharpoonup C)(o)$. Evaluation of this notion is immaterial for his generalization, and the reader can imagine e.g. fuzzy implication there. Accordingly to [9], three variants of *rule base validity* were defined based on:

1. simultaneous validity of all covering rules,

$$\mathcal{O}^+ = \{o \in \mathcal{O} : \mathcal{O}_R(o) > 0 \text{ and } \forall (A \rightharpoonup C) \in R : A(o) \otimes (A \rightharpoonup C)(o) = A(o)\},$$
$$\mathcal{O}^- = \{o \in \mathcal{O} : \mathcal{O}_R(o) > 0 \text{ and } \exists (A \rightharpoonup C) \in R : A(o) \otimes (A \rightharpoonup C)(o) < A(o)\},$$

2. the validity of the majority of covering rules,

$$\mathcal{O}^{+} = \{o \in \mathcal{O} : \mathcal{O}_R(o) > 0 \text{ and } \sum_{(A \rightharpoonup C) \in R} A(o) \otimes (A \rightharpoonup C)(o) > \sum_{(A \rightharpoonup C) \in R} A(o) \otimes \neg(A \rightharpoonup C)(o)\},$$

$$\mathcal{O}^{-} = \{o \in \mathcal{O} : \mathcal{O}_R(o) > 0 \text{ and } \sum_{(A \rightharpoonup C) \in R} A(o) \otimes (A \rightharpoonup C)(o) \leq \sum_{(A \rightharpoonup C) \in R} A(o) \otimes \neg(A \rightharpoonup C)(o)\},$$

3. and the relative validity of covering rules,

$$\mathcal{O}^{+}(o) = \frac{\sum_{(A \rightharpoonup C) \in R} A(o) \otimes (A \rightharpoonup C)(o)}{\sum_{(A \rightharpoonup C) \in R} A(o)},$$

$$\mathcal{O}^{-}(o) = \frac{\sum_{(A \rightharpoonup C) \in R} A(o) \otimes \neg(A \rightharpoonup C)(o)}{\sum_{(A \rightharpoonup C) \in R} A(o)}.$$

Note that the sets $\mathcal{O}^{+}$ and $\mathcal{O}^{-}$ are always crisp in the first two cases, and fuzzy in the third case. Crisp definitions of $\mathcal{O}^{+}$ and $\mathcal{O}^{-}$, called "Boolean" in [9], made sense for crisp consequents in fuzzy classification. However, we propose here their possible fuzzification using a concept of fuzzy equality:

4. the minimal validity of all covering rules, where

$$\mathcal{O}^{+}(o) = \min_{A \rightharpoonup C} \{1 - |A(o) \otimes (A \rightharpoonup C)(o) - A(o)|\},$$

$$\mathcal{O}^{-}(o) = 1 - \mathcal{O}^{+}(o)$$

if $\mathcal{O}_R(o) > 0$, and

$$\mathcal{O}^{+}(o) = \mathcal{O}^{-}(o) = 0$$

otherwise.

In this way, we can avoid unnatural crisp behavior when all rules have validity 1 and only one rule has validity 0.9 for a particular object. Similarly, we can fuzzify the validity of the majority of covering rules to the validity of Σ-majority in the following way:

5. the validity of Σ-majority of all covering rules, where

$$\mathcal{O}^{+}(o) = \min \left\{1, \frac{\sum_{(A \rightharpoonup C) \in R} A(o) \otimes (A \rightharpoonup C)(o)}{\sum_{(A \rightharpoonup C) \in R} A(o) \otimes \neg(A \rightharpoonup C)(o)}\right\},$$

$$\mathcal{O}^{-}(o) = 1 - \mathcal{O}^{+}(o)$$

if $\mathcal{O}_R(o) > 0$, and

$$\mathcal{O}^{+}(o) = \mathcal{O}^{-}(o) = 0$$

otherwise.

In addition to quality measures (4), we can use the variants of $\mathcal{O}^+$ and $\mathcal{O}^-$ to provide a novel definition of *rule base support* and *rule base confidence*:

$$\text{rbsupp} = \frac{|\mathcal{O}^+|}{|\mathcal{O}|},$$
$$\text{rbconf} = \frac{|\mathcal{O}^+|}{|\mathcal{O}^+| + |\mathcal{O}^-|}.$$

Even other interestingness measures might be generalized to the rule bases. However, it is not yet clear whether it is a reasonable direction. Further theoretical and empirical study would be appropriate here.

All of the above rule base quality measures might be used in rule base reductions. This leads to the last class of measures that were already used in rule base reduction.

6 Coverage-Based Measures

Coverage-based measures quantify the amount of source data or volume of input space that are in some way touched with the rules in the rule base.

A *dataset coverage by antecedents* introduced in [17] measures the fraction of data, for which some antecedent of rules from given rule base gets fired:

$$\text{acov}(R) = \frac{|\mathcal{O}_R|_\Sigma}{|\mathcal{O}|}. \tag{5}$$

The *probabilistic coverage*, introduced in [18], estimates the pcov(R) probability of firing a rule from rule base R on random input. Let $\iota \in I$ represent an input vector from an (infinite) space I of all possible inputs and let $p(\iota)$ represents a probability density function on I. (Obviously, $\mathcal{O} \subseteq I$.) The rule base R splits the input space I into two parts: I^+ denoting inputs, for which there is a rule in R that fires, and I^- denoting the rest. Thus $I = I^+ \cup I^-$ and $I^+ \cap I^- = \emptyset$. The probabilistic coverage pcov(R) then equals to

$$\text{pcov}(R) = P(I^+) = \int_{I^+} p(\iota)\mathrm{d}\iota.$$

In [18], pcov(R) is simply estimated with

$$\text{pcov}(R) = \frac{|\mathcal{O}_R^0|}{|\mathcal{O}|}, \tag{6}$$

where

$$\mathcal{O}_R^0 = \{o \in \mathcal{O} : \mathcal{O}_R(o) > 0\}.$$

Both acov(R) and pcov(R) were compared in [18]. Also, rule base reduction methods based on these two coverage measures were empirically evaluated. Both methods find in certain sense locally optimal subsets of rule bases and thus each method was winning on different data sets. For more details see [17,18].

7 Conclusion

We have surveyed, explored and generalized various rule base interestingness and quality measures. We have omitted to discuss the interpretability measures in [10], which are even in our broad discussion a little bit out of our scope. We hope we have provided such measures that correlate with the quality of inference mechanism, in which the measured rule base will be used, or that somehow summarize the quality of particular rules in the base. Interpretability for sure does not correlate with accuracy and is rather contrary to precision.

In the future, we would like to focus our research on the study of the properties of the rule base quality measures, to compare them, and to highlight their similarities, differences, benefits and drawbacks. Also, an experimental studies are needed here to obtain a feedback from real world applications.

We hope to inspire readers to further investigation of the various classes of measures proposed here.

Acknowledgment. This work was supported by the NPU II project LQ1602 IT4Innovations excellence in science.

References

1. Hájek, P., Havránek, T.: Mechanizing Hypothesis Formation. Mathematical Foundations for a General Theory. Springer-Verlag, Heidelberg (1978)
2. Hájek, P., Holeňa, M., Rauch, J.: The guha method and its meaning for data mining. J. Comput. Syst. Sci. **76**(1), 34–48 (2010)
3. Rauch, J.: Observational Calculi and Association Rules, vol. 469. Springer, Heidelberg (2012)
4. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp. 207–216 (1993)
5. Ralbovský, M.: Fuzzy GUHA. Ph.D. thesis, University of Economics, Prague (2009)
6. Yager, R.R.: A new approach to the summarization of data. Inf. Sci. **28**(1), 69–86 (1982)
7. Kacprzyk, J., Yager, R.R., Zadrożny, S.: A fuzzy logic based approach to linguistic summaries of databases. Int. J. Appl. Math. Comput. Sci. **10**(4), 813–834 (2000)
8. Wilbik, A., Keller, J.M.: A distance metric for a space of linguistic summaries. Fuzzy Sets Syst. **208**, 79–94 (2012)
9. Holeňa, M.: Measures of ruleset quality for general rules extraction methods. Int. J. Approximate Reasoning **50**(6), 867–879 (2009)
10. Gacto, M.J., Alcalá, R., Herrera, F.: Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures. Inf. Sci. **181**(20), 4340–4360 (2011)
11. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Data Bases, VLDB 1994, pp. 487–499. Morgan Kaufmann Publishers Inc., San Francisco (1994)
12. Geng, L., Hamilton, H.J.: Interestingness measures for data mining: a survey. ACM Comput. Surv. (CSUR) **38**(3), 9 (2006)

13. Burda, M.: Lift Measure for Fuzzy Association Rules, pp. 249–260. Springer International Publishing, Cham (2015)
14. Burda, M.: Interest measures for fuzzy association rules based on expectations of independence. Adv. Fuzzy Syst. **2014**, 2 (2014)
15. Han, J.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers Inc., San Francisco (2005)
16. Li, H., Dick, S.: A similarity measure for fuzzy rulebases based on linguistic gradients. Inf. Sci. **176**(20), 2960–2987 (2006)
17. Burda, M., Štěpnička, M.: Reduction of fuzzy rule bases driven by the coverage of training data. In: Proceedings of the 16th World Congress of the International Fuzzy Systems Association and 9th Conference of the European Society for Fuzzy Logic and Technology, IFSA-EUSFLAT (2015)
18. Rusnok, P.: Probabilistic coverage of linguistic if-then rules. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 798–803. IEEE (2016)

Particle Swarm Optimization with Fuzzy Dynamic Parameters Adaptation for Modular Granular Neural Networks

Daniela Sánchez, Patricia Melin(✉), and Oscar Castillo

Tijuana Institute of Technology, Tijuana, Mexico
danielasanchez.itt@hotmail.com,
{pmelin, ocastillo}@tectijuana.mx

Abstract. In this paper a new method for Modular Granular Neural Network (MGNN) optimization with a granular approach is presented. A Particle Swarm Optimization technique is proposed to perform the granulation of information with a fuzzy dynamic parameters adaptation to prevent stagnation. The proposed fuzzy inference system seeks to adjust some PSO parameters such as w, C_1 and C_2 to ensure that the parameters have adequate values depending on the current behavior of the particles. The objective of the proposed PSO is design optimal MGNN architectures. The modular granular neural networks are applied to human recognition based on iris biometrics, where a benchmark database is used and the objective function in this work is the minimization of the error of recognition.

1 Introduction

The combination of two or more intelligent techniques have allow to generate powerful intelligent systems overcoming limitations that each technique has individually [20, 21]. Nowadays, there are many works performing this kind of systems, where good results have been shown. Techniques such as fuzzy logic [25], neural networks [7], data mining [22], and granular computing [24], among others, have important parameters that depending of their values the final results can be affected. To perform the correct establish of parameters there are other important techniques known as optimization techniques, among the most important are: Genetic Algorithm (GA) [8], Particle Swarm Optimization (PSO) [11], Ant Colony Optimization (ACO) [5], Cuckoo Optimization Algorithm (COA) [17], Harmony Search (HS) [6], Gravitational Search Algorithm (GSA) [18], Bee Colony Optimization (BCO) [13]. In this paper different intelligent techniques are combined such as neural networks, fuzzy logic and particle swarm optimization. The proposed method was applied to human recognition based on iris biometrics. Optimizing some parameters of MNN such as; number of sub modules, percentage of information for the training phase and number of hidden layers (with their respective number of neurons) for each sub module and learning algorithm, the objective function the minimization of the error of recognition. This paper is organized as follows: The basic concepts used in this work are presented in Sect. 2. Section 3,

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_25

the general architecture of the proposed method is shown. Section 4 presents experimental results and in Sect. 5, the conclusions of this work are presented.

2 Basic Concepts

In this section to understand the proposed method, the basic concepts used in this research work are presented.

2.1 Modular Neural Networks

A Modular Neural Network (MNN) consists of several modules (artificial neural networks), where each module carrying out one sub-task of a global task and is based on the principle of divide and conquers. Each module works independently in its own domain and is build and trained for a specific subtask. The simpler subtasks are then accomplished by a number of the specialized local computational systems or models which are integrated together via an integrating unit [14]. The learning mode can be supervised or unsupervised [21]. This kind of neural network has a wide application area such as: pattern recognition, function approximation, clustering or time series prediction [1].

2.2 Fuzzy Logic

The term fuzzy logic was introduced in 1965 by L. Zadeh. Fuzzy logic (FL) has the ability to mimic the human mind to effectively employ modes of reasoning that are approximate rather than exact. This technique is a useful tool for modeling complex systems and deriving useful fuzzy relations or rules [15, 25]. The basic structure of a fuzzy inference system consists of three conceptual components: a rule base, which contains a selection of fuzzy rules, a database which defines the membership functions used in the rules, and a reasoning mechanism that performs the inference procedure [3]. These intelligent techniques have been successfully applied in different areas such as detection of edges, feature extraction, classification, and clustering [10, 12].

2.3 Granular Computing

L. Zadeh originally proposed Granular computing (GrC). A granule may be interpreted as one of the numerous small particles forming a larger unit [23], these granules are collections of entities that usually originate at the numeric level and are arranged together due to their similarity, functional or physical adjacency, coherency, or the like [9]. This concept has been applied in relevant fields such as bioinformatics, e-Business, security, machine learning, data mining, cluster analysis, databases and knowledge discovery [2, 16].

2.4 Particle Swarm Optimization

Particle Swarm Optimization (PSO) was developed by Kennedy and Eberhart in 1995 [11], this optimization technique is based on the social behaviors of a flock of birds or a schooling fish. This algorithm doesn't have any leader in their group or swarm, unlike other algorithms. The flocks achieve their best condition simultaneously through communication among members who already have a better situation or position. The member of the flock with better condition or position will inform it to its flocks and the others will move simultaneously to that place. Particle swarm optimization consists of a swarm of particles, where particle represent a potential solution [19].

3 Proposed Method

The proposed method consists in design optimal architectures of MGNNs using a PSO with a fuzzy dynamic parameter adaptation to prevent stagnation. The main idea is to find an optimal granulation of the information and optimal MGNNs architectures minimizing the error of recognition. The optimization algorithms seeks an optimal number of sub modules or granules, having as search space up to "m" sub modules or granules, percentage of information for the training phase, number of hidden layers (with their respective number of neurons) for each sub module and learning algorithm. Figure 1 shows the architecture of proposed method for MGNNs optimization.

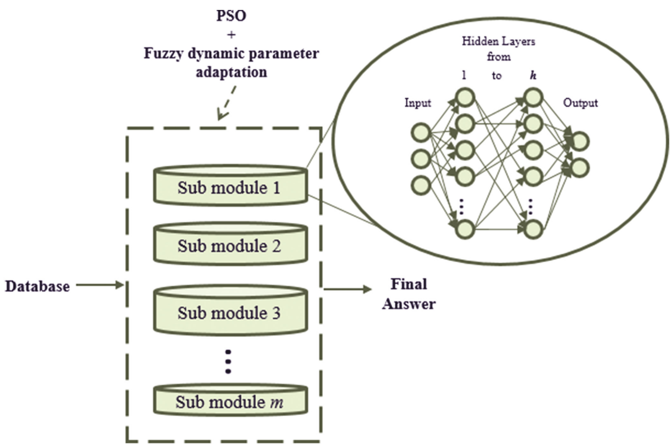


Fig. 1. Architecture of proposed method for MGNN optimization

The minimum and maximum values used for the MGNNs optimization are shown in Table 1. Those parameters are used to establish the search space of each optimization technique, for this work, the database used is described later.

Table 1. Values for MGNNs

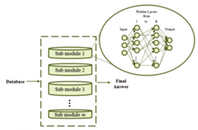
Parameters of MNNs	Minimum	Maximum
Modules (m)	1	10
Percentage of data for training	20	80
Error goal	0.000001	0.001
Learning algorithm	1	3
Hidden layers (h)	1	5
Neurons for each hidden layers	20	300

For the learning algorithm, 3 backpropagation algorithms are used to perform the modular neural networks simulations: Gradient descent with scaled conjugate gradient (SCG), Gradient descent with adaptive learning and momentum (GDX) and Gradient descent with adaptive learning (GDA).

3.1 Particle Swarm Optimization with Fuzzy Dynamic Parameters Adaptation

Particle swarm optimization has important parameters that help the convergence towards an optimal result, for example: The w parameter can facility exploration and exploitation depending of its value. The values of C_1 and C_2 are the cognitive and social components that influence the velocity of each particle. While w is a decreasing value, C_1 and C_2 have fixed values during an evolution. These parameters are usually initialized: to trial and error, depending of our experience or depending area of application.

The proposed PSO seeks to adjust these parameters depending of the actual population behavior using a fuzzy inference system to obtain an update of w, C_1 and C_2 based on number of iterations and actual w to avoid a stagnation in the result during a certain number of iterations. The pseudocode of the proposed optimization is shown in Fig. 2, where can be observed that before update velocity and position, the fuzzy inference system update w, C_1 and C_2. In Table 2, the initial parameters for the PSO are shown.



```
Initialize population of particles
while ( t < Max number of iterations)
    for each particle
        Calculate the fitness of the particle (evaluate the MGNN)
        If the fitness value is better than the best fitness value (pBest) in history
        set current value as the new pBest
    end for
    Choose the particle with the best fitness value of all the particles as the gBest
    for each particle
        Update C1, C2 and w using the fuzzy inference system
        Update particle velocity
        Update particle position
    end
end while
```

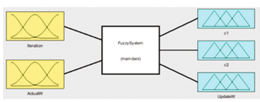


Fig. 2. Pseudocode of PSO with fuzzy dynamic parameters adaptation

Table 2. Initials parameters of the PSO

Parameter	Number
Particles	10
Maximum iterations	30
C_1	2
C_2	2
w	0.8

The proposed fuzzy inference system to adjust parameters is shown in Fig. 3, this fuzzy inference system has 2 inputs and 3 outputs. As inputs, iteration and actual w are represented and as outputs C_1, C_2 and w are represented to update their values. The range of each variable of the proposed fuzzy inference system are shown in Table 3, for each variable 3 triangular membership functions are used with 9 fuzzy rules.

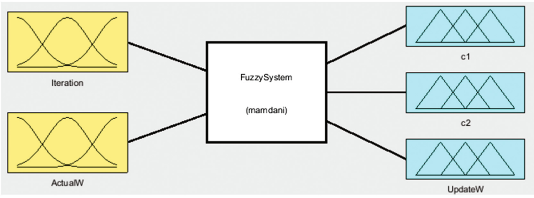


Fig. 3. Fuzzy inference system for the PSO with dynamic parameters adaptation

Table 3. Range of variables for the fuzzy inference system

Variable	Range
Iteration	1 to 10
Actual W	0.2 to 1.2
C_1	1 to 2
C_2	1 to 2
Update W	0.2 to 1.2

3.2 Iris Database

The database of human iris from the Institute of Automation of the Chinese Academy of Sciences was used [4]. Each person has 14 images (7 for each eye). For this work, only 77 persons were used. The image dimensions are 320 × 280, JPEG format. Figure 4 shows examples of the human iris images from CASIA database.

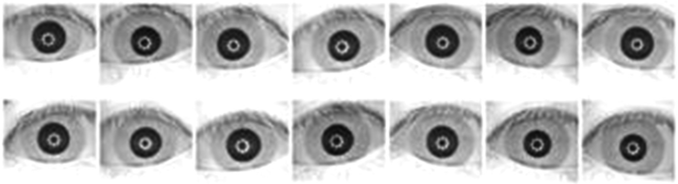


Fig. 4. Examples of the human iris images from CASIA database

4 Experimental Results

To compare with the proposed optimization, non-optimized trainings (MGNN architecture is randomly established) were also performed. 30 non-optimized trainings and 20 evolutions were performed. The achieve results are shown in this section.

4.1 Non-optimized Results

The 5 best non-optimized trainings are shown in Table 4. In training #25, the best results is obtained, where using 69% of data for the training phase a 98.05 of recognition rate is obtained. In Table 5, a summary of the results is shown.

Table 4. The first 5 results (Non-optimized)

Training	Images for training	Persons per module	Rec. Rate/Error
8	41% (3, 7, 8, 9, 12 and 14)	Module#1(1 to 11) Module #2(12 to 21) Module #3(22 to 25 Module #4(26 to 30) Module #5(31 to 40) Module #6(41 to 49) Module #7(50 to 61) Module #8(62 to 69) Module #9(70 to 75) Module #10(76 to 77)	93.51% (576/616)
14	42% (2, 4, 5, 6, 13 and 14)	Module #1(1 to 3) Module #2(4 to 19) Module #3(20 to 29) Module #4(30 to 34) Module #5(35 to 41) Module #6(42 to 54) Module #7(55 to 70) Module #8(71 to 77)	96.27% (593/616)

(*continued*)

Table 4. (*continued*)

Training	Images for training	Persons per module	Rec. Rate/Error
15	58% (1, 3, 5, 6, 8, 9, 13 and 14)	Module #1(1 to 7) Module #2(8 to 13) Module #3(14 to 30) Module #4(31 to 37) Module #5(38 to 59) Module #6(60 to 71) Module #7(72 to 77)	94.59% (437/462)
25	69% (1, 2, 3, 4, 5, 6, 7, 11, 13 and 14)	Module #1(1 to 6) Module #2(7 to 22) Module #3(23 to 31) Module #4(32 to 48) Module #5(49 to 53) Module #6(54 to 64) Module #7(65 to 66) Module #8(67 to 77)	98.05% (302/308)
30	64% (1, 3, 5, 7, 8, 9, 10, 11 and 14)	Module #1(1 to 15) Module #2(16 to 54) Module #3(55 to 77)	92.99% (358/385)

Table 5. Summary of non-optimized results

	Recognition Rate
Best	98.05%
Average	83.78%
Worst	43.56%

4.2 Optimized Results

The optimized results using the particle swarm optimization with the fuzzy dynamic parameters adaption are shown and compared in this section. 20 evolutions were performed. The 5 best results obtained by the proposed method are shown in Table 6. In evolutions #14, #15 and #19, the best results are obtained, where using 75% of data for the training phase a 99.57 of recognition rate is obtained. In Fig. 5, the convergence of evolution #14 is shown. In Table 7, a summary of the optimized results is shown. It can be observed that the best, the average and the worst recognition rate are better when the proposed optimization is used.

Table 6. The first 5 results (Optimized results, PSO)

Ev.	Images for training	Number of neurons per hidden layer	Persons per module	Rec. Rate/Error
2	79% (1, 2, 3, 4, 5, 6, 8, 11, 12, 13 and 14)	70,106,172,120,177 100,158,170,148,39 192,113,97,100,118 124,230,136,148,241 204,67,53,213,196 109,101,220,124,77 26,198,130,205,156 88,241,166,37,66 177,146,57,110,166 126,77,184,47,209	Module #1(1 to 15) Module #2(16 to 18) Module #3(19 to 21) Module #4(22 to 24) Module #5(25 to 33) Module #6(34 to 38) Module #7(39 to 53) Module #8(54 to 65) Module #9(66 to 71) Module #10 (72 to 77)	99.13% (0.0087)
6	76% (2, 3, 4, 5, 6, 7, 8, 9, 11, 13 and 14)	91,116,204,192,80 189,115,49,75,136 119,93,122,141,191 180,88,197,124,191 126,188,245,241,141	Module #1(1 to 21) Module #2(22 to 36) Module #3(37 to 55) Module #4(56 to 73) Module #5(74 to 77)	99.13% (0.0087)
14	75% (1, 3, 5, 6, 7, 8, 9, 10, 11 and 14)	231,136,170,77,66 34,130,108,141,25 115,64,43,139,102 42,38,101,86,97 194,149,72,108,81 102,22,148,214,51 212,78,60,216,38 168,138,192,99,231 132,173,129,112,170	Module #1(1 to 15) Module #2(16 to 21) Module #3(22 to 29) Module #4(30 to 31) Module #5(32 to 42) Module #6(43 to 49) Module #7(50 to 52)	99.57% (0.0043)

(*continued*)

Table 6. (*continued*)

Ev.	Images for training	Number of neurons per hidden layer	Persons per module	Rec. Rate/Error
			Module #8(53 to 64) Module #9(65 to 77)	
15	75% (1, 2, 3, 5, 6, 8, 10, 11, 12, 13 and 14)	218,211,140,51,168 108,62,23,166,111 210,100,30,21,27 146,82,49,130,159 89,243,84,193 33,69,146,57,89 174,43,107,201,157	Module #1(1 to 15) Module #2(16 to 17) Module #3(18 to 31) Module #4(32 to 43) Module #5(44 to 55) Module #6(56 to 71) Module #7(72 to 77)	99.57% (0.0043)
19	75% (1, 2, 3, 4, 5, 6, 7, 8, 10, 11, and 14)	132,76,168,42,183 183,217,196,132,192 78,167,36,67,89 183,148,43,111,214 66,181,167,110,29 170,79,141,150,71 215,125,148,128,225 209,57,37,151,61 82,148,183,59 72,201,178,215,211	Module #1(1 to 15) Module #2(16 to 20) Module #3(21 to 32) Module #4(33 to 37) Module #5(38 to 41) Module #6(42 to 48) Module #7(49 to 63) Module #8(64 to 69) Module #9(70 to 74) Module #10 (75 to 77)	99.57% (0.0043)

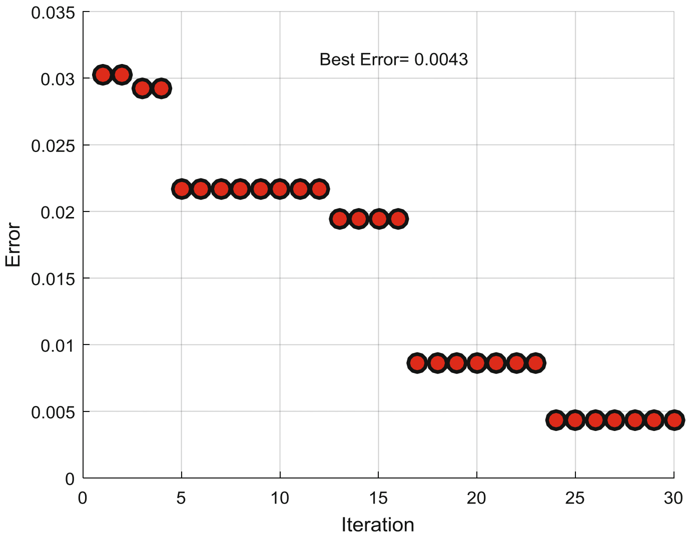


Fig. 5. Convergence of evolution #14 (PSO)

Table 7. Comparison of optimized results

	Recognition rate
Best	99.57%
Average	98.91%
Worst	98.27%

5 Conclusions

In this paper, a particle swarm optimization with a fuzzy dynamic parameter adaptation was proposed. This parameter adaptation arises with the purpose of establishing the most important parameters of the algorithm depending on the current behavior of the particles to avoid stagnation in the result (best solution) during the iterations. The optimization technique seeks to design MGNNs architectures and parameters such as number of sub modules, percentage of information for the training phase, number of hidden layers (with their respective number of neurons) for each sub module and learning algorithm. Non-optimized trainings were performed where their MGNN architecture was randomly established to compare with the results achieved by the proposed optimization and visibly the results obtained by the optimization overcome the non-optimized results. As future work a comparison with a simple PSO (without fuzzy dynamic parameter adaptation) will be performed, and other designs of fuzzy inference systems for the parameters adaptation will be proposed and compared.

References

1. Auda, G., Kamel, M.: Modular neural networks: a survey. Int. J. Neural Syst. **9**(2), 129–151 (1999)
2. Bargiela, A., Pedrycz, W.: The roots of granular computing. In: IEEE International Conference on Granular Computing (GrC), pp. 806–809 (2006)
3. Castillo, O., Melin, P.: Soft Computing for Control of Non-Linear Dynamical Systems. Springer, Heidelberg (2001)
4. Database of Human Iris. Institute of Automation of Chinese Academy of Sciences (CASIA). http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp. Accessed 12 Nov 2015
5. Dorigo, M., Gambardella, L.M.: Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans. Evol. Comput. **1**(1), 53–66 (1997)
6. Geem, Z.W., Yang, X.S., Tseng, C.L.: Harmony search and nature-inspired algorithms for engineering optimization. J. Appl. Math. **2013**, 438158:1–438158:2 (2013)
7. Hassoun, M.: Fundamentals of Artificial Neural Networks. A Bradford Book, Cambridge (2003)
8. Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975)
9. Jamal, A.: Granular computing. Int. J. Res. Cloud Eng. **2**(3), 29–40 (2015)
10. Jang, J., Sun, C., Mizutani, E.: Neuro-Fuzzy and Soft Computing. Prentice Hall, New Jersey (1997)
11. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of the IEEE International Joint Conference on Neuronal Networks. IEEE Press, pp. 1942–1948 (1995)
12. Khan, A., Bandopadhyaya, T., Sharma, S.: classification of stocks using self organizing map. Int. J. Soft Comput. Appl. **4**, 19–24 (2009)
13. Lucic, P., Teodorovic, D.: Bee system: modeling combinatorial optimization transportation engineering problems by swarm intelligence. In: Preprints of the TRISTAN IV Triennial Symposium on Transportation Analysis, pp. 441–445 (2001)
14. Melin, P., Castillo, O.: Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing: An Evolutionary Approach for Neural Networks and Fuzzy Systems, 1st edn., pp. 119–122. Springer (2005)
15. Okamura, M., Kikuchi, H., Yager, R., Nakanishi, S.: Character diagnosis of fuzzy systems by genetic algorithm and fuzzy inference. In: Proceedings of the Vietnam-Japan Bilateral Symposium on Fuzzy Systems and Applications, Halong Bay, Vietnam, pp. 468–473 (1998)
16. Qian, Y., Zhang, H., Li, F., Hu, Q., Liang, J.: Set-based granular computing: A lattice model. Int. J. Approx. Reasoning **55**, 834–852 (2014)
17. Rajabioun, R.: Cuckoo optimization algorithm. Appl. Soft Comput. J. **11**, 5508–5518 (2011)
18. Rashedi, E., Nezamabadi-Pour, H., Saryazdi, S.: GSA: a gravitational search algorithm. Inf. Sci. **179**, 2232–2248 (2009)
19. Rini, D.P., Shamsuddin, S.M., Yuhaniz, S.S.: Particle swarm optimization: technique, system and challenges. Int. J. Comput. Appl. **14**(1), 19–27 (2011)
20. Sánchez, D., Melin, P.: Hierarchical genetic algorithms for type-2 fuzzy system optimization applied to pattern recognition and fuzzy control. In: Recent Advances on Hybrid Approaches for Designing Intelligent Systems, pp. 19–35 (2014)
21. Saravanan, K., Sasithra, S.: Review on classification based on artificial neural networks. Int. J. Ambient Syst. Appl. (IJASA) **2**(4), 11–18 (2014)
22. Witten, I., Frank, E., Hall, E.: Fuzzy Logic for the Management of Uncertainty. Morgan Kaufmann, San Mateo (2011)

23. Yao, Y.Y.: On modeling data mining with granular computing. In: 25th International Computer Software and Applications Conference (COMPSAC), pp. 638–649 (2001)
24. Yao, Y.: Perspectives of granular computing. In: IEEE International Conference on Granular Computing (GrC), pp. 85–90 (2005)
25. Zadeh, L., Kacprzyk, J.: Fuzzy Logic for the Management of Uncertainty. Wiley-Interscience, New York (1992)

A Systematic Customer Oriented Approach based on Hesitant Fuzzy AHP for Performance Assessments of Service Departments

Ozlem Senvar(✉)

Department of Industrial Engineering, Marmara University, Istanbul, Turkey
ozlemsenvar@gmail.com

Abstract. Customer orientation is a business strategy in the lean business model that requires management and employees to focus on the changing demands and requirements of the customers.Improved business performance can be enhanced by .customer orientation. In this chapter, a systematic approach based on hesitant fuzzy AHP is proposed to deal with incomplete information due to the ambiguity to solve complex customer oriented multi criteria decision making problem of performance assessments of service departments.

Keywords: Customer orientation · Hesitant Fuzzy AHP · Hesitant Fuzzy Sets

1 Introduction

Customer orientation was defined as a state-like variable that measure an employee's attitude toward a customer and customer needs [1]. It has been emphasized that achieving customer orientation requires employees of a firm to understand which departments are most important to their customers [2, 3]. It is possible to define performance measurement as an essential element of effective planning, effective control and decision making.

It can provide feedback information for revealing progress, enhancing motivation and communication and diagnosing problems [4, 5]. From this standpoint, customer oriented performance measurement can be stated as a multi criteria decision making (MCDM) problem. Because of the imprecision in assessing the relative importance of attributes and the performance ratings of alternatives with respect to attributes, fuzzy multiple attribute decision-making (FMADM) techniques have been developed. It is possible to say that imprecision arises from some reasons such as; unquantifiable information, incomplete information, unobtainable information, partial ignorance, etc. Classical multi-attribute decision making (MADM) or MCDM techniques cannot effectively handle problems with such ambiguous information [5]. Fuzzy set theory can capture this type of ambiguity. Hesitant fuzzy sets characterize fuzziness by setting out all the possible values while assigning the membership degree of the elements of a set [6]. Linguistic hesitant fuzzy sets (LHFSs) have been used to represent qualitative preferences of the decision makers (DMs) and reflect their hesitancy and inconsistency. LHFSs are suitable for managing situations where the DMs are hesitant and inconsistent and decision making information takes the form of qualitative variables. HFSs are efficient so

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_26

as to study imprecise, uncertain or incomplete information or knowledge [7]. Because of their flexibility and efficiency, HFSs have attracted a great deal of attention [8]. Analytic hierarchy process-hesitant group decision making (AHP-HGDM) is an extension of AHP-group decision making (AHP-GDM). In AHP-HGDM, hesitant judgments that each may include several possible values are used to indicate the original judgments provided by the DMs. Generally, we assume that the hesitant judgments cannot be aggregated or revised. It is possible to define them as hesitant judgments to describe the hesitancy experienced by the DMs in decision making [9].

Singh [10] defined the *marketing mix* as a set of controllable variables that the firms influence the buyers' responses. The *marketing mix* has 4 elements which are *price, place, product* and *promotion*. In the literature, there are studies such as [11–15] emphasized that the 4 Ps' model is able to adapt perfectly and to continue to be the dominant paradigm in the new contexts. In this study, we handled the elements of classical marketing mix 4 Ps approach as criteria.

Oztaysi et al. [16] developed a hesitant fuzzy AHP technique involving decision makers' linguistic evaluations aggregated by ordered weighted averaging (OWA) operator and applied this technique to a multi-criteria supplier selection problem. Kahraman et al. [6] proposed hesitant fuzzy AHP technique for the humanitarian logistics warehouse location selection problem. Tuysuz and Simsek [17] developed hesitant fuzzy sets-based AHP method for a cargo company in Turkey.

This study aims to propose an evaluation framework for customer oriented performance rankings of departments (sales, delivery, quality, maintenance) of an organization via hesitant fuzzy AHP. To the best of our knowledge, customer oriented MCDM has not been handled using hesitant fuzzy AHP, yet. Hesitant fuzzy AHP have some advantages: Firstly, it deals with incomplete information due to the ambiguity. It provides a systematic approach in order to solve complex problems [6]. Moreover, it is flexible. According to the conditions of the organizations, new criteria can be included or removed. Therefore, it is possible to say that this method is applied for the performance evaluation problems [17]. The rest of the study is organized as follows: Sect. 2 briefly explains hesitant fuzzy AHP. Section 3 presents the real case application study with the results. For further directions, final section provides conclusion, recommendations.

2 Hesitant Fuzzy AHP

In the proposed hesitant fuzzy AHP method, we first determine the main and sub-criteria and the hierarchy for the problem, then make a multi-criteria evaluation of the alternatives to illustrate how the proposed model is used to solve it. The steps of the hesitant fuzzy AHP technique are given [6]:

Step 1. Construction of pairwise comparison matrices for criteria, sub-criteria and alternatives and collection of customers' evaluations using linguistic terms are performed.

Step 2. The linguistic terms which are given in Table 2 are transformed into triangular fuzzy numbers [18] and trapezoidal fuzzy numbers. These linguistic scales, which are based on the AHP's classical 1–9 scale, are given with their triangular and

Table 1. Linguistic scale for hesitant fuzzy AHP

Linguistic terms	Order number	Abbreviations	Triangular fuzzy numbers	Trapezoidal fuzzy numbers
Absolutely high importance	10	AHI	(7,9,9)	(7,9,9,9)
Very high importance	9	VHI	(5,7,9)	(5,7,7,9)
Essentially high importance	8	ESHI	(3,5,7)	(3,5,5,7)
Weakly high importance	7	WHI	(1,3,5)	(1,3,3,5)
Equally high importance	6	EHI	(1,1,3)	(1,1,1,3)
Exactly equal	5	EE	(1,1,1)	(1,1,1,1)
Equally low importance	4	ELI	(0.33,1,1)	(0.33,1,1,1)
Weakly low importance	3	WLI	(0.2,0.33,1)	(0.2,0.33,0.33,1)
Essentially low importance	2	ESLI	(0.14,0.2,0.33)	(0.14,0.2,0.2,0.33)
Very low importance	1	VLI	(0.11,0.14,0.2)	(0.11,0.14,0.14,0.2)
Absolutely low importance	0	ALI	(0.11,0.11,0.14)	(0.11,0.11,0.11,0.14)

trapezoidal fuzzy number representations in Table 2. The scale in Table 1 is sorted from the lowest (s_0) to the highest (s_g) assuming the customers' evaluations are varied between two terms i.e. s_i and s_j such that $s_0 \leq s_i \leq s_j \leq s_g$.

Each element ($\tilde{a}_{ij}^k$) of the pairwise comparison matrix $\tilde{A}^k$ is a fuzzy number corresponding to its linguistic term. The pairwise comparison matrix is given by;

$$\tilde{A}^k = \begin{bmatrix} 1 & \tilde{a}_{12}^k & \cdots & \tilde{a}_{1n}^k \\ \tilde{a}_{21}^k & 1 & & \tilde{a}_{2n}^k \\ & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{a}_{n1}^k & \tilde{a}_{n2}^k & & 1 \end{bmatrix} \quad (1)$$

where ($\tilde{a}_{ij}^k$) represents the k^{th} customers' evaluation on comparison of i^{th} element to j^{th} element.

Step 3. Examination of the consistency of each fuzzy pairwise comparison is performed. For checking the consistency of the fuzzy pairwise comparison matrices, pairwise comparison values are defuzzified by the graded mean integration approach [19]. Assume that $\tilde{A} = [\tilde{a}_{ij}]$ is a fuzzy positive reciprocal matrix and $\mathrm{A} = [a_{ij}]$ is its defuzzified positive reciprocal matrix. If the result of the comparisons of $A = [a_{ij}]$ is consistent, then it can imply that the result of the comparisons of $\tilde{A} = [\tilde{a}_{ij}]$ is also consistent [20]. According to the graded mean integration approach, a triangular fuzzy number $\tilde{A} = (l, m, u)$ can be transformed into a crisp number by employing Eq. 2:

$$A = \frac{l + 4m + u}{6} \tag{2}$$

Step 4. Conflicts are identified and the evaluations are renewed. The evaluations of the customers are checked for their closeness to each other. If the evaluations are not close then customers are informed of the need to discuss the situation and renew their evaluations. Closeness is the difference between order numbers of evaluations. The difference cannot be more than 2.

Step 5. Fuzzy envelope approach, which was proposed by Liu and Rodriguez [21], is used to combine customers' evaluations.

The scale given in Table 2 is sorted from the lowest (s_0) to the highest (s_g). Assume the customers' evaluations vary between two terms i.e. s_i and s_j. Then $s_0 \leq s_i \leq s_j \leq s_g$.

The parameters a and d of the trapezoidal fuzzy membership function $\tilde{A}$ = (a, b, c, d) are computed using Eqs. 3 and 4. Notably, a_L^i refers to the minimum element of the i^{th} trapezoidal fuzzy set, a_M^i refers to the middle element of the i^{th} trapezoidal fuzzy set. Similarly, a_M^{i+1} is the middle element of the $(i+1)^{th}$ trapezoidal fuzzy set. a_M^j represents the middle element of the j^{th} trapezoidal fuzzy set and a_R^j is the maximum element of the j^{th} trapezoidal fuzzy set.

$$a = \min\{a_L^i, a_M^i, a_M^{i+1} \ldots, a_M^j, a_R^j\} = a_L^i \tag{3}$$

$$d = \max\{a_L^i, a_M^i, a_M^{i+1}, \ldots, a_M^j, a_R^j\} = a_R^j \tag{4}$$

$$b = \begin{cases} a_M^i, & if\ i+1 = j \\ OWA_{w^2}\left(a_M^i, \ldots a_M^{\frac{i+j}{2}}\right), & if\ i+j\ is\ even \\ OWA_{w^2}\left(a_M^i, \ldots a_M^{\frac{i+j-1}{2}}\right), & if\ i+j\ is\ odd \end{cases} \tag{5}$$

$$c = \begin{cases} a_M^{j+1}, & if\ i+1 = j \\ OWA_{w^1}\left(a_M^j, a_M^{j-1} \ldots a_M^{\frac{i+j}{2}}\right), & if\ i+j\ is\ even \\ OWA_{w^1}\left(a_M^j, a_M^{j-1} \ldots a_M^{\frac{i+j+1}{2}}\right), & if\ i+j\ is\ odd \end{cases} \tag{6}$$

OWA operation requires a weight vector. Filev and Yager [22] defined the first and second type of weights using α parameter which belong to the unit interval [0, 1]. The first kind of weights $W^1 = (w_1^1, w_2^1, \ldots w_n^1)$ is defined as; $w_1^1 = \alpha_2, w_2^1 = \alpha_2(1 - \alpha_2), \ldots w_n^1 = \alpha_2(1 - \alpha_2)^{n-2}$.

The definition of the second kind of weights $W^2 = \left(w_1^2, w_2^2, \ldots w_n^2\right)$ is $w_1^2 = \alpha_1^{n-1}, w_2^2 = (1-\alpha_2)\alpha_1^{n-2}, \ldots w_n^2 = 1 - \alpha_1$.

Here, α_1 is $\frac{g-(j-i)}{g-1}$ and α_2 is $\frac{(j-i)-1}{g-1}$, where g is the number of terms in the evaluation scale, j is the rank of highest evaluation and i is the rank of lowest evaluation value of the given interval.

Step 6. The collaborative pairwise comparison matrix is created.

$$\tilde{C} = \begin{bmatrix} 1 & \tilde{C}_{12} & \ldots & \tilde{C}_{1n} \\ \tilde{C}_{21} & 1 & & \tilde{C}_{2n} \\ \vdots & & \ddots & \vdots \\ \vdots & \vdots & \ldots & \vdots \\ \tilde{C}_{n1} & \tilde{C}_{n2} & & 1 \end{bmatrix} \tag{7}$$

where $\tilde{c}_{ij} = (c_{ij}, c_{ij_{m1}}, c_{ij_{m2}}, c_{ij_u})$.

Since the fuzzy envelopes, obtained in the previous step are trapezoidal fuzzy numbers, reciprocal values are computed with Eq. 8.

$$\tilde{c}_{ji} = \left(\frac{1}{c_{ij}}, \frac{1}{c_{ij_{m2}}}, \frac{1}{c_{ij_{m1}}}, \frac{1}{c_{ij_l}}\right) \tag{8}$$

Step 7. Fuzzy geometric mean for each row ($\tilde{r}_i$) of the collaborative pairwise matrix is calculated using Eq. 9.

$$\tilde{r}_i = [\tilde{c}_{i1} \otimes \ldots \otimes \tilde{c}_{in}]^{\frac{1}{n}} \tag{9}$$

Step 8. The fuzzy weight ($\tilde{w}_i$) of each criteria (or alternative) is calculated using Eq. 10.

$$\tilde{w}_i = \tilde{r}_i \otimes [\tilde{r}_1 \oplus \tilde{r}_2 \oplus \ldots \oplus \tilde{r}_n]^{-1} \tag{10}$$

Step 9. The fuzzy performance scores of each alternative are computed. To this end, steps 1–7 are repeated for each pairwise comparison matrix formed according to the predetermined decision model. The final fuzzy score of each alternative is calculated using Eq. 11.

$$\tilde{S}_i = \sum\nolimits_{j=1}^{n} \tilde{w}_j \tilde{s}_j, \forall \mathrm{i} \tag{11}$$

where $\tilde{S}_i$ is the fuzzy performance score of alternative i; $\tilde{w}_j$ is the weight of the criteria j, and $\tilde{s}_j$ is the performance score of alternative i with respect to criteria j.

Step 10. The trapezoidal fuzzy numbers are defuzzified for determining the importance ranking of the alternatives with Eq. 12 [23].

$$\mathrm{D} = \frac{c_l + 2c_{m1} + 2c_{m2} + c_u}{6} \tag{12}$$

Step 11. The alternatives are ranked according to the defuzzified values and the alternative with the greatest score is chosen.

3 Application

This study is a real case study for a multinational company's Turkey branch. The considered company supplies chemical products, hand and power tools, and etc. Most of its customers are from automotive and construction sector. The motivation of the company is to make good decisions for finding the convenient solutions for its customers. In this study, a telephone survey was conducted among 238 clients that are categorized from Marmara (E1), Aegean and Mediterranean (E2), Blacksea (E3), Middle Anatolia, East Anatolia and South-East Anatolia (E4) regions of Turkey. Data obtained from the survey are used for evaluating and ranking performances of departments. Figure 1 illustrates the hierarchical structure of the problem. In this structure, the components of the classical marketing mix 4 Ps are determined as the criteria and the sub-criteria are the elements of each component. Maintenance, delivery, quality and sales are the departments that are determined as alternatives. The main goal of this study is to evaluate the performance of each department.

In this study, we handled the criteria, sub-criteria and alternatives which are given in Table 2. The compromised pairwise comparisons of main criteria are presented in Table 3. Using the OWA operations, HFLTSs are aggregated into trapezoidal fuzzy sets as in Table 4. Table 5 presents geometric means, normalized weights and the defuzzified weights of the main criteria. Tables 6, 7, 8 and 9 present the pairwise comparison matrices of the sub-criteria with respect to the main criteria price, place, product and promotion respectively. From Table 6, the trapezoidal fuzzy weights of the sub-criteria with respect to price, are obtained as (0.084,0.092,0.113,0.226), (0.508,0.703,0.897,1) and (0.058,0.096,0.113,0.156), respectively. From Table 7, the trapezoidal fuzzy weights of the sub criteria with respect to place, are obtained as

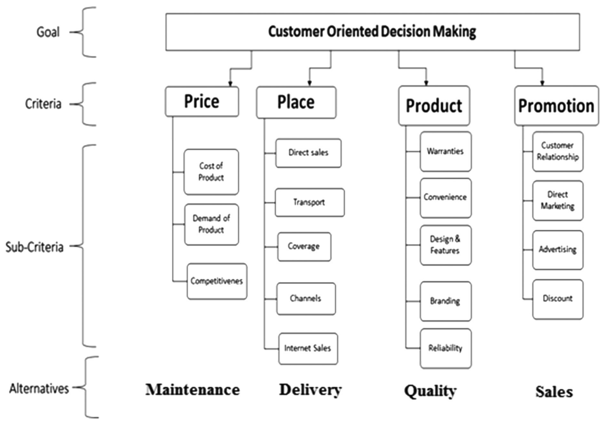


Fig. 1. Hierarchical structure of the problem

Table 2. Criteria, sub-criteria and alternatives

Elements of the problem	Abbreviations
Expert 1 from Marmara Region	E1
Expert 2 from Aegean and Mediterranean Regions	E2
Expert 3 from Black sea Region	E3
Expert 4 from Middle Anatolia, East Anatolia and South-East Anatolia	E4
Price	C1
Place	C2
Product	C3
Promotion	C4
Cost of product	C11
Demand of product	C12
Competitiveness	C13
Direct sales	C21
Transport	C22
Coverage	C23
Channels	C24
Internet sales	C25
Warranties	C31
Convenience	C32
Design & features	C33
Branding	C34
Reliability	C35
Customer relationship	C41
Direct marketing	C42
Advertising	C43
Discount	C44
Maintenance	A1
Delivery	A2
Quality	A3
Sales	A4

Table 3. Pairwise comparison of main criteria using HFLTS

Comparison of main criteria w.r.t. goal	C1	C2	C3	C4
C1	EE	Between WLI and EE	Between ELI and EHI	Between EHI and ESHI
C2	—	EE	Between EHI and ESHI	Between EHI and ESHI
C3	—	—	EE	Between EHI and WHI
C4	—	—	—	EE

Table 4. Aggregated HFLTS scores

Comparison of main criteria w.r.t. goal	C1	C2	C3	C4
C1	(1,1,1,1)	(0.2,0,926,1.074,1)	(0.333,1,1,3)	(1,2.778,3.222,7)
C2	(1,0.931,1.08,5)	(1,1,1,1)	(1,2.778,3.222,7)	(1,2.778,3.222,7)
C3	(0.333,1,1,1)	(0.143,0.31,0.36,1)	(1,1,1,1)	(1,1,3,5)
C4	(0.143,0.31,0.36,1)	(0.143,0.31,0.36,1)	(0.2,0.333,1,1)	(1,1,1,1)

Table 5. Geometric means, normalized weights and defuzzified weights of main criteria

	Geometric means	Normalized weights	Defuzzified weights
C1	(0.508,1.266,1.364,2.141)	(0.056,0.263,0.335,0.961)	0.369
C2	(1,1.637,1.83,3.956)	(0.110,0.340,0.449,1)	0.448
C3	(0.467,0.746,1.019,1.968)	(0.052,0.156,0.250,0.883)	0.291
C4	(0.253,0.423,0.6,1)	(0.028,0.088,0.147,0.449)	0.158

Table 6. Pairwise comparison of sub-criteria using HFLTS w.r.t. price

Comparison of sub-criteria w.r.t. price	C11	C12	C13s
C11	EE	Between ALI and VLI	Between EE and EHI
C12	—	EE	Between VHI and AHI
C13	—	—	EE

Table 7. Pairwise comparison of sub-criteria using HFLTS w.r.t. place

Comparison of sub-criteria w.r.t. place	C21	C22	C23	C24	C25
C21	EE	Between WLI and ELI	Between EHI and ESHI	WHI	ESLI
C22	—	EE	Between ESHI and VHI	Between WHI and ESHI	Between ESLI and ELI
C23	—	—	EE	Between WLI and ELI	Between ESLI and WLI
C24	—	—	—	EE	Between ELI and EHI
C25	—	—	—	—	EE

Table 8. Pairwise comparison of sub-criteria using HFLTS w.r.t. product

Comparison of sub-criteria w.r.t. product	C31	C32	C33	C34	C35
C31	EE	EHI	Between EHI and WHI	Between EHI and ESHI	WHI
C32	—	EE	Between EE and EHI	Between EHI and WHI	Between WHI and ESHI
C33	—	—	EE	Between EE and WHI	Between WHI and ESHI
C34	—	—	—	EE	Between EE and WHI
C35	—	—	—	—	EE

Table 9. Pairwise comparison of sub-criteria using HFLTS w.r.t. promotion

Comparison of sub-criteria w.r.t. promotion	C41	C42	C43	C44
C41	EE	Between WLI and ELI	Between ELI and EHI	ESHI
C42	—	EE	Between WHI and ESHI	ESHI
C43	—	—	EE	Between EHI and WHI
C44	—	—	—	EE

Table 10. Pairwise comparison of alternatives using HFLTS w.r.t. cost of product

Comparison of sub-criteria w. r.t. cost of product	A1	A2	A3	A4
A1	EE	Between VLI and WLI	Between ESLI and WLI	Between ESLI and WLI
A2	—	EE	Between ALI and VLI	Between WLI and ELI
A3	—	—	EE	Between EE and WHI
A4	—	—	—	EE

(0.046,0.14,0.264,0.669), (0.09,0.213,0.479,1), (0.016,0.038,0.078,0.209), (0.028, 0.055,0.114,0.412) and (0.067,0.278,0.446,1), respectively. From Table 8, the trapezoidal fuzzy weights of the sub criteria with respect to product, are obtained as (0.115,0.244,0.446,1), (0.088,0.189,0.333,0.726), (0.059,0.144,0.253,0.551), (0.032,0.107,0.196,0.369) and (0.02,0.064,0.111,0.369), respectively. From Table 9, the trapezoidal fuzzy weights of the sub criteria with respect to promotion, are obtained as (0.158,0. 413,0.546,1), (0.053,0.124,0.198,0.543), (0.091,0.246,0.332,0.778), and

Table 11. Trapezoidal fuzzy (global) weights of sub-criteria

Sub-criteria	Weight
C11	(0.005,0.024,0.038,0.217)
C12	(0.028,0.179,0.314,0.961)
C13	(0.003,0.024,0.038,0.15)
C21	(0.005,0.048,0.119,0.669)
C22	(0.01,0.073,0.215,1)
C23	(0.002,0.013,0.035,0.209)
C24	(0.003,0.019,0.051,0.412)
C25	(0.008,0.095,0.2,1)
C31	(0.006,0.038,0.112,0.883)
C32	(0.005,0.029,0.083,0.641)
C33	(0.003,0.022,0.063,0.487)
C34	(0.002,0.017,0.049,0.326)
C35	(0.001,0.01,0.028,0.326)
C41	(0.004,0.036,0.08,0.449)
C42	(0.001,0.011,0.029,0.244)
C43	(0.003,0.022,0.049,0.349)
C44	(0.001,0.005,0.017,0.124)

Table 12. Weights of alternatives

Sub-criteria	A1	A2	A3	A4
C11	0.015	0.018	0.047	0.039
C12	0.041	0.176	0.150	0.215
C13	0.003	0.023	0.023	0.027
C21	0.04	0.085	0.132	0.137
C22	0.026	0.237	0.102	0.065
C23	0.039	0.042	0.034	0.018
C24	0.073	0.078	0.066	0.034
C25	0.191	0.208	0.169	0.087
C31	0.157	0.168	0.141	0.073
C32	0.115	0.123	0.103	0.053
C33	0.087	0.093	0.078	0.04
C34	0.059	0.064	0.053	0.027
C35	0.056	0.06	0.051	0.026
C41	0.05	0.093	0.018	0.098
C42	0.025	0.026	0.037	0.047
C43	0.065	0.039	0.02	0.071
C44	0.022	0.024	0.02	0.01
Normalized total score	0.216	0.316	0.252	0.217

(0.029,0.058,0.113,0.276), respectively. The next step is to obtain the pairwise comparison matrices of alternatives with respect to each sub-criterion. In our case, there are 17 matrices of such comparisons. Due to the space limitations we only present one of them. Table 10 gives the pairwise comparison of alternatives using HFLTS with respect to cost of product. Table 11 indicates the trapezoidal fuzzy weights of the sub-criteria. Table 12 shows the defuzzified weights of sub-criteria and the final weights of the alternatives with respect to sub-criteria. The alternative with the highest normalized total score is the best alternative among the others. In Table 12, final performance weights of the departments are given. It is possible to see from the results of the study that according to the customers' perspective, delivery department (A2) exhibits the best performance according and they are satisfied with delivery processes of the company. Consecutively, quality department (A3) has second place, sales department (A4) has third place in the performance rankings. Finally, depending on customers' perspective, maintenance department (A1) has the worst performance.

4 Conclusion

Hesitant linguistic term sets provided the flexibility to reveal comparative linguistic expressions by using context-free grammar. Hence, dealing with incomplete information due to the vagueness of the criteria becomes easier. For further directions, this systematic approach can be utilized to solve other decision making problems with the same characteristics. Moreover, extensions of fuzzy sets such as intuitionistic fuzzy sets, interval type-2 fuzzy sets, multi fuzzy sets, etc. can be handled in the considered framework that we have established and can be compared with our results for obtaining more managerial insights.

References

1. Grizzle, J.-W., Zablah, A.-R., Brown, T.-J., Mowen, J.-C., Lee, J.-M.: Employee customer orientation in context: how the environment moderates the influence of customer orientation on performance outcomes. J. Appl. Psychol. **94**(5), 1227–1242 (2009). doi:10.1037/a0016404
2. Anderson, S., Pearo, L.-K., Widener, S.-K.: Drivers of service satisfaction: linking customer satisfaction to the service concept and customer characteristics. J. Serv. Res. **10**(4), 365–381 (2008). doi:10.1177/1094670508314575
3. Mathe, K., Halsell, S.-S., Roseman, M.: The role of customer orientation in the relationship between manager communications and customer satisfaction. J. Hosp. Tour. Res. **40**(2), 198–209 (2016). doi:10.1177/1096348013496278
4. Waggoner, D.-B., Neely, A.-D., Kennerley, M.-P.: The forces that shape organisational performance measurement systems: an interdisciplinary review. Int. J. Prod. Econ. **60**, 53–60 (1999). doi:10.1016/S0925-5273(98)00201-1
5. Mohaghar, A., Fathi, M.-R., Zarchi, M.-K., Omidian, A.: A combined VIKOR-fuzzy AHP approach to marketing strategy selection. Bus. Manag. Strategy **3**(1), 13–27 (2012). doi:10.5296/bms.v3i1.957

6. Kahraman, C., Bolturk, E., Onar, S.-C., Oztaysi, B., Goztepe, K.: Multi-attribute warehouse location selection in humanitarian logistics using hesitant fuzzy AHP. Int. J. Anal. Hierarchy Process **8**(2), 271–298 (2016). doi:10.13033/ijahp.v8i2.387. ISSN 1936-6744
7. Yu, D.-J., Li, D.-F.: Dual hesitant fuzzy multi-criteria decision making and its application to teaching quality assessment. J. Intell. Fuzzy Syst. **27**(4), 1679–1688 (2014). doi:10.3233/IFS-141134
8. Zhou, H., Wang, J.-Q., Zhang, H.-Y., Chen, X.-H.: Linguistic hesitant fuzzy multi-criteria decision-making method based on evidential reasoning. Int. J. Syst. Sci. **47**(2), 314–327 (2016). doi:10.1080/00207721.2015.1042089
9. Zhu, B., Xu, Z.: Analytic hierarchy process-hesitant group decision making. Eur. J. Oper. Res. **239**(3), 794–801 (2014). doi:10.1016/j.ejor.2014.06.019
10. Singh, M.: Marketing mix of 4P's for competitive advantage. J. Bus. Manag. **3**(6), 40–45 (2012). http://www.iosrjournals.org
11. Aldridge, A., Forcht, K., Pierson, J.: Get linked or get lost: marketing strategy for the internet. Internet Res. **7**(3), 161–169 (1997). doi:10.1108/10662249710171805
12. Bhatt, G.-D., Emdad, A.-F.: An analysis of the virtual value chain in electronic commerce. Logist. Inf. Manag. **14**(1/2), 78–85 (2001). doi:10.1108/09576050110362465
13. Allen, E., Fjermestad, J.: E-commerce marketing strategies: an integrated framework and case analysis. Logist. Inf. Manag. **14**(1/2), 14–23 (2001). doi:10.1108/09576050110360106
14. Moller, K.E.: Comment on: the marketing mix revisited: towards the 21st century marketing by E. Constantinides. J. Mark. Manag. **22**(3), 439–450 (2006). doi:10.1362/026725706776861181
15. Dominici, G.: From marketing mix to e-marketing mix: a literature overview and classification. Int. J. Bus. Manag. **4**(9), 17–24 (2009). http://www.ccsenet.org/journal.html
16. Oztaysi, B., Onar, S.-C., Bolturk E., Kahraman, C.: Hesitant fuzzy analytic hierarchy process. In: 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–7. IEEE (2015). doi:10.1109/FUZZ-IEEE.2015.7337948
17. Tuysuz, F., Simsek, B.: A hesitant fuzzy linguistic term sets-based AHP approach for analyzing the performance evaluation factors: an application to cargo sector. Complex Intell. Syst. **1**(1), 1–9 (2017). doi:10.1007/s40747-017-0044-x
18. Tan, R.-R., Aviso, K.-B., Huelgas, A.-P., Promentilla, M.-A.-B.: Fuzzy AHP approach to selection problems in process engineering involving quantitative and qualitative aspects. Process Saf. Environ. Prot. **92**(5), 467–475 (2014). doi:10.1016/j.psep.2013.11.005
19. Hsieh, C.-H., Chen, S.-H.: A model and algorithm of fuzzy product positioning. Inf. Sci. **121**(1), 61–82 (1999). doi:10.1016/S0020-0255(99)00050-X
20. Buckley, J.-J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985). doi:10.1016/0165-0114(85)90090-9
21. Liu, H., Rodríguez, R.-M.: A fuzzy envelope for hesitant fuzzy linguistic term set and its application to multicriteria decision making. Inf. Sci. **258**, 220–238 (2014). doi:10.1016/j.ins.2013.07.027
22. Filev, D., Yager, R.-R.: On the issue of obtaining OWA operator weights. Fuzzy Sets Syst. **94**(2), 157–169 (1998). doi:10.1016/S0165-0114(96)00254-0
23. Sahoo, N.-K., Mohanty, B.-S., Tripathy, P.-K.: Fuzzy inventory model with exponential demand and time-varying deterioration. Glob. J. Pure Appl. Math. **12**(3), 2573–2589 (2016). http://www.ripublication.com/gjpam.htm

Edge Detection Based on Ordered Directionally Monotone Functions

Mikel Sesma-Sara[1,2(✉)], Humberto Bustince[1,2], Edurne Barrenechea[1,2], Julio Lafuente[1], Anna Kolsesárová[3], and Radko Mesiar[4]

[1] Universidad Pública de Navarra, Campus Arrosadia s/n, 31006 Pamplona, Spain
{mikel.sesma,bustince,edurne.barrenechea,lafuente}@unavarra.es
[2] Institute of Smart Cities, Pamplona, Spain
[3] Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, Bratislava, Slovakia
anna.kolesarova@stuba.sk
[4] Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia
mesiar@math.sk

Abstract. We present an image edge detection algorithm that is based on the concept of ordered directionally monotone functions, which permit our proposal to consider the direction of the edges at each pixel and perform accordingly. The results of this method are presented to the EUSFLAT 2017 Competition on Edge Detection.

Keywords: Edge detection · Image processing · Ordered directionally monotonicity · Ordered directionally monotone functions

1 Introduction

The task of detecting the edges of objects in an image is complicated as the notion of edge is not explicitly characterized. We understand an edge as a *big enough jump between the intensity of a pixel and those of its neighbours*.

In the literature one can find many methods to detect edges [7,9,12] and there exist different approaches to construct an edge detector. In our proposal we follow the scheme given by Bezdek et al. in [3], in which the process is composed by four phases: conditioning, feature extracting, blending and scaling. Specifically, the edge detector presented in this paper consists of the following steps:

(ED1) Smoothening the original image with a Gaussian filter;
(ED2) Obtaining the feature image using ordered directionally monotone functions;
(ED3) Thinning the feature image using the Non-Maximum Suppression (NMS) procedure;

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_27

(ED4) Binarizing the feature image applying hysteresis to obtain the black and white edge image.

Additionally, some edge detectors in the literature only consider the information of the intensity changes between a pixel and its surroundings (see [1,4,9]), whereas others also take into account the direction in which those intensities change for each pixel (see [7,12]). This is also the case of this proposal. The edge detector presented in this paper is based on ordered directionally monotone functions, which enable to fusion information taking into account the most influential direction for each point.

This paper is organized as follows. In Sect. 2 we recall some preliminary notions and in Sect. 3 we present the concept of ordered directionally monotone functions. In Sect. 4 we present the edge detection algorithm and we finish with some conclusions.

2 Preliminaries

Let $n \in \mathbb{N}$ with $n \geq 2$. We refer to points in $[0,1]^n$ as $\mathbf{x} = (x_1, \dots, x_n) \in [0,1]^n$ and denote $\mathbf{0} = (0, \dots, 0)$ and $\mathbf{1} = (1, \dots, 1)$. The order relation we consider on $[0,1]^n$ is the partial order, which is defined as follows: Let $\mathbf{x}, \mathbf{y} \in [0,1]^n$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for every $i \in \{1, \dots, n\}$.

Let $\mathcal{S}_n$ be the symmetric group of order n, i.e., the set of all permutation operators of the set $\{1, \dots, n\}$. Let $\mathbf{x} \in [0,1]^n$ and $\sigma \in \mathcal{S}_n$, we denote as $\mathbf{x}_\sigma = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

The edge detection method we present in this work makes use of Choquet integrals [8]. The Choquet integral is a generalization of the Lebesgue integral, where instead of additive measures, fuzzy measures are used. Let us recall the definition of fuzzy measure.

Definition 1. *Let $N = \{1, 2, \dots, n\}$. A function $\mathfrak{m} : 2^N \to [0,1]$ is a fuzzy measure if it satisfies the following properties:*

(i) *For all $X, Y \subseteq N$, if $X \subseteq Y$, then $\mathfrak{m}(X) \leq \mathfrak{m}(Y)$;*
(ii) $\mathfrak{m}(\emptyset) = 0$*;*
(iii) $\mathfrak{m}(N) = 1$.

Thus, we recall the definition of Choquet integral.

Definition 2 ([2,10])**.** *Let $\mathfrak{m}\colon 2^N \to [0,1]$ be a fuzzy measure. The discrete Choquet integral with respect to $\mathfrak{m}$ is defined as the function $C_\mathfrak{m}\colon [0,1]^n \to [0,1]$, given, for all $\mathbf{x} = (x_1, \dots, x_n) \in [0,1]^n$, by*

$$C_\mathfrak{m}(\mathbf{x}) = \sum_{i=1}^{n} \left(x_{(i)} - x_{(i-1)}\right) \cdot \mathfrak{m}\left(A_{(i)}\right), \qquad (1)$$

where $\left(x_{(1)}, \dots, x_{(n)}\right)$ is an increasing permutation of $\mathbf{x}$, i.e., $x_{(1)} \leq \dots \leq x_{(n)}$, with the convention that $x_{(0)} = 0$, and $A_{(i)} = \{(i), \dots, (n)\}$ is the subset of indices of $n - i + 1$ largest components of $\mathbf{x}$.

The concept of overlap index is also used in this work, let us recall its definition.

Definition 3. *Let U be a nonempty universe and denote $FS(U)$ the set of fuzzy sets over U. An overlap index is a function $O : FS(U) \times FS(U) \to [0,1]$ such that*

(O1) $O(A,B) = 0$ *if and only if A and B have disjoint supports;*
(O2) $O(A,B) = O(B,A)$*;*
(O3) If $B \subset C$, then $O(A,B) \leq O(A,C)$ for all $A \in FS(U)$.

Furthermore, we recall the notion of directional monotonicity [6], where monotonicity is studied with respect to one ray (one vector in $\mathbb{R}^n$).

Definition 4 ([6])**.** *Let $\overrightarrow{r} = (r_1, \ldots, r_n) \in \mathbb{R}^n$ with $\overrightarrow{r} \neq \overrightarrow{0}$. A function $F\colon [0,1]^n \to [0,1]$ is said to be $\overrightarrow{r}$-increasing if for all points $\mathbf{x} \in [0,1]^n$ and for all $c > 0$ such that $\mathbf{x} + c\overrightarrow{r} \in [0,1]^n$ it holds that*

$$F(x_1 + cr_1, \ldots, x_n + cr_n) \geq F(x_1, \ldots, x_n)\ .$$

Analogously, one can define the notion of $\overrightarrow{r}$-decreasing function.

3 Ordered Directionally Monotone Functions

Directional monotone functions are monotone along some fixed direction, but that direction does not vary on the point that is being considered. In the case of ordered directionally monotone functions, the direction along which monotonocity is demanded depends on each particular point that the function takes as an input. In this section we recall the definition and a result that provides conditions under which the Choquet integral is ordered directionally monotone (see [5]).

Definition 5. *Let $F\colon [0,1]^n \to [0,1]$ and let $\overrightarrow{r} \in \mathbb{R}^n$ with $\overrightarrow{r} \neq \mathbf{0}$. F is said to be ordered directionally (OD) $\overrightarrow{r}$-increasing if for each $\mathbf{x} \in [0,1]^n$, and any permutation $\sigma \in \mathcal{S}_n$ with $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(n)}$ and any $c > 0$ such that*

$$1 \geq x_{\sigma(1)} + cr_1 \geq \cdots \geq x_{\sigma(n)} + cr_n \geq 0$$

it holds that

$$F(\mathbf{x} + c\overrightarrow{r}_{\sigma^{-1}}) \geq F(\mathbf{x}), \tag{2}$$

where $\overrightarrow{r}_{\sigma^{-1}} = (r_{\sigma^{-1}(1)}, \ldots, r_{\sigma^{-1}(n)})$.
The concept of OD $\overrightarrow{r}$-decreasing function can be defined analogously.

Theorem 1. *Let $\mathfrak{m}\colon 2^N \to [0,1]$ be a fuzzy measure and let $\overrightarrow{r} = (r_1, \ldots, r_n)$ be a non-null real vector. Then the Choquet integral $C_\mathfrak{m}\colon [0,1]^n \to [0,1]$ is an OD $\overrightarrow{r}$-increasing function if and only if for each permutation $\tau \in \mathcal{S}_n$ it holds that*

$$\sum_{i=1}^{n} r_i \mathfrak{m}_\tau(i) \geq 0,$$

where $\mathfrak{m}_\tau(1) = \mathfrak{m}(\{\tau(n)\})$, and for each $i \in \{2, \ldots, n\}$, $\mathfrak{m}_\tau(i) = \mathfrak{m}(\{\tau(n-i+1), \ldots, \tau(n)\}) - \mathfrak{m}(\{\tau(n-i+2), \ldots, \tau(n)\})$.

4 Edge Detection Using OD Monotone Functions

In this section we present an edge detection method that is based on OD monotone functions.

From the four steps, (ED1)–(ED4), described in the Introduction, in this work we focus on step (ED2): obtaining the feature image. We make use of OD monotone functions to assign to each pixel of the image a value that represents the magnitude of the gradient vector at that position. This step is further explained in Sect. 4.1.

Regarding the remaining steps of our edge detector, in (ED1) a Gaussian filter with deviation $\sigma = 1$ is used. In (ED3) we use Kovesis' MATLAB implementation [11] of the Non-Maximum Suppression (NMS) algorithm, a thinning algorithm presented in [7]. In (ED4) we use the automatic hysteresis algorithm in [13] to obtain a black/white image.

4.1 Step (ED2): Obtaining the Feature Image

We work with grayscale images normalized to $[0, 1]$ and to construct the feature image we follow the following method:

(1) Consider a pixel a_{22} and its 3×3 neighbourhood as in Fig. 1. Then compute

$$\begin{aligned} d_1 &= |a_{22} - a_{11}|, d_2 = |a_{22} - a_{12}|, d_3 = |a_{22} - a_{13}|, d_4 = |a_{22} - a_{23}|, \\ d_5 &= |a_{22} - a_{33}|, d_6 = |a_{22} - a_{32}|, d_7 = |a_{22} - a_{31}|, d_8 = |a_{22} - a_{21}|, \end{aligned} \tag{3}$$

i.e., the absolute values of the differences of intensities from the central pixel a_{22} and its neighbours.

(2) Take $\sigma \in \mathcal{S}_n$ such that it sorts the values in (3) in a decreasing order, i.e.,

$$d_{\sigma_{(1)}} \geq d_{\sigma_{(2)}} \geq \ldots \geq d_{\sigma_{(8)}},$$

and set the vector $\overrightarrow{r} = (d_{\sigma_{(1)}}, \ldots, d_{\sigma_{(8)}})$.

(3) Apply Theorem 1 to compute the value of the pixel of the feature image corresponding to the position of a_{22}. This is done by using the values d_i for $1 \leq i \leq 8$ in (1).

(4) Carry on (1), (2) and (3) taking now the next pixel until the process is repeated for each pixel of the original image.

Our proposal takes into account the direction in which the differences increase since when we sort the values in a decreasing way, the largest differences become the ones that influence the most.

4.2 Fuzzy Measures

The Choquet integral in Theorem 1 that we use for constructing the feature image needs a fuzzy measure. In this section we describe how fuzzy measures can be built in terms of overlap indices (see [14]).

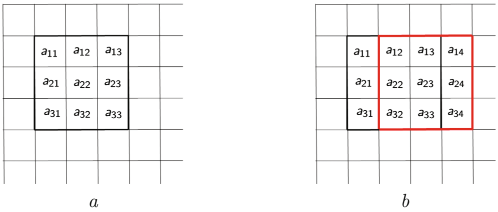


Fig. 1. *a:* 3×3 neighbourhood of pixel a_{22}. *b:* Next central pixel, a_{23}.

Let O be an overlap index and set the following fuzzy sets for each pixel:

$$\begin{aligned}
A_1 &= \{(A_1(1) = d_{\sigma_{(1)}}), (A_1(2) = 0), (A_1(3) = 0), \cdots, (A_1(8) = 0)\} \\
A_2 &= \{(A_2(1) = d_{\sigma_{(1)}}), (A_2(2) = d_{\sigma_{(2)}}), (A_2(3) = 0), \cdots, (A_2(8) = 0)\} \\
A_3 &= \{(A_3(1) = d_{\sigma_{(1)}}), (A_3(2) = d_{\sigma_{(2)}}), (A_3(3) = d_{\sigma_{(3)}}), \cdots, (A_3(8) = 0)\} \\
&\vdots \\
A_8 &= \{(A_8(1) = d_{\sigma_{(1)}}), (A_8(2) = d_{\sigma_{(2)}}), (A_8(3) = d_{\sigma_{(3)}}), \cdots, (A_8(8) = d_{\sigma_{(8)}})\},
\end{aligned} \tag{4}$$

Then, the function $\mathfrak{m} : 2^N \to [0, 1]$ given by

$$\mathfrak{m}(A_i) = \frac{O(A_i, A_8)}{O(A_8, A_8)} \text{ for } 1 \leq i \leq 8$$

is a fuzzy measure.

Note that the $\mathfrak{m}$ varies from pixel to pixel, depending upon the information that the neighbourhood of each pixel provides.

4.3 Edge Detection Algorithm

As it is stated above, the method presented in this paper is intended for grayscale images and the images in the ED17 competition dataset[1] are in color. Therefore, first we convert them to grayscale images computing the mean of their three color channels.

The reason to use $p \geq 1$ in Step 5 is that the values obtained in the previous step tend to be low and hence it is convenient to enlighten the result. For the results the used value is $p = 4$.

The overlap index used for the results submitted to the Edge Detection competition is given by:

$$O(A, B) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (A(u_i) B(u_i))^2}.$$

[1] http://irafm.osu.cz/edge2017/main.php.

Algorithm 1. Edge detection algorithm

Input: A grayscale image I and an overlap index, O.
Output: An edge image E in black and white

1: **for** each pixel (i, j) of I **do**
2: Consider its 3×3 neighbourhood (see Fig. 1).
3: Compute the absolute values of the 8 differences between $I(i, j)$ and its 8 surrounding pixels as in (3).
4: Proceed as in (ED2) (Sect. 4.1) with a fuzzy measure obtained as in Sect. 4.2 to get the pixel (i, j) of the feature image E.
5: Set $E(i, j) \leftarrow E(i, j)^{1/p}$, with $p \geq 1$, to enhance the feature image.
6: **end for**
7: Use the NMS algorithm to thin E with Kovesis' implementation [11].
8: Apply hysteresis to E to obtain a black and white image using the method in [13].

5 Conclusions

In this work we have presented an edge detection algorithm that is based on the novel concept of ordered directionally monotone functions and which takes into account the direction in which the intensity of the pixels varies at each point.

Acknowledgments. This work is supported by the research services of *Universidad* Publica de Navarra and by the project TIN2016-77356-P (AEI/FEDER, UE).

References

1. Barrenechea, E., Bustince, H., Lopez-Molina, C., De Baets, B.: Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images. IEEE Trans. Fuzzy Syst. **19**(5), 819–830 (2011)
2. Beliakov, G., Bustince, H., Calvo, T.: A Practical Guide to Averaging Functions. Studies in Fuzziness and Soft Computing. Springer (2016)
3. Bezdek, J., Chandrasekhar, R., Attikouzel, Y.: A geometric approach to edge detection. IEEE Trans. Fuzzy Syst. **6**, 52–75 (1998)
4. Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J.: Interval-valued fuzzy sets constructed from matrices: application to edge detection. Fuzzy Sets Syst. **160**(13), 1819–1840 (2009)
5. Bustince, H., Barrenechea, E., Sesma-Sara, M., Lafuente, J., Mesiar, R., Kolesárová, A.: Ordered directionally monotone functions. Justification and Application. Submitted to IEEE Transactions on Fuzzy Systems
6. Bustince, H., Fernandez, J., Kolesárová, A., Mesiar, R.: Directional monotonicity of fusion functions. Eur. J. Oper. Res. **244**(1), 300–308 (2015)
7. Canny, J.: Finding edges and lines in images. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA (1983)
8. Choquet, G.: Theory of capacities. Annales de l'institut Fourier **5**, 131–295 (1954)
9. Gonzalez-Hidalgo, M., Massanet, S., Mir, A., Ruiz-Aguilera, D.: On the choice of the pair conjunction-implication into the fuzzy morphological edge detector. IEEE Trans. Fuzzy Syst. **23**(4), 872–884 (2015)

10. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, Cambridge (2009)
11. Kovesi, P.D.: MATLAB and Octave functions for computer vision and image processing. http://www.peterkovesi.com/matlabfns/
12. Lopez-Molina, C., Bustince, H., Fernandez, J., Couto, P., De Baets, B.: A gravitational approach to edge detection based on triangular norms. Pattern Recogn. **43**, 3730–3741 (2010)
13. Medina-Carnicer, R., Munoz-Salinas, R., Yeguas-Bolivar, E., Diaz-Mas, L.: A novel method to look for the hysteresis thresholds for the Canny edge detector. Pattern Recogn. **44**(6), 1201–1211 (2011)
14. Paternain, D., Bustince, H., Pagola, M., Sussner, P., Kolesárová, A., Mesiar, R.: Capacities and overlap indexes with an application in fuzzy rule-based classification systems. Fuzzy Sets Syst. **305**, 70–94 (2016)

Adaptive Fuzzy Clustering of Multivariate Short Time Series with Unevenly Distributed Observations Based on Matrix Neuro-Fuzzy Self-organizing Network

Galina Setlak[1], Yevgeniy Bodyanskiy[2(✉)], Iryna Pliss[2], Olena Vynokurova[2], Dmytro Peleshko[3], and Illya Kobylin[2]

[1] Rzeszow University of Technology, 12 Al. Powstancow Warszawy, 35-959 Rzeszow, Poland
gsetlak@prz.edu.pl
[2] Kharkiv National University of Radio Electronics, 14 Nauky Avenue, Kharkiv 61166, Ukraine
{yevgeniy.bodyanskiy,iryna.pliss}@nure.ua, vynokurova@gmail.com
[3] University "IT Step Academy", 83a Zamarstynivska Street, Lviv 79019, Ukraine
dpeleshko@gmail.com

Abstract. In the paper the method of fuzzy clustering task for multivariate short time series with unevenly distributed observations is proposed. Proposed method allows to process the time series both in batch mode and sequential on-line mode. In the first case we can use the matrix modification of fuzzy C-means method, and in second case we can use the matrix modification of neuro-fuzzy network by T. Kohonen, which is learned using the rule "Winner takes more". Proposed fuzzy clustering algorithms are enough simple in computational implementation and can be used for solving of wide class of Big Data and Data Stream Mining problems. The effectiveness of proposed approach is confirmed by many experiments based on real data sets.

Keywords: Adaptive fuzzy clustering · Multivariate short time series · Unevenly distributed observations · Matrix neuro-fuzzy self-organizing network

1 Introduction

Clustering and segmentation task for time series has been well studied in the context of Data Mining [1–4] and at present there are many different algorithms for solving such tasks that are based on someone or other a priori assumptions.

However, there are situations when well-known and topical approaches for solving this task are inoperative in practical applications.

One of such tasks is fuzzy clustering for short time series with unevenly distributed in time observations [5]. The clustering task of incomplete time series with missed or nonpresented observations is sufficiently close to this problem [6].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_28

The distinction of these tasks is fact that objects of clustering are not separate observations. The object is the sample in total and the observations are recorded in unevenly instants of time, and generated clusters are overlapped in such way that each processed sample can belong to several classes [7,8]. At that it is supposed also, that all processed information is given in the form of a fixed data set, whose size is not changed in time.

The situation is essentially complicated if initial information is defined in the form of multivariate time series (i.e. two-dimensional fields of observations-features). The example of such two-dimensional data can be the electromagnetic, thermoelectric and optical fields, the areas of air pollution and contamination of water, the biomedical data sets and, first of all, the digital video signals, which form a discrete two-dimensional fields.

In the connections with that, it seems appropriate the spreading of the fuzzy clustering of short time series with unevenly distributed observations approach [5] to the situation when the data are fed to the processing in online mode in the form of multivariate information stream in the context of Data Stream Mining [9].

2 Fuzzy Probabilistic Clustering of Multivariate Short Time Series

Let an initial information is given in the form of the set of $(q \times n)$ matrices $X(k) = \{x_{ip}(k)\}$ (here $i = 1, 2, \ldots, n$, n is a number of a separate observation of q-dimensional multivariate sequences in k-th sample), $k = 1, 2, \ldots, N$, $p = 1, 2, \ldots, q$ - p-th coordinate of multivariate process), which contains N $(N > n)$ q-dimensional samples with unevenly distributed observations, at that p-th component of $X(k)$ can be presented in the form $(1 \times n)$-dimensional vector $x_p(k) = (x_{1p}(k), x_{2p}(k), \ldots, x_{np}(k))^T$. Unevenness of quantization can be presented as $\Delta t_i = t_i - t_{i-1} \neq \Delta t_{i+1} = t_{i+1} - t_i$, i.e. $\Delta t_i \neq const$. The example of such sample is shown in the Fig. 1.

Obviously, neither conventional Euclidean distance nor classical probabilistic criteria cannot be used for the estimation of distance between two samples $X(k)$ and $X(l)$. For the estimation of distance between one-dimensional time series in [5] authors have introduced, so-called, PS-distance, which is based on the presentation of these time series in the form of piecewise linear functions. In fact, this PS-distance estimates a distinction of analyzed samples forms (obliquity). At that, the distance between two time series, for example, $x_p(k)$ and $x_p(l)$ can be written in the form

$$\begin{aligned} d_{PS}^2(x_p(k), x_p(l)) &= \sum_{i=1}^{n-1} \left(\frac{x_{i+1,p}(k) - x_{ip}(k)}{t_{i+1} - t_i} - \frac{x_{i+1,p}(l) - x_{ip}(l)}{t_{i+1} - t_i} \right)^2 \\ &= \sum_{i=1}^{n-1} \left(\frac{x_{i+1,p}(k) - x_{ip}(k)}{\Delta t_{i+1}} - \frac{x_{i+1,p}(l) - x_{ip}(l)}{\Delta t_{i+1}} \right)^2, \end{aligned} \tag{1}$$

which is satisfied to all conditions for distance functions.

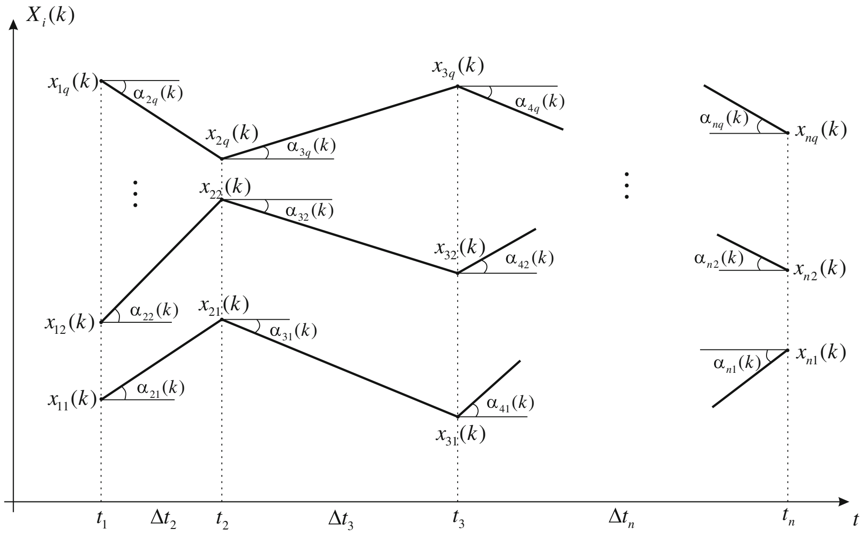


Fig. 1. Multivariate time series with unevenly distributed observations

Using the distance (1) authors of [5] have proposed the fuzzy clustering method, which is modified procedure of well-known fuzzy c-means algorithm (FCM) [7] for processing of one-dimensional time series with unevenly distributed observations.

It is easy to see that the distance's components (1) are in fact the first-order differences of discrete signals $x_p(k)$ and $x_p(l)$, i.e. this components are tangents of obliquity angles of piecewise linear functions:

$$\Delta x_{i+1,p}(k) = \frac{x_{i+1,p}(k) - x_{ip}(k)}{\Delta t_{i+1}} = tg\alpha_{i+1,p}(k), \tag{2}$$

$$\Delta x_{i+1,p}(l) = \frac{x_{i+1,p}(l) - x_{ip}(l)}{\Delta t_{i+1}} = tg\alpha_{i+1,p}(l). \tag{3}$$

Nevertheless, the time series, which are formed by differences, contains an one point less than initial sample, i.e. they contain $(n-1)$ observations instead n: $\Delta x_{2p}(k) = tg\alpha_{2p}(k), \Delta x_{3p}(k) = tg\alpha_{3p}(k), \ldots, \Delta x_{np}(k) = tg\alpha_{np}(k)$.

In the result of first-order difference computing (it is operation like first derivation in indiscrete (continuous) mode) the mean value of time series $x_p(k)$ is deleted. In this way for recovering of initial time series by its first-order differences it is necessary to add any initial observation into first-order differences sequence, for example, $x_{np}(k)$.

In this case a recovering of initial time series is performed by simple expression in the form

$$\begin{cases} x_{n-1,p}(k) = x_{np}(k) - \Delta x_{np}(k)\Delta t_n, \\ x_{n-2,p}(k) = x_{n-1,p}(k) - \Delta x_{n-1,p}(k)\Delta t_{n-1}, \\ \vdots \\ x_{1,p}(k) = x_{2p}(k) - \Delta x_{2p}(k)\Delta t_2. \end{cases} \tag{4}$$

Introducing $(1 \times n)$-dimensional vector

$$\tilde{x}_p(k) = (\Delta x_{2p}(k), \Delta x_{3p}(k), \ldots, \Delta x_{np}(k), x_{np}(k))^T$$

we can rewrite (1) in the conventional form

$$d^2_{PS}(x_p(k), x_p(l)) = \|\tilde{x}(k) - \tilde{x}(l)\|^2. \tag{5}$$

In this case we return to the standard Euclidean distance between first-order differences of the initial samples.

Further, using the distance (5) it is simply to implement any of fuzzy clustering methods [8]. For using evaluation idea of distances between time series by their first-order differences, let's introduce into consideration $q \times n$-dimensional matrix

$$\tilde{X}(k) = \begin{pmatrix} \Delta x_{21}(k) & \Delta x_{31}(k) & \cdots & \Delta x_{nl}(k) & x_{nl}(k) \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \Delta x_{ip}(k) & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \Delta x_{2q}(k) & \Delta x_{3q}(k) & \cdots & \Delta x_{nq}(k) & x_{nq}(k) \end{pmatrix}$$

and spherical norm instead of Euclidean distance in the form

$$D^2_{PS}(X(k), X(l)) = Tr(\tilde{X}(k) - \tilde{X}(l))(\tilde{X}(k) - \tilde{X}(l))^T. \tag{6}$$

This norm is an extension of distance (5) for matrix case. Based on the distance (6) we can provide the fuzzy clustering of samples array $\tilde{X}(1), \tilde{X}(2), \ldots, \tilde{X}(N)$.

After that, using technics of probabilistic fuzzy cluster analysis let's introduce into consideration the objective function in the form

$$\begin{aligned} E(u_j(k), \tilde{C}_j) &= \sum_{k=1}^{N} \sum_{j=1}^{m} u_j^{\beta}(k) D^2_{PS}(\tilde{X}(k), \tilde{C}_j) \\ &= \sum_{k=1}^{N} \sum_{j=1}^{m} u_j^{\beta}(k) Tr(\tilde{X}(k) - \tilde{C}_j)(\tilde{X}(k) - \tilde{C}_j)^T \end{aligned}$$

under constraints

$$\sum_{j=1}^{m} u_j(k) = 1 \quad \text{or} \quad \sum_{j=1}^{m} u_j(k) - 1 = 0, k = 1, 2, \ldots, N,$$

$$0 < \sum_{j=1}^{m} u_j(k) < N, j = 1, 2, \ldots, m$$

where $u_j(k)$ is membership level of matrix $\tilde{X}(k)$ to j-th cluster with matrix centroid $\tilde{C}_j$, m is number of clusters, which is set a priori, $\beta > 1$ is fuzzyfication parameter, which defines border "blurring" between clusters.

The result of clustering is $(N \times m)$-dimensional matrix $U = u_j(k)$, which is called fuzzy partition matrix, and centroid matrices $\tilde{C}_j$, $j = 1, 2, \ldots, m$.

Writing the Lagrange function in the form

$$\begin{aligned} L(u_j(k), \tilde{C}_j, \lambda(k)) = &\sum_{k=1}^{N} \sum_{j=1}^{m} u_j^{\beta}(k) Tr(\tilde{X}(k) - \tilde{C}_j)(\tilde{X}(k) - \tilde{C}_j)^T \\ &+ \sum_{k=1}^{N} \lambda(k) \left(\sum_{j=1}^{m} u_j(k) - 1 \right) \end{aligned} \tag{7}$$

(where $\lambda(k)$ are undetermined Lagrange multipliers) and solving system of Karush-Kuhn-Tucker equations

$$\begin{cases} \partial L(u_j(k), \tilde{C}_j, \lambda(k)) / \partial u_j(k) = \beta u_j^{\beta-1}(k) Tr(\tilde{X}(k) - \tilde{C}_j)(\tilde{X}(k) - \tilde{C}_j)^T + \lambda(k) = 0, \\ \partial L(u_j(k), \tilde{C}_j, \lambda(k)) / \partial \lambda(k) = \sum_{j=1}^{m} u_j(k) - 1 = 0, \\ \{\partial L(u_j(k), \tilde{C}_j, \lambda(k)) / \partial \tilde{C}_{jip}\} = -2 \sum_{k=1}^{N} u_j^{\beta}(k)(\tilde{X}(k) - \tilde{C}_j) = \mathbf{0} \end{cases}$$

(where $\{\partial L(u_j(k), \tilde{C}_j, \lambda(k)) / \partial \tilde{C}_{jip}\}$ -$(q \times n)$-dimensional matrix, which is formed by partial derivatives, $\mathbf{0}$ is matrix of the same dimensionality, which is formed by zeros), we can obtain fuzzy clustering algorithm in the form [9]

$$\begin{cases} u_j(k) = \frac{(Tr(\tilde{X}(k) - \tilde{C}_j)(\tilde{X}(k) - \tilde{C}_j)^T)^{\frac{1}{1-\beta}}}{\sum_{g=1}^{m} (Tr(\tilde{X}(k) - \tilde{C}_g)(\tilde{X}(k) - \tilde{C}_g)^T)^{\frac{1}{1-\beta}}}, \\ \lambda(k) = -\left(\sum_{g=1}^{m} (\beta Tr(\tilde{X}(k) - \tilde{C}_g)(\tilde{X}(k) - \tilde{C}_g)^T)^{\frac{1}{1-\beta}} \right)^{1-\beta}, \\ \tilde{C}_j = \frac{\sum_{k=1}^{N} u_j^{\beta}(k) \tilde{X}(k)}{\sum_{k=1}^{N} u_j^{\beta}(k)}. \end{cases} \tag{8}$$

This algorithm is close to Bezdek's clustering algorithm [7] when parameter $\beta = 2$ and is its extension for matrix version:

$$\begin{cases} u_j(k) = \frac{(Tr(\tilde{X}(k) - \tilde{C}_j)(\tilde{X}(k) - \tilde{C}_j)^T)^{-1}}{\sum_{g=1}^{m} (Tr(\tilde{X}(k) - \tilde{C}_g)(\tilde{X}(k) - \tilde{C}_g)^T)^{-1}}, \\ \tilde{C}_j = \frac{\sum_{k=1}^{N} u_j^2(k) \tilde{X}(k)}{\sum_{k=1}^{N} u_j^2(k)}. \end{cases} \tag{9}$$

Due to the fact that matrices $\tilde{C}_j$, $j = 1, 2, \ldots, m$ are centroids of clusters, which are formed by time series of first-order differences, for centroids recovering of initial data $\tilde{C}_j$ it is necessary to use expression (4).

3 Sequential On-Line Clustering of Multivariate Time Series Based on Modified Neuro-Fuzzy Network by T. Kohonen

The clustering methods (8), (9) are introduced in the assumption that all information is given in the form of fixed data array $X(1), X(2), \ldots, X(N)$ and is not changed in time. However, if discrete fields $X(k)$ are fed to the processing in sequential mode in the form of data stream, we can use approaches of Data Stream Mining, and first of all, adaptive methods [9,10].

The clustering neural networks, such as self-organizing maps of T. Kohonen [11,12], are the best for sequential data processing. Such network allows to provide crisp stream partition of vector observations. When initial information is fed in the form of $(q \times n)$-dimensional matrix observations under conditions of overlapping classes, we can use matrix neuro-fuzzy clustering network [13].

Using opportunity of the nonlinear programming recurrent algorithm by Arrow-Hurwicz-Uzava for searching of Lagrangian saddle point (7), we can write adaptive clustering procedure for multivariate time series with unevenly distributed observations in the form

$$\begin{cases} u_j(k) = \frac{(Tr(\tilde{X}(k)-\tilde{C}_j(k-1))(\tilde{X}(k)-\tilde{C}_j(k-1))^T)^{\frac{1}{1-\beta}}}{\sum_{g=1}^{m}(Tr(\tilde{X}(k)-\tilde{C}_g(k-1))(\tilde{X}(k)-\tilde{C}_g(k-1))^T)^{\frac{1}{1-\beta}}}, \\ \tilde{C}_j(k) = \tilde{C}_j(k-1) - \eta(k)\{\partial L(u_j(k), \tilde{C}_j, \lambda(k))/\partial \tilde{C}_{jip}\} \\ \qquad = \tilde{C}_j(k-1) + \eta(k) u_j^{\beta}(k)(\tilde{X}(k) - \tilde{C}_j(k-1)) \end{cases} \tag{10}$$

for arbitrary value of fuzzyfier β (here $\eta(k)$ is learning rate parameter) and

$$\begin{cases} u_j(k) = \frac{(Tr(\tilde{X}(k)-\tilde{C}_j(k-1))(\tilde{X}(k)-\tilde{C}_j(k-1))^T)^{-1}}{\sum_{g=1}^{m}(Tr(\tilde{X}(k)-\tilde{C}_g(k-1))(\tilde{X}(k)-\tilde{C}_g(k-1))^T)^{-1}}, \\ \tilde{C}_j(k) = \tilde{C}_j(k-1) + \eta(k) u_j^{2}(k)(\tilde{X}(k) - \tilde{C}_j(k-1)) \end{cases} \tag{11}$$

for $\beta = 2$.

It is easy to see that from the point of view of self-learning clustering network by T. Kohonen, the second recurrent relations in (10) and (11) are matrix modification of learning rule "Winner takes more" (WTM) [11], where the multiplier $u_j^{\beta}(k)$ fulfills the role of neighborhood function.

Hence, the architecture of self-organizing map with $(q \times n)$-dimensional matrix input and m neurons can be used for solving task of multivariate time series fuzzy clustering.

4 Simulation

For the efficiency confirmation of the proposed approach to clustering-segmentation of short time series with unevenly distributed observations, the task of clustering-segmentation of multivariate time series, which consist of three sequences - energy consumption, dry bulb temperature and dew point temperature from the New England Pool region [14] was solved. Results of the proposed approach allow increasing the quality of analysis and prediction of time series.

Time series consists of 2400 observations. For clustering this multivariate time series was divided by segments with 8 observations. For obtaining unevenly distributed observations in each segment, the 3rd and 5th observations were deleted from a data set.

Hence, according to (2) we obtain the data set in the form of the table "object-properties" with 300 observations and 6 properties. A number of clusters were $m = 3$ (morning, day and evening segments of energy consumption). All clustering algorithms were tested by the same data set. Average mean class error (MCE) was taken as the quality criterion of clustering results.

In the first experiment, we compared the performance of the clustering algorithms in the problem of classification when instances of all the available classes were present in the data set used for clustering, i.e., the number of classes was known a priori and equal to 3. The data sets were divided into the training and testing sets with 70% and 30% of data, respectively. For better performance of the recursive clustering algorithms, the data sets were randomly shuffled. The training sets were used for the initialization of the classifier through fuzzy clustering, and the testing sets were used for the comparison of the classification accuracy.

We used the learning rate $\eta = 0.01$ in the recursive procedures (10) and (11), and the "fuzzifier" parameter was taken $\beta = 2$. We performed 10 iterations for the batch clustering procedures, and 10 runs over the training data for the recursive clustering procedures. The experiment was repeated 50 times, and then average results were calculated. The final results are given in Table 1. They represent the percentage of the incorrectly classified objects from the testing data set.

Table 1. Results of clustering-segmentation of multivariate time series

Clustering procedures	$M\{MCE\}$
Proposed fuzzy clustering algorithm (8)	8.5%(25)
Proposed adaptive fuzzy clustering algorithm (10)	7.2%(22)
Matrix modification of fuzzy clustering method	12.1%(35)

As it can be seen from the obtained results, the proposed fuzzy clustering algorithms have the best quality of clustering (both batch and adaptive modes). We can also see that the results of adaptive modes of clustering algorithms have better quality than that of batch mode.

5 Conclusion

The adaptive fuzzy clustering approach of multivariate short time series with unevenly distributed observations or incomplete time series with missed or non-presented observations is proposed. For processing multivariate time series in batch mode we have proposed the matrix modification of fuzzy C-means method and for processing in sequential on-line mode - matrix neuro-fuzzy self-organizing network, which is learned using the rule "Winner takes more". Proposed matrix fuzzy clustering-segmentation algorithms for time series are enough simple in computational implementation and can be used in machine learning, data stream mining, big data processing tasks.

References

1. Liao, T.W.: Clustering of time series data-a survey. Pattern Recogn. **38**(11), 1857–1874 (2005). doi:10.1016/j.patcog.2005.01.025
2. Mitsa, T.: Temporal Data Mining. CRC Press, Boca Raton (2010)
3. Aggarwal, C.C., Reddy, C.K.: Data Clustering. Algorithms, and Applications. CRC Press, Boca Raton (2014)
4. Aggarwal, C.C.: Data Mining. Springer, New York (2015)
5. Möller-Levet, C.S., Klawonn, F., Cho, K.-H., Wolkenhauer, O.: Fuzzy clustering of short time series with unevenly distributed sampling points. Lecture Notes in Computers Science, vol. 2810, pp. 330–340. Springer, Heidelberg (2003). doi:10.1007/978-3-540-45231-7_31
6. Cruz, L.P., Vieira, S.M., Vinga, S.: Fuzzy clustering for incomplete short time series data. Lecture Notes in Artificial Intelligence, vol. 9273, pp. 353–359. Springer Int. Publishing, Switzerland (2015). doi:10.1007/978-3-319-23485-4_36
7. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)
8. Höppner, F., Klawonn, F., Kruse, R., Runkler, T.: Fuzzy Clustering Analysis: Methods for Classification, Data Analysis, and Image Recognition. Wiley, Chichester (1999)
9. Bifet, A.: Adaptive Stream Mining: Pattern Learning and Mining from Evolving Data Streams. IOS Press, Amsterdam (2010)
10. Bodyanskiy, Y., Skuratov, M., Volkova, V.: Adaptive matrix fuzzy c-means clustering. In: 19th East-West Fuzzy-Colloquium, pp. 96–103. HS, Zittau-Görlitz (2012)
11. Kohonen, T.: Self-Organizing Maps. Springer, Berlin (1995)
12. Haykin, S.: Neural Networks. A Comprehensive Foundation. Prentice Hall, Upper Saddle River (1999)
13. Bodyanskiy, Y., Skuratov, M., Volkova, V.: Matrix neuro-fuzzy self-organizing clustering network. Comput. Sci. Inf. Technol. Manag. Sci. **49**, 54–58 (2011). doi:10.2478/v10143-011-0042-1
14. Deoras, A.: Electricity load forecasting using neural networks. In: Electricity Load and Price Forecasting Webinar Case Study (2011). https://www.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study?s_tid=srchtitle

Learning in Comparator Networks

Łukasz Sosnowski[1(✉)] and Dominik Ślęzak[2]

[1] Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, 01-447 Warsaw, Poland
sosnowsl@ibspan.waw.pl
[2] Institute of Informatics, University of Warsaw,
ul. Banacha 2, 02-097 Warsaw, Poland
slezak@mimuw.edu.pl

Abstract. We discuss how to train and tune comparators aimed at multi-similarity-based classification of compound objects. The proposed approach is supported by a collection of techniques and algorithms for construction and use of comparator networks. The described methodology has been implemented as a software library and may be used for a variety of future applications.

1 Introduction

Learning is one of the most important processes, which makes the analytical methods intelligent. Comparator networks are focused on recognition tasks, which involve management, identification, classification and modelling of data objects that are intrinsically compound and described by means of various types of information. This approach is based on object similarities computed with respect to single features, which are then combined into the final result. Precisely, the result of utilization of a network for an input object consists of rankings based on its similarities to a set of so called reference objects. In this article we propose some methods in the field of learning such types of networks, by means of particular comparators and the overall network structures.

A comparator is a basic computational unit that models single aspect of the similarity between objects reflecting physical phenomena, processes or subsystems. They usually are arranged into comparator networks that are capable of aggregating and summarizing local information about similarity into a global measure of proximity between data objects. This provides us the ability to solve the problem at hand by decomposing it into simpler, localized steps that can be processed quickly. Then, local results from elementary units can be aggregated and processed in the network producing the overall, possibly generalized final result.

The learning procedure in this field is the one, which makes the whole solution better, easier and less time consuming. In the short term we can say about tuning the network solution. The goal is to make the process of recognition

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_29

with less comparison operations using few comparators on the one hand, but with enough accuracy. This goal can be divided into various smaller tasks. The first can be an increasing efficiency (by means of accuracy). The second task is to make architecture of the network as easy as possible. It means removing of comparators or moving them from layer to layer. In both cases the supervised learning is helpful. Using already known results we can eliminate comparators and then check if the final result is still good enough. Checking all possible configurations is quite costly, that is why we should use here some heuristic methods.

In the long term perspective, learning can be a way of adaptation of a solution to a changing environment. The learning procedure should be repeated as many times as we recognize that there are types of objects, which are not properly handled by our solution. Such new objects collected from previous execution of the learning process are the cases that can help the network to adapt to new situations and to be more general.

The article is composed as follows: Sect. 2 describes the nature of learning in general and refers to some well known approaches that can be relevant. Section 3 introduces the methodology of comparator networks. Section 4 contains the main topic of this paper. It explains the usage of standard learning methods in the field of comparator networks. It describes step by step how to apply some popular learning and what can be a target of such procedures. Section 5 presents a general case study, which shows how these methods can change the solution with respect to efficiency, time and other factors. The example is based on experiments reported in earlier publications, which, however, did not focus on the aspects of learning. We can find there the results, which have been achieved earlier, the short description of the methods used and new results after learning has been applied. Section 6 contains summary, some discussion and the future work.

2 Related Work

There are many learning methods and acquiring knowledge known in the decision systems. Their division is based on many criteria. The main factor is a strategy of learning and output, which is produced by the procedure [1].

The first widely described approach to machine learning is a *supervised learning*. The boundary condition is having proper data to use this procedure. The method can be used only if there is possibility to construct learning and testing sets. In practice, this assumption sometimes turns out to be too strong, which prevents the use of this method. However, assuming possession of relevant data, pre-processing is performed to improve data quality, remove duplicates, etc. Next the learning set is constructed, which has a valid output for the given input elements. Thus, n pairs of input and output examples are obtained, by which the algorithm approximates the function modeling the given phenomenon. The number of elements of the learning set is a 33% of the total quantity. Its elements are selected randomly, so as to ensure the independence of selection of objects.

The remained part of the initial set became the test set, which is necessary to evaluate the learned function. The next step is to choose the supervised learning technique (e.g. back propagation [2]).

A generalized supervised learning algorithm is shown in Fig. 1 in the form of a UML activity diagram. The procedure described is repeated many times for individual objects from the training set until the stop condition is met.

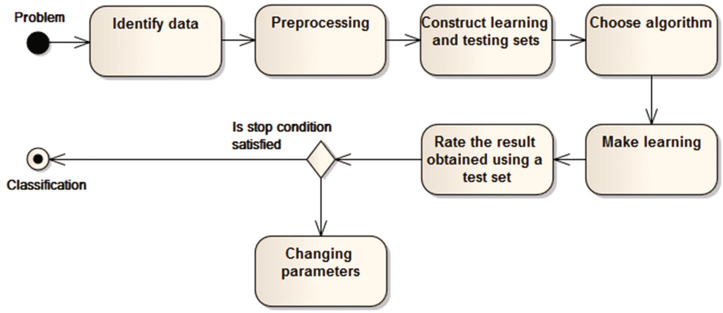


Fig. 1. The activity diagram (UML) of generalized supervised learning procedure

The second approach to learning that is complementary to the one described previously is *unsupervised learning*. The method used for problems, for which the test sets either do not exist or are difficult to construct. This method uses knowledge derived directly from data. It involves detection, analysis and modeling. It boils down to the clustering analysis based on various criteria, including the object similarity criteria. Discovering knowledge from the data allows to identify features, with respect to which the grouping should be performed. The problem of automatic grouping is a matter of high complexity. There is no single universal grouping algorithm for each data type. In each case, one should use the knowledge domain of the problem and select the algorithm individually. An additional difficulty is the unknown number of groups to be created as a result of the algorithm. However, there are certain conditions that a data group should fulfill, i.e. homogeneity within each of groups and heterogeneity between groups. The learning procedure is based on data availability, but without a decision-making label. As a criterion for assessing the correctness of grouping, it is used so-called *grouping quality indicators* [3].

Both in the literature and in practical applications there are known many other approaches to machine learning. They are apart of a hybrid approach that combines both of the above described cases or they are differing in one of the fundamental parts of the algorithm. Examples of methods are: *semi-supervised learning*, where learning can proceeds with a small sample of labeled data and a large sample of unlabeled data, *reinforcement learning* where the aim is to automatically acquire procedural knowledge based on interaction with the environment. Indicators for learning may be in this case the reaction of the environment triggered by the algorithm. Another example could be *transduction*, which is similar to supervised learning but the form of learning results is different. In this case, this is not a function, but the output values derived from the prediction based on input training data. A slightly different category of learning is *meta-learning*. The idea here is to use different learning methods to verify a set of hypotheses. Its purpose is to select the best learning method for the given problem [4].

3 Comparator Networks

There are many ways to implement object recognition solutions. The method considered in this paper is based on multi-similarity calculations, gathering many aspects of the similarity between pairs of objects. The objects belong to multi-dimensional space and are described by the similarity values between input objects and reference objects, measured on the given set of features. The result of the recognition is in a form of similarity vector, which shows the closeness between input object and reference points in the domain space. The units responsible for single-feature calculations will be called *comparators*. The networks allowing to process input objects through the layers of multiple comparators will be called *comparator networks*.

Comparator networks can play different roles depending on their settings. They can serve as multi-stage classifiers whose purpose is to limit the reference set of objects and identify the most probable candidate to be a final result. The scenario of processing in such networks is to compute relatively simple features at the first layers and to filter out the reference objects to only those that are the most promising in the final perspective. Particular comparators can be also specialized in recognition of different features based on the nature of sub-objects. The idea is that the similarity of parts of objects can help in resolving the similarity of the whole objects. Sometimes the knowledge only about parts is not enough but it brings us closer to the solution and having some additional domain knowledge the satisfactory result can be obtained.

From the mathematical perspective a comparator network can be interpreted as a calculation of a function:

$$\mu_{net}^{ref_{out}} : X \rightarrow [0,1]^{|ref_{out}|}, \tag{1}$$

which takes the input object $x \in X$ as an argument and ref_{out} is a reference set for the network's output layer. The target set of $\mu_{net}^{ref_{out}}$ is the space of proximity vectors. The proximity vector from the target space will be denoted by $\boldsymbol{v}$. Such a vector encapsulates information about similarities between a given input object x and objects from the reference set ref, by ordering the reference set, i.e. taking $ref = \{y_1, ..., y_{|ref|}\}$. In this way we get the value network's function of:

$$\mu_{net}^{ref}(x) = \langle SIM(x, y_1), ..., SIM(x, y_{|ref|})\rangle, \tag{2}$$

where $SIM(x, y_i)$ is the value of *global similarity* established by the network for an input object x and a reference object y_i. Global similarity depends on partial

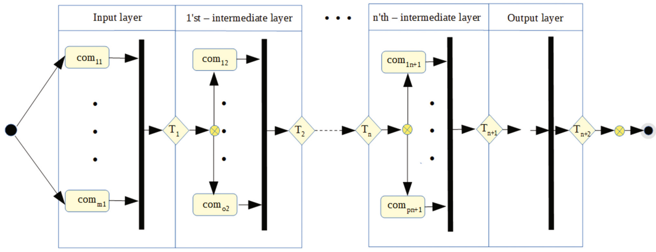


Fig. 2. General scheme of a comparator network in UML-like representation. Notation: com_{ji} – comparators, T_j – translators. Symbols: oval – comparator, thick vertical line – aggregator, rhombus – translator, encircled cross – projection module.

(local) similarities calculated by the elements of the network (unit comparators). Through application of aggregation and translation procedures at subsequent layers of the network these local similarities are ultimately leading to the global one. Particular elements of the network have been described in detail in our previous publications [5]. Only for remind reasons we put the scheme of the comparators network with all possible elements on the Fig. 2.

4 Comparator Learning

The issue of constructing a reference set is an inseparable element of the design of comparator-based solutions. It is a set of patterns, so it is important to get it as good as possible. Optimization of this set is reduced to the problem of grouping objects and operating in the reference set with only representatives of the group.

Designation of a reference set can be performed if we have a set of object instances. This procedure can be performed both for a fixed comparator and for a fixed comparator network. The similarity function of objects in a cluster is modeled by a comparator network (multiple local similarities). To perform a learning procedure, the cross-validation method is used with *leave-one-out* option, where self-organization of objects in clusters is performed.

A single iteration of learning has been presented in Fig. 3. The algorithm starts with an empty reference set. In this case, the first input object becomes automatically the one-element cluster and the representative of this cluster in the reference set. Each iteration of the algorithm determines the proximity vector relative to the current clusters representatives. If the returned vector is nonzero with correspondingly high similarity values, then the input object is assigned to the cluster represented by the reference object, for which the most similarity was obtained. The applied method can indicate more than one cluster simultaneously, which should result in cluster merging into one and selecting a new representative of the newly created cluster.

If the proximity vector is returned as the zero one, the object creates a new cluster and automatically supplying it. The size of the cluster corresponds

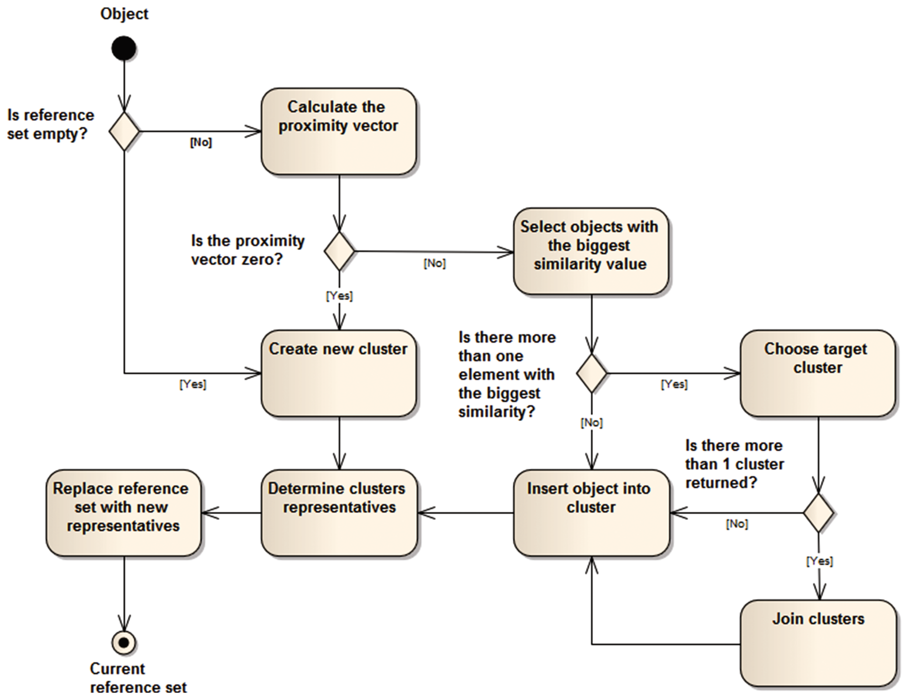


Fig. 3. Activity diagram in UML notation performs one iteration in unsupervised learning of set of reference objects

to the parameters of comparators p, defining the minimum acceptable resemblance at the output of the individual comparators. Therefore, if you want to get fewer clusters, automatically increase the tolerance for differences between objects within one cluster, lower the value of p.

An analogous procedure is used for each successive element from the cross validation set, until the object is exhausted. It should be noted that setting too stringent quality criteria for a similarity may result in the creation of separate clusters for each object. Conversely, if the parameter values set too low, it may happen that all the objects are in one cluster.

Another example of learning the compound object comparator is the optimization of the threshold parameter p. As defined in [6], the parameter p is the minimum acceptable value of the similarity in the proximity vector of the comparator. In order to automatically designate it, there should be uses the supervised method, where the learning is based on the training set. For this procedure we can use re-sampling method (with k at least 10). The data set is divided into two subsets. The first is called the learning set, which consists of $\frac{1}{3}$ of all elements, and the second is a testing set with quantity of $\frac{2}{3}$ of all elements. In the case of a single comparator, it is necessary to determine the limit value of similarity, below which no correct solution is found. One of the available methods to use is so called simple local search. The starting point of the algorithm is 1. The quality of the solution is determined by the function in a form:

$$f_{eval}(p) = \sum_{x \in X} (f_{recall}(x) - (1 - p)) \tag{3}$$

where $f_{recall} : X \rightarrow \{0, 1\}$ is a function returning 1 in case of achieving correct decision label from training set for a pair of objects (x, y) or 0 otherwise.

The higher the value of function (3) is the better quality of the solution can be achieved. Neighboring solutions are generated by modifying the parameter values by a fixed constant, such as 0.01 (addition or subtraction). The stop condition can be implemented as a monitoring of repetitive solutions. When the limit value is exceeded, the quality of the solution will start to swing (increase then decrease repeatedly). Upon detection of this phenomenon, processing should be discontinued.

Another area of a potential usage of the learning methods is construction of the network structure, by means of selecting features, which will be utilized by comparators. This issue relates directly to the concept of the object, its features, relationships and dependencies. Objects are defined by the description in the ontology [7], which operates on concepts and relationships between them. It specifies a set of features that can be used for network construction. The optimal selection of features for a given problem is therefore the problem of searching for minimal subsets of attributes, which can uniquely identify the objects or their classes. Such minimal subsets of attributes can be referred to as reducts in the rough set theory [8].

The process of selecting a network structure can be transformed into the problem of determining the significance of features and in particular the possibility of their elimination. The implementation was done by using evolutionary algorithms [9]. The learning procedure assumes the existence of a set of features for the processed objects (stored in the ontology). The task of the procedure is to select a minimum number of features and to indicate the location of the comparators in layers, ensuring that the network operates correctly with the established types of local aggregators, translators, and the number of layers.

Individuals represent the various configurations of the network structure, e.g. the comparator and its feature assigned to the particular layer. A single chromosome is in a form: $\boldsymbol{x} = \langle (i, k) \rangle$, where i is an identifier of the comparator and $k \in \{0, 1, ..., n-1\}$, where n is a cardinality of layers. Value k represents the number of a single layer, where 0 stands for no assignment to any layer. The order of the comparators in the layer is irrelevant due to concurrent processing. The comparator can be assigned to at most one layer.

Processing begins with a random selecting of the population. Individuals are randomized multiple times until a full population of correct individuals is formed. An improper individual is the one that can not be used to construct correct network, e.g. $\langle (1, 0), ..., (n-1, 0) \rangle$. Population size is parameter of the algorithm. In each iteration (called epoch) there are reproduction operations, application of genetic methods, evaluation and succession. The scheme of the applied genetic algorithm is consistent with evolutionary methods described in the literature. Reproduction was made with tournament selection, which accepts the l parameter that indicates the number of individuals drawn to the tournament. The number of individuals created in the reproduction process is also parameter of the algorithm. Reproduction creates a local population, on which the cross-over and mutation operations are performed. Two individuals are randomized to cross-over operation. The procedure involves randomly assigning a cut-off point where the elements of the two chromosomes are replaced. From the zero position to the drawn position, the genes of the first chromosome remain, and from the drawn position to the end of the genes of the second chromosome. This operation can, however, lead to the improper individual. In this case, an individual with a higher adaptation ratio is returned at the output. The mutation operation involves adding a random natural number to the current value and executing $mod\, n$, where n is the number of layers. Mutation is performed with a certain probability for each chromosome gene. One of the most important parts of the algorithm, individually adopted to the specifics of the network od comparators is the evaluation function (population fit assessment). The fit assessment values take into account both the final results of the recognition and the cost of performing the calculations by the comparators. The last value is calculated on the basis of the unit cost of the comparator's execution and the cardinality of reference set in given layer. The general form of the fit assessment function is as follows:

$$f_X(ch) = \sum_{x \in X} f_{recall}(x, ch) - \left(\frac{\sum_{i=1}^{n-1} \sum_{j=1}^{m} f_{ij}^{cost} * |ref_i|}{|C| * |ref|} - \alpha \right) \quad (4)$$

where ch is a chromosome representing a given network structure, X is a set of input objects from a learning set, C is a set of all considered comparators in a given problem, f_{ij}^{cost} is the function of the unit cost of execution the comparator with index j in the layer number i. Values of this function are from the interval $(0, 1)$, ref_i is a subset of the ref set occurring as a reference set in the layer with index i, and α is a positive value close to zero. In this way, the value of the penalty function (part in the second parenthesis (4)) is always less than 1, which makes the elements identified more accurately than the simple structure of the network, although the last element is important as well. In the given formula, f_{recall} expresses the performance of recognition for the individual from the learning set. Succession is performed in the same way as reproduction, e.g. using tournament selection. In this case both populations are combined and the individuals are selected for the tournaments. The tournament is played as many times as the size of the target population is. The population is also a parameter of the method. The stop condition of the algorithm refers to the number of epochs, for which there is no improvement in the solution.

5 Case Study

In this paper the automatic character recognition solution based on a comparator network is described. We designed a comparator network with three layers representing particular contexts. The first layer covers features of a very general nature. They can significantly reduce the cardinality of a reference set. The second layer relates to an in-depth analysis of the image, which gets the final answer (decision). The last layer in the classical way contains only an aggregator that performs a synthesis of the previously obtained results. The reference set consists of objects, which are images of characters grouped by different sizes and font types. The cardinality of the reference set grows with the number of types of recognized characters, languages, fonts, sizes, etc. In practice this is one of the main problems that there is a plenty of data that needs an efficient procedure to be preselected for further processing.

The described network has three comparators in the input layer: the font size, pixels distribution and an axis of colors. The representations of objects used for them are the following: the height of a character in pixels, a list of four quantities of the black pixel for each quarter of image, a two-string pattern created from the axis X and the axis Y of the image (in the middle) and representing pixel color changing. In the intermediate layer, there are four comparators: Upper approximation, Quadrangulation, Coherent area, Multi axis of colors. The first comparator compares images arising from the granulation process [10]. The upper approximation is represented by granules, which are activated if and only if there are black pixels inside. This means the comparison of images converted to a very low resolution ($m \times n$ – granulation parameters) and represented by an array of $\{0, 1\}$ values (*1* – activated, *0* – none).

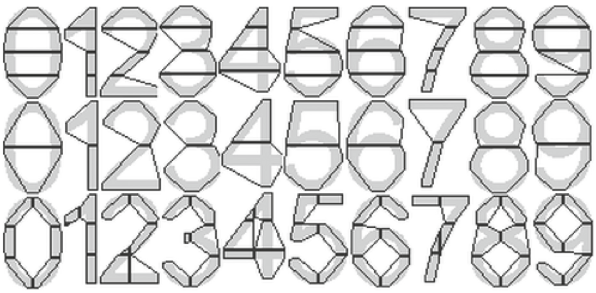


Fig. 4. Reference objects prepared for comparisons by the quadrangulation comparator with respect to their shapes (with granulation resolution parameters 1×4, 2×2, 2×3 displayed in consecutive rows).

Another comparator deals with the comparison of the geometric shapes arisen as a result of connection extremes points of subsequent granules (see Fig. 4). Further action is to take contours and to calculate the surface area of quadrangles. This value is required to calculate the final factor in a form of the quotient of the surface area of the quadrangle to the area of the whole image. The third comparator is prepared to compare coherent areas within objects. The last comparator in this layer (Multi axis of colors) compares string patterns created on the base of changing colors in the particular lines of image. There are two separate representations: one for the horizontal lines and second for the vertical ones. The string patterns have been created analogically to the one used in the input layer, but using each line (not only the middle one). After generation the string of signs for particular lines, the main pattern is created based on the signs, which have changed (from the previous state). This is in a form of two lists of string patterns to compare.

5.1 Results

First we have done the experiment with the manually designed network as we described above. Input set consists of 93 randomly generated images with single digit. The images contains different size and different font type. There was a combination of font sizes {*10, 14, 18, 24, 36, 48, 60*} and types of fonts {*Times New Roman, Arial, Verdana, Courier*}. Table 1 shows the achieved results. The time needed for the recognition on quad core INTEL Core i7 4700 HQ processor was 5661 ms.

Table 1. Results achieved for 93 digits performed for individual characters for all font types and all processed font sizes with the manually designed comparator network.

Character	Precision	Recall	F-score
0	0.91	1.00	0.95
1	0.94	1.00	0.97
2	1.00	1.00	1.00
3	1.00	1.00	1.00
4	0.87	1.00	0.93
5	0.80	1.00	0.89
6	1.00	1.00	1.00
7	0.91	1.00	0.95
8	0.88	1.00	0.93
9	1.00	1.00	1.00
ALL	0.92	1.00	0.96

After that we have used the learning procedure to determine best structure of the network with keeping quality of the recognition. We have numbered the comparators as: 1 – axis of colors, 2 – coherent area, 3 – quadrangulation, 4 – font size, 5 – upper approximation, 6 – multi axis of colors, 7 – pixels distribution. We have assigned the relative unit cost for each of them as follows: 1 – 0.1, 2 – 0.6, 3 – 1.0, 4 – 0.1, 5 – 0.3, 6 – 0.6, 7 – 0.2. We have used a learning structure algorithm made with genetic algorithm. Firstly we have achieved the results described with network structure and value of the fit assessment function. The results are shown in Table 2.

Table 2. Learned network structures ($\{1, 2, 3, 4, 5, 6, 7\}$ – identifiers of comparators, $\{0, 1, 2\}$ – identifiers of layers, where 0 means no assignment to any layer)

Network structure	Fit assessment value
1 -> 2, 2 -> 0, 3 -> 0, 4 -> 0, 5 -> 1, 6 -> 0, 7 -> 2	92.94796018367347
1 -> 2, 2 -> 2, 3 -> 0, 4 -> 2, 5 -> 1, 6 -> 0, 7 -> 2	92.9265316122449
1 -> 0, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 0, 6 -> 1, 7 -> 2	92.91076630612245
1 -> 1, 2 -> 1, 3 -> 2, 4 -> 0, 5 -> 2, 6 -> 2, 7 -> 2	92.870001
1 -> 2, 2 -> 2, 3 -> 0, 4 -> 2, 5 -> 2, 6 -> 1, 7 -> 2	92.9023479387755
1 -> 0, 2 -> 1, 3 -> 0, 4 -> 0, 5 -> 2, 6 -> 2, 7 -> 2	92.89632753061224
1 -> 2, 2 -> 2, 3 -> 2, 4 -> 0, 5 -> 0, 6 -> 1, 7 -> 0	92.90908263265307
1 -> 0, 2 -> 2, 3 -> 2, 4 -> 1, 5 -> 1, 6 -> 0, 7 -> 0	92.8938785510204
1 -> 2, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 1, 6 -> 0, 7 -> 2	92.94489895918367
1 -> 0, 2 -> 1, 3 -> 0, 4 -> 1, 5 -> 0, 6 -> 2, 7 -> 2	92.8869397755102

Table 3. Results achieved for networks structure in a form of: *1 -> 2, 2 -> 0, 3 -> 0, 4 -> 0, 5 -> 1, 6 -> 0, 7 -> 2* (first on the left), *1 -> 2, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 1, 6 -> 0, 7 -> 2* (in the middle), *1 -> 0, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 0, 6 -> 1, 7 -> 2* (on the right)

Digit	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
0	0.91	1.00	0.95	0.91	1.00	0.95	0.91	1.00	0.95
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	0.91	1.00	0.95	0.91	1.00	0.95	0.91	1.00	0.95
8	0.86	0.86	0.86	0.86	0.86	0.86	1.00	1.00	1.00
9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ALL	0.97	0.99	0.98	0.97	0.99	0.98	0.98	1.00	0.99

Testing the structures show that the time execution is much more shorter, e.g. for network *1 -> 2, 2 -> 0, 3 -> 0, 4 -> 0, 5 -> 1, 6 -> 0, 7 -> 2* – 2692 ms, *1 -> 2, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 1, 6 -> 0, 7 -> 2* – 2781 ms, *1 -> 2, 2 -> 2, 3 -> 0, 4 -> 2, 5 -> 1, 6 -> 0, 7 -> 2* – 4125 ms, *1 -> 0, 2 -> 0, 3 -> 0, 4 -> 2, 5 -> 0, 6 -> 1, 7 -> 2* – 4723 ms, etc. Table 3 shows the results of recognition for selected network structures.

6 Summary

We discussed how to train and tune comparator networks aimed at multi-similarity-based classification of compound objects. We introduced several techniques aimed at intelligent construction and use of the considered networks, particularly focused on validation whether all features used by the network units to compare input and reference objects are needed to assure good performance.

Further research should be focused on development of a framework for tuning aggregators. There is still place for learning on this area, both regarding the selection and the parameters of individual aggregators adapting various approaches to reaching consensus between different sources of similarity-based object rankings [11]. It is also possible to extend the environment for selecting features and reference objects that can together serve as an optimal "spine" for comparator network performance. With this respect, it is worth referring to feature and object selection methods based on rough sets and fuzzy similarities [12].

Basic examples of experiments reported in this paper show that appropriate methods of learning comparator network structures can lead to very interesting results. The achieved simpler structures can recognize objects in shorter time and with greater efficiency. Basically, the introduced techniques can detect that some of comparators are not necessary in the final model.

References

1. Ayodele, T.: Introduction to Machine Learning. INTECH Open Access Publisher (2010)
2. Szczuka, M., Ślęzak, D.: Feedforward neural networks for compound signals. Theoret. Comput. Sci. **412**(42), 5960–5973 (2011)
3. Rutkowski, L.: Computational Intelligence: Methods and Techniques. Springer, Heidelberg (2008)
4. Michalewicz, Z., Fogel, D.B.: How to Solve It - Modern Heuristics, 2nd edn. Springer, Heidelberg (2004)
5. Sosnowski, Ł.: Framework of compound object comparators. Intell. Decis. Technol. **9**(4), 343–363 (2015)
6. Sosnowski, L., Szczuka, M.S.: Recognition of compound objects based on network of comparators. In: Proceedings of FedCSIS 2016, Position Papers, pp. 33–40 (2016)
7. Staab, S., Maedche, A.: Knowledge portals: ontologies at work. AI Mag. **22**(2), 63–75 (2001)
8. Pawlak, Z., Skowron, A.: Rough sets: some extensions. Inf. Sci. **177**(1), 28–40 (2007)
9. Dasgupta, D., Michalewicz, Z.: Evolutionary Algorithms in Engineering Applications. Springer, Heidelberg (1997)
10. Pedrycz, W., Skowron, A., Kreinovich, V.: Handbook of Granular Computing. Wiley, Hoboken (2008)
11. Świechowski, M., Kacprzyk, J., Zadrożny, S.: A novel game playing based approach to the modeling and support of consensus reaching in a group of agents. In: Proceedings of SSCI 2016, pp. 1–8 (2016)
12. Benitez, M.J., Medina, J., Ślęzak, D.: δ-Information reducts and bireducts. In: Proceedings of IFSA-EUSFLAT-15 (2015)

Fuzzy φ-pseudometrics and Fuzzy φ-pseudometric Spaces

Alexander Šostak[1,2(✉)] and Raivis Bēts[1(✉)]

[1] Department of Mathematics, University of Latvia, Riga 1002, Latvia
raivis.bets@lu.lv
[2] Institute of Mathematics and Computer Science, UL,
Raina bulv. 29, Riga 1459, Latvia
sostaks@latnet.lv, sostaks@lanet.lv

Abstract. By replacing the axiom $m(x,x,t) = 1$ for all $x \in X, t > 0$ in the definition of a fuzzy pseudometric in the sense of George-Veeramani with a weaker axiom $m(x,x,t) = \varphi(t)$ for all $x \in X, t > 0$ where $\varphi : \mathbb{R}^+ \to (0,1]$ is a non-decreasing function, we come to the concept of a fuzzy φ-pseudometric space. Basic properties of fuzzy φ-pseudometric spaces and their mappings are studied. We show also an application of fuzzy φ-pseudometrics in the words combinatorics.

Keywords: Fuzzy pseudometric · Fuzzy φ-pseudometric · Supratopology · Cauchy sequences · Baire category theorem

1 Introduction

In 1942 K. Menger has introduced the concept of a statistical metric, see e.g. [11]. Basing on the concept of a statistical metric I. Kramosil and J.Michalek in [9] introduced the notion of a fuzzy metric.

In 1994 George and Veeramani [2], see also [3], slightly modified Kramosil-Michalek's definition of a fuzzy metric. This modification allows more natural examples of fuzzy metrics, in particular fuzzy metrics constructed from metrics; besides George-Veeramani fuzzy metrics are more appropriate for the definition and the study of the induced topological structure. In our work we generalize George-Veeramani definition of a fuzzy metric by weakening one of their axioms thus coming to a concept called a fuzzy φ-(pseudo)metric, where $\varphi : \mathbb{R}^+ \to [0,1]$ is a non-decreasing function such that $\lim_{t\to\infty} \varphi(t) = 1$. The motivation for such generalizing of George-Veeramani's definition will be presented in Remark 2.4. In Sect. 2 we study some properties of fuzzy φ-(pseudo)metrics. Two different topological-type structures on φ-(pseudo)metric spaces are discussed in Sect. 3. In the fourth section we consider the sequential structure of fuzzy φ-(pseudo)metric spaces, that is properties described by the behavior of sequences in such spaces. Uniformly continuous mappings of fuzzy φ-pseudometric spaces

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_30

are considered in Sect. 5. In particular, a certain version of Baire category theorem is established here. In Sect. 6 we construct a fuzzy φ-pseudometric on the set of infinite words from a sequence of partial ordinary pseudometrics on this set thus illustrating one of possible applications of fuzzy φ-pseudometrics.

2 Fuzzy φ-pseudometrics and Fuzzy φ-pseudometric Spaces

Definition 1 (George-Veeramani [2,3]). *A fuzzy pseudometric on a set X is a pair $(m, *)$ where $*$ is a continuous t-norm and $m : X \times X \times \mathbb{R}^+ \to (0, 1]$, satisfies the following conditions for all $x, y, z \in X$, $s, t \in \mathbb{R}^+ = (0, +\infty)$:*

(1FPM) $m(x, y, t) > 0$;
(2FPM) $m(x, y, t) = 1 \Longleftarrow x = y$;
(3FPM) $m(x, y, t) = m(y, x, t)$;
(4FPM) $m(x, z, t + s) \geq m(x, y, t) * m(y, z, s)$;
(5FPM) $m(x, y, -) : \mathbb{R}^+ \to [0, 1]$ *is continuous.*

We get definition of a fuzzy φ-metric by replacing (2FPM) *with a stronger axiom*

(2FM) $x = y \Longleftrightarrow m(x, y, t) = 1$

While being fully "satisfied" with axioms (1FPM), (3FPM), (4FPM), (5FPM), we suggest to replace axiom (2FPM) with a more general axiom (2φFPM) where $\varphi : (0, \infty) \to [0, 1]$ is a certain non-decreasing function. However, to be coherent, we have also to modify axiom (5FPM) by replacing it with a stronger axiom (5φFPM). Motivations for these changes and some examples are given below.

Definition 2. *A fuzzy φ-pseudometric on a set X is a triple $(m, *, \varphi)$, where $*$ is a continuous t-norm, $\varphi : (0, \infty) \to [0, 1]$ is a nondecreasing function such that $\lim_{t\to\infty} \varphi(t) = 1$ and $m : X \times X \times \mathbb{R}^+ \to (0, 1]$ is a mapping satisfying the following conditions for all $x, y, z \in X$, $s, t \in \mathbb{R}^+$:*

(1FPM) $m(x, y, t) > 0$;
(2φFPM) $m(x, x, t) = \varphi(t) \geq m(x, y, t)\ \forall x, y \in X$;
(3FPM) $m(x, y, t) = m(y, x, t)$;
(4FPM) $m(x, z, t + s) \geq m(x, y, t) * m(y, z, s)$;
(5φFPM) *function* $m(x, y, -) : \mathbb{R}^+ \to [0, 1]$ *is continuous and non-decreasing.*

A fuzzy φ-pseudometric m is called a fuzzy φ-metric if m satisfies axiom:

(2φFM) $m(x, y, t) = \varphi(t) \Leftrightarrow x = y$;

The quadruple $(X, m, *, \varphi)$ where $(m, *, \varphi)$ is a fuzzy φ-pseudometric on X, is called a fuzzy φ-pseudometric space.

We will usually abbreviate notations $(m, *, \varphi)$ and $(X, m, *, \varphi)$ and write just m and (X, m) in case when it will not lead to misunderstanding.

Remark 1. It is well-known that axiom (4FPM) combined with axiom (2FPM) implies that fuzzy pseudometric $m(x, y, -)$ is non-decreasing. On the other hand, axiom (4FPM) combined with axiom (2φFPM) does not allow to conclude that fuzzy φ-pseudometric $m(x, y,)$ is non-decreasing. Therefore, to be coherent, we request this explicitly in the modified axiom (5φFPM).

Remark 2. We came to the idea to replace axiom (2FPM) by some weaker axiom in the study of analytic problems in the theory of infinite words combinatorics. While the tools of fuzzy pseudometrics on the whole seem to be useful in this area, the "categoricity" of the axiom (2FPM) did not allow us to construct a "full bodied" fuzzy pseudometric, which would give "satisfactory" information about the distance between two infinite words. In order to solve this problem we introduced the notion of a *fragmentary fuzzy pseudometric* that obtained this name since they were constructed from "fragments" of ordinary metrics. Fragmentary fuzzy pseudometrics inspired us to introduce and to study a more general, and actually more natural concept of a fuzzy φ-pseudometric including fuzzy fragmentary metrics as a special case.

Besides these "practical" reasons, we see justification of the axiom (2φFPM) also from the "ideology" of "fuzzy mathematics". Namely, constituting that a distance between two equal objects should be the same for every $t \in \mathbb{R}^+$ and not be a subject of possible evaluation on each level t seems to be not natural in the context of defining "distance" by fuzzy pseudometrics. In this concern, we recall also the concept of an M-valued set, where an element need not be "fully equal" to itself, see [7]. Note, also that $\lim_{t\to\infty} m(x, x, t) = 1$ for every $x \in X$.

Remark 3. The inequality $m(x, x, t) \geq m(x, y, t)\ \forall x, y \in X$, that seems natural and is important for the work, does not follow from the rest of the axioms. Therefore we include it as a part of axiom (2φFPM). However, in case of the minimum t-norm it can be proved, referring to axioms (4FPM) and (3FPM), as follows: $m(x, x, t) \geq m(x, y, t) \wedge m(y, x, t) = m(x, y, t) \wedge m(x, y, t) = m(x, y, t)$.

Patterned after terminology used in the theory of fuzzy metric spaces we introduce the corresponding "φ-versions".

Definition 3 (cf [12])**.** *A fuzzy φ-pseudometric is called fuzzy ultra φ-pseudometric if for all $x, y, z \in X$, $t, s \in \mathbb{R}^+$ $m(x, y, t + s) \geq \min\{m(x, z, t), m(z, y, s)\}$.*

Definition 4 *(cf e.g.* [5,13]*). A fuzzy φ-pseudometric m on X is called strong if it satisfies the following stronger modification of axiom* (4FPM)*:*

(4^{s}FPM) $m(x, z, t) \geq m(x, y, t) * m(y, z, t)$ *for all* $x, y, z \in X$ *and for all* $t > 0$.

We justify this definition showing that (4^{s}FPM) is indeed stronger than (4FPM):

Proposition 1. *If $m : X \times X \times \mathbb{R}^+ \to (0, 1]$ satisfies axioms (1FPM), (2φFPM), (3FPM), (4^{s}FPM) and (5φFPM), then it is a fuzzy φ-pseudometric.*

Proof. Referring to axioms (4^sFPM) and (5φFPM) we get the following series of inequalities: $m(x,z,t+s) \geq m(x,y,t+s)*m(y,z,t+s) \geq m(x,y,t)*m(y,z,s)$, which hold for any $x,y,z \in X$ and any $t,s \in \mathbb{R}^+$. $\square$

Example 1. Let $\varphi(t) = 1$ for all $t \in \mathbb{R}^+$. Then the fuzzy φ-pseudometric is just the fuzzy pseudometric in the sense of George and Veeramani.

Example 2. Important fuzzy φ-pseudometric is defined by function $\varphi(t) = \frac{t}{1+t}$. It is used in the definition of fragmentary fuzzy pseudometrics considered in Sect. 6. By setting $\varphi_k(t) = \frac{t}{k+t}$ for $k > 0$ we obtain a parametrized family $\{\varphi_k(t) = \frac{t}{k+t} : k > 0\}$ of φ-pseudometrics.

We modify the construction of the standard fuzzy metric given in [2] to obtain a certain "optimal" fuzzy φ-pseudometric from a given pseudometric $d : X \times X \to \mathbb{R}^+$ and a fixed function $\varphi : \mathbb{R}^+ \to (0,1]$ such that $\lim_{t\to+\infty} \varphi(t) = 1$.

Proposition 2. *Let $d : X \times X \to \mathbb{R}^+$ be a pseudometric and $\varphi : (0,\infty) \to [0,1]$ be a non-decreasing function such that $\lim_{t\to\infty} \varphi(t) = 1$. Then by setting $m(x,y,t) = \frac{t\cdot\varphi(t)}{t+d(x,y)}$ we obtain a strong fuzzy φ-pseudometric for the product t-norm, and hence also for any weaker continuous t-norm.*

Proof. The validity of axioms (1FPM), (3FPM) and (5FPM) is obvious. To show axiom (2FPM) notice that for every $t > 0$ $m(x,x,t) = \frac{t\cdot\varphi(t)}{t+d(x,x)} = \varphi(t)$. Finally, to show axiom (4^sFPM) notice that the inequality $\frac{t\cdot\varphi(t)}{t+d(x,y)} \cdot \frac{t\cdot\varphi(t)}{t+d(y,z)} \leq \frac{t\cdot\varphi(t)}{t+d(x,z)}$ is equivalent to the following obvious inequality $t\cdot\varphi(t)\cdot(t+d(x,z)) \leq t^2 + t\cdot d(x,y) + t\cdot d(y,z) + d(x,y)\cdot d(y,z)$. $\square$

One can easily prove the following modification of Proposition 2 in case when the original pseudometric d is an ultra-pseudometric:

Proposition 3. *Let $d : X \times X \to \mathbb{R}^+$ be an ultra pseudometric and $\varphi : (0,\infty) \to [0,1]$ be a non-decreasing function such that $\lim_{t\to\infty} \varphi(t) = 1$. Then by setting $m(x,y,t) = \frac{t\cdot\varphi(t)}{t+d(x,y)}$ we obtain a strong fuzzy φ-pseudometric for the minimum t-norm, and hence also for any other t-norm.*

By revising definition of continuity for mappings of fuzzy pseudometric spaces [2], we come to the following

Definition 5. *A mapping $f : (X, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ is called continuous if for every $\varepsilon \in (0, \varphi(t))$, every $x \in X$ and every $t \in \mathbb{R}^+$ there exist $\delta \in (0, \varphi(t))$ and $s \in \mathbb{R}^+$ such that $m_2(f(x), f(y), t) > \varphi_2(t) - \varepsilon$ whenever $m_1(x,y,s) > \varphi_1(s) - \delta$.*

Noticing that the composition $g \circ f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_3, m_3, *_{m_3}, \varphi_3)$ of two continuous mappings $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$, $g : (X_2, m_2, *_{m_2}, \varphi_2) \to (X_3, m_3, *_{m_3}, \varphi_3)$ is continuous and that the identity mapping $id : (X, m, *, \varphi) \to (X, m, *, \varphi)$ is continuous, we get

Proposition 4. *Fuzzy φ-pseudometric space as objects and their continuous mappings as morphisms form a category* $\mathbf{F}\varphi\mathbf{PMS}$.

For strong fuzzy φ-pseudometrics the following stronger version of continuity will be useful:

Definition 6 (cf [4]). *A mapping* $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ *is strongly continuous at a point* $x \in X_1$ *if given* $\varepsilon \in (0, \varphi_2(t))$ *and* $t \in \mathbb{R}^+$ *there exists* $\delta \in (0, \varphi_1(t))$ *such that* $m_1(x, y, t) > \varphi_1(t) - \delta$ *implies* $m_2(f(x), f(y), t) > \varphi_2(t) - \varepsilon$. *A mapping* $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ *is strongly continuous if it is strongly continuous at each point* $x \in X$.

3 Topological Structure of a Fuzzy φ-pseudometric Space

Let $m : X \times X \times \mathbb{R}^+ \to (0, 1]$ be a fuzzy φ-pseudometric. Given a point $x \in X$, $t \in \mathbb{R}^+$ and $\varepsilon \in (0, \varphi(t))$, we define the ball with center x, at the level t and radius ε as follows: $B(x, \varepsilon, t) = \{y \mid m(x, y, t) > \varphi(t) - \varepsilon\}$. It is clear that $t \leq s \Longrightarrow B(x, \varepsilon, t) \subseteq B(x, \varepsilon, s)$ and $\varepsilon \leq \delta \Longrightarrow B(x, \varepsilon, t) \subseteq B(x, \delta, t)$.

In [2], [3] for *a fuzzy pseudometric* it is proved that the family $\mathcal{B} = \{B(x, \varepsilon, t) \mid x \in X, t \in (0, \infty), \varepsilon \in (0, 1)\}$ satisfies necessary conditions to be a base for some topology T_m on X and just with this topology the space (X, m) is considered. Unfortunately, we do not have the analogous theorem in case of fuzzy φ-pseudometrics. The problem is that one cannot guarantee that for every $y \in B(x, \varepsilon, t)$ there exists a ball $B(y, \delta, s)$ such that $B(y, \delta, s) \subseteq B(x, \varepsilon, t)$. Therefore the collection of all balls $\mathcal{B} = \{B(x, \varepsilon, t) \mid x \in X, t \in (0, \infty), \varepsilon \in (0, \varphi(t)\}$ generally does not satisfy the criteria to be a base for a topology. However, we use family $\mathcal{B}$ to construct two topological-type structures: a supratopology σ_m and a topology τ_m, that can be useful for the study of fuzzy φ-pseudometric spaces.

3.1 Supratopology σ_m induced by a fuzzy φ-pseudometric m

Let (X, m) be a fuzzy φ-metric space and let $\mathcal{B}_m$ be the collection of all open balls. Further, let σ_m be the family of all unions of balls from $\mathcal{B}_m$, that is

$$\sigma_m = \{U \subseteq X : \exists B_m(a_i, \varepsilon_i, t_i), i \in I \text{ such that } U = \bigcup_{i \in I} B_m(a_i, \varepsilon_i, t_i)\}.$$

The family σ_m is obviously a supratopology, that is σ_m is invariant under taking arbitrary unions, contains X and contains $\emptyset$ (as the union of the empty family of balls). However, it may fail to be a topology, due to the reasons explained above.

Below are some special cases when σ_m is closed under finite intersections, that is σ_m indeed is a topology.

Theorem 1 [2]. *Let* $\varphi(t) = 1$ *for all* $t \in (0, \infty)$. *Then for every* $B(x, \varepsilon, t)$ *and every* $y \in B(x, \varepsilon, t)$ *there exists a ball* $B(y, \delta, s)$ *such that* $B(y, \delta, s) \subseteq B(x, \varepsilon, t)$. *Thus* $\mathcal{B}$ *satisfies the criteria of a base of a topology and hence* σ_m *is a topology.*

Theorem 2. *If* $m : X \times X \times \mathbb{R}^+ \to (0, 1]$ *is a strong fuzzy ultra* φ*-pseudometric, then* $B(y, \varepsilon, t) \subseteq B(x, \varepsilon, t)$ *for every* $y \in B(x, \varepsilon, t)$ *and hence* σ_m *is a topology.*

Proof. Let $y \in B(x, \varepsilon, t)$, then $m(x, y, t) > \varphi(t) - \varepsilon$. Now, let $z \in B(y, \varepsilon, t)$. Then $m(x, z, t) \geq m(x, y, t) \wedge m(y, z, t) \geq (\varphi(t) - \varepsilon) \wedge (\varphi(t) - \varepsilon)$ and hence $z \in B(x, \varepsilon, t)$. □

One can easily verify the following two propositions:

Proposition 5. *If (X, m) is a fuzzy φ-metric space, then the induced supratopology σ_m is Hausdorff, that is for any two different points $x_1, x_2 \in X$ there exist $t > 0$ and $\varepsilon \in (0, \varphi(t))$ such that $B(x_1, \varepsilon, t) \cap B(x_2, \varepsilon, t) = \emptyset$.*

Proposition 6. *A mapping $f : (X_1, m_1, *_{m_1}, \varepsilon_1) \to (X_2, m_2, *_{m_2}, \varepsilon_2)$ is continuous (in the sense of Definition 5) if and only if the mapping of the corresponding supratopological spaces $f : (X_1, \sigma_{m_1}) \to (X_2, \sigma_{m_2})$ is continuous (that is $\forall V \in \sigma_{m_2} \Rightarrow f^{-1}(V) \in \sigma_{m_1}$).*

3.2 Topology τ_m Induced by a Fuzzy φ-pseudometric m

An alternative point of view on the topological structure of a fuzzy φ-pseudometric space is given in the following definition:

Definition 7. *Given a fuzzy φ-pseudometric space, a subset $U \subset X$ is called open if for every $x \in U$ there exists $B(x, r, t)$ such that $B(x, r, t) \subseteq U$.*

One can easily see that τ_m is a topology. From the definitions it is clear that $\tau_m \subseteq \sigma_m$. However, note that an open ball need not be open in this space.

3.3 Subsets of a Fuzzy φ-pseudometric Spaces

3.3.1 Compactness

Having two topological structures induced by a fuzzy φ-pseudometric, a supratopology σ_m and a topology τ_m, we must deal with two versions of compactness in a fuzzy φ-pseudometric space: σ_m-compactness and τ_m-compactness. Since $\tau_m \subseteq \sigma_m$, a σ_m-compact space is τ_m-compact. However, we do not know whether the converse is true. It is easy to notice that a compact subset in a fuzzy φ-pseudometric space is closed in the supratopology σ_m.

3.3.2 Boundedness

As different from compactness, definition of boundedness does not depend upon the choice of the induced topological structure.

Definition 8. *A set $A \subseteq X$ is called bounded if there exist $t > 0$ and $r \in (0, \varphi(t))$ such that $m(x, y, t) > \varphi(t) - r$ for all $x, y \in A$. A set $A \subseteq X$ is called bounded on a level t or t-bounded if there exists $r \in (0, \varphi(t))$ such that $m(x, y, t) > \varphi(t) - r$ for all $x, y \in A$. A set $A \subseteq X$ is called strongly bounded if it is bounded on every level $t > 0$.*

Patterned after the proof of Theorem 3.9 in [2] we easily get

Theorem 3. *σ_m-compact subsets of a fuzzy φ-pseudometric space are strongly bounded.*

3.3.3 Dense Subsets

Concerning density, being a topological property, we again have to distinguish two cases.

Definition 9. *A subset A of a fuzzy φ-pseudometric space (X, m) is called σ-dense if it is dense in σ_m that is if each open ball B(a,r,t) has a nonempty intersection with A: $A \cap B(a,r,t) \neq \emptyset \ \forall a \in X, \forall r \in (0, \varphi(t)),\ \forall t > 0$. A subset A is called τ-dense if it is dense in topology τ_m*

3.3.4 Closed Balls

By a closed ball with center x_0 at the level $t \in (0, +\infty)$ and radius $r \in (0, \varphi(t))$ we call the set $\bar{B}(x_0, t, r) = \{x \in X : m(x, x_0, t) \geq r\}$. One can easily notice that $\bar{B}(x_0, t, r)$ is closed in the supratopology σ_m and that $r' < r \Rightarrow \bar{B}(x_o, r', t) \subset B(x_0, r, t)$

4 Sequences in Fuzzy φ-pseudometric Spaces

4.1 Three Types of Convergence in Fuzzy φ-pseudometric Spaces

Let (X, m) be a fuzzy φ-pseudometric space, $(x_n)_{n\in\mathbb{N}}$ be a sequence in this space and $x_0 \in X$. We say that $(x_n)_{n\in\mathbb{N}}$ σ-converges to x_0 and write $\lim^{\sigma}_{n\to\infty} = x_0$ if $(x_n)_{n\in\mathbb{N}}$ converges in the supratopology σ_m that is if for every open ball $B(x_0, r, t)$ there exists $n_0 \in \mathbb{N}$ such that $x_n \in B(x_0, r, t)$ for all $n \geq n_0$. Given a sequence $(x_n)_{n\in\mathbb{N}}$ and a point $x_0 \in X$, we say that $(x_n)_{n\in\mathbb{N}}$ τ-converges to x_0 and write $\lim^{\tau}_{n\to\infty} = x_0$ if it converges in τ_m, that is if for every open set U containing x_0 there exists $n_0 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq n_0$. Besides, the specificity of the topological structure induced by fuzzy φ-pseudometrics provokes us to introduce an "intermediate" version of convergence in such spaces. Namely, given a sequence $(x_n)_{n\in\mathbb{N}}$ and a point $x_0 \in X$, we say that $(x_n)_{n\in\mathbb{N}}$ $\tau\sigma$-converges to x_0 and write $\lim^{\tau\sigma}_{n\to\infty} = x_0$ if for every open set U containing x_0 and for every $B(x_0, r, t) \subseteq U$ there exists $n_0 \in \mathbb{N}$ such that $x_n \in B(x_0, r, t)$ for all $n \geq n_0$.

From the definitions one can easily may be convinced in the following

Theorem 4. *If $\lim^{\sigma}_{n\to\infty} = x_0$, then $\lim^{\tau\sigma}_{n\to\infty} = x_0$, and if $\lim^{\tau\sigma}_{n\to\infty} = x_0$, then $\lim^{\tau}_{n\to\infty} = x_0$.*

Theorem 5. *Let (X, m) be a fuzzy φ-pseudometric space and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in this space. Then $\lim^{\sigma}_{n\to\infty} x_n = a$ if and only if $\lim_{n\to\infty} m(a, x_n, t) = \varphi(t)$ for each $t \in (0, \infty)$.*

Proof. Assume that $\lim^{\sigma}_{n\to\infty} x_n = a$. Given $t \in (0, \infty)$ and $r \in (0, \varphi(t))$, let $B(a, r, t)$ be the corresponding ball. Then we can choose $n_0 \in \mathbb{N}$ such that $x_n \in B(a, r, t)$ for all $n \geq n_0$, and hence $m(a, x_n, t) > \varphi(t) - r$ for all $n \geq n_0$. Since r and t were taken arbitrary, we conclude that $\lim_{n\to\infty} m(a, x_n, t) = \varphi(t)$ for every $t > 0$. Since, on the other hand, $m(x, x_n, t) \leq \varphi(t)$, we conclude, that $\lim_{n\to\infty} m(x, x_n, t) = \varphi(t)$.

Assume now that $\lim^{\sigma}_{n\to\infty} x_n \neq a$. Then there exist a ball $B(a,r,t)$ such that $x_n \notin B(a,r,t)$ for infinitely many $n \in \mathbb{N}$. However, this means that $m(a,x_n,t) \leq \varphi(t) - r$ for infinitely many $n \in \mathbb{N}$, and hence either $\lim_{n\to\infty} m(a,x_n,t) \neq \varphi(t)$, or $\lim_{n\to\infty} m(a,x_n,t)$ does not exist. □

From here and applying Theorem 4 we get

Corollary 1. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a fuzzy φ-pseudometric space (X,m). If $\lim_{n\to\infty} m(a,x_n,t) = \varphi(t)$ for a point $a \in X$, then $\lim^{\sigma}_{n\to\infty} x_n = \lim^{\tau}\sigma_{n\to\infty} x_n = \lim^{\tau}_{n\to\infty} x_n = a$.*

4.2 Completeness of Fuzzy φ-pseudometric Spaces

Definition 10. *A sequence $(x_n)_{n\in\mathbb{N}}$ in a fuzzy φ-pseudometric space is called a Cauchy sequence if for each $t \in (0,\infty)$ and $\varepsilon \in (0,\varphi(t))$ there exists $n_0 \in \mathbb{N}$ such that $m(x_n,x_m,t) > \varphi(t) - \varepsilon$ for all $n,m \geq n_0$.*

One can easily notice that a σ-convergent sequence is Cauchy. A fuzzy φ-pseudometric space is called complete if every Cauchy sequence in it σ-converges.

Example 3. One can easily see that a sequence (x_n) converges in the standard fuzzy φ-pseudometric $m_d(x,y,t) = \frac{t\cdot\varphi(t)}{t+d(x,y)}$ if and only if it converges in the underlying crisp metric $d(x,y)$. Hence a pseudometric space (X,d) is complete if and only if the fuzzy φ-pseudometric (X,m_d) space is complete.

4.3 Fuzzy φ-pseudometric Version of a Baire Theorem

Impossibility to use intersection axiom for open sets in supratopology σ_m on one hand, and the probable "non-openness" of open balls in topology τ_m make it doubtful to get a full-bodied version of the Baire Category theorem, neither in σ_m nor in τ_m. To overcome this obstacle, we introduce the concept of a valuably open set and with its help get a certain restricted version of Baire category theorem. An open set U of the space (X,σ_m) is called *valuably open* if for every ball $B(x_0,r,t)$ having non-empty intersection with U there exists a ball $B(x_1,r_1,t_1) \subseteq B(x_0,r,t) \cap U$ for some $x_1 \in B(x_0,r,t), r_1 \in (0,\varphi(t_1))$ and $t_1 > 0$.

Theorem 6. *Let (X,m) be a complete fuzzy-φ-pseudometric space. The intersection of a countable family of dense valuably open sets in the supratopological space (X,σ_m) is dense.*

Proof. Let (X,m) be a fuzzy φ-pseudometric space and $D_1 \supseteq D_2 \supseteq D_3 \supseteq \dots D_n \dots$ be a sequence of valuably open dense subsets of this space. Further, let $U \subseteq X$ be an open subset of X. We have to prove that $U \cap (\bigcap_n D_n) \neq \emptyset$. Referring to Sect. 3.3.4 we choose an open ball $B_0 = B(x_0,r_0,t_0)$ such that $\bar{B}(x_0,r_0,t_0) \subseteq U$. Without loss of generality we may assume that $r_0 < 1, t_0 < 1$. Since the set D_1 is dense, $D_1 \cap B_0 \neq \emptyset$, and since D_1 is valuably open, we can find $B(x_1,r_1,t_1) = B_1$ such that $\bar{B}(x_1,r_1,t_1) \subseteq D_1 \cap B_0$. Without loss of generality we assume that $r_1 < \frac{1}{2}$, $t_1 < \frac{1}{2}$. Since the set D_2 is dense, $D_2 \cap B_1 \neq \emptyset$, and we can find $B(x_2, r_2, t_2) = B_2$ such that $\bar{B}(x_2, r_2, t_2) \subseteq D_2 \cap B_1$. Without loss of generality we assume that $r_2 \leq \frac{1}{3}$, $t_2 \leq \frac{1}{3}$. We continue such procedure by induction and in the result obtain a sequence of points $x_0, x_1, x_2, \ldots, x_n, \ldots$ and a sequence of open balls $B_0(x_0, r_0, t_0) \supseteq B_1(x_1, r_1, t_1) \supseteq B_2(x_2, r_2, t_2) \supseteq \ldots \supseteq B_n(x_n, r_n, t_n) \ldots$, where $r_n \leq \frac{1}{n+1}$, $t_n \leq \frac{1}{n+1}$.

We claim that the constructed sequence $x_0, x_1, x_2, \ldots, x_n, \ldots$ is Cauchy. Indeed let $t > 0$ and $\varepsilon > 0$ be given. First, by continuity of the t-norm, find $\delta \in (0,1)$ such that $(1-\delta) * (1-\delta) \geq 1 - \varepsilon$. Further, find $n_0 \in \mathbb{N}$ such that $\frac{1}{n_0} < t$ and $\frac{1}{n_0} < \delta$. Then for $n, k \geq n_0$ we have $m(x_n, x_k, t) \geq m(x_{n_0}, x_n, t) * m(x_{n_0}, x_k, t) \geq (1-\delta) * (1-\delta) \geq 1 - \varepsilon \quad \forall n, k \geq n_0$. Thus the sequence $x_0, x_1, x_2, \ldots, x_n, \ldots$ is Cauchy. Since the fuzzy φ-pseudometric space (X, m) is complete this sequence σ-converges.

Let $\lim^{\sigma}_{n\to\infty} x_n = a$. Take some $n \in \mathbb{N}^+$. Since all elements x_k of this sequence for $k \geq n$ are contained in the closed ball $\bar{B}_n = \bar{B}(x_n, r_n, t_n)$, we conclude that the limit is contained in $\bar{B}(x_n, r_n, t_n) \cap D_{n-1}$ for all $n \geq 1$. Hence $U \cap (\bigcap_n D_n) \neq \emptyset$, that is $\bigcap_n D_n$ is dense in X.

4.4 Sequentiality Properties of Fuzzy φ-pseudometric Spaces

Recall that a topological space (X, T) is called sequential if its subset A is closed whenever it contains the limits of all convergent sequence lying in this subset. It is well-known and easy to prove, that each metric space is sequential. We extend the concept of sequentially to the case of supratopological spaces and show here that the supratopology induced by a fuzzy φ-pseudometric is sequential.

Theorem 7. *Let (X, m) be a fuzzy φ-pseudometric space. Then the induced supratopology σ_m is sequential.*

Proof. Assume that A is a subset of the space (X, m) which is not closed. Then its complement is not open and hence there exists a point $a \in X \setminus A$ such that for every $t > 0$ and every $r \in (0, \varphi(t))$ it holds $B(a, r, t) \cap A \neq \emptyset$. We fix t and for every $n \in \mathbb{N}$ choose a point $x_n \in B(a, \frac{1}{n}, t) \cap A$. From the construction it is clear that $\{x_n : n \in \mathbb{N}\} \subseteq A$ and $\lim^{\sigma}_{n\to\infty} x_n = a \notin A$. The obtained contradiction completes the proof. □

5 Uniform Continuity for Mappings of Fuzzy φ-pseudometric Spaces

For a mapping of fuzzy φ-pseudometric spaces in a natural way we define the property of strong uniform continuity, cf Definition 3.3. in [6].

Definition 11. *A mapping $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ is called strongly uniformly continuous if for every $t \in \mathbb{R}^+$ $\varepsilon \in (0, \varphi_2(t))$ there exists $\delta \in (0, \varphi_1(t))$ such that $m_1(x, y, t) > \varphi_1(t) - \delta$ implies $m_2(f(x), f(y), t) > \varphi_2(t) - \varepsilon$ for all $x, y \in X_1$.*

Proposition 7. *A mapping $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ is strongly uniformly continuous if and only if for every $t \in \mathbb{R}^+$ and every $\beta > 0$ there exists $\alpha > 0$ such that $\frac{1}{m_1(x,y,t)} - \varphi_1(t) < \alpha$ implies $\frac{1}{m_2(f(x),f(y),t)} - \varphi_2(t) < \beta$.*

Proof. Notice that the definition of a uniform continuity of the mapping $f : (X_1, m_1, *_{m_1}, \varphi_1) \to (X_2, m_2, *_{m_2}, \varphi_2)$ means that for every $t \in \mathbb{R}^+$ and for all $\varepsilon \in (0, \varphi_2(t))$ there exists $\delta \in (0, \varphi_1(t))$ such that $m_1(x, y, t) > \varphi_1(t) - \delta \Rightarrow m_2(f(x), f(y), t) > \varphi_2(t) - \varepsilon \forall x, y \in X_1$. The last inequality can be rewritten as $\frac{1}{m_1(x,y,t)} < \frac{1}{\varphi_1(t)-\delta} \Rightarrow \frac{1}{m_2(f(x),f(y),t)} < \frac{1}{\varphi_2(t)-\varepsilon} \forall x, y \in X_1$. Now we can reformulate the definition of uniform continuity as follows: for every $t \in \mathbb{R}^+$ and for all $\varepsilon \in (0, \varphi_2(t))$ there exists $\delta \in (0, \varphi_1(t))$ such that $\frac{1}{m_1(x,y,t)} < 1 + \frac{\delta}{\varphi_1(t)-\delta} := \alpha \Rightarrow \frac{1}{m_2(f(x),f(y),t)} < 1 + \frac{1}{\varphi_2(t)-\varepsilon} := \beta \ \forall x, y \in X_1$. To complete the proof notice that equalities $\frac{\delta}{\varphi_1(t)-\delta} = \alpha$ and $\frac{1}{\varphi_2(t)-\varepsilon} = \beta$ establish bijections between $(0, \varphi_1(t))$ and $\mathbb{R}^+$ and $(0, \varphi_2(t))$ and $\mathbb{R}^+$ respectively. □

Definition 12. *Let $(X, m, *, \varphi)$ be a fuzzy φ-metric space. A mapping $f : X \to X$ is called contractive, if there exists $k \in (0, 1)$ such that $\frac{1}{m(f(x),f(y),t)} - \varphi(t) \leq k\left(\frac{1}{m(x,y,t)} - \varphi(t)\right)$*

One can easily notice that a contractive mapping $f : X \to X$ is uniformly continuous. In our further research we plan to study the problem of existence and uniqueness of a fixed point for certain contractive mappings.

6 Fuzzy Ultra φ-metric on the Set of Infinite Words

In [1] we defined a fragmentary fuzzy metric on the set of infinite words as a useful tool for the study of problems in words combinatorics. In this section we show how the fragmentary fuzzy metric considered in [1] can be viewed as a fuzzy φ-metric for a special chosen mapping φ.

Let X be the set of infinite words. We define a sequence $\{d_n \mid n \in \mathbb{N} \cup \{0\}\}$ of ultra pseudometrics on X as follows. Let $x = (x_0, x_1, x_2, \ldots), y = (y_0, y_1, y_2, \ldots) \in X$ and let $\chi_i(x, y) = 0$ if $x_i = y_i$ and $\chi_i(x, y) = 1$ if $x_i \neq y_i$. We define: $d_0(x, y) = \chi_0(x, y)$; $d_1(x, y) = \chi_0(x, y) + \frac{\chi_1(x,y)}{2}$; $d_2(x, y) = \chi_0(x, y) + \frac{\chi_1(x,y)}{2} + \frac{\chi_2(x,y)}{2^2}$; $\ldots$ $d_n(x, y) = \sum_{i=0}^{n} \frac{\chi_i(x,y)}{2^i}$; $\ldots$

Proposition 8. *Every d_n is an ultra pseudometric.*

Proof. Obviously every $\frac{\chi_i(x,y)}{2^i}$ is an ultra pseudometric. From here we conclude that every $d_n(x, y)$ is an ultra pseudometric by induction referring to the following easily provable Lemma:

Lemma 1. *Let $d_1, d_2 : X \times X \to \mathbb{R}^+$ be ultra pseudometrics. Assume that $d_1(x, y) \in \{0\} \cup [a, 1]$ for any $x, y \in X$ and that $d_2(x, y) \in [0, \frac{a}{2}]$. Then $d = d_1 + d_2 : X \times X \to [0, 1]$ is an ultra pseudometric.*

Basing on this sequence of ultra pseudometrics we construct the sequence of mappings on the set of all infinite words:
$\mu_0(x,y,t) = \frac{t}{t+1+d_0(x,y)}$, $\mu_1(x,y,t) = \frac{t}{t+1+d_1(x,y)}$, $\mu_2(x,y,t) = \frac{t}{t+1+d_2(x,y)}$, ...,
$\mu_n(x,y,t) = \frac{t}{t+1+d_n(x,y)}$, ...

Further, we define the following family of mappings:
$m_0(x,y,t) = \mu_0(x,y,t)$; $m_1(x,y,t) = \mu_1(x,y,t) \vee \mu_0(x,y,1)$; $m_2(x,y,t) = \mu_2(x,y,t) \vee \mu_1(x,y,2)$; ...; $m_n(x,y,t) = \mu_n(x,y,t) \vee \mu_{n-1}(x,y,n)$; ...

From the construction it is clear that $m_n(x,z,t) \geq m_n(x,y,t) \wedge m_n(y,z,t)$, $m_n(x,y,t) \leq m_{n+1}(x,y,t)$, $m_n(x,y) \leq m_n(x,x)$ and $m_n(x,x) = \frac{n}{n+1}$ for all $x, y \in X$, $t \in \mathbb{R}^+$ and $n \in \mathbb{N}$.

Finally, we construct a mapping $\mathfrak{m} : X \times X \times \mathbb{R}^+ \to [0,1]$ as follows:

$$\mathfrak{m}(x,y,t) = \begin{cases} m_0(x,y,t) & \text{if } 0 < t \leq 1 \\ m_1(x,y,t) & \text{if } 1 < t \leq 2 \\ m_2(x,y,t) & \text{if } 2 < t \leq 3 \\ & \dots \\ m_n(x,y,t) & \text{if } n < t \leq n+1 \\ & \dots \end{cases}$$

We define mapping $\varphi : (0,\infty) \to (0,1]$ by setting $\varphi_f(t) = \frac{t}{t+1}$. Obviously, φ_f is non-decreasing and $\lim_{t\to+\infty} = 1$. Now, from the construction of the mapping $\mathfrak{m}$ one easily get

Proposition 9. $\mathfrak{m} : X \times X \times \mathbb{R}^+ \to [0,1]$ *is a fuzzy strong ultra* φ_f*-metric.*

We illustrate the shape in the initial interval $(0,3]$ of the fuzzy metric $\mathfrak{m}$ describing the distance between infinite words $x = (x_0x_1x_2...)$ and $y = (y_0y_1, y_2...)$ on dependence of the values x_0, x_1, x_2, y_0, y_1, and y_2.

1. The case $x_0 = y_0, x_1 = y_1, x_2 = y_2$. Then $\mathfrak{m}(x,y,t) = \frac{t}{t+1}$ for $t \in (0,3]$.
2. The case $x_0 = y_0, x_1 = y_1, x_2 \neq y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+1} & \text{if } 0 < t \leq 2 \\ \frac{2}{3} & \text{if } 2 < t \leq \frac{5}{2} \\ \frac{t}{t+\frac{5}{4}} & \text{if } \frac{5}{2} < t \leq 3 \end{cases}$$

3. The case $x_0 = y_0, x_1 \neq y_1, x_2 = y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+1} & \text{if } 0 < t \leq 1 \\ \frac{1}{2} & \text{if } 1 < t \leq \frac{3}{2} \\ \frac{t}{t+\frac{3}{2}} & \text{if } \frac{3}{2} < t \leq 3 \end{cases}$$

4. The case $x_0 = y_0, x_1 \neq y_1, x_2 \neq y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+1} & \text{if } 0 < t \leq 1 \\ \frac{1}{2} & \text{if } 1 < t \leq \frac{3}{2} \\ \frac{t}{t+\frac{3}{2}} & \text{if } \frac{3}{2} < t \leq 2 \\ \frac{4}{7} & \text{if } 2 < t \leq \frac{7}{3} \\ \frac{t}{t+\frac{7}{4}} & \text{if } \frac{7}{3} < t \leq 3 \end{cases}$$

5. The case $x_0 \neq y_0, x_1 = y_1, x_2 = y_2$. Then $m(x, y, t) = \frac{t}{t+2}$ for $t \in (0, 3]$.
6. The case $x_0 \neq y_0, x_1 = y_1, x_2 \neq y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+2} & \text{if } 0 < t \leq 2 \\ \frac{1}{2} & \text{if } 2 < t \leq \frac{9}{4} \\ \frac{t}{t+\frac{9}{4}} & \text{if } \frac{9}{4} < t \leq 3 \end{cases}$$

7. The case $x_0 \neq y_0, x_1 \neq y_1, x_2 = y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+2} & \text{if } 0 < t \leq 1 \\ \frac{1}{3} & \text{if } 1 < t \leq \frac{5}{4} \\ \frac{t}{t+\frac{5}{2}} & \text{if } \frac{5}{2} < t \leq 3 \end{cases}$$

8. The case $x_0 \neq y_0, x_1 \neq y_1, x_2 \neq y_2$. Then

$$\mathfrak{m}(x,y,t) = \begin{cases} \frac{t}{t+2} & \text{if } 0 < t \leq 1 \\ \frac{1}{3} & \text{if } 1 < t \leq \frac{5}{4} \\ \frac{t}{t+\frac{5}{2}} & \text{if } \frac{5}{4} < t \leq 2 \\ \frac{4}{9} & \text{if } 2 < t \leq \frac{11}{5} \\ \frac{t}{t+\frac{11}{4}} & \text{if } \frac{11}{5} < t \leq 3 \end{cases}$$

Acknowledgment. The authors are grateful to the anonymous referees for pointing out some misprints and other minor defects noticed in the first version of the paper.

References

1. Bēts, R., Šostak, A.: Fragmentary fuzzy pseudometrics: basics of the theory and applications in combinatorics of words. Baltic J. Modern Comput. **4**, 826–845 (2016)
2. George, A., Veeramani, P.: On some results in fuzzy metric spaces. Fuzzy Sets Syst. **64**, 395–399 (1994)
3. George, A., Veeramani, P.: On some results of analysis for fuzzy metric spaces. Fuzzy Sets Syst. **90**, 365–368 (1997)
4. Gregori, V., López-Crevillén, A., Morillas, S.: On continuity and uniform continuity in fuzzy metric spaces. In: Proceedings of the Workshop in Applied Topology WiAT 2009, pp. 85–91 (2009)
5. Gregori, V., Morillas, S., Sapena, A.: On a class of completable fuzzy metric spaces. Fuzzy Sets Syst. **161**, 2193–2205 (2010)
6. Gregori, V., Morillas, S., Sapena, A.: On fixed-point theorems in fuzzy metric spaces. Fuzzy Sets Syst. **125**, 245–252 (2002)
7. Höhle, U.: M-valued sets and sheaves over integral commutative cl-monoids. In: Rodabaugh, S.E., Klement, E.P., Höhle, U. (eds.) Applications of Category Theory to Fuzzy Sets, pp. 33–72. Kluwer Academic Press, Dordrecht (1992)
8. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
9. Kramosil, I., Michalek, J.: Fuzzy metrics and statistical metric spaces. Kybernetika **11**, 336–344 (1975)

10. Mashhourr, A.S., Allam, A.A., Mahmoud, F.S., Khadr, F.H.: On supratopological spaces. Indian J. Pure Appl. Math. **14**, 502–510 (1983)
11. Menger, K.: Probabilistic geometry. Proc. N.A.S. **27**, 226–229 (1951)
12. Sapena, A.: A contribution to the study of fuzzy metric spaces. Appl. Gen. Topol. **2**, 63–76 (2001)
13. Sapena, A., Morillas, S.: On strong fuzzy metrics. In: Proceedings of the Workshop in Applied Topology WiAT 2009, pp. 135–141 (2009)

Generalized Net Modelling of the Intuitionistic Fuzzy Evaluation of the Quality Assurance in Universities

Evdokia Sotirova[1(✉)], Todor Petkov[1], and Maciej Krawczak[2]

[1] Intelligent Systems Laboratory,
University "Prof. Dr. Assen Zlatarov", Burgas, Bulgaria
{esotirova, todor_petkov}@btu.bg
[2] Polish Academy of Sciences Warsaw School of Information Technology,
Systems Research Institute, Warsaw, Poland
krawczak@ibspan.waw.pl

Abstract. In the paper is proposed a method for evaluation of the quality assurance in universities and scientific organizations. The evaluation of the quality is based on criteria, which measure different aspects of university activities and consists of different sub-criteria. For the assessment the theory of intuitionistic fuzzy sets is used. The obtained intuitionistic fuzzy estimations reflect the degree of each criterion' satisfaction, and non-satisfaction. We also consider a degree of uncertainty that represents such cases wherein is no information about sub-criteria of the current criterion. The generalized model gives possibility for algorithmization of the methodology of forming the quality evaluations is constructed. It provides the possibility for the algorithmization of the process of forming the evaluation of the quality assurance in universities.

Keywords: Generalized nets · Intuitionistic fuzzy sets · University quality

1 Introduction

In a series of research papers, the authors have studied some of the most important processes of functioning of universities, [5–12]. In particular, Generalized Nets, [1, 2], are used to describe the process of student assessment, [5, 9, 10] where the assessments can be represented in an intuitionistic fuzzy form. The concept of Intuitionistic Fuzzy Set was delivered in [3, 4].

The purpose of the present paper is to offer a generalized net model with intuitionistic fuzzy assessments of the process of quality assurance in universities and scientific organizations. The quality in higher education is related to teaching and learning, including the learning environment and relevant links to research and innovation [13]. The research is a continuation of previous investigations of the authors into the modelling of a basic processes and functions of a typical university.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_31

2 Proposed Assessment Model

The evaluation of the quality assurance in higher education institutions and scientific organizations is based on criteria, which measure different aspects of university activities including teaching staff and academic potential, policy for quality assurance, prestige, innovation, scientific activities and etc. The final assessment is provided in the range from 0 to usually 100 points.

2.1 Determination of the Criterion Assessment

Let us consider a group of n criteria, the criteria are labeled as follows $j = 1, 2, \ldots, n$. Every criterion consists of c_j sub-criteria $i_j = 1, 2, \ldots, c_j$.

The assessment, which estimates a summative account of the *j-th* criterion, is formed on the basis of a set of intuitionistic fuzzy estimations $\langle \mu^j, \nu^j \rangle$ of real numbers from the set $[0, 1] \times [0, 1]$, related to the respective sub-criteria. These intuitionistic fuzzy estimations reflect the degree of satisfaction μ^j, or non-satisfaction ν^j, for each criterion.

The degree of uncertainty $\pi = 1 - \mu - \nu$ represents such cases wherein the university can not provide information on the criterion. Within the paper the ordered pairs were defined in the sense of intuitionistic fuzzy sets.

Initially, when there has not been information obtained for the criterion' assessment, then the estimation $\langle \mu_0^j, \nu_0^j \rangle$ is given by the initial values $\langle 0, 0 \rangle$. For $k \geq 0$, the current (k)-st estimation of the j-th criterion is obtained on the basis of the previous estimations according to the recurrence relation involved in the following formula (1), $j = 1, 2, \ldots, n$.

$$\langle \mu_k^j, \nu_k^j \rangle = \left\langle \frac{(k-1).\mu_{k-1}^j + m_{i_j}^j}{k}, \frac{(k-1).\nu_{k-1}^j + n_{i_j}^j}{k} \right\rangle, \tag{1}$$

where:

- $\langle \mu_{k-1}^j, \nu_{k-1}^j \rangle$ is the previous estimation of the j-th criterion on the basis of the estimations of the already evaluated sub-criteria,
- $\left\langle m_{i_j}^j, n_{i_j}^j \right\rangle$ is the estimation of the i_j sub-criterion of the j-th criterion, for $m_{i_j}^j, n_{i_j}^j \in [0, 1]$, $m_{i_j}^j + n_{i_j}^j \leq 1$, and $j = 1, 2, \ldots, n$, $i_j = 1, 2, \ldots, c_j$.
- $m_{i_j}^j$ and $n_{i_j}^j$ are calculated according (2) and (3) in the following way:

$$m_{i_j}^j = \begin{cases} \frac{p_{i_j}}{p_{\max}^j}, \text{if the sub - criterion } i_j \text{ is described,} \\ 0, \text{ if it is no information about sub - criteria } i_j \end{cases}, \tag{2}$$

$$n_{i_j}^j = \begin{cases} \frac{p_{\max}^j - p_{i_j}}{p_{\max}^j}, \text{if the sub - criterion } i_j \text{ is described,} \\ 0, \text{ if it is no information about sub - criteria } i_j \end{cases}, \tag{3}$$

where:

- p_{i_j} are the points for the description of the criterion i_j,
- $p^j_{\max}$ is the maximal possible number of points for the criterion i_j.

Therefore, the degree of uncertainty, in this case, is equal to 1, when there is no information about sub-criteria i_j of the j-th criterion.

2.2 Determination of the Final Assessment for the University

The final assessment for the quality assurance in a university can be calculated according (4) in the following way:

$$\langle \mu, \nu \rangle = \left\langle \frac{\sum_{j=1}^{n} \mu^j}{n}, \frac{\sum_{j=1}^{n} \nu^j}{n} \right\rangle, \tag{4}$$

where:

- $\langle \mu^j, \nu^j \rangle$ is the estimation of the j-th criterion, $j = 1, 2, \ldots, n$,
- n is the number of criteria in criteria system.

3 Generalized Net Model

The GN-model with intuitionistic fuzzy estimations of the quality assurance in universities (see Fig. 1) contains four transitions and fifteen places, collected in two groups and related to the two types of the tokens that will enter respective types of places:

- α-tokens and c-places represent the criteria,
- β-tokens and e-places represent universities and their estimations.

For brevity, we shall use the notation α- and β-tokens instead of α_k- and β_l-tokens, where k and l are numerations of the respective tokens.

Initially the α- and β-tokens remain, respectively, in places c_3 and e_4 with initial characteristics:

x_0^α = “Current criteria system for quality assurance of higher schools”,

x_0^β = “Name and current status of the university u_k, $k = 1, 2, \ldots, m$”.

The new criterion/criteria system and new higher schools enter the net via place c_1 and e_1 respectively. These tokens have initial characteristics

“Criterion or new criteria system”

in place c_1, and

“Name and current status of a university”

in place e_1.

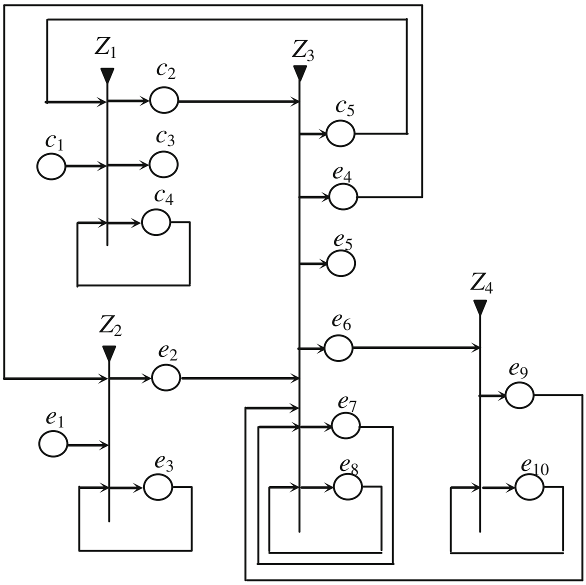


Fig. 1. GN model with intuitionistic fuzzy estimations of the quality assurance in universities

The GN contains the following set of transitions:

$$A = \{Z_1, Z_2, Z_3, Z_4\},$$

and they represent, respectively:

- Z_1 – The activities with criteria system;
- Z_2 – Determination of the university for evaluation;
- Z_3 – Process of evaluation of the criteria;
- Z_4 – Process of evaluation of the sub-criteria.

The forms of the transitions are the following.

$$Z_1 = \langle\{c_1, c_4, c_5\}, \{c_2, c_3, c_4\}, r_1, \vee(c_1, c_4, c_5)\rangle$$

where:

$$r_1 = \begin{array}{c|ccc} & c_2 & c_3 & c_4 \\ \hline c_1 & False & False & True \\ c_4 & W^c_{4,2} & W^c_{4,3} & True \\ c_5 & False & False & True \end{array},$$

$W^c_{4,2}$ = "There is a criteria system with criteria than will be used for assessment of the quality assurance";

and $W^c_{4,3}$ = "There is a rejected criteria system".

The α-token that enters place c_4 (from place c_1) unites with α-token (in place c_4). When the predicate $W^c_{4,2}$ has truth-value "True", the token α-token in place c_4 generates new α-token that enters place c_2 with characteristic

"Criteria system"

When the predicate $W^c_{4,3}$ has truth-value "True", the token α-token in place c_4 generates new α-token that enters place c_3 with characteristic

"Rejected criteria system".

$$Z_2 = \langle \{e_1, e_3, e_4\}, \{e_2, e_3\}, r_2, \vee(e_1, e_3, e_4) \rangle$$

where:

$r_2 =$

	e_2	e_3
e_1	*False*	*True*
e_3	$W^e_{3,2}$	*True*
e_4	*False*	*True*

,

and $W^e_{3,2}$ = "The university for evaluation is determined".

The β-token that enters place e_3 (from place e_1) unites with β-token (in place e_3). When the predicate $W^e_{3,2}$ has truth-value "True", the β-token in place e_3 generates new β-token that enters place e_2 with characteristic

"Name and current status of chosen university u_k".

$$Z_3 = \langle \{c_2, e_2, e_7, e_8, e_9\}, \{c_5, e_4, e_5, e_6, e_7, e_8\}, r_3, \vee(\wedge(c_2, e_2), e_7, e_8, e_9) \rangle$$

where:

$r_3 =$

	c_5	e_4	e_5	e_6	e_7	e_8
c_2	*False*	*False*	*False*	*False*	*False*	*True*
e_2	*False*	*False*	*False*	*False*	*False*	*True*
e_7	*False*	$W^e_{7,4}$	$W^e_{7,5}$	*False*	$W^e_{7,7}$	*False*
e_8	$W^c_{8,5}$	*False*	*False*	$W^e_{8,6}$	*False*	$W^e_{8,8}$
e_9	*False*	*False*	*False*	*False*	*True*	*False*

,

and

$W^e_{7,4} = W^e_{7,5} = W^c_{8,5}$ = "The final university's assessment on all criteria is evaluated",

$W^e_{7,7} = \neg\, W^e_{7,5}$,

$W^e_{8,6}$ = "The criterion j for assessment is chosen", for j = 1, 2, …, n,

$W^e_{8,8}$ = "There is at least one more evaluation criterion".

The α-token that enters place e_8 (from place c_2) do not obtains new characteristic. When the predicate $W^e_{8,5}$ has truth-value "True", the α-token from place e_8 enters place c_5 without new characteristic.

The β-token that enters place e_8 does not obtain new characteristic.

The β-token that enters place e_6 (from place e_8) obtains characteristic "Criterion j, university u_k", for $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$.

The β-tokens that enter places e_4 (from place e_7) and e_7 (from place e_9) do not obtain new characteristic.

The β-token that enters place e_5 (from place e_7) obtains characteristic "Final assessment according formula (4), university u_k", for $k = 1, 2, \ldots, m$.

$$Z_4 = \langle \{e_6, e_{10}\}, \{e_9, e_{10}\}, r_4, \vee(e_6, e_{10}) \rangle$$

where:

$$r_4 = \begin{array}{c|cc} & e_9 & e_{10} \\ \hline e_6 & False & True \\ e_{10} & W^e_{10,9} & W^e_{10,10} \end{array},$$

and $W^e_{10,9}$ = "The intuitionistic fuzzy estimations for all sub-criteria of the criterion j is evaluated",

$W^e_{10,10} = \neg W^e_{10,9}$.

The β-token from place that e_6 enters place e_{10} without new characteristic.

The β-token that enters place e_9 obtains characteristic "Intuitionistic fuzzy estimation for criterion j is evaluated according formula (1)".

4 Conclusion

In the present research the methodology of assessment the quality assurance in universities and scientific organizations with intuitionistic fuzzy estimations is given. The constructed generalized model gives possibility for algorithmization of the methodology of forming the quality evaluations related to learning and teaching in higher education, innovations, and research and governance activities.

Acknowledgments. The authors are thankful for the support provided by the Bulgarian National Science Fund under Grant Ref. No. DN-02/10 "*New Instruments for Knowledge Discovery from Data, and their Modelling*".

References

1. Atanassov, K.: Generalized nets. World Scientific, Singapore (1991)
2. Atanassov, K.: On Generalized Nets Theory. "Prof. Marin Drinov"Academic Publishing House, Sofia (2007)
3. Atanassov, K.T.: On Intuitionistic Fuzzy Sets Theory. Springer, Heidelberg (2012)
4. Atanassov, K.: Intuitionistic Fuzzy Sets. Springer, Heidelberg (1999)
5. Melo-Pinto, P., Kim, T., Atanassov, K., Sotirova, E., Shannon, A., Krawczak, M.: Generalized net model of e-learning evaluation with intuitionistic fuzzy estimations. In: Issues in the Representation and Processing of Uncertain and Imprecise Information, Warszawa, pp. 241–249 (2005)
6. Shannon, A., Langova-Orozova, D., Sotirova, E., Petrounias, I., Atanassov, K., Krawczak, M., Melo-Pinto, P., Kim, T.: Generalized Net Modelling of University Processes. KvB Visual Concepts Pty Ltd, Monograph No. 7, Sydney (2005)
7. Shannon, A., Atanassov, K., Sotirova, E., Langova-Orozova, D., Krawczak, M., Melo-Pinto, P., Petrounias, I., Kim, T.: Generalized Nets and Information Flow Within a University, Warsaw School of Information Technology, Warszawa (2007)
8. Shannon, A., Sotirova, E., Petrounias, I., Atanassov, K., Krawczak, M., Melo-Pinto, P., Kim, T.: Intuitionistic fuzzy estimations of lecturers' evaluation of student work. In: First International Workshop on Intuitionistic Fuzzy Sets, Generalized Nets & Knowledge Engineering, University of Westminster, London, pp. 44–47, 6–7 September 2006
9. Shannon, A., Sotirova, E., Petrounias, I., Atanassov, K., Krawczak, M., Melo-Pinto, P., Kim, T.: Generalized net model of lecturers' evaluation of student work with intuitionistic fuzzy estimations. In: Second International Workshop on Intuitionistic Fuzzy Sets, Banska Bystrica, Slovakia, 3 December 2006, Notes on IFS, vol. 12, No. 4, pp. 22–28 (2006)
10. Shannon, A., Sotirova, E., Atanassov, K., Krawczak, M., Melo-Pinto, P., Kim, T.: Generalized Net Model for the Reliability and Standardization of Assessments of Student Problem Solving with Intuitionistic Fuzzy Estimations, Developments in Fuzzy Sets, Generalized Nets and Related Topics. Applications, vol. 2, System Research Institute, Polish Academy of Science, pp. 249–256 (2008)
11. Shannon, A., Dimitrakiev, D., Sotirova, E., Krawczak, M., Kim, T.: Towards a Model of the Digital University: Generalized Net Model of a Lecturer's Evaluation with Intuitionistic Fuzzy Estimations, Cybernetics and Information Technologies, Bulgarian Academy of Sciences, vol. 9, No 2, pp. 69–78 (2009)
12. Shannon, A., Sotirova, E., Hristova, M., Kim, T.: Generalized net model of a student's course evaluation with intuitionistic fuzzy estimations in a digital university. Proc. Jangjeon Mathe. Soc. **13**(1), 31–38 (2010)
13. Standards and Guidelines for Quality Assurance in the European Higher Education Area. http://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf

How to Calibrate a Questionnaire for Risk Measurement?

Jana Špirková(✉) and Pavol Král'

Faculty of Economics, Matej Bel University, Banská Bystrica, Slovakia
{jana.spirkova,pavol.kral}@umb.sk

Abstract. Utility functions content parameters related to risk aversion coefficients which represent natural extensions of utility function properties. They measure how much utility we gain (or lose) as we add (or subtract) from our wealth. We set up these parameters for a person based on her/his answers to a questionnaire constructed to identify individual risk behavior. Calibration of such a questionnaire, and subsequently of utility functions, is based on an expected utility maximization of different alternatives of investment strategies. In the paper, we present questionnaire calibration methodology which we illustrate using absolute and relative risk aversion coefficients of two selected utility functions which have common, as well as different properties.

Keywords: Questionnaire · Utility function · Risk measurement · Calibration · Premium

1 Introduction

The term *utility* is the economist's way of measuring pleasure or happiness and how it relates to the decisions that people make. Utility measures the benefits (or drawbacks) from consuming goods or services or from working. Although utility is not directly measurable, it can be inferred from the decisions that people make.

Our interest in utility theory was primarily inspired by the monographs [6,9], where authors introduced a model for determination of maximal and minimal premium in non-life insurance using personal utility functions. An alternative approach to the utility theory can be found for example, in [4,5,10,11] or [14].

The most common way how to determine a personal utility function, are responses to a suitably constructed questionnaire. Choice under uncertainty is often characterized as the maximization of expected utility. If we intend to use in practice this relatively simple model of risk measurement, we must consider that seriousness and uncertainty of responses depend on the situation, the form of questions asked, the time which the respondents have, and on many other psychological and social factors. In other words, we need to calibrate various parameters of our questionnaire to be able to capture personal utility functions

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_32

of respondents as close as possible. There are several ways how to deal with this problem. For example, we can utilize theory of fuzzy questionnaires, e.g. [1–3]. In the paper, we present alternative, simple and straightforward, methodology of calibrating a questionnaire with a small number of questions based on two selected utility functions, where parameters of the questionnaire are determined using expected utility maximization, see [12,13].

2 Preliminaries

Firstly, we present a definition and basic properties of a utility function.

Definition 1 (Utility function). *A function $u(x)$ of the input x is a utility function if*

1. *it is continuous,*
2. *it is quasi-concave, i.e. $\{x|u(x) \geq k\}$ is a convex set for each $k \in R$,*
3. *it is monotone, i.e. $u(x) \geq u(y)$ if $x \geq y$, or strictly monotone, i.e. $u(x) > u(y)$ if $x > y$.*

2.1 Risk Aversion Coefficients

The utility function has two key properties - a slope and concavity or convexity. A higher curvature of $u(x)$ represents a higher risk aversion or higher risk seeking.

If we can specify the relationship between utility and wealth in a function, the risk aversion coefficient measures how much utility we gain (or lose) as we add (or subtract) from our wealth. The first derivative of the utility function according to wealth $u'(x)$ should provide a measure of this, but it is specific to an individual and cannot be easily compared across individuals with different utility functions.

To get around this problem, K.J. Arrow and J.W. Pratt proposed that we look at the second derivative of the utility function, which measures how the change in utility (as wealth changes) itself changes as a function of wealth level, and divide it by the first derivative to arrive at the Arrow-Pratt measure of absolute risk aversion ARA defined as follows:

Let $u(x)$ be a utility function and $x > 0$ be a wealth. Then the coefficient of absolute risk aversion is given by

$$ARA(x) = -\frac{u''(x)}{u'(x)}. \tag{1}$$

The advantage of this formulation is that it can be compared across different persons with different utility functions to draw conclusions about differences in risk aversion across people.

The following propositions express equivalence between the size of ARA and approach to risk:

- $ARA(x) > 0 \Leftrightarrow$ the utility function expresses risk aversion approach to risk,
- $ARA(x) < 0 \Leftrightarrow$ the utility function expresses risk seeker approach to risk,
- $ARA(x) = 0 \Leftrightarrow$ the utility function expresses risk neutral approach to risk.

The Arrow-Pratt-De Finetti measure of relative risk aversion (RRA) or coefficient of relative risk aversion is defined as follows.

Let $u(x)$ be a utility function and $x > 0$ be a wealth. Then the coefficient of relative risk aversion is given by

$$RRA(x) = -x \cdot \frac{u''(x)}{u'(x)}. \tag{2}$$

Following [8], we can explain a distinction between absolute and relative risk aversion coefficient. An absolute risk aversion coefficient explains how we react to absolute changes in wealth, whereas a relative risk aversion coefficient explains how we react to proportional changes in wealth.

Decreasing ARA implies that the amount of wealth that we are willing to put at risk increases as wealth increases, whereas decreasing RRA indicates that the proportion of wealth that we are willing to put at risk increases as wealth increases. With constant ARA, the amount of wealth that we expose to risk remains constant as wealth increases, whereas the proportion of wealth remains unchanged with constant relative risk aversion. Finally, we stand willing to risk smaller and smaller amounts of wealth, as we get wealthier, with increasing absolute risk aversion, and decreasing proportions of wealth with increasing relative risk aversion.

2.2 Expected Utility

Expected utility is calculated by the well-known formula

$$E\left[u\left(X\right)\right] = \sum_{i=1}^{n} u(x_i) \cdot p_i, \tag{3}$$

where $X = (x_1, x_2, \ldots, x_n)$ is a vector of the possible alternatives, p_i is the probability of an alternative x_i, $i = 1, 2, \cdots, n$ and $u(x) : R \to R$ is an increasing continuous utility function.

The criterion of expected utility maximization corresponds to the preference order $\preceq$, for which

$$X \preceq Y \equiv E\left[u(X)\right] \preceq E\left[u(Y)\right] \tag{4}$$

for a utility function u. The relation (4) means that among two alternatives X, Y we prefer the alternative with a larger expected utility. If $u(x)$ is monotone, the rule (4) preserves monotonicity.

2.3 Maximal Premium Model

In general, our respondent has two alternatives - to buy insurance or not. Suppose that he owns a capital w, which he values wealth by the utility function u.

If he is insured against a loss X for a gross annual premium GP, he has a certain situation and his decision to buy insurance gives him the utility value $u(w-GP)$. If he is not insured, it means an uncertain situation for insured. In this case, the expected utility is $E[u(w-X)]$. Based on Jensen's inequality, we get

$$E[u(w-X)] \leq u(E[w-X]) \leq u(w-GP). \tag{5}$$

Since utility function u is an increasing continuous function, $GP \leq P^{max}$, where P^{max} denotes the maximum premium to be paid. This so-called *zero utility premium* is the solution to the following utility equilibrium equation

$$E[u(w-X)] = u(w-P^{max}). \tag{6}$$

2.4 Power and Exponential Utility Functions

Our goal is to determine the utility functions of potential clients of an insurance company, therefore we consider that they are risk averse and we calibrate their utility functions to reflect the risk aversion. To illustrate the calibration process, we have chosen two well known families of utility functions, power and exponential functions. These functions have different absolute and relative risk aversion coefficients

Power function is given by

$$u(x) = \frac{x^{1-\alpha}}{1-\alpha} \;\; if \, \alpha > 1. \tag{7}$$

The corresponding absolute risk aversion coefficient is given by

$$ARA(x)_{power} = \frac{\alpha}{x}, \tag{8}$$

relative risk aversion coefficient is given by

$$RRA(x)_{power} = \alpha. \tag{9}$$

Exponential function in the form

$$u(x) = \frac{1}{\alpha} \cdot (1 - \exp(-\alpha \cdot x)) \quad if \;\, \alpha > 0. \tag{10}$$

has the corresponding risk aversion coefficients defined as follows:

$$ARA(x)_{exp} = \alpha, \tag{11}$$

$$RRA(x)_{exp} = x \cdot \alpha, \tag{12}$$

3 Calibration of the Questionnaire and the Selected Utility Functions

Forms of utility functions (7) and (10), mentioned above, are determined by the coefficient α, which in general represents the approach to risk. For the purpose of calibration of these utility functions, we can use a questionnaire like the one defined in [7].

Questionnaire. Suppose that you are going to invest 17,000 euros and you have a choice between four different investment strategies for a three-year investment. Which of these alternatives would you prefer (A respondent has to choose only one of the alternatives.)?

- Alternative A_1: in the best case profit 1,700 euros (10.00 %),[1] in the worst case profit 550 euros (3.24 %).
- Alternative A_2: in the best case profit 2,600 euros (15.29 %), in the worst case zero profit (but no loss), (0.00 %).
- Alternative A_3: in the best case profit 4,000 euros (23.53 %), in the worst case loss 1,700 euros (−10.00 %).
- Alternative A_4: in the best case profit 6,500 euros (38.53 %), in the worst case loss 4,000 euros (−23.53 %).

The above mentioned questionnaire has three basic parameters, namely the number of questions (N), probabilities associated with the best and the worst case (p_b, p_w), and a fundamental investment (w). As a part of our calibration process, we would like to find appropriate values for these parameters.

The number of questions in our questionnaire determines the number of risk groups in our population we are able to identify. If we suspect that there is a huge risk diversity among people in our population, we would like to increase the number of questions. On the other hand, with too many question some individuals could face the Paradox of Choice being unable to choose an alternative. We need to take account also time needed to respond to a questionnaire with too many questions. In our opinion, the number of questions should be between four and ten.

Probabilities do not seem to be present in the questionnaire. Without providing explicitly information about probability of the occurrence of the best or the worst case in each question of the questionnaire such probability is implicitly set to be 0.5 or could be self-imposed by a respondent causing an error in identification of her/his risk group. In order to avoid such situation we have to provide these probabilities as a part of the questionnaire and they should be the same for all alternatives. To minimize a possible problem of different understanding of provided values of the probability by respondents, we could add a

[1] Percentage of possible profit or loss sequentially in all alternatives.

referential trial with identical probabilities of random events. Feasible probabilities are those allowing us to obtain at least one maximal expected utility for each alternative.

The last parameter of the questionnaire, w, is strictly context dependent and is set individually for each respondent based on her/his preferences. We included it into our calibration process in order to eliminate infeasible triplets (N, p_b, w).

In the paper, we illustrate the proposed calibration process mainly using two triplets, namely $(N = 4, p_b = 0.8, w = 17{,}000)$ and $(N = 4, p_b = 0.8, w = 13{,}500)$.

In the next step of our calibration process we, for each (N, p_b, w), evaluate expected utilities for all alternatives and for some selected risk aversion coefficients α in the utility functions (7) and (10). On the basis of the Eq. (13), we evaluate expected utilities and determine maximal expected utilities by

$$E[u(w,\alpha)] = p \cdot u(x_1,\alpha) + (1-p) \cdot u(x_2,\alpha) \tag{13}$$

for the best case x_1 and the worst case x_2, for all alternatives, for each α.
Thus, we answer two questions:

1. How the expected utilities change if we change our fundamental investment in the questionnaire?
2. How the expected utilities change if we change probabilities (the same for all alternatives) of occurring the best and the worst case?

We start with analysis of the power utility function (7).
Using (7) and (13) the expected utility is given by

$$E[u(w,\alpha)] = p \cdot \frac{((1+\frac{i_p}{100}) \cdot w)^{1-\alpha}}{1-\alpha} + (1-p) \cdot \frac{((1+\frac{i_{1-p}}{100}) \cdot w)^{1-\alpha}}{1-\alpha}, \tag{14}$$

where

- i_p - profit in the best case which occurs with the probability p,
- i_{1-p} - profit or loss in the worst case which occurs with the probability $1-p$.

Because the partial derivative $\frac{\partial E[u(w,\alpha)]}{\partial w}$ is positive, with increasing fundamental investment expected utilities increase, too. But, expected utility to be monotone increasing according to α, we need to work with a fundamental investment satisfying the condition $w \geq \frac{1}{1+\frac{i_{1-p}}{100}}$.

Table 2 in Appendix shows expected utilities determined using the power utility function with the same probability $p_b = 0.8$ in all alternatives. From formula (9), where $RRA_{power} = \alpha$, it is clear that a fundamental investment has not the impact on maximal expected utility, so the maximal expected utility remains on the same level of α. That means, the change of the fundamental investment does not affect the choice of parameter α of the power utility function. The level of maximal expected utilities remains on the same level as is highlighted at Table 2, i.e., to $\alpha = 3$ corresponds the maximal expected utility of the alternative A_4, to $\alpha = 5$ corresponds alternative A_3, to $\alpha = 18$ belongs alternative A_2, and finally, to $\alpha = 20$ alternative A_1 (in the case if we assume maximal $\alpha = 20$).

From positive partial derivative $\frac{\partial E[u(w,\alpha)]}{\partial p}$, it is obvious that with increasing probability p the expected utility increases, too (see Tables 3 and 4).

In the case of the exponential function (10), using the same parameters, results are listed in Tables 5 and 6. It is obvious that a fundamental investment impacts maximal expected utility in this case.

If we construct tables analogous to Tables 2 and 5 for different triples (N, p_b, w), we can then choose a triple providing us with a table which is the most similar to the ideal solution, i.e., giving as a table with as good separation of αs as possible. Table 1 illustrates the ideal solution assuming four alternatives A_1, A_2, A_3, A_4 and $\alpha \in \{1, 2, 3, \ldots, 12\}$. Black cells represent maximal expected utilities for Similarity between the ideal table and the table corresponding to a triple (N, p_b, w) is given as follows:

$$\mathrm{SIM}(\mathrm{table}_{\mathrm{ideal}}, \mathrm{table}_{(N,p_b,w)}) = \frac{\#\ \text{cells with matching colors}}{\#\ \text{alternatives} \times \#\ \text{levels of}\ \alpha}. \tag{15}$$

Table 1. An ideal solution for four alternatives and $\alpha \in \{1, 2, 3, \ldots, 12\}$

α	A_1	A_2	A_3	A_4
1				■
2				■
3				■
4			■	
5			■	
6			■	
7		■		
8		■		
9		■		
10	■			
11	■			
12	■			

4 Maximal Premium

Once our questionnaire is calibrated we can use it to identify a utility function of a respondent. Then, on the basis of individual personal utility functions, we can determine maximum premium of our respondent - client of an insurance company will be willing to pay for insurance of his wealth on the basis of the following model.

Our client has 17,000 euros and wants to insure his wealth worth 12,000 euros. The maximum premium P^{max} is calculated using equilibrium Eq. (6) as follows

$$P^{max} = w - u^{-1}\left(E\left[u\left(w - X\right)\right]\right). \tag{16}$$

The previous function gives a maximum premium determined with respect to the power utility function (7) as follows

$$P^{max}_{power} = w - \left((w - X)^{1-\alpha} \cdot p^* + w^{1-\alpha} \cdot (1 - p^*)\right)^{1/(1-\alpha)}, \tag{17}$$

and with respect to the exponential function (10) as follows

$$P^{max}_{exp} = \frac{\ln(1 - p^* + p^* \cdot \exp(\alpha x))}{\alpha}, \tag{18}$$

where p^* is the probability of occurrence of insured event.

On the basic of maximal expected utilities using power function (7), we can calibrate utility functions by aggregating of α s to maximum, average or minimum as follows $\alpha = 3$, $\alpha = 2.5$ and $\alpha = 2$, see Table 2. The Maximal expected utilities and corresponding a for exponential utility function (10) of the alternative A3 are given in Table 7 corresponding maximal gross annual premium is described in Table 8.

Similarly, using maximal expected utilities evaluated by exponential function (10), we can calibrate utility functions by aggregating of α s to maximum, average or minimum, consequently $\alpha = 2$, $\alpha = 1.5$ and $\alpha = 1$, see Tables 5 and 6. Corresponding maximal gross annual premiums are described in Table 9.

The selection of maximal (minimal) α could be interpreted as focus on profits (market share).

5 Conclusions

In this paper, we introduced methodology how to calibrate questionnaires aimed for risk measurement, where parameters of the questionnaire are determined using expected utility maximization and similarity to ideal solution. For simplicity, we restricted ourselves to two basic families of utility functions - power and exponential utility functions. A further generalization of the presented methodology will be the object of future research.

Acknowledgement. Jana Špirková has been supported by the Project VEGA no. 1/0093/17 Identification of risk factors and their impact on products of the insurance and savings schemes.

Appendix[2]

Table 2. Expected utilities according to power utility function (7) with $p = 0.8$ and $w = 17{,}000$ €

	A_1	A_2	A_3	A_4
x_1	1.870	1.960	2.100	2.350
x_2	1.755	1.700	1.530	1.300
α	Expected utility			
2	−0.5417676006	−0.5258103241	−0.5116713352	**−0.4942716858**
3	−0.1468542974	−0.1387253581	−0.1334215603	**−0.1316025619**
4	−0.0531129349	−0.0489855239	**−0.0474083611**	−0.0508922003
5	−0.0216261081	−0.0195385907	**−0.0194081802**	−0.0240641956
6	−0.0093995620	**−0.0083486432**	−0.0086885586	−0.0130056075
7	−0.0042589103	**−0.0037327835**	−0.0041531599	−0.0076975207
⋮	⋮	⋮	⋮	⋮
10	−0.0004585870	**−0.0003956201**	−0.0005955985	−0.0021362131
11	−0.0002251476	**−0.0001948224**	−0.0003324839	−0.0014663374
⋮	⋮	⋮	⋮	⋮
17	−0.0000037796	**−0.0000036229**	−0.0000142122	−0.0001879100
18	−0.0000019532	**−0.0000019283**	−0.0000086842	−0.0001360248
19	**−0.0000010138**	−0.0000010340	−0.0000053344	−0.0000988139
20	**−0.0000005284**	−0.0000005582	−0.0000032912	−0.0000720071

Table 3. Maximal expected utilities for power utility function (7) of the alternative A_4

p	$w = 1.35$	$w = 1.70$	$E[u(1.35)]$	$E[u(1.70)]$
	α	α		
0.99	10	10	−0.0012378760	−0.0001551021
0.98	8	8	−0.0040690280	−0.00080904694
0.97	7	7	−0.0079685412	−0.0019957530
0.96	7	7	−0.0093087163	−0.0023311511
0.95	6	6	−0.0169412028	−0.0053443186
0.94	6	6	−0.0185613638	−0.0058550712
0.93	5	5	−0.0346051536	−0.0137506856
0.92	5	5	−0.0366030002	−0.0145440325
0.91	5	5	−0.0386008468	−0.0153373794
0.90	5	5	−0.0405387623	−0.0161307264
0.80	3	3	−0.2087725571	−0.1316025619
0.70	2	2	−0.6658316508	−0.5286415712

[2] All computations were made with fundamental investments divided by 10,000.

Table 4. Maximal expected utilities for power utility function (7) of the alternative A_3

p	$w = 1.35$	$w = 1.70$	$E[u(1.35)]$	$E[u(1.70)]$
	α	α		
0.99	18	18	−0.0000311904	−0.0000006202
0.98	15	15	−0.0001477490	−0.0000058666
0.97	14	14	−0.0002799662	−0.0000139965
0.96	12	12	−0.0007407520	−0.0000587266
0.95	11	11	−0.0012830466	−0.0001280853
0.94	11	11	−0.0014196868	−0.0001417119
0.93	10	10	−0.0023818293	−0.0002993874
0.92	9	9	−0.0040249082	−0.0006370676
0.91	9	9	−0.0042672553	−0.0006753900
0.90	8	8	−0.0072339231	−0.0014417376
0.80	5	5	−0.0487809975	−0.0194081802
0.70	–	–	–	–

Table 5. Expected utilities evaluated by the exponential utility function (10) with $p = 0.8$, $w = 17{,}000$ €

	A_1	A_2	A_3	A_4
x_1	1.870	1.960	2.100	2.350
x_2	1.755	1.700	1.530	1.300
α	Expected utility			
1	0.8421196221	0.8507765585	0.8587277239	**0.8691983116**
2	0.4875086673	0.4887262351	**0.4893129997**	0.4889345313
3	0.3320124220	**0.3321816075**	0.3321667948	0.3317525635
4	0.2498424572	**0.2498655774**	0.2498451037	0.2497076270
5	0.1999799036	**0.1999829890**	0.1999765524	0.1999386001
6	0.1666639888	**0.1666643862**	0.1666627810	0.1666529085
7	0.1428567748	**0.1428568231**	0.1428564579	0.1428539442
⋮	⋮	⋮	⋮	⋮
10	0.099999998917254	**0.099999998926022**	0.0999999954	0.0999999548
11	**0.09090909074923**	0.090909090740082	0.0909090900	0.0909090797
12	**0.08333333330946**	0.083333333306239	0.0833333332	0.0833333305

Table 6. Expected utilities evaluated by the exponential utility function (10) with $p = 0.8$, $w = 13{,}500$ €

	A_1	A_2	A_3	A_4
x_1	1.485	1.556	1.668	1.866
x_2	1.394	1.350	1.215	1.032
α	Expected utility			
1	0.7691819288	0.7793693088	0.7897588479	**0.8049481551**
2	0.4733242579	0.4754747117	0.4769647210	**0.4777275235**
3	0.3292167186	0.3296675663	**0.3298022929**	0.3293299448
4	0.2492842098	**0.2493779101**	0.2493593029	0.2490795745
5	0.1998670281	**0.1998862829**	0.1998698106	0.1997561365
6	0.1666408921	**0.1666447898**	0.1666379173	0.1665966451
7	0.1428519951	**0.1428527681**	0.1428503876	0.1428360749
⋮	⋮	⋮	⋮	⋮
10	0.099999953908425	**0.099999958602084**	0.0999998897	0.0999993400
11	0.09090908107190	**0.090909081766032**	0.0909090616	0.0909088773
12	0.08333333121209	**0.08333331279217**	0.0833333254	0.0833332636
13	**0.07692307646153**	0.076923076454629	0.0769230748	0.0769230540

Table 7. Maximal expected utilities and corresponding α for exponential utility function (10) of the alternative A_3

p	$w = 1.35$	$w = 1.70$	$E[u(1.35)]$	$E[u(1.70)]$
	α	α		
0.99	18	18	−0.0000311904	−0.0000006202
0.98	15	15	−0.0001477490	−0.0000058666
0.97	14	14	−0.0002799662	−0.0000139965
0.96	12	12	−0.0007407520	−0.0000587266
0.95	11	11	−0.0012830466	−0.0001280853
0.94	11	11	−0.0014196868	−0.0001417119
0.93	10	10	−0.0023818293	−0.0002993874
0.92	9	9	−0.0040249082	−0.0006370676
0.91	9	9	−0.0042672553	−0.0006753900
0.90	8	8	−0.0072339231	−0.0014417376
0.80	5	5	−0.0487809975	−0.0194081802
0.70	–	–	–	–

Table 8. Maximal premium for A_4 according to (7), $p = 0.8$, $w = 17{,}000$ €

α	3	2.5	2
p^*	Maximal premium (€)		
0.001	89.06	59.46	40.70
0.002	176.73	118.40	81.21
0.003	263.05	176.83	121.53
0.004	348.05	234.76	161.65
0.005	431.78	292.19	201.58
⋮	⋮	⋮	⋮
0.1	5,144.02	4,179.56	3,290.32
0.2	7,363.28	6,478.72	5,513.51
⋮	⋮	⋮	⋮
0.9	11,754.69	11,698.62	11,620.25
1.0	12,000.00	12.000.00	12,000.00

Table 9. Maximal premium for A_4 according to (10), $p = 0.8$, $w = 17{,}000$ €, and $w = 13{,}500$ €

w	1.70	1.35		
α	1	2	1.5	1
p^*	Maximal premium (€)			
0.001	23.17	49.87	33.58	23.17
0.002	46.30	99.24	66.99	46.30
0.003	69.36	148.13	100.24	69.36
0.004	92.38	196.55	133.32	92.38
0.005	115.34	244.50	166.23	115.34
⋮	⋮	⋮	⋮	⋮
0.1	2,086.48	3,471.53	2,725.13	2,086.48
0.2	3,811.88	5,500.78	4,654.00	3,811.88
⋮	⋮	⋮	⋮	⋮
0.9	11,275.58	11,523.34	11,418.93	11,275.58
1.0	12,000.00	12,000.00	12,000.00	12,000.00

References

1. Bouchon-Meunier, B., Lesot, M.-J., Marsala, C.: Modeling and management of subjective information in a fuzzy setting. Int. J. Gen Syst **42**(1), 3–19 (2013)
2. Bouchon-Meunier, B.: Useful information and questionnaires. Inf. Control **32**(4), 368–378 (1976)
3. Diaz, J., Rifqi, M., Bouchon-Meunier, B., Jhean-Larose, S., Denhiére, G.: Imperfect answers in multiple choice questionnaires. In: Dillenbourg, P., Specht, M. (eds.) Times of Convergence. Technologies Across Learning Contexts. EC-TEL 2008, Lecture Notes in Computer Science, vol. 5192. Springer, Heidelberg (2008)
4. Harrison, G.W., Lau, M.I., Rutström, E.E.: Estimating Risk Attitudes in Denmark: A Field Experiment, Department of Economics, College of Business Administration, University of Central Florida, USA (Harrison and Rutström), and Centre for Economic and Business Research, pp. 1–64. Denmark, Copenhagen (2004)
5. Chiappori, P.A., Paiella, M.: Relative risk aversion is constant:evidence from panel data. J. Eur. Econ. Assoc. **9**(6), 1021–1052 (2011)
6. Kaas, K., Goovaerts, M., Dhaene, J., Denuit, M.: Modern Actuarial Risk Theory. Kluwer Academic Publishers, Boston (2001)
7. Melicherčík, I.: Some useful optimization problems in portfolio theory. http://www.defm.fmph.uniba.sk/ludia/melichercik/papers/jarna_skola_2012.pdf
8. Risk Aversion. http://pages.stern.nyu.edu/~adamodar/New_Home_Page/risk/riskaversion.htm
9. Rotar, I.V.: Actuarial Models: The Mathematics of Insurance. Chapman & Hall/CRC Press, Boca Raton, London, New York (2007)
10. Špirková, J.: Mixture utility in general insurance. In: Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Methods Communications in Computer and Information Science, vol. 80, pp. 712–721. Springer (2010)
11. Špirková, J., Král', P.: Aggregation functions and personal utility functions in general insurance. Acta Polytech. Hung. **7**(4), 7–23 (2010)
12. Špirková, J.: Calibration of utility function. In: Aggregation Functions in Theory and in Practise, Advances in Intelligent systems and Computing, vol. 228, pp. 261–269. Springer (2013)
13. Špirková, J.: Calibration of utility function and mixture premium. In: Laurent, A., et al. (ed.) IPMU 2014, Part I, CCIS 442, pp. 486–495. Springer International Publishing Switzerland (2014)
14. Wakker, P.P.: Explaining the characteristics of the power (CRRA), utility family. Health Econ. **17**, 1329–1344 (2008)

Diagnostic Inference with the Dempster-Shafer Theory and a Fuzzy Input

Ewa Straszecka(✉)

Institute of Electronics, Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 16 Akademicka St., 44-100 Gliwice, Poland
ewa.straszecka@polsl.pl

Abstract. The present paper proposes a diagnosis support inference in which input evidence are fuzzy sets. Diagnostic rules are formulated as fuzzy focal elements in the Dempster-Shafer theory. An inclusion measure is used to evaluate matching knowledge with evidence and to calculate belief of the diagnosis. Data simulated for two diagnostic situations show that the method allow for using linguistic values as a diagnostic information.

Keywords: Dempster-Shafer theory · Fuzzy sets · Diagnosis support

1 Introduction

The Dempster-Shafer theory of evidence [2,5] with fuzzy focal elements [6] can be applied in medical diagnosis support [7]. This implementation uses crisp values of symptoms, for instance measurements, as inputs for the diagnostic process. However, many researchers nowadays work on possibility of introducing fuzzy evidence in to a diagnostic inference (e.g. [4]). Such an approach to diagnosis support is reasonable and prospective since linguistic values, like "low", "increased" or "acute" are natural for symptoms description. Therefore, it should be investigated how to introduce these values as fuzzy inputs into the inference. To this end, fuzzy inputs (i.e. evidence) should be matched to diagnostic rules (i.e. knowledge). This matching is possible by means of similarity measures [1,3] that compare membership functions in diagnostic rule premises to input functions. Yet, in diagnosis support it is clear which membership function is a pattern and which is an observation. The former is the membership function in a rule premise, as it is a source of knowledge. The latter is used for an evidence representation. Thus, a measure of inclusion of the latter to the former is probably more accurate to evaluate belief in the diagnosis than a similarity factor. Once the inclusion of the evidence in the premise is satisfactory, the conclusion of the rule, i.e. the diagnosis, is confirmed. Thus, the inclusion should influence the belief measure of the diagnosis [9]. In this way, beliefs of competitive diagnoses

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_33

can be determined and the final diagnosis can be chosen as the one with the greatest belief.

Though fuzzy databases are not in everyday use, they will surely develop in future. In the meantime, the proposed method can be also useful in following problems:

- combining expert's knowledge and data-driven rules
- comparing two databases to enable a knowledge transfer
- selecting training data for a knowledge-based system of a diagnosis support.

The present paper proposes a diagnosis support inference by means of the Demster-Shafer theory extended for fuzzy focal elements [7]. Yager's inclusion [8] is chosen as an instance of the inclusion used in belief calculations. The concept is demonstrated on simulated examples, yet they imitate real diagnosis support circumstances that are experienced by the author.

2 Methods

The Dempster-Shafer theory extended for fuzzy focal elements preserves the basic definition of the basic probability assignment (bpa) [2]:

$$m(f) = 0, \sum_{s_i \in S, i=1,\ldots,n} m(s_i) = 1, \tag{1}$$

where S in the set of focal elements and s_i is the fuzzy focal element. The only difference between (1) and the classical definition is that s_i are rule premises including membership functions (mfs). The diagnostic rule is of the form:

$$s_i^{(l)}\text{: IF } V_j \text{ i s} A_j^{(l)} \wedge \ldots \wedge V_k \text{ is } A_k^{(l)} \text{ then } D_l \tag{2}$$

where (l) is the index of diagnosis, V_j, V_k, are linguistic variables and $A_k^{(l)}$ their linguistic values. The latter are represented by the $\mu_j^{(l)}(v_j)$, $\mu_k^{(l)}(v_k)$, mfs. The focal element can be single - when it refers to one variable V_j, or complex when it concerns several variables $V_j, \ldots, V_k$. The belief in the diagnosis is calculated as [9]:

$$Bel(D_l) = \sum_{s_i \in S} I(s \subset s_i^{(l)}) m(s_i), \tag{3}$$

where s denotes an observation (an evidence). If s_i is single then s is represented by the single fuzzy set A_k defined by the $\mu_k(v_k)$. When s_i is complex, also s refers to several mfs $\mu_j(v_j), \ldots, \mu_k(v_k)$. Each of the latter functions are matched with the premise using the inclusion operator. Let us choose the inclusion as [8]:

$$I(\mu_1(x), \mu_2(x)) = \min_x \left(\bar{\mu_1}(x) \vee \mu_2(x) \right), \tag{4}$$

where $\bar{\mu_1}(x)$ denotes the complement of $\mu_1(x)$. If the s_i is single in (3), the calculation of Bel is straightforward. If it is complex, then the conjunction of inclusion measure values for the appropriate mfs is used. Let us notice that the intersection of mfs for the fuzzy rule premise (2) is related to different variables, so indeed the inclusion for the complex premise is performed individually for each variable. Thus, the minimum of inclusion measures is related to the global minimum of mfs over all variables. These ensures that if mfs for s and s_i are identical then $Bel = 1$. A a single empty intersection:

$$A_k \cap A_k^{(l)} = \emptyset \Longrightarrow \min_x(\mu_k(v_k), \mu_k^{(l)}(v_k)) = 0 \tag{5}$$

is sufficient for rejecting the s_i from the Bel calculation. When the empty sets result from all intersections $s \cap s_i = \emptyset$, $i = 1, \ldots, n$, then $Bel = 0$.

The mfs in rules can be proposed by an expert or found from data. Similarly the m values (1) can be determined. The evidence mfs are fuzzy inputs for the diagnosis, which are expected to be introduced by a human diagnostician. These functions are not necessarily defined by a user, it is enough if he/she choose a linguistic value that will be linked to an appropriate shape of the function.

3 Simulated Data

The simulated diagnosis is based on three symptoms: V_1, V_2 and V_3. Mfs of symptoms relevant to low values are assigned to the D_1 diagnosis and high values to the D_2. The symptoms V_2 and V_3 are correlated to model a situation that often happens in the diagnosis. Four rules are formulated for D_l, $l = 1, 2$:

$s_i^{(l)}$: IF V_j is $A_i^{(l)}$ then D_l, $i = 1, 2, 3$,
$s_4^{(l)}$: IF V_2 is $A_2^{(l)}$ and V_3 is $A_3^{(l)}$ then D_l.

Samples of data are simulated, each of them includes 100 data sets. Every data set contain 400 cases, i.e. 200 for each of two diagnoses. Each variable data is generated for chosen parameters of the normal distribution $N(\bar{x}, \sigma)$, where $\bar{x}$ - mean, σ - variance. For the diagnosis D_1 the v_1 variable data are simulated for $\bar{v}_1 = 1$, $\sigma_1 = 1$; for v_2: $\bar{v}_2 = 1$, $\sigma_2 = 2$; for v_3: $\bar{v}_3 = 1$, $\sigma_3 = 3$. For the D_2 diagnosis v_k values, $k = 1, 2, 3$ have the same parameters $\bar{v}_k = 5$, $\sigma_k = 1$. The variables v_2 and v_3 are correlated and the correlation coefficient is $r \geq 0.2$. Normality of data is verified by the Matlab® Liliefors test.

Mfs that are built for focal elements, i.e. premises of the above mentioned rules, are determined as Gaussian-like shapes:

$$\mu_k^{(l)}(v_k) = \exp\left((v_k - \bar{v_k})/(2\sigma_k)\right), \tag{6}$$

where $k = 1, 2, 3$ is the variable index, while l=1,2 is the index of the diagnosis. These functions are "knowledge" mfs.

Since the input mfs are needed, they are created from the simulated data. Each time the mean and the variance of the sample is found and then put into the (6) formula. These means and variances are generally different from the original values assumed during simulation. The smaller is the number of cases in a sample, the greater is the difference. The following numbers of cases are taken from the sample: 50, 100, 150 and 200. Hence, various shapes of mfs are obtained. Obviously, more numerous case sets are concerned as more reliable data to create mfs inputs.The input mfs are denoted as $\mu_k^{(l)_n}(v_k)$, where $n = 50, 100, 150, 200$.

The inclusion factors (4) were next calculated between the input mfs and the mfs of the both diagnoses, i.e. $I(\mu_k^{(1)_n}, \mu_k^{(1)})$ and $I(\mu_k^{(2)_n}, \mu_k^{(2)})$, $n = 50, 100, 150, 200$. Values of the inclusion factors were tested for statistical difference of means. The hypothesis of different means was rejected with $\alpha = 0.05$ for $n = 200$ and $\alpha = 0.01$ for $n = 150, 200$. Thus, the more numerous samples are reliable data. Anyway, both for numerous or for scarce data mfs from assumed $\bar{v}_k$, σ_k and from data-driven parameters are different. However, this difference should not change an interpretation of symptoms.

Two bpas are considered for the focal elements. The first is uniform: $m_l(s_i^{(l)}) = 0.25$, $i = 1, \ldots 4$, $l = 1, 2$. It suits as an easy-to follow example of calculations. The second bpa illustrates real diagnosis circumstances experienced by the author: $m_l(s_1^{(l)}) = 0.3$, $m_(s_2^{(l)}) = 0.25$, $m_l(s_3^{(l)}) = 0.25$, $m_l(s_4^{(l)}) = 0.2$, $l = 1, 2$. Such a distinction among single and complex focal elements appear when the bpa is data-driven from a frequency of (crisp) symptom's occurrence [6]. Complex focal elements for correlated variables are usually less significant.

4 Experiment

It is important to find out to what extent the shape of the input mfs may influence the belief in the diagnosis. It is obvious that mfs input by a human user will be roughly shaped. The same concerns the data driven input mfs. Even if they are built for several measurements, the number of data in medical diagnosis is usually low, so these mfs are approximated anyway. If the inclusion measure value strongly depends on mfs shapes, the whole idea of the inference could be ruined.

During the experiment simulated fuzzy inputs s_k $k = 1, 2, 3$ are matched with the fuzzy focal elements $s_i^{(l)}$ $i = 1, \ldots, 4, l = 1, 2$ from (2). These means that the inclusion measure value (4) for input mfs $\mu_k^{(l)_n}(v_k)$, $n = 50, 100, 150, 200$ and knowledge mfs $\mu_k^{(l)}(v_k)$ are calculated. Afterwards, the $Bel(D_1)$ and $Bel(D_2)$ values (3) are found. In this way it can be tested what is the influence of a mf shape on the value of the inclusion measure and – in the following on the Bel value.

The Figs. 1 and 2 show exemplary mfs for chosen samples of 50 and 200 cases, respectively. The both $\mu_k^{(2)_{50}}(v_k)$ and $\mu_k^{(2)_{200}}(v_k)$ do not vary much from $\mu_k^{(2)}(v_k)$, $k = 1, 2, 3$. The same occurs for $\mu_1^{(1)_{50}}(v_1)$ and $\mu_1^{(1)_{200}}(v_1)$. On the other hand, differences between $\mu_k^{(1)_{50}}(v_k)$ and $\mu_k^{(1)}(v_k)$ as well as $\mu_k^{(1)_{200}}(v_k)$ and $\mu_k^{(1)}(v_k)$ $k = 2, 3$ seem significant. Apparently, the shape is more influenced by the variance than by the number of cases. Now, there is the question how these shape changes influence Bel.

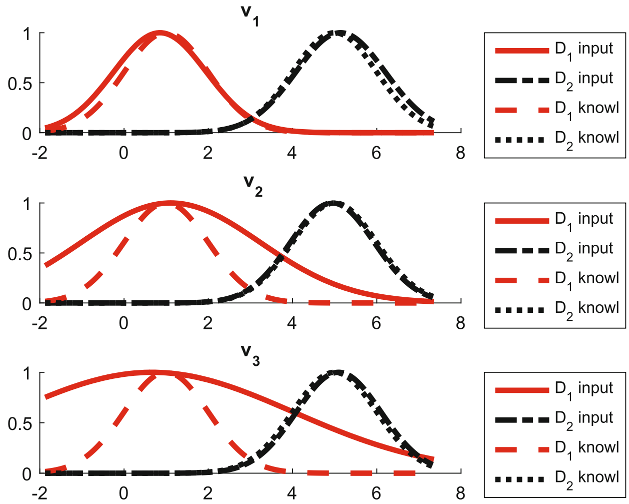


Fig. 1. Mfs for the knowledge base in comparison to mfs for fuzzy inputs for 50-case samples ($\mu_k^{(1)50}(v_k)$ – solid, $\mu_k^{(1)}(v_k)$ – dashed, $\mu_k^{(2)50}(v_k)$ – dash-and-dotted, $\mu_k^{(2)}(v_k)$ – dotted line).

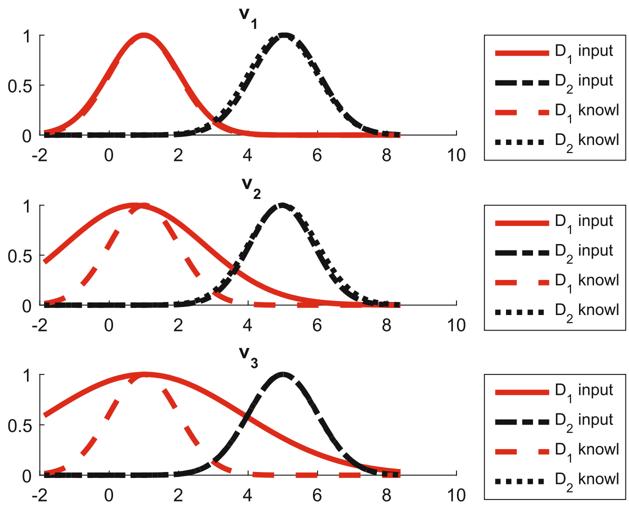


Fig. 2. Mfs for the knowledge base in comparison to mfs for fuzzy inputs for 200-case samples ($\mu_k^{(1)200}(v_k)$ – solid, $\mu_k^{(1)}(v_k)$ – dashed, $\mu_k^{(2)200}(v_k)$ – dash-and-dotted, $\mu_k^{(2)}(v_k)$ – dotted line).

Still, before it is evaluated, let us consider even more mfs changes. The means of various variable distributions also influence mfs. Therefore, the input mfs are shifted by the distance of one and next of two standard deviations towards the

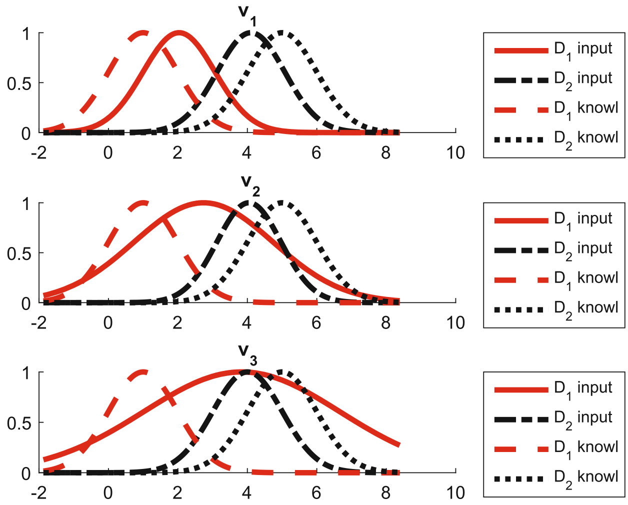


Fig. 3. Mfs in a knowledge base in comparison to mfs for fuzzy inputs with one-standard-deviation shifts ($\mu_k^{(1)200}(v_k)$ – solid, $\mu_k^{(1)}(v_k)$ – dashed, $\mu_k^{(2)200}(v_k)$ – dash-and-dotted, $\mu_k^{(2)}(v_k)$ – dotted line).

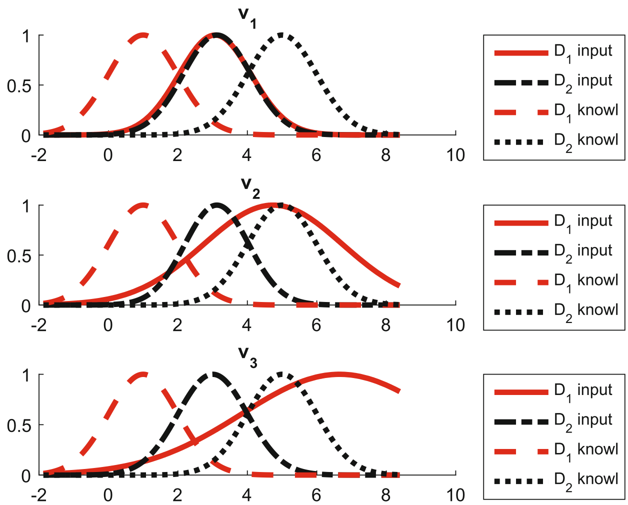


Fig. 4. Mfs in a knowledge base in comparison to mfs for fuzzy inputs with two-standard-deviation shifts ($\mu_k^{(1)200}(v_k)$ – solid, $\mu_k^{(1)}(v_k)$ – dashed, $\mu_k^{(2)200}(v_k)$ – dash-and-dotted, $\mu_k^{(2)}(v_k)$ – dotted line).

center, to model a disturbed diagnosis. Such changes of mfs for chosen 200-case samples are illustrated in Figs. 3 and 4.

This time differences are clear and for the two-standard-deviation shift even $\mu_k^{(1)200}(v_k)$ and $\mu_k^{(2)200}(v_k)$ are more similar to each other then to $\mu_k^{(1)}(v_k)$

or $\mu_k^{(2)}(v_k)$. This trial should test whether the rules are capable of responding with low Bel, i.e. a doubtful diagnosis for dubious inputs.

After designing mfs and computation of inclusion measure values (4), $Bel(D_1)$ and $Bel(D_2)$ are calculated. The Bel mean values and variances for 100 samples, evaluated for input mfs resulted from different number of cases, without the shift and with the one-standard-deviation shift are in Figs. 5 and 6, respectively. The results are obtained for the uniform bpa (see Sect. 3). The $Bel(D_1)$ value is calculated both for $\mu_k^{(1)n}(v_k)$, $k = 1, 2, 3$ (the diagram denoted "input D_1, diag D_1") as well as for $\mu_k^{(2)n}(v_k)$ $k = 1, 2, 3$ (input D_2, diag D_1), $n = 50, 100, 150, 200$. The $Bel(D_2)$ is calculated in analogous way. It is observable that Bel for diagnoses competitive to original (lower diagrams) are much smaller (note 10^{-4} factor on vertical axes). On the contrary, for mfs with two-standard-deviations shift, the Bel presented in Fig. 7 are almost equal or even higher for the competitive diagnoses. This indicate that the method is sensitive for a change of data.

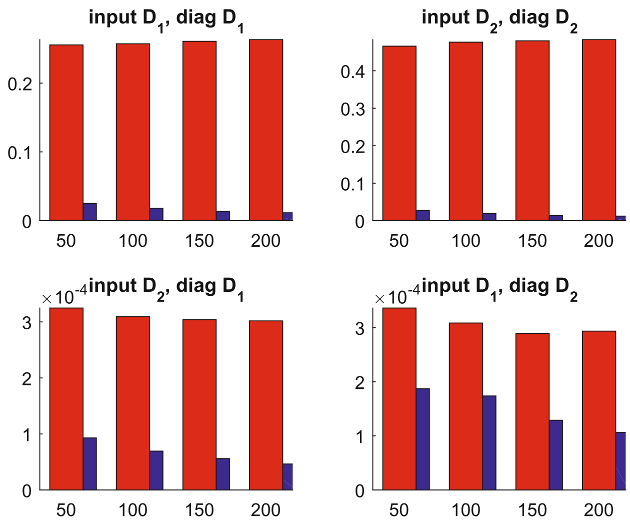


Fig. 5. Means (bars in front) and variances of Bel values for mfs without shift. Mfs determined for 50, 100, 150 and 200 data cases.

Dependencies are observable regardless the number of cases used for mfs determination, hence a rough shape of input mfs should not influence the effectiveness of the inference.

Results obtained for the bpa of the "real diagnosis" (see Sect. 3) are in Table 1. It can be noticed that smaller number of cases in a sample causes greater variance, but general conclusions from mean beliefs are the same. When mfs

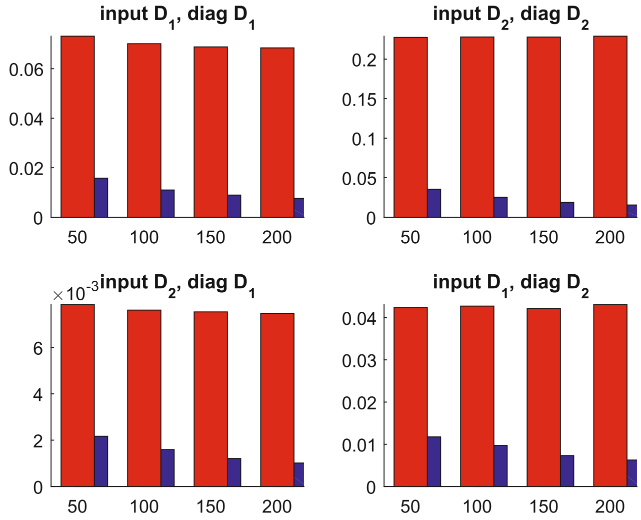


Fig. 6. Means (bars in front) and variances of *Bel* values for mfs with one-standard-deviation shift. Mfs determined for 50, 100, 150 and 200 data cases.

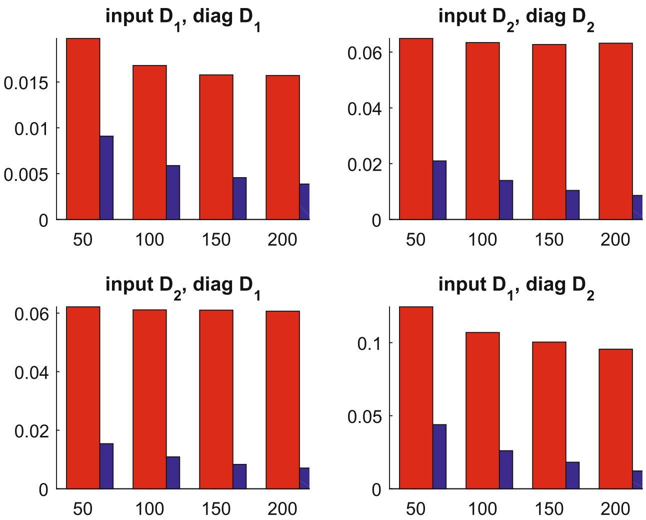


Fig. 7. Means (bars in front) and variances of *Bel* values with two-standard-deviation shift. Mfs determined for 50, 100, 150 and 200 data cases.

are shifted towards more dubious, the belief decrease, so for the two-standard-deviation cases the beliefs $Bel(D_1)$ and $Bel(D_2)$ are almost equal. Hence, also for this bpa the method is both resistant to rough fuzzy input and sufficiently sensitive.

Table 1. *Bel* values

	nb of cases in sample	Input D_1 diag D_1 $\bar{x} \pm \sigma$	Input D_2 diag D_2 $\bar{x} \pm \sigma$	Input D_2 diag D_1 $\bar{x} \pm \sigma$	Input D_1 diag D_2 $\bar{x} \pm \sigma$
No shift	50	0.272 ± 0.025	0.467 ± 0.027	3.3e-4 ± 9e-5	3.5e-4 ± 1.8e-4
	100	0.273 ± 0.018	0.477 ± 0.019	3.1e-4 ± 7e-5	3.2e-4 ± 1.7e-4
	150	0.277 ± 0.014	0.480 ± 0.014	3.1e-4 ± 6e-5	3.0e-4 ± 1.3e-4
	200	0.279 ± 0.012	0.483 ± 0.012	3.1e-4 ± 5e-5	3.0e-4 ± 1.0e-4
1 std shift	50	0.085 ± 0.018	0.229 ± 0.035	0.008 ± 0.002	0.041 ± 0.011
	100	0.081 ± 0.013	0.228 ± 0.025	0.008 ± 0.002	0.041 ± 0.009
	150	0.080 ± 0.010	0.228 ± 0.019	0.008 ± 0.001	0.041 ± 0.007
	200	0.080 ± 0.009	0.229 ± 0.015	0.008 ± 0.001	0.042 ± 0.006
2 std shift	50	0.023 ± 0.010	0.066 ± 0.021	0.063 ± 0.015	0.122 ± 0.041
	100	0.020 ± 0.007	0.064 ± 0.014	0.062 ± 0.011	0.107 ± 0.024
	150	0.019 ± 0.005	0.063 ± 0.010	0.062 ± 0.008	0.101 ± 0.017
	200	0.019 ± 0.005	0.063 ± 0.009	0.061 ± 0.007	0.097 ± 0.012

5 Discussion and Conclusions

The paper is the first proposition of the author to use the Dempster-Shafer theory extended for fuzzy sets as a tool of reasoning with fuzzy evidence. An inference is analogous to crisp inputs, but the inclusion factor is introduced to the belief measure calculation. Using the inclusion instead of the similarity of fuzzy sets appears more appropriate for matching evidence with knowledge in diagnostic rules. In the diagnosis support we know which fuzzy set is a diagnostic pattern and which is a representation of evidence. The latter is often imperfect, as its membership function is input by a user or calculated from limited number of data cases. Still, results of simulated diagnosis seem to be promising. Observable, but not too significant changes in membership functions do not influence the correct diagnosis. On the contrary, if membership functions on inputs are dubious, the belief of the diagnosis is low. These trials indicate that the method is both sufficiently robust and sensitive.

The method is here presented on simulation data. However, it is ready to implement and test for real problems of medical diagnosis. Such an implementation is justified. Usually, mean values of blood pressure or glucose level are considered during a diagnosis. It would be probably more correct to use membership functions describing the measurements instead.

Thorough tests of the method are planned in the nearest future. First of all – for benchmark medical databases in which part of data will be replaced by fuzzy numbers. It may be also interesting to use this method to evaluate a generalization of knowledge by calculating beliefs for rules found for two data sets. Such a test may help to avoid unjustified extrapolation of diagnoses for symptom values that are not provided in training data.

Acknowledgement. This research was supported by statutory funds of the Institute of Electronics, Silesian University of Technology.

References

1. Chen, S.-M., Yeh, M.-S., Hsiao, P.-Y.: A comparison of similarity measures of fuzzy values. Fuzzy Sets Syst. **72**, 79–89 (1995)
2. Dempster, A.P.: A generalisation of Bayesian inference. J. Royal Stat. Soc. **30**, 205–247 (1968)
3. Lee, S.H., Pedrycz, W.: Design of similarity measures for fuzzy sets on the basis of distance measure. Int. J. Fuzzy Syst. **11**(2), 67–72 (2009)
4. Rakus-Anderson E.: The parametric modification of fuzzy sets constituting a content of medical linguistic variables. In: Fuzzy Systems Conference, FUZZ-IEEE 2007, IEEE International, pp. 1-6 (2007)
5. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
6. Straszecka, E.: Combining uncertainty and imprecision in models of medical diagnosis. Inf. Sci. **176**, 3026–3059 (2006)
7. Straszecka, E.: Combining knowledge from different sources. Expert Syst. **27**(1), 40–52 (2010)
8. Yager, R.: Generalized probabilities of fuzzy events from fuzzy belief structures. Inf. Sci. **28**, 45–62 (1982)
9. Yen, J.: Generalizing the Dempster-Shafer theory to fuzzy sets. IEEE Trans. Syst. Man Cybern. **20**, 559–570 (1990)

Analyzing Feedback Mechanisms in Group Decision Making Problems

Atefeh Taghavi[1], Esfandiar Eslami[2], Francisco Javier Cabrerizo[3(✉)], and Enrique Herrera-Viedma[3]

[1] Graduate University of Advanced Technology, Kerman, Iran
taghavi.atefe@gmail.com
[2] Shahid Bahonar University of Kerman, Kerman, Iran
Esfandiar.Eslami@uk.ac.ir
[3] Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain
{cabrerizo,viedma}@decsai.ugr.es

Abstract. As reaching the maximum consensus degree in the group decision making problems is very important, many consensus reaching processes have been proposed in the literature. An important step within a consensus reaching process is the feedback mechanism, in which the experts involved in the decision problem under consideration are advised to modify their opinions in order to increase the level of consensus achieved. Therefore, many different feedback mechanisms have been proposed in the existing literature. The aim of this study is to present three of them and analyze their strengths and weaknesses. To do so, an illustrative example is provided.

Keywords: Consensus · Group decision making · Feedback mechanism

1 Introduction

In group decision making (GDM) problems several experts are gathering to rank a set of alternatives by means of evaluating them. In particular, it is assumed that there is a collection of feasible alternatives, $X = \{x_1, x_2, \ldots, x_n\}$, $n \geq 2$, and a group of experts, $E = \{e_1, e_2, \ldots, e_m\}$, $m \geq 2$. Each expert, based on her/his expertise, tries to give a preference value for x_i over x_j. We assume in this study that fuzzy preference relations are utilized to represent the experts' preferences, as they are the most common preference representation format [8,12,16,18].

The most important part of a GDM problem is the consensus reaching process, which is defined as a negotiation process composed by several consensus rounds, where, following the advice given by a moderator, the experts accept to change their preferences [2,11]. The agreement degree achieved in each round of the consensus reaching process is known by the moderator via the computation of some consensus measures. Given the importance of obtaining an

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_34

accepted solution by the whole group of experts, the consensus has attained a great attention and it is the major goal of GDM problems [4,7,9,11,13].

In a consensus reaching process, to improve the consensus level achieved among the group of experts, some recommendations have to be sent to the experts. In the first consensus approaches developed [1,5,11,13], a moderator was the person who provided the recommendations to the experts. The moderator's objective in each consensus round is to address the consensus reaching process towards success by reaching the maximum consensus level that is possible and reducing the number of experts outside of the consensus. The drawback of these consensus reaching processes based on the moderator is that she/he can introduce some subjectivity in the process. To avoid it, consensus reaching processes substituting the moderator's actions by an automatic feedback mechanism were proposed [7,10,11,15,17]. These new approaches compute proximity measures to assess the proximity between the individual experts' preferences and the collective one. They allow us to know which alternatives are the most controversial, which preferences have the highest disagreement between the group, how much they should change to influence the consensus degree, and so on. It makes more effective and efficient the consensus reaching process.

Since the feedback mechanism is an important step within a consensus reaching process, the aim of this study is to analyze some feedback mechanisms that have been proposed in the literature in order to find out their strengths and weaknesses. In particular, three different feedback mechanisms are analyzed in this study. The first one is a basic feedback mechanism in which both consensus and proximity measures are used to guide the consensus reaching process [10]. The second one is an adaptive feedback mechanism that adapts its behavior to the agreement achieved in each discussion round [15]. The third one is a feedback mechanism adjusting the amount of advice required by each expert depending on her/his own importance level. These feedback mechanisms are analyzed as they represented a novelty when they were proposed.

The rest of this study is set out as follows. Section 2 contains preliminaries and some general considerations about GDM and consensual processes. In Sect. 3, we introduce the three feedback mechanisms analyzed in this study. A practical example is illustrated in Sect. 4, and an analysis of the three feedback mechanisms is also provided. The last section contains some concluding remarks and future studies.

2 Preliminaries

In this section, we review some basic concepts which are needed in the sequel. For more details, see [10,15,17]. In whole paper, it is assumed that we have m experts who decide about n alternatives. In addition, each expert expresses his/her opinions using a fuzzy preference relation.

Definition 1. *A fuzzy preference relation P on a set of alternatives X is a fuzzy set on the Cartesian product $X \times X$, i.e., it is characterized by a membership function $\mu_P : X \times X \longrightarrow [0,1]$.*

Each value p_{ij} of the $n \times n$ matrix $P = (p_{ij})$ represents the preference degree of the alternative x_i over x_j; $p_{ij} = 0.5$ indicates indifference between x_i and x_j ($x_i \sim x_j$); when x_i is absolutely preferred to x_j we have $p_{ij} = 1$; and $p_{ij} > 0.5$ indicates that x_i is preferred to x_j ($x_i \succ x_j$). It is also assumed that $p_{ii} = 0.5$ $\forall i \in \{1, \ldots, n\}$. Since p_{ii} does not matter, it is usually written as '–' instead of 0.5 [13].

Once the fuzzy preference relations have been given by the experts, two kinds of measures can be computed [10]: consensus degrees, which are used to measure the actual level of consensus in the process, and proximity measures, which give information about how close to the collective solution every expert is.

The agreement achieved among all of the experts is calculated based on the coincidence concept [6]. To do so, consensus degrees given at three different levels of a fuzzy preference relation are determined [3]: pairs of alternatives, alternatives and relation.

– For each pair of experts (e_k, e_l) $(k = 1, \ldots, m-1, l = k+1, \ldots, m)$ a similarity matrix, $SM^{kl} = (sm^{kl}_{ij})$, is defined as:

$$sm^{kl}_{ij} = 1 - |p^k_{ij} - p^l_{ij}| \ . \tag{1}$$

– A consensus matrix, $CM = (cm_{ij})$, is calculated by aggregating all the $(m-1) \times (m-2)$ similarity matrices by using the arithmetic mean as the aggregation function, ϕ. It should be noted that depending on the nature of the GDM problem, different aggregation operators could be used:

$$cm_{ij} = \phi(sm^{kl}_{ij}),\ (k = 1, \ldots, m-1, l = k+1, \ldots, m) \ . \tag{2}$$

– After that the consensus matrix is computed, the consensus degrees are obtained at three different levels of a fuzzy preference relation:
  1. The consensus degree on pairs of alternatives (x_i, x_j), called cp_{ij}, is defined to measure the consensus degree among all the experts on that pair of alternatives. In this case, it is expressed by the element of the CM:

$$cp_{ij} = cm_{ij} \ . \tag{3}$$

  2. The consensus degree on alternatives x_i, called ca_i, is defined to measure the consensus degree among all the experts on that alternative:

$$ca_i = \frac{\sum_{j=1; j \neq i}^{n} (cp_{ij} + cp_{ji})}{2(n-1)} \ . \tag{4}$$

  3. The consensus degree on the relation, called cr, expresses the global consensus degree among all the experts' opinions. It is computed as the average of all the consensus degrees for the alternatives:

$$cr = \frac{\sum_{i=1}^{n} ca_i}{n} \ . \tag{5}$$

The consensus degree of the relation, cr, is the value used to control the consensus situation. The closer cr is to 1, the greater the agreement among all the experts' opinions.

Proximity measures are used to obtain the degree of agreement between each individual expert's preferences and the group ones. For this purpose, a collective fuzzy preference relation is first calculated. The collective preference, $P^c = (p^c_{ij})$, is computed by means of the aggregation of all individual preference relations, $\{P^1, P^2, \ldots, P^m\} : p^c_{ij} = \phi(p^1_{ij}, p^2_{ij}, \ldots, p^m_{ij})$, in which ϕ is an appropriate aggregation operator. As one of the feedback mechanism analyzed in this study has into account the importance level of each expert [17], the weighted arithmetic average is used as ϕ in order to give more importance to the preferences given by the most relevant experts. Therefore, at first we have to give a weight to each of the experts.

The proximity measures are also computed at the three level of a fuzzy preference relation:

1. Similarity measure on pairs of alternatives. The proximity measure of an expert, e_k, on the pair of alternatives, (x_i, x_j), to the group one, denoted as pp^k_{ij}, is computed as:
$$pp^k_{ij} = 1 - |p^k_{ij} - p^c_{ij}| \ . \tag{6}$$
2. Similarity measure on alternatives. The proximity measure of an expert, e_k, on the alternative, x_i, to the group one, denoted as pa^k_i, is computed as:
$$pa^k_i = \frac{\sum_{j=1, j\neq i}^{n} (pp^k_{ij} + pp^k_{ji})}{2(n-1)} \ . \tag{7}$$
3. Similarity measure on the relation. The proximity measure of an expert, e_k, on his/her preference relation to the group one, denoted as pr^k, is computed as:
$$pr^k = \frac{\sum_{i=1}^{n} pa^k_i}{n} \ . \tag{8}$$

This measure structure will allow us to find out the consensus state of the process at different levels. For instance, we will be able to identify which experts are close to the consensus solution, or in which alternatives the experts are having more trouble to reach consensus.

3 Description of the Feedback Mechanisms

In this section, we describe the main characteristics of the three feedback mechanisms that are analyzed in this study.

3.1 Basic Feedback Mechanism

The first feedback mechanism introduced is based on the consensus reaching process presented in [10]. Here, the authors proposed a consensus approach providing tools to support the consensus processes in presence of incomplete information. However, we are only interested in the feedback mechanism presented in this paper. In addition, in order to unify the methods in a way that analyzing them together be more meaningful, we ignore the consistency part presented in this paper. We denote this feedback mechanism as basic because it does not consider other criteria and it is always applied in the same way.

This feedback mechanism is applied in two steps: (i) identification of the preference values that should be modified, and (ii) generation of advice.

1. *Identification of the preference values.* The experts whose proximity measure on the relation is lower than a satisfaction threshold, $\overline{pr} = \sum_{k=1}^{m} pr^k/m$, should change their opinions:

$$EXPCH = \{k \mid pr^k < \overline{pr}\} \ . \tag{9}$$

   The alternatives that the above experts should consider to change, ALT, is the set of alternatives whose proximity measure is lower than a satisfaction threshold, $\overline{pa} = \sum_{k=1}^{m} pa_i^k/m$, i.e.:

$$ALT = \{(k,i) \mid e_k \in EXPCH \wedge pa_i^k < \overline{pa}\} \ . \tag{10}$$

   Finally, the preference values for every alternative and expert that should be changed are:

$$APS = \{(k,i,j) \mid (k,i) \in ALT \wedge pp_{ij}^k < \overline{pp}\} \tag{11}$$

   where $\overline{pp} = \sum_{k=1}^{m} pp_{ij}^k/m$.
2. *Generation of advice.* We must find out the direction of the change to be recommended in each case, that is, the direction of change to be applied to the preference degree p_{ij}^k, with $(k,i,j) \in APS$. On the one hand, if $p_{ij}^k < p_{ij}^c$, then the expert should increase this preference degree. On the other hand, if $p_{ij}^k > p_{ij}^c$, then the expert should decrease this preference degree. We also assumed that the expert increases or decreases the preference value with a value equal to 0.05.

3.2 Adaptive Feedback Mechanism

The second feedback mechanism described is based on the consensus reaching process proposed in [15], where the authors presented an adaptive consensus support model in a multigranular fuzzy linguistic context. Here, we adapt this feedback mechanism to deal with fuzzy preference relations.

This feedback mechanism is adaptive because it distinguishes three levels of consensus for searching the preferences: very low, low, and medium consensus. Each level implies a different search policy to identify the preferences with low agreement degree. The first level of consensus, the lowest level of agreement, all experts will be advised to modify all the preferences values identified in disagreement. If the level of consensus is greater, the search will be limited to the preference values in disagreement of those experts furthest from the group. If the global consensus degree, cr, is less than a consensus threshold γ, i.e. $cr < \gamma$, then a new consensus round is applied until the level of agreement is sufficient or we reach the maximum number of consensus rounds. So, the system establishes three different preference search procedure (PSp).

The adaptive search for preferences consists of two processes: (i) choose the most suitable PSp, and (ii) apply the PSp. To choose the most suitable PSp, two parameters θ_1 and θ_2 are fixed at the beginning of the consensus process.

- PSp for very low consensus, PSp^{VL}, if: $cr \leq \theta_1$; in which the level of pairs of alternatives is considered.
- PSp for low consensus, PSp^{L}, if: $\theta_1 < cr \leq \theta_2$; in which the level of alternative is considered.
- PSp for medium consensus, PSp^{M}, if: $\theta_2 < cr < \gamma$; in which the level of preference relation is considered.

After determining the PSp for each consensus level, each PSp finds out a set of preferences, $PREFECH_k = \{(i,j) \mid i,j \in \{1,2,\ldots,n\}, i \neq j\}$, to be changed by each expert e_k in the next discussion round.

1. The procedure for PSp^{VL} to find out the set of preferences to be changed by e_k, $PREFECH_k^{VL}$, is as follows.
   - Computing $\rho = \overline{cp}$ as follows:

$$\overline{cp} = \sum_{i=1}^{n} (\sum_{j=1, i \neq j}^{n} cp_{ij})/(n^2 - n) . \tag{12}$$

   - The set of preference values $PREFECH_k^{VL}$ to be changed by each expert e_k is:

$$P = \{(i,j) \mid cp_{ij} < \rho,\ i,j = 1,\ldots,n\} \tag{13}$$

$$PREFECH_k^{VL} = P . \tag{14}$$

2. PSp^{L} finds out the set of preferences to be changed by e_k, $PREFECH_k^{L}$, as follows:
   - The alternatives to be changed, X^{ch}, are identified as:

$$X^{ch} = \{i \mid ca_i < cr\} . \tag{15}$$

   - The pairs of alternatives to be changed are identified as:

$$P = \{(i,j) \mid i \in X^{ch} \wedge cp_{ij} < \overline{cp}\} . \tag{16}$$

   - The proximity of the alternatives that should be changed is computed for all experts:

$$\{pa_i^k \mid i \in X^{ch}\} \ \forall e_k \in E . \tag{17}$$

   - Computing the proximity threshold, β, to identify the experts that will be required to modify the identified pairs of alternatives:

$$\beta = \overline{pa}_i = \sum_{k=1}^{m} pa_i^k / m . \tag{18}$$

– The sets of preference values that are required to be modified are:

$$PREFECH_k^L = \{(i, j) \in P \mid pa_i^k < \overline{pa}_i\} \ . \tag{19}$$

3. For PSp^M, computing the set of preferences which should be changed by e_k, $PREFECH_k^M$ is as follows:
   – At first, (15)–(18) in PSp^L are carried out.
   – The proximity threshold to be used in identifying the experts required to modify the identified pairs of alternatives in disagreement is computed as follows:

$$\{\overline{pp}_{ij} = \sum_{k=1}^{m} pp_{ij}^k / m \mid (i, j) \in P\} \ . \tag{20}$$

   – The sets of preference values that are required to be modified are:

$$PREFECH_k^M = \{(i, j) \in P \mid pa_{ij}^k < \overline{pa}_i \wedge pp_{ij}^k < \overline{pp}_{ij}\} \ . \tag{21}$$

After identifying the preferences which should be changed, the model will suggest increasing or decreasing the current assessments like the first feedback mechanism presented in this study.

3.3 Feedback Mechanism Based on Experts' Importance

The third feedback mechanism presented is based on the consensus reaching process introduced in [17]. This feedback mechanism is based on the importance degrees of the experts. In this way, it is assumed that those experts with lower knowledge level on the problem will need more advice than others with higher importance. Hence, according to the experts' importance degree, they are classified into three groups: (i) high-importance experts, E_{high}, (ii) medium-importance experts, E_{med}, and (iii) low-importance experts, E_{low}. This classification is done via a fuzzy matching mechanism whose parameters depend on the problem dealt with. Therefore, each group of experts is a fuzzy set characterized by a membership function and two parameters λ_1 and λ_2 are established as membership thresholds. In such a way, using the importance degree of each expert, he/she can be classified in a particular group of experts. According to it, three different advising strategies are defined.

1. *Identify low-importance experts' controversial preferences.*
   – Establishing a threshold, α_1, as follows:

$$\alpha_1 = \sum_{i=1}^{n} (\sum_{j=1, j \neq i}^{n} cp_{ij}) / (n^2 - n) \ . \tag{22}$$

   – Identifying the pairs of alternatives, P, with a consensus degree smaller than the threshold α_1:

$$P = \{(i, j) \mid cp_{ij} < \alpha_1\} \ . \tag{23}$$

- The set of the controversial preferences, PCH^k_{low}, that should be changed by each expert $e_k \in E_{low}$ is:

$$PCH^k_{low} = P\ . \tag{24}$$

2. *Identify medium-importance experts' controversial preferences.*
   - Initially, the alternatives to be changed, XCH, are obtained:

$$XCH = \{i \mid ca_i < \alpha_2\} \tag{25}$$

   where $\alpha_2 = \sum_{i=1}^{n} ca_i/n$.
   - The pairs of alternatives to be changed are:

$$P = \{(i,j) \mid i \in XCH \wedge cp_{ij} < \alpha_1\}\ . \tag{26}$$

   - The set of preference values, PCH^k_{med}, that are required to be modified is:

$$PCH^k_{med} = \{(i,j) \in P \mid pa^k_i < \beta_1\} \tag{27}$$

   where, $\beta_1 = \sum_{k=1}^{m} pa^k_i/m$, $e_k \in E_{med}$.
3. *Identify high-importance experts' controversial preferences.*
   - The alternatives to be changed are:

$$XCH = \{i \mid ca_i < \alpha_2\}\ . \tag{28}$$

   - The pairs of alternatives to be changed:

$$P = \{(i,j) \mid i \in XCH \wedge cp_{ij} < \alpha_1\}\ . \tag{29}$$

   - The set of preference values that are required to be modified is:

$$PCH^k_{high} = \{(i,j) \in P \mid pa^k_i < \beta_1 \wedge pp^k_{ij} < \beta_2\} \tag{30}$$

   where, $\beta_2 = \sum_{k=1}^{m} pp^k_{ij}/m$, $e_k \in E_{high}$.

Finally, after identifying all the controversial preferences, the generation of advices is going on like the both previous methods.

4 Experimental Study

In this section, we highlight the main characteristics of the three feedback mechanism described in this study and show their advantages and drawbacks. To do so, an illustrative example is solved by using them.

Let us suppose that four experts $E = \{e_1, e_2, e_3, e_4\}$ provide the following fuzzy preference relations over a set of four alternatives $X = \{x_1, x_2, x_3, x_4\}$:

$$P^1 = \begin{pmatrix} - & 0.90 & 0.80 & 0.90 \\ 0.10 & - & 0.70 & 0.80 \\ 0.20 & 0.30 & - & 0.60 \\ 0.10 & 0.20 & 0.30 & - \end{pmatrix} \quad P^2 = \begin{pmatrix} - & 0.10 & 0.30 & 0.20 \\ 0.90 & - & 0.80 & 0.90 \\ 0.70 & 0.40 & - & 0.50 \\ 0.80 & 0.10 & 0.70 & - \end{pmatrix}$$

$$P^3 = \begin{pmatrix} - & 0.40 & 0.30 & 0.30 \\ 0.50 & - & 0.20 & 0.20 \\ 0.60 & 0.80 & - & 0.40 \\ 0.70 & 0.80 & 0.40 & - \end{pmatrix} \quad P^4 = \begin{pmatrix} - & 0.40 & 0.30 & 0.10 \\ 0.40 & - & 0.20 & 0.20 \\ 0.20 & 0.90 & - & 0.50 \\ 0.90 & 0.70 & 0.60 & - \end{pmatrix}$$

In addition, the following values of the parameters used in each feedback mechanism are assumed:

- Each feedback mechanism is run for 8 discussion rounds.
- The following weights are established for each expert: $e_1 = 0.35$, $e_2 = 0.25$, $e_3 = 0.20$, and $e_4 = 0.20$.
- In the adaptive feedback mechanism, θ_1 is equal to 0.80, and θ_2 is equal to 0.83.
- In the feedback mechanism based on experts' importance, λ_1 and λ_2 are equal to 0.25 and 0.35, respectively. Therefore, the low important experts are: $[e_3, e_4]$; the medium important expert is: $[e_2]$; and the high important experts is: $[e_1]$.

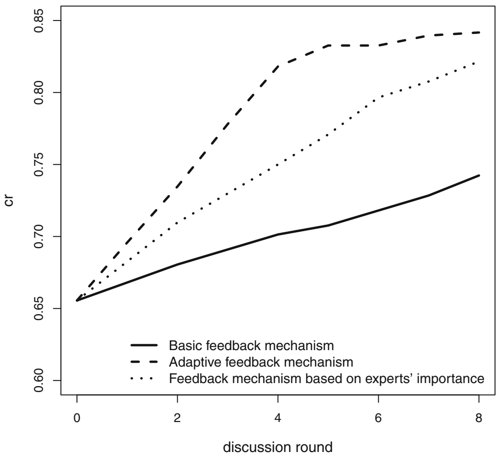


Fig. 1. The values of the *cr* obtained in successive discussion rounds

Figure 1 shows the performance of the three feedback mechanisms quantified in terms of the *cr* obtained in successive discussion rounds. Here, it should be pointed out that the value corresponding to the discussion round 0 is the *cr* obtained according to the initial fuzzy preference relations. Due to it, it is equal in the three approaches. In addition, Tables 1, 2 and 3 shows both the number of changes suggested to each expert and the total number of changes suggested by the basic feedback mechanism, the adaptive feedback mechanism and the feedback mechanism based on experts' importance, respectively.

In the following, we analyze the performance of the three approaches based on this example:

Table 1. Basic feedback mechanism

	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8
Changes of e_1	3	4	3	3	3	5	5	5
Changes of e_2	3	2	0	2	0	0	0	0
Changes of e_3	0	0	0	0	0	0	0	0
Changes of e_4	0	0	4	0	0	0	0	3
Total changes	6	6	7	5	3	5	5	8

Table 2. Adaptive feedback mechanism

	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8
Changes of e_1	8	8	7	5	3	2	0	0
Changes of e_2	8	8	7	5	6	1	1	1
Changes of e_3	8	8	7	5	0	0	0	0
Changes of e_4	8	8	7	5	0	1	0	0
Total changes	32	32	28	20	9	4	1	1

- *Basic feedback mechanism.* In this approach, the improvement of the consensus achieved among the experts remains constant through the discussion rounds as it is always applied in the same way. However, this approach presents a lower convergence to the consensus than the others two approaches. As an advantage, it suggests a lower number of changes to the experts (see Table 1). It facilitates that experts accept to change their preferences.
- *Adaptive feedback mechanism.* This approach presents a faster convergence to the consensus. The most notable improvement is at the first rounds of the discussion process. It is due to the consensus in this rounds is low and many modifications are suggested to the experts. When the consensus is high, lower changes are recommended, but the consensus continues increasing. The disadvantage of this approach is that many changes are advised when the consensus level is low (see Table 2) and, therefore, the experts could reject them.
- *Feedback mechanism based on experts' importance.* This approach also presents a fast convergence to the consensus. Furthermore, it recommends a low number of modifications to the most important experts (see Table 3). Consequently, the most considerable experts' opinions never will be strongly modified during the consensus reaching process. However, it also suggests a high number of changes to the low important experts during all the discussion rounds and, therefore, they might reject them. In addition, there could exists a limit scenario where the tyranny of the minority is accomplished if the excellence group is very small inside the set of experts [17].

Table 3. Feedback mechanism based on experts' importance

	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8
Changes of e_1	1	1	1	0	0	2	0	0
Changes of e_2	2	2	2	3	4	5	0	0
Changes of e_3	12	12	12	12	12	12	12	0
Changes of e_4	12	12	12	12	12	12	12	0
Total changes	27	27	27	27	28	31	24	24

5 Concluding Remarks

In this study we have introduced and analyzed three feedback mechanisms proposed in the literature to highlight their main features. All of these approaches overcome the problem of the moderator, giving the way to use an automatic system to calculate and send customized advice to the experts if there is not enough consensus. The adaptive feedback mechanism and the feedback mechanism based on experts' importance present a better performance as they advice to the experts according to different criteria as the consensus level achieved and the experts' importance, respectively. As a consequence, they present a faster convergence to the consensus than the basic feedback mechanism. However, they also suggest many more modifications that the basic one.

We have shown an only one example to analyze the performance of each feedback mechanism. It helps to understand the characteristics of each approach, but, for a better comparison, more examples should be analyzed. Therefore, as future study, we suggest to solve several GDM problems with different number of experts and alternatives, and with different values of the parameters utilized by each approach. In addition, the performance of other feedback mechanisms as, for instance, the one based on the so called action rules [14], should be also analyzed.

Acknowledgments. The authors would like to acknowledge FEDER funds under grants TIN2013-40658-P and TIN2016-75850-R.

References

1. Bordogna, G., Fedrizzi, M., Pasi, G.: A linguistic modeling of consensus in group decision making based on OWA operators. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **27**, 126–133 (1997)
2. Butler, C.T., Rothstein, A.: On Conflict and Consensus: A Handbook on Formal Consensus Decision Making. Creative Commons, Tahoma Park (2006)
3. Cabrerizo, F.J., Moreno, J.M., Pérez, I.J., Herrera-Viedma, E.: Analyzing consensus approaches in fuzzy group decision making: advantages and drawbacks. Soft. Comput. **14**, 451–463 (2010)
4. Eklund, P., Rusinowska, A., de Swart, H.: Consensus reaching in committees. Eur. J. Oper. Res. **178**, 185–193 (2007)

5. Herrera, F., Herrera-Viedma, E., Verdegay, J.L.: A rational consensus model in group decision making using linguistic assessments. Fuzzy Sets Syst. **88**, 31–49 (1997)
6. Herrera, F., Herrera-Viedma, E., Verdegay, J.L.: Linguistic measures based on fuzzy coincidence for reaching consensus in group decision making. Int. J. Approx. Reason. **16**, 309–334 (1997)
7. Herrera-Viedma, E., Herrera, F., Chiclana, F.: A consensus model for multi-person decision making with different preference structures. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **32**, 394–402 (2002)
8. Herrera-Viedma, E., Herrera, F., Chiclana, F., Luque, M.: Some issues on consistency of fuzzy preference relations. Eur. J. Oper. Res. **154**, 98–109 (2004)
9. Herrera-Viedma, E., Martínez, L., Mata, F., Chiclana, F.: A consensus support system model for group decision-making problems with multigranular linguistic preference relations. IEEE Trans. Fuzzy Syst. **13**, 644–658 (2005)
10. Herrera-Viedma, E., Alonso, S., Chiclana, F., Herrera, F.: A consensus model for group decision making with incomplete fuzzy preference relations. IEEE Trans. Fuzzy Syst. **15**, 863–877 (2007)
11. Herrera-Viedma, E., Cabrerizo, F.J., Kacprzyk, J., Pedrycz, W.: A review of soft consensus models in a fuzzy environment. Inf. Fusion **17**, 4–13 (2014)
12. Kacprzyk, J.: Group decision making with a fuzzy linguistic majority. Fuzzy Sets Syst. **18**, 105–118 (1986)
13. Kacprzyk, J., Fedrizzi, M., Nurmi, H.: Group decision making and consensus under fuzzy preferences and fuzzy majority. Fuzzy Sets Syst. **49**, 21–31 (1992)
14. Kacprzyk, J., Zadrozny, S., Ras, Z.W.: How to support consensus reaching using action rules: a novel approach. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **18**, 451–470 (2010)
15. Mata, F., Martínez, L., Herrera-Viedma, E.: An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context. IEEE Trans. Fuzzy Syst. **17**, 279–290 (2009)
16. Orlovski, S.A.: Decision-making with fuzzy preference relations. Fuzzy Sets Syst. **1**, 155–167 (1978)
17. Pérez, I.J., Cabrerizo, F.J., Alonso, S., Herrera-Viedma, E.: A new consensus model for group decision making problems with non-homogeneous experts. IEEE Trans. Syst. Man Cybern. Syst. **44**, 494–498 (2014)
18. Tanino, T.: Fuzzy preference orderings in group decision making. Fuzzy Sets Syst. **12**, 117–131 (1984)

A Statistical Study for Quantifier-Guided Dominance and Non-Dominance Degrees for the Selection of Alternatives in Group Decision Making Problems

J.M. Tapia(✉), M.J. del Moral, S. Alonso, and E. Herrera-Viedma

University of Granada, 18071 Granada, Spain
jmtaga@ugr.es

Abstract. In a group decision making problem the selection process is decisive to find a solution. In these problems there is a widespread agreement to use fuzzy preference relations to express different preferences about possible alternatives. Previous papers have proposed different selection methods in this context. An usual way is the use of a ranking method to obtain a classification of the alternatives. One of the methods used is based on two choice degrees: quantifier guided dominance degree and quantifier guided non-dominance degree. This paper presents a limited comparative study about the application of the two previously cited quantifier guided choice degrees. By using statistical tools, it is concluded that both choice degrees can offer significantly different rankings of alternatives. In addition, it has been observed that the variability of the alternatives in the ranking obtained by dominance choice degree is generally greater, which may facilitate a better discrimination between different alternatives.

Keywords: Group decision making · Fuzzy preferences · Dominance choice degree · Non-Dominance choice degree

1 Introduction

In a Group Decision Making (GDM) problem, a group of individuals must choose an solution from several possibilities. These individuals, usually called experts, can express their opinions by means of comparisons over all the possibilities, called alternatives.

A selection process is applied in order to obtain the final solution to the GDM problem [1]. In this situation, different selection policies can be applied [2]. The aim of this paper is to present a limited comparative study between two proposed [3] quantified guided choice degrees of alternatives: a Quantified Guided Dominance Degree (QGDD) and a Quantified Guided Non-Dominance Degree (QGNDD), both extend Orlovsky's non-dominance concept [4]. These choice degrees are used in the process for obtaining a ranking of alternatives from the best to the worst according to experts' opinions.

Different statistical tools are used in this study. A two-sample statistical test allows us to establish whether the application of the two choice degrees is different. The limited comparative study carried out is focused on the different ranking of alternatives

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_35

obtained by choice degrees and the variability of these results. This behaviour is further analysed in order to better discriminate between both choice degrees. In this process it is needed the coefficient of variation.

The paper is structured as follows. In Sect. 2, we introduce essential concepts: the GDM problem and the selection process. In Sect. 3 we describe the design of the experiment used to develop our study. We present the results obtained in this study in Sect. 4. Finally, Sect. 5 concludes the paper.

2 Preliminaries

2.1 The GDM Problem

A GDM problem consists in deciding the best alternative from a set of possible alternatives $X = \{x_1, \ldots, x_n\}$ according to the preferences expressed by a group of experts $E = \{e_1, \ldots, e_m\}$. Different preference methods were compared in [5], where it was concluded that pairwise comparison methods are 'better' than non-pairwise ones. There are different representation formats that experts may use to express their opinions, but fuzzy preference relations are one of the most used [4, 6–8].

A *fuzzy preference relation* P on a finite set of alternatives X is characterised by a membership function

$$\mu_P : X \times X \to [0,\ 1]\ with\ \mu_P(x_i, x_j) = p_{ij},$$

verifying

$$p_{ij} + p_{ji} = 1 \forall i,j \in \{1, \ldots, n\}.$$

If cardinality of X is small, the fuzzy preference relation may be denoted by the matrix $P = (p_{ij})$, $p_{ij} \in [0, 1]$ indicates the degree of preference for x_i over x_j; where 0 is minimum and 1 is maximum.

Before a final solution can be obtained in a GDM problem, a process is applied called *the selection process* [9–11]. This process supplies the final solution according to the preferences expressed by the experts [7].

2.2 Selection Process

The selection process have two different steps [12, 13]: (i) *aggregation* of individual preferences in a collective preference and (ii) *exploitation* of this collective preference.

Aggregation phase. In this phase a collective preference relation, $P^C = (p_{ij}^c)$, is obtained by means of the aggregation of all individual fuzzy preference relations $\{P^1, P^2, \ldots, P^m\}$, and shown the global preference between every pair of alternatives (x_i, x_j) according to the majority of experts' opinions. Different families of aggregation operators have been studied [14–21]. Among them, the Yager's Ordered Weighted Averaging (OWA) operator is the most widely used [20].

The aggregation operation by way of a quantifier guided OWA operator, ϕ_Q, is calculated as follows:

$$p_{ij}^{c} = \phi_Q\left(p_{ij}^{1}, \ldots, p_{ij}^{m}\right) = \sum_{k=1}^{m} w_k \cdot p_{ij}^{\sigma(k)},$$

where σ is a permutation function such that

$$p_{ij}^{\sigma(k)} \geq p_{ij}^{\sigma(k+1)}, \forall k = 1, \ldots, m-1$$

and Q is a fuzzy linguistic quantifier [22] that indicates the concept of fuzzy majority and it is used to obtain the vector of ϕ_Q, $W = (w_1,\ldots,w_n)$.

Exploitation phase. In this phase, the global information about the alternatives is converted into a global ranking of them. The global ranking is calculated by applying two choice degrees of alternatives to the collective fuzzy preference relation [23]: the *quantifier guided dominance degree* and the *quantifier guided non-dominance degree.*

1. *Quantifier guided dominance degree:* For the alternative xi we obtain the quantifier guided dominance degree, $QGDD_i$, as follows:

$$QGDD_i = \phi_Q(p_{ij}^{c}, j = 1, \ldots, n).$$

   It is used to quantify the dominance that alternative x_i has over all the others alternatives in a fuzzy majority sense.
2. *Quantifier guided non-dominance degree:* For the alternative x_i we deduce the quantifier guided non-dominance degree, $QGNDD_i$, through the expression:

$$QGNDD_i = \phi_Q(1 - p_{ij}^{s}, j = 1, \ldots, n).$$

   where $p_{ji}^{s} = max\left\{p_{ji}^{c} - p_{ij}^{c}, 0\right\}$.

In this situation $QGNDD_i$ is used to quantify the degree in which each alternative is not dominated by a fuzzy majority of the rest of alternatives.

Finally, the solution is obtained by applying one of these two choice degrees and making a ranking of alternatives from the best to the worst according to the valuation degrees.

3 Statistical Comparative Study: Experimental Design

In a GDM problem, the exploitation phase plays a fundamental role in order to solve the decision problem. It is therefore worth conducting research to determine whether or not the use of different choice degrees could affect the selection process. Furthermore, if this was the case, the decision of one choice degree or another could prove to be an important decision tool.

In this paper the hypothesis that we are testing can be stated as follows:

The application of the QGDD and QGNDD choice degrees in GDM problems do not produce significant differences in the final ranking of alternatives.

Although with the application of different choice degrees, different final ranking of alternatives have been observed [3], no study on the subject has been presented to date.

To test the hypothesis above, fifty pairs of dominance and non-dominance degrees, are obtained from fuzzy collective preference relations randomly generated for each alternatives (n = 4, 6, 8), and therefore we ended having repeated measurements on a single sample [24].

We have to analyse two related samples. The usual parametric test used in these cases is the t test. However, this test requires for its application two conditions: the assumption of normality and independence of the distribution of the difference scores in the population from which the random sample of fuzzy preference relation is drawn. On the one hand, we consider unrealistic in our context these assumptions, as no evidence can be provided to support them. On the other hand, by not requiring these stringent assumptions we can achieve greater generality with our conclusions. Therefore, we conclude that nonparametric test are the most appropriate in our experimental study [24–26].

In this context, continuous data and two related samples, the main nonparametric tests available are the sign test and the Wilcoxon signed-rank test [24, 26–28]. The sign test calculates the differences between two variables and classifies the differences as positive, negative, or zero (tied). If two variables have the same distribution (null hypothesis), the median of the differences between the two variable scores is zero, i.e. if the null hypothesis is true we would expect about half the differences to be negative and half to be positive. The null hypothesis is rejected if ‘too few’ differences of one sign occur. An alternative test to apply in this context is the Wilcoxon signed-rank test, which takes into account information from the sign of the differences and their magnitude so that they are appropriately ranked in order of absolute magnitude. Since this test incorporates more information about the data, it is more powerful than the sign test, and therefore preferable to use in our limited study. For a more detailed information, see [29].

For a study that quantitatively describe information about the variability of the QGDD and the QGNDD, we used the coefficient of variation (CV), also known as relative standard deviation (RSD). CV is a measure of the dispersion of a data distribution that is defined as the ratio between the standard deviation and the mean.

$$CV = \frac{\sigma_Y}{\bar{y}}$$

With

$$\sigma_Y = \sqrt{\sigma_Y^2},\ \sigma_Y^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2,\ \text{and}\ \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

where $y_i \in \{y_1, \ldots, y_N\}$ a set of numerical data for variable Y.

The value of the CV is independent of the unit in which the measurement was taken, so it is very useful to compare different distributions with different units or different means [24].

The interpretation used in this study is as follows:

A higher value of CV implies greater data dispersion and greater variability, and, in this situation, greater discrimination into classes.

4 Statistical Comparative Study: Experimental Results

4.1 Wilcoxon Signed-Rank Test: Experimental Results

A total of fifty pairs of degrees (QGDD & QGNDD) obtained from randomly generated fuzzy collective preference relations for each one of the possible alternatives ($n = 4, 6, 8$), and we ended having repeated measurements on a single sample.

In the following, we summarize the percentage of cases that were found to be significantly different according to the Wilcoxon matched-pairs signed-ranks statistical test.

Table 1. Wilcoxon signed-ranks statistical test results

	4 alternatives	6 alternatives	8 alternatives	Global
p-value	0.00	0.00	0.00	0.00
%	100	100	100	100

We assume that two measures with test *p*-value under the null hypothesis lower than or equal to 0.05 (α) will be considered as significantly different; we refer to it as the test being significant and therefore we conclude that the hypothesis tested is to be rejected.

Table 1, shows the percentage of tests with *p*-value lower than or equal to 0.05 (α) for each one number of alternatives used in our experimental study.

In summary, we conclude that the hypothesis tested is to be rejected. Therefore, the application of different choice degrees affects significantly the final ranking of alternatives.

4.2 Descriptive Study: Experimental Results

For the rest of the study we generated randomly collective preference relations (P^C) as follows:

– 4 alternatives: 60 P^C, 240 pairs of choice degrees.
– 6 alternatives: 40 P^C, 240 pairs of choice degrees.
– 8 alternatives: 30 P^C, 240 pairs of choice degrees.

Maintenance of the ranking of alternatives: Experimental results. Based on the comparison of data obtained for QGDD and QGNDD Figure summarised the coincidence in the first (best) alternative in both degrees and in other positions.

Table 2. Coincidence first alternative in percentage

%	4 alt	6 alt	8 alt	Global
1°	75	62,5	63,3	68
Other	25	37,5	36,7	32

As we can observe in Table 2, the coincidence in the first (best) alternative for both choice degrees is greater than 60% but smaller than 80% in all cases.

Coefficient of variation (CV): Experimental results. For a collective preference relation (PC), we can calculate QGDD and QGNDD values for each alternative, and later obtain CV's in both situations

For comparing the results, we use the difference:

$$CV_{QGDD} - CV_{QGNDD}$$

Results obtained are shown in next Table.

Table 3. Differences between CV's in percentages

%	Global	4 alt	6 alt	8 alt
Negative	3	4	0	3
Positive	97	96	100	97
	100	100	100	100

From Table 3 we can conclude that the dispersion is higher in QGDD than in QGNDD.

To continue our study about the distribution of the differences between the QGDD and QGNDD results, we define the variation rate as the ratio between CVQGDD - CVQGNDD and CVQGNDD and express the results as percentages.

$$Variation\,Rate = \frac{CV_{QGDD} - CV_{QGNDD}}{CV_{QGNDD}}$$

Tables 4, 5 and 6 depicted in Figs. 1 and 2 shown the distribution of the variation rate for each possible alternatives (n = 4, 6, 8).

Example. Let P^C be a random collective preference relation for 4 alternatives:

$$P^C = \begin{pmatrix} 0. & 0.87615 & 0.97178 & 0.59443 \\ 0.12385 & 0. & 0.52458 & 0.24997 \\ 0.02822 & 0.47542 & 0. & 0.90911 \\ 0.40557 & 0.75003 & 0.09089 & 0. \end{pmatrix}$$

Table 7 shows the quantifier guided degrees $QGDD_i$ and $QGNDD_i$ for this case in which 4 alternatives have been considered.

Table 4. Distribution of the variation rate for 4 alternatives

Intervals	Absolute frequency	Relative frequency*100%
−50 – −25	1	2
−25 – 0	2	3
0–25	17	28
25–50	23	38
50–75	9	15
75–100	6	10
100–125	2	3
	60	100 aprox

Table 5. Distribution of the variation rate for 6 alternatives

Intervals	Absolute frequency	Relative frequency*100%
0–30	14	35
30–60	19	48
60–90	6	15
90–120	0	0
120–180	0	0
180–210	0	0
210–240	1	3
	40	100 aprox

Table 6. Distribution of the variation rate for 8 alternatives

Intervals	Absolute frequency	Relative frequency*100%
−15 – 0	1	3
0–15	1	3
15–30	9	30
30–45	10	33
45–60	4	13
60–75	5	17
	30	100 aprox

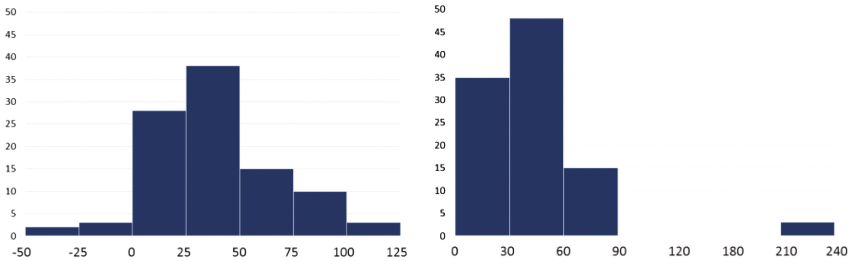


Fig. 1. Distribution for 4 and 6 alternatives

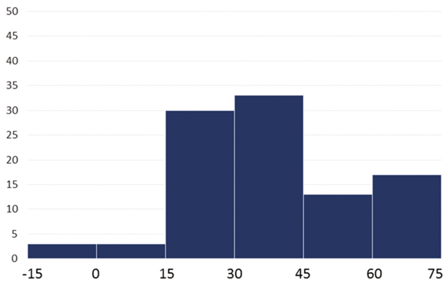


Fig. 2. Distribution for 8 alternatives

Table 7. Quantifier guided degrees $QGDD_i$ and $QGNDD_i$

$QGDD_i$		Position	QGNDD1		Position
$QGDD_1$	2,44235	1	QGNDD1	3	1
$QGDD_2$	0,89840	4	QGNDD2	1,74764	4
$QGDD_3$	1,41276	2	QGNDD3	2,00729	2
$QGDD_4$	1,24649	3	QGNDD4	1,99292	3
CV_{QGDD}	0,38323		CV_{QGNDD}	0,21976	

The final ranking of alternatives for both choice degrees is:

$$X1 > X3 > X4 > X2$$

Then we compare the differences in the ranking of alternatives between QGDD and QGNDD.

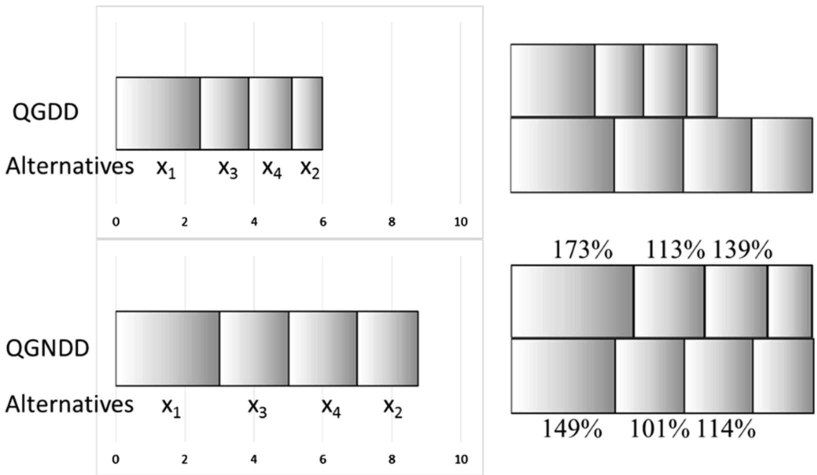


Fig. 3. Ranking of alternatives

As shown in Fig. 3, when normalizing these values the differences between the alternatives are visually greater in the ranking of alternatives obtained with QGDD than in the one obtained with QGNDD. To see it accurately, the ratio between two consecutive alternatives for each ranking is also included.

5 Conclusion

In this paper we have considered two choice degrees used in the selection process for GDM problems, QGNN and QGNDD. We have presented a comparative experimental study based on the use of the nonparametric Wilcoxon statistical test and the coefficient of variation. The results has shown that the compared choice degrees produce significant different solutions, i.e. ranking of alternatives. The subsequent analysis of the data allows us to conclude that the rankings of alternatives obtained by QGDD generally present higher variability than the rankings obtained by QGNDD and, therefore, allow a greater discrimination among the alternatives.

In future, we will address this problem from a theoretical point of view and we will apply other statistical tools.

Acknowledgements. The authors would like to acknowledge FEDER financial support from the Projects TIN2013-40658-P and TIN2016-75850-R.

References

1. Chiclana, F., Herrera, F., Herrera-Viedma, E., Martinez, L.: A note on the reciprocity in the aggregation of fuzzy preference relations using OWA operators. Fuzzy Sets Syst. **137**(1), 71–83 (2003)
2. Pedrycz, W., Parreiras, R., Ekel, P.: Fuzzy Multicriteria Decision-Making. Models, Methods and Applications. Wiley, Chichester (2011)
3. Chiclana, F., Herrera, F., Herrera-Viedma, E., Poyatos, M.: A classification method of alternatives for multiple preference ordering criteria based on fuzzy majority. J. Fuzzy Math. **4**, 801–813 (1996)
4. Orlovsky, S.A.: Decision-making with a fuzzy preference relation. Fuzzy Sets Syst. **1**(3), 155–167 (1978)
5. Millet, I.: The effectiveness of alternative preference elicitation methods in the analytic hierarchy process. J. Multi-Criteria Decis. Anal. **6**(1), 41–51 (1997)
6. Kacprzyk, J.: Group decision making with a fuzzy linguistic majority. Fuzzy Sets Syst. **18** (2), 105–118 (1986)
7. Chiclana, F., Herrera, F., Herrera-Viedma, E.: Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. Fuzzy Sets Syst. **97** (1), 33–48 (1998)
8. Chiclana, F., Herrera, F., Herrera-Viedma, E.: A note on the internal consistency of various preference representations. Fuzzy Sets Syst. **131**(1), 75–78 (2002)
9. Carlsson, C., Ehrenberg, D., Eklund, P., Fedrizzi, M., Gustafsson, P., Lindholm, P., Merkuryeva, G., Riissanen, T., Ventre, A.G.S.: Consensus in distributed soft environments. Eur. J. Oper. Res. **61**(1–2), 165–185 (1992)

10. Perez, I.J., Cabrerizo, F.J., Herrera-Viedma, E.: A mobile decision support system for dynamic group decision-making problems. IEEE Trans. Syst. Man Cybernet. Part A Syst. Hum. **40**(6), 1244–1256 (2010)
11. Kacprzyk, J., Fedrizzi, M., Nurmi, H.: Group decision making and consensus under fuzzy preferences and fuzzy majority. Fuzzy Sets Syst. **49**(1), 21–31 (1992)
12. Kacprzyk, J., Nurmi, H., Fedrizzi, M.: Consensus Under Fuzziness, vol. 10. Kluwer Academic Publishers, Dordrecht (1997)
13. Herrera, F., Herrera-Viedma, E., Verdegay, J.L.: Choice processes for non-homogeneous group decision making in linguistic setting. Fuzzy Sets Syst. **94**(3), 287–308 (1998)
14. Chiclana, F., Herrera-Viedma, E., Herrera, F., Alonso, S.: Some induced ordered weighted averaging operators and their use for solving group decision making problems based on fuzzy preference relations. Eur. J. Oper. Res. **182**(1), 383–399 (2007)
15. Dubois, D., Prade, H.: A review of fuzzy set aggregation connectives. Inf. Sci. **36**(1–2), 85–121 (1985)
16. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht (1994)
17. Herrera, F., Herrera-Viedma, E., Chiclana, F.: A study of the origin and uses of the ordered weighted geometric operator in multicriteria decision making. Int. J. Intell. Syst. **18**(6), 689–707 (2003)
18. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty and Information. Prentice Hall, New Jersey (1988)
19. Xu, Z.S., Da, Q.L.: An overview of operators for aggregating information. Int. J. Intell. Syst. **18**(9), 953–969 (2003)
20. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. Syst. Man Cybernet. **18**(1), 183–190 (1988)
21. Yager, R.R.: Quantifier guided aggregation using OWA operators. Int. J. Intell. Syst. **11**(1), 49–73 (1996)
22. Zadeh, L.A.: A computational approach to fuzzy quantifiers in natural languages. Comput. Math Appl. **9**(1), 149–184 (1983)
23. Herrera, F., Herrera-Viedma, E.: Choice functions and mechanisms for linguistic preference relations. Eur. J. Oper. Res. **120**(1), 144–161 (2000)
24. Siegel, S.: Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York (1956)
25. Lehmann, E.: Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, San Francisco (1975)
26. Wilcoxon, F.: Individual comparisons by ranking methods. Biom. Bull. **1**(6), 80–83 (1945)
27. Del Moral Avila, M.J., Tapia Garcia, J.M.: Tecnicas estadisticas aplicadas. Grupo Editorial Universitario, Granada (2006)
28. Rohatgi, V.K., Saleh, A.K.M.E.: An Introduction to Probability and Statistics, 2nd edn. Wiley-Interscience, New York (2011)
29. Chiclana, F., Garcia, J.M.T., del Moral Avila, M.J., Herrera-Viedma, E.: A statistical comparative study of different similarity measures of consensus in group decision making. Inf. Sci. **221**, 110–123 (2013)

Using Bibliometrics and Fuzzy Linguistic Modeling to Deal with Cold Start in Recommender Systems for Digital Libraries

Alvaro Tejeda-Lorente[1(✉)], Juan Bernabé-Moreno[1], Carlos Porcel[2], and Enrique Herrera-Viedma[1]

[1] Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain
atejeda@decsai.ugr.es

[2] Department of Computer Science, University of Jaén, Jaén, Spain

Abstract. Every recommender system approach suffers the cold start problem to a greater or lesser extent. To soften this impact, the more common solution is to find the way of populating users profiles either using hybrid approach or finding external data sources. In this paper, we present a fuzzy linguistic approach that using bibliometrics aids to soft or remove the necessity of interaction of users providing them with personalized profiles built beforehand, thus reducing the cold start problem. To prove the effectiveness of the system, we conduct a test involving some researchers, aiming to build their profiles automatically. The results obtained proved to be satisfactory for the researchers.

Keywords: Recommender system · Cold start · Fuzzy linguistic modeling · Digital library

1 Introduction

In the era of Big Data, the amount of information generated on every field in the Web is growing constantly leading to the well known information overload problem [6]. Recommender Systems (RSs) appear as a natural solution to this problem providing personalized recommendations to their users filtering out the non valuable information for them [9]. However, a key feature which defines a recommender system are the user profiles that allow them to provide users with recommendations that are suiting them better. One of the main problems RSs have is the *cold start* problem, that is, when a user or item is new to the system and nothing is known about him/her or it. This is a known problem that has been addressed several times in the literature [23]. However, the importance of providing a suitable solution to this problem is scaling due to the growing presence in the Web of systems requiring personalization as well as real time interaction.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_36

In the academic world, research is a field of main importance. It is an over specialized field where each specialty is quite specific. In a previous work, we presented REFORE, a quality-based fuzzy linguistic recommender system for researchers [24] which main purpose was to aid researchers by keeping them up-to-date regarding the new articles that might be relevant for them. In order to solve the cold start problem, researchers must select the top 5 articles of their profiles that show better the current research topics they are interested on. Afterwards, they should provide them with a linguistic assessment of the importance of each of them. Same process is required for the keywords that define better his current research interest.

To deal with the cold start problem in most of the systems, it requires a previous set up of the user profile done by users themselves, requiring time and a fix and complete profile into the system [17]. REFORE was conceived as a system which delivered to researchers every month an email with the most relevant papers for them from this same month, working as a service to keep the authors up-to-date on their research topics. However, the necessity of a previous and good set up stopped new users to try the tool. Therefore, to reduce the cold start problem removing the interaction of the user to the minimum is basic to provide the system with a better acceptance degree among users. It will allow us to expand REFORE to other possibilities such like being used without any necessity of registration.

The proposed approach provides automatically profiles for users based on their names by extracting the information from their historic research trajectory. This profile is used by REFORE reducing the cold start problem impact that forced users to provide all of this information beforehand.

The aim of this paper is to present a semi automatic fuzzy linguistic solution for the cold start problem in REFORE. This solution that could be applied in whatever bibliographic database from a University Digital Library allowing us to open REFORE to a wider research community by reducing the necessity of a previous profile set up done the users. The major innovations and contributions of the solution include:

1. The provision of automatic profiles extracted from the authors name who are using the system dealing with (incomplete information) and transforming novelty, frequency and quality on users interest.
2. The ability of using the system within less than 1 min without the necessity of registering on it, providing the researchers the possibility of obtaining the more relevant publications for themselves.
3. The system uses fuzzy linguistic modeling to improve the user-system interactions.

The paper is organized as follows: In Sect. 2 the background is presented, that is the basis of recommender systems, fuzzy linguistic modeling and other approach to the cold start problem; Sect. 3 presents the new method for the user profile generation; Sect. 4 addresses the validation of the system, and Sect. 5 offers conclusions based on the study findings.

2 Background

With this section we provide the needed background information to describe our system. First, we will present a short description of recommender systems, then a brief explanation of the cold start problem followed by a description of the fuzzy linguistic modeling.

2.1 Recommender Systems

RSs produce personalized recommendations as output or guide users in a personalized way through a wide range of possible options [2]. Well known examples of successful use of RSs are given in e-commerce [3,19], health [7] or learning [15]. In order to do that, the system must have knowledge from users. This knowledge can be obtained from different sources and has to be related directly or indirectly with the recommendations provided by the system. That can be done in an *implicit way* through the normal functioning of the system, i.e.: Ratings from a movie, geographical proximity to a shop, preferences regarding tastes, etc., or in a *explicit way* when users are required to provide the information manually to the system [8]. Some systems, as for example some movie recommenders, force users to fill some questionnaires or to rate certain selected movies before any recommendation could be received in order to avoid the cold start problem.

Different categorizations have been proposed for recommender systems based on the approach followed to generate recommendations, being the one who split them on two categories the more extended: content-based and collaborative [9]. Content-based recommender systems are based on the similarity of an user profile with an item profile, meanwhile in collaborative recommender systems the recommendations are generated based on the ratings or behavior provided by other similar users.

On the one hand, collaborative systems [10] use to perform better in some domains adding diversity to the recommendations. However, those systems required an important amount of information gathered from the users behavior making them relatively weaker when dealing with the cold start problem. On the other hand, content-based approaches [20] perform better with new users where their taste is rapidly defined. They have the problem of lacking diversity and serendipity. Each approach has advantages and disadvantages, the combination of the both in a hybrid system tend to mitigate the problems they have [2], e.g.: content-based deal in a better way with the cold start problem, so combining this approach with the collaborative benefits from the advantages of the both. However these recommender systems tend to fail when little is known about users information needs.

2.2 Cold Start Problem

Cold start problem is present in certain information systems where the lack of knowledge affect the system purpose [23]. It is particularly present in recommender systems where the basic functioning is based on the amount of knowledge accumulated over users or items [17]. We found two main variants: new user cold start problem and new item cold start problem. Since most of the time recommender systems use historical ratings as part of user profiles or item profiles, the issue is present on the both sides.

Different approaches have been trying to deal with the cold start problem with different results. On the one hand, hybrid recommendation approaches, where the mix of different recommendation techniques are used in order to take advantage of each other's strengths [22]. E.g.: softening the problem of collaborative filtering with content-based support [4,8]. On the other hand, since each new item or user introduced in the system presents a problem, implicitly or explicitly populating profiles is a common solution. User profiles are enriched through information either inferred through some technique or provided by users themselves, meanwhile item profiles are done through rich metadata descriptions or accelerating the rating acquisitions [25]. E.g.: In [17], authors used the binary classifier C4.5 [16] and Naive Bayes algorithm [21] in a previous phase to build user profiles.

Le Hoan Son in [23] present a review of different algorithms and their effectiveness against cold start. A classification into three groups is proposed:*making use of additional data sources*, *selecting the most prominent groups of analogous users* and *enhancing the prediction using hybrid methods*. Results showed a better performing of the algorithm denominated *new heuristic similarity model* [18] which belong to the second group and has no need of additional information.

The basic idea of the most of the solutions is to extract some information not provided in the moment the user profile is set based on the rest of the available information. e.g.: Clustering of users based on common characteristics like geographical information, or adding information from an external source.

2.3 Fuzzy Linguistic Modeling

Information is not always able to be evaluated in a quantitative manner, in some occasions it has to be assessed in a qualitative way. The fuzzy linguistic modeling is based on the concept of *linguistic variable* [26] which has proven good results for modeling qualitative information in many problems [14]. Some classic solutions when it comes to fuzzy linguistic modeling are: classic fuzzy linguistic modeling [1,26] and ordinal fuzzy linguistic modeling [5].

A typical problem when it comes to fuzzy linguistic modeling is the loss of information that use to happen with approaches like classical and ordinal [26]. In [12] authors present the 2-tuple approach for fuzzy linguistic modeling. It consists on a continuous model of representation of information that allows to reduce the typical information loss.

Let $S = \{s_0, ..., s_g\}$ be a linguistic term set with odd cardinality, where the mid term represents an indifference value and the rest of the terms are symmetrically related to it. We assume that the semantics of labels are given by means of triangular membership functions and consider all terms distributed on a scale on which a total order is defined [11]. If a symbolic method aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, ..., g\}$, then an approximation function is used to express the result in S. β is represented by means of 2-tuples (s_i, α_i), where $s_i \in S$ represents the linguistic label of the information, and α_i is a numerical value expressing the value of the translation from the original result β to the closest index label, i, in the linguistic term set ($s_i \in S$). This model defines $\Delta(\beta) = (s_i, \alpha)$ and $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ as a set of transformation functions between numeric values and 2-tuples.

In order to establish the computational model a definition of a negation, comparison and aggregation operators is needed. Using the transformation functions above described Δ and Δ^{-1} that avoid the loss of information, any of the existing aggregation operators can be easily extended for dealing with linguistic 2-tuples [12].

When modeling the information a problem arises if different uncertainty degrees on the phenomenon are perceived. In order to deal with that matter an important parameter to determine known as the "granularity of uncertainty" is needed, i.e., the cardinality of the linguistic term set S [13]. In [13] a multi-

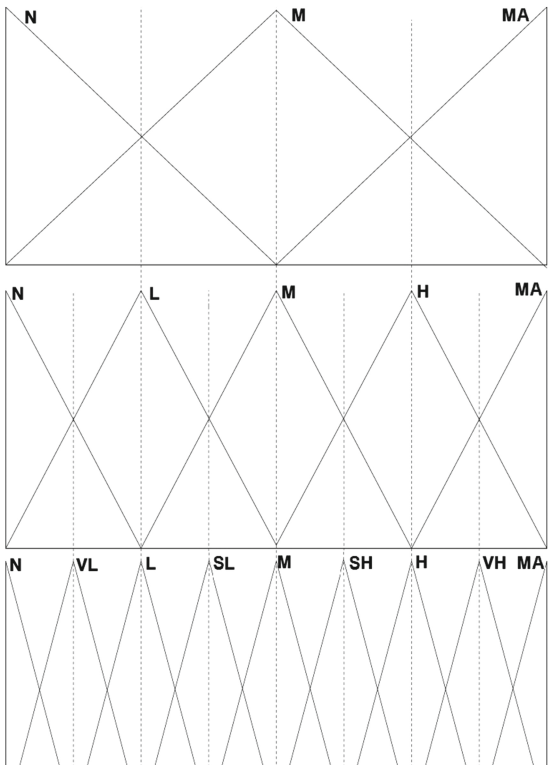


Fig. 1. Linguistic hierarchy of 3, 5 and 9 labels.

granular 2-tuple fuzzy linguistic modeling based on the concept of linguistic hierarchy is proposed.

A *Linguistic Hierarchy, LH,* is a set of levels *l(t,n(t))*, where each level t is a linguistic term set with different granularity $n(t)$ from the remaining hierarchy levels. The levels are ordered according to their granularity, i.e., a level $t+1$ provides a linguistic refinement of the previous level t. We can define a level from its predecessor level as: $l(t, n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$. A graphical example of a three level linguistic hierarchy is shown in Fig. 1. Using this LH, the linguistic terms in each level are the following:

- $S^3 = \{a_0 = Null = N,\ a_1 = Medium = M,\ a_2 = Maximum = MA\}$.
- $S^5 = \{b_0 = None = N,\ b_1 = Low = L,\ b_2 = Medium = M,\ b_3 = High = H,\ b_4 = Maximum = MA\}$
- $S^9 = \{c_0 = None = N,\ c_1 = Very\_Low = VL,\ c_2 = Low = L,\ c_3 = Slightly\_Low = SL,\ c_4 = Medium = M,\ c_5 = Slightly\_High = SH,\ c_6 = High = H,\ c_7 = Very\_High = VH,\ c_8 = Maximum = MA\}$

In [13] authors remark that the family of transformation functions between labels from different levels is bijective, guarantying that the transformations between levels are produced without loss of information in a linguistic hierarchy.

3 Proposal Description

In this section we present an automatic academic profiles builder for users in order to deal with the cold start problem using multi-granular fuzzy linguistic modeling and bibliometrics. We work over REFORE, a recommender system for researchers introduced in our previous work [24]. First, we will see the architecture and approach followed. Then, we will go through the representation of the information as well as the resources. We will conclude with profile formation of the researchers.

3.1 System Concepts

The approach works based on the following concepts:

- Researchers does not want to spend time building a profile. However there is a better predisposition to adjust one done beforehand.
- *Cold Star* problem is solved adding extra information from users from additional data sources. The system extracts all the necessary information from Scopus based on the name of the researcher.
- Due to the nature of research itself, authors needs of information or investigation interests are very specific.
- The use of bibliometric quality measures worked in REFORE [24] as part of the recommendation approach. Authors tend to public their best works in the best possible journals.

– Novelty. To follow the hype. Authors will be more interested in topics related with their last works.
– Frequency. The more recurrent topics from authors will mark an important lines
– Authorship. Authors tend to be first authors in those works they lead.

REFORE profiles key needs for its correct functioning are two: keywords for the main search and papers for the bibliometrics filters applied. The process followed by the profile builder after a researcher name is introduced is shown on Fig. 2 and consists on:

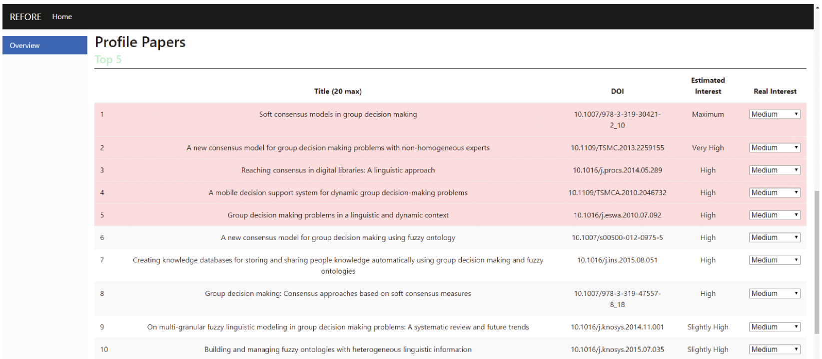


Fig. 2. Operating scheme

1. Query the Scopus API [1] for author information being papers the most important.
2. Estimate the importance degree of the paper on the author profile. Details explained below.
3. Extracting keywords from all the user papers weighted with the importance of each.

3.2 Information Representation

In order to represent the information we stayed with the same linguistic hierarchy defined in REFORE [24] but using only two levels. The concepts to asses in this work are the following:

– The *Importance degree* of keywords for the users, which is assessed in S^5.
– The *Relevance degree* of a paper for a user, which is assessed in S^9.

We propose to use a linguistic hierarchy which linguistics term sets are:

[1] https://dev.elsevier.com/scopus.html.

- $S^5 = \{b_0 = None = N, b_1 = Low = L, b_2 = Medium = M, b_3 = High = H, b_4 = Maximum = MA\}$
- $S^9 = \{c_0 = None = N, c_1 = VeryLow = VL, c_2 = Low = L, c_3 = SlightlyLow = SL, c_4 = Medium = M, c_5 = SlightlyHigh = SH, c_6 = High = H, c_7 = VeryHigh = VH, c_8 = Maximum = MA\}$

Level 1 is used to represent the importance degree of keywords ($S_1 = S^9$) and for the predicted relevance degrees we use the level 2 ($S_2 = S^5$).

3.3 Profiles Construction

Author names are required in the format used for publishing. Afterwards, Scopus API is used to look for authors profiles and retrieve them together with their research history. The system performs a quick analysis on the papers splitting user profiles in two: *Papers* and *Keywords*.

In REFORE, *papers* were split in two groups, *selected* and *non-selected*. This classification was used for the filtering process. In order to provide the same classification, the following characteristics have been considered to estimate the individual importance degree of each one.

Given a paper P, P_{iu} is the paper i from user u estimated aggregating the following paper characteristics related to the user:

- Quality, given by the *Impact Factor*:

$$IF_{iu} \begin{cases} IF_{iJ} & if\ J = ranked\ journal \\ 0.4 & Otherwise \end{cases} \tag{1}$$

  where J is the source where the paper i from user u was published. We consider authors tend to publish their better works in the best journals, giving slightly less importance to conferences.
- *Novelty*, given by the publication date:

$$No_{iu} \begin{cases} Current\_Year - Year(P_{iu}) * 0.5 & if\ Current\_Year - Y > 5 \\ 0.4 & Otherwise \end{cases} \tag{2}$$

  where $Year(P_{iu})$ is the year when the paper P_{iu} was published. We considered that authors tend to be more interested on the research lines they are working in the present.
- *Authorship*, given by the occupied position on the authors line in the paper:

$$Au_{iu} \begin{cases} 2 & if\ Current\_Year - Y > 5 \\ 0.4 & Otherwise \end{cases} \tag{3}$$

  where being the first author in the list is considered as being the lead carrier of work.

Before aggregating, all results are normalized within the interval $[0, 1]$. Different weights distribution are applied: 20% for the IF, 50% for No and 30% the Au.

Given the necessity of operating in the following steps with the linguistic set of keywords which are expressed with labels from level 2 of our linguistic hierarchy, that is $S_1 = S^9$, a linguistic transformation is needed. We followed the representation model described in [12] to transform the aggregated values to their linguistic labels belonging to the level 3 of our linguistic hierarchy, i.e.: $S_1 = S^9$. Level 3 was chosen to allow users a bigger margin when reviewing the papers profile.

On the other hand, *keywords* are used for the similarity estimation in REFORE. In order to obtain the importance degree on each of them we applied the linguistic weighted average (see Definition 1). The importance degree results as the average of each keyword appearing on user papers, each keyword inherits the importance degree of the paper it appears.

Definition 1. Linguistic Weighted Average Operator *[24]. Let* $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ *be a set of linguistic 2-tuples and* $W = \{(w_1, \alpha_1^w), ..., (w_n, \alpha_n^w)\}$ *be their linguistic 2-tuple associated weights. The 2-tuple linguistic weighted average* $\overline{x}_l^w$ *is:*

$$\overline{x}_l^w[((r_1, \alpha_1), (w_1, \alpha_1^w))...((r_n, \alpha_n), (w_n, \alpha_n^w))] = \Delta(\frac{\sum_{i=1}^n \beta_i \cdot \beta_{W_i}}{\sum_{i=1}^n \beta_{W_i}}), \quad (4)$$

with $\beta_i = \Delta^{-1}(r_i, \alpha_i)$ and $\beta_{W_i} = \Delta^{-1}(w_i, \alpha_i^w)$.

As above mentioned keywords are expressed withing the level 2 of the linguistic hierarchy used. So in order to be used by REFORE a transformation between level is performed.

4 Experiments and Approach Evaluation

In this section we present the evaluation of the proposed approach. Due to the nature of REFORE, the system objective of this approach proposition is to alleviate the cold start problem. No comparison with other approaches is possible since no standard data set is used. Thus, in this study we only perform online experiments, i.e., practical studies where a group of researchers indicate their optimum profile. Users input is compared to the estimated one by our approach.

In order to test the effectiveness of the approach followed and after adjusting the different parameters to the optimum weight, i.e.: Quality, novelty and authorship, the experiment is performed over a set of users from REFORE. Other approaches test the validity of their solutions for cold start problem through evaluating recommendations, in this work we propose a direct evaluation of the profiles by the own users. So our experiment consists on showing to the users the profiles built for them together with the estimated importance values for each paper and keyword. Afterwards, the user is inquired for the real ones.

For that purpose a section in REFORE was created for the test group (see Fig. 3).

The test group consisted on 20 researchers with different profiles, going from a more junior research profiles to senior ones. After loading their profiles from

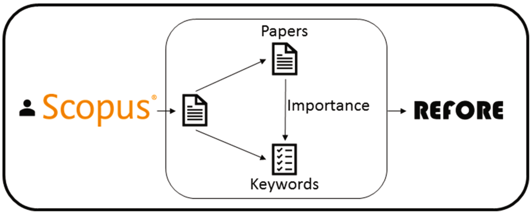


Fig. 3. Evaluation page

Scopus extracting papers and keywords, the system estimates their importance degree following the steps indicated in Sect. 3 and shows them to the user.

When it comes to recommender systems, the most common measures for accuracy are precision, recall or F1 [20]. However, in this work we left to users the evaluation of their own estimated profile, so in order to measure how accurate the system is mirroring user interests we use the **Mean Absolute Error (MAE)** [10]. In particular, we defined it in a linguistic framework:

$$MAE = \Delta(g \times \frac{\sum_{i=1}^{n} abs(\Delta^{-1}(p_i, \alpha_{pi}) - \Delta^{-1}(r_i, \alpha_{ri}))}{n}) \quad (5)$$

with $MAE \in S_1 \times [-0.5, 0.5]$, and where g is the granularity of the term set used to express the relevance degree, i.e. S^9, n is the the number of cases in the test set, (p_i, α_{pi}) is the predicted 2-tuple linguistic value for paper or keyword i and (r_i, α_{ri}) the real one.

We evaluated both, papers and keywords separately, importance degree for keywords and relevance degree for papers, obtaining the following MAE results:

- **Keywords**: 0.127
- **Papers**: 0.104

We observe that the profiles estimations generated with the proposed approach are in line with the real users preferences, softening or removing the necessity of their interaction with the system to establish a previous user profile.

5 Concluding Remarks

In this paper we propose a fuzzy linguistic approach based on bibliometrics to deal with the cold start problem for researchers present in the REFORE system.

As we experienced in REFORE, the first barrier users find to use a recommender system is the building of their own profiles. The idea of automatize the profile construction will support the system as well as will enable the creation of different system more oriented to real time interactions.

A user historic research record is retrieved from Scopus. The system split keywords and the rest of meta-information from the papers from authors. Afterwards, the importance degree of each paper and keyword is obtained based on the idea of novelty, frequency and quality. We have applied the approach over our previous recommender system REFORE in a real environment with satisfactory results. Those results, showed that the approach performance was better within the top 5 element of the each list, keywords and papers.

As future work, we consider to study the inclusion of the automatic profiling on a real time recommender system for researcher with no need of previous registration.

Acknowledgments. This paper has been developed with the FEDER financing under Projects TIN2013-40658-P and TIN2016-75850-R

References

1. Bordogna, G., Pasi, G.: A fuzzy linguistic approach generalizing boolean information retrieval: A model and its evaluation. J. Am. Soc. Inf. Sci. **44**(2), 70 (1993)
2. Burke, R.: Hybrid web recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web. LNCS, vol. 4321, pp. 377–408 (2007)
3. Chen, P.Y., Wu, S.Y.: Does collaborative filtering technology impact sales? empirical evidence from amazon.com (2007). https://ssrn.com/abstract=1002698
4. Claypool, M., Gokhale, A., Miranda, T.: Combining content-based and collaborative filters in an online newspaper. In: Proceedings of the ACM SIGIR 1999 Workshop on Recommender Systems-Implementation and Evaluation, pp. 714–720 (1999)
5. Delgado, M., Herrera, F., Herrera-Viedma, E., Martinez, L.: Combining numerical and linguistic information in group decision making. Inf. Sci. **107**(1–4), 177–194 (1998)
6. Edmunds, A., Morris, A.: The problem of information overload in business organizations: a review of the literature. Int. J. Inf. Manage. **20**, 17–28 (2000)
7. Esteban, B., Tejeda-Lorente, A., Porcel, C., Arroyo, M., Herrera-Viedma, E.: TPLUFIB-WEB: A fuzzy linguistic web system to help in the treatment of low back pain problems. Knowl. Based Syst. **67**, 429–438 (2014)
8. Good, N., Schafer, J., Konstan, J., Borchers, A., Sarwar, B., Herlocker, J., Riedl, J.: Combining collaborative filtering with personal agents for better recommendations. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence, AAAI/IAAI, pp. 439–446 (1999)
9. Hanani, U., Shapira, B., Shoval, P.: Information filtering: Overview of issues, research and systems. User Model. User-Adap. Inter. **11**, 203–259 (2001)
10. Herlocker, J., Konstan, J., Terveen, L., Riedl, J.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. **22**(1), 5–53 (2004)
11. Herrera, F., Herrera-Viedma, E.: Choice functions and mechanisms for linguistic preference relations. Eur. J. Oper. Res. **120**, 144–161 (2000)
12. Herrera, F., Martínez, L.: A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans. Fuzzy Syst. **8**(6), 746–752 (2000)
13. Herrera, F., Martínez, L.: A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multiexpert decision-making. IEEE Trans. Syst. Man Cybern. Part B Cybern. **31**(2), 227–234 (2001)
14. Herrera-Viedma, E., López-Herrera, A., Luque, M., Porcel, C.: A fuzzy linguistic irs model based on a 2-tuple fuzzy linguistic approach. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **15**(2), 225–250 (2007)

15. Hsu, M.: A personalized english learning recommender system for ESL students. Expert Syst. Appl. **34**, 377–408 (2008)
16. Kotsiantis, S.B.: Supervised machine learning: A review of classification techniques. In: Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies, pp. 3–24. IOS Press, Amsterdam, The Netherlands (2007)
17. Lika, B., Kolomvatsos, K., Hadjiefthymiades, S.: Facing the cold start problem in recommender systems. Expert Syst. Appl. **41**(4, Part 2), 2065–2073 (2014)
18. Liu, H., Hu, Z., Mian, A., Tian, H., Zhu, X.: A new user similarity model to improve the accuracy of collaborative filtering. Knowl. Based Syst. **56**, 156–166 (2014)
19. Long-Sheng, C., Fei-Hao, H., Mu-Chen, C., Yuan-Chia, H.: Developing recommender systems with the consideration of product profitability for sellers. Inf. Sci. **178**(4), 1032–1048 (2008)
20. Lops, P., Gemmis, M., Semeraro, G.: Content-based recommender systems: State of the art and trends. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook, pp. 73–106. Springer (2011)
21. Rish, I.: An empirical study of the naive bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, Vol. 3, pp. 41–46. IBM New York (2001)
22. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 253–260. ACM (2002)
23. Son, L.H.: Dealing with the new user cold-start problem in recommender systems: A comparative review. Inf. Syst. **58**, 87–104 (2016)
24. Tejeda-Lorente, A., Porcel, C., Bernabé-Moreno, J., Herrera-Viedma, E.: Refore: A recommender system for researchers based on bibliometrics. Appl. Soft Comput. **30**, 778–791 (2015)
25. Tejeda-Lorente, A., Porcel, C., Peis, E., Sanz, R., Herrera-Viedma, E.: A quality based recommender system to disseminate information in a university digital library. Inf. Sci. **261**, 52–69 (2014)
26. Zadeh, L.: The concept of a linguistic variable and its applications to approximate reasoning. Part I, Inf. Sci. 8, 199–249 (1975), Part II, Inf. Sci. 8, 301–357 (1975), Part III. Inf. Sci. 9(1975), 43–80 (1975)

Type 2 Fuzzy Control Charts Using Likelihood and Deffuzzification Methods

Hatice Ercan Teksen(✉) and Ahmet Sermet Anagün

Eskisehir Osmangazi University, Eskisehir, Turkey
ercanhatice@gmail.com

Abstract. Besides control charts are used in many fields, they are important because the process gives information about the product's situation. Thanks to control charts, necessary precautions are taken by noticing abnormal and normal situations of process and/or product. It is considered that at this point the most important and critical thing is that there will be loss of information about the expert opinions. It can be said that this situation is more common especially for the qualitative data. To prevent losses of data like this and so on and to transform linguistic expressions into crisp data, it is needed to take advantage of fuzzy logic that is commonly used recently. Although some studies about creating control charts by fuzzy sets have been done recently, all of them are done only by using type 1 fuzzy sets. However, it is known that much of the data used in daily life cannot be expressed by type 1 fuzzy number. Some data may be more suitable for type 2 fuzzy numbers. In this study, type 2 fuzzy control charts are obtained by using the methods of defuzzification and likelihood. The results are compared with the classical control charts This study aims to use type 2 fuzzy sets in control charts as a new approach.

Keywords: Interval type 2 fuzzy control charts · Defuzzification · Likelihood

1 Introduction

Control charts are a tool that provides insight into the process, provides process evaluation, and monitors the progress. Control charts consist of the center line, upper and lower control limits obtained from the data. These limits allow to see changes in the process and obtain information about abnormal situations of the process.

The control charts developed by Shewhart in the 1920's at Bell Laboratories were later used in a variety of fields [1]. Control charts have been developed in order to have an idea about the product/process according to the product/process exhibited and to check whether an unexpected situation has occurred in the process.

Classic control charts are used in many fields. Fuzzy control charts have also recently been found, in literature. The use of fuzzy numbers in control charts is especially advantageous for control charts for attributes. Properties such as color, smell, cracks in a surface, brightness, etc. are quality features that cannot be expressed by any measure or number. Besides the ability to measure attributes data with fuzzy numbers, more than one quality characteristic can be used different from classical control charts, and more flexible control limits can be created using expert opinions [2]. As a result of

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_37

available literature reviews, studies on fuzzy control graphs have been limited to type 1 fuzzy sets [2–10, 12–16].

Faraz and Shapiro have studied Xort-S charts in their work and have benefited from LR-fuzzy numbers for this control chart [3]. Gülbay and Kahraman have identified the UCL and LCL as type 1 fuzzy sets and compared data with fuzzy UCL and LCL. Accordingly, results such as in control, rather in control, out of control and rather out of control interpreted. So it is said that the fuzzy control charts are more flexible than classical control charts [2]. Senturk, Erginel, in their study, have calculated the X_{ort}, R and S values and control limits as fuzzy sets. They used the α-cut method in the decision phase [4]. Alaeddini et al. have drawn Xort control charts using triangular fuzzy numbers and interpreted the results using clustering analysis [5]. Cheng has calculated the average values from triangular fuzzy numbers. He has drawn two separate control graphs using the distance to the possibility and the necessity [6]. Using LR-fuzzy numbers, UCL and LCL are also expressed as fuzzy sets. Then the fuzzy set is interpreted according to the determined α–level [7]. In another study, Gülbay et al. similarly used α-cut fuzzy control charts for linguistic data [8]. Shu and Wub have found triangular fuzzy R and Xort values and compared each of the numbers in two charts to see that they are in control or not [9]. Kaya and Kahraman have used triangular and trapezoidal fuzzy numbers for process capability analysis [10]. In the Asai study, he pointed out that fuzzy logic can be used in control charts for categorized data [11]. Similarly, Woodall et al. and Laviolette et al. (1998) studied fuzzy control charts using categorized data [12, 13] The linguistic terminology as a quality characteristic is first used in Wang and Raz and Raz and Wang's work [14, 15]. Kanagawa et al. Developed Wang and Raz's work to gain the literature on fuzzy probability and fuzzy membership approaches [16].

Type 1 fuzzy numbers have been missing to identify linguistic expressions, while type 2 fuzzy numbers have begun to be used. First, the type 2 fuzzy numbers generated by Zadeh's work in 1975 fuzzifing membership functions, has provided more realistic and more relevant data [17]. In the accessible literature, there was no control chart made with the type 2 fuzzy sets, which could better express the linguistic terms of the expert opinions and the linguistic terms [18].

In this study, control charts are obtained with type 2 fuzzy numbers in which a fuzzy number is represented by multiple membership degrees instead of a single membership degree. The study includes creating control charts with the calculation of defuzzification and likelihood using interval type 2 trapezoidal fuzzy sets.

The study is designed in the following order. In Sect. 2, arithmetic operations of trapezoidal interval type 2 fuzzy sets arithmetic operations is mentioned. In Sect. 3, deffuzzification and likelihood methods for trapezoidal interval type 2 fuzzy sets are explained. In Sect. 4, defuzzification and likelihood approach are applied to fuzzy control charts with trapezoidal interval type 2 fuzzy numbers. In Sect. 5, the results of the approaches are compared and interpreted with a numerical example. Finally in Sect. 6, concluding remarks are given.

2 Interval Type 2 Fuzzy Sets

In this section, some information about interval type 2 fuzzy sets is given and some arithmetic operations related to trapezoidal interval type 2 fuzzy sets are mentioned.

Zadah (1975) noted that type 2 fuzzy numbers differ from type 1 numbers in that membership functions are fuzzy and shown as follows [17].

$$\tilde{\tilde{A}} = \left\{ (x,u), \mu_{\tilde{\tilde{A}}}(x,u) \middle| \forall x \in X, \forall u \in J_x \subseteq [0,1], 0 \leq \mu_{\tilde{\tilde{A}}}(x,u) \leq 1 \right\}$$

where J_x symbolizes an interval [0,1]. If all $\mu_{\tilde{\tilde{A}}}(x,u) = 1$, $\tilde{\tilde{A}}$ is called an interval type 2 fuzzy set [19].

$\tilde{\tilde{A}}_i = \left(\left(a_{i1}^U, a_{i2}^U, a_{i3}^U, a_{i4}^U; H_1\left(\widetilde{A_i^U}\right), H_2\left(\widetilde{A_i^U}\right)\right), \left(a_{i1}^L, a_{i2}^L, a_{i3}^L, a_{i4}^L; H_1\left(\widetilde{A_i^L}\right), H_2\left(\widetilde{A_i^L}\right)\right)\right)$ is illustration of trapezoidal interval type 2 fuzzy set where a_{ik}^m is reference point of the interval type 2 fuzzy set $\tilde{\tilde{A}}_i$, k = 1,2,3,4, m = U, L (U defines upper membership function and L defines lower membership function) and $1 \leq i \leq n$. $H_j\left(\widetilde{A_i^m}\right) \in [0,1]$: denotes the membership value of the element $a_{i(j+1)}^m$, j = 1, 2, m = U, L and $1 \leq i \leq n$.

Some arithmetic operations related to trapezoidal interval type 2 fuzzy sets are given in following.

The addition operation between the trapezoidal interval type 2 fuzzy sets $\tilde{\tilde{A_1}}$ and $\tilde{\tilde{A_2}}$ is defined as follow:

$$\begin{aligned}\tilde{\tilde{A_1}} + \tilde{\tilde{A_2}} = &\left(\left(a_{11}^U + a_{21}^U, a_{12}^U + a_{22}^U, a_{13}^U + a_{23}^U, a_{14}^U + a_{24}^U; \min\left(H_1\left(\widetilde{A_1^U}\right); H_1\left(\widetilde{A_2^U}\right)\right),\right.\right.\\ &\left.\min\left(H_2\left(\widetilde{A_1^U}\right); H_2\left(\widetilde{A_2^U}\right)\right)\right), \left(a_{11}^L + a_{21}^L, a_{12}^L + a_{22}^L, a_{13}^L + a_{23}^L, a_{14}^L + \right.\\ &\left.\left. a_{24}^L; \min\left(H_1\left(\widetilde{A_1^L}\right); H_1\left(\widetilde{A_2^L}\right)\right), \min\left(H_2\left(\widetilde{A_1^L}\right); H_2\left(\widetilde{A_2^L}\right)\right)\right)\right)\end{aligned} \quad (1)$$

The subtraction operation between the trapezoidal interval type 2 fuzzy sets $\tilde{\tilde{A_1}}$ and $\tilde{\tilde{A_2}}$ is defined as follow:

$$\begin{aligned}\tilde{\tilde{A_1}} - \tilde{\tilde{A_2}} = &\left(\left(a_{11}^U - a_{24}^U, a_{12}^U - a_{23}^U, a_{13}^U - a_{22}^U, a_{14}^U - a_{21}^U; \min\left(H_1\left(\widetilde{A_1^U}\right); H_1\left(\widetilde{A_2^U}\right)\right),\right.\right.\\ &\left.\min\left(H_2\left(\widetilde{A_1^U}\right); H_2\left(\widetilde{A_2^U}\right)\right)\right), \left(a_{11}^L - a_{24}^L, a_{12}^L - a_{23}^L, a_{13}^L - a_{22}^L, a_{14}^L - \right.\\ &\left.\left. a_{21}^L; \min\left(H_1\left(\widetilde{A_1^L}\right); H_1\left(\widetilde{A_2^L}\right)\right), \min\left(H_2\left(\widetilde{A_1^L}\right); H_2\left(\widetilde{A_2^L}\right)\right)\right)\right)\end{aligned} \quad (2)$$

The multiplication operation between the trapezoidal interval type 2 fuzzy sets $\widetilde{\widetilde{A}}_1$ and $\widetilde{\widetilde{A}}_2$ is defined in Eq. (3) and the arithmetic operations between the trapezoidal interval type 2 fuzzy sets $\widetilde{\widetilde{A}}_i$ and the crisp value k is defined in Eq. (4):

$$\widetilde{\widetilde{A}}_1 * \widetilde{\widetilde{A}}_2 = \left(\left(a_{11}^U * a_{21}^U, a_{12}^U * a_{22}^U, a_{13}^U * a_{23}^U, a_{14}^U * a_{24}^U; \min\left(H_1\left(\widetilde{A_1^U}\right); H_1\left(\widetilde{A_2^U}\right)\right), \min\left(H_2\left(\widetilde{A_1^U}\right); H_2\left(\widetilde{A_2^U}\right)\right)\right),\right.$$
$$\left.\left(a_{11}^L * a_{21}^L, a_{12}^L * a_{22}^L, a_{13}^L * a_{23}^L, a_{14}^L * a_{24}^L; \min\left(H_1\left(\widetilde{A_1^L}\right); H_1\left(\widetilde{A_2^L}\right)\right), \min\left(H_2\left(\widetilde{A_1^L}\right); H_2\left(\widetilde{A_2^L}\right)\right)\right)\right) \tag{3}$$

$$k * \widetilde{\widetilde{A}}_i = \left(\left(k * a_{i1}^U, k * a_{i2}^U, k * a_{i3}^U, k * a_{i4}^U; H_1\left(\widetilde{A_i^U}\right), H_2\left(\widetilde{A_i^U}\right)\right),\right.$$
$$\left.\left(k * a_{i1}^L, k * a_{i2}^L, k * a_{i3}^L, k * a_{i4}^L; H_1\left(\widetilde{A_i^L}\right), H_2\left(\widetilde{A_i^L}\right)\right)\right) \tag{4}$$

3 Comparison Methods for Type 2 Fuzzy Sets

It is quite difficult to operate with type 2 fuzzy sets. For this reason, some defuzzification methods for interval type 2 fuzzy sets have been developed. These methods often define type 2 fuzzy set as a type 1 fuzzy set and then use one of the defuzzification methods for type 1 fuzzy sets.

This section discusses the two approaches mentioned in the literature. In the following Chen and Lee's likelihood approach and Kahraman et al.'s defuzzification approach for interval type 2 fuzzy sets are given.

3.1 Chen and Lee's Likelihood Approach

Chen and Lee proposed ranking method for type 2 fuzzy sets. The following equation is likelihood of $\widetilde{A}_s^U \geq \widetilde{A}_t^U$ [20].

$$P\left(\widetilde{A}_s^U \geq \widetilde{A}_t^U\right) =$$
$$max\left(1 - max\left(\frac{\sum_{k=1}^4 max\left(a_{tk}^U - a_{sk}^U, 0\right) + \left(a_{t4}^U - a_{s1}^U\right) + \sum_{j=1}^2 max\left(H_j\left(\widetilde{A}_t^U\right) - H_j\left(\widetilde{A}_s^U\right), 0\right)}{\sum_{k=1}^4 \left|a_{tk}^U - a_{sk}^U\right| + \left(a_{s4}^U - a_{s1}^U\right) + \left(a_{t4}^U - a_{t1}^U\right) + \sum_{j=1}^2 \left|H_j\left(\widetilde{A}_t^U\right) - H_j\left(\widetilde{A}_s^U\right)\right|}, 0\right), 0\right) \tag{5}$$

Then the ranking values for upper and lower membership functions are defined as follows:

$$Rank\left(\breve{A}_i^m\right) = \frac{1}{n(n-1)}\left(\sum_{k=1}^n p(\left(\widetilde{A}_s^m \geq \widetilde{A}_t^m\right) + \frac{n}{2} - 1\right) \tag{6}$$

where m = U, L and n is a number of sets.

Finally the ranking values of the interval type 2 fuzzy set $\tilde{\tilde{A}}_i$ can be calculated by following equation:

$$Rank\left(\tilde{A}_i\right) = \frac{Rank\left(\breve{A}_i^U\right) + Rank\left(\breve{A}_i^L\right)}{2} \tag{7}$$

3.2 Kahraman et al.'s Defuzzification Approach

Kahraman et al. proposed ranking methods for triangular and trapezoidal interval type 2 fuzzy sets. (see Eq. (8)) [21].

$$DTraT = \frac{\left[\frac{\left(a_{i4}^U - a_{i1}^U\right) + \left(H_1^U * a_{i2}^U - a_{i1}^U\right) + \left(H_2^U * a_{i3}^U - a_{i1}^U\right)}{4} + a_{i1}^U\right] + \left[\frac{\left(a_{i4}^L - a_{i1}^L\right) + \left(H_1^L * a_{i2}^L - a_{i1}^L\right) + \left(H_2^L * a_{i3}^L - a_{i1}^L\right)}{4} + a_{i1}^L\right]}{2} \tag{8}$$

DTraT is abbreviation for the trapezoidal type 2 fuzzy sets.

4 Type 2 Fuzzy Control Charts

The advantage of using fuzzy sets is that linguistic values can represent numerical values. Because of this advantage, fuzzy sets can be used in many areas, control charts are one of these areas. In particular, the fuzzy approach is suitable for control graphs for attributes where the data are linguistic, categorical and based on human opinion.

Classical control charts for attributes can be obtained for fraction rejected as nonconforming to specifications, ember of nonconforming items, number of nonconformities and number of nonconformities per unit. For crisp values, control limits for number of nonconformities are calculated by the following equations:

$$\mathrm{CL} = \bar{c} \tag{9}$$

$$\mathrm{LCL} = \bar{c} - 3\sqrt{\bar{c}} \tag{10}$$

$$\mathrm{UCL} = \bar{c} + 3\sqrt{\bar{c}} \tag{11}$$

where $\bar{c}$ is the mean of the nonconformities. In the literature, fuzzy control charts are used for type 1 fuzzy sets. In this study, each sample is expressed as a type 2 trapezoidal fuzzy numbers

$$\left(\left(a_{i1}^U, a_{i2}^U, a_{i3}^U, a_{i4}^U; H_1\left(\widetilde{A_i^U}\right), H_2\left(\widetilde{A_i^U}\right)\right), \left(a_{i1}^L, a_{i2}^L, a_{i3}^L, a_{i4}^L; H_1\left(\widetilde{A_i^L}\right), H_2\left(\widetilde{A_i^L}\right)\right)\right)$$

In type 2 trapezoidal fuzzy case, the center line, $\widetilde{CL}$, given in Eq. (12)

$$\left(\left(\overline{a_1^U},\overline{a_2^U},\overline{a_3^U},\overline{a_4^U};min\left(H_1\left(\widetilde{A_l^U}\right)\right),min\left(H_2\left(\widetilde{A_l^U}\right)\right)\right),\left(\overline{a_1^L},\overline{a_2^L},\overline{a_3^L},\overline{a_4^L};min\left(H_1\left(\widetilde{A_l^L}\right)\right),min\left(H_2\left(\widetilde{A_l^L}\right)\right)\right)\right)$$
$$=\left(\left(\begin{array}{c}\left(\frac{\sum_{i=1}^{m}a_{i1}^U}{m},\frac{\sum_{i=1}^{m}a_{i2}^U}{m},\frac{\sum_{i=1}^{m}a_{i3}^U}{m},\frac{\sum_{i=1}^{m}a_{i4}^U}{m};min\left(H_1\left(\widetilde{A_l^U}\right)\right),min\left(H_2\left(\widetilde{A_l^U}\right)\right)\right),\\ \left(\frac{\sum_{i=1}^{m}a_{i1}^L}{m},\frac{\sum_{i=1}^{m}a_{i2}^L}{m},\frac{\sum_{i=1}^{m}a_{i3}^L}{m},\frac{\sum_{i=1}^{m}a_{i4}^L}{m};min\left(H_1\left(\widetilde{A_l^L}\right)\right),min\left(H_2\left(\widetilde{A_l^L}\right)\right)\right)\end{array}\right)\right) \tag{12}$$

where m is the number of fuzzy samples.

After $\widetilde{CL}$ is calculated, $\widetilde{UCL}$ and $\widetilde{LCL}$ are calculated as in Eqs. (13) and (14), respectively

$$\begin{array}{l}\left(\left(\overline{a_1^U}+3\sqrt{\overline{a_1^U}},\overline{a_2^U}+3\sqrt{\overline{a_2^U}},\overline{a_3^U}+3\sqrt{\overline{a_3^U}},\overline{a_4^U}+3\sqrt{\overline{a_4^U}};\min\left(H_1\left(\widetilde{A_i^U}\right)\right),\min\left(H_2\left(\widetilde{A_i^U}\right)\right)\right),\right.\\ \left.\left(\overline{a_1^L}+3\sqrt{\overline{a_1^L}},\overline{a_2^L}+3\sqrt{\overline{a_2^L}},\overline{a_3^L}+3\sqrt{\overline{a_3^L}},\overline{a_4^L}+3\sqrt{\overline{a_4^L}};\min\left(H_1\left(\widetilde{A_i^L}\right)\right),\min\left(H_2\left(\widetilde{A_i^L}\right)\right)\right)\right)\end{array} \tag{13}$$

$$\left(\begin{array}{c}\left(\overline{a_1^U}-3\sqrt{\overline{a_4^U}},\overline{a_2^U}-3\sqrt{\overline{a_3^U}},\overline{a_3^U}-3\sqrt{\overline{a_2^U}},\overline{a_4^U}-3\sqrt{\overline{a_1^U}};\min\left(H_1\left(\widetilde{A_i^U}\right)\right),\min\left(H_2\left(\widetilde{A_i^U}\right)\right)\right),\\ \left(\overline{a_1^L}-3\sqrt{\overline{a_4^L}},\overline{a_2^L}-3\sqrt{\overline{a_3^L}},\overline{a_3^L}-3\sqrt{\overline{a_2^L}},\overline{a_4^L}+3\sqrt{\overline{a_1^L}};\ \min\left(H_1\left(\widetilde{A_i^L}\right)\right),\min\left(H_2\left(\widetilde{A_i^L}\right)\right)\right)\end{array}\right) \tag{14}$$

The relationship between control limits, which are calculated as a type 2 trapezoidal fuzzy sets, and data is determined by Chen and Lee's approach. The relationship is actually a probability value that allows two type 2 trapezoid fuzzy sets to be compared with each other. In this study, the probability is designed to determine the uncontrolled situation and the rules have been developed accordingly. In other words, the likelihood that the data is greater than the upper control limit and the likelihood that the lower control limit is greater than the data are calculated. Decision rules and flowchart are shown in Fig. 1.

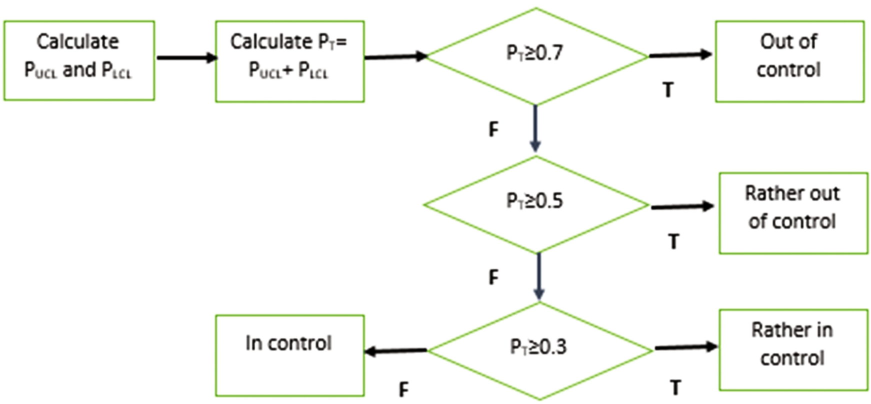


Fig. 1. Flowcharts and decision rules for type 2 fuzzy control charts

Furthermore, control charts are drawn with type 2 trapezoidal fuzzy sets deffuzzified by Kahraman and et al. Control limits, which are calculated as type 2 trapezoidal fuzzy sets, and data are deffuzzified and then drawn as a classical control chart.

5 Numerical Example

Samples of 250 units are taken every 2 h to control number of nonconformities. Data collected from 30 subgroups shown in Table 1 are crisp and in Table 2 linguistic values. These data are transformed into type 2 trapezoidal fuzzy sets and operations are done with type 2 trapezoidal fuzzy sets.

Table 1. Crisp value for numerical example

No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Crisp value	30	25	9	6	38	22	6	40	13	12	6	32	13	51	40
No	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Crisp value	40	41	39	18	28	34	18	30	25	36	18	10	32	23	8

$\widetilde{CL}$, $\widetilde{LCL}$ and $\widetilde{UCL}$ are determined using Eqs. (12) and (14) and the type 2 trapezoidal fuzzy $\widetilde{CL}$, $\widetilde{LCL}$ and $\widetilde{UCL}$ calculated are given below.

$$\widetilde{CL} = ((18.13, 22.67, 26.93, 32.07; 0.63, 0.59), (19.37, 23.67, 26.00, 30.30; 0.48, 0.45))$$
$$\widetilde{LCL} = ((1.14, 7.10, 12.65, 19.29; 0.63, 0.59), (32.57, 38.26, 41.30, 46.81; 0.48, 0.45))$$
$$\widetilde{UCL} = ((30.91, 36, 95, 42.50, 49.05; 0.63, 0.59), (32.57, 38.26, 41, 30, 46.81; 0.48, 0, 45))$$

Tables 3 and 4 are generated with the use of Eq. 3. Table 3 shows the likelihood that the data is greater than the upper control limit and Table 4 shows the likelihood that the lower control limit is greater than the data. In both tables, column, which is expressed as the average likelihood (P), shows the calculation of Eq. (6) where n = 2.

Chen and Lee proposed ranking methods for fuzzy multiple attributes group-decision making. In fuzzy multiple attributes group-decision making case, each attribute is compared to all other attributes, so the ranking equation is formed as Eq. (6). However, in this study, the fuzzy data is compared to the UCL or the LCL, so n = 2.

Figure 2a and b shows graphical results of Tables 3 and 4, respectively. Using Tables 3 and 4, it is decided about the process by the rules in Fig. 2a. according to the results, it is said that the data 4, 7, 11 and 14 are out of control; data 3, 17 and 30 are rather out of control; data 1, 2, 6, 9, 12, 13, 19–26, 28, 29 are in control; the others are rather in control.

Another method used in the study is the deffuzzification method for interval type 2 fuzzy sets. Trapezoidal type 2 fuzzy data and control limits are deffuzzified using Eq. (8). Defuzzified data are given Table 5.

Table 2. Linguistic value for numerical example

No	Between	Approximately	No	Between	Approximately
1		30	16		40
2	20–30		17	32–50	
3	5–12		18		39
4		6	19	15–21	
5		38	20		28
6	20–24		21	32–35	
7	4–8		22	10–25	
8	36–44		23		30
9	11–15		24		25
10	10–13		25	31–41	
11		6	26	10–25	
12		32	27	5–14	
13		13	28	28–35	
14	50–52		29	20–25	
15	38–41		30		8

Table 3. Chen and Lee's likelihood value for data and UCL

No	$P\left(\widetilde{A}_i^U \geq \widetilde{A}_{UCL}^U\right)$	$P\left(\widetilde{A}_i^L \geq \widetilde{A}_{UCL}^L\right)$	Av. P	No	$P\left(\widetilde{A}_i^U \geq \widetilde{A}_{UCL}^U\right)$	$P\left(\widetilde{A}_i^L \geq \widetilde{A}_{UCL}^L\right)$	Av. P
1	0,063	0,025	0,044	16	0,444	0,534	0,489
2	0,045	0,007	0,026	17	0,624	0,639	0,632
3	0	0	0	18	0,397	0,339	0,368
4	0	0	0	19	0,063	0,045	0,054
5	0,395	0,381	0,388	20	0,075	0,026	0,051
6	0	0	0	21	0,211	0,146	0,178
7	0	0	0	22	0,026	0	0,013
8	0,448	0,407	0,427	23	0,055	0	0,027
9	0	0	0	24	0,004	0	0,002
10	0	0	0	25	0,274	0,238	0,256
11	0	0	0	26	0	0	0
12	0,109	0,057	0,083	27	0	0	0
13	0	0	0	28	0,109	0,058	0,084
14	0,859	0,912	0,886	29	0	0	0
15	0,346	0,278	0,312	30	0	0	0

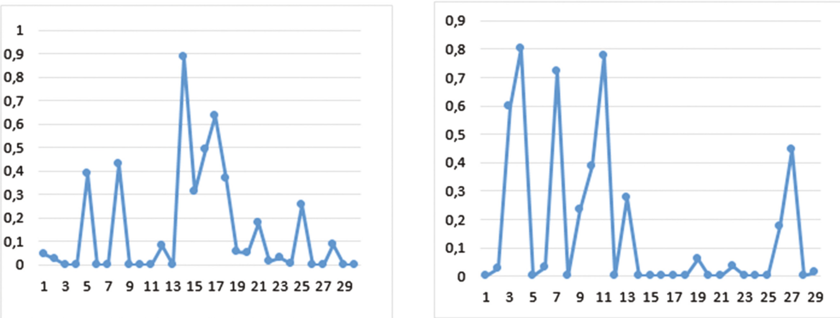


Fig. 2. (a) Average likelihood of comparing; (b) Average likelihood of comparing data and upper control limit data and lower control limit

Table 4. Chen and Lee's likelihood value for data and LCL

No	$P\left(\widetilde{A}^U_{LCL} \geq \widetilde{A}^U_i\right)$	$P\left(\widetilde{A}^L_{LCL} \geq \widetilde{A}^L_i\right)$	Av. P	No	$P\left(\widetilde{A}^U_{LCL} \geq \widetilde{A}^U_i\right)$	$P\left(\widetilde{A}^L_{LCL} \geq \widetilde{A}^L_i\right)$	Av. P
1	0	0	0	16	0	0	0
2	0,044	0,012	0,028	17	0	0	0
3	0,561	0,631	0,596	18	0	0	0
4	0,788	0,814	0,801	19	0,081	0,036	0,059
5	0	0	0	20	0	0	0
6	0,042	0,016	0,029	21	0	0	0
7	0,707	0,733	0,720	22	0,058	0,014	0,036
8	0	0	0	23	0	0	0
9	0,233	0,233	0,233	24	0	0	0
10	0,413	0,360	0,386	25	0	0	0
11	0,750	0,802	0,776	26	0,177	0,169	0,173
12	0	0	0	27	0,443	0,445	0,444
13	0,294	0,261	0,277	28	0	0	0
14	0	0	0	29	0,028	0,000	0,014
15	0	0	0	30	0,546	0,516	0,531

Table 5. Deffuzzification of data and control limits

No	xi	No	xi	No	Xi
1	23,82	11	5,72	21	29,26
2	20,86	12	27,75	22	20,15
3	7,94	13	11,15	23	26,23
4	5,49	14	46,02	24	22,03
5	35,10	15	32,32	25	29,89
6	16,88	16	36,46	26	14,63
7	5,67	17	37,82	27	8,95
8	33,55	18	33,21	28	26,30
9	11,20	19	19,09	29	19,62
10	9,87	20	25,70	30	8,74
CL	19,14	LCL	7,68	UCL	30,58

The control chart is drawn by using the classic control graph method, after defuzzification, it is shown in Fig. 3 for numerical example. Control limits are also obtained by the defuzzification method.

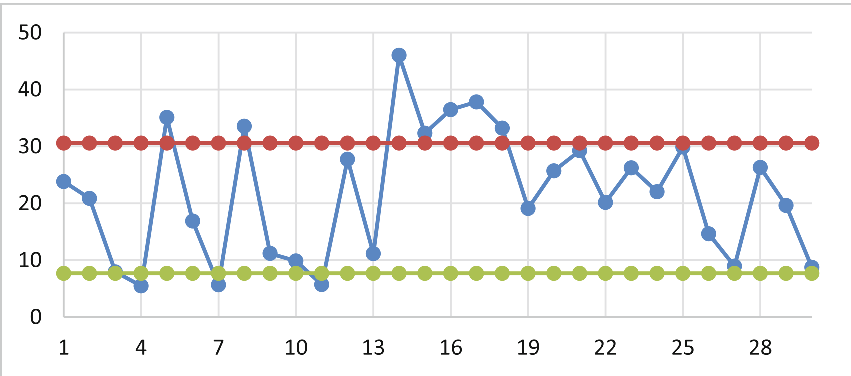


Fig. 3. Control charts with deffuzzified data

According to deffuzzification method, data 4, 5, 7, 8, 11, 14–18 are out of control and the others are in control.

Eventually, the control charts obtained with the crisp data were compared with the other results. The Minitab program is used for the classic control chart and the results are shown in Fig. 4.

It is shown that in Fig. 4, data 3, 4, 7, 8, 11, 14–17 and 30 are out of control and the others are in control.

Finally, the comparison of the results of each method is summarized in Table 6. Abbreviation for IC, RIN, ROC and OC express in control, rather in control, rather out of control and out of control, respectively. Results of data 8, 11 and 15, which are

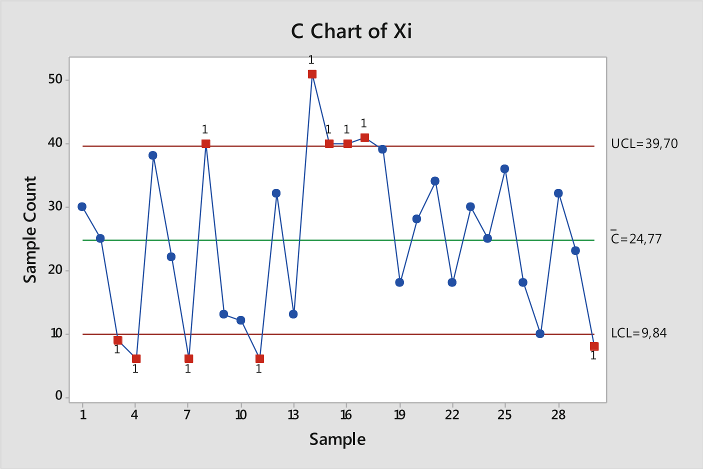


Fig. 4. Classic c control charts for crisp data

Table 6. Comparison of three approach

Sample No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Likelihood approach	IC	IC	ROC	OC	RIC	IC	OC	RIC	IC	RIC	OC	IC	IC	OC	RIC
Deffuzification method	IC	IC	IC	OC	OC	IC	OC	OC	IC	IC	OC	IC	IC	OC	OC
Clasic control charts approach	IC	IC	OC	OC	IC	IC	OC	OC	IC	IC	OC	IC	IC	OC	OC
Sample No	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Likelihood approach	RIC	ROC	RIC	IC	IC	IC	IC	IC	IC	IC	IC	RIC	IC	IC	ROC
Deffuzification method	OC	OC	OC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC
Clasic control charts approach	OC	OC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	IC	OC

obtained with likelihood approach, are different from same data obtained with classical control charts approach. Also for some data, the process control chart is stretched with "rather in control" and "rather out of control" expressions On the other hand, results of data 3, 5, 18 and 30, which are calculated with deffuzzification approach, are different from same data calculated with classical control charts approach.

6 Conclusions

The starting point of this study is the lack of any control charts made with the interval type 2 fuzzy sets in the accessible literature. For this reason, type 2 fuzzy control charts have been obtained by using likelihood and defuzzification methods for interval trapezoidal type 2 fuzzy sets and are clarified with a numerical example. Deffuzzification method evaluates the process as “in control” and “out of control”, just like the classical method. On the other hand, the likelihood approach evaluates the process as more flexible and determines new expressions such as “rather in control” and “rather out of control” for control charts.

In future studies, different trapezoidal interval type 2 fuzzy sets can be applied with these methods and also different trapezoidal interval type 2 fuzzy set approaches can be compared with the control charts methods applied in this study.

References

1. Shewhart, W.A.: Economic Control of Quality of Manufactured Product. D. Van Nosrrand Inc., Princeton (1931)
2. Gülbay, M., Kahraman, C.: An alternative approach to fuzzy control charts: direct fuzzy approach. Inf. Sci. **177**(6), 1463–1480 (2007)
3. Faraz, A., Shapiro, A.F.: An application of fuzzy random variables to control charts. Fuzzy Sets Syst. **161**(20), 2684–2694 (2010)
4. Şentürk, S., Erginel, N.: Development of fuzzy X–R and X–S control charts using α-cuts. Inf. Sci. **179**, 1542–1551 (2009)
5. Alaeddini, A., Ghanzanfari, M., Nayeri, M.A.: A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts. Inf. Sci. **179** (11), 1769–1784 (2009)
6. Cheng, C.B.: Fuzzy process control: construction of control charts with fuzzy numbers. Fuzzy Sets Syst. **154**(2), 287–303 (2005)
7. Gülbay, M., Kahraman, C.: Development of fuzzy process control charts and fuzzy unnatural pattern analyses. Comput. Stat. Data Anal. **51**, 434–451 (2006)
8. Gülbay, M., Kahraman, C., Ruan, D.: α-cut fuzzy control charts for linguistic data. Int. J. Intell. Syst. **19**, 1173–1196 (2004)
9. Shu, M.H., Wu, H.C.: Fuzzy X and R control charts: fuzzy dominance approach. Comput. Ind. Eng. **613**, 676–686 (2011)
10. Kaya, I., Kahraman, C.: Process capability analyses based on fuzzy measurements and fuzzy control charts. Expert Syst. Appl. **38**, 3172–3184 (2011)
11. Asai, K.: Fuzzy Systems for Management. IOS Press, Amsterdam (1995)
12. Woodall, W., Tsui, K.L., Tucker, G.L.: A Review of Statistical and Fuzzy Control Charts Based on Categorical Data, Frontiers in Statistical Quality Control, vol. 5. Physica-Verlag, Heidelberg (1997)
13. Laviolette, M., Seaman, J.W., Barrett, J.D., Woodall, W.H.: A probabilistic and statistical view of fuzzy methods, with discussion. Technometrics **37**, 249–292 (1995)
14. Wang, J.H., Raz, T.: On the construction of control charts using linguistic variables. Int. J. Prod. Res. **28**, 477–487 (1990)
15. Raz, T., Wang, J.H.: Probabilistic and membership approaches in the construction of control charts for linguistic data. Prod. Plan. Control **1**, 147–157 (1990)
16. Kanagawa, A., Tamaki, F., Ohta, H.: Control charts for process average and variability based on linguistic data. Int. J. Prod. Res. **2**, 913–922 (1993)
17. Zadeh, L.A.: The concept of linguistic variable and its application to approximate reasoning-I. Inf. Sci. **8**(3), 338–353 (1975)
18. Mendel, J.M., John, R.I.B.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**(2), 117–127 (2002)

19. Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **29**, 133–143 (1985)
20. Chen, S.-M., Lee, L.-W.: Fuzzy multiple attributes group-decision making based on the ranking values and the arithmetic operations of interval type-2 fuzzy sets. Expert Syst. **37**, 824–833 (2010)
21. Kahraman, C., Öztayşi, B., Sarı, I.U., Turanoğlu, E.: Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. Knowl. Based Syst. **59**, 48–57 (2014)

Linked Open Data: Uncertainty in Equivalence of Properties

Nhuan D. To[1], Marek Z. Reformat[1(✉)], and Ronald R. Yager[2]

[1] University of Alberta, Edmonton T6G 1H9, Canada
marek.reformat@ualberta.ca
[2] Iona College, New Rochelle, NY 10801, USA

Abstract. Linked Open Data (LOD) is a graph-based repository of data that uses data representation format called Resource Description Framework (RDF). The basic piece of RDF data is a triple subject-property-object. LOD seen as a network of interconnected pieces of data creates an environment suitable for developing methods enabling learning processes that rely on data integration. Application of frequentionistic-based approaches to integrate data leads to identification of pieces of information that are consistent and frequently used. An essential element of such methods is the ability to identify similar pieces of data. In reality, multiple sources of information use different vocabularies to represent relations (properties) existing between data. That introduces a challenge for data integration methods.

In this paper, we propose a simple approach to determine degrees of equivalences between relations (properties) defined by different LOD vocabularies. We process numbers of occurrences of matching pairs of RDF triples in order to determine intervals representing lower and upper levels of property equivalences. As the result, we obtain a graph of equivalent properties where interval-based strength of edges represent degrees of similarity between properties. A case study illustrating the details of the approach and a validation experiment are included.

Keywords: RDF data · Property equivalence · Possibility theory

1 Introduction

One of the important contributions of the Semantic Web [1] is a graph-based form of data representation called Resource Description Framework (RDF) [16]. RDF data model treats each piece of information as an RDF triple: subject-property-object [9,16]. The application of RDF for data representation has become a very popular way of representing data on the web [13]. Over time, the term Linked Open Data (LOD) has been used to describe the network of data sources that use RDF triples for information representation [2].

LOD is beneficial for various semantic applications such as web search engines, web browsers, information retrieval systems, and reasoning engines.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_38

LOD, and graph-based data formats, are suitable environments for development of learning focused methods and approaches. An example could be a learning process that involves continuous collecting and aggregating pieces of data. Further, the accumulated data would be processed in order to identify importance of individual pieces of data, and to determine strength of relations between them. Assuming a simple frequentionistic approach, we can say that elements and connections between data that occur more often become more pronounced and 'stronger'. For the case of LOD, an integration of data from multiple RDF repositories poses a number of challenges regarding data processing and analysis [7]. One of such challenges associated with diversity of LOD is multitude of vocabularies used in the datasets. In order to manage interconnectivity between pieces of data and diversity of relations between them, we postulate that construction of a graph that includes degrees of equivalence between data relations is an essential step in development of learning focused systems. We use the term 'equivalence' to comply with the terminology of LOD that defines two properties as equivalent it their semantics is the same. For LOD, an integration of data from RDF repositories poses a number of challenges regarding processing and analysis of data [7]. One of such challenges associated with the diversification of LOD is a multitude of vocabularies used in the datasets.

In this paper, we propose a method that evaluates degrees of equivalence between properties defined by different LOD vocabularies. The method uses a simple approach to determine a number of occurrences of pairs of RDF triples, subject-property-object, with the same subjects and objects. The obtained occurrence numbers are processed and elements of possibility theory [5,15] are used to express lower and upper limits of degrees of equivalences. The application of possibility theory allows us to handle a number of difficulties when dealing with real data sources: (1) inability to compare all pairs of properties; (2) existence of inconsistency among pieces of data as the consequence of incoherent semantics or simple spelling errors; and (3) unregulated nature of posted information resulting in unpredicted diversity of data semantics. The lower and upper limits of equivalence are determined based on the available data. The limits determine a range that potentially contains the most realistic degree of equivalence. The method is used to evaluate equivalencies of properties defined by three well-known RDF datastore – *DBpedia* [17], *Wikidata* [18], and *YAGO* [19] – which define and use their own LOD vocabularies.

2 Background

2.1 Related Work

Property alignment, which is crucial for data integration tasks in LOD has received a significant attention of many researchers. Different techniques have been used to address the task. In the schema-based (profile-based) method [3], information about a label, domain and range is utilized. In this approach, a similarity of property is acquired via string matching using different similarity measures such as Jaccard similarity coefficient, cosine similarity or WordNet based methods processing words that occur in the property names.

In the instance-based approach (also called schema-independent) [6,11] the content of instances that belong to classes and the properties that appear in these instances are consulted. Some researchers used a mixture of these techniques. For example, [14] uses four different basic measures, including a string edit distance, a WordNet based method, a profile, and an instance based technique to determine a level of similarity between properties.

The closest to the method proposed here is the approach presented in [6,10]. The authors utilize the concept of equivalent properties that is defined by OWL. However, there are some differences between their method and the method described here. In [6], the authors analyze the results of statistical processing of number of matching subject-object tuples for a pair of investigated properties $<p_1, p_2>$. A degree of equivalence is calculated as the ratio of a number of pairs of RDF triples with the same subject and object that have $<p_1, p_2>$ as their properties to a number of RDF triples that have p_1 and p_2 as their properties but they only have matching subjects. We start with instances of a given class c from *DBpedia*, for example *Person* and find instances – via the property *owl:sameAs* – from *Wikidata* and *YAGO*. For each pair of *owl:sameAs* instances, we find all RDF triples with the same object from both datastores. This decreases the search space and increases effectiveness of the approach.

2.2 Possibility Theory

Introduced by [15] and fully developed by [5] possibility theory is a suitable vehicle to handle incomplete information. Even if similar to the probability theory it differs in using two sets of functions: possibility and necessity measures, instead of just one measure as in the probability theory. Here, we present the basic definitions and concepts that are used in our approach for similarity evaluation. More information on possibility theory can be found in [4,8].

Let us assume a finite set of states, S. A possibility distribution function is:

$$\pi(s) : S \rightarrow < 0, 1 > \tag{1}$$

that s represents a current state of knowledge. Possibility theory appraises what elements of S are plausible and what elements are not, what is 'normal' and what is not. The state s is expressed to be impossible as:

$$\pi(s) = 0 \tag{2}$$

or totally possible (plausible):

$$\pi(s) = 1 \tag{3}$$

This allows for expressing complete knowledge, when some state s_0 is possible $\pi(s_0) = 1$, and other states s are impossible $\pi(s_0) = 0$. Total ignorance is expressed as $\pi(s_0) = 1$ for all s from S. Therefore, degrees of possibility and necessity for a given subset of states S_{sub} can be computed as below: possibility:

$$\Pi(S_{sub}) = sup_{s \in S_{sub}} \pi(s) \tag{4}$$

necessity:

$$N(S_{sub}) = inf_{s \notin S_{sub}} 1 - \pi(s) \tag{5}$$

The duality of possibility-necessity is:

$$N(S_{sub}) = 1 - \Pi(S'_{sub}) \tag{6}$$

where S' represents the complement of S. Possibility measures satisfy the basic property of:

$$\Pi(S^a_{sub} \cup S^a_{sub}) = max(\Pi(S^a_{sub}), \Pi(S^b_{sub})) \tag{7}$$

while necessity measures satisfies the dual property:

$$N(S^a_{sub} \cap S^a_{sub}) = min(N(S^a_{sub}), N(S^b_{sub})) \tag{8}$$

3 Property Equivalence Evaluation

3.1 Concept

The proposed process of identifying a degree of equivalence between properties from different vocabularies and data repositories is based on the idea of using triples representing well-know entities. Hereafter, we use the term objects instead of entities in compliance with the RDF terminology. In a nutshell, we compare definitions of the well-know subjects, denoted *reference subjects*, with the definitions of the same subjects from different data repositories. Based on the result of that comparison, we evaluate a degree of equivalence of properties that exist in the subjects' definitions. The principle idea seems simple, however the core of our approach lays in a way how we process the results for a given property. We are interested in the lower and upper boundaries of degree of equivalence, i.e., in necessity and possibility values of equivalence.

3.2 Phase I: Direct Evaluation

In the first phase of the approach, we compare *reference* subjects with subjects from other RDF data-stores: A, B, C, and so on. Figure 1 is an illustration of the process we follow. Let us consider two data-stores: *reference* and A. The triples taken from them are compared: if both subject and object of a triple from *reference* match the subject and object of a triple from A, we assume that properties of these triples are potentially equivalent.

Let S_{ref} is a set of K subjects from the *reference* data-store:

$$S_{ref} = \{s^1_{ref}, \ldots, s^k_{ref}, \ldots s^K_{ref}, \}. \tag{9}$$

Based on this set, we create a set of unique properties existing in S_{ref}:

$$P_{ref} = \{p^1_{ref}, \ldots, p^r_{ref}, \ldots, p^R_{ref}\}. \tag{10}$$

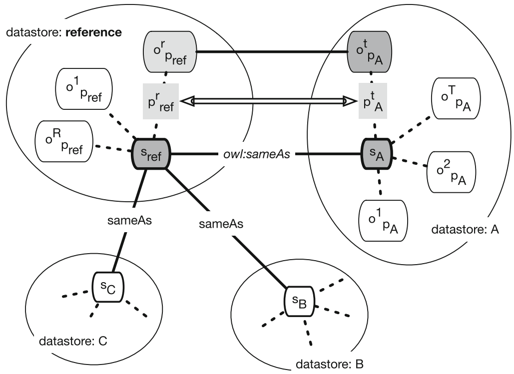


Fig. 1. Comparison of triples from different data-stores

For a given property p_{ref} (for simplicity of notation, let us drop the superscript), we select all subject from S_{ref} that contain this property. Such a set of subjects with p_{ref} is called $S_{p_{ref}}$, and its cardinality is $|S_{p_{ref}}| = N$.

For every subject $s_{p_{ref}}$ from $S_{p_{ref}}$, we find an equivalent subject s_A from the data-store A using the property *owl:sameAs*. Also, we identify all triples that have s_A as their subjects. Among them, we look for triples

$$s_A - p_A^t - o_{A,p_A}^t \tag{11}$$

such that

$$o_{p_{ref}}^r = o_{A,p_A}^t \tag{12}$$

where $o_{p_{ref}}^r$ is the object in the triple

$$s_{ref} - p_{ref}^r - o_{p_{ref}}^r \tag{13}$$

from S_{ref}, Fig. 1. That means that if the triples represented by Eqs. 11 and 13 have the same subjects and objects, then we can assume that p_{ref}^r and p_A^t are equivalent. At the current development stage, the similarity of objects $o_{p_{ref}}^r$ and o_{A,p_A}^t is determined during a simple string matching process. We envision that in the future, we will apply more advanced methods for determining similarity between objects. We keep track how many times we encounter matching pairs of triples. We store this number of occurrences as $M(p_{ref}^r, p_A^t)$.

Once such a process is finished for all subjects $s_{p_{ref}}$ from $S_{p_{ref}}$, we obtain multiple pairs $<p_{ref}, p_A>$ together with their number of occurrences $M(p_{ref}, p_A)$.

Without loss of generality, let us assume that we deal with Q such pairs for the property p_{ref}:

$$\begin{aligned} &M(p_{ref}, p_A^1) \\ &\dots \\ &M(p_{ref}, p_A^q) \\ &\dots \\ &M(p_{ref}, p_A^Q) \end{aligned} \tag{14}$$

We can say that M's represent numbers of occurrences of triples supporting a statement that: p_{ref} is equivalent to p_A^1, p_{ref} is equivalent to p_A^q …, and p_{ref} is equivalent to p_A^Q. These numbers are used to determine a degree of equivalence of p_{ref} and p_A^q.

To determine the necessity value, we take a number of times when subjects/objects of triples from *reference* and A match, and divide it by a number of triples in *reference* with the property p_{ref}. The formula is:

$$N_{EQ(p_{ref},p_A^q)} = \frac{M(p_{ref},p_A^q)}{Total} \tag{15}$$

where $Total$ is a number of triples with p_{ref}. The possibility is determined when a maximum possible number of matching pairs $<p_{ref},p_A^q>$ is taken into consideration. The formula is:

$$\Pi_{EQ(p_{ref},p_A^q)} = \frac{Total - SUM}{Total} \tag{16}$$

where

$$SUM = \sum_{i=1,\ldots,q-1,q+1,\ldots,K} M(p_{ref},p_A^i) \tag{17}$$

represents a number of the found matching pairs different than $<p_{ref},p_A^q>$. Using a simple arithmetics, we can demonstrate that $\Pi_{EQ(p_{ref},p_A^q)}$ is always larger or equal than $N_{EQ(p_{ref},p_A^q)}$.

3.3 Phase II: Indirect Evaluation

In the previous phase, we determine the values of necessity and possibility based on the comparison of triples from other data stores, for example A and B, to the triples from the *reference* data-store, i.e., we obtain equivalence values for 'pairs' *reference-A* and *reference-B*. In the phase II, we determine values of equivalence between triples from non-reference data-stores, i.e., for pairs A-B.

As before, i.e., without loss of generality, we assume we have the following numbers of occurrences for a number of pairs $<p_{ref},p_A^q>$ and $<p_{ref},p_B^j>$:

$$\begin{array}{ll} M(p_{ref},p_A^1) & M(p_{ref},p_B^1) \\ \ldots & \\ M(p_{ref},p_A^q) & M(p_{ref},p_B^r) \\ \ldots & \\ M(p_{ref},p_A^Q) & M(p_{ref},p_B^R) \end{array} \tag{18}$$

Let us pick up two pairs, again for simplicity we omit superscripts, and determine a degree of equivalence for the properties p_A and p_B. We have two values of occurrences for each pair:

$$M(p_{ref},p_A) \qquad M(p_{ref},p_B) \tag{19}$$

The calculations of necessity and possibility are done in reference to the property – p_A or p_B – that provides a larger value of $M(p_{ref}, \ldots)$. We use two simple equations:

$$larger = argmax(M(p_{ref}, p_A), M(p_{ref}, p_B)) \tag{20}$$

to identify an index of property with a larger value of M, and

$$smaller = argmin(M(p_{ref}, p_A), M(p_{ref}, p_B)) \tag{21}$$

to find an index of the property with a smaller value of M.

The calculations follow the pessimistic and optimistic views regarding numbers of matching pairs of triples. In the worst case scenario, we assume that $M(p_{ref}, p_{smaller})$ is at its maximum – there are no more matching pairs for such a pair of properties, and that we found a minimum number of matching pairs for $<p_{ref}, p_{larger}>$. As the result, we denote the value of necessity as:

$$N_{EQ(p_A, p_B)} = \frac{M(p_{ref}, p_{smaller})}{Total - SUM_{larger}} \tag{22}$$

with

$$SUM_{larger} = \sum_{i=1,\ldots,q-1,q+1,\ldots,S} M(p_{ref}, p^i_{larger}). \tag{23}$$

As before $Total$ is a number of triples with the property p_{ref}, while SUM_{larger} represents a total number of matching pairs for all other properties excluding p_{larger}. The possibility, on the other hand, is seen as an optimistic scenario. We assume that all triples from the data-store A that do not have matching triples from the *reference* data-store can potentially be matched to triples with the property $p_{smaller}$.

The formula is:

$$\Pi_{EQ(p_A, p_B)} = \frac{Total - SUM_{smaller}}{Total - SUM_{larger}} \tag{24}$$

with

$$SUM_{smaller} = \sum_{i=1,\ldots,r-1,r+1,\ldots,J} M(p_{ref}, p^i_{smaller}) \tag{25}$$

that represents a total number of matching pairs for all other properties excluding $p_{smaller}$.

In general, the values of SUM_{larger} and $SUM_{smaller}$ could be such that $\Pi_{EQ(p_A, p_B)}$ becomes more than 1. Thus, we modify the original equation to

$$\Pi_{EQ(p_A, p_B)} = min\{1.00, \frac{Total - SUM_{smaller}}{Total - SUM_{larger}}\}. \tag{26}$$

Using a simple arithmetics we can demonstrate that $\Pi_{EQ(p^q_A, p^r_B)}$ is always larger than $N_{EQ(p^q_A, p^r_B)}$ as long as:

$$\frac{M(p_{ref}, p_{smaller})}{M(p_{ref}, p_{larger})} < \frac{M(p_{ref}, p_{smaller}) + SUM_{larger}}{M(p_{ref}, p_{larger}) + SUM_{smaller}} \tag{27}$$

4 Case Study and Experiment

4.1 Wikidata as Reference Source

The equivalence evaluation process uses a set of reference subjects. As it has been indicated earlier, RDF triples with these subjects provide a set of properties that are treated as the base against which we compare other properties.

In our case study, we use three data-stores: *Wikidata*, *DBpedia* and *YAGO*. In order to illustrate the proposed approach, we evaluate equivalences of two *wikidata* properties: $P569$ ($dateOfBirth$) and $P106$ (*occupation*). The obtained values of $M(p_{wiki}, p^q_{dp})$ and $M(p_{wiki}, p^r_{yago})$ are shown in Table 1.

Table 1. Equivalent properties – frequency of occurrence of RDF triples with a given property from $Wikidata$ (left column), and frequency of occurrence of properties from *DBpedia* (centre column) and *YAGO* (rigth column) that match the *Wikidata* property.

wikidata.org	dbpedia.org	yago-knowledge.org
for .../P569 (dateOfBirth) N = **23629**	M($P569, p_{onto/birthDate}$) = **7974** M($P569, p_{prop/birthDate}$) = **7037** M($P569, others$) = **11**	M($P569, p_{wasBornOnDate}$) = **10091** M($P569, p_{en/birthdate}$) = **2641** M($P569, others$) = **4**
for .../P106 (occupation) N = **34576**	M($P106, p_{onto/occupation}$) = **1768** M($P106, p_{prop/occupation}$) = **983** M($P106, others$) = **248**	M($P106, p_{en/occupation}$) = **2772** M($P106, p_{type}$) = **1325** M($P106, others$) = **246**

4.2 Property Matching: Wikidata-DBpedia

Let us start with evaluating equivalence of the property $P569$ (*dateOfBirth*) from *Wikidata* in respect to properties from *DBpedia*. As we can see in Table 1, there are two properties from *DBpedia*: $p_{onto/birthDate}$ and $p_{prop/birthDate}$ that have large values of $M(P569, p_{...})$'s. In the table, there is also the value of $M(P569, others)$ representing a cumulative number of matches to other properties. The similar situation is for the property $P106$ (*occupation*). The calculated values of necessity and possibility for both *Wikidata* properties are presented in Table 2.

Table 2. Necessity and possibility values for *Wikidata* and *DBpedia* properties.

	$p_{onto/birthDate}$	$p_{prop/birthDate}$
$P569$	$N_{EQ} = 0.3375$	$N_{EQ} = 0.2978$
	$\Pi_{EQ} = 0.7017$	$\Pi_{EQ} = 0.6621$
	$p_{onto/occupation}$	$p_{prop/occupation}$
$P106$	$N_{EQ} = 0.0511$	$N_{EQ} = 0.0284$
	$\Pi_{EQ} = 0.9644$	$\Pi_{EQ} = 0.9417$

As it can be seen, the necessity values for the property $P569$ are higher than for $P106$. A quick look at the M values in Table 1 explains this. There are more occurrences of $P106$, and smaller number of triples with matching subjects and objects – $M(P106, p_{...})$'s. The values of possibilities are higher for $P106$ – this means that a better process of identifying matching pairs of subject-object would potentially lead to less uncertainty.

4.3 Property Matching: Wikidata-YAGO

In a very similar way, the calculations are done for determining equivalence between $P569$ (*birthDate*) and $P106$ (*occupation*) and *YAGO*'s properties. The necessity and possibility values of equivalence are in Table 3. As before, the necessity values for the property $P569$ are higher than for $P106$, while the values of possibilities are higher for $P106$.

Table 3. Necessity and possibility values for *Wikidata* and *YAGO* properties.

	$p_{wasBornOnDate}$	$p_{en/birthdate}$
$P569$	$N_{EQ} = 0.4271$	$N_{EQ} = 0.1118$
	$\Pi_{EQ} = 0.8881$	$\Pi_{EQ} = 0.5728$
	$p_{en/occupation}$	p_{type}
$P106$	$N_{EQ} = 0.0802$	$N_{EQ} = 0.0383$
	$\Pi_{EQ} = 0.9546$	$\Pi_{EQ} = 0.9127$

4.4 Property Matching: Dbpedia-YAGO

Finally, we evaluate levels of equivalence between *DBpedia* properties and *YAGO* properties. Based on the values from Table 1 and the approach presented in Sect. 3.3, we obtain the estimated equivalence values, Table 4. The table contains values for all possible combinations between two properties from *DBpedia* and two from *YAGO*. The levels of uncertainty are the lowest for the pairs $<p_{wasBornOnDate} : p_{onto/birthDate}>$ and $<p_{wasBornOnDate} : p_{prop/birthDate}>$. As in the previous cases, the values for the properties related to *occupation* indicate high uncertainty, i.e., low confidence in the obtained results.

As we can see in all presented cases, the values of necessity and possibility are very much dependent on available data and quality of a comparison process. However, despite a simple method used to determine numbers of matching RDF triples, the proposed approach of processing these numbers is capable of calculating ranges of degree of equivalence between properties.

Table 4. Necessity and possibility values for *DBpedia* and *YAGO* properties.

	$p_{wasBornOnDate}$	$p_{en/birthdate}$
$p_{onto/birthDate}$	$N_{EQ} = 0.3800$	$N_{EQ} = 0.1593$
	$\Pi_{EQ} = 0.7902$	$\Pi_{EQ} = 0.8162$
	$p_{wasBornOnDate}$	$p_{en/birthdate}$
$p_{prop/birthDate}$	$N_{EQ} = 0.3354$	$N_{EQ} = 0.1688$
	$\Pi_{EQ} = 0.7455$	$\Pi_{EQ} = 0.8651$
	$p_{en/occupation}$	p_{type}
$p_{onto/occupation}$	$N_{EQ} = 0.0536$	$N_{EQ} = 0.0397$
	$\Pi_{EQ} = 1.0000$	$\Pi_{EQ} = 0.9464$
	$p_{en/occupation}$	p_{type}
$p_{prop/occupation}$	$N_{EQ} = 0.0298$	$N_{EQ} = 0.0311$
	$\Pi_{EQ} = 0.9464$	$\Pi_{EQ} = 1.0000$

4.5 Visualization of Results

The calculations presented above can be visualized as a graph where nodes are properties and edges between them represent equivalence relations. Each edge is labeled with an interval $[N, \Pi]$ representing a range of degrees of equivalence. The graph is shown in Fig. 2. It includes relations calculated in the phase I of the approach (solid lines), and the equivalence values obtained indirectly – phase II – represented as dashed lines.

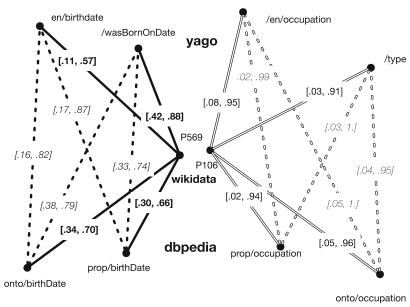


Fig. 2. Property equivalence graph

4.6 Validation Experiment

A set of validation tests has been performed to determine capabilities of the proposed method. For this purpose, we have selected a set of subjects from a single class *Person* from *DBpedia*. We have run SPARQL queries at *DBpedia*, *Wikidata* and *YAGO* endpoints to extract RDF triples representing the same subject using *owl:sameAs* property. The obtained set contains 50,000 RDF triples of the category *Person* from *DBpedia*, and additional 100,000 equivalent entities from the datastores *Wikidata* and *YAGO*.

For the evaluation purposes, we use values of necessity calculated only for the matching pairs of properties. We introduce a threshold value α. If the necessity value calculated for a pair of properties is larger than α, we conclude the equivalence pair is found. We have created a 'reference set of equivalent properties'. This set contains pairs of properties from *DBpedia* and *Wikidata* that have been identified by experts as equivalent; the properties of each pair are connected by a special relation *owl:equivalentProperty*. The pairs we found are compared with the pairs in the reference set to determine such data mining performance measures as precision, recall, and F1 [12]. The experiments are done with values of α changing from 0.00 to 1.00, Fig. 3.

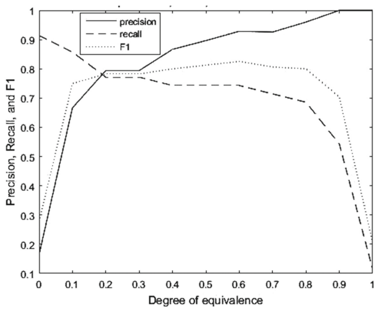


Fig. 3. Performance evaluation: equivalence of *DBpedia* and *Wikidata* properties

5 Conclusion

RDF data repositories that constitute Linked Open Data (LOD) create an opportunity to develop methods that enable data integration and their full utilization. Further, those methods can be seen as the base for developing algorithms suitable for learning and knowledge extraction.

One of important components enabling data integration is related to determining equivalence of different properties used by RDF stores. The approach presented here evaluates the upper and lower limits of property equivalences. The approach utilizes elements of possibility theory to estimate these limits. It is able to cope with imprecision and uncertainty caused by inability to compare all possible pairs of properties in a comprehensive and thorough way. The ultimate goal is to construct a graph of properties (relations) where edges represent degrees of equivalence, Fig. 2. Analysis of such a graph would lead to discovering relations of different types between properties.

References

1. Berners-Lee, T., Hendler, J.: Scientific publishing on the semantic web. Nature **410**, 1023–1024 (2001)
2. Bizer, C., Heath, T., Berners-Lee, T.: Linked data - the story so far. Int. J. Semant. Web Inf. Sys. **5**(3), 1–22 (2009)
3. Cheatham, M., Hitzler, P.: The properties of property alignment. In: Proceeding of 9th International Conference on Ontology Matching, vol. 1317, pp. 13–24, Italy (2014)
4. Dubois, D., Prade, H., Harding, E.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York (1988)
5. Dubois, D., Prade, H.: Possibility theory and its applications: a retrospective and prospective view. In: Decision Theory and Multi-Agent Planning, vol. 482, pp. 89–109. Springer (2006)
6. Gunaratna, K., Thirunarayan, K., Jain, P., Sheth, A., Wijeratne, S.: A statistical and schema independent approach to identify equivalent properties on linked data. In: ISEM 2013, Graz, Austria (2013)
7. Hossein Zadeh, P.D., Reformat, M.: Context-aware similarity assessment within semantic space formed in linked data. J. Ambient Intel. Humanized Comp. **4**(4), 515–532 (2013)
8. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty, and Information. Prentice Hall, Englewood Cliffs (1988)
9. Lassila, O., Swick, R.: Resource description framework (RDF) model and syntax specification, 14 February 2017. www.w3.org/TR/1999/REC-rdf-syntax-19990222
10. Liu, Y., Chen, S.-H., Gu, J.-G.: Property alignment of linked data based on similarity between functions. Int. J. Database Theory App. **8**(4), 191–206 (2015)
11. Nguyen, K., Ichise, R., Le, B.: SLINT: a schema-independent linked data interlinking system. In: 7th International Conference on Ontology Matching, USA, pp. 1–12 (2012)
12. Powers, D.M.W.: Evaluation: from precision, recall and F-Measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Tech. **2**(1), 37–63 (2011)
13. Shadbolt, N., Hall, W., Berners-Lee, T.: The semantic web revisited. IEEE Intel. Syst. **21**, 96–101 (2006)
14. Tran, Q., Ichise, R., Ho, B.: Cluster-based similarity aggregation for ontology matching. In: 6th International Conference on Ontology Matching, Germany, pp. 142–147 (2011)
15. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. **100**, 9–34 (1999)
16. https://www.w3.org/RDF. Accessed 14 Mar 2017
17. http://wiki.dbpedia.org. Accessed 14 Mar 2017
18. https://www.wikidata.org/. Accessed 14 Mar 2017
19. http://www.yago-knowledge.org. Accessed 14 Mar 2017

Power Means in Success Likelihood Index Method

Emilio Torres-Manzanera[1(✉)], Susana Montes[1], Irene Díaz[2], Lucía Zapico[2], and Baltasar Gil[3]

[1] Department of Statistics and O.R., University of Oviedo, Oviedo, Spain
{torres,montes}@uniovi.es

[2] Department of Computer Sciences, University of Oviedo, Oviedo, Spain
{sirene,uo220125}@uniovi.es

[3] ESM Research Institute, Oviedo, Spain
baltasargil@esm.es

Abstract. The Successive Likelihood Index Method establishes the degree of liability, and therefore the corresponding compensation, of the various errors that have caused an accident. From an expert judgment, the successive likelihood index of each error is calculated by a weighted arithmetic mean of their opinions. In this work we have considered other averaging functions for aggregating this information and we have studied their behavior. In particular, we have studied in detail the case of power means applied to the accident of the oil tanker Aegean Sea.

Keywords: Weighted power means · Success Likelihood Index · Aegean Sea

1 Introduction

From the World War II, the reliability analysis techniques have been widely used, in order to reduce the incidence rate of severe accidents in many circumstances. Initially, these techniques were focused on technical aspects of the design and quality of the machinery. However, some investigations proved that human error was the most common cause of failures in a lot of situations (see, for instance, [1,4,18]). Thus, for instance, IEEE published a report about human reliability in 1972. From then, a large number of studies about the incorporation of human factors in risk assessment have been developed (see, for instance, [2,8,9,11,13,21]).

Among the proposed human reliability analysis methods, the Success Likelihood Index Method (SLIM) [20] is one of the most widely used. This method was derived from multi-attribute utility theory. It was initially investigated by the U.S. Nuclear Regulatory Commission in 80 ages, as a method for using expert judgments to estimate human error probabilities in nuclear power plants [10]. From it, the Success Likelihood Index Methodology was developed and evaluated by Brookhaven National Laboratory, as a method of obtaining human reliability estimates from expert judges. A detailed description of this method can be seen in [6].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_39

It is proven to be a simple an flexible decision-analytic approach [14,15] and it has been studied, compared and combined with other approaches in several more recent papers (see, for instance, [5,19]). It is based on the idea that the success likelihood of a task depends on the combined effects of a set of performance shaping factors. The way to do this combination of the effects is given by a weighted arithmetic mean. However, there are a lot of different ways of aggregating these effects by means of different aggregation functions [3].

Thus, our main purpose is to analyze the behavior of these indexes in accordance to the aggregation function considered for combining the different effects. In particular, we have focused in some specific functions such as the power means family.

This work is organized as follows. In the following section we recall some definitions and results concerning aggregation functions and we are going to introduce the SLIM. Section 3 is devoted to analyze the differences between the obtained output for the real data of a famous accident. In Sect. 4 we continue the study of the influence of the parameters in a more general environment. Finally, in Sect. 5 we provide some conclusions and open problems.

2 Preliminaries

2.1 SLIM

The steps of SLIM described in [6], following the description of the method given at [10], are:

1. Constitution of the group of experts and first approach to the case analysis.
2. Definition and selection of the performance shaping factors for the case of analysis.
3. Assignment of weighting factor for each performance shaping factor.
4. Scoring of each performance shaping factor.
5. Calculating of the success likelihood index.
6. Conversion of the index in human error probability.

According to [19], the first step establishes criteria to choose experts and give the experts and in-depth description to ensure that all experts share a common understanding of the given task. Then they identify the relevant performance shaping factors (PSF) to the event of interest. Any of them is denoted by PSF_i for $i = 1, \ldots, n$. Then, a weight (w_i with $0 \leq w_i \leq 1$) is associated to any PSF_i. Once this is done, it is necessary to score each PSF_i for each task T_j, with $j = 1, \ldots, m$ with a value from 1 to 9, depending on the PSF characteristics.

Usually, this information is summarized by means of a table, which describes the tasks and the PSFs with their weights and scores. Thus, it is easier to obtain the Success Likelihood Index (SLI) associated to any task j, which is usually defined by means of the weighted arithmetic mean as follows:

$$SLI_j = \sum_{i=1}^{n} w_i T_{ij} \tag{1}$$

for any $j \in \{1, 2, \ldots, m$, where T_{ij} is the scale rating of task j on the i-th PSF.

The conversion of the SLI in human error probability (HEP) is giving by means of the following formula:

$$LogP = a \cdot SLI + b \tag{2}$$

where P represents the probability and a and b are values such that they are calculated from the SLIes of two tasks where the HEP is already known.

In order to better understand the method we are going to consider it for a famous and unfortunate accident.

Example 1. Aegean Sea was a double-bottom Greek-flagged oil tanker. On December 3, 1992, it was en route to Repsol refinery in A Coruña, Spain when it suffered an accident off the Galician coast. This accident caused the cargo of crude oil to be spilled. Then, it affected the coast resulting in ecosystem damage, as well as damage to the fishing and tourist industries in the region. As the ship had successfully passed all required tests and revisions, a detailed study of the causes of the accident was given by the Spanish authorities.

Table 1 shows the human performance factors and their associated weights, which can vary from 0% to 100% and the sum of all of them has to be 100%. In order to apply the SLIM, it is also necessary to describe the tasks and their associated scores with respect to any PSF. This information is described in Table 2. Both tables are based on specialist opinion in this accident [12].

After defining the weight for any PSF and their corresponding scores for any task, it is possible to obtain one SLI per task by applying (1). Thus, for example,

$$SLI_1 = \sum_{i=1}^{9} w_i T_{i1} = (0.1509)3 + (0.1132)2 + \cdots + (0.0566)2 = 3.45 \tag{3}$$

Analogously, we can obtain the associated indexes for the remaining tasks. They are shown in Table 3.

When we consider the average mark of experts, most people would use arithmetic mean, or perhaps its weighted version in order to associate the opinions with the degree of importance. While this procedure is certainly the simplest and most intuitive averaging function, its use is often not warranted: for instance, in some Olympic sports the judges' marks are trimmed before averaging them.

The study of average (also called means) is very rich in both mathematical and practical sense. Arithmetic means and power means are classic functions for combining several values into a single value. This process is called aggregation. At the next subsection we will present the most relevant properties using formal definitions, following the notation of [3].

2.2 Averaging Functions

Let us briefly recall the main definition related to aggregating functions and, in particular, averaging functions. A detailed studied about these kind of functions can be found in [3,7].

Table 1. PSF values

PSFs		w_i (%)
Meteorology	(PSF_1)	15.09
Methods	(PSF_2)	11.32
Overconfidence	(PSF_3)	11.32
Risk perception	(PSF_4)	13.21
DUS		
Fatigue	(PSF_5)	7.55
Stress	(PSF_6)	9.43
Communication		
Pilot	(PSF_7)	18.87
Company	(PSF_8)	7.55
Routine	(PSF_9)	5.66

Table 2. Scores for the tasks

PSFs	Tasks (T_j)			
	Finished maneuver	Lack of reaction	Deficient communication	Absence of pilot
PSF_1	3	4	7	7
PSF_2	2	4	5	6
PSF_3	3	6	5	2
PSF_4	3	5	5	5
PSF_5	5	2	8	7
PSF_6	4	2	6	7
PSF_7	3	5	2	1
PSF_8	8	8	2	7
PSF_9	2	5	4	2

Table 3. SLIM table

	Tasks (T_j)			
	Finished maneuver	Lack of reaction	Deficient communication	Absence of pilot
SLI	3.45	4.57	4.77	4.64

Definition 1. *An aggregation function is a function of $n > 1$ arguments that maps the (n-dimensional) cube onto an interval $\mathbb{I} = [a, b]$, $f : \mathbb{I}^n \to \mathbb{I}$, with the properties*

1. $f(a, a, \ldots, a) = a$ *and* $f(b, b, \ldots, b) = b$ *for any* $a, b \in \mathbb{I}$.
2. $\boldsymbol{x} \leq \boldsymbol{y}$ *implies* $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{I}^n$.

As we mentioned, typically in SLIM $\mathbb{I} = [1, 9]$. In this methodology, the way to combine the different scores is by means of the weighted arithmetic mean. This is a particular case of aggregation function, the averaging functions.

Definition 2. *An aggregation function f has averaging behavior if for every $\boldsymbol{x} \in \mathbb{I}^n$ it is bounded by*

$$\min(\boldsymbol{x}) \leq f(\boldsymbol{x}) \leq \max(\boldsymbol{x}) \tag{4}$$

Definition 3. *An aggregation function f is said to be idempotent if for every $t \in \mathbb{I}$ the output for input $(t, t, \ldots, t)$ is $f(t, t, \ldots, t) = t$.*

Because of monotonicity of f, idempotency is equivalent to averaging behavior.

An example of averaging function is the weighted arithmetic mean, which is the function

$$M_{\boldsymbol{w}}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i x_i \tag{5}$$

where $\boldsymbol{w} = (w_1, \ldots, w_n)$ is a weighting vector, that is, $w_i \in [0, 1]$ for any $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} w_i = 1$.

The weighted arithmetic mean has a lot of interesting properties: it is an additive function, it fulfills the Jensen inequality, it is a kernel aggregation function, it is a shift-invariant function, etc. However, it is not the only interesting averaging function. Thus, given a weighting vector $\boldsymbol{w}$,

– the weighted geometric mean is the function

$$G_{\boldsymbol{w}}(\boldsymbol{x}) = \prod_{i=1}^{n} x_i^{w_i}, \forall \boldsymbol{x} \in \mathbb{I}^n \tag{6}$$

– the weighted harmonic mean is the function

$$H_{\boldsymbol{w}}(\boldsymbol{x}) = \left(\sum_{i=1}^{n} \frac{w_i}{x_i} \right)^{-1}, \forall \boldsymbol{x} \in \mathbb{I}^n \tag{7}$$

We have that

$$H_{\boldsymbol{w}}(\boldsymbol{x}) \leq G_{\boldsymbol{w}}(\boldsymbol{x}) \leq M_{\boldsymbol{w}}(\boldsymbol{x}), \tag{8}$$

for any vector $\boldsymbol{x} \in \mathbb{I}^n$ and weighting vector $\boldsymbol{w}$ and the equality is obtained if and only if $\boldsymbol{x} = (t, t, \ldots, t)$.

A further generalization of the arithmetic mean is a family called power means (also called root-power means), defined by

Definition 4. *Given a weighting vector $\boldsymbol{w}$ and $r \in \mathbb{R}$, the weighted power mean is the function*

$$M_{\boldsymbol{w},[r]}(\boldsymbol{x}) = \begin{cases} \left(\sum_{i=1}^{n} w_i x_i^r\right)^{1/r} & \text{if } r \neq 0 \\ G_{\boldsymbol{w}}(\boldsymbol{x}) & \text{if } r = 0 \end{cases} \tag{9}$$

The family of weighted power means is augmented to $r = -\infty$ and $r = \infty$ by using the limiting cases

$$M_{\boldsymbol{w},[-\infty]}(\boldsymbol{x}) = \lim_{r\to-\infty} M_{\boldsymbol{w},[r]}(\boldsymbol{x}) = \min(\boldsymbol{x}), \tag{10}$$

$$M_{\boldsymbol{w},[\infty]}(\boldsymbol{x}) = \lim_{r\to\infty} M_{\boldsymbol{w},[r]}(\boldsymbol{x}) = \max(\boldsymbol{x}). \tag{11}$$

However min and max are not themselves power means. The limiting case of the weighted geometric mean is also obtained as

$$M_{\boldsymbol{w},[0]}(\boldsymbol{x}) = \lim_{r\to 0} M_{\boldsymbol{w},[r]}(\boldsymbol{x}) = G_{\boldsymbol{w}}(\boldsymbol{x}). \tag{12}$$

Of course, the family of weighted power means includes the following special cases $M_{\boldsymbol{w},[1]}(\boldsymbol{x}) = M_{\boldsymbol{w}}(\boldsymbol{x})$ and $M_{\boldsymbol{w},[-1]}(\boldsymbol{x}) = H_{\boldsymbol{w}}(\boldsymbol{x})$. Another special case is the weighted quadratic mean

$$M_{\boldsymbol{w},[2]}(\boldsymbol{x}) = Q_{\boldsymbol{w}}(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{n} w_i x_i^2}. \tag{13}$$

The main properties of weighted power mean are:

- The weighted power mean $M_{\boldsymbol{w},[r]}$ is a strictly increasing aggregation function, if all $w_i > 0$ and $0 < r < \infty$;
- $M_{\boldsymbol{w},[r]}$ is a continuous function on $[0, \infty)^n$;
- $M_{\boldsymbol{w},[r]}$ is an asymmetric idempotent function (symmetric if all $w_i = \frac{1}{n}$);
- $M_{\boldsymbol{w},[r]}$ is a homogeneous function, i.e., $M_{\boldsymbol{w},[r]}(\lambda\boldsymbol{x}) = \lambda M_{\boldsymbol{w},[r]}(\boldsymbol{x})$ for all $\boldsymbol{x} \in [0, \infty)^n$ and for all $\lambda \in \mathbb{R}$;
- Weighted power means are comparable: $M_{\boldsymbol{w},[r]}(\boldsymbol{x}) \leq M_{\boldsymbol{w},[s]}(\boldsymbol{x})$ if $r \leq s$; this implies the geometric-arithmetic mean inequality;
- $M_{\boldsymbol{w},[r]}$ has absorbing element (always $a = 0$) if and only if $r \leq 0$ (and all weights w_i are positive);
- $M_{\boldsymbol{w},[r]}$ does not have neutral element.

3 Generalized SLIM Case Study Application

Applying this classic root-power means to the experts' mark of the Aegean Sea accident, we observe the results given at Table 4. For $r = 1$, we obtain the weighted arithmetic mean, $M_{\boldsymbol{w}}$, which is the default formula used by SLIM.

Graphically, this information is represented in Fig. 1.

We can see as depending on the selected averaging function, the importance of the human errors vary considerably. For instance, for the case of the weighted geometric mean ($r = 0$), the most relevant error is "Deficient communication", but for the weighted quadratic mean ($r = 2$), this critical error changes to "Absence of pilot". We can also notice that this error, the "Absence of pilot", is the one which is more affected by the choice of the aggregation function.

Table 4. The SLI for each task varies depending on the power means

r	Tasks			
	Finished maneuver	Lack of reaction	Deficient communication	Absence of pilot
Minimum	2	2	2	1
−10.00	2.37	2.39	2.29	1.18
−9.00	2.41	2.44	2.32	1.20
−8.00	2.46	2.49	2.36	1.23
−7.00	2.51	2.57	2.42	1.27
−6.00	2.57	2.67	2.49	1.32
−5.00	2.64	2.81	2.60	1.39
−4.00	2.72	3.00	2.75	1.50
−3.00	2.81	3.25	2.98	1.68
−2.00	2.92	3.57	3.32	2.01
H_w −1.00	3.05	3.93	3.78	2.66
G_w 0.00	3.22	4.27	4.30	3.68
M_w 1.00	3.45	4.57	4.77	4.64
Q_w 2.00	3.77	4.82	5.15	5.26
3.00	4.15	5.05	5.44	5.63
4.00	4.56	5.25	5.67	5.87
5.00	4.94	5.45	5.86	6.03
6.00	5.29	5.64	6.02	6.15
7.00	5.57	5.81	6.16	6.24
8.00	5.81	5.97	6.29	6.31
9.00	6.02	6.12	6.39	6.38
10.00	6.18	6.26	6.49	6.43
Maximum	8	8	8	7

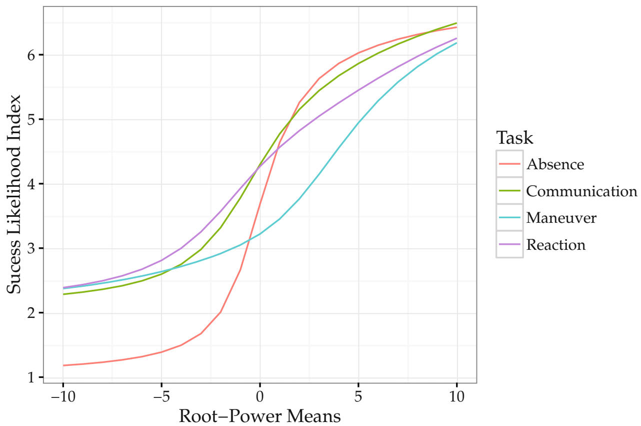


Fig. 1. Different power means provide different SLI for each task.

4 Generalized SLI Method for the Weighted Power Mean

As we could see for the previous case study, there are important difference in accordance to the chosen weighted power mean. We could think these differences are caused by the differences among the different weights and their relationship with the different scores. But this is not the real reason, as we will see along this section.

Proposition 1. *Let* $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ *be a weighting vector and let* $(T_{1j}, T_{2j}, \ldots, T_{nj})$ *be the vector in* $([1,9] \cap \mathbb{Z})^n$ *representing the scores for the task* T_j*. If* $T_{ij} = T_{kj}$ *for any* $i, k \in \{1, 2, \ldots, n\}$ *then the success likelihood index for this task based on any weighted power mean is equal to* T_{1j}*.*

Proof. Since (9), we are going to consider two cases, depending on the values of r. Moreover, we are going to denote by $SLI_j^{[r]}$ the SLI associated to the task j with respect to the weighted power mean $M_{\boldsymbol{w},[r]}$. Thus, for any task j, we have that if $r = 0$, $SLI_j^{[0]} = G_{\boldsymbol{w}}(\boldsymbol{T_j})$. Thus, if we apply (8) we have that $SLI_j^{[0]} \leq M_{\boldsymbol{w}}(\boldsymbol{T_j}) = \sum_{i=1}^{n} w_i T_{ij} = \sum_{i=1}^{n} w_i T_{1j} = T_{1j} \sum_{i=1}^{n} w_i = T_{1j}$ and also that $SLI_j^{[0]} \geq H_{\boldsymbol{w}}(\boldsymbol{T_j}) = \left(\sum_{i=1}^{n} \frac{w_i}{T_{ij}} \right)^{-1} = \left(\frac{1}{T_{1j}} \sum_{i=1}^{n} w_i \right)^{-1} = T_{1j}$.

Otherwise, if $r \neq 0$, we have that $SLI_j^{[r]} = M_{\boldsymbol{w},[r]}(\boldsymbol{T_j}) = \left(\sum_{i=1}^{n} w_i T_{ij}^r \right)^{1/r} = \left(T_{1j}^r \sum_{i=1}^{n} w_i \right)^{1/r} = T_{1j}$. □

Thus, if the scores are constant, the SLI coincides with them for any task and it is the same for any value of r. In order to analyse the behavior for the same mean with a different variability, we are going to consider three different cases: when there is no with no variability in the scores, when an intermediate variability is present, and when the maximum variability is obtained with the same mean. A statistical study of this outputs will help to decide which weighted power mean represents the best choice for a given problem. The original idea of this method of comparison may be found in [17].

Thus, we have considered the case $n = 100$ with three different tasks, which are shown in Table 5. In this table we have also considered three different kind of weighting vectors: $\boldsymbol{w} = (1/100, 1/100, \ldots, 1/100)$, $\boldsymbol{w'} = (100/5050, 99/5050, \ldots, 1/5050)$ and $\boldsymbol{w''} = (1/5050, 2/5050, \ldots, 100/5050)$.

As we could see previously, the SLIes for any task depend on the considered weighted power mean used to calculate them. The values for these indexes for any $r \in [-100, 100]$ are given in Figs. 2, 3 and 4, for the weighting vectors $\boldsymbol{w}$, $\boldsymbol{w'}$ and $\boldsymbol{w''}$, respectively.

Table 5. PSF values, weights, tasks and scores for the experimental study

PSFs	w_i (%)	w'_i (%)	w''_i (%)	Variability		
				No	Medium	High
				Tasks		
				T_1	T_2	T_3
PSF_1	1	1000/505	10/505	5	3	1
PSF_2	1	990/505	20/505	5	3	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
PSF_{50}	1	510/505	500/505	5	3	1
PSF_{51}	1	500/505	510/505	5	7	9
PSF_{52}	1	490/505	520/505	5	7	9
⋮	⋮	⋮	⋮	⋮	⋮	⋮
PSF_{100}	1	10/505	1000/505	5	7	9

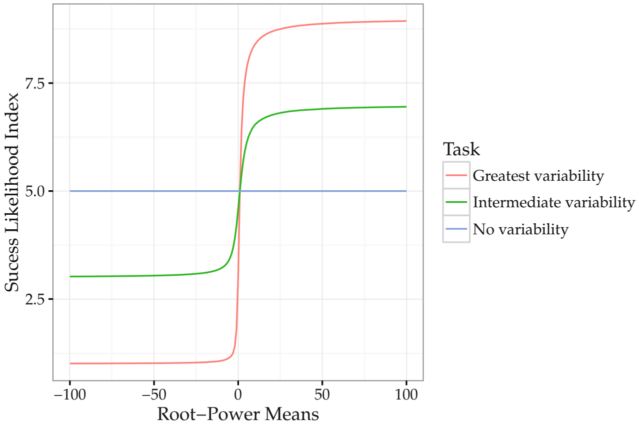


Fig. 2. SLI versus r-power mean for the weighting vector $\boldsymbol{w}$

We can see as, in all the studied cases, if the variability is increasing, the influence of the value of r is greater. Moreover, we can see the differences are significant.

Moreover, if we consider these data as a sample of possible scores under different situations, we can apply the Wilcoxon test for any pair of tasks. Then, we have found evidences of significant differences (p-value < 0.05) for any weighting vector, as it is detailed in Table 6.

Thus, it is clear that the choice of the power mean is an essential step for determining the values of the SLIes and therefore it should be done after a detailed study about the most appropriate for each situation.

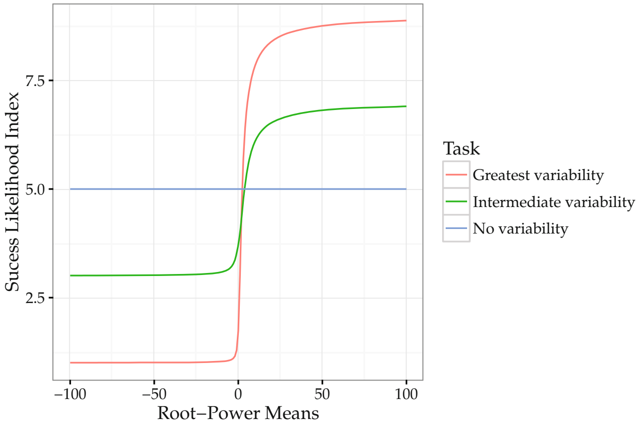


Fig. 3. SLI versus r-power mean for the weighting vector $\boldsymbol{w}'$

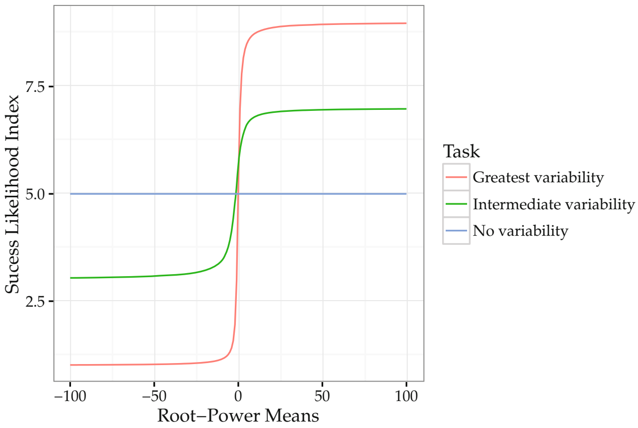


Fig. 4. SLI versus r-power mean for the weighting vector $\boldsymbol{w}''$

Table 6. P-values for Wilcoxon tests

	$T_1 - T_2$	$T_1 - T_3$	$T_2 - T_3$
$\boldsymbol{w}$	$5.3 \cdot 10^{-4}$	$5.6 \cdot 10^{-8}$	$1.0 \cdot 10^{-9}$
$\boldsymbol{w}'$	$2.2 \cdot 10^{-8}$	$2.1 \cdot 10^{-9}$	$6.4 \cdot 10^{-10}$
$\boldsymbol{w}''$	$4.5 \cdot 10^{-3}$	$5.0 \cdot 10^{-3}$	$1.4 \cdot 10^{-9}$

5 Conclusion

In this paper we have analyzed the behavior of the SLIes produced by a family of averaging function, the power mean. After the study of a real case, we have simulated different conditions, to illustrate how the outputs cannot be considered statistically equivalent and therefore, the choice of the parameter value r becomes a critical step in the use of the power means as a tool for the success likelihood index method.

Thus, this work extends the SLIM for a wider family of averaging function, instead of just for the weighted arithmetic mean, in order to be able to consider the most appropriate function based on its specific properties.

This study can be enlarged by analyzing the behavior of other type of averaging functions, in a similar way as the done in [16] and also to study if the choice of a different aggregation function could make sense in some scenarios.

Acknowledgment. The research in this communication has been supported in part by MINECO-TIN2014-59543-P and by PCTI-*Estancias de personal investigador en empresas del Principado de Asturias*. Their financial supports are gratefully acknowledged.

References

1. Aneziris, O.N., Papazoglou, I.A., Mud, M.L., Damen, M., Kuiper, J., Baksteen, H., Ale, B.J., Bellamy, L.J., Hale, A.R., Bloemhoff, A., Post, J.G., Oh, J.: Towards risk assessment for crane activities. Saf. Sci. **46**, 872–884 (2008)
2. Apostolakis, G.E., Bier, V.M., Mosleh, A.: A critique of recent models for human error rate assessment. Reliab. Eng. Syst. Saf. **22**, 201–217 (1988)
3. Beliakov, G., Bustince, H., Calvo, T.: A Practical Guide to Averaging Functions, 1st edn. Springer, Cham (2015)
4. Bellamy, L.J., Geyer, T.A.W., Astley, J.A.: Evaluation of the human contribution to pipework and in-line equipment failure frequencies. HSE Contract Research Report No. 15/1989. UK Health and Safety Executive, Bootle, Merseyside
5. Calixto, E.: Human Reliability Analysis in Gas and Oil Reliability Engineering, 2nd edn. Gulf Professional Publishing, Boston (2016)
6. Calixto, E., Lima, G., Firmino, P.: Comparing SLIM, SPAR-H and Bayesian network methodologies. Open J. Saf. Sci. Technol. **3**(2), 31–41 (2013)
7. Calvo, T., Mayor, G., Mesiar, R. (eds.): Aggregation Operators: New Trends and Applications. Springer, Heidelberg (2007)
8. Castiglia, F., Giardina, M.: Analysis of operator human errors in hydrogen refuelling stations: comparison between human rate assessment techniques. Int. J. Hydrog. Energy **38**, 1166–1176 (2013)
9. Desmorat, G., Guarnieri, F., Besnard, D., Desideri, P., Loth, F.: Pouring CREAM into natural gas: the introduction of common performance conditions into the safety management of gas networks. Saf. Sci. **54**, 1–7 (2013)
10. Embrey, D.E., Humphreys, P.C., Rosa, E.A., Kirwan, B., Rea, K.: SLIM-MAUD: an approach to assessing human error probabilities using structured expert judgement. US Nuclear Regulatory Commission, NUREG/CR-3518, Washington DC (1984)

11. Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczkowski, A., Wreathall, J.: Expert elicitation approach for performing ATHEANA quantification. Reliab. Eng. Syst. Saf. **83**, 207–220 (2004)
12. Gil, B., Calvo, P., Garca, C., Camblor, M.: Factores Humanos para la Investigacin de Accidentes Martimos, Ed. ESM (2003)
13. Groth, K.M., Swiler, L.P.: Bridging the gap between HRA research and HRA practice: a Bayesian network version of SPAR-H. Reliab. Eng. Syst. Saf. **115**, 33–42 (2013)
14. Khan, F.I., Amyotte, P.R., DiMattia, D.G.: HEPI: a new tool for human error probability calculation for offshore operation. Saf. Sci. **44**, 313–334 (2006)
15. Kirwan, B.: A Guide to Practical Human Reliability Assessment. Taylor & Francis, London (1994)
16. Troiano, L., Rodríguez-Muiz, L.J.: A statistical study of differences and similarities among aggregation functions. Log. J. IGPL **19**(2), 415–424 (2011)
17. Troiano, L., Rodríguez-Muiz, L.J., Marinaro, P., Díaz, I.: Statistical analysis of parametric t-norms. Inf. Sci. **257**, 138–162 (2014)
18. Silva, V.A.: O Planejamento de Emergèncias em Refinarias de Petróleo Brasileiras: Um Estudo dos Planos de Refinarias Brasileiras e uma Análise de Acidentes em Refinarias no Mundo e a Apresentaçäo de uma Proposta de Relaçäo de Canários Acidentais para Planejamento. Dissertaçäo (Mestrado em Sistemas de Gestão) Universidade Federal Fluminense, Niterói (2003)
19. Tu, J., Lin, W., Lin, Y.: A Bayes-SLIM based methodology for human reliability analysis of lifting operations. Int. J. Ind. Ergon. **45**, 48–54 (2015)
20. Vestrucci, P.: The logistic model for assessing human error probabilities using the SLIM method. Reliab. Eng. Syst. Saf. **21**, 89–196 (1988)
21. Zimolong, B.: Empirical evaluation of THERP, SLIM and ranking to estimate HEPs. Reliab. Eng. Syst. Saf. **35**, 1–11 (1992)

Three Dimensional Intercriteria Analysis over Intuitionistic Fuzzy Data

Velichka Traneva[1], Stoian Tranev[1], Eulalia Szmidt[2], and Krassimir Atanassov[1,3](✉)

[1] "Prof. Asen Zlatarov" University, "Prof. Yakimov" Blvd, 8000 Bourgas, Bulgaria
veleka13@gmail.com, tranev@abv.bg

[2] Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warsaw, Poland
szmidt@ibspan.waw.pl

[3] Department of Bioinformatics and Mathematical Modelling, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 105, 1113 Sofia, Bulgaria
krat@bas.bg

Abstract. In the paper is extended two dimensional intercriteria analysis over intuitionistic fuzzy data to three dimensional and will be discussed possibility for application of this analysis as an illustration of the application of the intercriteria analysis.

Keywords: Intercriteria analysis · Intuitionistic fuzzy index matrix · Intuitionistic fuzzy pair

AMS Classification: 03E72

1 Introduction

The concept of intercriteria analysis was introduced in [4,7]. The intercriteria analysis is based on the apparatus of the Index Matrices (IMs, [4]) and of Intuitionistic Fuzzy Sets (IFSs, [3]).

The paper is a continuation of the papers [1,6,7,9–13,16] and we for the first time discuss the possibility, the data, that will be processed by three dimensional intercriteria analysis, to be intuitionistic fuzzy pairs (IFP, see [8]) or more general intuitionistic fuzzy data, saved in 3D-intuitionistic fuzzy index matrix (3D-IFIM, [19]).

2 Basic Definitions

2.1 Short Notes on Intuitionistic Fuzzy Pairs

Let us started with some remarks on Intuitionistic Fuzzy Logic from [3,8]. The IFP is an object with the form $\langle a, b\rangle$, where $a, b \in [0, 1]$ and $a + b \leq 1$, that

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_40

is used as an evaluation of some object or process. Its components (a and b) are interpreted as degrees of membership and non-membership, or degrees of validity and non-validity, or degree of correctness and non-correctness, etc. Let us have two IFPs $x = \langle a, b\rangle$ and $y = \langle c, d\rangle$.

Let us have two IFPs $x = \langle a, b\rangle$ and $y = \langle c, d\rangle$. In [8] were defined following operations:

$$\neg x = \langle b, a\rangle$$
$$x\&y = \langle \min(a, c), \max(b, d)\rangle$$
$$x \vee y = \langle \max(a, c)), \min(b, d)\rangle$$
$$x + y = \langle a + c - a.c, b.d\rangle$$
$$x.y = \langle a.c, b + d - b.d\rangle$$
$$x@y = \langle \tfrac{a+c}{2}, \tfrac{b+d}{2}\rangle.$$

and relations

$$\begin{array}{lll} x < y & \text{iff} & a < c \text{ and } b > d \\ x > y & \text{iff} & a > c \text{ and } b < d \\ x \geq y & \text{iff} & a \geq c \text{ and } b \leq d \\ x \leq y & \text{iff} & a \leq c \text{ and } b \geq d \\ x = y & \text{iff} & a = c \text{ and } b = d. \end{array}$$

2.2 Short Remarks on Index Matrices

The concept of Index Matrix (IM) was discussed in a series of papers and collected in [4].

Definition of 3D-Extended Index Matrix (3D-EIM). Let $\mathcal{X}$ be a fixed set of objects (real numbers, numbers 0 or 1, logical variables, propositions or predicates, intuitionistic fuzzy pairs (IFPs), function and etc.). Let $\mathcal{I}$ be a fixed sets of indices and

$$\mathcal{I}^n = \{\langle i_1, i_2, ..., i_n\rangle | (\forall j : 1 \leq j \leq n)(i_j \in \mathcal{I})\} \text{ and } \mathcal{I}^* = \bigcup_{1\leq n\leq\infty} \mathcal{I}^n.$$

By 3D-extended IM (3D-EIM) [4,17], with index sets K, L and H $(K, L, H \subset I^*)$ and elements from the set $\mathcal{X}$ we denote the object :

$$[K, L, H, \{a_{k_i,l_j,h_g}\}] = \left\{ \begin{array}{c|ccccc} h_g & l_1 & \ldots & l_j & \ldots & l_n \\ \hline k_1 & a_{k_1,l_1,h_g} & \vdots & a_{k_1,l_j,h_g} & \ldots & a_{k_1,l_n,h_g} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ k_i & a_{k_i,l_1,h_g} & \ldots & a_{k_i,l_j,h_g} & \ldots & a_{k_i,l_n,h_g} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ k_m & a_{k_m,l_1,h_g} & \ldots & a_{k_m,l_j,h_g} & \ldots & a_{k_m,l_n,h_g} \end{array} \middle| h_g \in H \right\},$$

where $K = \{k_1, k_2, \ldots, k_m\}$, $L = \{l_1, l_2, \ldots, l_n\}$, $H = \{h_1, h_2, \ldots, h_f\}$, and for $1 \leq i \leq m,\ 1 \leq j \leq n,\ 1 \leq g \leq f : a_{k_i,l_j,h_g} \in \mathcal{X}$.

In [4,17,19], different operations, relations and operators are defined over IMs. For the needs of the present research, we will introduce the definitions of some of them. With $3D-EIM_{\mathcal{R}}$ we denote the set of all $3D$-EIMs with elements real numbers, with $3D-EIM_{\{0,1\}}$ – the set of all $(0,1)$-3D-EIMs with elements 0 or 1, with $3D-EIM_{\mathcal{P}}$ – the set of all $3D$-EIMs with elements – predicates and , with $3D-EIM_{\mathcal{IFP}}$ – the set of all 3D-EIMs with elements – IFPs.

We can define the evaluation function V that juxtaposes to this 3D-EIM a new one with elements – IFPs $\langle \mu, \nu \rangle$, where $\mu, \nu, \mu + \nu \in [0, 1]$. The new IM, called Intuitionistic Fuzzy IM (IFIM), contains the evaluations of the variables, propositions, etc., i.e., it has the form

$$\begin{aligned} V([K, L, H, \{a_{k_i,l_j,h_g}\}]) &= [K, L, H, \{V(a_{k_i,l_j,h_g})\}] \\ &= [K, L, H, \{\langle \mu_{k_i,l_j,h_g}, \nu_{k_i,l_j,h_g} \rangle\}] \end{aligned}$$

$$= \left\{ \begin{array}{c|ccc} h_g \in H & l_1 & \dots & l_n \\ \hline k_1 & \langle \mu_{k_1,l_1,h_g}, \nu_{k_1,l_1,h_g} \rangle & \dots & \langle \mu_{k_1,l_n,h_g}, \nu_{k_1,l_n,h_g} \rangle \\ \vdots & \vdots & \dots & \vdots \\ k_i & \langle \mu_{k_i,l_1,h_g}, \nu_{k_i,l_1,h_g} \rangle & \dots & \langle \mu_{k_i,l_n,h_g}, \nu_{k_i,l_n,h_g} \rangle \\ \vdots & \vdots & \dots & \vdots \\ k_m & \langle \mu_{k_m,l_1,h_g}, \nu_{k_m,l_1,h_g} \rangle & \dots & \langle \mu_{k_m,l_n,h_g}, \nu_{k_m,l_n,h_g} \rangle \end{array} \quad | h_g \in H \right\},$$

where for every $1 \leq i \leq m, 1 \leq j \leq n,\ 1 \leq g \leq f$: $V(a_{k_i,l_j,h_g}) = \langle \mu_{k_i,l_j,h_g}, \nu_{k_i,l_j,h_g} \rangle$ and $0 \leq \mu_{k_i,l_j,h_g}, \nu_{k_i,l_j,h_g}, \mu_{k_i,l_j,h_g} + \nu_{k_i,l_j,h_g} \leq 1$.

Aggregation Operations over 3D-EIM. Let the 3D-EIM $A = [K, L, H, \{a_{k_i,l_j,h_g}\}]$ be given, where $K, L, H \subset \mathcal{I}^*$, and let $k_0 \notin K, l_0 \notin L, h_0 \notin H$. Let $\circ : \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}$ and $* : \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}$.

Let

$$\circ \in \begin{cases} \{“+”, “\times”, “average”, “max”, “min”\}, & \text{if } A \in 3D-EIM_{\mathcal{R}}, \\ \{“max”, “min”\}, & \text{if } A \in 3D-EIM_{\{0,1\}} \\ \{“\wedge”, “\vee”\}, & \text{if } A \in 3D-EIM_{\mathcal{P}} \\ & \text{or } A \in 3D-EIM_{\mathcal{IFP}} \end{cases}$$

In the case of $3D-EIM_{\mathcal{IFP}}$, in aggregation operations can participate aggregating pair operations $(\circ, *)$ whose elements are applied respectively on the first and second element of IFP, where

$$\begin{aligned} (\circ, *) \in \{&(min, max)(min, average), (min, min), (average, average), \\ &(average, min), (max, min)\}. \end{aligned}$$

Therefore when $A \in 3D-EIM_{\mathcal{IFP}}$, operations "$(\circ, *)$" are defined for the intuitionistic fuzzy pairs $\langle a, b \rangle$ and $\langle c, d \rangle$, elements of A by

$$\langle a, b \rangle (\circ, *) \langle c, d \rangle = \langle \circ(a, c), *(b, d) \rangle.$$

In all other cases, we use only one operation $(\circ)$.

In [18] were defined aggregation operations. We will recall the following definition:

$(\circ) - \alpha_H$**-aggregation**

$$\alpha_{(H,\circ)}(A, h_0) = \left\{ \begin{array}{c|c} l_j & h_0 \\ \hline k_1 & \underset{1\le g\le f}{\circ} a_{k_1,l_1,h_g} \\ k_2 & \underset{1\le g\le f}{\circ} a_{k_2,l_2,h_g} \\ \vdots & \vdots \\ k_m & \underset{1\le g\le f}{\circ} a_{k_m,l_n,h_g} \end{array} \mid l_j \in L \right\}$$

$$= \begin{array}{c|cccc} & l_1 & l_2 & \dots & l_n \\ \hline k_1 & \underset{1\le g\le f}{\circ} a_{k_1,l_1,h_g} & \underset{1\le g\le f}{\circ} a_{k_1,l_2,h_g} & \cdots & \underset{1\le g\le f}{\circ} a_{k_1,l_n,h_g} \\ k_2 & \underset{1\le g\le f}{\circ} a_{k_2,l_1,h_g} & \underset{1\le g\le f}{\circ} a_{k_2,l_2,h_g} & \cdots & \underset{1\le g\le f}{\circ} a_{k_2,l_n,h_g} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k_m & \underset{1\le g\le f}{\circ} a_{k_m,l_1,h_g} & \underset{1\le g\le f}{\circ} a_{k_m,l_2,h_g} & \cdots & \underset{1\le g\le f}{\circ} a_{k_m,l_n,h_g} \end{array}.$$

3 Three Dimensional Intercriteria Analysis Applied over Intuitionistic Fuzzy Data

In this section we extended two-dimensional intercriteria analysis from [12] applied over intuitionistic fuzzy data to three dimensional.

Let us have the set of objects $O = \{O_1, O_2, ..., O_n\}$ that must be evaluated by criteria from the set $C = \{C_1, C_2, ..., C_m\}$ in the index $h_g \in H$ for $1 \le g \le f$, where H is the third fixed scale and h_g is its element. For example, index set H can be interpreted as a time-scale and its elements h_g – as time-moments.

Let us have an 3D-EIM

$$A = \left\{ \begin{array}{c|ccccccc} h_g & O_1 & \cdots & O_i & \cdots & O_j & \cdots & O_n \\ \hline C_1 & a_{C_1,O_1,h_g} & \cdots & a_{C_1,O_i,h_g} & \cdots & a_{C_1,O_j,h_g} & \cdots & a_{C_1,O_n,h_g} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_k & a_{C_k,O_1,h_g} & \cdots & a_{C_k,O_i,h_g} & \cdots & a_{C_k,O_j,h_g} & \cdots & a_{C_k,O_n,h_g} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_l & a_{C_l,O_1,h_g} & \cdots & a_{C_l,O_i,h_g} & \cdots & a_{C_l,O_j,h_g} & \cdots & a_{C_l,O_n,h_g} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_m & a_{C_m,O_1,h_g} & \cdots & a_{C_m,O_i,h_g} & \cdots & a_{C_m,O_j,h_g} & \cdots & a_{C_m,O_n,h_g} \end{array} | h_g \in H \right\},$$

where for every p, q $(1 \le p \le m,\ 1 \le q \le n)$:

(1) C_p is a criterion, taking part in the evaluation,
(2) O_q is an object, being evaluated.
(3) a_{C_p,O_q,h_g} is a variable, formula or $a_{C_p,O_q,h_g} = \langle \alpha_{C_p,O_q,h_g}, \beta_{C_p,O_q,h_g} \rangle$ is an IFP, that is comparable about relation R with the other a-objects, so that for each i, j, k, g: $R(a_{C_k,O_i,h_g}, a_{C_k,O_j,h_g})$ is defined. Let $\overline{R}$ be the dual relation of R in the sense that if R is satisfied, then $\overline{R}$ is not satisfied and vice versa. For example, if "R" is the relation "<", then $\overline{R}$ is the relation ">", and vice versa.

For each index h_g $(1 \leq g \leq f)$ let $S^{\mu}_{k,l,g}$ be the number of cases in which

$$\langle \alpha_{C_k,O_i,h_g}, \beta_{C_k,O_i,h_g} \rangle \leq \langle \alpha_{C_k,O_j,h_g}, \beta_{C_k,O_j,h_g} \rangle$$

and

$$\langle \alpha_{C_l,O_i,h_g}, \beta_{C_l,O_i,h_g} \rangle \leq \langle \alpha_{C_l,O_j,h_g}, \beta_{C_l,O_j,h_g} \rangle,$$

or

$$\langle \alpha_{C_k,O_i,h_g}, \beta_{C_k,O_i,h_g} \rangle \geq \langle \alpha_{C_k,O_j,h_g}, \beta_{C_k,O_j,h_g} \rangle$$

and

$$\langle \alpha_{C_l,O_i,h_g}, \beta_{C_l,O_i,h_g} \rangle \geq \langle \alpha_{C_l,O_j,h_g}, \beta_{C_l,O_j,h_g} \rangle$$

are simultaneously satisfied.

Let $S^{\nu}_{k,l,g}$ be the number of cases in which

$$\langle \alpha_{C_k,O_i,h_g}, \beta_{C_k,O_i,h_g} \rangle \geq \langle \alpha_{C_k,O_j,h_g}, \beta_{C_k,O_j,h_g} \rangle$$

and

$$\langle \alpha_{C_l,O_i,h_g}, \beta_{C_l,O_i,h_g} \rangle \leq \langle \alpha_{C_l,O_j,h_g}, \beta_{C_l,O_j,h_g} \rangle,$$

or

$$\langle \alpha_{C_k,O_i,h_g}, \beta_{C_k,O_i,h_g} \rangle \leq \langle \alpha_{C_k,O_j,h_g}, \beta_{C_k,O_j,h_g} \rangle$$

and

$$\langle \alpha_{C_l,O_i,h_g}, \beta_{C_l,O_i,h_g} \rangle \geq \langle \alpha_{C_l,O_j,h_g}, \beta_{C_l,O_j,h_g} \rangle$$

are simultaneously satisfied. We can see, that $S^{\mu}_{k,l,g} + S^{\nu}_{k,l,g} \leq \frac{n(n-1)}{2}$ for each g, so that $1 \leq g \leq f$.

Now, for every k, l, g, such that $1 \leq k < l \leq m$, $n \geq 2$ and g is fixed, we define

$$\mu_{C_k,C_l} = 2\frac{S^{\mu}_{k,l}}{n(n-1)}, \quad \nu_{C_k,C_l} = 2\frac{S^{\nu}_{k,l}}{n(n-1)}.$$

Hence,

$$\mu_{C_k,C_l,h_g} + \nu_{C_k,C_l,h_g} = 2\frac{S^{\mu}_{k,l,g}}{n(n-1)} + 2\frac{S^{\nu}_{k,l,g}}{n(n-1)} \leq 1.$$

Therefore, $\langle \mu_{C_k,C_l,h_g}, \nu_{C_k,C_l,h_g} \rangle$ is an IFP.

Now, we can construct the IM

$$R = \left\{ \begin{array}{c|ccc} h_g & C_1 & \cdots & C_m \\ \hline C_1 & \langle \mu_{C_1,C_1,h_g}, \nu_{C_1,C_1,h_g} \rangle & \cdots & \langle \mu_{C_1,C_m,h_g}, \nu_{C_1,C_m,h_g} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ C_m & \langle \mu_{C_m,C_1,h_g}, \nu_{C_m,C_1,h_g} \rangle & \cdots & \langle \mu_{C_m,C_m,h_g}, \nu_{C_m,C_m,h_g} \rangle \end{array} |h_g \in H \right\}$$

that determines the degrees of correspondence between criteria $C_1, \ldots, C_m$.

Let apply aggregation operation to the 3D-IM R $= [K, K, H, \{a_{k_i,l_j,h_g}\}]$ $(K, H \subset \mathcal{I}^*)$ and let $h_0 \notin H$. Let $\circ : \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}$ and $* : \mathcal{X} \times \mathcal{X} \longrightarrow \mathcal{X}$.

Let

$$\langle \circ, * \rangle \in \{\langle min, max \rangle, \langle max, min \rangle, \langle average, average \rangle\}.$$

Follow [18] we used aggregation operations as follows:

$(\circ) - \alpha_H$**-aggregation**

$$\alpha_{(H,\circ)}(R, h_0) = \left\{ \begin{array}{c|c} k_i & h_0 \\ \hline k_1 & \underset{1 \le g \le f}{\circ} a_{k_1,k_1,h_g} \\ k_2 & \underset{1 \le g \le f}{\circ} a_{k_2,k_2,h_g} \\ \vdots & \vdots \\ k_m & \underset{1 \le g \le f}{\circ} a_{k_m,k_m,h_g} \end{array} \mid k_i \in K \right\}$$

$$= \begin{array}{c|cccc} & k_1 & k_2 & \ldots & k_m \\ \hline k_1 & \underset{1 \le g \le f}{\circ} a_{k_1,k_1,h_g} & \underset{1 \le g \le f}{\circ} a_{k_1,k_2,h_g} & \cdots & \underset{1 \le g \le f}{\circ} a_{k_1,k_m,h_g} \\ k_2 & \underset{1 \le g \le f}{\circ} a_{k_2,l_1,h_g} & \underset{1 \le g \le f}{\circ} a_{k_2,l_2,h_g} & \cdots & \underset{1 \le g \le f}{\circ} a_{k_2,k_m,h_g} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k_m & \underset{1 \le g \le f}{\circ} a_{k_m,k_1,h_g} & \underset{1 \le g \le f}{\circ} a_{k_m,k_2,h_g} & \cdots & \underset{1 \le g \le f}{\circ} a_{k_m,k_m,h_g} \end{array}.$$

Therefore, finally, R obtains the form

$$R = \begin{array}{c|cc} h_g & C_1 & \cdots \\ \hline C_1 & \langle \underset{1 \le g \le f}{\circ} \mu_{C_1,C_1,h_g}, \underset{1 \le g \le f}{*} \nu_{C_1,C_1,h_g} \rangle & \cdots \\ \vdots & \vdots & \ddots \\ C_m & \langle \underset{1 \le g \le f}{\circ} \mu_{C_m,C_1,h_g}, \underset{1 \le g \le f}{*} \nu_{C_m,C_1,h_g} \rangle & \cdots \end{array}$$

$$\begin{array}{cc} \cdots & C_m \\ \hline \cdots & \langle \underset{1\le g\le f}{\circ} \mu_{C_1,C_m,h_g}, \underset{1\le g\le f}{*} \nu_{C_1,C_m,h_g} \rangle \\ \ddots & \vdots \\ \cdots & \langle \underset{1\le g\le f}{\circ} \mu_{C_m,C_m,h_g}, \underset{1\le g\le f}{*} \nu_{C_m,C_m,h_g} \rangle \end{array},$$

where $\langle \circ, * \rangle \in \{\langle min, max \rangle, \langle max, min \rangle, \langle average, average \rangle\}$.

If the pair $\langle \circ, * \rangle = \langle min, max \rangle$ is used in this aggregation operation, then we obtain pessimistic forecast of intercriteria correlation coefficient values. With pair $\langle \circ, * \rangle = \langle max, min \rangle$, then optimistic evaluations are acquired. With pair $\langle \circ, * \rangle = \langle average, average \rangle$, we obtain the averaged estimate of the intercriteria correlation coefficients.

4 Conclusion

In the presented research, a three dimensional intercriteria analysis over intuitionistic fuzzy data is discussed. Intercriteria Analysis can be applied over intuitionistic fuzzy data to determine possible correlations between the pairs of criteria. In a next research of the authors, the above described constructions will be extend to the case of 3-dimensional multilayer IMs and will be applied to practical data.

In future, follow [14, 15] for the Kendall rank correlation coefficient between two IFSs, we will present another approach to 3-dimensional intercriteria analysis applied over intuitionistic fuzzy data.

Acknowledgments. The authors are thankful for the support provided by the Bulgarian National Science Fund under Grant Ref. No. DFNI-I-02-5 "InterCriteria Analysis: A New Approach to Decision Making" and the project of Asen Zlatarov University under Ref. No. NIS-401/2017 "Modern methods of optimization and business management".

References

1. Angelova, N., Atanassov, K., Riecan, B.: Intercriteria analysis of the intuitionistic fuzzy implication properties. In: Proceedings of the 11th International Workshop on Intuitionistic Fuzzy Sets, Banská Bystrica, Slovakia, 30 October 2015. Int. J. Notes Intuitionistic Fuzzy Sets, Vol. 21(5), pp. 20–23 (2015)
2. Atanassov, K.: Generalized index matrices, Comptes rendus de l'Academie Bulgare des Sciences, Vol. 40(11), pp. 15–18 (1987)
3. Atanassov, K.: On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012)
4. Atanassov, K.: Index Matrices: Towards an Augmented Matrix Calculus. Springer, Cham (2014)

5. Atanassov, K., Atanassova, V., Gluhchev, G.: InterCriteria Analysis: Ideas and problems. Notes on Intuitionistic Fuzzy Sets, Vol. 21(1), pp. 81–88 (2015)
6. Atanassov, K., Atanassova, V., Chountas, P., Mitkova, M., Sotirova, E., Sotirov, S., Stratiev, D.: Intercriteria analysis over normalized data. In: Proceedings of the 8th IEEE Conference Intelligent Systems, Sofia, pp. 136–138, 4–6 September 2016
7. Atanassov, K., Mavrov, D., Atanassova, V.: Intercriteria decision making: a new approach for multicriteria decision making, based on index matrices and intuitionistic fuzzy sets. In: Issues in Intuitionistic Fuzzy Sets and Generalized Nets, Vol. 11, pp. 1–8 (2014)
8. Atanassov, K., Szmidt, E., Kacprzyk, J.: On intuitionistic fuzzy pairs. Notes on Intuitionistic Fuzzy Sets, Vol. 19(3), pp. 1–13 (2013)
9. Atanassova, V., Doukovska, L., Mavrov, D., Atanassov, K.: InterCriteria decision making approach to EU member states competitiveness analysis: Temporal and threshold analysis. In: Angelov, P., et al. (eds.) Intelligent systems 2014, pp. 95–106. Springer, Cham (2015)
10. Atanassova, V., Doukovska, L., Atanassov, K., Mavrov, D.: InterCriteria decision making approach to EU member states competitive analysis. In: Proceedings of 4th International Symposium on Business Modeling and Software Design, Luxembourg, Grand Duchy of Luxembourg, pp. 289–294, 24–26 June 2014. ISBN 978-989-758-032-1
11. Atanassova, V., Mavrov, D., Doukovska, L., Atanassov, K.: Discussion on the threshold values in the InterCriteria Decision Making approach. Notes on Intuitionistic Fuzzy Sets, Vol. 20(2), pp. 94–99 (2014)
12. Bureva, V., Sotirova, E., Atanassova, V., Angelova, N., Atanassov, K.: Intercriteria analysis over intuitionistic fuzzy data (2017, in press)
13. Stratiev, D., Sotirov, S., Shishkova, I., Nedelchev, A., Sharafutdinov, I., Vely, A., Mitkova, M., Yordanov, D., Sotirova, E., Atanassova, V., Atanassov, K., Stratiev, D.D., Rudnev, N., Ribagin, S.: Investigation of relationships between bulk properties and fraction properties of crude oils by application of the intercriteria analysis. Pet. Sci. Technol. **34**(13), 1113–1120 (2016)
14. Szmidt, E., Kasprzyk, J.: The kendall rank correlation between intuitionistic fuzzy sets. In: Proceedings of the World Conference on Soft Computing, San Francisco, 23–26 May 2011
15. Szmidt, E., Kasprzyk, J.: The spearman and kendall rank correlation coefficients between intuitionistic fuzzy sets. In: EUSFLAT-LFA, pp. 521–528. Atlantis Press (2011)
16. Todinova, S., Mavrov, D., Krumova, S., Marinov, P., Atanassova, V., Atanassov, K., Taneva, S.: Blood plasma thermograms dataset analysis by means of intercriteria and correlation analyses for the case of colorectal cancer. Int. J. Bioautomation **20**(1), 115–124 (2016)
17. Traneva, V., Atanassov, K., Rangasami, P.: On 3-dimensional index matrices. Int. J. Res. Sci. Indexed in i-scholar **1**(2), 64–68 (2014)
18. Traneva, V., Sotirova, E., Bureva, V., Atanassov, K.: Aggregation operations over 3-dimensional extended index matrices. Adv. Stud. Contemp. Math. **25**(3), 407–416 (2015)
19. Traneva, V.: On 3-dimensional intuitionistic fuzzy index matrices. Notes Intuitionistic Fuzzy Sets **20**(4), 59–64 (2014)

$\boldsymbol{M}$-bornologies on $\boldsymbol{L}$-valued Sets

Ingrīda Uļjane[1,2(✉)] and Alexander Šostak[1,2]

[1] Department of Mathematics,
University of Latvia, Zellu street 25, Riga 1002, Latvia
{ingrida.uljane,aleksandrs.sostaks}@lu.lv
[2] Institute of Mathematics and CS UL, Raina bulv. 29, Riga 1459, Latvia

Abstract. We develop an approach to the concept of bornology in the framework of many-valued mathematical structures. It is based on the introduced concept of an M-bornology on an L-valued set (X, E), or an LM-bornology for short; here L is an iccl-monoid, M is a completely distributive lattice and $E : X \times X \to L$ is an L-valued equality on the set X. We develop the basics of the theory of LM-bornological spaces and initiate the study of the category of LM-bornological spaces and appropriately defined bounded "mappings" of such spaces.

Keywords: Bornology · L-valued set · Bounded L-fuzzy set · Fuzzy function · LM-valued bornology

1 Introduction and Motivation

1.1 Bornologies and Bornological Spaces

In order to apply the conception of boundedness to the case of a general topological space, Hu S.T. introduced the notions of a bornology and a bornological space [10,11]. A *bornology* on a set X is a family $\mathfrak{B} \subseteq 2^X$ such that

(1B) $\bigcup\{U \mid U \in \mathfrak{B}\} = X$;
(2B) if $U \subseteq V$ and $V \in \mathfrak{B}$ then $U \in \mathfrak{B}$;
(3B) if $U, V \in \mathfrak{B}$ then $U \cup V \in \mathfrak{B}$.

The pair $(X, \mathfrak{B})$ is called *a bornological space* and the sets belonging to $\mathfrak{B}$ are called *bounded*. Given bornological spaces $(X, \mathfrak{B}_X)$ and $(Y, \mathfrak{B}_Y)$, a mapping $f : (X, \mathfrak{B}_X) \to (Y, \mathfrak{B}_Y)$ is called *bounded* if $f(A) \in \mathfrak{B}_Y$ for every $A \in \mathfrak{B}_X$.

Remark. In the original definition of bornology instead of axioms (1B) the following axiom [(1′B)] $\forall x \in X \;\; \Rightarrow \{x\} \in \mathfrak{B}$ was used. It is easy to see that under assumption of axiom (2B), axiom (1B) is equivalent to axiom (1′B).

Important examples of bornological spaces $(X, \mathfrak{B})$ are: a metric space and its bounded subsets; a topological space and its relatively compact subsets; a uniform space and its totally bounded subsets.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_41

At present the theory of bornological spaces is developed in various directions. Most of the research involving bornologies is done in the context of topological linear spaces and topological algebras, see e.g. [8,22], that is in cases when the underlying set, besides topology, is endowed with an algebraic structure. However, a notable research work, is being done in the field of bornologies without referring to the algebraic structure of the underlying set. General bornological spaces play a key role in research of convergence in hyperspaces [2,3,15], in optimization theory [5] and in study of topologies on function spaces [4,17].

The principal aim of this work is to make a contribution to bornological theories in "*Fuzzy Mathematics*", that is in the context of *many-valued mathematical structures*. At present there are only few works dealing with this problem, see [1,18,24,25]. It is quite different from the situation in "*Fuzzy Topology*", which is a very well developed area of theoretical mathematics and besides has important applications to other fields of mathematics, (see e.g. [13,16], etc.).

1.2 L-bornologies or Bornologies on L-power-sets

An L-bornology on a set X, where L is a complete infinitely distributive lattice, was defined in [1] as a subset $\mathfrak{B}$ of the family L^X such that

(1LB) $\bigvee\{B \mid B \in \mathfrak{B}\} = 1_{L^X}$;
(2LB) if $A \leq B$ and $B \in \mathfrak{B}$ then $A \in \mathfrak{B}$;
(3LB) if $A, B \in \mathfrak{B}$ then $A \vee B \in \mathfrak{B}$.

The pair $(X, \mathfrak{B})$ is called *an L-bornological space.*

In [1] also a stronger version of the axiom (1LB) was considered:
(1′LB) $\forall x \in X \;\Rightarrow\; \chi_{\{x\}} \in X$, where $\chi_{\{x\}}$ is the characteristic function of $\{x\}$. The corresponding structure $\mathfrak{B}$ is called a strong L-bornology. In case when 1_L is an isolated element in L, L-bornology and strong L-bornology are equivalent.

1.3 M-valued Bornologies on Powersets

An alternative approach to the problem of fuzzification of the concept of bornology was presented in [24,25]. An *M-valued bornology* on a set X where M is a complete frame, is a mapping $\mathcal{B} : 2^X \to M$ such that

(1MB) $\forall x \in X \;\Rightarrow\; \mathcal{B}(\{x\}) = 1_M$ where 1_M is the top element of M;
(2MB) if $A \subseteq B$ then $\mathcal{B}(A) \geq \mathcal{B}(B)$;
(3MB) for every $A, B \subseteq X$ it holds $\mathcal{B}(A \cup B) \geq \mathcal{B}(A) \wedge \mathcal{B}(B)$.

The pair $(X, \mathcal{B})$ is called an *M-valued bornological space.*

In this paper we consider the concept of an M-valued bornology on the L-power-set of a set X defined as a mapping $\mathcal{B} : L^X \to M$ satisfying conditions analogous to the properties (1MB) – (3MB). Thus this approach is the "roof" for the approaches sketched above. However, having an intention to develop our approach in the sufficiently general framework, we will assume that, instead of an ordinary set X, we deal with a many-valued set, that is a set endowed with a many-valued equality. As an additional incentive to develop our approach in the framework of many-valued sets we consider the role of many-valued sets in the theory of fuzzy topologies: just in this framework many interesting results and examples are obtained, in particular applications in other fields see e.g. [13].

The paper has the following structure. After recalling in Sect. 2 basic concepts needed in the paper, in Sect. 3, we introduce the concept of an M-bornology on an L-valued set, or an LM-bornology for short, and discuss some properties of LM-bornologies. Our next aim is to develop, what can be called, the categorical basics of the theory of LM-valued bornological spaces. This problem is being considered in Sect. 5. However, to approach this problem, we have first to distinguish the class of potential morphisms for this category. Our LM-bornologies are defined on L-valued sets and just fuzzy functions are used as "natural" morphisms between L-valued sets. Therefore, before introducing the category of LM-bornological spaces and bounded fuzzy functions in Sect. 5, we present a brief introduction in the theory of fuzzy functions in Sect. 4.

2 Prerequisites: The Context of the Work

2.1 Lattices and Iccl-Monoids

In this work, two objects, **L** and **M**, will play the fundamental role. By $\mathbf{L} = (L, \leq_L, \wedge_L, \vee_L)$ we denote a complete lattice that is a lattice in which arbitrary suprema (joins) and infima (meets) exist. In particular, the top 1_L and the bottom 0_L elements in L exist and $0_L \neq 1_L$. A lattice $(L, \leq_L, \wedge_L, \vee_L)$ is called infinitely distributive or a frame if $\wedge$ distributes over arbitrary joins:
$\alpha \wedge_L \left(\bigvee_i \beta_i\right) = \bigvee_i \left(\alpha \wedge_L \beta_i\right) \quad \forall \alpha \in L,\ \forall \{\beta_i : i \in I\} \subseteq L.$
In the sequel we usually omit the subscript L in notation of $\leq, \wedge, \vee$.

By an integral commutative complete lattice monoid or *iccl-monoid* for short, following [12], we call a tuple $(L, \leq, \wedge, \vee, *)$ where $(L, \leq, \wedge, \vee)$ is a complete lattice and $(L, *)$ is a commutative associative monoid such that

(1cl) $*$ is monotone: $\alpha \leq \beta \Longrightarrow \alpha * \gamma \leq \beta * \gamma$ for all $\alpha, \beta, \gamma \in L$;
(2cl) $*$ distributes over arbitrary joins:
$\alpha * \left(\bigvee_{i\in I} \beta_i\right) = \bigvee_{i \in I}(\alpha * \beta_i)\ \forall \alpha \in L$, and $\forall \{\beta_i \mid i \in I\} \subseteq L$;
(3cl) $\alpha * 1_L = \alpha\ \forall \alpha \in L$.

In an iccl-monoid a further binary operation $\mapsto$, residuation, is defined:
$\alpha \mapsto \beta = \bigvee \{\lambda \in L \mid \lambda * \alpha \leq \beta\}\ \forall \alpha, \beta \in L$. Residuation is connected with operation $*$ by Galois connection, see [6]: $\alpha * \beta \leq \gamma \Longleftrightarrow \alpha \leq (\beta \mapsto \gamma)$.

By **M** we denote a complete completely distributive lattice $(M, \leq_M, \wedge_M, \vee_M)$ whose bottom and top elements are 0_M and 1_M respectively. Actually we use not the original definition of complete distributivity, see e.g. [6, Definition I-2-8], but its characterization given by G.N. Raney [21]. Namely, given a complete lattice M and $\beta, \alpha \in M$ following [21], see also [6, Excercise IV-3-31], we introduce the so called *"wedge below"* relation $\lhd$ on M:
$\beta \lhd \alpha \Longleftrightarrow (\text{if } K \subseteq M \text{ and } \alpha \leq \bigvee K \text{ then } \exists \gamma \in K,\ \beta \leq \gamma)\,.$

As shown by G.N. Raney [21] a lattice M is completely distributive if and only if relation $\lhd$ has *approximation property*, that is $\alpha = \bigvee\{\beta \in M \mid \beta \lhd \alpha\}$ for every $\alpha \in M$. Moreover, relation $\lhd$ has the following properties [6,21]:

($\lhd$ 1) $\beta \lhd \alpha$ implies $\beta \leq \alpha$;
($\lhd$ 2) $\gamma \leq \beta \lhd \alpha \leq \delta$ implies $\gamma \lhd \delta$;
($\lhd$ 3) $\beta \lhd \alpha$ implies that there exists $\gamma \in L$ such that $\beta \lhd \gamma \lhd \alpha$.

Let $M^0 = \{\alpha \in M | \alpha \lhd 1_M\}$. Obviously $1_M \in M^0$ iff 1_M is not isolated in M, that is $1_M \neq \bigvee\{\alpha | \alpha \lhd 1_M\}$.

2.2 *L*-relations, *L*-valued Equalities and *L*-valued Sets

Given sets X, Y and an iccl-monoid L, by an L-relation between X and Y we call a mapping $R : X \times Y \to L$. In case $X = Y$, an L-relation $E : X \times X \to L$ is called *an L-valued equality* if it is reflexive, that is $E(x, x) = 1_L$ for every $x \in X$; symmetric, that is $E(x, y) = E(y, x)$ for all $x, y \in X$; transitive, that is $E(x, y) * E(y, z) \leq E(x, z)$ for all $x, y, z \in X$. A pair (X, E), where $E : X \times X \to L$ is an L-valued equality on X, is called an L-valued, or a many-valued, set.[1]

A fuzzy set A in an L-valued set (X, E) is called *extensional* [12,14] if $A(x) * E(x, x') \leq A(x')\ \forall x, x' \in X$. The smallest extensional fuzzy set $e(A)$ in (X, E) that is larger than A or equal to A ($A \leq e(A)$) is called the extensional hull of A. Explicitly the extensional hull of A can be defined by $e(A)(x) = \bigvee_{x' \in X} (E(x, x') * A(x'))$, (see e.g. [12,14]). In particular, identifying an element x_0 with the characteristic function $\chi_{\{x_0\}}$, we get the extensional hull of the point x_0, sometimes called a fuzzy singleton: $e(x_0)(x) = E(x_0, x)$.

3 *M*-bornologies on *L*-valued Sets

3.1 *LM*-bornologies: Basic Definitions

Definition 1. *By an M-bornology on the L-valued set (X, E), or just an LM-bornology on (X, E) for short, we call a mapping $\mathcal{B} : L^X \to M$ satisfying the following conditions:*

(1LMB) $\forall\ \alpha \in M^0\ \exists \mathfrak{U} \subseteq L^X$ *such that* $\bigvee \mathfrak{U} = 1_L$ *and* $\mathcal{B}(U) \geq \alpha\ \forall U \in \mathfrak{U}$;
(2LMB) $A \leq B\ \Rightarrow \mathcal{B}(A) \geq \mathcal{B}(B)\ \forall\ A, B \in L^X$;
(3LMB) $\mathcal{B}(A \vee B) \geq \mathcal{B}(A) \wedge \mathcal{B}(B)\ \forall\ A, B \in L^X$;
(4LMB) $\mathcal{B}(e(A)) = \mathcal{B}(A)$.

The triple $(X, E, \mathcal{B})$ is called an LM-bornological space.

Remark (1) The value $\mathcal{B}(A)$ is interpreted as the degree of boundedness of a fuzzy set $A \in L^X$ in the LM-bornology $\mathcal{B}$.

[1] The concepts called here an L-relation and L-valued equality under different names and with different degrees of generality appear in many papers, see e.g. [26,27], etc.

(2) The axioms (2LMB) and (3LMB) together are equivalent to the axiom (3^LMB) $\mathcal{B}(A \vee B) = \mathcal{B}(A) \wedge \mathcal{B}(B)$ for all $A, B \in L^X$.
However, we prefer to split axiom (3^LMB) into two separate axioms.
(3) In case when the relation E is crisp, condition (4LMB) is trivial. In particular, if besides M=2 is the two-element lattice, then our definition reduces to the definition of an L-bornology considered in Subsection 1.2. On the other hand, if L =2 and E is still crisp, then we come to the definition of an M-valued bornology considered in Subsection 1.3. Finally, in case if both lattices are two-element, $L = M = 2$, we obtain definition of a bornology, Subsection 1.1.

An LM-bornology is called *strong*, if it satisfies a stronger version of (1LMB): (1′LMB) $\bigvee\{U \mid U \in L^X, \mathcal{B}(U) = 1_M\} = 1_L$. If 1_M is an isolated element in M, the concepts of an LM-bornology and a strong LM-bornology are equivalent.

3.2 Lattice of *LM*-bornologies

We introduce a partial order relation $\preceq$ on the set $\mathfrak{B}(L, M, X, E)$ of all LM-bornologies on an L-valued set (X, E) by setting for $\mathcal{B}_1, \mathcal{B}_2 \in \mathfrak{B}(L, M, X, E)$:

$$\mathcal{B}_1 \preceq \mathcal{B}_2 \Longleftrightarrow \mathcal{B}_1(A) \geq \mathcal{B}_2(A) \ \forall A \in L^X,$$

and say in this case that $\mathcal{B}_1$ is coarser than $\mathcal{B}_2$, and $\mathcal{B}_2$ is finer than $\mathcal{B}_1$. We show that the partially ordered set $\mathfrak{B}(L, M, X, E \preceq)$ is a complete lattice.

By setting $\mathcal{B}_\perp(A) = 1_M$ for all $A \in L^X$ we get the coarsest LM-bornology on an L-valued set (X, E_X), that is $\mathcal{B}_\perp$ is the bottom of $\mathfrak{B}(L, M, X, E, \preceq)$. We define the finest element in $\mathcal{B}_\top$ as follows. Let $S \subseteq X$ and $\lambda \in L$. We define a fuzzy set $S^\lambda : X \to L$ by setting $S^\lambda(x) = \lambda$ if $x \in S$ and $S^\lambda(x) = 0_L$ otherwise. Then $\mathcal{B}_\top$ defined by

$$\mathcal{B}_\top(A) = \begin{cases} 1_M & \text{if } \exists S \subseteq X, |S| < \aleph_0, \exists \lambda \in L \text{ such that } A \leq e(S^\lambda); \\ 0_M & \text{otherwise.} \end{cases}$$

is the finest LM-bornology on (X, E), that is the top of $\mathfrak{B}(L, M, X, E, \preceq)$.

Further, given a family $\{\mathcal{B}_i : L^X \to M \mid i \in I\}$ of LM-bornologies, we define $\curlyvee_{i \in I} \mathcal{B}_i =: \mathcal{B}_0 : L^X \to M$ by setting $\mathcal{B}_0(A) = \bigwedge_{i \in I} \mathcal{B}_i(A)$.

One can easily see that the mapping $\curlyvee_{i \in I} \mathcal{B}_i : L^X \to M$ thus obtained is an LM-bornology on (X, E). Besides, from the construction it is clear that $\curlyvee_{i \in I} \mathcal{B}_i$ is the lower upper bound of the family $\mathcal{B}_i, i \in I$, in $\mathfrak{B}(L, M, X, E, \preceq)$.

This already guarantees that the family $\mathfrak{B}(L, M, X, E, \preceq, \curlyvee)$ of LM-bornologies is a complete join semi-lattice. Notice however, that the pointwise supremum $\bigvee_i(\mathcal{B}_i(A))$ of the family $\{\mathcal{B}_i : L^X \to M \mid i \in I\}$ need not be an LM-bornology (axiom (3LMB) may be violated). Therefore the infimum $\curlywedge$ of the family $\{\mathcal{B}_i : L^X \to M \mid i \in I\}$ in the partially ordered set $\mathfrak{B}(L, M, X, E, \preceq, \curlyvee)$ is defined by $\curlywedge_{i \in I} \mathcal{B}_i = \curlyvee \{\mathcal{B} \in \mathfrak{B}(L, M, X, E, \preceq, \curlyvee) \mid \mathcal{B} \preceq \mathcal{B}_i \ \forall i \in I\}$.

3.3 Decomposition of an LM Bornology into Level L-bornologies

Given an LM-bornology $\mathcal{B} : L^X \to M$ on a set (X, E) and $\alpha \in M$, let $\mathcal{B}_\alpha = \{A \in L^X \mid \mathcal{B}(A) \geq \alpha\}$. One can easily see that the family $\mathcal{B}_\alpha$ is an L-bornology Sect. 1.2 on the set (X, E). Further, since in a completely distributive lattice every element is the supremum of a family of wedge-below elements, the family of α-levels $\{\mathcal{B}_\alpha \mid \alpha \in M\}$ of an M-bornology is lower semi continuous, in the sense that $\mathcal{B}_\alpha = \bigcap\{\mathcal{B}_\beta \mid \beta \lhd \alpha, \beta \in M\}$ for every $\alpha \in M$. In particular, $\mathcal{B}_{0_M} = L^X$ since $0_M \lhd 0_M$. In the special case when $M = [0,1]$ is the unit interval with the "less" ordering $<$, we have $\mathcal{B}_\alpha = \bigcap\{\mathcal{B}_\beta \mid \beta < \alpha\}$ for every $\alpha > 0_M$, and $\mathcal{B}_{0_M} = L^X$. Thus every LM-bornology can be characterized by its lower semi-continuous decomposition into level L-bornologies.

3.4 Construction of an M-bornology from a Family of L-bornologies on an L-valued Set

As opposite to the previous subsection, we present here a construction of an LM-bornology from an indexed family of crisp L-bornologies.

Let $\mathbf{M} = (M, \leq, \wedge, \vee)$ be a complete completely distributive lattice, and let K be a subset of M, such that K is $\lhd$-approximative, that is $\lambda = \sup\{\alpha \in K \mid \alpha \lhd \lambda\}$ for each $\lambda \in L$.

Since M is completely distributive, we conclude that K satisfies the $\lhd$-interpolation property, that is $\alpha \lhd \beta, \alpha, \beta \in L, \Rightarrow \exists \gamma \in K, \alpha \lhd \gamma \lhd \beta$. In particular, one can take $K = M$ or $K = \mathrm{CP}(M)$, where $\mathrm{CP}(M)$ is the set of all coprimes of the lattice M. (Note that 0_M need not belong to the family K since 0_M is the supremum of the empty family.) Further, let a non-increasing family of L-bornologies on an L-valued set X be given $\{\mathcal{C}_\alpha \mid \alpha \in K\}$ such that $e(A) \in \mathcal{C}_\alpha$ whenever $A \in \mathcal{C}_\alpha$. Given $A \in L^X$ we define $\mathcal{B}(A) = \bigvee\{\alpha \in K \mid A \in \mathcal{C}_\alpha\}$.

Proposition 1. *The mapping $\mathcal{B} : L^X \to M$ defined above is an LM-bornology.*

Proof. Since every $\mathcal{C}_\alpha$ is an L-bornology, and hence $\bigvee \mathcal{C}_\alpha = L^X$, axiom (1LMB) for $\mathcal{B}$ is ensured by the properties of the set K. To show the validity of axiom (2LMB) let $A \leq B$; $A, B \in L^X$. Then $\mathcal{B}(A) = \bigvee\{\alpha \in K \mid A \in \mathcal{C}_\alpha\} \geq \bigvee\{\alpha \in K \mid B \in \mathcal{C}_\alpha\} = \mathcal{B}(B)$. To verify axiom (3′LMB) let $A, B \in L^X$ and assume that

$$\mathcal{B}(A) = \bigvee \{\alpha \in K \mid A \in \mathcal{C}_\alpha\} := \lambda,\ \mathcal{B}(B) = \bigvee \{\alpha \in K \mid B \in \mathcal{C}_\beta\} := \mu$$

for some $\lambda, \mu \in L$, but there exists $\nu \lhd \lambda \wedge \mu$, $\nu \in K$ such that $\mathcal{B}(A \vee B) \leq \nu$. By the properties of the set K we can find $\alpha \in K$ such that $\nu \lhd \alpha \lhd \lambda \wedge \mu$.

Then $A \in C_\alpha$ and $B \in C_\alpha$, and hence $A \vee B \in \mathcal{C}_\alpha$. However this means that $\mathcal{B}(A \vee B) \geq \alpha$ The obtained contradiction shows that $\mathcal{B}(A \vee B) \geq \mathcal{B}(A) \wedge \mathcal{B}(B)$. The validity of (4LMB) for $\mathcal{B}$ follows from the validity of the corresponding axiom for $\mathcal{C}_\alpha$. □

Proposition 2. $\mathcal{B}_\alpha = \bigcap\{\mathcal{C}_\beta \mid \beta \in K,\ \beta \lhd \alpha\}$ *for every* $\alpha \in M$.

Proof. The inclusion $\mathcal{B}_\alpha \supseteq \bigcap\{\mathcal{C}_\beta \mid \beta \in K,\ \beta \lhd \alpha\}$ is clear from the construction of the LM-bornology $\mathcal{B}$. Conversely, if $A \notin \mathcal{B}_\alpha$, then there exists $\beta \lhd \alpha,\ \beta \in K$ such that $A \notin \mathcal{C}_\beta$. Hence $A \notin \bigcap\{\mathcal{C}_\beta \mid \beta \in K,\ \beta \lhd \alpha\}$. □

Given a family of L-bornologies $\{\mathcal{C}_\alpha \mid \alpha \in K\}$ on an L-valued a set (X, E), let the M-bornology $\mathcal{B} : L^X \to M$ be defined as above. We define a new family of L-bornologies $\{\bar{\mathcal{C}}_\alpha \mid \alpha \in M\}$ by setting $\bar{\mathcal{C}}_\alpha := \bigcap\{\mathcal{C}_\beta \mid \beta \lhd \alpha, \beta \in K\}$, where $K \subseteq M$ is $\lhd$-approximative. Further, let an $\bar{\mathcal{B}} : L^X \to M$ be constructed from this family as above, that is $\bar{\mathcal{B}}(A) = \bigvee\{\alpha \in M \mid A \in \bar{\mathcal{C}}_\alpha\}$ for every $A \in L^X$. Then for every $\alpha \in L$ $\bar{\mathcal{B}}_\alpha = \bigcap_{\beta \lhd \alpha, \beta \in M} \bar{\mathcal{C}}_\alpha = \bigcap_{\gamma \lhd \beta, \gamma \in K} \left(\bigcap_{\beta \lhd \alpha} \mathcal{C}_\beta\right) = \bigcap_{\gamma \lhd \alpha, \gamma \in K} \mathcal{C}_\gamma = \mathcal{B}_\alpha$, and hence $\bar{\mathcal{B}} = \mathcal{B}$.

In order to define a category with LM-bornological spaces as objects we have to specify its morphisms. Since our general framework for LM-bornologies is formed by L-valued sets, we take so called fuzzy functions as morphisms for the category of LM-bornological spaces. The reason is that fuzzy functions are well coordinated with L-valued equalities and can be interpreted as morphisms between L-valued sets. In case of ordinary sets, a fuzzy function between them in a natural way can be interpreted as an ordinary function. Therefore, before touching in Sect. 5 the categorical aspects of LM-bornological spaces, we give a brief introduction into the theory of fuzzy functions in the next section.

4 Fuzzy Functions

The concept of a fuzzy function was (independently) introduced in [7,9]. Further fuzzy functions were studied and applied by different authors, see [19,20,23], etc.

Definition 2. *An L-relation $R : X \times Y \to L$ is called a fuzzy function from an L-valued set (X, E_X) to an L-valued set (Y, E_Y) if*

(1ff) $R(x, y) * E_Y(y, y') \le R(x, y')\ \forall x \in X,\ \forall y, y' \in Y$;
(2ff) $E_X(x, x') * R(x, y) \le R(x', y)\ \forall x, x' \in X,\ \forall y \in Y$;
(3ff) $R(x, y) * R(x, y') \le E_Y(y, y')\ \forall x \in X,\ \forall y, y' \in Y$.

For a fuzzy function we use notations both $R : X \times Y \to L$ and $R : (X, E_X) \to (Y, E_Y)$ giving preference to the one which is more convenient in the context.

Remark 1. Let E_X and E_Y be ordinary equalities $=_X$ and $=_Y$ on the sets X and Y respectively. Then an ordinary function $f : X \to Y$ can be realized as a fuzzy function $R_f : (X, =_X) \to (Y, =_Y)$ by setting $R_f(x, y) = \begin{cases} 1_L & \text{if } f(x) = y \\ 0_L & \text{otherwise.} \end{cases}$

Composition of L-relations $R : X \times Y \to L$ and $S : Y \times Z \to L$ is the L-relation $S \circ R : (X, E_X) \to (Z, E_Z)$ defined by $(S \circ R)(x, z) = \bigvee_{y \in Y}(R(x, y) * S(y, z))$. In [9,23], it is shown that composition of two fuzzy functions is a fuzzy function and that $E_X : (X, E_X) \to (X, E_X)$ is the identical morphism. Thus we obtain:

Proposition 3. [9,23] *L-valued sets as objects and fuzzy functions as morphisms constitute a category denoted $\mathcal{F}\mathbf{SET}(L)$.*

Definition 3. [20,23] *Given a fuzzy function $R : X \times Y \to L$ we define its measure of its soundedness by $\mu(R) = \bigwedge_x \bigvee_y R(x, y)$. A fuzzy function R is called sound if $\mu(R) = 1$.*

Given L-valued sets (X, E_X) and (Y, E_Y) and a fuzzy function $R : X \times Y \to L$, we define its degree of surjectivity by $\sigma(R) = \bigwedge_y \bigvee_x R(x, y)$. A fuzzy function R is called surjective if $\sigma(R) = 1$.

If $R : (X, E_X) \to (Y, E_Y)$, $S : (Y, E_Y) \to (Z, E_Z)$ are fuzzy functions, then $\mu(S \circ R) \geq \mu(R) * \mu(S)$ and $\sigma(S \circ R) \geq \sigma(R) * \sigma(S)$ [23]. Hence, composition of sound functions is sound and composition of surjective functions is surjective.

Let $R : X \times Y \to L$ be a fuzzy function. Then referring to the generalization of the Zadeh's extension principle [28], we come to the following definition:

Definition 4. [9,23] *The forward power-set operator $R^{\to} : L^X \to L^Y$ induced by the fuzzy function R is defined by $R^{\to}(A)(y) = \bigvee_x (R(x, y) * A(x))$ $\forall A \in L^X$, $\forall y \in Y$. Fuzzy set $R^{\to}(A) \in L^Y$ is called the image of the fuzzy set A under the L-relation $R : X \times Y \to L$.*

The backward power-set operator $R^{\leftarrow} : L^Y \to L^X$ induced by the fuzzy function R is defined $R^{\leftarrow}(B)(x) = \bigvee_y R(x, y) * B(y)$ $\forall B \in L^Y$, $\forall x \in X$. Fuzzy set $R^{\leftarrow}(B) \in L^X$ is called the preimage of the fuzzy set $B \in L^Y$ under L-relation $R : X \times Y \to L$.

Proposition 4. [20,23] *Let $R : (X, E_X) \to (Y, E_Y)$ be a sound fuzzy function. Then:*

(1) $R^{\to}\left(\bigvee_{i \in I}(A_i)\right) = \bigvee_{i \in I} R^{\to}(A_i)$ $\forall \{A_i \mid i \in I\} \subseteq L^X$;
(2) $R^{\to}(A_1 \wedge A_2) \leq R^{\to}(A_1) \wedge R^{\to}(A_2)$ $\forall A_1, A_2 \in L^X$;
(3) $R^{\leftarrow}\left(\bigwedge_{i \in I} B_i\right) = \bigwedge_{i \in I}(R^{\leftarrow} B_i)$.
(4) $R^{\leftarrow}\left(\bigvee_{i \in I} B_i\right) = \bigvee_{i \in I}(R^{\leftarrow} B_i)$ $\forall \{B_i : i \in I\} \subseteq L^Y$
(5) $A \leq R^{\leftarrow}(R^{\to}(A)$ $\forall A \in L^X$;
(6) *If R is surjective and B is extensional, than $R^{\to}(R^{\leftarrow}(B) = B$;*
(7) *If R is surjective then $R^{\to}(\lambda_X) = \lambda_X$ for each $\lambda \in L$.*

5 Category $\mathcal{F}$BORN (L, M) of LM-bornological Spaces

5.1 Bounded Fuzzy Functions of LM-bornological Spaces

Definition 5. *A fuzzy function $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}_Y)$ is called bounded if $\mathcal{B}_Y \circ R^{\to} \geq \mathcal{B}_X$.*

Explicitly the boundedness of a fuzzy function $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}_Y)$ means that $\mathcal{B}_Y(R^{\to}(A)) \geq \mathcal{B}_X(A)$ for every $A \in L^X$. Since composition of two bounded fuzzy functions is obviously bounded and the identical fuzzy function $E_X : (X, E_X, \mathcal{B}_X) \to (X, E_X, \mathcal{B}_X)$ is bounded, we get.

Proposition 5. *LM-bornological spaces and bounded fuzzy functions form a category which will be denoted $\mathcal{F}$BORN(L, M).*

Since composition of sound fuzzy functions is sound, we can distinguish a useful subcategory $\mathcal{FS}$**BORN** of the category $\mathcal{F}$**BORN** whose objects are the same as in $\mathcal{F}$**BORN** and whose morphisms are sound fuzzy functions.

When studying properties of categories of LM-bornologies the construction of LM-bornologies from LM-bornology bases will be useful.

Definition 6. *Let $\mathcal{L}$ be a subset of L^X closed under finite unions and extensional hulls. A mapping $\mathcal{C} : \mathcal{L} \to M$ is called an LM-bornology base if*

(1LMBB) $\forall\ \alpha \in M^0\ \exists \mathfrak{U} \subseteq \mathcal{L}$ *such that* $\bigvee \mathfrak{U} = 1_{L^X}$ *and* $\mathcal{C}(U) \geq \alpha\ \forall U \in \mathfrak{U}$;
(2LMBB) $\mathcal{C}(A \vee B) = \mathcal{C}(A) \wedge \mathcal{C}(B)\ \forall A, B \in \mathcal{L}$;
(3LMBB) $\mathcal{C}(e(A)) = \mathcal{C}(A)\ \forall A \in \mathcal{L}$.

The proofs of the following statements are straightforward and we omit them:

Proposition 6. *Let $\mathcal{C} : \mathcal{L} \to M$ be an LM-bornology base on X. By setting $\mathcal{B}_\mathcal{C}(A) = \bigvee\{\mathcal{C}(D) : D \in \mathcal{L}, A \leq D\}$ we obtain an LM-bornology $\mathcal{B}_\mathcal{C} : L^X \to M$. Besides $B_\mathcal{C}$ is the finest LM-bornology such that $B_\mathcal{C}(A) \geq \mathcal{C}(A)$ for all $A \in \mathcal{L}$.*

Theorem 1. *Let $(X, E_X, \mathcal{B}_\mathcal{C})$ be an LM-bornological space, where $\mathcal{C} : \mathcal{L} \to M$ is a base for LM-bornology $\mathcal{B}_\mathcal{C}$. Then a fuzzy function $R : (X, E_X, \mathcal{B}_\mathcal{C}) \to (Y, E_Y, \mathcal{B}_Y)$ is bounded if and only if $\mathcal{B}_Y(R^\to(A)) \geq \mathcal{B}_\mathcal{C}(A)$ for every $A \in \mathcal{L}$.*

5.2 Preimages of LM-bornologies and Initial LM-bornologies Induced by Families of Sound Fuzzy Functions

Theorem 2. *Let $R : (X, E_X) \to (Y, E_Y, \mathcal{B}_Y)$ be a sound fuzzy function and let $\mathcal{L} := \{A = R^\leftarrow(D) \mid D \in L^Y\}$. Then $\mathcal{C}_X = \mathcal{B}_Y \circ R^\leftarrow : \mathcal{L} \to M$ is a base for an LM-bornology $\mathcal{B}_X$ on (X, E_X). Besides $\mathcal{B}_X$ is the coarsest LM-bornology for which the fuzzy function $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}_Y)$. is bounded.*

Sketch of the proof. In order to verify axiom (1LMBB) for the mapping $\mathcal{C}_X = \mathcal{B}_Y \circ R^\leftarrow : \mathcal{L} \to M$ let $\alpha \in M^0$ and take a family $\mathcal{V} = \{V_\xi \mid \xi \in \Xi\} \subseteq L^Y$ such that $\bigvee_\xi V_\xi = 1_{L^X}$ and $\mathcal{B}_Y(V_\xi) \geq \alpha\ \forall \xi \in \Xi$. Let $\mathfrak{U} = \{U_\xi \mid U_\xi = R^\leftarrow V_\xi, V_\xi \in \mathcal{V}\}$, then $\mathcal{C}_X(U_\xi) \geq \alpha$ for all $\xi \in \Xi$. Moreover, referring to the soundedness property of R one can show that the family $\mathfrak{U}$ is an (extensional) cover of the set (X, E).

The proof of the axiom (2LMBB) follows easily from the definition of $\mathcal{C}_X$. Finally we establish (3MLBB) by applying the general easily provable fact that for every $B \in L^Y$ the preimage $R^\leftarrow(B)$ is extensional.

The fact that $\mathcal{B}_X$ is the coarsest LM-bornology for which $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}_Y)$ is bounded is clear from our construction. □

We call the LM-bornology $\mathcal{B}_X$ constructed above *the preimage of the LM-bornology $\mathcal{B}_Y$ under fuzzy function* $R : (X, E_X) \to (Y, E_Y, \mathcal{B}_Y)$ and denote it by $R^\leftarrow(\mathcal{B}_Y)$.

From the previous theorem we easily can prove the existence of the initial LM-bornology induced by a family of functions:

Theorem 3. *For every family $R_i : (X, E_X) \to (Y_i, E_{Y_i}, \mathcal{B}_{Y_i})$ of sound fuzzy functions, there exists the initial LM-bornology on (X, E_X), that is the coarsest LM-bornology $\mathcal{B}_X$ for which all fuzzy functions $R_i : (X, E_X, \mathcal{B}_X) \to (Y_i, E_{Y_i}, \mathcal{B}_{Y_i})$ are bounded.*

We can summarize the above results as follows:

Theorem 4. *Every source $R_i : (X, E_X) \to (Y_i, E_{Y_i}, B_{Y_i}), i \in I$ has a unique initial lift $R_i : (X, E_X, \mathcal{B}_X) \to (Y_i, E_{Y_i}, B_{Y_i}), i \in I$ in the category $\mathcal{FS}$**BORN** (L, M) of sound LM-valued bornological spaces. In particular, products exist in the category $\mathcal{FS}$**BORN**(L, M).*

5.3 Images of LM-bornologies and Final LM-bornologies Induced by a Family of Fuzzy Functions

Let $(X, E_X, \mathcal{B}_X)$ be an LM-bornological space, (Y, E_Y) be an L-valued set and let $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y)$ be a sound fuzzy function. We define a mapping $\mathcal{B}_Y : L^Y \to M$ by setting $\mathcal{B}_Y = \mathcal{B}_X \circ R^{\leftarrow} : L^Y \to M$, or explicitly $\mathcal{B}_Y(B) = \mathcal{B}_X(R^{\leftarrow}(B))$ for every $B \in L^Y$. It is clear that $\mathcal{B}_Y : L^Y \to M$ satisfies property (2LMB), because this property holds for the LM-bornology $\mathcal{B}_Y : L^Y \to M$. Property (3LMB) for $\mathcal{B}_Y : L^Y \to M$ can be established referring to Proposition 4. Noticing that the image $R^{\to}(A)$ of a fuzzy set $A \in L^X$ is extensional, and applying property (4LMB) for $\mathcal{B}_X : L^X \to M$ we can establish the validity of the property (4LMB) for $\mathcal{B}_Y : L^Y \to M$.

So, the mapping $\mathcal{B}_Y := R^{\to}(\mathcal{B}_X) = \mathcal{B}_X \circ R^{\to}$ constructed above satisfies the properties (2LMB), (3LMB), (4LMB) of an LM-bornology on the L-valued set (Y, E_Y).

To establish property (1LMB) we additionally assume that $R : (X, E_X) \to (Y, E_Y)$ is surjective. In this case we fix $\alpha \in M^0$ and referring to the axiom (1LMB) of the LM-bornology $\mathcal{B}_X$ find a family of L-fuzzy sets $\mathfrak{U} \subseteq L^X$ such that $\mathcal{B}_X(U) \geq \alpha$ for every $U \in \mathfrak{U}$ and $\bigvee \{U \mid U \in \mathfrak{U}\} = 1_{L^X}$. For every $U \in \mathfrak{U}$ let $V_U = R^{\to}(U)$ and let $\mathcal{V} = \{V_u \mid U \in \mathfrak{U}\}$. Then, since $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y)$ is surjective by Proposition 4, we have $\bigvee (V_U \mid U \in \mathfrak{U}\} = \bigvee (R^{\to} U \mid U \in \mathfrak{U}\} = R^{\to} \left(\bigvee (U \mid U \in \mathfrak{U}\}\right) = R^{\to}(1_{L^X}) = 1_{L^Y}$,

Since, by definition of $\mathcal{B}_Y$, for every $V_U \in \mathcal{V}$ it holds $\mathcal{B}_Y(V_U) = \mathcal{B}_X(U) \geq \alpha$, the validity of axiom (1LMB) for $\mathcal{B}_X$ is established. We summarize the obtained results in the following.

Theorem 5. *$\mathcal{B}_Y =: R^{\to}(\mathcal{B}_X)$ is the finest LM-bornology for which the sound surjective fuzzy function $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}_Y)$ is bounded.*

Let now $R : (X, E_X) \to (Y, E_Y, \mathcal{B}_Y)$ be *an arbitrary fuzzy function*. Then the constructed LM-bornology $R^{\to}(\mathcal{B}_X)$ may not satisfy axiom (1LMB). Therefore, to get the finest LM-fuzzy topology $\mathcal{B}'_Y : L^Y \to M$, for which $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}'_Y)$ is bounded, we must "strengthen" the mapping $R^{\to}(\mathcal{B}_X)$ by the finest LM-fuzzy topology $\mathcal{B}_Y^{\top}$ on (Y, E_Y) and define $\mathcal{B}'_Y = \mathcal{R}^{(\to)}\mathcal{B}_X \curlywedge \mathcal{B}_Y^{\top}$. Now one can easily prove the following.

Theorem 6. *$\mathcal{B}'_Y = \mathcal{R}^{\rightarrow}(\mathcal{B}_X) \curlywedge \mathcal{B}_Y^{\top}$ is an LM-bornology on the L-valued set (Y, E_Y). Besides it is the finest LM-bornology on (Y, E_Y) for which fuzzy function $R : (X, E_X, \mathcal{B}_X) \to (Y, E_Y, \mathcal{B}'_Y)$ is bounded.*

Let now $\{R_i : (X_i, E_{X_i}, \mathcal{B}_{X_i}) \to (Y, E_Y) \mid i \in I\}$ be *a family of fuzzy functions*. For every $i \in I$ let ${\mathcal{B}_Y^i}' : L^Y \to M$ be the LM-bornology induced by the fuzzy function $R_i : (X_i, E_{X_i}, \mathcal{B}_{X_i}) \to (Y, E_Y)$. Further, let $\mathcal{B}_Y := \curlywedge_{i \in I} R_i^{\rightarrow}(\mathcal{B}_{X_i}) : L^Y \to M$. Analyzing the construction, we can conclude that $\mathcal{B}_Y$ is the finest LM-bornology on the L-valued set (Y, E_Y) for which all fuzzy functions R_i are bounded.

Applying the terminology of the category theory we get the following

Theorem 7. *Every sink $\{R_i : (X_i, E_{X_i}, \mathcal{B}_{X_i}) \to (Y, E_Y) \mid i \in I\}$ in the category $\mathcal{F}$**BORN**(L, M) has a single final lift $R_i : (X_i, E_{X_i}, \mathcal{B}_{X_i}) \to (Y, E_Y, \mathcal{B}_Y), i \in I$. In particular, coproducts exist in the category $\mathcal{F}$**BORN**(L, M).*

5.4 Subcategories of the Category $\mathcal{FS}$BORN(L, M)

Category. $\mathcal{FS}$BORN(L,2) Let M be the two-element lattice $2 = \{0, 1\}$ and let $E : X \times X \to L$ be an L-valued equality on X. Then the $L2$-bornology on a set X is just a family $\mathcal{B} \subseteq L^X$ such that (1 L**2**) $\bigvee\{A \mid A \in \mathfrak{B}\} = 1_{L^X}$; (2 L**2**) $B \leq A, A \in \mathcal{B} \Rightarrow B \in \mathcal{B}$; (3 L**2**) $A_1, A_2 \in \mathcal{B} \Rightarrow A_1 \vee A_2 \in \mathcal{B}$ (4 L**2**) $e(A) = A$.

Now, let **BORN**$(L, 2)$ be the subcategory of $\mathcal{FS}$**BORN**$(L, 2)$ whose objects are $L2$-bornological spaces such that L is endowed with the crisp equality and whose morphisms are fuzzy functions $R_f : (X, =_X, \mathcal{B}_X) \to (Y, =_Y, \mathcal{B}_Y)$ determined by ordinary functions $f : X \to Y$ (see Remark 1) and such that $\mathcal{B}_Y(f(A)) \geq \mathcal{B}_X(A)$. The category **BORN**$(L, 2)$ thus obtained is actually the category of L-bornological spaces, see Subsection 1.2.

Category. $\mathcal{FS}$BORN(2,M) Let $L = 2$ be the two-element lattice. Then 2^X is just the family of all subsets of a set X. Further, let X be considered with the crisp equality $=$. Then the $2M$-bornology on X is a mapping $\mathcal{B} : 2^X \to M$ such that (1 **2**M) $\mathcal{B}(\{x\}) = 1_M$ for every $x \in X$; (2 **2**M) $A \subseteq B \Rightarrow \mathcal{B}(B) \leq \mathcal{B}(A)\ \forall A, B \in \mathfrak{B}$; (3 **2**M) $\mathcal{B}(A \cap B) \geq \mathcal{B}(A) \wedge \mathcal{B}(B)\ \forall A, B \in \mathfrak{B}$.

Now, let **BORN**$(2, M)$ be the subcategory of $\mathcal{FS}$**BORN**$(2, M)$ whose objects are $2M$-bornological spaces and whose morphisms are fuzzy functions $R_f : (X, =_X, \mathcal{B}_X) \to (Y, =_Y, \mathcal{B}_Y)$ determined by ordinary functions $f : X \to Y$ and such that $\mathcal{B}_Y(f(A)) \geq \mathcal{B}_X(A)$. The category **BORN**$(L, 2)$ thus obtained is actually the category of M-valued bornological spaces, see Subsection 1.3.

Acknowledgement. The authors are grateful to an anonymous referee for pointing out some misprints and other minor defects noticed in the first version of the paper.

References

1. Abel, M., Šostak, A.: Towards the theory of L-bornological spaces. Iran. J. Fuzzy Syst. **8**, 19–28 (2011)
2. Beer, G., Naimpally, S., Rodrigues-Lopes, J.: S-topologies and bounded convergences. J. Mathe. Anal. Appl. **339**, 542–552 (2008)
3. Beer, G., Levi, S.: Gap, excess and bornological convergence. Set-Valued Anal. **16**, 489–506 (2008)
4. Beer, G., Levi, S.: Strong uniform continuity. J. Math. Anal. Appl. **350**, 568–589 (2009)
5. Beer, G., Levi, S.: Total boundedness and bornology. Topol. Appl. **156**, 1271–1288 (2009)
6. Gierz, G., Hoffman, K.H., Keimel, K., Lawson, J.D., Mislove, M.W., Scott, D.S.: Continuous Lattices and Domains. Cambridge University Press, Cambridge (2003)
7. Demirci, M.: Fuzzy functions and their fundamental properties. Fuzzy Sets Syst. **106**, 239–246 (1999)
8. Hogle-Nled, H.: Bornology and Functional Analysis. Math. Studies, vol. 26. North-Holland, Amsterdam (1977)
9. Höhle, U., Porst, H.E., Šostak, A.: Fuzzy functions: a fuzzy extension of the category SET and some related categories. Appl. Gen. Topol. **1**(1), 115–127 (2000)
10. Hu, S.-T.: Boundedness in a topological space. J. Math. Pures Appl. **78**, 287–320 (1949)
11. Hu, S.-T.: Introduction to General Topology. Holden-Day, San-Francisko (1966)
12. Höhle, U.: M-valued sets and sheaves over integral commutative cl-monoids. In: Rodabaugh, S.E., Klement, E.P., Höhle, U. (eds.) Applications of Category Theory to Fuzzy Sets, pp. 33–72. Kluwer Academic Press (1992)
13. Höhle, U.: Many Valued Topology and its Application. Kluwer Academic Publishers, Boston, Dodrecht, London (2001)
14. Klawonn, F.: Fuzzy points, fuzzy relations and fuzzy functions. In: Novák, V., Perfilieva, I. (eds.) Discovering the World with Fuzzy Logic, pp. 431–453. Springer, Berlin (2000)
15. Lechicki, A., Levi, S., Spakowski, A.: Bornological convergence. J. Math. Anal. Appl. **297**, 751–770 (2004)
16. Yingming, L., Maokang, L., Topology, F.: Advances in Fuzzy Systems - Applications and Topology. World Scientif, Singapore, New Jersey, London, Hong Kong (1997)
17. Osçağ, S.: Bornologies and bitopological function spaces. Filomat **27**(7), 1345–1349 (2013)
18. Paseka, J., Solovyov, S.A., Stehlik, M.: Lattice-valued bornological systems. Fuzzy Sets and Systems (to appear)
19. Perfilieva, I.: Fuzzy function: Theoretical and practical point of view. In: Proceedings 7th Conference European Society for Fuzzy Logic and Technology, EUSFLAT 2011, Aix-Les-Bains, France, pp. 480–486. Atlantis Press, 18–22 July 2011
20. Perfilieva, I., Šostak, A.: Fuzzy functions: Basics of the Theory and Applications to Topology (2015). Preprint
21. Raney, G.N.: A subdirect-union representation for completely distibutive complete lattices. Proc. Amer. Math. Soc. **4**, 518–522 (1953)
22. Schaefar, H.H.: Topological Vector Spaces. Springer, New York (1970)
23. Šostak, A.: Fuzzy functions and an extension of the category L-TOP of Chang-Goguen L-topological spaces. In: Proceedings 9th Prague Topological Symposium, pp. 271–294 (2001)

24. Šostak, A., Uļjane, I.: Bornological structures in the context of L-fuzzy sets. In: 8th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2013), Atlantis Premium Proceedings, pp. 481-488 (2013)
25. Šostak, A., Uļjane, I.: L - valued bornologies on powersets. Fuzzy Sets Syst. **294**, 93–104 (2016)
26. Valverde, L.: On the structure of F-indistinguishibility operators. Fuzzy Sets Syst. **17**, 313–328 (1985)
27. Zadeh, L.: Similarity relations and fuzzy orderings. Inf. Sci. **3**, 177–200 (1971)
28. Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning I, II, III. Inf. Sci. **8-9**, 199–257, 301–357, 43–80 (1975)

Reduced IFAM Weight Matrix Representation Using Sparse Matrices

Marek Vajgl(✉)

Centre of Excellence IT4Innovations,
Institute for Research and Applications of Fuzzy Modeling,
University of Ostrava, 30. dubna 22, Ostrava, Czech Republic
marek.vajgl@osu.cz
http://irafm.osu.cz

Abstract. The implicative fuzzy associative memories (IFAM) is a tool used to store patterns in a database and to recall desired pattern upon a presentation. The original IFAM model has been later updated to simplify the weight matrix construction. As a result of this improvement, model internally contains only significant values. This article describes how sparse matrix used to capture model's weight matrix can be used to reduce memory-space consumption.

Keywords: IFAM · Weight matrix · Sparse matrix · Implicative fuzzy associative memory

1 Introduction and Motivation

In this contribution, we present a continuation of our research of lattice-based models of implicative fuzzy associative memories (IFAM). We have previously discovered that the weight matrix representing model's internal memory can be obtained using more simplified way than for general IFAM models. Here we are going to present that the weight matrix captured using sparse matrices can significantly reduce memory consumption of the whole IFAM model.

Generally, *associative memories* are memory models with an ability to store submitted patterns and recall them latter upon presentation, even for damaged, noisy, or incomplete inputs. The first presentation of associative memory mechanism has been introduced using matrix associative memories [1,3]. *Fuzzy* approach has been introduced by Kosko's *Fuzzy Associative Memories (FAM)* [4] as a single-layer feedforward neural net containing nonlinear matrix-vector product. FAM has been later extended to increase a storage capacity (e.g. [6]). Another important improvement has been done by adding so-called learning implication rules [5,7], that afterwards leads to *implicative fuzzy associative memory* (IFAM) with *implicative fuzzy learning*.

In [2], theoretical background of IFAM was discussed and examples that demonstrate the ability of IFAM to work with incomplete and noisy information

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_42

were given. The model of IFAM presented there stores patterns using two internal matrices – weight matrix W and bias matrix θ. Based on [2], we have done a theoretical research justifying that IFAM uses the notion and properties of a *fuzzy preorder relation* [8]. This leads to discovering requirements to be fulfilled by input data to ensure insensitivity to a certain type of noise. Another result of this work was an improvement of a learning process to obtain weight matrix W. This model will be referred as *r-IFAM*. Further, following research [9] has shown another ability to reduce the content of the weight matrix W to contain only binary $\{0; 1\}$ values. This model will be referred as *b-IFAM*.

Theoretically, IFAM models work well. However, from the application point of view, IFAM models are internally very memory consuming. When applied in image processing (that means when database patterns to be learned are images), the weight matrix size W contains (number_of_pixels)^2 elements, what becomes a very high number even for images with small resolution. Nevertheless, our extension of r-IFAM and b-IFAM models shows potential how the total occupied size in memory by weight matrix can be reduced using sparse matrices.

The paper organization is as follows: in the second section we introduce previously published IFAM later used in our experiments; the third section explains the *sparse matrix* term and its properties; in the fourth section we present results of our experiments of usage sparse matrices together with IFAM models; the last, fifth section, states conclusion.

2 IFAM Evolution

This section describes the evolution of our research over IFAM. For application purposes we demonstrate the behavior in the field of image processing. So, we have a database of grayscale image patterns $\{\mathbf{x}^1, \ldots, \mathbf{x}^p\}$, each is converted from matrix representation to a n-dimensional vector of pixels with values in $[0; 1]$.

2.1 Implicative Fuzzy Associative Memory – IFAM

By default, implicative fuzzy associative memory finds a fuzzy relation that connects corresponding finite set of inputs with given finite set of outputs. By associativity we mean an ability to represent an association between an input and an output.

As mentioned before, the default IFAM model has been taken from [2]. There presented, fuzzy associative memory is a one level neural network endowed with Pedrycz's neurons with thresholds, whose input-output relation is represented by

$$\mathbf{y}_i = \bigvee_{j=1}^{p} (w_{i,j} \text{ t } \mathbf{x}_j) \vee \theta_i, \; i = 1, \ldots, m, \tag{1}$$

where all elements are from $[0, 1]$, t is a t-norm, $W = (w_{i,j})$ is a connecting fuzzy relation or (in terms of neural networks) is a $n \times n$ synaptic *weight matrix*, θ is a threshold (bias), n represents the length of data and p represents the number of constituent input patterns.

The weight matrix W of IFAM is in the form using the adjoint to t implication $\rightarrow$:

$$w_{i,j} = \bigwedge_{k=1}^{p} (\mathbf{x}_j^k \rightarrow \mathbf{y}_i^k), \tag{2}$$

where $(\mathbf{x}^k, \mathbf{y}^k)$, $k = 1, \ldots, p$, are couples of input-output patterns.

Another part of the model is the bias vector calculated as:

$$\theta_i = \bigwedge_{k=1}^{p} \mathbf{x}_i^k. \tag{3}$$

For further research we aim at the autoassociative fuzzy implicative memory model. The model is based on the Łukasiewicz t-norm $a \text{ t } b = \max(0, a + b - 1)$, and the adjoint implication $a \rightarrow b = \min(1, 1 - a + b)$. By autoassociativity we mean that the input-output patterns are the same objects (images).

Such autoassociative IFAM defined by W and θ returns output image y for the input given by x in accordance with (1).

2.2 Reduced IFAM – r-IFAM

The first research was inspired by the ability of the IFAM to process the noisy input. Let $\tilde{\mathbf{x}}$ is an eroded version of the input image $\mathbf{x}$, that is $\tilde{\mathbf{x}} \leq \mathbf{x}$. In [2] there has been claimed that IFAM is able to remove (or at least suppress) the eroded noise of $\tilde{\mathbf{x}}$, what obviously cannot be true for "zero-eroded" $\tilde{\mathbf{x}}$. We have investigated minimum requirements of $\tilde{\mathbf{x}}$ allowing successful retrieval of $\mathbf{x}$ from IFAM.

In [8], we have built a bridge between IFAM model and fuzzy preorder with eigen sets claiming that IFAM with the weight matrix is a fuzzy preorder relation. Then, if images in an input (learning) database represent normal fuzzy sets, we have achieved following (for this contribution related) conclusions:

- each constituent input image $\mathbf{x}^k$, $k = 1, \ldots, p$, can be retrieved using the simpler version of (1), i.e.

$$\mathbf{y}_i = \bigvee_{j=1}^{p} (\mathbf{x}_j^k \text{ t } w_{ij}),\ i = 1, \ldots, n, \tag{4}$$

  which does not involve bias θ;
- each constituent input image $\mathbf{x}^k$, $k = 1, \ldots, p$, can be retrieved, if the synaptic weight matrix W is equal to Q^r, for which holds (see [8] for details about core elements):

$$Q^r(x, y) = \begin{cases} \mathbf{x}_i^y, & \text{if } x_y \text{ is a core element for } x, \\ 1, & \text{if } x = y, \\ 0, & \text{otherwise .} \end{cases} \tag{5}$$

Both results are important. The first one states that we can omit θ bias matrix without impact at the result. The second one is more important. It states that (a) when the weight matrix is not calculated, but is constructed according to (5), it is tremendously saving computational complexity (already proved in [8]); (b) as the weight matrix might contain mostly 0 values, it might be also saving memory occupation complexity.

2.3 Binary IFAM – b-IFAM

We had continued our research in this field aimed at IFAM retrieval mechanism according to two related systems of fuzzy equations - [9]. We have extended the idea of 'binary skeleton' and have obtained new results, which generally lead to more simplified version of weight matrix W containing now only binary values $\{0; 1\}$. The weight matrix equals to Q^b and is now obtained as

$$Q^b(x, y) = \begin{cases} 1, & \text{if } x_y \text{ is a core element for } x \text{ and } \mathbf{x}_i^y = 1, \\ 0, & \text{otherwise .} \end{cases} \tag{6}$$

As a result, Q^b contains less non-zero values than Q^r, so this approach again decreases required memory used to store IFAM model.

3 Sparse Matrix Application

Experiments created over theoretical background typically need implementation in some environment using a programming language. There are several common approaches applicable to store matrix values across different programming languages. Generally, to store matrix, two dimensional array structure can be used. When declaring such structure, both dimensions of the structure must be known and set. The runtime environment then allocates memory for the whole structure at once. A weight matrix of IFAM model is very large. If dataset image dimension is $n \times m$, weight matrix W will contain $(n \times m)^2$ elements. For a quite small image 200×200 pixels the weight matrix W contains 16×10^8 elements. If every element takes 32b (typical value of "float" data type), the whole weight matrix will occupy 6GB of RAM. Even if less consuming data type is used, it is very hard to fit even 1Mpx image into nowadays computer available RAM.

Simple solution might be found in a usage of sparse matrices. Simply say, a sparse matrix allocates memory for an element only if its value differs from a default value (a default value is typically set to 0). For the case when matrix contains mostly default values and only low number of elements has different value, this approach reduces significantly required memory space.

On the other hand, the sparse matrix may have a disadvantage in performance. If a weight matrix is created as a whole structure (first approach), it can be very simply and quickly accessed. As an internal structure of sparse matrix is more complex, it's usage will very probably cause some performance hit. Also, sparse matrix representation creates some memory space overhead to capture knowledge about the occupied cells. A size of such overhead depends on the internal sparse matrix implementation, and is a typically a trade off between speed an overhead – larger overhead improves processing speed. Simply say, for our case we need $2 \times n + (\frac{m}{n})$ memory cells to store one element. That means, the sparse matrix will reduce total occupied memory size only if it has filled less than 50% elements with non-default value.

In our contribution, we have done implementation of sparse matrix approach for weight matrix W to prove if it can significantly decrease required memory space and therefore increase resolution of images which are used as an IFAM dataset.

4 Experiments

Original implementations of r-IFAM and b-IFAM algorithms were done in Matlab programming environment. However, to work with sparse matrices, some more "common" programming environment has been used and all the aforementioned IFAM approaches were implemented in C# programming language. The image dataset remains the same as for r-IFAM and b-IFAM publications. We have dataset of 48 images rescaled to different resolutions. To fulfill r/b-IFAM requirement let every image is a normal fuzzy set, all of them has been adjusted to fulfill this condition. The simple results are presented in Table 1. This input set contained up to 48 random grayscale images with resolution 40×30 pixels. The model has been progressively learned from 1 up to 48 patterns. After 1, 5, 10, 20 and 48 learned patterns we have measured mean error (that is a simple difference between relayed pattern and model response), utilization of weight matrix (that is how many elements of weight matrix contain non-zero value in comparison of full weigh matrix size) and total required learning time.

The results are quite surprising.

Firstly, the standard IFAM model is much more susceptible to obtain degenerated artifacts. Generally, the more images in the learning dataset, the lower quality of the reconstructed images is obtained. However, the reconstruction quality of IFAM model decreases much faster then for the r-IFAM or b-IFAM model. It should be expected that as IFAM's weight matrix is the most complex, it should preserve the most information for recall. However, as can be also seen in Fig. 1, the IFAM mode (first row in the figure) is the fastest loosing the quality of reconstructed image.

Secondly, the memory space saved by an usage of sparse matrix is tremendous. As can be seen in Table 1, the utilization of weight matrix is extremely low and growing up very slowly. For a comparison, for 120×90 image, the full weight matrix of IFAM model contains 116,640,000 elements. The reduced IFAM (r-IFAM) model needs only 172,744 elements. The binary IFAM model (b-IFAM) needs only 14,113 elements.

At last, the surprise in processing time did not occur. As the IFAM model uses fixed double array structure for weight matrix, it is (for bigger sets) the faster solution than implementations based on sparse matrix. Here we can state, that

Table 1. Comparison of IFAM, r-IFAM and b-IFAM models. Mean error is average error between input pattern and result. W utilization is % of non-default values in weight matrix. Learning time is time needed to learn the model in seconds on i7 CPU.

Dataset size	IFAM type	Mean error	W utilization	Learning time (s)
1	IFAM	0.00000	1.000%	0.472
	r-IFAM	0.00000	0.166%	0.266
	b-IFAM	0.00000	0.084%	0.263
5	IFAM	0.28398	1.000%	4.343
	r-IFAM	0.01977	0.499%	3.955
	b-IFAM	0.00107	0.085%	3.968
10	IFAM	0.37813	1.000%	13.569
	r-IFAM	0.10136	0.916%	14.625
	b-IFAM	0.02049	0.106%	14.586
20	IFAM	0.46044	1.000%	47.064
	r-IFAM	0.17696	1.582%	58.622
	b-IFAM	0.11328	0.182%	55.984
48	IFAM	0.49981	1.000%	248.260
	r-IFAM	0.24735	3.664%	324.147
	b-IFAM	0.11982	0.233%	312.966

for patterns with a small resolution the standard approach should be preferred (from the computation point of view), however, for higher resolution patterns the sparse-matrix based solution might be the only way how the IFAM weigh matrix can be fitted into memory.

Fig. 1. Example of result disruption of different models according to number of learning dataset. Dataset image size 120×90 pixels. Dataset size from left to right: 1, 3, 6, 11. Models from top to down: IFAM, r-IFAM, b-IFAM.

The Fig. 1 shows comparison of results obtained after recall of input image from learning dataset. Again, it is obvious that the more learned images, the lower is the quality of the recall. However, again, it is obvious that the recall quality of IFAM model decreases more than for r-IFAM or b-IFAM models.

5 Summary

We have presented a result from an application point of view of previous research made in the field of implicative fuzzy associative memories. In our previous research we have improved IFAM model by updating the way of weight matrix construction. According to those changes, a sparse matrix become a good tool to capture weight matrix values in the memory. The results show that the usage of sparse matrix for weight matrix can significantly reduce required memory space. Moreover, for the tested set, the reduced and binary IFAM models show better performance (in the meaning of learning/recalling patterns) than the original IFAM model.

Acknowledgment. This research was supported by the project "LQ1602 IT4Innovations excellence in science".

References

1. Steinbuch, K.: Die lernmatrix. Kybernetik **1**, 36–45 (1961)
2. Sussner, P., Valle, M.-E.: Implicative fuzzy associative memories. IEEE Trans. Fuzzy Syst. **14**, 793–807 (2006)
3. Taylor, W.: Eletrical simulation on some nervous system functional activities. Inf. Theory **15**(1), 166–177 (2004)
4. Kosko, B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall, Englewood Cliffs (1992)
5. Cheng, Q., Fan, Z.-T.: The stability problem for fuzzy bidirectional associative memories. Fuzzy Sets Syst. **132**(1), 83–90 (2002)
6. Chung, F.-L., Lee, T.: On fuzzy associative memory with multiple-rule storage capacity. IEEE Trans. Fuzzy Syst. **4**(4), 375–384 (1996)
7. Junbo, F., Fan, J., Yan, S.: A learning rule for fuzzy associative memories. In: Proceedings of the IEEE International Conference on Neural Networks, IEEE World Congress on Computational Intelligence, vol. 7, pp. 4273–4277 (1994)
8. Vajgl, M., Perfiljeva, I.: Autoassociative fuzzy implicative memory on the platform of fuzzy preorder. In: 16th World Congress of the International Fuzzy Systems Association (IFSA) and the 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT): Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology, 30 June 2015, Gijon, Spain. Atlantis Press (2015). s. 1598-1603. ISBN 978-94-62520-77-6
9. Perfiljeva, I., Vajgl, M.: A new fast model of a fuzzy associative memory. In: The 12th International FLINS Conference: Uncertainty Modelling in Knowledge Engineering and Decision Making, 24 August 2016, Roubaix. World Scientific, Singapore (2016). s. 270-275. ISBN 978-981-3146-96-9

A Note on Intuitionistic Fuzzy Modal-Like Operators Generated by Power Mean

Peter Vassilev(✉) and Simeon Ribagin

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

Abstract. In this paper we propose new type of intuitionistic fuzzy modal-like operators generated by the application of the power mean. We study some of their properties and establish some relations between them.

Keywords: Intuitionistic fuzzy modal operators · Necessity · Possibility · Power mean

1 Introduction

Following [1,2], we give the definitions of the basic concepts and of the basic operations, relations and operators over IFSs.

Let us have a fixed universe X and its subset A. The set

$$A^* = \{\langle x, \mu_A(x), \nu_A(x)\rangle \mid x \in X\},$$

where

$$0 \le \mu_A(x) + \nu_A(x) \le 1 \tag{1}$$

is called IFS and the mappings $\mu_A : X \to [0,1]$ and $\nu_A : X \to [0,1]$ represent the degree of membership (validity, etc.) and non-membership (non-validity, etc.), respectively.

The mapping $\pi_A : X \to [0,1]$, given for $x \in X$, by

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$$

is called degree of uncertainty or hesitancy degree.

Further, we accept the premise that for every element $x \in X$, we can view the range of theoretically possible values of $\mu_A(x)$ as $[\mu_A(x), 1-\nu_A(x)]$, and for $\nu_A(x)$ as $[\nu_A(x), 1-\mu_A(x)]$,

We remind that the modal operators $\Box$ (necessity) and $\Diamond$ (possibility) are given by:

$$\Box\,(A) = \{\langle x, \mu_A(x), 1-\mu_A(x)\rangle | x \in X\}; \tag{2}$$

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_43

$$\Diamond(A) = \{\langle x, 1 - \nu_A(x), \nu_A(x)\rangle | x \in X\}. \tag{3}$$

For the last five years, there has been a renewed interest in modal operators see e.g., [3,5,6]. Further we will make use of the following definition:

Definition 1 ([4]). *A power mean of two non-negative numbers x, y is given by:*

$$M_p(x, y) = \left(\frac{x^p + y^p}{2}\right)^{\frac{1}{p}} \tag{4}$$

Some special particular cases are given by:

$$M_{-\infty}(x, y) \stackrel{\text{deff}}{=} \min(x, y)$$

$$M_{\infty}(x, y) \stackrel{\text{deff}}{=} \max(x, y)$$

$$M_0(x, y) = \sqrt{xy},$$

the last obtained as the limit when $p \to 0$.

2 The Modal-Like Operators Generated by the Power Mean

Making note that the possible values that $\mu_A(x)$ can take are in the interval $[\mu_A(x), 1 - \nu_A(x)]$, and $\nu_A(x)$ in the interval $[\nu_A(x), 1 - \mu_A(x)]$, we can rewrite $\Box(A)$ and $\Diamond(A)$ as

$$\Box(A) = \{\langle x, \mu_A(x), M_{\infty}(\nu_A(x), 1 - \mu_A(x)) | x \in X\};$$

$$\Diamond(A) = \{\langle x, M_{\infty}(\mu_A(x), 1 - \nu_A(x)), \nu_A(x)\rangle | x \in X\}.$$

Now, we are ready to propose the following operators:

Definition 2. *We define the operator "M_p-necessity" by*

$$\Box_{M_p}(A) = \{\langle x, \mu_A(x), M_p(\nu_A(x), 1 - \mu_A(x))\rangle | x \in X\}; \tag{5}$$

and the operator "M_p-possibility" by

$$\Diamond_{M_p}(A) = \{\langle x, M_p(\mu_A(x), 1 - \nu_A(x)), \nu_A(x)\rangle\} | x \in X\}. \tag{6}$$

The definition is correct since from the properties of the power mean we have $M_p(x, y) \geq M_q(x, y)$ for $p \geq q$ and,

$$\nu_{\Box_{M_p}(A)} = \left(\frac{\nu_A(x)^p + (1 - \mu_A(x))^p}{2}\right)^{\frac{1}{p}} \leq \max(\nu_A(x), 1 - \mu_A(x)) \leq 1$$

and hence $\mu_{\Box_{M_p}}(A) + \nu_{\Box_{M_p}}(A) \leq 1$;

$$\mu_{\Diamond_{M_p}(A)} = \left(\frac{\mu_A(x)^p + (1 - \nu_A(x))^p}{2}\right)^{\frac{1}{p}} \leq \max(\mu_A(x), 1 - \nu_A(x)) \leq 1$$

and hence $\mu_{\Diamond_{M_p}(A)} + \nu_{\Diamond_{M_p}(A)} \leq 1$.

Definition 3 (c.f. [1, p. 18, (2.9)]**).** *The negation of an intuitionistic fuzzy set A denoted by $\neg A$ is given by*

$$\neg A = \{\langle x, \nu_A(x), \mu_A(x)\rangle | x \in X\}.$$

Theorem 1. *For the newly defined operators we have:*

$$\neg \Box_{M_p}(\neg A) = \Diamond_{M_p}(A)$$

$$\neg \Diamond_{M_p}(\neg A) = \Box_{M_p}(A),$$

The property

$$\Box_{M_p}(A) \subseteq A \subseteq \Diamond_{M_p}(A)$$

is also preserved.

Proof. We have $\neg A = \{\langle x, \nu_A(x), \mu_A(x)\rangle | x \in X\}$, hence

$$\Box_{M_p}(\neg A) = \{\langle x, \nu_A(x), M_p(\mu_A(x), 1 - \nu_A(x))\rangle | x \in X\} = \neg \Diamond_{M_p}(A)$$

Thus,

$$\neg \Box_{M_p}(\neg A) = \neg\neg \Diamond_{M_p}(A) = \Diamond_{M_p}(A)$$

The second relation is proved similarly.

For the third we have to observe that: $\mu_{\Box_{M_p}(A)} = \mu_A \leq \mu_{\Diamond_{M_p}(A)}$ and $\nu_{\Diamond_{M_p}(A)} = \nu_A \leq \nu_{\Box_{M_p}(A)}$, and, hence,

$$\Box_{M_p}(A) \subseteq A \subseteq \Diamond_{M_p}(A)$$

Remark 1. It is noteworthy that for $p \neq \infty$, the other properties of the modal operators are not preserved:

$$\Diamond_{M_p} \Diamond_{M_p}(A) \neq \Diamond_{M_p}(A)$$

$$\Box_{M_p} \Box_{M_p}(A) \neq \Box_{M_p}(A)$$

This is indeed so, due to the fact that

$$\nu_{\Box_{M_p} \Box_{M_p}(A)} \geq \nu_{\Box_{M_p}(A)}$$

and

$$\mu_{\Diamond_{M_p} \Diamond_{M_p}(A)} \geq \mu_{\Diamond_{M_p}(A)},$$

where in the general case the inequality is strict.

Theorem 2. *For some of the other properties to be fulfilled additional requirements have to be requested, for instance for every two IFS A, B and $p \geq 0$ we have:*

$$\Box_{M_p}(A \cap B) = \Box_{M_p}(A) \cap \Box_{M_p}(B) \Leftrightarrow A \subseteq B \vee B \subseteq A$$

$$\Diamond_{M_p}(A \cup B) = \Diamond_{M_p}(A) \cup \Diamond_{M_p}(B) \Leftrightarrow A \subseteq B \vee B \subseteq A$$

Proof. We obtain sequentially:

$$\Box_{M_p}(A \cap B) = \Box_{M_p}\{\langle x, \min(\mu_A(x), \mu_B(x)), \max(\nu_A(x), \nu_B(x)) | x \in X\}$$

$$= \{\langle x, \min(\mu_A(x), \mu_B(x)), M_p(\max(\nu_A(x), \nu_B(x)), 1 - \min(\mu_A(x), \mu_B(x))) | x \in X\}$$

On the other hand,

$$\Box_{M_p}(A) \cap \Box_{M_p}(B)$$
$$= \{\langle x, \min(\mu_A, \mu_B), \max(M_p(1 - \mu_A, \nu_A), M_p(1 - \mu_B, \nu_B)) | x \in X\}$$

We have to show that if $A \subseteq B$ or $B \subseteq A$ then

$$M_p(\max(\nu_A, \nu_B), 1 - \min(\mu_A, \mu_B))$$

$$= \max(M_p(1 - \mu_A, \nu_A), M_p(1 - \mu_B, \nu_B))$$

Let $A \subseteq B$. Then, $\min(\mu_A, \mu_B) = \mu_A$ and $\nu_A = \max(\nu_A, \nu_B)$, hence

$$(1 - \mu_A)^p \geq (1 - \mu_B)^p$$

$$(\nu_A)^p \geq (\nu_B)^p$$

and thus,

$$(1 - \mu_A)^p + (\nu_A)^p \geq (1 - \mu_B)^p + (\nu_B)^p$$

which completes the proof.

The other property is proved in the same manner.

Remark 2. The condition $A \subseteq B \vee B \subseteq A$ is significant in securing the desired property. Consider IFS $A = \{\langle x, 0.2, 0.1\rangle | x \in X\}$ and $B = \{\langle x, 0.5, 0.5\rangle | x \in X\}$. We have for $p = 1$

$$\Box_{M_1}(A \cap B) = \{\langle x, 0.2, 0.65\rangle\}$$

and, on the other hand,

$$\Box_{M_1}(A) \cap \Box_{M_1}(B) = \{\langle x, 0.2, 0.5\rangle\}$$

and thus the property is not fulfilled.

Further we introduce another modal-like operator.

Theorem 3. *The operator* $\boxed{\Diamond}_p$: *IFS*$(X) \to$ *IFS*(X) *given by*

$$\boxed{\Diamond}_p(A) = \{\langle x, M_p(\mu_A(x), 1 - \nu_A(x)), M_p(\nu_A(x), 1 - \mu_A(x))\rangle | x \in X\}; \quad (7)$$

is well defined for $p \leq 1$.

Proof. For $p = 1$, we have $\mu_{\boxed{\Diamond}_p}(A) + \nu_{\boxed{\Diamond}_p}(A) = 1$. Hence, for $q < p$ we will have $\mu_{\boxed{\Diamond}_q}(A) + \nu_{\boxed{\Diamond}_q}(A) \leq 1$.

Remark 3. For this operator it is fulfilled:

$$\neg \boxed{\Diamond}_p(\neg A) = \boxed{\Diamond}_p(A) \tag{8}$$

Theorem 4. *The three operators satisfy (for $p \leq 1$) the following relation*

$$\Box_{M_p}(A) \subseteq \boxed{\Diamond}_p(A) \subseteq \Diamond_{M_p}(A) \tag{9}$$

Proof. We have

$$\mu_{\Box_{M_p}(A)} \leq \mu_{\boxed{\Diamond}_p(A)} = \mu_{\Diamond_{M_p}(A)}$$

and also

$$\nu_{\Diamond_{M_p}(A)} \leq \nu_{\boxed{\Diamond}_p(A)} = \nu_{\Box_{M_p}(A)}.$$

Thus, (9) is fulfilled.

Remark 4. It is clear that $\boxed{\Diamond}_p$ is in some way very similar to the operator $F_{\alpha,\beta}$, since it also increases the value of both the membership and non-membership degrees. In particular, we have to solve the following system of equations in order to find the explicit connection:

$$\begin{cases} \mu + \alpha(1 - \mu - \nu) = M_p(\mu, 1 - \nu) \\ \nu + \beta(1 - \mu - \nu) = M_p(\nu, 1 - \mu) \end{cases} \tag{10}$$

For instance, it is easy to verify that $\boxed{\Diamond}_1 = F_{\frac{1}{2},\frac{1}{2}}$. However, for different values of p, it becomes more difficult to find the solution in simple form.

Further, a way of extending the above operators inspired by the extended modal operator $F_{\alpha,\beta}$ is by making them dependent on two parameters $\alpha \in [0, 1]$ and $\beta \in [0, 1]$. For example, we may introduce the following operator:

$$\boxed{\Diamond}_{p;\alpha,\beta}(A) = \{\langle x, \alpha M_p(\mu_A(x), 1-\nu_A(x)), \beta M_p(\nu_A(x), 1-\mu_A(x))\rangle | x \in X\} \tag{11}$$

There are many more ways to introduce such operators as in (11). In such a way it may even be possible to extend $\boxed{\Diamond}_{p;\alpha,\beta}(A)$ to values of p greater than 1, with an appropriate restriction on the parameters α and β.

3 Conclusion

In near future, we will study in more details the properties of the introduced intuitionsitic fuzzy modal-like operators. Their relation to the existing extended modal operators and their possible use in decision making processes, involving results from the application of InterCriteria Analysis (ICrA), will be investigated.

Acknowledgement. The present research has been supported by the Bulgarian National Science Fund under Grant Ref. No. DFNI-I-02-5 "InterCriteria Analysis: A New Approach to Decision Making".

References

1. Atanassov, K.T.: On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012)
2. Atanassov, K.T.: Properties of the intuitionistic fuzzy extended modal operators. Notes IFS **21**(3), 1–4 (2015)
3. Atanassova, V., Doukovska, L.: Compass-and-straightedge constructions in the intuitionistic fuzzy interpretational triangle: two new intuitionistic fuzzy modal operators. Notes IFS **23**(2), 1–7 (2017)
4. Bullen, P.S.: Handbook of Means and Their Inequalities. Kluwer Academic Publishers, Dordrecht (2003)
5. Yilmaz, S., Çuvalcioğlu, G.: On intuitionistic fuzzy modal operators. Notes IFS **22**(5), 27–36 (2016)
6. Yilmaz, S., Atanassova, V., Atanassov, K., Çuvalcioğlu, G.: Some relationships between new intuitionistic fuzzy modal operators. Notes IFS **22**(5), 37–42 (2016)

On Power Mean Generated Orderings Between Intuitionistic Fuzzy Pairs

Peter Vassilev(✉) and Todor Stoyanov

Institute of Biophysics and Biomedical Engineering,
Bulgarian Academy of Sciences, Acad. G. Bonchev Str. Bl. 105, 1113 Sofia, Bulgaria
peter.vassilev@gmail.com, tstoyanov72@gmail.com

Abstract. In this paper we revisit the topic of orderings between intuitionistic fuzzy pairs and then provide a more general point of view in their introduction. This would allows us to use less strict orderings in producing similarity scores for objects whose evaluations are in the form of intuitionistic fuzzy pair.

Keywords: Intuitionistic fuzzy pairs · Similarity · Orderings · Power mean

1 Introduction

There are many orderings defined on intuitionistic fuzzy pairs. We refer the interested reader to [1,5,6], for some examples. Further we will focus on a previous idea of ours where the $\preceq_\mu$ ordering between intuitionistic fuzzy pairs was introduced [3]. We revisit the idea and after making some observations we provide a general approach to introducing orderings. This is implemented with the help of the power mean. It is possible to use other means as well as the main property of any mean applied to two numbers is to produce a value which lies between the minimal and the maximal number. However, in the current paper we focus, for simplicity, on the power mean.

Further we remind some of the definitions we will require.

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

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

This concept is very important in practice since many methods implementing intuitionistic fuzzy techniques produce estimates in the form of IFPs. The first component usually signifies validity, similarity, agreement, while the second signifies falsity, distance, disagreement, etc. In order to choose between two IFPs an ordering (or another ranking method) must be used.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_44

The classical partial ordering is given by:

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

$$u \le v,$$

iff

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

It is easily seen that $\le$ is only partial ordering, since it is evident that it is transitive, reflexive and antisymmetric but there exist u and v, for which conditions (2) are not satisfied. For instance, the pairs $\langle 0.3, 0.4 \rangle$ and $\langle 0.4, 0.5 \rangle$ are not comparable under the above ordering.

We now turn our attention to the power mean and how it can be used to introduce different orderings.

Definition 3 ([4, p. 198]). *A power mean of two non-negative numbers* x, y *is given by:*

$$M_p(x, y) = \left(\frac{x^p + y^p}{2} \right)^{\frac{1}{p}} \tag{3}$$

Special cases of the power mean are the following: $M_{-\infty}(x, y) = \min(x, y)$, $M_0(x, y) = \sqrt{xy}$, $M_\infty(x, y) = \max(x, y)$.

Another important property of the power mean is the following [4, p. 198]:

$$M_p(x, y) \le M_q(x, y) \text{ for } p \le q.$$

Further, we will require the following definition.

Definition 4. *We will say that a partial ordering* $\preceq_i$ *is stricter than partial ordering* $\preceq_j$*, when* $u \preceq_i v$*, implies* $u \preceq_j v$*, but not vice versa.*

2 The Proposed Power Mean Orderings and Some Results

It is not hard to see that (2) can be rewritten as

$$\begin{cases} u_1 \le v_1 \\ \min(u_2, 1 - u_1) \ge \min(v_2, 1 - v_1) \end{cases}$$

which rewritten with the help of the power mean becomes

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

The ordering defined in [3] given by $u \preceq_\mu v$ for two IFPs $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$, which is given by

$$\begin{cases} u_1 \leq v_1 \\ u_2(1-u_1) \geq v_2(1-v_1). \end{cases} \tag{5}$$

can be restated with the help of the power mean to its equivalent as

$$\begin{cases} u_1 \leq v_1 \\ M_0(u_2, 1-u_1) \geq M_0(v_2, 1-v_1). \end{cases} \tag{6}$$

As can be seen from the above we have obtained two of the existing orderings only with the help of the power mean and different values of p.

This gives us the idea to introduce an ordering $\preceq_{\mu;M_p}$.

Definition 5. *Given $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$ we shall say that u is first component biased power mean based with value p less or equal to v and write $u \preceq_{\mu;M_p} v$ if*

$$\begin{cases} u_1 \leq v_1 \\ M_p(u_2, 1-u_1) \geq M_p(v_2, 1-v_1). \end{cases} \tag{7}$$

The fact that we are only applying the power mean to the second components of the IFPs for these orderings is due to the practical implementations of our current algorithm since the first component is accepted as degree of similarity and is presumed precisely evaluated, while the second has a level of imprecision tied to it. It is, of course, entirely possible to consider similar ordering with respect to the second elements and even to both but this is out of the scope of the current paper.

Below we consider some of the properties of these orderings.

Remark 1. In the case $\preceq_{\mu;M_\infty}$ the partial ordering degenerates to complete, since we are now only comparing the first elements of the pairs as the second inequality becomes equivalent to the first.

Theorem 1. *Let $p \geq 0$. Then we have $\preceq_{\mu;M_{-\infty}}$ is stricter than $\preceq_{\mu;M_p}$.*

Proof. From (4) we have $\min(u_1, v_1) = u_1$ and $\min(u_2, v_2) = v_2$. Hence,

$$\begin{cases} 1-u_1 \geq 1-v_1 \\ u_2 \geq v_2 \end{cases}$$

Thus, we have

$$\begin{cases} (1-u_1)^p \geq (1-v_1)^p \\ (u_2)^p \geq (v_2)^p, \end{cases}$$

Adding the left and right hand sides we get

$$(1-u_1)^p + (u_2)^p \geq (1-v_1)^p + (v_2)^p$$

Therefore,

$$\frac{((1-u_1)^p+(u_2)^p)}{2} \geq \frac{(1-v_1)^p+(v_2)^p}{2}$$

and lastly,

$$\left(\frac{((1-u_1)^p+(u_2)^p)}{2}\right)^{\frac{1}{p}} \geq \left(\frac{(1-v_1)^p+(v_2)^p}{2}\right)^{\frac{1}{p}}.$$

which is exactly $M_p(u_2, 1-u_1) \geq M_p(v_2, 1-v_1)$, and therefore, we have that $\preceq_{\mu;M_{-\infty}}$ implies $u \preceq_{\mu;M_p} v$.

The reverse is not true, as for instance

$$\langle 0.45, 0.51\rangle \preceq_{\mu;M_1} \langle 0.47, 0.52\rangle$$

but they are not comparable under the standard ordering.

It is clear that any $\preceq_{\mu;M_p}$ is stricter than $\preceq_{\mu;M_\infty}$. In some sense, it seems likely that larger values of p relax the strictness of the ordering, e.g., $\preceq_{\mu;M_q}$ is stricter than $\preceq_{\mu;M_p}$, for all $q < p$. However, at present we have no definitive proof of that claim and it is not unlikely that counterexamples can be found.

We provide only the following partial result which is easy to prove.

Theorem 2. *The ordering $u \preceq_{\mu;M_0} v$ is stricter than $\preceq_{\mu;M_1} v$.*

Proof. We already saw that if u, v are IFPs, then $u \preceq_{\mu;M_{-\infty}} v$ implies both $u \preceq_{\mu;M_0} v$ and $u \preceq_{\mu;M_1} v$. Let us suppose that this is not the case. Hence, we have:

$$\begin{cases} u_1 \leq v_1 \\ u_2 \leq v_2. \end{cases} \tag{8}$$

By assumption $u \preceq_{\mu;M_0} v$, which means:

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

This is equivalent to

$$u_2 \geq v_2 \frac{1-v_1}{1-u_1} \tag{9}$$

We need to prove that:

$$u_2 + 1 - u_1 \geq v_2 + 1 - v_1$$

This will be true according to (9) if

$$v_1 - u_1 + v_2 \frac{1-v_1}{1-u_1} - v_2 \geq 0$$

The last is equivalent to

$$v_1 - u_1 - v_2 \left(1 - \frac{1-v_1}{1-u_1}\right) \geq 0$$

which is

$$(v_1 - u_1)\left(1 - \frac{v_2}{1 - u_1}\right) \geq 0$$

But we have $v_1 \geq u_1$, so this inequality will be true if $1 - \frac{v_2}{1-u_1} \geq 0$, i.e., $1 - u_1 - v_2 \geq 0$. Returning to (8), we see that

$$1 - u_1 \geq 1 - v_1,$$

and using (1) we obtain

$$1 - u_1 - v_2 \geq 1 - v_1 - v_2 \geq 0$$

This completes the proof.

Definition 6. *For two IFPs u and v let us call the set of $P \subset R$ their ordering base iff for every $p \in P$.*

$$u \preceq_{\mu;M_p} v \text{ or } v \preceq_{\mu;M_p} u$$

Theorem 3. *For any two IFPs u and v there exists an ordering base P, which is non-empty.*

Proof. Since either $u_1 \leq v_1$ or $v_1 \leq u_1$, we have that at least $\preceq_{\mu;M_\infty}$ is a member of P.

Corollary 1. *From Theorem 2 it follows that if $0 \in P$, then $1 \in P$.*

All of the above considered partial orderings can be used in defining relations between intuitionistic fuzzy sets (IFSs). Indeed, the relation between two IFSs A and B,

$$A \leq B,$$

is generated by the standard partial order between IFPs $\leq$, or in the above notation $\preceq_{\mu;M_{-\infty}}$. This is not surprising, considering the fact that we can think of the IFS as collections of IFPs which are labelled by elements of a universe set X. Thus, results concerning orderings over IFPs can be easily transferred to IFS theory.

3 Conclusion

In the present work we have proposed a new consistent way of introducing partial orderings between IFPs. In future work we will investigate orderings generated by other generalized means (e.g. Lehmer's, Gini's, etc.) and study their properties.

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

References

1. Atanassov, K.T.: On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012)
2. Atanassov, K., Szmidt, E., Kacprzyk, J.: On intuitionistic fuzzy pairs. Notes Intuitionistic Fuzzy Sets **19**(3), 1–13 (2013)
3. Vassilev, P., Stoyanov, T.: On a new ordering between intuitionistic fuzzy pairs. In: Proceedings of the 8th European Symposium on Computational Intelligence and Mathematics, ESCIM 2016, pp. 77–80 (2016)
4. Bullen, P.S.: Handbook of Means and Their Inequalities. Kluwer Academic Publishers, Dordrecht (2003)
5. De Miguel, L., Bustince, H., Fernandez, J., Indurain, E., Kolesarova, A., Mesiar, R.: Construction of admissible linear orders for interval-valued atanassov intuitionistic fuzzy sets with an application to decision making. Inf. Fusion **27**, 189–197 (2016)
6. Marinov, E.: π-ordering and index of indeterminacy for intuitionistic fuzzy sets. In: Proceedings of the 12th International Workshop on IFS and GN, IWIFSGN13, Warsaw, October 2013

Dynamical Behaviors of Fuzzy *SIR* Epidemic Model

Renu Verma(✉), S.P. Tiwari, and Ranjit Kumar Upadhyay

Department of Applied Mathematics, Indian Institute of Technology (ISM), Dhanbad 826004, India
renuismmaths@gmail.com

Abstract. In this paper, we propose and analyze a fuzzy SIR model with an asymptotic transmission rate. Specifically, the fuzziness is due to the consideration of the disease transmission rate, additional death due to disease and rate of recovery from infection as fuzzy sets. Further, a comparative study of the equilibrium points of the disease for the classical and fuzzy models are performed. We study the fuzzy basic reproduction number for groups of infected individuals with different virus loads and compare with a basic reproduction number for the classical model. Finally, a program based on the basic reproduction value $\mathcal{R}_0^f$ of disease control is suggested and the numerical simulations are carried out to illustrate the analytical results.

Keywords: SIR model · Fuzzy expected value · Fuzzy basic reproduction number · Stability

1 Introduction

Epidemiological models have recently begun to explore the part of treatment functions within their dynamic equations. The central idea is that as the number of infectious individuals I, increases, society's resources are organized to counter the potential spread of the infection. Hence a treatment function $T(I)$ is suggested, where $T(I)$ works to restore infectious individuals and hence to reduce the value of the time derivative $\frac{dI}{dt}$. Wang [13] includes the piecewise linear treatment function into an SIR model. Zhang and Liu [14], Li et al. [5], and Hu et al. [4] involve the same treatment function as Wang [13] but apply it in different settings. Zang and Liu use an SIS model. Li et al. consider with an SIR model with saturated incidence rate. Barros et al. [2], Barros and Bassanezi [1], and Bassanezi and Barros [3] proposed a new approach to treat an ecological model using fuzzy parameters in differential equations that describes the dynamical system. In this case, the solution of the set of equations is found to be so-called fuzzy expected value. The recent study on fuzzy parameter done in [6,7,9] described interval valued parameter set in a harvested prey-predator model and an epidemic model with fuzzy parameters which they have studied

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_45

for computer network. Recently, Mondal et al. [8] have studied the dynamical behavior of an epidemic model with fuzzy transmission and Verma et al. [12] have studied the Fuzzy epidemic model for the spread of influenza virus and its possible control.

The organization of this paper is as follows: Sect. 2 deals with the development of the mathematical model where the parameters β, ϵ and Υ are fuzzy set. We analyze the fuzzy system and interpretation of SIR fuzzy model and presents the existence and stability analysis of the fuzzy model system. In Sect. 3, we present a new definition of the fuzzy basic reproduction value which is different from the classical model and provides some conditions for the disease control in fuzzy epidemic system. In Sects. 4 and 5, we give some numerical simulations to verify our results and conclusion.

2 Proposed Fuzzy Model

In the SIR model, it has been described that the dynamics of directly transmitted disease with interaction between susceptible and infected individuals in the absence of vital dynamics (i.e., the rates of birth and mortality are not considered). In this section, we propose a fuzzy SIR model. Following is the system of differential equations describing the proposed model.

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \beta SI - \mu S \\ \frac{dI}{dt} &= \beta SI - (\mu + \epsilon + \Upsilon)I \\ \frac{dR}{dt} &= \Upsilon I - \mu R, \end{aligned} \tag{1}$$

and $S + I + R = 1$, with $S(0) = S_0$, $I(0) = I_0$, $R(0) = R_0$.

A fuzzy SIR model corresponding to Eq. (1) describe as follows:

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \beta(\sigma)SI - \mu S \\ \frac{dI}{dt} &= \beta(\sigma)SI - (\mu + \epsilon(\sigma) + \Upsilon(\sigma))I \\ \frac{dR}{dt} &= \Upsilon(\sigma)I - \mu R, \end{aligned} \tag{2}$$

The parameter $\Lambda, \beta, \mu, \epsilon, \Upsilon$ are positive constants. S is the proportion of susceptible individuals, I is the proportion of infected individuals and R who have been removed from the possibility of infection through full immunity at each instant, fuzzy number β is the transmission coefficient of the disease, fuzzy number ϵ is the additional disease rate, and fuzzy number Υ is the rate of recovery from infection, Λ is the influx of individuals into the susceptible and μ is the natural death rate, σ is the virus load.

2.1 Analysis of the Fuzzy System

In this section, we study the concepts of contact rate $\beta(\sigma)$, additional death rate $\epsilon(\sigma)$, recovery rate from $\Upsilon(\sigma)$ and virus-load $\Gamma(\sigma)$. Towards the end, we provide an analysis and interpretation of the proposed fuzzy model.

We assume that the population in this fuzzy model is given by the infected individuals and disease-induced mortality as the function of the accessible virus. Let $\beta = \beta(\sigma)$ be the chance of transmission to turn out in a meeting between a susceptible and an infected individuals with the amount of virus σ. Then there will be the maximum chance of disease transmission when the virus-load is maximum. Following [2], fuzzy membership function of the transmission parameter $\beta(\sigma)$ is given by

$$\beta(\sigma) = \begin{cases} 0 & \text{if } \sigma < \sigma_m \\ \frac{\sigma - \sigma_m}{\sigma_0 - \sigma_m} & \text{if } \sigma_m \leq \sigma \leq \sigma_0 \\ 1 & \text{if } \sigma_0 < \sigma < \sigma_M. \end{cases}$$

From above, it is clear that if the virus-load is low then the disease transmission will be negligible and that there is a minimum virus-load σ_m is required. Moreover, there should be a certain amount of virus say, σ_0, where the transmission rate is maximum and equal to unity. Again, the amount of virus is always limited by σ_M for each disease. The diagram for membership function of $\beta(\sigma)$ is given in Fig. 1.

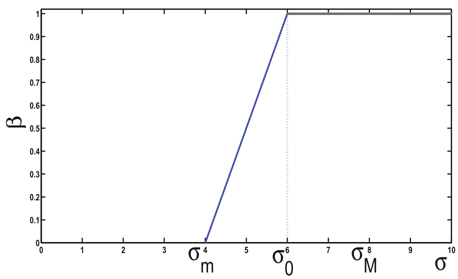


Fig. 1. Membership function of $\beta = \beta(\sigma)$.

The addition death rate can also be assumed to be a fuzzy number as it occurs due to the infection of the disease. When the disease transmission is negligible for low virus load. There is no transmission of disease due to infection, say Υ_0. Also, it is an increasing function of σ. When the amount of virus is at its highest level, i.e., $\sigma_0 < \sigma$, the death will be higher. We assume that maximum value of the additional death is $(1 - \eta)$, $(\eta \geq 0)$. In view of this, the fuzzy membership function of $\epsilon(\sigma)$ is given by

$$\epsilon(\sigma) = \begin{cases} \frac{(1-\eta)-\epsilon_0}{\sigma_0} \sigma + \epsilon_0 & \text{if } 0 \leq \sigma \leq \sigma_0 \\ 1 - \eta & \text{if } \sigma_0 < \sigma, \end{cases}$$

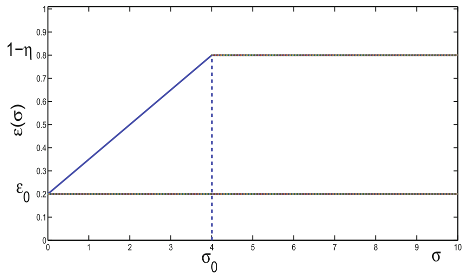


Fig. 2. Membership function of $\epsilon = \epsilon(\sigma)$.

where $0 < \epsilon_0 < 1$, is the lowest additional death rate. The diagram for membership function of additional death rate $\epsilon(\sigma)$ is given in Fig. 2.

Now, $\Upsilon = \Upsilon(\sigma)$ is the recovery rate from infection. The higher the virus load, the longer it will take to recover from infection, i.e., it is a decreasing function of σ. Thus the fuzzy membership function of $\Upsilon(\sigma)$ is given by

$$\Upsilon(\sigma) = \left\{ \frac{(\Upsilon_0 - 1)}{\sigma_M}\sigma + 1 \text{ if } 0 \leq \sigma \leq \sigma_M, \right.$$

where $\Upsilon_0 > 0$ is the lowest recovery rate. The diagram for membership function of additional death rate $\Upsilon(\sigma)$ is given in Fig. 3.

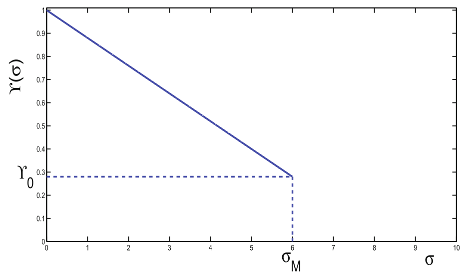


Fig. 3. Membership function of $\Upsilon = \Upsilon(\sigma)$.

We also assume that the virus load of the studied group Σ may be different for different individuals and so, Σ can be viewed as a linguistic variable with classification given by an expert according to the studied group. Each classification is modeled by a fuzzy number whose membership function is given as under.

$$\Gamma(\sigma) = \begin{cases} 0 & \text{if } \sigma < \hat{\sigma} - \delta \\ \frac{\sigma - \hat{\sigma} + \delta}{\delta} & \text{if } \hat{\sigma} - \delta \leq \sigma \leq \hat{\sigma} \\ -\frac{\sigma - \hat{\sigma} - \delta}{\delta} & \text{if } \hat{\sigma} < \sigma \leq \hat{\sigma} + \delta \\ 1 & \text{if } \sigma > \hat{\sigma} + \delta. \end{cases}$$

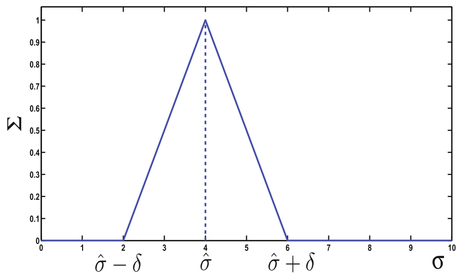


Fig. 4. Membership function of $\Sigma = \Sigma(\sigma)$.

Here the parameter $\hat{\sigma}$ is a central value and δ gives the dispersion of each one of the fuzzy sets assumed by σ. For a fixed $\hat{\sigma}$, $\Gamma(\sigma)$ can has a linguistic meaning, given by an expert, such as weak, medium and high. The diagram for membership function of Σ is given in Fig. 4.

3 Existence, Stability Analysis and Bifurcation of the Fuzzy Model

In this section, we study the existence and stability analysis of the non-negative equilibrium point of the fuzzy model (2). Further, we also study the bifurcation of the same. We refer all these theorem motivated by this paper [11].

For the following equation of fuzzy model (2), we have

$$\frac{dS}{dt} = \Lambda - \beta(\sigma)SI - \mu S$$
$$\frac{dI}{dt} = \beta SI - (\mu + \epsilon(\sigma) + \Upsilon(\sigma))I$$
$$\frac{dR}{dt} = \Upsilon(\sigma)I - \mu R,$$

There are two non-negative equilibrium point of fuzzy model system (2). The existence and the stability conditions for them are as follows.

(i) The disease free equilibrium point $E_1 = (\frac{\Lambda}{\mu}, 0, 0)$ exists on the boundary of the first octant.
(ii) The nontrivial equilibrium $E^*(S^*, I^*, R^*)$ exists if and only if there is a positive solution to the following set of equations.

$$\frac{dS}{dt} = 0,$$
$$\Lambda - \beta(\sigma)SI - \mu S = 0,$$
$$\frac{dI}{dt} = 0,$$
$$\beta(\sigma)SI - (\mu + \epsilon(\sigma) + \Upsilon(\sigma))I = 0,$$

$$\frac{dR}{dt} = 0,$$
$$\Upsilon(\sigma)I - \mu R = 0.$$

Thus there are two equilibrium points $E_1 = (\frac{\Lambda}{\mu}, 0, 0)$ and $E^* = \left(\frac{\mu+\epsilon(\sigma)+\Upsilon(\sigma)}{\beta(\sigma)}, \frac{\Lambda(\sigma}{\mu+\epsilon(\sigma)+\Upsilon(\sigma)} - \frac{\mu}{\beta(\sigma)}, 1 - S - I\right)$.

For the two equilibrium points one for disease free and another for endemic in epidemiology, we have the following.

Theorem 1. *The system has a disease-free equilibrium* $(\frac{\Lambda}{\mu}, 0, 0)$ *and a unique endemic equilibrium* $\left(\frac{\mu+\epsilon(\sigma)+\Upsilon(\sigma)}{\beta(\sigma)}, \frac{\mu}{\beta(\sigma)}(\mathcal{R}_0 - 1), 1 - S - I\right)$, *where* $\mathcal{R}_0 = \frac{\beta(\sigma)\Lambda}{\mu(\mu+\epsilon(\sigma)+\Upsilon(\sigma))}$ *is the basic reproduction number. The endemic equilibrium exists only when* $\mathcal{R}_0 > 1$.

Theorem 2. *The disease-free equilibrium* $(\frac{\Lambda}{\mu}, 0, 0)$ *is locally asymptotically stable when* $\mathcal{R}_0 < 1$ *and unstable when* $\mathcal{R}_0 > 1$.

Now, in order to study the behavior of solution near the equilibrium points, we need to compute the variational matrix of the fuzzy model system (2). Let $V(S, I, R)$ be the variational matrix of the fuzzy model system (2) at the point (S, I, R). Then

$$V = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

where

$$\begin{aligned}
a_{11} &= \beta(\sigma)I - \mu, \\
a_{12} &= -\beta(\sigma)S, \\
a_{13} &= 0, \\
a_{21} &= \beta(\sigma)I, \\
a_{22} &= \beta(\sigma)S - (\mu + \epsilon(\sigma) + \Upsilon(\sigma)), \\
a_{23} &= 0, \\
a_{31} &= 0, \\
a_{32} &= \Upsilon(\sigma), \\
a_{33} &= -\mu.
\end{aligned}$$

Thus for E_1, the eigenvalues are $-\mu$ and $\left(\frac{\beta(\sigma)\Lambda}{\mu} - (\mu + \epsilon(\sigma) + \Upsilon(\sigma))\right)$. Therefore, the equilibrium point E_1 is asymptotically stable provided $\beta(\sigma)\Lambda < \mu(\mu + \epsilon(\sigma) + \Upsilon(\sigma))$. Also, E_1 is a saddle point if $\beta(\sigma)\Lambda > \mu(\mu + \epsilon(\sigma) + \Upsilon(\sigma))$.

The stability study of the proposed fuzzy model shows that E_1 is unstable while E^* is asymptotically stable for $\beta(\sigma)\Lambda < \mu(\mu + \epsilon(\sigma) + \Upsilon(\sigma))$.

Now, consider σ^* such that $\beta(\sigma^*)\Lambda = \mu\left[\mu + \epsilon(\sigma^*) + \gamma(\sigma^*)\right]$ is the bifurcation value. Then

$$\sigma^* = \frac{\sigma_0\ \sigma_m\ \sigma_M \Lambda + \left(\mu^2 + \mu\ \epsilon_0 + \mu\right)\sigma_0\ \sigma_M\left(\sigma_0 - \sigma_m\right)}{\sigma_0\ \sigma_M \Lambda - \mu\left[\sigma_M(1-\eta-\epsilon_0) + (\gamma_0 - 1)\sigma_0\right]\left(\sigma_0 - \sigma_m\right)},$$

where $\sigma_m \leq \sigma^* \leq \sigma_0$ (Fig. 5).

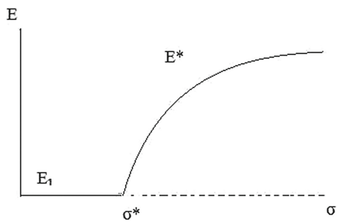


Fig. 5. Bifurcation diagram.

Theorem 3. *For the virus load $\sigma = \sigma^*$, the value of bifurcation parameter of the proposed fuzzy model* (2) *the model has only one unstable equilibrium E_1 if $\sigma < \sigma^*$ and it has an asymptotically equilibrium E^* if $\sigma > \sigma^*$.*

3.1 Fuzzy Basic Reproduction Number

The basic reproduction number $\mathcal{R}_0$ is obtained through the analysis of the stability of the trivial equilibrium point. For the classical SIR model, $\mathcal{R}_0 = \frac{\beta\Lambda}{\mu(\mu+\epsilon+\Upsilon)}$, and $\mathcal{R}_0(\sigma)$ is not a fuzzy set as it can be greater than 1. However, the maximum value of $\mathcal{R}_0(\sigma)$ is $\frac{\Lambda}{\mu(\mu+\epsilon(\sigma)+\Upsilon(\sigma))}$. Thus $\epsilon_0\mathcal{R}_0(\sigma) \leq 1$, whereby $\epsilon_0\mathcal{R}_0(\sigma)$ is a fuzzy set and hence $FEV\left[\epsilon_0\mathcal{R}_0(\sigma)\right]$ is well-defined. In view of this, we introduce the fuzzy basic reproduction number as under.

Definition 1. *The* **fuzzy basic reproduction number** *is given by*

$$\mathcal{R}_0^f = \frac{1}{\epsilon_0} FEV\left[\epsilon_0\mathcal{R}_0(\sigma)\right],$$

where $\mathcal{R}_0(\sigma) = \frac{\beta(\sigma)\Lambda}{\mu(\mu+\epsilon(\sigma)+\Upsilon(\sigma))}$

Now, from the Definition of fuzzy expected value,

$$FEV\left[\epsilon_0\mathcal{R}_0(\sigma)\right] = \sup_{0\leq\alpha\leq 1} \inf\left[\alpha, k(\alpha)\right], \tag{3}$$

where $k(\alpha) = \mu\left\{u : \epsilon_0\mathcal{R}_0(\sigma) \geq \alpha\right\} = \mu(X)$ is a fuzzy measure.
Further, to obtain $FEV\left[\epsilon_0\mathcal{R}_0(\sigma)\right]$, we have to define a fuzzy measure μ. For this purpose, the possibility measure is given by

$$\mu(X) = \sup_{\sigma\in X} \Gamma(\sigma), \qquad X \subset R. \tag{4}$$

From $FEV\,[\epsilon_0\mathcal{R}_0]$, it is clear that $\frac{\beta(\sigma)\Lambda}{\mu(\mu+\epsilon(\sigma)+\Upsilon(\sigma))}$ is not decreasing with σ, whereby the set X is an interval of the form $[\sigma', \sigma_M]$, where σ' is the solution of the following equation.

$$\epsilon_0 \frac{\beta(\sigma)\Lambda}{\mu\,(\mu+\epsilon(\sigma)+\Upsilon(\sigma))} = \alpha. \tag{5}$$

Thus

$$k(\alpha) = \mu\,[\sigma', \sigma_M] = \sup_{\sigma'\leq\sigma\leq\sigma_M} \Gamma(\sigma) \tag{6}$$

where, $k(0) = 1$ and $k(1) = \Gamma(\sigma_M)$.

This measure indicates that the infectivity of a group is the one presented by the individual belonging to the group with the maximal infectivity. Now, in order to determine $FEV\,[\epsilon_0 R_0]$ we assume that the amount of the virus load Σ of a group of individuals is a linguistic and is divided into three classes: "weak (Σ_-)", "medium (Σ_-^+)" and "strong (Σ^+)". Each classification is a fuzzy number based on the values σ_m, σ_0 and σ_M that come into view in the definition of β (Fig. 6).

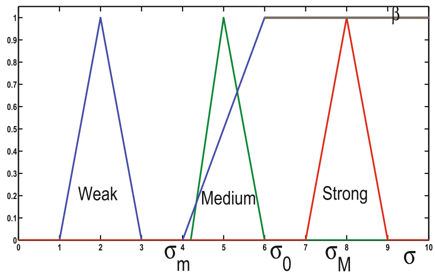


Fig. 6. Classification of linguistic variables.

Case (a) Weak virus load (Σ_-) is defined for $\hat{\sigma}+\delta < \sigma_m$. As $\hat{\sigma}+\delta < \sigma'$, we have FEV $[\epsilon_0\mathcal{R}_0] = 0 < \epsilon_0 \Leftrightarrow \mathcal{R}_0^f < 1$, which makes it possible to conclude that the disease will be extinct.

Case (b) Strong virus load (Σ^+) is defined for $\hat{\sigma}-\delta > \sigma_0$ and $\hat{\sigma}+\delta < \sigma_M$. Thus from (3), we have

$$k(\alpha) = \begin{cases} 1 & \text{if } 0 \leq \alpha < \epsilon_0 \frac{\Lambda}{\mu(\mu+\epsilon(\hat{\sigma})+\Upsilon(\hat{\sigma}))}, \\ \Gamma(\sigma') & \text{if } \epsilon_0 \frac{\Lambda}{\mu(\mu+\epsilon(\hat{\sigma})+\Upsilon(\hat{\sigma}))} \leq \alpha < \epsilon_0 \frac{\Lambda}{\mu(\mu+\epsilon(\hat{\sigma}+\delta)+\Upsilon(\hat{\sigma}+\delta))}, \\ 0 & \text{if } \epsilon_0 \frac{\Lambda}{\mu(\mu+\epsilon(\hat{\sigma}+\delta)+\Upsilon(\hat{\sigma}+\delta))} \leq \alpha \leq 1. \end{cases}$$

Obviously, if $\delta > 0$, k is continuous and decreasing function with $k(0) = 1$ and $k(1) = 0$. Hence $FEV\,[\epsilon_0 \mathcal{R}_0]$ is the fixed point of k and

$$\epsilon_0 \frac{\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))} < FEV\,[\epsilon_0 \mathcal{R}_0] < \epsilon_0 \frac{\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}+\delta) + \Upsilon(\hat{\sigma}+\delta))}.$$

or

$$\frac{\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))} < \mathcal{R}_0^f < \frac{\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))},$$

or that, $\mathcal{R}_0^f > 1$, which indicates that the disease will be endemic.
Case (c) Medium virus load Σ_-^+ is defined for $\hat{\sigma} - \delta > \sigma_m$ and $\hat{\sigma} + \delta < \sigma_0$. Therefore again from (3),

$$k(\alpha) = \begin{cases} 1 & \text{if } 0 < \alpha \le \epsilon_0 \frac{\beta(\hat{\sigma})\Lambda}{\mu(\mu+\epsilon(\hat{\sigma})+\Upsilon(\hat{\sigma}))}, \\ \Gamma(\sigma') & \text{if } \epsilon_0 \frac{\beta(\hat{\sigma})\Lambda}{\mu(\mu+\epsilon(\hat{\sigma})+\Upsilon(\hat{\sigma}))} < \alpha \le \epsilon_0 \frac{\beta(\hat{\sigma}+\delta)\Lambda}{\mu(\mu+\epsilon(\hat{\sigma}+\delta)+\Upsilon(\hat{\sigma}+\delta))}\ , \\ 0 & \text{if } \epsilon_0 \frac{\beta(\hat{\sigma}+\delta)\Lambda}{\mu(\mu+\epsilon(\hat{\sigma}+\delta)+\Upsilon(\hat{\sigma}+\delta))} < \alpha \le 1\ . \end{cases}$$

Similar to Case (b), we have

$$\frac{\beta(\hat{\sigma})\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))} < \mathcal{R}_0^f < \frac{\beta(\hat{\sigma}+\delta)\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}+\delta) + \Upsilon(\hat{\sigma}+\delta))}.$$

Thus in any case, we have

$$\epsilon_0 \frac{\beta(\hat{\sigma})\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))} < FEV\,[\epsilon_0 \mathcal{R}_0] < \epsilon_0 \frac{\beta(\hat{\sigma}+\delta)\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}+\delta) + \Upsilon(\hat{\sigma}+\delta))}$$

or

$$\frac{\beta(\hat{\sigma})\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}) + \Upsilon(\hat{\sigma}))} < \frac{FEV\,[\epsilon_0 \mathcal{R}_0]}{\epsilon_0} < \epsilon_0 \frac{\beta(\hat{\sigma}+\delta)\Lambda}{\mu\,(\mu + \epsilon(\hat{\sigma}+\delta) + \Upsilon(\hat{\sigma}+\delta))}$$

or

$$\mathcal{R}_0(\hat{\sigma}) < \mathcal{R}_0^f < \mathcal{R}_0(\hat{\sigma}+\delta).$$

As the function $\mathcal{R}_0(\sigma) = \frac{\beta(\sigma)\Lambda}{\mu(\mu+\epsilon(\sigma)+\Upsilon(\sigma))}$ is an increasing and continuous function. According to the Intermediate Value Theorem [10], there exists a unique $\bar{\sigma}$, with $\hat{\sigma} < \bar{\sigma} < (\hat{\sigma}+\delta)$ such that

$$\mathcal{R}_0^f = \mathcal{R}_0(\bar{\sigma}) > \mathcal{R}_0(\hat{\sigma}),$$

or that, there exists only one virus load $(\bar{\sigma})$ such that the basic reproduction number $\mathcal{R}_0$ and the fuzzy basic reproduction number $\mathcal{R}_0^f$ coincide. Furthermore, the average number of secondary cases $(\mathcal{R}_0^f)$ is higher than the number of secondary cases $\mathcal{R}_0(\hat{\sigma})$ due to the medium amount of infection.

3.2 Disease Control in Fuzzy Epidemic System

In this section, we analyze the control of the estimation of the disease in the population using the fuzzy basic reproduction number $\mathcal{R}_0^f = \mathcal{R}_0(\bar{\sigma})$. In the proposed fuzzy system, spread of disease depend not only on the variable σ but also on the transmission coefficient β, additional disease rate ϵ as well as the rate of recovery Υ. In the following, we describe some of the following cases about the existence and stability of the disease in the system. It is to be pointed out here that the proposed fuzzy system represents a family of systems depending on the parameter σ. In order to simplify these family of systems by a unique system of equations with the same outcome, our result shows that there is one value of σ, i.e., the bifurcation value σ^*.

(i) **Weak amount of infection:** In this case, $\bar{\sigma} < \hat{\sigma} + \delta \leq \sigma_m$, whereby the fuzzy basic reproduction number $\mathcal{R}_0^f$ is zero and the disease will be vanish in the population.

(ii) **Medium amount of infection:** In this case,
- if $\sigma^* > \sigma$, then fuzzy basic reproduction number $\mathcal{R}_0^f$ is less than the unity and the system will be free from disease; and
- if $\sigma^* < \sigma$, then fuzzy basic reproduction number $\mathcal{R}_0^f$ is greater than the unity and the system will become endemic in the population.

(iii) **Strong amount of infection**: In this case, $\bar{\sigma} > \hat{\sigma} > \hat{\sigma} + \delta \geq \sigma_0$, whereby the fuzzy basic reproduction number $\mathcal{R}_0^f(\sigma) > 1$ and the disease will invade. Now, the assumption of $\mathcal{R}_0^f$ is related to control policies to stop the spread of influenza;
  (1) $\mathcal{R}_0^f$ can be reduced by increasing σ_m (or, increasing σ^*).
  (2) Since, $\bar{\sigma} \in (\hat{\sigma}, \hat{\sigma} + \delta)$, if the amount of median virus load is very less then the value of R_0^f can reduce. For example, by using the medicine or separation of infected individuals (decreasing δ).

4 Numerical Simulation

In this paper we use an iterative method to solve the numerical simulation. For numerical simulation we consider the parameter value $\beta(\sigma) = 0.3$, $\Lambda = 1.2$, $\mu = 0.2$, $\gamma(\sigma) = 0.2$, $\epsilon(\sigma) = 0.1$.

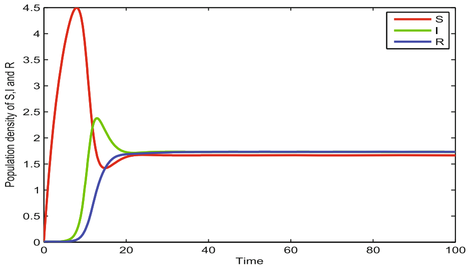


Fig. 7. Dynamical behavior of the system with $\beta(\sigma) = 0.3$, $\mu = 0.2$, $\epsilon(\sigma) = 0.1$.

5 Conclusion

In this paper we have studied theoretically an SIR epidemic model both in crisp and fuzzy system. Numerical methods are employed to solve the system and the behavior of the SIR models with respect to time are observed which is illustrate in Fig. 7. The stability of the system can also be observed from Fig. 7. Interestingly, the fuzzy basic reproduction number $\mathcal{R}_0^f$ is a function of disease spreading virus, while in case of crisp system basic reproduction number is not a direct function of virus. Thus it may be considered that fuzzy model is more flexible and balanced than the crisp system.

References

1. Barros, L.C., Bassanezi, R.C.: About fuzzy dynamical systems: theory and applications, Ph.D. thesis, University of Campinas, Săo Paulo, brazil (1997)
2. Barros, L.C., Bassanezi, R.C., Leite, M.B.F.: The SI epidemiological models with a fuzzy transmission parameter. Comput. Math. Appl. **45**, 1619–1628 (2003)
3. Bassanezi, R.C., Barros, L.C.: A simple model of life expectancy with subjective parameters. Kybernets **24**, 91–98 (1995)
4. Hu, Z., Liu, S., Wang, H.: Backward bifurcation of an epidemic model with standard incidence rate and treatment rate. Nonlinear Anal.: Real World Appl. **9**, 2302–2312 (2008)
5. Li, X., Li, W., Ghosh, M.: Stability and bifurcation of an SIR epidemic model with nonlinear incidence and treatment. Appl. Math. Comput. **210**, 141–150 (2009)
6. Massad, E., Ortega, N.R.S., Barros, L.C., Struchiner, C.J.: Fuzzy Logic in Action: Applications and Epidemiology and Beyond. In: Studied in Fuzziness and Soft Computing, vol. 232 (2008)
7. Mishra, B.K., Pandey, S.K.: Fuzzy epidemic model for the transmission of worms in computer network. Nonlinear Anal.: Real World Appl. **11**, 4335–4341 (2010)
8. Mondal, P.K., Jana, S., Haldar, P., Kar, T.K.: Dynamical behavior of an epidemic model in a fuzzy transmission. Int. J. Uncertainty, Fuzziness Knowl.-Based Syst. **23**, 651–665 (2015)
9. Pal, D., Mahaptra, G.S., Samanta, G.P.: Optimal harvesting of prey-predator system with interval biological parameters: a bioeconomic model. Math. Biosci. **241**, 181–187 (2013)
10. Stewart, J.: Calculus-Early Transcendentals. Books/Cole publishing Company, USA (1999)
11. Upadhyay, R.K., Roy, P., Rai, V.: Deciphering dynamics of epidemic spread: the case of influenza virus. Int. J. Bifurcat. Chaos. **24**, 145–164 (2014)
12. Verma, R., Tiwari, S.P., Upadhyay, R.K.: Fuzzy epidemic model for the spread of influenza virus and its possible control. Iranion J. Fuzzy Syst. (Communicated)
13. Wang, W.: Backward bifurcation of an epidemic model with treatment. Math. Biosci. **201**, 58–71 (2006)
14. Zhang, X., Liu, X.: Backward bifurcation and global dynamics of an SIS epidemic model with general incidence rate and treatment. Nonlinear Anal.: Real World Appl. **10**, 565–575 (2009)

Optimal Parameter Ranges in Fuzzy Inference Systems, Applied to Spatial Data

Jörg Verstraete[1,2(✉)] and Weronika Radziszewska[1]

[1] Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warszawa, Poland
{jorg.verstraete,weronika.radziszewska}@ibspan.waw.pl
[2] Centro Singular de Investigación en Tecnoloxías da Información, University of Santiago de Compostela, Rúa de Jenaro de la Fuente Domínguez, 15782 Santiago de Compostela, Spain
jorg.verstraete@usc.es

Abstract. Processing of spatial data can benefit from the use of fuzzy inference systems, and such systems have been proposed to deal with the map overlay problem for gridded data. The development of fuzzy inference system for solving spatial problems poses specific challenges due to the type of data and specific properties of the spatial context. In this contribution, we take into account that a spatial dataset can exhibit a big variety in different areas and determine the most possible ranges for the variables in the rulebase system in a more appropriate and dynamic way. In addition, we show how the construction and application of a rulebase can be modified in order to handle this changed definition of the most possible ranges.

Keywords: Fuzzy inference · Gridded data · Parameter ranges

1 Introduction

The use of spatial data plays a big role in many fields of research. Spatial data is a broad term that covers all data that carries a spatial aspect; this usually relates to real world locations using coordinates and appropriate projections. Examples of such data are in fields such as environmental research and climatic research, referring to data on e.g. concentration of different pollutants, temperature or humidity. There are different representation models for spatial data [6,7]; the two main models are feature based models and raster based models. Feature based models allow for the representation of real world entities, such as locations, roads or areas using basic geometric structures. However, for representing e.g. a numerical property that has a different value on different locations, a different model is required. Such a property theoretically has a precise value on every location but as this is impossible to represent, the models used for this therefor need to approximate the real situation. Most commonly used are raster-based

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_46

models, also known as grid-based models. A raster or grid is overlayed with the region of interest and partitions it in a number of cells. These cells tend to be chosen to be regular, usually square or rectangular, but use of irregular shapes is possible (e.g. partitioning a country in its provinces). With every cell, a value that is considered representative is associated: for a property such as temperature, this will be a weighted average; for a property such as population, this will be a sum. The size of the cells determines the resolution: a grid consisting of cells that represent areas of $1\,\text{km} \times 1\,\text{km}$ has a much higher resolution than a grid with cells that represent $5\,\text{km} \times 5\,\text{km}$ and thus can capture much more detail regarding the spatial distribution of the property.

In many analysis, it is necessary to combine data stemming from different sources: in epidemiology for example, temperature can be combined with humidity and salinity to identify areas where certain bacteria can thrive [4]; air pollution data can be combined with population data to determine the exposure of people to pollutants [2,3]. This requires overlaying the different datasets, an operation which is called a map overlay; for current algorithms to combine spatial data, we refer to [1,5]. Geographic Information Systems (GIS) help the users to do this, by already correctly georeferencing the data. However, overlaying different datasets poses several challenges, particular for gridded data. A method using artificial intelligence, in particular a rulebase system, to perform the map overlay on gridded data was developed; this concept was presented in [9], the resulting algorithm in [10]. In this contribution, attention will go to optimizing the definitions of the antecedents of this rulebase system, by allowing the variable spaces to be different for each datapair in both training set and data set.

The next section explains the map overlay problem for grids in more detail; Sect. 3 details how the map overlay problem is translated into a rulebase. The methodology to define antecedents is elaborated on in Sect. 4, along with its impact on rulebase construction and application. The article finishes with concluding remarks in Sect. 5.

2 Map Overlay Problem for Grids

A raster or grid allows for the modelling of a numeric property by partitioning the region of interest into a number of cells, thus discretizing the two-dimensional space. The defintion of the raster used is dependent on many factors, ranging from the way data are gathered, the way models generate the data, the representation deemed most practical for typical applications or even historical factors. When performing a map overlay, researchers intend to combine data from different sources, and as such the representation of this data may differ. In the case of grids, this can result in incompatible grids: grids where there is not a one-to-one mapping between the grid cells, as shown on Fig. 1. This can be as one grid is shifted compared to the other, or because it uses a different cell size; it could be rotated, or any combination of those. Note that this is not a problem of coordinate transformation: the grids are already georeferenced and located in the correct spot; they are just defined differently. In order to perform a map

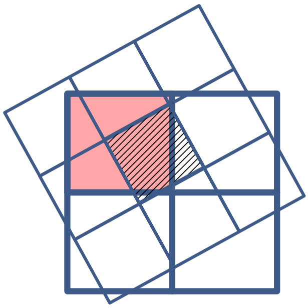


Fig. 1. Illustration of the map overlay problem for ill-aligned grids: how to relate the hashed area with the shaded area?

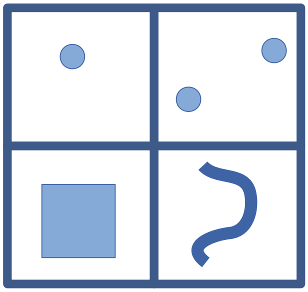


Fig. 2. Examples of underlying distributions that can result in the same value for a grid cell: if the value associated with the features in each of the above cells is the same, the value of the grid cell will be the same.

overlay of grids, the common approach is to remap one grid onto the other (this is called regridding), in order to achieve a one-one mapping of the cells. This is a non-trivial step, as the underlying distribution of the data in a grid is not known: at the level of the grid, the cell is the smallest unit. Whether the data inside the grid is in reality localized in one area (e.g. in the case of pollutions sources: a single factory), along lines (e.g. for pollution sources: roads) or uniform over the region is not known. This is illustrated on Fig. 2, while the map overlay problem for gridded data is illustrated on Fig. 1.

As such, assumptions are made to perform this remapping. Most commonly used approaches assume either a uniform distribution in each grid cell, resulting in areal weighting; a smooth distribution over the region of interest, requiring spatial smoothing; or a spatial statistical correlation [1,8]. In [9], we proposed solving the map overlay problem using proxy data in combination with a fuzzy inference system. Proxy data is known data that has a spatial correlation with the data of the grid that is remapped onto the other grid: the proxy data holds additional information on the spatial distribution and can help to improve the regridding process. The algorithm was presented in [10]. In the next section, the translation of the problem to a rulebase system will be considered.

3 From Map Overlay to Rulebase

3.1 Parameters

As mentioned in [9], the regridding problem can be considered a problem of spatial disaggregation onto on irregular grid whose cells partition the cells of the input (Fig. 3). As such, we consider here the problem of spatial disaggregation of an input grid A onto an output grid B.

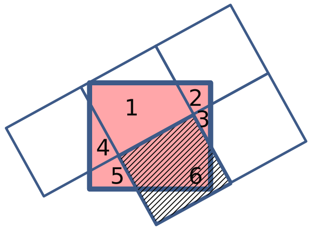


Fig. 3. Remapping the data from the shaded area can be done by performing a spatial disaggregation into the numbered segments. The appropriate cells can later be combined to form the output gridcells (e.g. the hashed cell).

The cells of a grid X will be denoted x_i, with associated values $f(x_i)$. For the discussion, it is assumed that we dissagregate a grid A representing data that is cumulative: redistributing the data of a cell a_i means dividing it in portions that sum up to the value $f(a_i)$. In the rulebase approach, proxy data C is used to help provide information on the underlying distribution; C is assumed to also be a gridded dataset, but the approach can be modified to work with a feature-based datasets as proxy data.

The disaggregation problem of A into B basically means finding values w_{b_i} such that

$$\forall b_i \in B, \exists a_j \in A, w_{b_i} \in [0,1] : f(b_i) = w_{b_i} \times f(a_j)$$
$$\forall a_j \in A : f(a_j) = \sum_{b_i \in B | b_i \cap a_j \neq \emptyset} w_{b_i} \times f(b_i)$$

The second condition is equivalent to

$$\forall a_j \in A : 1 = \sum_{b_i \in B | b_i \cap a_j \neq \emptyset} w_{b_i}$$

The rulebase approach determines these values w_i by considering proxy data, translated into antecedents.

IF (property of C relating to w_{b_i}) IS high AND ...
    THEN w_{b_i} IS high
IF (property of C relating to w_{b_i}) IS low AND ...
    THEN w_{b_i} IS low

This rulebase will be executed for every output cell b_i, so for every weight w_{b_i} that needs to be determined. The first task here is to define the parameters of the rulebase system; these will match a *property of* C *relating to* w_{b_i}. Note that the grid C does not have to match grid A or grid B, and it does not have to be a partitioning of either of those. To define a property that relates C to the underlying distribution of A, it is possible to use the definition of the geometries of the cells. If C for example holds similar data as A, but e.g. from a different source, the amount of overlap of the cells of C, combined with the value of those cells, can serve as an indication for the distribution of the data. Other examples are possible; using overlap, distance, topology or other operations and combinations. In this article, we consider the aforementioned example of overlap and will denote the property that relates to b_i as $g(b_i)$:

$$x^i = g(b_i) = \sum_{c_j \in C} \frac{S(c_j \cap b_i)}{S(c_j)} \times f(c_j) \tag{1}$$

where $S()$ is the notation of the surface area, $c_j \cap b_i$ is the notation for the geometry created by the intersection of the geometries of c_j and b_i and $f(c_j)$ is the notation of the value associated with c_j. Once a parameter for use in the antecedent is defined, the construction of the rulebase using training dataset can be started.

3.2 Rules

Construction of a rulebase from examples has been described in [11], and starts from a partitioning of the input spaces and output space. The examples are taken from a training set; this is a set that has known values for input as well as output parameters. The input parameter in our case study is the value-weighted overlap of cells of C with the target cell b_i. Typically, and also in our case study, it is therefor first necessary to find the domain of most possible values. An obvious choice for this is to calculate the value for all datapairs in the training set, and use the minimum and maximum of these to define the range, Fig. 4 illustrates which cells are used to calculate the value of the hashed cell. The training set will consist of input grid, output grid and proxy grid, denoted A', B' and C', with elements respectively a'_j, b'_i and c'_k.

$$x^- = \min_{b'_i \in B'} g(b'_i) \tag{2}$$

$$x^+ = \max_{b'_i \in B'} g(b'_i) \tag{3}$$

The linguistic terms high, low and medium will be defined over the interval $[x^-, x^+]$ as shown on Fig. 5. The output parameters y_i are the weights, and by definition they are in the range $[0, 1]$. The construction of the rulebase from datapairs is done by constructing a rule for each datapair (x_i, y_i) yielding the rule:

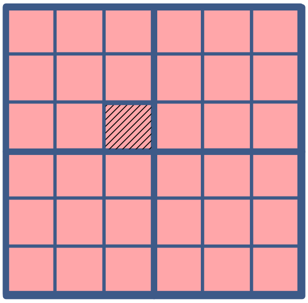


Fig. 4. The shaded area indicates all the values that will be considered for determining the minimum and maximum possible range for the hashed grid cell.

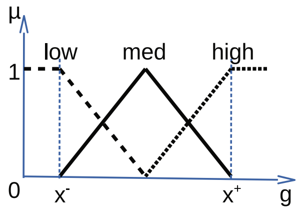


Fig. 5. Example of a natural partitioning to define the fuzzy sets for the linguistic terms for low, medium and high in the range $[x^-, x^+]$.

IF x IS L_j^x THEN y IS L_k^y

Here, x and y are the names of the variables for which x_i and y_i are the values and L_j^x, respectively L_k^y the best matching linguistic term defined on the domain of the variable. The rule has a weight which is assigned the lowest membership grade of any of the values in the antecedent or consequent of the rule for its considered membership term. In the last stage of the rulebase construction, duplicate rules are removed and only those with the highest weights are kept.

3.3 Spatial Aspects and Impact

Consider the dataset on Fig. 6. The datapairs are of the form (x_i, y_i), with x_i as defined above and y_i as the weight. In particular, consider the datapairs that stem from cell a_1 which disaggregated into two cells b_{1_1} and b_{1_2}. If the value of C is assumed to be 1, then the calculated value $g(b_i)$ reverts to the amount of overlap; the overlap with C is graphically illustrated by the shaded area. The training dataset has values for $x_{b_{1_1}}$, $x_{b_{1_2}}$ and for the output $y_{b_{1_1}}$ and $y_{b_{1_2}}$ (the output values are not illustrated on the figure and are arbitarily chosen for the example):

$$(x_{b_{1_1}}, y_{b_{1_1}}) = (0.1, 0.2)$$
$$(x_{b_{1_2}}, y_{b_{1_2}}) = (0.2, 0.8)$$

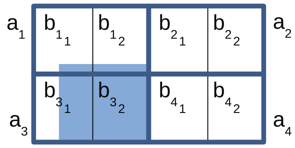


Fig. 6. Example used to explain the impact of the range selection.

The cell b_{1_1} just slightly overlaps with C, as indicated by the shaded area on Fig. 6, resulting in $x_{b_{1_1}} = 0.1$ and $x_{b_{1_2}} = 0.2$. Both of these evaluate to LOW, using the linguistic terms defined on Fig. 5 and the values x^- and x^+ defined as in Eqs. 2 and 3. At the same time, the output values are $y_{b_{i_1}} = 0.2$ and $y_{b_{i_2}} = 0.8$, as this is how the data is distributed; these evaluate to LOW and HIGH. Following the algorithm, these two datapairs would yield the rules:

```
IF x IS LOW y IS LOW
IF x IS LOW y IS HIGH
```

The differences between the evaluation of x will cause the rules to have different weights, and this in favour of the second rule, which is the more consistent one considering the connection between grids C and A.

Similarly, examples can be found that will add the rules

```
IF x IS LOW y IS LOW
IF x IS HIGH y IS HIGH
```

and

```
IF x IS HIGH y IS LOW
IF x IS HIGH y IS HIGH
```

This means that the rulebase constructed will have four rules that are not distinguishing enough, causing difficulties in differentiating values and resulting in outcomes that have a high uncertainty. Even applying the same datapairs as input to this rulebase results in evaluations that make the output value match with both HIGH and LOW, causing the system to have to consider all results and after aggregation yielding a solution that has higher uncertainty.

The origin of the problem is the fact that the range of most possible value was chosen too wide. For the datapair in the first example, it is impossible that either $x_{b_{1_1}}$ or $x_{b_{1_2}}$ are greater than 0.3, as the total weighted overlap of a_1 with C is 0.3. The biggest possible value is 1, which occurs in cell b_{3_2}. The choice of the range $[0, 1]$ is not suitable for this datapair and causes the construction of rules that are less effective at solving the problem.

4 Customizing the Most Possible Ranges

4.1 Local Most Possible Range

If the grids are examined closer, grid C is used to help steer the underlying spatial distribution of the data. As such, higher values in C, indicate high values in A but the actual definition of *high* should not be so strictly interpreted: the connection between C and A is only required locally, and should be considered relative for all cells of B that form the partitioning of the same cell in A. High as such would have to refer to "the highest of those cells involved". In order to solve this problem in the rulebase creation and application, it is necessary to change the definition of the most possible ranges. Considering the original spatial datasets, the issue arises as the minimum possible range was chosen as the lowest possible value that occurs, while the maximum was the highest possible value that occurs. While this is true for the entire grid, the spatial dataset can exhibit quite a different local behaviour. As such we can consider first defining the most possible range individually for a given datapair, by limiting the area from which this minimum and maximum are considered. For a given output cell b_i, one could consider all the cells that are within a given distance, or even more individual, all the cells that are covered by the same input cell a_j. Both these options are illustrated on Fig. 7.

$$x^-_{b'_i} = \min_{b'_j \in B' | b'_j \cap a'_k \in A' \neq \emptyset} g(b'_i) \quad \text{WHERE} \quad b'_i \cap a'_k \neq \emptyset$$
$$x^+_{b'_i} = \max_{b'_j \in B' | b'_j \cap a'_k \in A' \neq \emptyset} g(b'_i) \quad \text{WHERE} \quad b'_i \cap a'_k \neq \emptyset$$

As a result, a given datapair (x^{b_i}, y^{b_i}) will have its own range for x^{b_i}: $[x^-_{b_i}, x^+_{b_i}]$. The range for the output variable y^{b_i} in this example can be shared between all datapairs, and remains $[0, 1]$.

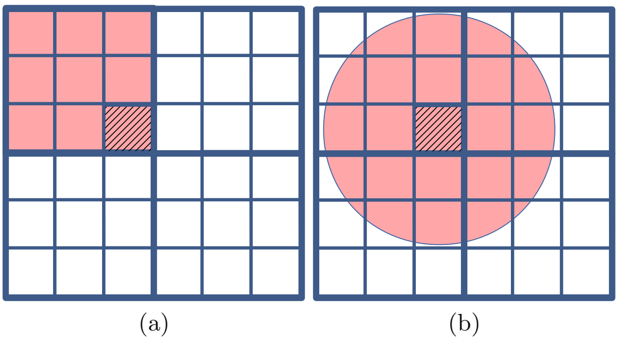


Fig. 7. Two examples to illustrate the cells considered for determining the minimum and maximum possible range for the hashed cell. These cells are indicated by the shaded area: (a) using those cells that overlap the same input, (b) using the cells that overlap a given neighbourhood.

In order to apply the rulebase with the algorithm in [11], all datapairs need to have the same most possible range for a variable. It is possible to define a common range, and rescale every value of x^{b_i} to match this common range. As an arbitrary choice, the range $[0, 1]$ was selected; the scaled function that will be used determine values as input for the rulebase is $g'(b_i)$ and is defined as:

$$g'(b_i) = \frac{x^{b_i} - x^-_{b_i}}{x^+_{b_i} - x^-_{b_i}} \tag{4}$$

Using this definition, the rulebase construction method can be applied as in the previous section. A special case can be added for the situation where $x^-_{b_i} = x^+_{b_i}$, but this is trivial. The same example on Fig. 6, the two datapairs $(x_{b_{1_1}}, y_{b_{1_1}}) = (0.1, 0.2)$ and $(x_{b_{1_2}}, y_{b_{1_2}}) = (0.2, 0.8)$ will now first have to be rescaled. This yields:

$$g'(b_{1_1}) = \frac{x^{b_{1_1}} - x^-_{b_{1_1}}}{x^+_{b_{1_1}} - x^-_{b_{1_1}}} = \frac{0.1 - 0.1}{0.2 - 0.1} = 0$$

$$g'(b_{1_2}) = \frac{x^{b_{1_2}} - x^-_{b_{1_2}}}{x^+_{b_{1_2}} - x^-_{b_{1_2}}} = \frac{0.2 - 0.1}{0.2 - 0.1} = 1$$

In an example where there would be more than two cells partitioning the input cell a_j, other cells can have values that differ from 0 or 1. This transformation of the input variable results in two new datapairs that will be used for the rulebase construction: $(0, 0.2)$ and $(1, 0.8)$. These datapairs will evaluate respectively to (LOW, LOW) and (HIGH, HIGH), yielding the rules:

```
IF x IS LOW y IS LOW
IF x IS HIGH y IS HIGH
```

While it still is possible for different rules to be constructed, such rules will stem from datapairs that are inconsistent with the general assumption on the relation between A and C; rather than from side-effects caused by the definitions used.

4.2 Estimated Most Possible Range

The above approach for determining the range is defined by means of the calculated values for neighbouring cells. However, given that the most possible range is a property which is connected to the output cell for which the value is calculated, a most possible range can be calculated rather than derived from the parameter values of neighbouring cells. This is the most difficult approach, as the calculated range has to be defined in an appropriate way for all the output cells. For the ongoing example of the value-weighted overlap, a suitable range will be proposed.

Consider the an output grid B and a proxy grid C. The spatial distribution of gridded data is not known, and therefor this is also the case for the data in grid C. The different options for the spatial distribution of the data in this grid can therefor help to determine the limits. The value-weighted overlap, which is used as the value in the rulebase, matches the assumption that data in the gridcells of C is uniformly distributed. However, a different distribution, as illustrated on Fig. 8a can be considered: the underlying spatial distribution of grid C is such that it lies outside of the cell b_i that is considered (shaded cell on Fig. 8a). This is only possible for cells c_k that have a partial overlap with b_i; what remains is the value-weighted overlap of the cells of C that are fully contained (hashed cell on Fig. 8a). Such a distribution effectively results in the lowest possible value for the parameter for b_i.

$$x^-_{b'_i} = \sum_{c_j \in C | c_j \subset b_i} \frac{S(c_j \cap b_i)}{S(c_j)} \times f(c_j)$$
$$= \sum_{c_j \in C | c_j \subset b_i} f(c_j)$$

Similarly, the spatial distribution of the data in the cells of C could be such that it maximizes the amount of data mapped in the considered cell b_i. This is illustrated on Fig. 8b. The underlying spatial distribution of the partially overlapping cells is in this case such that the data is fully contained in the considered cell b_j; the maximum therefore is the sum of all overlapping cells (all nine cells that overlap with the shaded cell on Fig. 8b).

$$x^+_{b'_i} = \sum_{c_j \in C | c_j \cap b_i \neq \emptyset} f(c_j)$$

This range the benefit that it effectively evaluates the value (value-weighted overlap) against the realistically possible values. On the other hand, the property has its limits: when there are no partially overlapping cells between b_j and C, the value will evaluate to the maximum. Similarly, when the cells of C are very big compared to those of B, the range may be skewed too much towards high values, resulting in the evaluation of the value of b_j to be rather low.

If we consider as an example that all cells of the proxy grid in Fig. 8 have an associated value 1, then for the shown shaded output cell b'_i, the most possible range would be defined by the values:

$$x^-_{b'_i} = \sum_{c_j \in C | c_j \subset b_i} f(c_j) = 1$$
$$x^+_{b'_i} = \sum_{c_j \in C | c_j \cap b_i \neq \emptyset} f(c_j) = 9$$

Estimated ranges are more difficult to construct than local parameters, and are more prone to unforeseen side-effects. However, a properly defined estimated range can capture an expert's interpretation on how grids relate to one another in much more advanced ways. It is difficult to illustrate the estimated ranges on a simple example as the one on Fig. 6; for this small example it would perform similar as to the global range, but this does not illustrate the benefits.

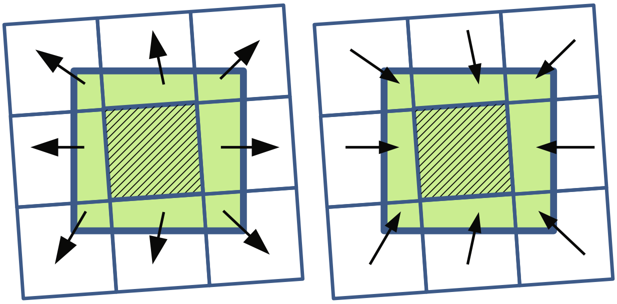


Fig. 8. Interpretation of the estimated minimum and maximum, as respectively contained cells (a) and overlapping cells (b).

Using estimated ranges to construct and evaluate a rulebase is similar to using local ranges and require the rescaling of the value as mentioned in Sect. 4.1. The assessment whether or not the rules generated from the examples are consistent with the assumed connection with the proxy data is more difficult than when using local ranges.

5 Conclusion

In this article, we presented a way of using variable most-possible ranges in both the construction and application of fuzzy rulebase systems. The use of variable most-possible ranges allow for the fuzzy rulebase to still differentiate between values when there are *local* differences between the datapair, an aspect that becomes very tangible and literal when dealing with spatial data. In addition to showing how such local ranges can be used in fuzzy rulebase systems, we also constructed two categories of variable most-possible ranges. The first is similar in concept to the typical partitioning of the input space of the variable, but limiting it to a subset of the datapairs. The second effectively computes a most-possible range for the given variable, using specific algorithms designed for the application. Small examples were used to illustrate the impact in the construction of the rulebase.

Acknowledgments. This work has received financial support from the Consellera de Cultura, Educacin e Ordenación Universitaria (accreditation 2016–2019, ED431G/08) and the European Regional Development Fund (ERDF).

References

1. Gotway, C.A., Young, L.J.: Combining incompatible spatial data. J. Am. Stat. Assoc. **97**(458), 632–648 (2002)
2. Horabik, J., Nahorski, Z.: Improving resolution of a spatial air pollution inventory with a statistical inference approach. Climatic Change **124**, 575–589 (2014)

3. Hryniewicz, O., Nahorski, Z., Verstraete, J., Horabik, J., Jonas, M.: Compliance for uncertain inventories via probabilistic/fuzzy comparison of alternatives. Climatic Change **124**(3), 519–534 (2014)
4. Martinez-Urtaza, J., Bowers, J.C., Trinanes, J., DePaola, A.: Climate anomalies and the increasing risk of vibrio parahaemolyticus and vibrio vulnificus illnesses. Food Res. Int. **43**(7), 1780–1790 (2010)
5. Mugglin, A.S., Carlin, B.P., Gelfand, A.E.: Fully model-based approaches for spatially misaligned data. J. Am. Stat. Assoc. **95**(451), 877–887 (2000)
6. Rigaux, P., Scholl, M., Voisard, A.: Spatial Databases with Applications to GIS. Morgan Kaufman Publishers, San Francisco (2002)
7. Shekhar, S., Chawla, S.: Spatial Databases: A Tour. Pearson Educations, Upper Saddle River (2003)
8. Tobler, W.R.: Smooth pycnophylactic interpolation for geographic regions. J. Am. Stat. Assoc. **74**(367), 519–536 (1979)
9. Verstraete, J.: Solving the map overlay problem with a fuzzy approach. Climatic Change **124**, 591–604 (2014)
10. Verstraete, J.: The spatial disaggregation problem: simulating reasoning using a fuzzy inference system. IEEE Trans. Fuzzy Syst. **PP**(99), 1 (2016)
11. Wang, L.X., Mendel, J.M.: Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern. **22**(6), 1414–1427 (1992)

On Finite-Valued Bimodal Logics with an Application to Reasoning About Preferences

Amanda Vidal[1(✉)], Francesc Esteva[2], and Lluis Godo[2]

[1] Institute of Computer Science (ICS - CAS), Pod vodárenskou věží 2, 182 07 Prague, Czech Republic
amanda@cs.cas.cz

[2] Artificial Intelligence Research Institute (IIIA - CSIC), Campus de la UAB s/n, 08193 Bellaterra, Spain
{esteva,godo}@iiia.csic.es

Abstract. In a previous paper by Bou et al., the minimal modal logic over a finite residuated lattice with a necessity operator $\Box$ was characterized under different semantics. In the general context of a residuated lattice, the residual negation $\neg$ is not necessarily involutive, and hence a corresponding possibility operator cannot be introduced by duality. In the first part of this paper we address the problem of extending such a minimal modal logic with a suitable possibility operator $\Diamond$. In the second part of the paper, we introduce suitable axiomatic extensions of the resulting bimodal logic and define a logic to reason about fuzzy preferences, generalising to the many-valued case a basic preference modal logic considered by van Benthem et al.

Keywords: Many-valued modal logic · Necessity and possibility modal operators · Finite residuated lattice · Reasoning about graded preferences

1 Introduction

Theoretical studies of fuzzy or many-valued modal logics have attracted an increasing attention in the last years, both following general and foundational approaches e.g. in [4,9,14,15], as well as focusing on particular families of fuzzy logics, mainly those based on Gödel logic [6–8,10,11], Łukasiewicz logic [3,12] or more recently on Product logic [17].

In particular, in [4] the authors study in depth minimal modal logics with a necessity operator $\Box$ (and canonical truth-constants) over a finite residuated lattice, considering different classes of many-valued Kripke frames and getting complete axiomatizations with respect to them.

In the first part of this paper, Sect. 2, we address the problem of extending those minimal modal logics with a possibility modal operator $\Diamond$. Note that in the

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_47

general context of a residuated lattice, if the residual negation $\neg$ is not involutive, then $\Box$ and $\Diamond$ are not dual in the usual sense ($\Diamond$ is not definable as $\neg\Box\neg$).

In the second part, in Sect. 3 we define suitable axiomatic extensions of the above fuzzy bi-modal logics, and then in Sect. 4 we define a logic to reason about fuzzy preferences, generalising to the many-valued case one of the preference modal logics considered by van Benthem et al. in [1].

2 The Minimal Bimodal Logic of a Finite Residuated Lattice

We start from basic definitions in [4], with which we assume the reader certain familiarity. Through the following sections, we will be assuming $\mathbf{A} = (A, \wedge, \vee, \odot, \rightarrow, 0, 1)$ to denote a *finite* (bounded, integral, commutative) residuated lattice, and we will consider its canonical expansion $\mathbf{A^c}$ by adding a new constant $\overline{a}$ for every element $a \in A$ (canonical in the sense that the interpretation of $\overline{a}$ in $\mathbf{A^c}$ is a itself.) The logic associated with $\mathbf{A^c}$ will be denoted by $\boldsymbol{\Lambda}(\mathbf{A^c})$, and its logical consequence relation $\models_{\mathbf{A^c}}$ is defined as follows: for all sets $\Gamma \cup \{\varphi\} \subseteq \mathbf{Fm}$ of formulas built in the usual way from a set of propositional variables $\mathcal{V}$ in the language of residuated lattices (possibly including constants from $\{\overline{a} : a \in A\}$),

$$\Gamma \models_{\mathbf{A^c}} \phi \quad \iff \quad \forall h \in \mathrm{Hom}(\mathbf{Fm}, \mathbf{A}^c),\ \text{if } h[\Gamma] \subseteq \{1\} \text{ then } h(\phi) = 1, \qquad (1)$$

where $\mathrm{Hom}(\mathbf{Fm}, \mathbf{A}^c)$ denotes the set of evaluations of formulas on $\mathbf{A^c}$.

In order to introduce the minimum bimodal logic over $\mathbf{A^c}$, let us consider the modal language **MFm** being the expansion of **Fm** with two modal operators $\Box$ and $\Diamond$. Kripke-style semantics for the bimodal logic is defined as follows.

Definition 1. *An* **A**-Kripke model *is a triple* $\mathfrak{M} = \langle W, R, e\rangle$ *where*

- W *is a set of worlds,*
- $R\colon W \times W \rightarrow A$*, is an* A*-valued accessibility relation between worlds,*
- $e\colon W \times \mathcal{V} \rightarrow A$ *is the evaluation of the model, and it is uniquely extended to formulas by letting* $e(w, \overline{a}) = a$ *for every* $a \in A$*,* $e(w, \varphi\star\psi) := e(w, \varphi)\star e(w, \psi)$ *for any propositional connective* $\star$ *in the language,*[1] *and*

$$e(w, \Box\varphi) := \bigwedge_{w\in W} \{R(v, w) \rightarrow e(w, \varphi)\}, \quad e(w, \Diamond\varphi) := \bigvee_{w\in W} \{R(v, w)\&e(w, \varphi)\}.$$

We let $\mathbb{BM}_{\mathbf{A}}$ *denote the class of all* **A***-Kripke models.*

Observe that the above values are always well-defined because the lattice is finite.

We say that, in a Kripke model $\mathfrak{M}$, a formula φ follows from a set of premises Γ, and write $\Gamma \Vdash_{\mathfrak{M}} \varphi$, whenever for any $v \in W$ such that $e(v, \gamma) = 1$

[1] For the sake of clarity, we use the same symbol (e.g. $\odot, \rightarrow$) both as syntactic connective in the language **MFm** and as the corresponding algebraic operation.

for all $\gamma \in \Gamma$, it holds that $e(v, \varphi) = 1$ too. Whenever we have a class of models $\mathbb{C}$, we will write $\Gamma \Vdash_{\mathbb{C}} \varphi$ meaning that φ follows from Γ in all the models of the class.[2]

As usual, in any deductive system used along this article (including the ones defined in the above lines), we will omit writing $\emptyset$ whenever the set of premises is empty, and simply write $\vdash \varphi$.

2.1 Axiomatization

In this section we present a modal axiomatic system over a finite residuated lattice $\mathbf{A}$, as an extension of the axiomatic modal system presented in [4, Definition 4.6], called $\boldsymbol{\Lambda}(\mathsf{Fr}, \mathbf{A^c})$, and that will be shown to be complete with respect to $\Vdash_{\mathbb{BM}_{\mathbf{A}}}$ defined above.

Before proceeding to the definition of the axiomatic system, observe that a sort of symmetric version of the axiom (Ax_a) in [4, Propsition 3.10] is valid in all $\mathbf{A}$-Kripke models. Namely, for every $a \in A$,

$$\Vdash_{\mathbb{BM}_{\mathbf{A}}} \Box(\varphi \to \overline{a}) \leftrightarrow (\Diamond\varphi \to \overline{a}).$$

This follows immediately from the fact that in any residuated lattice $\mathbf{A}$, for any $X \cup \{a\} \subseteq A$, it holds that $\bigwedge_{x \in X} \{x \to a\} = \bigvee_{x \in X} \{x\} \to a$ whenever the corresponding inf. and sup. exist (which is our case since the algebra is finite).

It looks then natural to consider that formula as a member of the axiomatic system, and as we prove below, this one is indeed the only formula referring to $\Diamond$ that we need to consider in order to get a complete axiomatic system for $\Vdash_{\mathbb{BM}_{\mathbf{A}}}$.

Definition 2. *Let* $\mathsf{BM}_{\mathbf{A}}$ *be the deductive system given by:*

1. *The axiomatic basis of* $\boldsymbol{\Lambda}(\mathsf{Fr}, \mathbf{A^c})$*, i.e.:*
   - *an axiomatic basis for* $\boldsymbol{\Lambda}(\mathbf{A^c})$
   - *modal axioms for* $\Box$*:*
     $\Box\overline{1}$, $\quad(\mathsf{MD})\quad (\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$, $\quad(\mathsf{Ax}_a)\quad \Box(\overline{a} \to \varphi) \leftrightarrow (\overline{a} \to \Box\varphi)$
2. *The axiom schemata*
   $(\Box\Diamond_a)\quad \Box(\varphi \to \overline{a}) \leftrightarrow (\Diamond\varphi \to \overline{a})$*, for each* $a \in A$
3. *The rules of the basis for* $\boldsymbol{\Lambda}(\mathbf{A^c})$ *and the Monotonicity rule:*
   $(\mathsf{Mon})\quad$ *from* $\varphi \to \psi$ *derive* $\Box\varphi \to \Box\psi$

We will denote by $\vdash_{\mathsf{BM}_{\mathbf{A}}}$ the corresponding notion of proof, and by $Th(\mathsf{BM}_{\mathbf{A}})$ the set of theorems of the logic $\mathsf{BM}_{\mathbf{A}}$, i.e. $Th(\mathsf{BM}_{\mathbf{A}}) = \{\varphi \in \mathbf{MFm} : \vdash_{\mathsf{BM}_{\mathbf{A}}} \varphi\}$.

In order to prove completeness of the previous logic with respect to the relation $\Vdash_{\mathbb{BM}_{\mathbf{A}}}$ we will resort to the usual canonical model construction. However, we need to define a canonical model different from the one used

[2] This logical consequence is usually referred to as the *local* modal logic arising from a class of Kripke models, in contrast with the *global* one that considers truth in the whole model. It is out of the scope of this work to introduce and study the global modal logic over residuated lattices.

in [4, Lemma 4.8] in order to capture the behaviour of the $\Diamond$ operator. Before doing so, let us state a useful lemma that will allow to move from deductions in the modal logic $\mathsf{BM}_{\mathbf{A}}$ to deductions in the underlying propositional logic $\boldsymbol{\Lambda}(\mathbf{A^c})$.

Lemma 1. *For any $\Gamma \cup \{\varphi\} \subseteq MFm$, $\Gamma \vdash_{\mathsf{BM}_{\mathbf{A}}} \varphi$ iff $Th(\mathsf{BM}_{\mathbf{A}}) \cup \Gamma \vdash_{\boldsymbol{\Lambda}(\mathbf{A})} \varphi$.*[3]

Proof. Right-to-left direction is immediate, since $\mathsf{BM}_{\mathbf{A}}$ expands $\boldsymbol{\Lambda}(\mathbf{A^c})$. The other direction is easily proved by induction on the length of the proof of φ from Γ, observing that the rule (Mon), the only new inference rule added to $\boldsymbol{\Lambda}(\mathbf{A^c})$ in doing the modal expansion, only applies to theorems of the logic. □

If $\Gamma \not\vdash_{\mathsf{BM}_{\mathbf{A}}} \varphi$, the previous result allows us to obtain a non-modal homomorphism h that evaluates to 1 the formulas in Γ and all theorems of $\mathsf{BM}_{\mathbf{A}}$, and does not do the same for φ. This is the reason behind the definition of the canonical model that follows.

Definition 3. *The **canonical Kripke model** of $\mathsf{BM}_{\mathbf{A}}$ is the $\mathbf{A}$-valued model $\mathfrak{M}^{\mathfrak{c}} = \langle W^c, R^c, e^c \rangle$ where:*

- $W^c := \{h \in Hom(\mathbf{MFm}, \mathbf{A})\colon h(Th(\mathsf{BM}_{\mathbf{A}})) = \{1\}\}$,
- $R^c vw := \bigwedge_{\psi \in Fm} (v(\Box\psi) \to w(\psi)) \wedge (w(\psi) \to v(\Diamond\psi))$,
- $e^c(v, p) = v(p)$, *for any propositional variable* p.

As usual, the key fact in using the previously defined model to prove completeness is that it enjoys the so-called *truth lemma*, ensuring that the behaviour of e^c coherently extends to all formulas.

Lemma 2 (Truth lemma). *For any $v \in W^c$ and any modal formula φ, it holds that $e^c(v, \varphi) = v(\varphi)$.*

The previous lemma can be proved by structural induction, the only non trivial cases being the formulas beginning by a modality. One of the inequalities (for both modalities) is easy to prove, as shown next.

Lemma 3. *For any formula $\varphi \in Fm$, and any $v \in W^c$, the following hold:*
1. $v(\Diamond\varphi) \geqslant \bigvee_{w \in W^c} \{R^c(v, w) \odot w(\varphi)\}$,
2. $v(\Box\varphi) \leqslant \bigwedge_{w \in W^c} \{R^c(v, w) \to w(\varphi)\}$.

Proof. We prove the first inequality, the other can be proved analogously. Applying the definition of $R^c(v, w)$ and the monotonicity of $\odot$ in any residuated lattice, it is possible to prove the following inequality for any $v, w \in W^c$:

$$R^c(v, w) \odot w(\varphi) = \bigwedge_{\psi \in Fm} (v(\Box\psi) \to w(\psi)) \wedge (w(\psi) \to v(\Diamond\psi)) \odot w(\varphi)$$
$$\leqslant (w(\varphi) \to v(\Diamond\varphi)) \odot w(\varphi) \leqslant v(\Diamond\varphi).$$

Since this holds for any world w, we have $\bigvee_{w \in W^c} \{R^c(v, w) \odot w(\varphi)\} \leqslant v(\Diamond\varphi)$. □

[3] We do not detail this issue here due to lack of space and interest, but for the interested reader, it should be clear that the language from the right side of this equivalence counts with an extended -countable- set of variables that capture the modal formulas.

As for the converse inequalities, it is worth to first prove a powerful technical lemma (cf. [16, Lemma 6.12]) that generalizes and provides a more modular and scalable proof of the truth lemma compared to that in [4, Lemma 4.8]

Lemma 4. *Let $v \in W^c$ and $\varphi \in Fm$ be such that for all $w \in W^c$ it holds that $R^c(v, w) \leqslant w(\varphi)$. Then $v(\Box\varphi) = 1$.*

Proof. For the sake of a clearer notation, let $\sigma\colon Fm \to Fm$ be given by

$$\sigma(\psi) = (\overline{v(\Box\psi)} \to \psi) \wedge (\psi \to \overline{v(\Diamond\psi)}).$$

In this way, we have that $R^c vw = \bigwedge_{\psi \in Fm} w(\sigma(\psi))$. Now, observe that by definition, $R^c(v, w) \leqslant w(\varphi)$ if and only if

$$\text{for all } a \in A, \text{ if } a \leqslant R^c vw \text{ then } a \leqslant w(\varphi). \tag{2}$$

By hypothesis, this holds for each $w \in W^c$. Unfolding all the definitions, this means that for any $w \in Hom(\mathbf{Fm}, \mathbf{A})$ such that $w(Th(\mathsf{BM}_\mathbf{A})) = \{1\}$, and for any $a \in A$, if $a \leqslant w(\sigma(\psi))$ for all $\psi \in Fm$ then $a \leqslant w(\varphi)$. Clearly, we can now formulate this fact in terms of the propositional consequence relation $\models_\mathbf{A}$:

$$Th(\mathsf{BM}_\mathbf{A}) \cup \{\overline{a} \to \sigma(\psi)\colon \psi \in Fm\} \models_{\mathbf{A}^c} \overline{a} \to \varphi. \tag{3}$$

Since the propositional logic is finitary, then for each $a \in A$ there is a finite set of formulas Σ_a,[4] such that (3) holds iff

$$Th(\mathsf{BM}_\mathbf{A}) \cup \{\overline{a} \to \bigwedge_{\psi \in \Sigma_a} \sigma(\psi)\} \models_{\mathbf{A}^c} \overline{a} \to \varphi.$$

Let $\Sigma := \bigcup_{a \in A} \Sigma_a$, which is clearly finite. Since $\models_{\mathbf{A}^c} \bigwedge_{\psi \in \Sigma} \sigma(\psi) \to \bigwedge_{\psi \in \Sigma_a} \sigma(\psi)$, we have for each $a \in A$, $Th(\mathsf{BM}_\mathbf{A}) \cup \{\overline{a} \to \bigwedge_{\psi \in \Sigma} \sigma(\psi)\} \models_{\mathbf{A}^c} \overline{a} \to \varphi$, from where

$$Th(\mathsf{BM}_\mathbf{A}) \models_{\mathbf{A}^c} \bigwedge_{\psi \in \Sigma} \sigma(\psi)) \to \varphi,$$

by taking for each $h \in Hom(\mathbf{Fm}, \mathbf{A})$ such that $h(Th(\mathsf{BM}_\mathbf{A}),) = 1$, the constant $a = h(\bigwedge_{\psi \in \Sigma} \sigma(\psi))$ in the deduction above.

We can now successively apply completeness of $\Lambda(\mathbf{A})$ w.r.t $\models_\mathbf{A}$, Lemma 1, (Mon) rule and then Lemma 1 and non-modal completeness again to get

$$Th(\mathsf{BM}_\mathbf{A}) \models_{\mathbf{A}^c} \Box(\bigwedge_{\psi \in \Sigma} \sigma(\psi)) \to \Box\varphi.$$

To conclude the proof it suffices to check that $v(\Box(\bigwedge_{\psi \in \Sigma} \sigma(\psi)) = 1$. By axiom (MD), $v(\Box(\bigwedge_{\psi \in \Sigma} \sigma(\psi)) = v(\bigwedge_{\psi \in \Sigma} \Box\sigma(\psi))$, so we only need to check

[4] Note that for a finite set of formulas Θ, $\bigwedge_{\theta \in \Theta} \theta$ is a formula in the language too.

$v(\Box\sigma(\psi)) = 1$ for each $\sigma \in \Sigma$. This is proved by the following chain of equalities:

$$\begin{aligned} v(\Box\sigma(\psi)) =& v(\Box((\overline{v(\Box\psi)} \rightarrow \psi) \wedge (\psi \rightarrow \overline{v(\Diamond\psi)}))) \\ =& v(\Box(\overline{v(\Box\psi)} \rightarrow \psi) \wedge \Box(\psi \rightarrow \overline{v(\Diamond\psi)})) \\ =& v((\overline{v(\Box\psi)} \rightarrow \Box\varphi) \wedge (\Diamond\varphi \rightarrow \overline{v(\Diamond\psi)})) \\ =& (\overline{v(\Box\psi)} \rightarrow v(\Box\varphi)) \wedge (v(\Diamond\varphi) \rightarrow \overline{v(\Diamond\psi)}) = 1. \end{aligned}$$ □

We have now the two main pieces to provide a clear proof of the truth lemma.

Proof of Lemma 2. Let us prove the converse inequalities of Lemma 3. Let $\varphi = \Box\psi$ for some ψ. Since we already know that $e(v, \Box\psi) \geqslant v(\Box\psi)$, to prove the equality is enough to prove that for all $a \in A$, if $a \leqslant e(v, \Box\psi)$ then $a \leqslant v(\Box\psi)$. Thus, let $a \in A$ be such that

$$a \leqslant e(v, \Box\psi) = \inf\{R^c(v, w) \rightarrow e(w, \psi) : w \in W^c\}.$$

By the induction hypothesis, it is enough to prove that $a \leqslant R^c(v, w) \rightarrow w(\psi)$ for all $w \in W$. By residuation, $R^c(v, w) \leqslant a \rightarrow w(\psi)$, and so,

$$R^c(v, w) \leqslant w(\overline{a} \rightarrow \psi) \quad \text{for all } w \in W^c.$$

Lemma 4 implies that $v(\Box(\overline{a} \rightarrow \psi)) = 1$. Then, by axiom ($\mathsf{Ax}_a$) we get that $a \rightarrow v(\Box\psi) = 1$, and so, $a \leqslant v(\Box\psi)$.

In a very similar way we can prove the analogous result for $\Diamond$. Let $\varphi = \Diamond\psi$ for some ψ. To check that $e(v, \Diamond\psi) \geqslant v(\Diamond\psi)$ is enough to prove that for all $a \in A$, if $a \geqslant e(v, \Diamond\psi)$ then $a \geqslant v(\Diamond\psi)$. Thus, let $a \in A$ be such that

$$a \geqslant e(v, \Diamond\psi) = \sup\{R^c(v, w) \odot e(w, \psi) : w \in W^c\}.$$

By induction, this is equivalent to $a \geqslant R^c(v, w) \odot w(\psi)$ for all $w \in W$. By residuation, $R^c(v, w) \leqslant w(\psi) \rightarrow a$, and so,

$$R^c(v, w) \leqslant w(\psi \rightarrow \overline{a}) \quad \text{for all } w \in W^c.$$

From Lemma 4, we know that $v(\Box(\psi \rightarrow \overline{a})) = 1$. Then, by axiom ($\Box\Diamond_a$) we get that $v(\Diamond\psi) \rightarrow a = 1$, and so, $a \geqslant v(\Diamond\psi)$. □

Completeness of $\mathsf{BM}_{\mathbf{A}}$ is now a corollary of the Truth Lemma and Lemma 1.

Theorem 1 (Completeness of $\mathsf{BM}_{\mathbf{A}}$). *For any $\Gamma \cup \{\varphi\} \subseteq \mathbf{MFm}$, $\Gamma \vdash_{\mathsf{BM}_{\mathbf{A}}} \varphi$ iff $\Gamma \Vdash_{\mathbb{BM}_{\mathbf{A}}} \varphi$.*

Proof. Soundness was already justified before Definition 2. Concerning completeness, let $\Gamma \nvdash_{\mathsf{BM}_{\mathbf{A}}} \varphi$. From Lemma 1, there is a homomorphism h from MFm into A evaluating to 1 all theorems of $\mathsf{BM}_{\mathbf{A}}$ and all elements in Γ, and such that $h(\varphi) < 1$. Then, h is by definition a world of the canonical model of $\mathsf{BM}_{\mathbf{A}}$. Using the truth lemma, we know that $e^c(h, \gamma) = h(\gamma) = 1$ for all $\gamma \in \Gamma$, while $e^c(h, \varphi) = h(\varphi) < 1$, and so, the canonical model serves to prove that $\Gamma \nVdash_{\mathbb{BM}_{\mathbf{A}}} \varphi$.

3 Some Useful Axiomatic Extensions

It is reasonable to ask ourselves whether some interesting frame and model conditions can be characterized by means of adding some axiom schemata to the system $\mathsf{BM}_{\mathbf{A}}$. While a systematic study of these properties is far from being developed in the context of modal fuzzy logics, we can still obtain some results for particularly interesting conditions. Motivated by the application to preference modelling in Sect. 4, we will study the classes of transitive, reflexive and symmetric models, and also the class of models whose accessibility relation is *crisp* (i.e., evaluated only on $\{0, 1\}$).

Even though most of the literature addresses fuzzy relations as those evaluated over the interval $[0, 1]$, there is no motivation for that restriction in general, and most of the conventions, notions and results known for fuzzy relations are preserved in the more general context of relations evaluated over bounded integral residuated lattices (i.e., those where there exists a top and bottom elements that coincide with the usual constants 0 and 1). From this reflection, the definition of reflexive, $\odot$-transitive and symmetric **A**-Kripke models is immediate: an **A**-Kripke model $\mathfrak{M}$ is:

R: *Reflexive* when $R(v, v) = 1$ for all $v \in W$.
S: *Symmetric* when $R(v, w) = R(w, v)$ for any $v, u \in W$.
T: $\odot$-*Transitive* when $R(v, w) \odot R(w, u) \leqslant R(v, u)$ for any $v, u, w \in W$.

If $\mathbb{P}$ denotes one or more of the previous conditions, we will denote by $(\mathbb{P})\mathbb{BM}_{\mathbf{A}}$ the class of **A**-Kripke models satisfying the conditions from $\mathbb{P}$. It is not hard to see that the well-known modal axioms that characterize the previous frame conditions in classical modal logic also characterize their corresponding many-valued counterpart defined above.

Proposition 1. *Let* X *be one or more of the following pairs of axiom schematas:*

- $(\mathsf{T}\Box)$: $\Box\varphi \to \varphi$ *and* $(\mathsf{T}\Diamond)$: $\varphi \to \Diamond\varphi$ *(reflexivity)*
- $(\mathsf{B}\Box)$: $\Diamond\Box\varphi \to \varphi$ *and* $(\mathsf{B}\Diamond)$: $\varphi \to \Box\Diamond\varphi$ *(symmetry)*
- $(\mathsf{4}\Box)$: $\Box\varphi \to \Box\Box\varphi$ *and* $(\mathsf{4}\Diamond)$: $\Diamond\Diamond\varphi \to \Diamond\varphi$ *(transitivity)*

Then let $(\mathsf{X})\mathsf{BM}_{\mathbf{A}}$ *be the axiomatic extension of* $\mathsf{BM}_{\mathbf{A}}$ *with the axioms from* X*, and let* $\mathbb{P}$ *be the model conditions corresponding to the axioms in* X*. Then, for any* $\Gamma \cup \{\varphi\} \subseteq MFm$*,* $\Gamma \vdash_{(\mathsf{X})\mathsf{BM}_{\mathbf{A}}} \varphi$ *iff* $\Gamma \Vdash_{(\mathbb{P})\mathbb{BM}_{\mathbf{A}}} \varphi$*.*

Proof. Soundness is easy to check in all three cases. Concerning completeness, it is just necessary to take into account that the canonical model for $(\mathsf{X})\mathsf{BM}_{\mathbf{A}}$ is defined in the same way as the one for $\mathsf{BM}_{\mathbf{A}}$ but taking into account the new equations arising from the additional axioms in the definition of the worlds of the model (that now need to validate them). Under this consideration, reflexivity follows immediately from the definition of $R^c(v, v)$. Indeed,

$$R^c(v, v) = \bigwedge_{\psi \in Fm} (v(\Box\psi) \to v(\psi)) \wedge (v(\psi) \to v(\Diamond\varphi)),$$

and due to the reflexivity axioms ($T\Box$) and ($T\Diamond$) it follows that $R^c(v, v) = 1$.

As for symmetry, assume towards a contradiction that for some $v, w \in W^c$, $R^c(v, w) \not\leqslant R^c(w, v)$. By definition, this means that there is some formula φ such that $R^c(v, w) \not\leqslant (w(\Box\varphi) \to v(\varphi)) \wedge (v(\varphi) \to w(\Diamond\varphi))$, and thus, at least one of the following situations must hold:

(1) $R^c(v, w) \not\leqslant w(\Box\varphi) \to v(\varphi)$,
(2) $R^c(v, w) \not\leqslant v(\varphi) \to w(\Diamond\varphi)$.

It is easy to show that any of the previous conditions leads to a contradiction with the symmetry axioms. For if (1) were to hold, necessarily we would also have $R^c(v, w) \not\leqslant w(\Box\varphi) \to v(\Diamond\Box\varphi)$ due to axiom ($\mathsf{B}\Diamond$), but this contradicts the definition of $R^c(v, w)$ (since $R^c(v, w) \leqslant w(\psi) \to v(\Diamond\varphi)$ for all ψ). The second possible situation (2) is handled in the same way by resorting to axiom ($\mathsf{B}\Box$).

The case of $\odot$-transitivity is a bit more cumbersome but equally simple. Assume towards a contradiction that there are some $v, w, u \in W^c$ for which $R^c(v, w) \odot R^c(w, u) \not\leqslant R^c(v, u)$. Then, by definition of $R^c(w, u)$, there is some formula φ such that $R^c(v, w) \odot R^c(w, u) \not\leqslant (v(\Box\varphi) \to u(\varphi)) \wedge (u(\varphi) \to v(\Diamond\varphi)$. As above, there are two possible cases:

(1') $R^c(v, w) \odot R^c(w, u) \not\leqslant (v(\Box\varphi) \to u(\varphi))$ or
(2') $R^c(v, w) \odot R^c(w, u) \not\leqslant (u(\varphi) \to v(\Diamond\varphi)$.

We can again show that none of the previous conditions can hold. We will show the first one, the other is done analogously (using the dual ($4\Diamond$) axiom). Observe that by Axiom ($4\Box$), together with the fact that $\to$ is decreasing in the first component, (1') implies that $R^c(v, w) \odot R^c(w, u) \not\leqslant (v(\Box\Box\varphi) \to u(\varphi)) = \bigwedge_{z \in W^c}(R^c(v, z) \to z(\Box\varphi)) \to u(\varphi)$. In particular (resorting again to the anti-monotonicity of $\to$), letting $z = w$, we get $R^c(v, w) \odot R^c(w, u) \not\leqslant (R^c(v, w) \to w(\Box\varphi)) \to u(\varphi)$ By the residuation law it follows $R^c(v, w) \odot (R^c(v, w) \to w(\Box\varphi)) \not\leqslant R^c(w, u) \to u(\varphi)$ and thus, we get the contradictory statement $w(\Box\varphi) \not\leqslant R^c(w, u) \to u(\varphi)$. □

Interestingly enough, the previous completeness results allow us to characterize the class of models with fuzzy $\odot$-preorders (i.e. reflexive and $\odot$-transitive models) with axioms ($\mathsf{T}\Box, \mathsf{T}\Diamond$) and ($\mathsf{B}\Box, \mathsf{B}\Diamond$). However, if we further add axioms ($4\Box, 4\Diamond$) we do not get an axiomatization of the class of models with a universal relation (i.e., for which $R(v, w) = 1$ for all v, w), in contrast with what happens in the classical case.

This lack of expressibility can be solved, when the underlying truth-value algebra **A** enjoys certain nice properties, by combining the previous axiomatic extensions with a system complete with respect to the models whose accessibility relation is crisp, i.e. evaluated only over $\{0, 1\}$, Indeed, it is possible to provide such an axiomatic system whenever **A** is subdirectly irreducible (SI),[5] with the same approach that the one followed in [4]. For the sake of simplicity, we will

[5] For the interested reader, see eg. [5] for an insight on the importance of this kind of algebras.

focus in this paper in the particular case of $\mathbf{A}$ being a linearly ordered residuated lattice, which is always a SI residuated lattice. The key idea is the fact that any (finite) SI residuated lattice $\mathbf{A}$, and so, any linearly ordered one, has a unique coatom k, i.e. a unique element $k < 1$ such that, for any $a \in A$, if $a < 1$ then $a \leqslant k$. Since it coincides almost exactly with the proof of [4, Theorem 4.22], we do not detail here the proof of the following result.

Theorem 2. *Let* $\mathbf{A}$ *be a finite linearly ordered residuated lattice, and* $\mathbb{C}_{\mathbf{A}}$ *be the class of crisp* $\mathbf{A}$*-Kripke models. Define the logic* $\mathsf{CBM}_{\mathbf{A}}$ *as the extension of* $\mathsf{BM}_{\mathbf{A}}$ *with the axiom schemata*

– $(\Box \mathsf{k})$ $\quad \Box(\overline{k} \vee \varphi) \to (\overline{k} \vee \Box\varphi)$,

and let $\vdash_{\mathsf{CBM}_{\mathbf{A}}}$ *denote the corresponding notion of proof. Then, for any* $\Gamma \cup \{\varphi\} \subseteq MFm$, $\Gamma \vdash_{\mathsf{CBM}_{\mathbf{A}}} \varphi$ *iff* $\Gamma \Vdash_{\mathbb{C}_{\mathbf{A}}} \varphi$.

As a direct corollary we get the following result.

Corollary 1. *Let* $\mathsf{S5BM}_{\mathbf{A}}$ *be the axiomatic extension of* $\mathsf{BM}_{\mathbf{A}}$ *with the axioms* $(\mathsf{T}\Box), (\mathsf{T}\Diamond), (\mathsf{B}\Box), (\mathsf{B}\Diamond), (4\Box), (4\Diamond)$ *and* $(\Box\mathsf{k})$, *and consider the class* $\mathbb{U}_{\mathbf{A}}$ *of universal* $\mathbf{A}$*-Kripke models. Then, if* $\mathbf{A}$ *is a finite linearly ordered residuated lattice, for any* $\Gamma \cup \{\varphi\} \subseteq MFm$, *we have* $\Gamma \vdash_{\mathsf{S5BM}_{\mathbf{A}}} \varphi$ *iff* $\Gamma \Vdash_{\mathbb{U}_{\mathbf{A}}} \varphi$.

4 Modelling Fuzzy Preferences

In this section, as a matter of illustrating application, we show how the logical machinery developed in the previous sections can be used to devise a logical framework to represent and reason with fuzzy preferences.

We take as starting point van Benthem et al.'s work [1] where, among other logics, the authors consider a basic (classical) modal logic of weak and strict preference interpreted in ordered models of possible worlds, provide a complete axiomatization, and show how global preferences between propositions can be defined by lifting the world ordering to an ordering between sets of worlds. Actually they consider different possibilities to define such global preferences based on (crisp) preference models, i.e. structures $M = (W, \preceq, e)$, where $\preceq$ is preorder on the set of worlds and e is a valuation. The language contains two modal operators, a global S5 modality A and a S4 modality $\Box$, where $\mathsf{A}\varphi$ reads that φ is true in all the worlds, while $\Box\varphi$ reads that φ is true in all the worlds that are more preferred (in the sense of $\preceq$) than the current world. Then, one possibility to encode that "ψ is weakly preferred to φ" is by the formula

$$\varphi \leqslant_{\forall\exists} \psi := \mathsf{A}(\varphi \to \Diamond\psi),$$

to be interpreted as expressing that for any world where φ is true, there is a more preferred world where ψ holds.

In what follows we show how the above framework can be faithfully generalised to deal with both fuzzy propositions and preferences, taking values in a linearly ordered finite residuated lattice **A**. The reason to restrict ourselves to linearly ordered algebras **A** is due to the need of using a global modal operator, for which we only have an axiomatization in such a case, see previous section.

Thus, for modelling preferences we consider a language **PFm** expanding **MFm** with two additional unary operators A and E, which will the role of global operators. The intended semantics is given by the class of reflexive and $\odot$-transitive **A**-Kripke models, that we will call **A**-preference models.

Definition 4. *An* **A***-preference model* $\mathfrak{P}$ *is a triple* $\mathfrak{P} = \langle W, R, e\rangle$ *such that*

- W *is a set of worlds,*
- $R\colon W \times W \to A$ *is an* A*-valued fuzzy pre-order, i.e. a reflexive and* $\odot$*-transitive relation between worlds,*
- $e\colon W \times \mathcal{V} \to A$ *is a* **A***-evaluation of variables that is uniquely extended to formulas of* **PFm** *as in Definition 1 for the propositional connectives and operators* $\Box$ *and* $\Diamond$*, and for the new operators is extended as follows:*

$$e(v, \mathsf{A}\varphi) = \bigwedge_{w \in W} \{e(w, \varphi)\} \qquad e(v, \mathsf{E}\varphi) = \bigvee_{w \in W} \{e(w, \varphi)\}.$$

We will denote by $\mathbb{P}_{\mathbf{A}}$ the class of **A**-preference models, and use $\Vdash_{\mathbb{P}_{\mathbf{A}}}$ with the analogous meaning it had for **A**-valued Kripke models in the previous sections.

After the work developed in the previous sections, it is very natural the way to provide an axiomatic system complete with respect to $\Vdash_{\mathbb{P}_{\mathbf{A}}}$.

Theorem 3 (Completeness). *Let* $\mathsf{P}_{\mathbf{A}}$ *be the deductive system given by:*

- *The axioms and rules of* $(\mathsf{T4})\mathsf{BM}_{\mathbf{A}}$ *for the* $\Box$ *and* $\Diamond$ *operators*
- *The axioms and rules of* $\mathsf{S5BM}_{\mathbf{A}}$ *for the* A, E *operators*
- *The* inclusion *axiom schematas:* $\mathsf{A}\varphi \to \Box\varphi$, $\Diamond\varphi \to \mathsf{E}\varphi$

Denoting by $\vdash_{\mathsf{P}_{\mathbf{A}}}$ *its corresponding notion of proof, then for any* $\Gamma \cup \{\varphi\} \subseteq PFm$*, we have* $\Gamma \vdash_{\mathsf{P}_{\mathbf{A}}} \varphi$ *iff* $\Gamma \Vdash_{\mathbb{P}_{\mathbf{A}}} \varphi$.

Proof. Soundness is a simple exercise. As for completeness, analogously to the approach to prove completeness for the minimal bimodal logic $\mathsf{BM}_{\mathbf{A}}$ in Sect. 2, one can build a corresponding canonical model $\mathfrak{P}^c = (W^c, R_1^c, R_2^c, e^c)$, with $W^c := \{h \in Hom(\mathbf{PFm}, \mathbf{A})\colon h(Th(\mathsf{P}_{\mathbf{A}})) = 1\}$ and this time with two accessibility relations R_1^c and R_2^c, one for the pair of operators $(\Box, \Diamond)$ and another for the operators (A, E). By Proposition 1, it follows that R_1^c is a fuzzy $\odot$-preorder (it is reflexive and $\odot$-transitive), while R_2^c is a crisp equivalence relation. The corresponding truth-lemma (analogous to Lemma 2) shows that if $\Gamma \not\vdash_{\mathsf{P}_{\mathbf{A}}} \varphi$ there is world $v_0 \in W^c$ for which $e(v_0, \gamma) = 1$ for all $\gamma \in \Gamma$ and $e(v_0, \varphi) < 1$. Note that this does not prove yet the claim of the theorem, since R_2^c is not guaranteed to be the universal binary relation on W^c. So a bit more elaboration is needed.

Due to the inclusion axioms, it is immediate to see that $R_1^c(v, w) \leqslant R_2^c(v, w)$ for any pair of worlds, so in particular

$$\text{if } R_1^c(v, w) > 0 \text{ then } R_2^c(v, w) = 1. \tag{4}$$

At this point, we can consider the submodel $\mathfrak{P}^c_{v_0}$ generated by v_0 with respect to R_2, i.e., the model whose universe is $W^c(v_0) = \{w \in W^c\colon R_2(v_0, w) = 1\}$, whose relations are the restrictions of R^c_1 and R^c_2 to $W^c(v_0)$, and whose evaluation of variables is the same. We only need to check that the truth-evaluations in the submodel (in the worlds from $W^c(v_0)$) and in the original model are the same. Note that R^c_2 on $W^c(v_0)$ is total. This can be proved by induction on the complexity of the formula, being immediate for the cases concerning non-modal connectives. As for the modal operators, first observe that for any $u, w \in W$, if $u \in W^c(v_0)$ and $w \notin W^c(v_0)$ it follows that $R^c_2(u, w) = 0$ (since R^c_2 is an equivalence relation), and hence $R^c_1(u, w) = 0$ as well. Then, for any $u \in W^c(v_0)$:

$$e^c_v(u, \Box\varphi) = \bigwedge_{w \in W^c(v_0)} \{R^c_1(u, w) \to e^c(w, \varphi)\} = \bigwedge_{w \in W^c} \{R^c_1(u, w) \to e^c(w, \varphi)\},$$

$$e^c_v(u, \mathsf{A}\varphi) = \bigwedge_{w \in W^c(v_0)} e^c(w, \varphi) = \bigwedge_{w \in W^c : R^c_2(u,w)=1} e^c(w, \varphi).$$

A similar argument can be done for $\Diamond$ and E. This concludes the proof, since the resulting model $\mathfrak{P}^c_{v_0}$ is an **A**-preference model in which there is a world v_0 satisfying Γ and not φ. $\square$

From the above, in the frame of the $\mathsf{P}_{\mathbf{A}}$ logic one can represent the (weak) preference of a proposition ψ over another φ by the expression $\mathsf{A}(\varphi \to \Diamond\psi)$. This preference between propositions actually enjoys the properties of a fuzzy $\odot$-preorder, which justifies in a sense the meaningfulness of this choice. Indeed,

- Reflexivity: $\mathsf{A}(\varphi \to \Diamond\varphi)$ is valid in $\mathbb{P}_{\mathbf{A}}$, since $\varphi \to \Diamond\varphi$, i.e. axiom $(4\Diamond)$, is valid in $\mathbb{P}_{\mathbf{A}}$.
- $\odot$-Transitivity: one can show that

$$\mathsf{A}(\varphi \to \Diamond\psi) \odot \mathsf{A}(\psi \to \Diamond\chi) \to \mathsf{A}(\varphi \to \Diamond\chi) \tag{5}$$

  is also a valid formula in $\mathbb{P}_{\mathbf{A}}$. Namely, this follows by first showing that the following formula expressing a form of monotonicity for $\Diamond$ holds true in $\mathbb{P}_{\mathbf{A}}$:

$$\mathsf{A}(\varphi \to \psi) \to \mathsf{A}(\Diamond\varphi \to \Diamond\psi).$$

  This leads to $\mathsf{A}(\psi \to \Diamond\chi) \to \mathsf{A}(\Diamond\psi \to \Diamond\Diamond\chi)$, but since $\Diamond\Diamond\chi \to \Diamond\chi$ holds true, we get $\mathsf{A}(\varphi \to \Diamond\psi) \odot \mathsf{A}(\psi \to \Diamond\chi) \to \mathsf{A}(\varphi \to \Diamond\psi) \odot \mathsf{A}(\Diamond\psi \to \Diamond\chi)$, and by axiom K for A, it follows the validity of

$$\mathsf{A}(\varphi \to \Diamond\psi) \odot \mathsf{A}(\Diamond\psi \to \Diamond\chi) \to \mathsf{A}(\varphi \to \Diamond\chi),$$

  that directly allows us to show the validity of (5).

5 Conclusions

In this paper we have been concerned with completing the notion of the minimal modal logic over a finite residuated lattice (and with truth-constants) from [4] to get a full-fledged modal logic with both a necessity and a possibility operators. The gain in expressibility has been used, as a matter of example, to define a many-valued counterpart of a modal logic studied in [1] to reason about preferences between propositions in such a many-valued setting. As for future work, there are several interesting open research issues that are left open, among them:

– axiomatization of modal expansions of logics arising from varieties generated by a finite residuated lattice, probably without resorting to canonical truth-constants;
– better understanding of the expressibility of general frame/model conditions in residuated lattice-based modal logics;
– deepening on the axiomatization of the logic arising from crisp Kripke models over non SI residuated lattices;
– general study of a larger set of preference relations definable in the many-valued context introduced in this work, along the line of [1,13].

Acknowledgements. Vidal acknowledges support by the joint project Austrian Science Fund (FWF) I1897-N25 and Czech Science Foundation (GACR) 15-34650L, and by the institutional grant RVO:67985807. Esteva and Godo acknowledge support by the FEDER/MINECO project TIN2015-71799-C2-1-P.

References

1. van Benthem, J., Girard, P., Roy, O.: Everything else being equal: A modal logic for ceteris paribus preferences. J. Philos. Logic **38**, 83–125 (2009)
2. van Benthem, J., Liu, F.: Dynamic logic of preference upgrade. J. Appl. Non-Class. Logic **17**(2), 157–182 (2007)
3. Bou, F., Cerami, M., Esteva, F.: Finite-valued Łukasiewicz modal logic is PSPACE-complete. In: Proceedings of IJCAI 2011, pp. 774–779 (2011)
4. Bou, F., Esteva, F., Godo, L., Rodríguez, R.: On the minimum many-valued modal logic over a finite residuated lattice. J. Logic Comput. **21**(5), 739–790 (2011)
5. Burris, S., Sankappanavar, H.P.: A course in Universal Algebra. The Millennium (2012). update edition
6. Caicedo, X., Metcalfe, G., Rodriguez, R., Rogger, J.: Decidability of order-based modal logics. J. Comput. Syst. Sci. (2016, in Press)
7. Caicedo, X., Rodríguez, R.O.: Standard Gödel modal logics. Stud. Logica **94**(2), 189–214 (2010)
8. Caicedo, X., Rodríguez, R.O.: Bi-modal Gödel logic over [0, 1]-valued kripke frames. J. Logic Comput. **25**(1), 37–55 (2015)
9. Cintula, P., Noguera, C., Rogger, J.: From kripke to neighborhood semantics for modal fuzzy logics. In: Carvalho, J., et al. (eds.) Proceedings of IPMU 2016, Part II. CCIS 611, pp. 123–135 (2016)
10. Fan, T.F.: Fuzzy bisimulation for Gödel modal logic. IEEE Trans. Fuzzy Syst. **23**(6), 2387–2396 (2015)
11. Fan, T.F., Liau, C.J.: Logical characterizations of regular equivalence in weighted social networks. Artif. Intell. **214**, 66–88 (2014)
12. Hansoul, G., Teheux, B.: Extending Łukasiewicz logics with a modality: Algebraic approach to relational semantics. Stud. Logica **101**(3), 505–545 (2013)
13. Liu, F.: Reasoning About Preference Dynamics. Synthese Library, vol. 354. Springer, Heidelberg (2011)

14. Marti, M., Metcalfe, G.: A hennessy-milner property for many-valued modal logics. In: Goré, R., Kooi, B.P., Kurucz, A. (eds.) Advances in Modal Logic, Vol. 10, pp. 407–420. College Publications (2014)
15. Priest, G.: Many-valued modal logics: a simple approach. Rev. Symbolic Logic **1**(2), 190–203 (2008)
16. Vidal, A.: On modal expansions of t-norm based logics with rational constants. Ph.D. Dissertation, University of Barcelona (2015)
17. Vidal, A., Esteva, F., Godo, L.: On modal extensions of product fuzzy logic. J. Logic Comput. **27**(1), 299–336 (2017)

Improving Supervised Classification Algorithms by a Bipolar Knowledge Representation

Guillermo Villarino[1(✉)], Daniel Gómez[1], and J. Tinguaro Rodríguez[2]

[1] Facultad de Estudios Estadísticos, Universidad Complutense, Avenida Puerta de Hierro s/n, 28040 Madrid, Spain
gvillari@ucm.es, dagomez@estad.ucm.es

[2] Facultad de Ciencias Matemáticas, Universidad Complutense, Plaza de Las Ciencias 3, 28040 Madrid, Spain
jtrodrig@mat.ucm.es

Abstract. The aim of supervised classification algorithms is to assign objects/items to known classes. Before carrying out the final assignment, many classification algorithms obtain a soft score (probability, fuzzy, possibility, cost...) between each item and each class. In order to improve this final decision, we build a bipolar probabilistic model that considers some extra information about the dissimilarity structure between the classes. We present here some improvements for several supervised classification algorithms such as random forest, decision trees and neural networks for binary classification problems.

Keywords: Supervised classification models · Bipolar models · Soft information

1 Introduction

Classification is one of the most important topics in statistics since the beginning of supervised learning algorithms. It is possible to find a huge number of supervised classification algorithms and applications based on the problem, type of data, characteristics or efficiency [20, 21].

In a supervised classification scheme, the main aim is to classify a set of items into classes based on a training sample that is used to train the algorithm. Once the algorithm has been trained, the algorithm classifies the new items into some of the known classes. Nevertheless, we should differentiate in many classification algorithms between the internal learning process that assigns a score (probability, fuzzy, possibility) to the pair (item, class) and the final decision (this item belong to this class) that is made from the former.

In this paper, we are going to model this soft information that appears in many classification algorithms into a bipolar way with the main aim of fixing some of the deficiencies in the learning process and improving the classification performance. We will focus on several algorithms such as CART [6], Random

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_48

Forest (RF) [7] and Neural Networks [26,31] but other classification algorithms could also benefit from our approach.

For the experimental analysis we firstly fit these three algorithms and compute the estimated probabilities of belonging to each class in the training step. The negative probability is calculated once a dissimilarity matrix (which we shall introduce later) has been found, and eventually both positive and negative probabilities are aggregated to achieve a final bipolar probability for each class. The final decision is given, in a usual way, by taking into account this new probability value.

The suitability of the proposed bipolar approach is evaluated for each algorithm in the framework of standard classification. Specifically, our new method is tested on 10 datasets selected from the KEEL dataset repository [3], and it is supported by a proper statistical analysis, as suggested in the literature (see e.g. [9,12,13]).

The remainder of this paper is organized as follows. Section 2 describes the base idea of bipolar knowledge. Then, in Sect. 3 we introduce the probabilistic bipolar models in the frame of supervised classification problems. Two aggregation functions are proposed to provide the final bipolar decision. Finally, the experimental framework along with the respective experimental analysis are presented in Sects. 4 and 5. We summarize the paper with the main concluding remarks in Sect. 6.

2 Bipolar Knowledge

Human beings have the ability and the propensity to reason and make decisions on the basis of positive and negative affects. In other words, information or alternatives are usually judged not only in terms of plausibility or utility, but also attending to their intrinsic positive or negative character, as when people make decisions by weighing up the good (pros) and bad (cons) sides of the available alternatives and choosing that with the best balance between those sides or poles. Due to the presence of two poles, this human judgment feature is usually referred to as bipolarity.

In recent years, many authors (see for instance [5,10,24,25]) have emphasized the relevance of introducing a bipolar approach in soft knowledge representation formalisms as fuzzy logic [34] or possibility theory [11] in order to enhance their expressive and representational power. On the other hand, our knowledge about our own methods of reasoning has been a fruitful source of inspiration in the development of artificial intelligence (AI) since its foundation in the middle of last century. In fact, important areas of this discipline have been closely concerned with the replication of some mental capacities of the human beings, like playing chess or proving theorems.

For example, expert systems and decision support systems replicate the human ability to make correct decisions in many fields of activity (such as medical care, advertising, control tasks, disaster management, etc.) based on explicit knowledge about these realities. In the same way, machine learning and data mining techniques (see e.g. [18]) have found a wide success by trying to replicate the human ability to discover patterns and to learn from examples or experience (usually assumed to be contained in databases). However, it is important to remark that this success would have not been possible without finding some inspiration in our understanding of how humans manage to solve problems in these fields and neither without the development of formal models enabling an effective representation of this knowledge either. In consonance with these ideas, in this paper we propose to introduce a bipolar approach in the framework of soft rule-based classification systems, as e.g. fuzzy classification systems or random forests.

3 Probabilistic Bipolar Models in Supervised Classification Problems

Given a classification scheme, let us denote by $\{C_1, \ldots, C_k\}$ the set classes, and let us denote by $X = \{x_1, \ldots, x_n\}$ the set of items that has to be classified. Most of the classification users focus on the final result of the classification process, only considering the final class to which an item $x \in X$ has been classified. In this sense, classifiers are usually viewed as functions

$$C\colon X \longrightarrow \{C_1, \ldots, C_k\}.$$

As we say in the introduction most classification process goes through many steps to arrive to the final decision, and the soft information appears in a natural way in intermediary steps. Taking into account this, let us going to denote by a probabilistic classifier C as a function

$$C\colon X \longrightarrow [0,1]^k.$$

The probabilistic classifiers gives us for each object x, a probability vector $C(x) = (p_1, \ldots, p_k)$ where p_i represents the probability given by the algorithm to which object x belongs to the class C_i. Obviously, in a probabilistic scheme we could impose that $\sum_{i=1}^{k} p_i = 1$. Let us note that many classical algorithms can be addressed in this class if we gather the information given by the algorithm in a intermediary step.

The output that we take from a probabilistic classifier and the structure and relationships that exist among classes can be used to represent the knowledge in a bipolar way. In this work, we are going to consider that the vector of probabilities given by a probabilistic classifier is the positive information that we have for an item x. So let us denote by $p^+ = \left(p_1^+, \ldots, p_k^k\right) = (p_1, \ldots, p_k)$. Obviously, for each item x, we have $\sum_{i=1}^{k} p_i^+ = 1$.

Let us represent by p_i^- the probability of item x to belong a dissimilarity class of C_i and by $D = (d_{i,j})$ the dissimilarity matrix, where $d_{i,j}$ represent the degree dissimilarity between classes C_i and C_j. Let us observed that for any dissimilarity matrix D, $d_{ii} = 0$ for all $i = 1, \ldots, k$, and this matrix is in general non symmetric.

In this work, we built the negative probabilities from this dissimilarity measure and from the positive probabilities. As we have previously mentioned, the negative probability represents the probability of belongings to a dissimilar class, so the following equation holds.

$$p_i^- = \sum_{r=1}^{k} d_{ir}\ p_r^+$$

Remark 1. Let us note that the negative probability represents the probability of membership to a dissimilar class, and then this probability does not come from a probability distribution.

Once we have defined for each class and for each item, the positive and negative information, in this work we propose an aggregation process to transform the pair (p_i^+, p_i^-) into one probability that we have called bipolar probability p_i^{bi}. Once these probabilities are obtained for all items and for all classes, we will assign an object x to the class in which the maximum bipolar probability is reached. In this sense a crisp bipolar classifier $C^{bi}\colon X \longrightarrow \{C_1, \ldots, C_k\}$ can be built from a probabilistic one C, in the following way:

$$\begin{array}{ccccc} C^{bip} : X \longrightarrow & [0,1]^k \longrightarrow & [0,1]^k x [0,1]^k \longrightarrow & [0,1]^k \longrightarrow & \{C_1, \ldots, C_k\} \\ x \longrightarrow & C(x) = (p_1, \ldots, p_k) \longrightarrow & (p_1^+, \ldots, p_k^+) x (p_1^-, \ldots, p_k^-) \longrightarrow & (p_1^{bip}, \ldots, p_k^{bip}) \longrightarrow & C_{arg\ Max\{p_1^{bip}, \ldots, p_k^{bip}\}} \end{array}$$

From now on, the key process is the manner on how to aggregate the positive and negative information in order to build this bipolar probability. In this work we have considered two aggregation models that will be shown with more detail.

Let us note that this is our first bipolar approach, but other ways for dealing with positive and negative information will be explored in future works.

In a classification problem, let $X = \{x_1, \ldots, x_n\}$ be the set of items and let $\{C_1, \ldots, C_k\}$ be the set of classes. If C is a probabilistic classifier and D a dissimilarity matrix between the set of classes, it is possible to built (as we have explained previously) for each item x_l, the vector of positive probabilities $p^{+l} = (p_1^{+,l}, \ldots, p_k^{+,l})$ and the vector of negative probabilities $p^{-l} = (p_1^{-,l}, \ldots, p_k^{-,l})$. Without lost of generality, we will denote by $p^+ = (p_1^+, \ldots, p_k^+)$ and by $p^- = (p_1^-, \ldots, p_k^-)$ the positive and negative probability vector respectively for a given item x_l. The question now is how to aggregate for a given class C_i the values p_i^+ p_i^- into one. Obviously different ways of aggregation will give different bipolar classifiers. In this work we have studied two, that we define below.

Definition 1. *Let p_i^+, p_i^- be the positive and negative probability of one item into one class C_i, we define the additive bipolar probability of the object x into class C_i as*

$$p_i^{bipad} = Max\{0, p_i^+ - p_i^-\}.$$

Previous model presents an additive aggregation of positive and negative information. A $p_i^{bipad} > 0$ represent the positive gap between the probability of belonging to the class C_i and the probability of belong to a dissimilar class of C_i. A value zero of p_i^{bipad} gives an idea that, for the item i, there are more negative indications than positive and thus the bipolar classifier should not assign the item to this class.

In the following definition, we present an alternative way to aggregate the positive and negative information into one that we have called logistic bipolar probability according to the usual notation of the exponential aggregation function used.

Definition 2. *Let p_i^+, p_i^- be the positive and negative probability of one item into one class C_i, we define the logistic bipolar probability of the object x into class C_i as*

$$p_i^{biplog} = \begin{cases} 1 - e^{-\frac{p_i^+}{p_i^-}} & if\ p_i^- > 0 \\ 1 & otherwise \end{cases}$$

To conclude this section, the bipolar classifier is defined from the bipolar probability.

Definition 3. *Given a probabilistic classifier $C\colon X \longrightarrow [0,1]^k$ over an universe X. And given a dissimilarity matrix D that permits to built for each item x the bipolar information $(p^+, p^-) = \big((p_1^+, p_1^-), \ldots, (p_k^+, p_k^-)\big)$ from the probabilistic classifier $C(x)(p_1, \ldots, p_k)$. The crisp bipolar classifier is defined as follow:*

$$C^{bip}(x) = C_r \ if\ and\ only\ if\ p_r^{bip} = Max\{p_j^{bip}; j = 1..k\}.$$

where p^{bip} is the bipolar probability result of aggregation the vectors p^+ and p^-.

From now on, we will denote by C^{bipad} to the additive bipolar classifier built from the probabilistic classifier C, and by C^{biplog} we will denote the logistic bipolar classifier.

4 Experimental Framework

This section is aimed to analyze the behaviour of our bipolar knowledge representation approaches when applied on recognized classifiers such as CART [6], Random Forest [7] and Neural Networks [26,31].

We firstly present the real world classification datasets selected for the experimental study. Next, we briefly describe the different classifiers and also the configuration of the GA that we will use in the experimental analysis. Finally, we present the statistical analysis used to evaluate the results.

4.1 Data Sets

We have selected a benchmark of 10 datasets from the KEEL dataset repository [3], which are publicly available on the corresponding web site, including general information about them, partitions for the validation of the experimental results and so on.

A 5-folder cross-validation model was considered in order to carry out the different experiments. That is, the considered datasets have been split into 5 random partitions of data, each one with 20% of the patterns, and we employ a combination of 4 of them (80%) to train the system and the remaining one to test it.

Table 1 summarizes the properties of the selected datasets, showing for each dataset the number of examples (#Ex.), the number of attributes (#Atts.), type (Real/Integer/Natura) and the imbalance ratio once the dataset has been transform into a binary classification problem.

Table 1. Summary description for the employed datasets.

Id.	Data-set	#Ex.	#Atts.	(R/I/N)	#IR
aut	Autos	159	25	(15/0/10)	2.58
con	Contraceptive	1473	9	(6/0/3)	3.43
fla	flare	1066	25	(15/0/10)	2.58
eco	ecoli	336	7	(7/0/0)	3.34
gla	Glass	214	9	(9/0/0)	2.05
lin	Lymphography	148	18	(3/0/15)	1.43
shu	Shuttle	2175	9	(0/9/10)	5.44
thy	Thyroid	720	21	(6/0/15)	18.1
yea	Yeast	1484	8	(8/0/0)	5.1
car	Car	159	25	(15/0/10)	3.5

We must point out that the thyroid dataset presents a high IR, because of its previous multiclass imbalance distribution. We consider that fact as an opportunity to asses our bipolar approaches when dealing with high imbalance datasets.

4.2 Algorithms Considered as Base Classifier

We have selected three well known classifiers for our first experiment. On one hand, we have considered the CART [6] algorithm because of its simplicity, on the other hand we have chosen a more complex rule based classifier such as Random Forest [7] as well as one of the most powerful machine learning parametric algorithms, Neural Networks [26,31].

For the learning process, we have used the following configuration of the GAs used for the learning of dissimilarity structures, that provided a solution in a feasible amount of time. We indicate the values that have been considered for the parameters of the evolutionary tuning of our bipolar proposal:

– Population Size: 50 individuals.
– Number of iterations: 20

- Mutation Chance: 0.01.
- Elitism: About 20% of the population size.

4.3 Statistical Test for Performance Comparison

In this paper, we use some hypothesis validation techniques in order to give statistical support to the analysis of the results. We will use non-parametric tests because the initial conditions that guarantee the reliability of the parametric tests cannot be fulfilled, which implies that the statistical analysis loses credibility with these parametric tests [9].

Specifically, we employ the Wilcoxon rank test [32] as a non-parametric statistical procedure for making pairwise comparisons between two algorithms. For multiple comparisons, we use the Friedman aligned rank test, since it is recommended in the literature [9,12,13] to detect statistical differences among a group of results, and the Holm post-hoc test [17] to find the algorithms that reject the equality hypothesis with respect to a selected control method.

The post-hoc procedure allows us to know whether a hypothesis of comparison of means could be rejected at a specified level of significance α. Furthermore, we compute the adjusted p-value (APV) in order to take into account that multiple tests are conducted. In this manner, we can compare directly the APV with respect to the level of significance α in order to be able to reject the null hypothesis.

These tests are suggested in the studies presented in [9,12,13], where it is recommended their use in the field of machine learning. A complete description of these tests with many considerations and recommendations and even the software used to run this analysis can be found on the website: http://sci2s.ugr.es/sicidm/.

To analyse the suitability of our bipolar proposal we want to show empirically whether our methodology enhances the results of the base classifier without the bipolar tuning step. In this sense, for each base classifier we have two aggregation functions.

Therefore, the main aim is to compare, for each classifier, the perform reached by the three approaches: Classifier without tuning, Classifier + additive bipolar and Classifier + logistic bipolar.

5 Experimental Results

This section is aimed to show the capacity of enhancement of our two new bipolar approaches with respect to the base algorithms which this final decision tuning method are applied to.

This experimental analysis is carried out by training the base classifier in such a way that the best parameters for each training set will be selected in terms of the kappa statistic. It means, we are selecting the optimal base classifier configuration among the grid of parameters considered. We have selected a bootstraped training with 25 samples.

Once the classifier is trained, we ran the GA in order to find the best dissimilarity structure for this training set. Finally, the predictions in test data for the three models (base classifier, base classifier + Additive bipolar tuning and base classifier + logistic bipolar tuning) are computed.

This experimental study has been carried out using R Software. Specifically, we used the *caret* package [19] for the classifiers training and the *genalg* package [33] to run the GA.

Results are grouped, for each base algorithm, in pairs for training and test, where the best global result for each dataset in test is stressed in **bold-face**.

We observe from the results of Tables 2, 3 and 4 the general good behaviour of the bipolar tuning method, at least in one of the aggregation approaches, since it enhances the performance of the initial base algorithms.

Table 2. Results in train and test achieved by the genetic bipolar approaches applied to the CART algorithm.

	CART		CARTbipAdd		CARTbipLog	
	Tr.	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*
aut	0.730	0.738	0.730	0.738	0.730	0.738
car	0.793	0.724	0.793	**0.733**	0.793	**0.733**
con	0.328	0.241	0.337	0.295	0.339	**0.298**
eco	0.714	0.598	0.714	0.598	0.714	0.598
fla	0.584	0.529	0.595	**0.530**	0.595	**0.530**
gla	0.714	0.531	0.714	0.531	0.714	0.531
lin	0.659	0.514	0.659	0.514	0.659	0.514
shu	0.995	0.993	0.995	0.993	0.995	0.993
thy	0.890	0.823	0.890	0.823	0.890	0.823
yea	0.521	0.436	0.524	**0.447**	0.524	**0.447**
Mean	0.693	0.613	0.695	0.620	0.695	**0.621**

When the bipolar method is applied to the CART classifier, see Table 2, improvements are achieved in 4 out of 10 datasets and very similar results are reached by both aggregations. By contrast, for the RF classifier it is clear that the additive aggregation of both positive and negative information outperforms the logistic approach.

In case of NNet, as we can see from the results of Table 4, in general terms the logistic aggregation seems to reach better results than the additive one.

Statistical Analysis. In order to detect significant differences among the results of the different approaches, we carry out the Friedman aligned rank test. This test obtains a low p-value for all the three algorithms, which implies that there are significant differences between the results. For this reason, we can apply

Table 3. Results in train and test achieved by the genetic bipolar approaches applied to the RF algorithm.

	RF		RFbipAdd		RFbipLog	
	Tr.	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*
aut	1.000	0.760	1.000	**0.787**	1.000	0.779
car	1.000	0.796	1.000	**0.805**	1.000	0.804
con	0.868	0.235	0.884	0.252	0.884	**0.259**
eco	1.000	0.578	1.000	**0.581**	1.000	0.580
fla	0.646	0.549	0.672	0.571	0.673	**0.575**
gla	1.000	0.666	1.000	**0.701**	1.000	0.691
lin	0.986	0.684	1.000	**0.702**	1.000	0.677
shu	1.000	**0.996**	1.000	0.995	1.000	0.995
thy	1.000	0.885	1.000	**0.899**	1.000	0.883
yea	1.000	**0.503**	1.000	0.495	1.000	0.491
Mean	0.950	0.665	0.956	**0.679**	0.956	0.673

Table 4. Results in train and test achieved by the genetic bipolar approaches applied to the NNet algorithm.

	NNet		NNetbipAdd		NNetbipLog	
	Tr.	*Tst*	*Tr.*	*Tst*	*Tr.*	*Tst*
aut	0.686	0.568	0.703	**0.610**	0.703	**0.610**
car	0.992	0.954	0.995	0.948	0.995	**0.949**
con	0.254	0.235	0.382	0.271	0.388	**0.285**
eco	0.662	0.555	0.693	**0.581**	0.692	**0.581**
fla	0.597	0.589	0.626	**0.595**	0.633	0.589
gla	0.568	0.442	0.647	0.476	0.645	**0.480**
lin	0.895	**0.759**	0.905	0.744	0.906	0.745
shu	0.962	0.955	0.980	**0.966**	0.981	0.965
thy	0.856	0.628	0.933	0.813	0.941	**0.817**
yea	0.518	0.467	0.558	**0.529**	0.562	0.528
Mean	0.699	0.615	0.742	0.653	0.745	**0.655**

a post-hoc test to compare our methodology against the remaining approaches. Specifically, a Holm test is applied using the best approach (the one with the lower ranking) as control method and computing the adjusted p-value (APV). The statistical analysis, shown on Table 5, reflects that the bipolar method outperforms the base classifier with a high level of confidence for all the algorithms considered.

Table 5. Average rankings of the algorithms (Aligned Friedman), associated p-values and Holm test APV for each algorithm.

Algorithm	Ranking CART	Ranking RF	Ranking NNet
"Ref"	21.9	22.2	23.4
"BipAdd"	12.35	9.45	11.9
"BipLog"	12.25	14.85	11.2
p-val	0.0271	0.0255	0.0295
Holm APV	**0.0284***	**0.0024***	**0.0038***

Table 6. Wilcoxon test to compare the bipolar tuning approaches (R^+) against the base classifier (R^-).

Comparision	R^+	R^-	Ex. p-val	Asymp. p-val
CARTbipAdd vs. CART	44.5	10.5	0.094	0.074
CARTbipLog vs. CART	44.5	10.5	0.094	0.074
RFbipAdd vs. RF	51.0	4.0	0.013	0.014
NNetbipAdd vs. NNet	50.0	5.0	0.019	0.019
NNetbipLog vs. NNet	48.0	7.0	0.037	0.032

The statistical analysis, which is carried out by means of a Wilcoxon test, see Table 6, clearly reflects the superiority of our new methodology with an acceptable p-value. In case of RF, only the additive aggregation could be considered better than the reference in statistical terms.

6 Discussion and Final Remarks

In this paper we have presented a bipolar knowledge representation-based classifier, a proposal to tune the classification given by any classification algorithm. In order to do so, we have presented a GA based method to find the optimal dissimilarity structure between the classes and also two new approaches for the aggregation of both positive and negative information.

Along the experimental study, we have learned several lessons:

- The bipolar method allows to improve the results of the three base machine learning algorithms considered in this work.
- Both additive and logistic aggregations outperform the results of the base classifier in case of CART and NNet, and only the additive aggregation method enhances the behaviour of RF in statistical terms.
- Comparing the two different aggregation methods, there is no clear winner, in fact it is highly dependent on the base algorithm considered as well as on the dataset of application.

These results allow us to conclude that our new methodology is a suitable solution to confront binary classification problems by incorporating the bipolar knowledge representation framework to the soft information given by any algorithm.

Acknowledgment. This research has been partially supported by the Government of Spain, grant TIN2009-07901.

References

1. Amo, A., Montero, J., Molina, E.: Representation of consistent recursive rules. Eur. J. Oper. Res. **130**, 29–53 (2001)
2. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: SIGMOD Conference, pp. 207–216 (1993)
3. Alcalá-Fdez, J., Fernandez, A., Luengo, J., Derrac, J., García, S., Sánchez, L., Herrera, F.: KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. J. Multiple-Valued Log. Soft Comput. **17**(2–3), 255–287 (2011)
4. Alcalá-Fdez, J., Alcala, R., Herrera, F.: A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. IEEE Trans. Fuzzy Syst. **19**(5), 857–872 (2011)
5. Atanassov, K.T.: Intuitionistic Fuzzy Sets Theory And Applications. Physica-Verlag, Heidelberg, New York (1999)
6. Breiman, L.: Classification and Regression Trees. Kluwer Academic Publishers, New York (1984)
7. Breiman, L.: Random forests. Mach. Learn. **40**, 5–32 (2001)
8. Cordon, O., del Jesús, M.J., Herrera, F.: A proposal on reasoning methods in fuzzy rule-based classification systems. Int. J. Approximate Reasoning **20**(1), 21–45 (1999)
9. Demsar, J.: Statistical comparisons of classifiers over multiple datasets. J. Mach. Learn. Res. **7**, 1–30 (2006)
10. Dubois, D., Prade, H.: An introduction to bipolar representations of information and preference. Int. J. of Intell. Syst. **23**(8), 866–877 (2008)
11. Dubois, D., Prade, H.: A bipolar possibilistic representation of knowledge and preferences and its applications. In: Fuzzy Logic and Applications, vol. 3849, pp. 1–10 (2006)
12. García, S., Herrera, F.: An extension on "statistical comparisons of classifiers over multiple datasets" for all pairwise comparisons. J. Mach. Learn. Res. **9**, 2677–2694 (2008)
13. García, S., Fernandez, A., Luengo, J., Herrera, F.: Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power. Inform. Sci. **180**(10), 2044–2064 (2010)
14. Goldberg, D.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Boston (1989)
15. Gómez, D., Montero, J., Yanez, J.: A coloring fuzzy graph approach for image classification. Inf. Sci. **176**(24), 3645–3657 (2006)
16. Holland, J.H.: Adaptation in Natural and Artificial Systems. The University of Michigan Press, Ann Arbor (1975)

17. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Stat. **6**, 65–70 (1979)
18. Hullermeier, E.: Fuzzy methods in machine learning and data mining: Status and prospects. Fuzzy Sets Syst. **156**(3), 387–406 (2005)
19. Kuhn, M.: Building predictive models in R using the caret package. J. Stat. Softw. **28**(5), 1–26 (2008). doi:10.18637/jss.v028.i05
20. Kumar, R., Verma, R.: Classification algorithms for data mining: a survey. Int. J. Innovations Eng. Technol. **2**, 7–14 (2012)
21. Lim, T.S., Loh, W.Y., Shih, Y.S.: A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. Mach. Learn. **40**, 203–228 (2000)
22. Lucasius, C., Kateman, G.: Understanding and using genetic algorithms Part 1. Concepts, properties and context. Chemometr. Intell. Syst. **10**, 1–33 (1993)
23. Lucasius, C., Kateman, G.: Understanding and using genetic algorithms Part 2. Representation, configuration and hybridization. Chemometr. Intell. Lab. Syst. **25**, 99–145 (1994)
24. Montero, J., Gómez, D., Bustince, H.: On the relevance of some families of fuzzy Sets. Fuzzy Sets Syst. **158**(22), 2429–2442 (2007)
25. Montero, J., Bustince, H., Franco, C., Rodríguez, J.T., Gómez, D., Pagola, M., Fernandez, J., Barrenechea, E.: Paired structures in knowledge representation. Knowl.-Based Syst. **100**, 50–58 (2016)
26. Ripley, B.D.: Pattern Recognition and Neural Networks. Cambridge (1996)
27. Rodríguez, J.T., Vitoriano, B., Montero, J.: A general methodology for data-based rule building and its application to natural disaster management. Comput. Oper. Res. **39**(4), 863–873 (2012)
28. Rodríguez, J.T., Vitoriano, B., Gómez, D., Montero, J.: Classification of disasters and emergencies under bipolar knowledge representation. In: Vitoriano, B., Montero, J., Ruan, D. (eds.) Decision Aid Models for Disaster Management and Emergencies, Atlantis Computational Intelligence Systems, vol. 7, pp. 209–232 (2013)
29. Sivanandam, S.N., Deepa, S.N.: Introduction to Genetic Algorithms. Springer, Heidelberg (2007)
30. Trillas, E., Moraga, C., Guadarrama, S., Cubillo, S., Castineira, E.: Computing with antonyms. Forging New Frontiers: Fuzzy Pioneers I, vol. 217, pp. 133–153 (2007)
31. Venables, W.N., Ripley, B.D.: Modern Applied Statistics with S, 4th edn. Springer, New York (2002)
32. Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics **1**, 80–83 (1945)
33. Willighagen, E.: genalg: R Based Genetic Algorithm. (2005). http://cran.r-project.org/
34. Zadeh, L.A.: Fuzzy-Logic. Comput. **21**(4), 83–93 (1988)

Edge Detection Based on the Fusion of Multiscale Anisotropic Edge Strength Measurements

Gang Wang(✉) and Bernard De Baets

KERMIT, Research Unit Knowledge-based Systems, Ghent University,
Coupure Links 653, 9000 Ghent, Belgium
{gang.wang,bernard.debaets}@ugent.be

Abstract. Edge detection plays an essential role in many computer vision tasks, but there is limited literature on the fusion of multiscale edge strength measurements. In this paper, we extend an edge detector using both isotropic and anisotropic Gaussian kernels in multiscale space to obtain the multiscale anisotropic edge strength measurements (AESMs). Subsequently, we propose a fusion scheme of multiscale AESMs based on geometric mean. This scheme inherits the merits of the isotropic/anisotropic Gaussian kernel based method and suppress the scale-space diffusion at the same time. Experimental results on example images in the EUSFLAT Edge Detection Competition dataset illustrate that the proposed method outperforms the widely used Canny method and the state-of-the-art isotropic/anisotropic Gaussian kernel method.

Keywords: Edge detection · Anisotropic gaussian kernel · Multiscale edge measurement fusion

1 Introduction

Edges are fundamental image features and play an essential role in many computer vision tasks, such as noise reduction, image segmentation, object tracking and active contour detection [4]. The output binary edge line is usually defined as the one-pixel wide centerline of edge areas where the intensity exhibits significant discontinuities. Over the past several decades, edge detection has received great attention and various methods have been developed, including differentiation-based methods [2,5,9,10], machine learning methods [3], fuzzy transform methods [7], etc.

However, there are several obstacles that make edge detection still a challenging problem. First, the image sometimes contains noise and textures that also have significant local contrast and will lead to spurious detection results. Secondly, some edges may be situated adjacent to other structures, like parallel edges and thin ridges, which will decay the edge strength measurement. Moreover, the width and the steepness of the edge areas are highly variable.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_49

In order to deal with these problems, anisotropic scale space representations and fusion of multiscale edge strength measurements (ESMs) have gained attention. Bao et al. [1] detected edges using scale multiplication and improved the pioneering work of Canny [2]. However, in their work only two scales are considered. Lopez-Molina et al. [5] proposed an edge detector based on restricted multiscale Sobel operators. Because of the inherent isotropy of the Sobel operator, this detector shows a limited performance for adjacent edges. Shui and Zhang [9] proposed an isotropic/anisotropic Gaussian kernel (IAGK) based edge detector that can improve both the noise robustness and the detection of irregular edges compared with Canny's method. Zhang et al. [10] further developed this method with automatic calculation of the anisotropic factor under the principles of high signal-to-noise ratio, fine localization and high edge resolution. However, edges in images usually exhibit multiscale characteristics. The edges with narrow width are suitable to be detected at fine scales, while the ones with large width are suitable to be detected at coarse scales. Unfortunately, all the aforementioned methods show problems when they are applied in multiscale space. If we use an edge detector with a small scale, wide edges will produce a weak response. Moreover, this kind of small scale edge detector is not noise-robust. If we employ a large scale edge detector, the diffusion effect will cause strong responses for non-edge structures. Besides, this will also lead to more serious interference between adjacent structures. To our knowledge, an effective fusion scheme of multiscale ESMs is still lacking.

With the purpose of tackling the fusion of multiscale ESMs, in this paper, after elaborating related works in Sect. 2, we further extend the detector based on isotropic and anisotropic Gaussian kernel in multiscale space to obtain the multiscale AESMs. We also propose a fusion scheme for multiscale AESMs in Sect. 3. Experimental results on example images in the EUSFLAT Edge Detection Competition dataset are given in Sect. 4, while the conclusions are presented in Sect. 5.

2 Related Work

Shui and Zhang [9] proposed an edge detector based on an IAGK. It attains noise reduction while maintaining a high edge resolution. It also employs a directional anisotropic Gaussian (DAG) kernel, which can be written as

$$g(\mathbf{x};\sigma,\rho,\theta)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{R}_{-\theta}\begin{bmatrix}\rho^2 & 0\\ 0 & \rho^{-2}\end{bmatrix}\mathbf{R}_{\theta}\mathbf{x}\right), \tag{1}$$

with

$$\mathbf{R}_\theta=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix},$$

where the superscript T denotes the transpose, $\mathbf{x}=[x,y]^T$ is the two dimensional planar coordinate, $\sigma^2(\sigma>0)$ represents the scale, $\theta\in[0,\pi[$ denotes the orientation and $\rho\geq 1$ is the anisotropy factor. Note that this kernel reduces to an isotropic Gaussian kernel when $\rho=1$.

From Eq. (1), the directional derivative of the anisotropic Gaussian kernel is derived as

$$g'(\mathbf{x};\sigma,\rho,\theta) = -\frac{\rho^2[\cos\theta,\sin\theta]\mathbf{x}}{\sigma^2}g(\mathbf{x};\sigma,\rho,\theta)\,. \tag{2}$$

Since a single directional derivative of the anisotropic Gaussian kernel can only extract the intensity variation along the specific orientation, a family of kernels combining all possible orientations needs to be used to measure the intensity variation in all orientations. Because of the effect of edge stretch, the ESMs obtained by both isotropic and anisotropic Gaussian kernels should be fused to produce the final ESM.

3 Proposed Method

We first modify the DAG kernel in Eq. (1) by employing a new anisotropy factor φ instead of the original ρ:

$$g(\mathbf{x};\sigma,\varphi,\theta) = \frac{1}{2\pi\varphi\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{R}_{-\theta}\begin{bmatrix}1 & 0\\ 0 & \varphi^{-2}\end{bmatrix}\mathbf{R}_\theta\mathbf{x}\right), \tag{3}$$

where $\varphi \geq 1$ and the other arguments are the same as Eq. (1). Note that the function in Eq. (3) is different from the corresponding forms in [6,9,10]. Equation (3) can better show the relationship between the scales in the x and y direction, while the conventional ones bind the two scales together with the ρ in Eq. (1). Thus the formula in Eq. (3) can be used more explicitly in multiscale space. Similarly, the directional derivative of the anisotropic Gaussian kernel is derived from Eq. (3) as follows

$$g'(\mathbf{x};\sigma,\varphi,\theta) = -\frac{[\cos\theta,\sin\theta]\mathbf{x}}{\sigma^2}g(\mathbf{x};\sigma,\varphi,\theta)\,. \tag{4}$$

Note that the revised directional derivative of the anisotropic Gaussian kernel in Eq. (4) is also different from Eq. (2). Conventionally, in order to reduce the effect of scale decay, the edge detection filter is normalized by making the sum of the positive and negative part of the kernel be 1, respectively. Since the sum of the positive part of the kernel is equal to that of the negative part, and the sum of each part can be calculated as

$$C = \int_0^{+\infty}\left(\int_{-\infty}^{+\infty} g'(\mathbf{x};\sigma,\varphi,\theta)\,\mathrm{d}y\right)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma}\,, \tag{5}$$

the normalized directional derivative of the anisotropic Gaussian kernel is given by

$$\hat{g}'(\mathbf{x};\sigma,\varphi,\theta) = -\sqrt{2\pi}\frac{[\cos\theta,\sin\theta]\mathbf{x}}{\sigma}g(\mathbf{x};\sigma,\varphi,\theta)\,. \tag{6}$$

Considering that a digital image is a discrete signal in the integer lattice, the directional derivative of the anisotropic Gaussian kernel should be discretized. We can get both the discrete Gaussian function and the normalized directional derivative of the anisotropic Gaussian kernel by sampling the formulae in Eqs. (3) and (6) in the 2D integer coordinate $\mathbb{Z}^2$:

$$g(\mathbf{m};\sigma_j,\varphi_k,\theta_p)=\frac{1}{2\pi\varphi_k\sigma_j^2}\exp\left(-\frac{1}{2\sigma_j^2}\mathbf{m}^T\mathbf{R}_p^T\begin{bmatrix}1 & 0\\ 0 & \varphi_k^{-2}\end{bmatrix}\mathbf{R}_p\mathbf{m}\right)$$
$$\hat{g}'(\mathbf{m};\sigma_j,\varphi_k,\theta_p)=-\sqrt{2\pi}\frac{[\cos\theta_p,\sin\theta_p]\mathbf{m}}{\sigma_j}g(\mathbf{m};\sigma_j,\varphi_k,\theta_p)\,, \tag{7}$$

with

$$\mathbf{R}_p=\begin{bmatrix}\cos\theta_p & \sin\theta_p\\ -\sin\theta_p & \cos\theta_p\end{bmatrix},$$
$$\theta_p=(p-1)\pi/P,\quad p=1,2,3,...,P,$$

where $\mathbf{m}=[m,n]^T\in\mathbb{Z}^2$ represents the 2D image coordinate, $\sigma_j\in\mathbf{\Sigma}\ \{\sigma_1,\sigma_2,...,\sigma_J\}$ with $0<\sigma_1<\sigma_2<...<\sigma_J$, $\varphi_k\in\mathbf{\Phi}\ \{\varphi_1,\varphi_2,...,\varphi_K\}$ with $\varphi_k\geq 1$ and $\theta_p\in\mathbf{\Theta}\ \{\theta_1,\theta_2,...,\theta_P\}$ with $\theta_p\in[0,\pi[$ denote the square root of scale, anisotropy factor and orientation, respectively.

Therefore, a family of discrete normalized directional derivative of the anisotropic Gaussian kernels combining all possible scales, anisotropy factors and orientations are used to filter the image to obtain all the ESMs:

$$E(\mathbf{m};\sigma_j,\varphi_k,\theta_p)=\left|I*g'(\mathbf{m};\sigma_j,\varphi_k,\theta_p)\right|\,, \tag{8}$$

where $*$ denotes the convolution operation, $|\ |$ represents the absolute value and I stands for the intensity image.

Shui and Zhang [9] has pointed that the directional derivative of the anisotropic Gaussian kernel produces edge stretch effect while the isotropic one does not. It means that the anisotropic edge filters can produce large responses for edges and some non-edge structures. We also note that compared with anisotropic filters, isotropic ones produce smaller but comparable ESM for edge structures. From the definition of geometric mean we can infer that if one number in the list is nearly zero, the geometric mean of the list is nearly zero. If all the number in the list are large, the geometric mean of the list is also large. Hence, in order to reduce the edge stretch effect of anisotropic ESM while enhancing the isotropic ESM, the isotropic ESM and anisotropic ESM are fused using their geometric mean. As a result, the ESM for a specific scale is given by:

$$E_s(\mathbf{m};\sigma_j)=\sqrt{E_m(\mathbf{m};\sigma_j)\cdot E_i(\mathbf{m};\sigma_j)}\,, \tag{9}$$

with

$$E_m(\mathbf{m};\sigma_j)=\max_{\varphi_k\in\mathbf{\Phi}}\ \max_{\theta_p\in\mathbf{\Theta}}\ E(\mathbf{m};\sigma_j,\varphi_k,\theta_p)$$
$$E_i(\mathbf{m};\sigma_j)=\max_{\theta_p\in\mathbf{\Theta}}\ E(\mathbf{m};\sigma_j,1,\theta_p)\,. \tag{10}$$

With respect to the fusion of multiscale ESMs, we note several properties of the multiscale edge filters. First, small scale edge filters can produce large responses for almost all edges. But their responses for valleys between neighboring edges are nearly zero. Secondly, compared with small scale edge filters, some selected large scale filters can produce larger responses for not only edges but also valleys between neighboring edges. Thus, the maxima of the multiscale ESMs show large values for both edges and some valleys. Besides, if we set the minimum of the scale set $\mathbf{\Sigma}$ in Eq. (7) large enough, most of the noise points and thin textures can be suppressed. But thin edges will be maintained because the anisotropic edge filters have a high edge resolution. Considering the properties of geometric mean again, we also fuse the multiscale ESMs using the geometric mean. In this way, the response of valleys can be suppressed while the response of edge structures are maintained. Thus the fused multiscale ESMs is obtained by

$$E_f(\mathbf{m}) = \sqrt{E_m(\mathbf{m}) \cdot E_i(\mathbf{m}; \sigma_1)}\,, \tag{11}$$

with

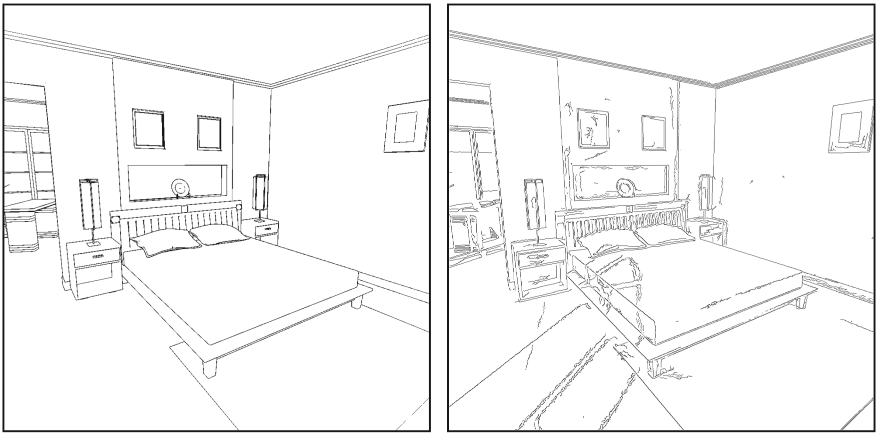
$$E_m(\mathbf{m}) = \max_{\sigma_j \in \mathbf{\Sigma}} \ E_s(\mathbf{m}; \sigma_j)\,. \tag{12}$$

In the procedure of the postprocessing, non-maxima suppression [8] is first applied to the fused ESM $E_f(\mathbf{m})$. For each pixel, the fused ESM and direction are used in the non-maxima suppression procedure. If the fused ESM at either side of the current position is less than the fused ESM at the pixel in the gradient direction, the pixel intensity is retained; otherwise, the intensity of the current position would be set to zero.

Subsequently, double threshold and hysteresis segmentation [2] are employed to produce the binary edge map. The hysteresis processing requires an upper threshold and a lower threshold. The upper threshold is specified in terms of the percentile of the histogram of the contrast equalized edge map while the ratio of the lower threshold to the upper threshold is selected as 0.4. Segmentation of edge pixels is realized in two steps. All the pixels in the maxima set whose ESM exceeds the upper threshold are first labelled as strong edge pixels. The pixels whose ESM between the upper threshold and the lower threshold are labelled as weak edge candidates. For each of those candidates, if there is a path to connect it with a strong edge pixel, it is considered as a real weak edge pixel. Hence both the strong and weak edge pixels form the binary edge map.

4 Experimental Validation

To evaluate the performance of the proposed scheme, we apply our method to example images, which are referred to as *Ex1*, *Ex2* and *Ex3*, in the EUSFLAT Edge Detection Competition dataset. They all have the resolution of $1024 \times 1024 \times 3$. We also compare our method with two selected edge detectors, i.e., the widely used Canny method [2] and the state-of-the-art IAGK based method [9]. For a fair comparison, the same postprocessing procedures are applied to all the



(a) (b)

Fig. 1. The edge ground truth of image *Ex3* and the detection result of the proposed method. (a) Ground truth; (b) Edge detection result.

methods. In addition, experiments are conducted with either the recommended parameters mentioned in the original papers or optimally tuned ones. In this paper we adopt the widely used F-measure to get a quantitative performance evaluation:

$$F_\eta = \frac{\text{PREC} \cdot \text{REC}}{\eta\text{PREC} + (1-\eta)\text{REC}} \tag{13}$$

with

$$\text{PREC} = \frac{\text{TP}}{\text{TP}+\text{FP}} \quad \text{and} \quad \text{REC} = \frac{\text{TP}}{\text{TP}+\text{FN}}, \tag{14}$$

where TP, FP and FN are the percentages of true positives, false positives and false negatives, respectively, and η is a control parameter with the value of 0.5 in this paper.

As shown in Table 1, the proposed method outperforms the Canny method and IAGK based method in all cases. The acceptable values of F-measure mean that most of the edges have been extracted. However, in Fig. 1 we can note that shadows usually lead to spurious edges. This is because they also have significant local intensity variation, i.e., they also have the properties of edges. We should note that this kind of spurious edges is subjective to users, since some other users may consider them as real edges.

Table 1. Values of the F-measure for different selected methods and example images.

Method	*Ex1*	*Ex2*	*Ex3*
Canny	0.8759	0.8795	0.8047
IAGKs	0.8669	0.9075	0.8035
Proposed	0.8857	0.9097	0.8227

5 Conclusions

In this paper, in order to fuse the ESMs in multiscale space, we have further extended the edge detector based on isotropic/anisotropic Gaussian kernels in multiple scale space to obtain the multiscale AESMs. Afterwards we proposed a fusion scheme of multiscale AESMs based on the geometric mean. Experimental results on example images in the EUSFLAT Edge Detection Competition dataset show that the proposed method outperforms the widely used Canny method and the state-of-the-art IAGK based method.

References

1. Bao, P., Zhang, L., Wu, X.: Canny edge detection enhancement by scale multiplication. IEEE Trans. Pattern Anal. Mach. Intell. **27**(9), 1485–1490 (2005)
2. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
3. Dollár, P., Zitnick, C.L.: Fast edge detection using structured forests. IEEE Trans. Pattern Anal. Mach. Intell. **37**(8), 1558–1570 (2014)
4. Li, Y., Wang, S., Tian, Q., Ding, X.: A survey of recent advances in visual feature detection. Neurocomputing **149**, 736–751 (2015)
5. Lopez-Molina, C., De Baets, B., Bustince, H., Sanz, J., Barrenechea, E.: Multiscale edge detection based on Gaussian smoothing and edge tracking. Knowl.-Based Syst. **44**, 101–111 (2013)
6. Lopez-Molina, C., Vidal-Diez De Ulzurrun, G., Baetens, J.M., Van Den Bulcke, J., De Baets, B.: Unsupervised ridge detection using second order anisotropic Gaussian kernels. Sig. Process. **116**, 55–67 (2015)
7. Perfilieva, I., Hodáková, P., Hurtík, P.: Differentiation by the F-transform and application to edge detection. Fuzzy Sets Syst. **288**, 96–114 (2016)
8. Rosenfeld, A., Thurston, M.: Edge and curve detection for visual scene analysis. IEEE Trans. Comput. **C–20**(5), 562–569 (1971)
9. Shui, P.L., Zhang, W.C.: Noise-robust edge detector combining isotropic and anisotropic Gaussian kernels. Pattern Recogn. **45**(2), 806–820 (2012)
10. Zhang, W., Zhao, Y., Breckon, T.P., Chen, L.: Noise robust image edge detection based upon the automatic anisotropic Gaussian kernels. Pattern Recogn. **63**, 193–205 (2017)

Fuzzy MCDA Without Defuzzification Based on Fuzzy Rank Acceptability Analysis

Boris Yatsalo[1(✉)] and Luis Martinez[2]

[1] Institute of Cybernetic Intelligent Systems, National Research Nuclear University MEPHI, Obninsk-Moscow, Russian Federation
yatsalo@gmail.com
[2] Department of Computer Science, University of Jaen, Jaén, Spain
martin@ujaen.es

Abstract. Multi-criteria decision analysis (MCDA) in the fuzzy environment needs not only in implementation of functions of fuzzy variables but also inevitably leads to ranking fuzzy quantities. The use of simplification and defuzzification methods at different stages of fuzzy MCDA (FMCDA) results in a loss of information and does not meet the concept of fuzzy decision analysis that "*the decision taken in the fuzzy environment must be inherently fuzzy*". In this contribution, a new approach to FMCDA is suggested, in which fuzzy criterion values and fuzzy weight coefficients are considered as fuzzy numbers (FNs) of a general type. Ranking alternatives is based on a novel methodological approach, fuzzy rank acceptability analysis (FRAA), for ranking FNs, whose use within FMCDA forms the fuzzy multicriteria acceptability analysis (FMAA) and implements a consistent approach to fuzzy decision analysis providing both ranking alternatives and the degree of confidence for each alternative to have the corresponding rank. Properties of FRAA ranking and integration of FRAA with a fuzzy extension of MAVT (FMAVT) as an example are considered and discussed along with the overestimation problem, which can arise when implementing FMCDA. The outcomes of FMAVT application for analysis of a multicriteria problem within the case study on land-use planning are considered and compared with the results by (classical) MAVT method.

Keywords: Multi-criteria decision analysis · Fuzzy numbers · Fuzzy preference relations · Ranking fuzzy numbers · Fuzzy acceptability analysis · Fuzzy MCDA · Fuzzy MAVT · Overestimation problem · Multicriteria land-use planning

1 Introduction

In many situations, uncertainties are related to imprecision and vagueness provided by experts' elicitation of objective values and subjective judgments [1–4]. In such cases, the use of fuzzy set theory within decision analysis and, specifically in multi-criteria decision analysis (MCDA), provides a suitable approach to handling three main MCDA problematiques: choice, ranking, and sorting [5, 6].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_50

The treatment and analysis of uncertainties by using fuzzy sets not only need dealing with fuzzy models and functions of fuzzy variables, but inevitably implies the comparison and ranking of fuzzy quantities (in most cases, ranking of fuzzy numbers, FNs). In fuzzy MCDA (FMCDA) [2, 4], ranking of fuzzy quantities is the key stage within all types of multicriteria problems. In addition, ranking of FNs is considered as one of the key problems of fuzzy sets theory [7] and has been extensively investigated in different directions [3, 7–11].

The management of uncertainties by fuzzy modeling implies that "*in situations where fuzzy sets are a suitable way of representing uncertainty, the decision taken must be inherently fuzzy*" [12]. Taking into account the indicated concept, this contribution aims at developing an approach to FMCDA providing ranking alternatives along with a fuzzy measure (degree of confidence) for each alternative to have the corresponding rank. This approach to FMCDA, fuzzy multicriteria acceptability analysis (FMAA), is based on original approach to ranking FNs, fuzzy rank acceptability analysis (FRAA) [13]. FRAA, as an analytical method, can be based on different fuzzy preference relations and extends the idea of stochastic multicriteria acceptability analysis, SMAA [17], which uses Monte Carlo simulations, and Probabilistic MAA, ProMAA [18].

This paper is structured as follows. Section 2 revises briefly different fuzzy concepts, including FNs, fuzzy preference relations, ranking FNs, introduces fuzzy rank acceptability indices (FRAIs) and FRAA approach to ranking FNs. Section 3 introduces the fuzzy multicriteria acceptability analysis (FMAA), considers the overestimation problem when implementing FMCDA, and introduces FMAVT–FMAA methods. The case study on land-use planning is considered in Sect. 4 along with comparison and discussion of the results by several MCDA methods. Eventually, Sect. 5 concludes this paper.

2 Preliminaries

This section fixes basic notions of fuzzy numbers (FNs), fuzzy preference relations, and fuzzy ranking, and introduces fuzzy rank acceptability analysis (FRAA), which is a basis for fuzzy multicriteria acceptability analysis (FMAA) in FMCDA.

2.1 Fuzzy Numbers: Comparison and Ranking

Definition 1. A FN Z is a convex normal and restricted fuzzy set in $\mathbb{R}$ with a continuous or upper-continuous membership function $\mu_Z(x)$.
It means that there exist two real numbers, c_1 and c_2, $c_1 \leq c_2$, such that:

$$Z = \{(x,\ \mu_Z(x)) :\ \mu_Z(x) > 0\ x \in (c_1,\ c_2),\ \mu_Z(x) = 0,\ x \notin [c_1,\ c_2]. \tag{1}$$

$\mathbb{F}$ denotes the set of FNs according to (1).

Remarks. If $c_1 = c_2 = c$, $Z = c$ is a singleton and $\mu_Z(c) = 1$. The condition $\mu_Z(c_1) = {}_Z(c_2) = 0$ is, strictly speaking, not necessary and is often used for convenience, stressing the most often usage of FNs in applications.

Definition 2. For FN $Z \in \mathbb{F}$ and $\alpha \in (0, 1]$. An α-cut of Z is defined as [19]:

$$Z_\alpha = [A_\alpha,\ B_\alpha] = \{x \in \mathbb{R} :\ \mu_Z(x) \geq \alpha\}. \tag{2}$$

For $\alpha = 0$, consider $[A_0, B_0] = [c_1, c_2]$ (according to (1)), then FN Z can be identified with the family of intervals

$$Z = \{[A_\alpha,\ B_\alpha]\} \quad (\alpha \in [0,\ 1]). \tag{3}$$

For FNs $Z_i = \{[A^i_\infty,\ B^i_\infty]\}$ and $Z_j = \{[A^j_\infty,\ B^j_\infty]\}$, according to standard fuzzy arithmetic [3, 22]:

$$Z_{ij} = Z_i - Z_j = \left\{\left[A^{ij}_\infty,\ B^{ij}_\infty\right]\right\} = \left\{\left[A^i_\infty - B^j_\infty,\ B^i_\infty - A^j_\infty\right]\right\}. \tag{4}$$

In this contribution, comparison of two FNs is based on their difference Z_{ij} by using fuzzy preference relation.

Definition 3. Fuzzy preference relation R is a fuzzy relation on $\mathbb{F} \times \mathbb{F} : R = ((Z_i,\ Z_j),\ \mu_R(Z_i,\ Z_j))$, where membership function $\mu_R(Z_i, Z_j)$ ($\in$ [0, 1]) indicates the degree of preference Z_i over Z_j.

An important property of a fuzzy preference relation R is *reciprocity*, often used in ranking methods [8, 15]:

$$\mu_R\left(Z_i,\ Z_j\right) = 1 - \mu_R\left(Z_j,\ Z_i\right). \tag{5}$$

Note that from (5): $\mu_R(Z, Z) = 0.5$.

Definition 4. Let R be a fuzzy preference relation on $\mathbb{F} \times \mathbb{F}$. For any $Z_i,\ Z_j \in \mathbb{F}$, their fuzzy ranking is defined as:

$$\begin{aligned} &Z_i \succcurlyeq Z_j \quad \text{if } \mu_R\left(Z_i,\ Z_j\right) \geq 0.5, \quad Z_i \succ Z_j \quad \text{if } \mu_R\left(Z_i,\ Z_j\right) > 0.5, \\ &\text{and} \quad Z_i \sim Z_j \quad \text{if } \mu_R\left(Z_i,\ Z_j\right) = 0.5. \end{aligned} \tag{6}$$

To simplify further presentations, the following denotations are also used for fuzzy preference relation $R(Z_i, Z_j) \equiv (Z_i \geq Z_j) \equiv (Z_j \leq Z_i)$:

$$\mu_{ij} = \mu_R\left(Z_i,\ Z_j\right) = \mu_R\left(Z_i \geq Z_j\right) = \mu_R\left(Z_j \leq Z_i\right). \tag{7}$$

Note also that the symbols $\geq,\ \leq$ used here for notational purposes are different from the symbols $\succcurlyeq,\ \preccurlyeq$, which are associated with ranking FNs.

Basic axioms or requirements, which specify the reasonable properties of ordering FNs by ranking method $\mathbb{M}$ have been presented in [9, 14, 16] (below, the 5 main axioms are considered):

A_1 (*reflexivity*): For any finite subset $\mathcal{A}$ of $\mathbb{S}$ and $A \in \mathcal{A}$, $A \succcurlyeq A$ by $\mathbb{M}$: on $\mathcal{A}$;

A_2 (*antisymmetry*): For an arbitrary finite subset $\mathcal{A}$ of $\mathbb{S}$ and $(A,\ B) \in \mathcal{A}^2$, if $A \succcurlyeq B$ and $B \succcurlyeq A$ by $\mathbb{M}$ on $\mathcal{A}$, we should have $A \sim B$ by $\mathbb{M}$ on $\mathcal{A}$;

A_3 (*transitivity*): For an arbitrary finite subset $\mathcal{A}$ of $\mathbb{S}$ and $(A,\ B,\ C) \in \mathcal{A}^3$, if $A \succcurlyeq B$ and $B \succcurlyeq C$ by $\mathbb{M}$ on $\mathcal{A}$, we should have $A \succcurlyeq C$ by $\mathbb{M}$ on $\mathcal{A}$;

A_4 (*distinguishability*): For an arbitrary finite subset $\mathcal{A}$ of $\mathbb{S}$ and $(A,\ B) \in \mathcal{A}^2$, if *inf* supp $(A) >$ *sup* supp (B), we should have $A \succcurlyeq B$ by $\mathbb{M}$ on $\mathcal{A}$:

A_5 (*absence of rank reversal*): Let $\mathbb{S}$ and $\mathbb{S}'$ be two arbitrary finite sets of FNs in which method $\mathbb{M}$ can be applied, and A and B are in $\mathbb{S} \cap \mathbb{S}'$, then: $A \succ B$ by $\mathbb{M}$ on $\mathbb{S}'$ iff $A \succ B$ by $\mathbb{M}$ on $\mathbb{S}$.

The main classes of fuzzy ranking methods (apart of linguistic approaches) have been presented in [3, 9, 10] and include *Defuzzification based ranking methods*; *ranking methods based on the distance to a reference set*; and *ranking methods based on pairwise comparison*.

In this contribution, the new fuzzy multi-criteria approach, FMAA (fuzzy multi-criteria acceptability analysis), is based on the FRAA (fuzzy rank acceptability analysis) that belongs to latter type of the indicated methods for ranking FNs and can use different preference relations. In our proposal, the Yuan's [8] fuzzy preference relation is considered.

2.2 Yuan's Fuzzy Preference Relation

The Yuan's fuzzy preference relation [8], $R_Y(Z_i, Z_j)$, with the fuzzy measure $\mu_Y(Z_i, Z_j)$, which presents the degree preference of FN Z_i over Z_j, is based on the assessing difference $Z_{ij} = Z_i - Z_j$, $Z_{ij} = \{[A_\alpha, B_\alpha]\}$. For evaluation of $\mu_Y(Z_i, Z_j)$, the value $S_Y^+(Z_{ij})$ is assessed by the expression

$$S_Y^+\left(Z_{ij}\right) = \int_0^1 (B_\alpha \theta(B_\alpha) + A_\alpha \theta(A_\alpha))d\alpha, \tag{8}$$

where $\theta(x)$ is the Heaviside function and can be interpreted as a "distance of the positive part of FN $Z = Z_{ij}$ to the axis OY". Adjusted "total area" under membership function of FN Z_{ij} is assessed as

$$S_Y\left(Z_{ij}\right) = S_Y^+\left(Z_{ij}\right) + S_Y^+\left(Z_{ji}\right). \tag{9}$$

For Yuan's fuzzy preference relation, $R_Y(Z_i, Z_j)$, $\mu_Y(Z_i, Z_j)$ is defined as follows:

$$\mu_{ij} = \mu_Y\left(Z_i,\ Z_j\right) = S_Y^+\left(Z_{ij}\right)/S_Y\left(Z_{ij}\right), \quad S_Y\left(Z_{ij}\right) > 0; \tag{10}$$

where S_Y^+ and S_R are assessed according to (8), (9); for singletons $Z_i = c_i$, $Z_j = c_j$, $c_{ij} = c_i - c_j$:

$$\mu_{ij} = \mu_R\left(Z_i,\ Z_j\right) = \left\{\left(1 \text{ if } c_{ij} > 0\right),\ \left(0 \text{ if } c_{ij} < 0\right),\ \left(0.5 \text{ if } c_{ij} = 0\right)\right\}. \tag{11}$$

Yuan's preference relation satisfies the reciprocity property (5). Consider (without proof) the following kind of *distinguishability* property, which is important for comparison of input and output FNs within fuzzy modeling.

Definition 5. For two FNs, $Z_i = [A^i_\alpha,\ B^i_\alpha]$, i = 1, 2, we use the denotation $Z_2 \trianglelefteq Z_1$ if

$$A^2_\alpha \leq A^1_\alpha \quad \textit{and} \quad B^2_\alpha \leq B^1_\alpha \quad \text{for all } \alpha \in [0,\ 1]. \tag{12}$$

Lemma 1. If $Z_2 \trianglelefteq Z_1$, then $Z_2 \preccurlyeq Z_1$ for Yuan's preference relation.

The Yuan's preference relation satisfies all the main axioms for ranking FNs, A_1–A_5 [10]. The detailed analysis of Yuan's preference relation and its comparison with the *integral* preference relation [20, 21], which is intransitive but can be considered as the closest to Yuan's one, is presented in [22].

2.3 Fuzzy Rank Acceptability Indices

Let $\mathbf{Z} = \{Z_i,\ i = 1, \ldots, n\} \subset \mathbb{F}$ be a finite set of FNs.

Definition 6. Fuzzy rank statement (FRS) is a statement of the type:

$$F_{ik} = (Z_i,\ k) = \{Z_i \text{ has a rank } k\}, \quad i,\ k = 1, \ldots, n, \tag{13}$$

i.e., $k - 1$ FNs from the set $\mathbf{Z}$ can be considered as exceeding Z_i in the sense of the chosen ranking, and Z_i can be considered as exceeding remaining $n - k$ FNs. The following model is suggested for formalization of FRSs [20, 21].

First, FRS, F_{ik}, is considered as an element of the fuzzy set $\mathbf{Z} \times \mathbb{N}_n$, $\mathbb{N}_n = \{1, \ldots, n\}$; the membership function on the fuzzy set $\mathbf{Z} \times \mathbb{N}_n$, $\mu(F_{ik})$, is provided below. Second, the following model is suggested for formalization of FRSs with the use of fuzzy relation $R(Z_i, Z_j) \equiv (Z_i \geq Z_j) \equiv (Z_j \leq Z_i)$:

$$F_{i1} = \left\{ \mathop{\wedge}_{j \neq i}^{n} (Z_i \geq Z_j) \right\}, \tag{14}$$

$$F_{i2} = \mathop{\vee}_{l \neq i}^{n} \left((Z_i \leq Z_l) \mathop{\wedge}_{j \neq i,\, j \neq l}^{n} (Z_i \geq Z_j) \right), \tag{15}$$

$$F_{ik} = \left\{ \mathop{\vee}_{\substack{(l_1 < l_2 < \cdots < l_{k-1}) \\ l_s \neq i,\ s = 1, \ldots, k-1}} \left(\left(\mathop{\wedge}_{s=1}^{k-1} (Z_i \leq Z_{l_s}) \mathop{\wedge}_{\substack{j \neq i,\ j \neq l_s, \\ s = 1, \ldots, k-1}}^{n} (Z_i \geq Z_j) \right) \right) \right\}, \tag{16}$$

$$F_{in} = \left\{ \mathop{\wedge}_{j \neq i}^{n} (Z_i \leq Z_j) \right\}. \tag{17}$$

Definition 7. Fuzzy rank acceptability index (FRAI), $\mu(i, k)$, is defined as a fuzzy measure of FRS F_{ik}, $\mu(i, k) = \mu_R(F_{ik})$, i.e., an index (degree of confidence), which describes acceptability of FN Z_i with a rank k (for a given fuzzy preference relation R).

FRAIs can be assessed based on the classical approach to assessing the measure of logical expressions (14)–(17) and the notations (7) [20, 21]:

$$\mu(i,\ 1) = \mu(F_{i1}) = \mathop{\wedge}_{j\neq i}^{n} \mu_{ij}, \tag{18}$$

$$\mu(i,\ 2) = \mu(F_{i2}) = \mathop{\vee}_{l\neq i}^{n} \left(\mu_{li} \mathop{\wedge}_{j\neq i,\, j\neq l}^{n} \mu_{ij} \right), \tag{19}$$

$$\mu(i,\ k) = \mu(F_{ik}) = \mathop{\vee}_{\substack{(l_1 < l_2 < \cdots < l_{k-1}) \\ l_s \neq i,\ s=1,\ldots,k-1}} \left(\left(\mathop{\wedge}_{s=1}^{k-1} \right) \mu_{l_s i} \wedge \left(\mathop{\wedge}_{\substack{j\neq i,\, j\neq l_s, \\ s=1,\ldots,k-1}}^{n} \mu_{ij} \right) \right), \tag{20}$$

$$\mu(i,\ n) = \mu(F_{in}) = \mathop{\wedge}_{j\neq i}^{n} \mu_{ji}. \tag{21}$$

Note that the model for FRSs are consistent with (mathematical) intuition and can be further processed within fuzzy and probabilistic approaches to their evaluation.

2.4 Fuzzy Rank Acceptability Analysis

The fuzzy rank acceptability analysis (FRAA) is an approach to ranking FNs that represents a systematical implementation of the concept of fuzzy decision analysis that "the decision taken in the fuzzy environment must be inherently fuzzy" [12].

Let $\mathbf{Z} = \{Z_i,\ i = 1,\ldots,n\} \subset \mathbb{F}$ be a finite set of FNs and $R = R(Z_i,\ Z_j)$ a fuzzy preference relation with a membership function $\mu_{ij} = \mu_R(Z_i,\ Z_j)$. Using FRAIs $\mu(i, k) = \mu_R(F_{ik})$, the matrix $\mathbf{M} = \{\mu(i, k)\}$, $i, k = 1,\ldots,n$, can be built and used by decision makers to choose the most appropriate FNs/alternatives within a decision problem.

The analysis of matrix **M** by using a simple implementation of the FRAA approach within a choice problem is similar to the use of stochastic, SMAA [17], and probabilistic ProMAA [18], MCDA methods. Here, a systematic approach to FRAA is considered.

Definition 8. A FN Z_m from a set $\mathbf{Z} = \{Z_i,\ i = 1,\ldots,n\} \subset \mathbb{F}$ has a $FRAA_R$ rank k, denoted by $r(Z_m) = k$, if it has the maximal FRAI for that k:

$$\mu_R(m,\ k) = \max_{i=1,\ldots,n} \mu_R(i,\ k), \quad 1 \leq m, \quad k \leq n. \tag{22}$$

The following properties of FRAA ranking based on Yuan's preference relation, $FRAA_Y$, have the place [22].

Lemma 2. Let for a set of FNs $\mathbf{Z} = \{Z_i,\ i = 1,\ldots,n\}$ and preference relation R, $Z_n \prec Z_{n-1} \prec \cdots \prec Z_1$ without violation of transitivity (i.e., for all $1 \leq k < m \leq n$, $Z_m \prec Z_k$), then, Z_k has the $FRAA_R$ rank k: $r(Z_k) = k$, $k = 1,\ldots,n$.

Proposition 1. $FRAA_Y$ ranking and ranking by Yuan's preference relation coincide, and $FRAA_Y$ ranking satisfies axioms A_1–A_5.

Despite the results of ranking FNs by $FRAA_Y$ and by Yuan's preference relation R_Y coincide, FRAA provides a measure/degree of confidence for each rank based on FRAIs ensuring a greater clarity and reliability within ranking FNs and presents a consistent approach to fuzzy ranking and fuzzy decision analysis.

To demonstrate the possibility and advantages of FRAA ranking, the following example is provided. Consider the sets RS_1, RS_2, and RS_3 (relatively stable ranking) of triangular FNs (23)–(25) (Fig. 1).

$$RS_1 : Z_1 = (1, 2, 6), \quad Z_2 = (0, 3, 4.3), \quad Z_3 = (0.5, 2.5, 4.5), \tag{23}$$

$$RS_2 : Z_1 = (1, 2, 4.7), \quad Z_2 = (0, 3, 4.3), \quad Z_3 = (0.5, 2.5, 4.5), \tag{24}$$

$$RS_3 : Z_1 = (1, 2, 4.7), \quad Z_2 = (0, 2.71, 4.3), \quad Z_3 = (0.5, 2.5, 4.5). \tag{25}$$

Fig. 1 Fuzzy numbers of the set RS_2.

In Table 1, ranking FNs by $FRAA_Y$ and centroid index, *CI* (defuzzification based method) [3, 9] is presented. According to Table 1, ranks of FNs for the three sets by $FRAA_Y$ completely differ from *CI* ranking method, which is, in contrast to $FRAA_Y$ ranking, insensitive to changing FNs of the sets RS_1, RS_2, and RS_3.

Table 1 Ranking FNs of the sets RS_1, RS_2, and RS_3 (27)–(29) by the two methods

Ranking method	RS_1: Ranks			RS_2: Ranks			RS_3: Ranks		
	1	2	3	1	2	3	1	2	3
$FRAA_Y$	Z_1 0.536	Z_2 0.517	Z_3 0.517	Z_2 0.517	Z_3 0.517	Z_1 0.519	Z_3 0.516	Z_2 0.501	Z_1 0.501
CI	Z_1	Z_3	Z_2	Z_1	Z_3	Z_2	Z_1	Z_3	Z_2

3 Fuzzy Multicriteria Acceptability Analysis

The concept of fuzzy multicriteria acceptability analysis (FMAA) is the implementation of FRAA within FMCDA. Such an approach to FMCDA provides both ranking alternatives and degree of confidence for each rank based on FRAIs.

FMAA can be implemented in different FMCDA methods. In this proposal, integration of FMAA with MAVT, FMAVT–FMAA method, is introduced.

In the beginning, consider implementation of FMAA for an FMCDA method in a general case and discuss the problem of overestimation [26], which can arise at some steps of FMCDA models application.

3.1 FMCDA and the Overestimation Problem

In a general case, an MCDA/MADM method is based on evaluation of an entire/overall criterion V and can be presented by a model

$$V(\boldsymbol{a}_i) = F(\boldsymbol{a}_i,\ \boldsymbol{w},\ \boldsymbol{a},\ \boldsymbol{p});\quad (\boldsymbol{a},\ \boldsymbol{w},\ \boldsymbol{p}) \in \boldsymbol{U}, \tag{26}$$

here $F(.)$ is a function for determination of the overall criterion V, $\boldsymbol{a} = (\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n)$ is a vector of alternatives, $\boldsymbol{a}_i = (C_{i1},\ldots,C_{im})$, C_{ik} is the criterion value of the alternative $\boldsymbol{a}_i$ for criterion k, $i = 1,\ldots,n$, $k = 1,\ldots,m$; $\boldsymbol{w} = (w_1,\ldots,w_m)$ presents a vector of weight coefficients, $\boldsymbol{p}$ is a set of model parameters, and $\boldsymbol{U}$ is a set of restrictions and coupling.

Integration of FMAA with the fuzzy extension of MCDA model (26), where criterion values and weight coefficients are FNs, is based on implementation of FRAA for the set of FNs $\boldsymbol{Z} = \{Z_i = V(\boldsymbol{a}_i),\ i = 1,\ldots,n\}$. Corresponding differences of overall values/FNs, $Z_{ij} = Z_i - Z_j$, are presented by the expression

$$Z_{ij} = F(\boldsymbol{a}_i,\ \boldsymbol{w},\ \boldsymbol{a},\ \boldsymbol{p}) - F\left(\boldsymbol{a}_j,\ \boldsymbol{w},\ \boldsymbol{a},\ \boldsymbol{p}\right). \tag{27}$$

The following methodological approach to implementation of FMAA based on the expression (27) should be pointed out.

Each weight coefficient w_k, $k = 1,\ldots,m$ (and, in a general case, criterion values of alternatives) occurs in the expression (27) at least two times. Thus, due to dependent variables/FNs in the components of the formula (27), the overestimation problem should be taken into account [23].

To present briefly the overestimation problem within fuzzy modeling, consider the following expressions (here all the variables are positive FNs):

$$Z_O = wa - wb,\quad Z_T = w(a - b), \tag{28}$$

Analysis of the indicated formulas (O and T mean overestimation and transformation correspondingly), leads to the following outcome: in a general case, $supp(Z_T) \subset supp(Z_O)$. Thus, model Z_O leads to an overestimation.

For fuzzy extensions of classical additive MAVT model [24], the transformation of the source expression as for Z_T in (28) can be implemented. For some other MCDA methods with more complicated expressions $F(.)$ (26), e.g., for TOPSIS and PROMETHEE [6], there is no such a simple transformation, and corresponding transformation method(s) (reduced or/and general/extended transformation methods) [23] should be used.

There may be two points of view concerning implementation and justification of the two approaches within FMCDA:

– Model-O, when no approach to avoid overestimations is used, and

– Model-T, where transformations of the source expression(s) at the steps of FMAA implementation (as for the case (28)) or/and a transformation method(s) [23] are implemented.

Authors recommend using Model-T approach. For research purposes, taking into account pros and cons concerning Model-O and Model-T approaches, we implement below both models within the case study analysis.

3.2 Implementation of FMAA in FMCDA

Implementation of FMAA within FMCDA is considered here on a fuzzy extension of additive MAVT (multi-attribute value theory) model [5, 6, 24] as an example.

For additive MAVT, expression (26) has the following form:

$$V(\boldsymbol{a}_i) = \sum\nolimits_{k=1}^{m} w_k V_k(a_{ik}); \quad w_k > 0, \quad \sum\nolimits_{k=1}^{m} w_k = 1, \tag{29}$$

where $V_k(x)$ is a partial value function, and w_k is a weight coefficient (reflecting a scaling factor) for criterion k, $k = 1,\ldots,m$; $V(\boldsymbol{a}_i)$ is interpreted as an overall value of the alternative $\boldsymbol{a}_i$, $i = 1,\ldots,n$, within additive MAVT.

Implementation of FNs, fuzzy logic, linguistic variables and other methods of fuzzy sets theory has been widely used in fuzzy MCDA, see, e.g., an comprehensive survey [25]. However, there are only several works, where MAUT/MAVT models in a fuzzy environment are mentioned [26, 27], and no consistent extensions of fuzzy MAVT/MAUT are discussed.

In FMAVT, criterion values for alternative i and criterion k, a_{ik}, and weight coefficients, w_k, are considered as FNs of a general type, $i = 1,\ldots,n$, $k = 1,\ldots,m$, and overall (fuzzy) value $V(\boldsymbol{a}_i)$ is determined according to (29) based on implementation of functions of FNs.

The recommended approach for assigning fuzzy weight coefficients w_j in FMAVT is a modification [23] of the *Swing* weighting (scaling) process for MAVT [5, 24], F-Swing.

Ranking FNs $\boldsymbol{Z} = \{Z_i = V(\boldsymbol{a}_i), i = 1,\ldots,n\}$, where $\boldsymbol{A} = \{\boldsymbol{a}_i\}$ is a set of alternatives, can be implemented by different ranking methods, including defuzzification ones [3, 9]. Here, FMAVT is extended to FMAVT–FMAA based on implementation of FRAA to the set of FNs $\boldsymbol{Z}$.

FRAA approach (22), (18)–(21) is based on assessing fuzzy measures $\mu_{ij} = {}_R(Z_i, Z_j)$ (7), where μ_{ij} is a function of the difference $Z_{ij} = Z_i - Z_j$. In this contribution, Yuan's fuzzy preference relation R is used according to (10). There may be two distinct approaches to determination of Z_{ij} within FMAVT–FMAA.

Within Model-O approach, $\{Z_i = V(\boldsymbol{a}_i), i = 1,\ldots,n\}$ are assessed according to (29), and Z_{ij}, $i, j = 1,\ldots,n$ is determined directly as $Z_{ij} = Z_i - Z_j$:

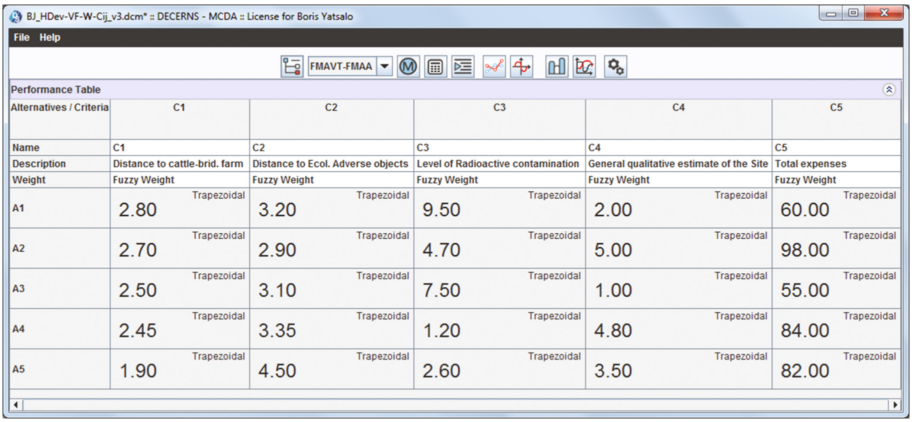
$$Z_{ij}(O) = \left(\sum_{k=1}^{m} w_k V_k(a_{ik})\right) - \left(\sum_{k=1}^{m} w_k V_k(a_{jk})\right) \tag{30}$$

For Model-O, as in the case of expression (28), overestimation has the place. Within Model-T approach, Z_{ij} is determined as

$$Z_{ij}(T) = \sum_{k=1}^{m} w_k \left(V_k(a_{ik}) - V_k(a_{jk}) \right). \tag{31}$$

4 Application of FMAVT–FMAA to a Case Study on Land-Use Planning

In this section, FMAVT–FMAA is implemented in a case study on multicriteria land-use planning (the choice of an area for housing development) [28]. Within this case study, 11 criteria are used. On the first stage, taking into account a set of constrains imposed by experts/stakeholders (distance to rivers, lakes, roads, forest, household centers, and some other objects of the vector map of land-use), a conjunctive screening process based on six criteria is implemented (with the use of GIS-functions for building buffer zones and overlays) to exclude inappropriate areas. On the second stage, five criteria (C_1 - distance from stockyards (*maximize*), C_2 - distance from ecologically adverse objects (*maximize*), C_3 - level of radioactive contamination (*minimize*), C_4 - general qualitative assessment of the local landscape/site quality (*maximize*), C_5 – total expenses (*minimize*)) are used within the MCDA problem created by experts [28]. The performance table and weight coefficients by F-Swing method are represented in Figs. 2 and 3. The results of ranking alternatives for FMAVT–FMAA and MAVT (29) are presented in Table 2.



Alternatives / Criteria	C1	C2	C3	C4	C5
Name	C1	C2	C3	C4	C5
Description	Distance to cattle-brid. farm	Distance to Ecol. Adverse objects	Level of Radioactive contamination	General qualitative estimate of the Site	Total expenses
Weight	Fuzzy Weight	Fuzzy Weight	Fuzzy Weight	Fuzzy Weight	Fuzzy Weight
A1	Trapezoidal 2.80	Trapezoidal 3.20	Trapezoidal 9.50	Trapezoidal 2.00	Trapezoidal 60.00
A2	Trapezoidal 2.70	Trapezoidal 2.90	Trapezoidal 4.70	Trapezoidal 5.00	Trapezoidal 98.00
A3	Trapezoidal 2.50	Trapezoidal 3.10	Trapezoidal 7.50	Trapezoidal 1.00	Trapezoidal 55.00
A4	Trapezoidal 2.45	Trapezoidal 3.35	Trapezoidal 1.20	Trapezoidal 4.80	Trapezoidal 84.00
A5	Trapezoidal 1.90	Trapezoidal 4.50	Trapezoidal 2.60	Trapezoidal 3.50	Trapezoidal 82.00

Fig. 2 Performance table: housing development case study

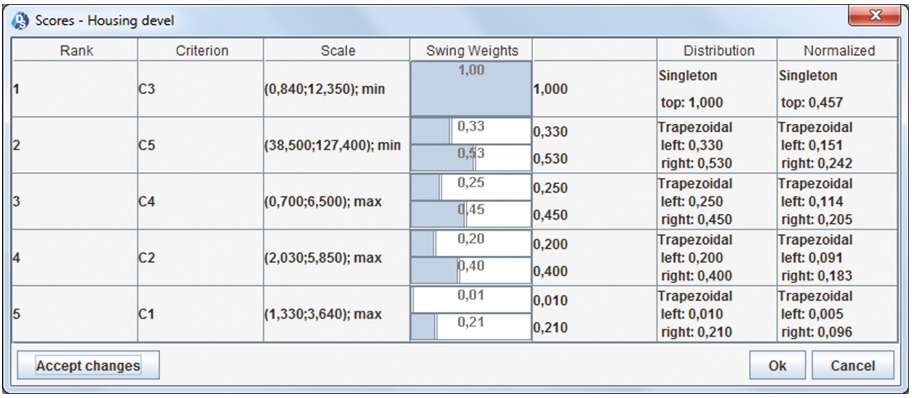


Fig. 3 Setting criterion weights with the use of F-Swing method

Table 2 Ranking alternatives by FMAVT–FMAA-T/O methods and MAVT

Method/alternative	A1	A2	A3	A4	A5
FMAVT–FMAA-T	5 (0.622)	3 (0.634)	4 (0.622)	1 (0.506)	2 (0.506)
FMAVT–FMAA-O	5 (0.602)	3 (0.578)	4 (0.602)	2 (0.508)	1 (0.508)
MAVT	5	3	4	2	1

Within this case study, criterion values and weight coefficients are considered as trapezoidal FNs symmetrical regarding the mean value (centroid index [3]), which is used as corresponding value in MAVT. For FMAVT–FMAA, the Model-O and Model-T are assessed along with the FRAIs (as the degree of confidence that corresponding alternative has rank k). Non-linear partial value functions for all the criteria are used.

For the considered case study and different models suggested for its analysis, ranks of alternatives according to MAVT and FMAVT–FMAA as well for Model-O and Model-T coincide for alternatives A1, A2, and A3, and differ for A4 and A5. However, the degrees of confidence for ranks of alternatives A4 and A5 are close to 0.5, and it means these alternatives may be considered as equivalent with the ranks 1 and 2.

The indicated approaches to multicriteria analysis (MAVT and FMAVT–FMAA-T/O) are, in fact, different and can lead to different outcomes. Further analysis of the case study on housing development at the stage of sensitivity analysis demonstrates difference in ranking alternatives by Model-O and Model-T when changing partial value functions for FMAVT–FMAA.

Comparison of the output results according to MAVT and FMAVT (in general, by a classical MCDA method and its "fuzzy analogue") presents a methodological problem. There may be several approaches/recommendations to "solve" or analyze such a problem. One of them is discussed below.

In the case of FMAVT and MAVT, the mean value (centroid index) for fuzzy weight coefficients, used in FMAVT, can be used in MAVT (this may be justified within swing approach for most cases, when the change of ranges for all criterion values may be considered as insignificant); partial value functions for MAVT can be built not in local, but in a global scale/range [3] taking into account the range of left and right points for *supports* of fuzzy criterion values.

- The developed version of FMAVT–FMAA model and corresponding computer system has the following features:
- Experts can set for criterion values and weight coefficients the following types of fuzzy numbers: singleton, triangular, trapezoidal, and piecewise-linear (the last type of FNs can be used for approximation of any input FN with non-linear membership function);
- Different types of partial value functions may be used: linear, piecewise-linear, and non-linear/exponential;
- Implementation of value function sensitivity analysis (changing value function by mouse with effective representation of changing the matrix of FRAIs and ranking alternatives);
- Graphical and table demonstration of the output results.

After analysis of this and several other scenarios, we can conclude that FMAVT–FMAA approach demonstrates adequate results for examined case studies and may be considered as a validated MCDA method for uncertainty treatment and analysis.

5 Conclusions

This contribution presents a consistent approach to ranking both FNs and alternatives within fuzzy MCDA (FMCDA), which is in full agreement with the basic concept of fuzzy decision analysis that *in fuzzy environment decision taken should be inherently fuzzy*.

Original approach to ranking FNs, FRAA (fuzzy rank acceptability analysis), which is based on fuzzy rank acceptability indices (FRAIs), allows using different fuzzy preference relations for ranking FNs and provides a degree of confidence for each rank obtained, and FMAA (fuzzy multicriteria acceptability analysis) as an embedding of FRAA into MCDA have been introduced and discussed.

FMAA has been integrated with the fuzzy extension of MAVT into FMAVT–FMAA method, which allows using FNs of a general type for criterion values and weight coefficients, and no defuzzification method is implemented within FMAVT–FMAA. Comparison of FMAVT–FMAA outputs with results for classical MAVT method has been discussed along with the overestimation problem, which arises within FMCDA.

FRAA/FMAA concept is in agreement with human intuition and has a rational and consistent interpretation with corresponding analogues in probability theory. In addition, FRAIs can be rapidly calculated without complexity for users/experts even if they do not go into the details of the algorithms.

The use of FRAA for ranking FNs and its implementation in MCDA can be fruitful for uncertainty treatment and decision analysis as well as an adjustment of FRAA concept to other types of fuzzy sets.

References

1. Bellman, R.E., Zadeh, L.A.: Decision-making in a fuzzy environment. Manag. Sci. **17**, 141–163 (1970)
2. Chen, S.J., Hwang, C.L.: Fuzzy Multiple Attribute Decision Making: Methods and Applications, vol. 375. Springer, Berlin (1992)
3. Wang, X., Ruan, D., Kerre, E.E.: Mathematics of Fuzziness Basic Issues. Springer, Berlin (2009)
4. Pedrycz, W., Ekel, P., Parreiras, R.: Fuzzy Multicriteria Decision-Making: Models, Methods and Applications. Wiley, Chichester (2010)
5. Belton, V., Stewart, T.: Multiple Criteria Decision Analysis: An Integrated Approach. Kluwer Academic Publishers, Dordrecht (2002)
6. Figueira, J., Greco, S., Ehrgott, M. (eds.): Multiple Criteria Decision Analysis. State of the Art Surveys. Springer, New York (2005)
7. Matarazzo, B., Munda, G.: New approaches for the comparison of LR fuzzy numbers: a theoretical and operational analysis. Fuzzy Sets Syst. **118**, 407–418 (2001)
8. Yuan, Y.: Criteria for evaluating fuzzy ranking methods. Fuzzy Sets Syst. **44**, 139–157 (1991)
9. Wang, X., Kerre, E.E.: Reasonable properties for the ordering of fuzzy quantities (I). Fuzzy Sets Syst. **118**(3), 375–385 (2001)
10. Wang, X., Kerre, E.E.: Reasonable properties for the ordering of fuzzy quantities (II). Fuzzy Sets Syst. **118**(3), 387–405 (2001)
11. Wang, Z.-X., Mo, Y.-N.: Ranking fuzzy numbers based on ideal solution. In: Cao, B.-Y., Zhang, C.-Y., Li, T.-F. (eds.) Fuzzy Information and Engineering. Advances in Soft Computing, vol. 54, pp. 201–209. Springer, Berlin (2009)
12. Tong, R.M., Bonissone, P.P.: A linguistic approach to decision making with fuzzy sets. IEEE Trans. Syst. Man Cybern. **10**, 716–723 (1980)
13. Yatsalo, L., Martinez, A.: Novel approach to ranking fuzzy numbers based on fuzzy acceptability analysis. In: Zeng, X., Lu, J., Kerre, E.E., Martinez, L., Koehl, L. (eds.) Proceedings of the 12th International FLINS2016 Conferences on Uncertainty Modelling in Knowledge Engineering and Decision Making, pp. 75–80. World Scientific, Roubaix (2016). doi:10.1142/9789813146976_0015. ISBN 978-981-316-96-9.24-26.08
14. Basu, K.: Fuzzy revealed preference theory. J. Econ. Theory **32**(2), 212–227 (1984)
15. Nakamura, Kazuo.: Preference relations on a set of fuzzy utilities as a basis for decision making. Fuzzy Sets Syst. **20**(2), 147–162 (1986)
16. Delgado, M., Verdegay, J.L., Vila, M.A.: A procedure for ranking fuzzy numbers using fuzzy relations. Fuzzy Sets Syst. **26**(1), 49–62 (1988)
17. Salminen, P., Lahdelma, R.: SMAA-2: stochastic multicriteria acceptability analysis for group decision making. Oper. Res. **49**(3), 444–454 (2001)
18. Yatsalo, B., Gritsyuk, S., Mirzeabasov, O.A., Vasilevskaya, M.: Uncertainty treatment within multicriteria decision analysis with the use of acceptability concept. Control Big Syst. **32**, 5–30 (2011). (in Russian)
19. Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. J. Syst. Sci. **9**(6), 613–626 (1978)
20. Yatsalo, B.: A new approach to fuzzy multi-criteria acceptability analysis. In: Alonso, J.M., Bustince, H., Reformat, M. (eds.), Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology. Series AISR, vol. 89, pp. 947–952 (2015). ISBN 978-94-62520-77-6. ISSN 1951-6851 (2015)

21. Yatsalo, B., Martinez, L.: A novel approach to ranking fuzzy numbers based on fuzzy acceptability analysis. In: Zeng, X., Lu, J., Kerre, E.E., Martinez, L., Koehl, L. (eds.), Proceedings of the 12th International FLINS2016 Conference on Uncertainty Modelling in Knowledge Engineering and Decision Making, pp. 75–80. World Scientific, Roubaix (2016). doi:10.1142/9789813146976_0015. ISBN 978-981-316-96-9.24-26.08
22. Yatsalo, B., Martinez, L.: Fuzzy multi-criteria acceptability analysis: a new approach to multi-criteria decision analysis under fuzzy environment. Expert Syst. Appl. **84**, 262–271 (2017)
23. Hanss, M.: Applied Fuzzy Arithmetic. Springer, Berlin (2005)
24. Keeney, R.L., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley, New York (1976)
25. Kahraman, C., Cevik Onar, S., Oztaysi, B.: Fuzzy multicriteria decision-making: a literature review. Int. J. Comput. Intell. Syst. **8**(4), 637–666 (2015)
26. Ashour, O.M., Okudan Kremer, Gl.E.: A simulation analysis of the impact of FAHP-MAUT triage algorithm on the emergency department performance measures. Expert Syst. Appl. **40** (1), 177–187 (2013)
27. Jimnez, A., Mateos, A., Sabio, P.: Dominance intensity measure within fuzzy weight oriented MAUT: an application. Omega Manag. Sci. Environ. Issues **41**(2), 397–405 (2013)
28. Yatsalo, B., Didenko, V., Tkachuk, A., Gritsyuk, S., Mirzeabasov, O., Slipenkaya, V., Babutski, A., Pichugina, I., Sullivan, T., Linkov, I.: Multi-criteria spatial decision support system DECERNS: application to land use planning. Int. J. Inf. Syst. Soc. Change **1**, 11–30 (2010)

A Portfolio of Minimum Risk in a Hybrid Uncertainty of a Possibilistic-Probabilistic Type: Comparative Study

Alexander Yazenin and Ilia Soldatenko(✉)

Tver State University, 33 Zhelyabova Str., Tver, Russia
Yazenin.AV@tversu.ru, soldis@yandex.ru

Abstract. We investigate a minimum risk portfolio model under conditions of a hybrid uncertainty of a possibilistic-probabilistic type with weak and strong triangular norms (t-norms) describing the interaction of fuzzy factors of the model. For the case of the weakest t-norm, a formula for variance is derived, which makes it possible to estimate the risk of the portfolio. An equivalent crisp analog of the model is constructed and demonstrated on a numerical example.

Keywords: Minimum risk portfolio · Hybrid uncertainty · Weakest t-norm · Strongest t-norm · Possibility · Necessity · Expected possibility · Crisp second moments

1 Introduction

The work investigates the behavior of a set of investment opportunities under conditions of hybrid uncertainty of a possibilistic-probabilistic type with extreme t-norms describing the interaction of fuzzy factors in the minimum risk portfolio model. To this end, a mathematical model of a minimum risk portfolio, depending on the type of t-norm, has been developed, equivalent crisp analogues have been constructed. Based on the results, we study a set of feasible portfolios and construct sets of quasi-efficient portfolios.

To handle the uncertainty of probabilistic type in decision-making process we use a principle based on an expected possibility, which completely corresponds to the classical approach of Markowitz. As a result, we have possibilistic function that represents portfolio return. Uncertainty of possibilistic type is removed by imposing requirements on the possibility/necessity of fulfilling an investor's constraints on an acceptable level of portfolio return which follows the usual practice (not only in the financial area). We use indirect method in order to build equivalent crisp analogue which in turn is a problem of mathematical programming that can be solved by a set of classical approaches.

The work develops the results from [1,2] in which a similar problem was solved in min-related environment. In the present work we investigate the case

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_51

when all fuzzy parameters of the model are mutually T_W-related, that is we use the weakest (drastic) t-norm in order to aggregate fuzzy information. After that we make comparative analysis of these two results.

2 Necessary Concepts and Notations

We introduce a number of definitions and concepts from the possibility theory following [3–5]. Let $(\Gamma,\ \mathrm{P}(\Gamma),\ \tau)$ and $(\Omega,\ \mathrm{B},\ \mathrm{P})$ be possibility and probability spaces, where Ω is a sample space with $\omega \in \Omega$, Γ – a pattern space with elements $\gamma \in \Gamma$, B – σ-algebra of events, $\mathrm{P}(\Gamma)$ is the discrete topology on Γ, $\tau \in \{\pi,\ \nu\}$, π and ν – measures of possibility and necessity, respectively, and P – probability measure, E^1 – number line.

Definition 1. *Fuzzy random variable $Y(\omega,\gamma)$ is a real function $Y : \Omega\times\Gamma \to E^1$ σ-measurable for each fixed γ, where*

$$\mu_Y\left(\omega, t\right) = \pi\{\gamma \in \Gamma : Y\left(\omega,\ \gamma\right) = t\}$$

is called its distribution function.

It follows from Definition 1 that the distribution function of a fuzzy random variable depends on a random parameter, that is, it is a random function.

Definition 2. *Let $Y(\omega,\gamma)$ be a fuzzy random variable. Its expected value $E[Y]$ is a fuzzy variable with possibility distribution function*

$$\mu_{E[Y]}\left(t\right) = \pi\{\gamma \in \Gamma : E\left[Y\left(\omega,\gamma\right)\right] = t\},$$

where E is the expectation operator

$$E\left[Y\left(\omega,\gamma\right)\right] = \int_{\Omega} Y\left(\omega,\gamma\right) P(d\omega).$$

In this case, the distribution function of the expected value of a fuzzy random variable is no longer dependent on a random parameter and is therefore deterministic. We define second central moment following [6]. Let X and Y be fuzzy random variables.

Definition 3. *A covariance of fuzzy random variables X and Y is defined as:*

$$cov\left(X,Y\right) = \frac{1}{2}\int_0^1 \Big(cov\left(X_\omega^-\left(\alpha\right), Y_\omega^-\left(\alpha\right)\right) + cov\left(X_\omega^+\left(\alpha\right), Y_\omega^+\left(\alpha\right)\right)\Big)d\alpha, \quad (1)$$

where $X_\omega^-\left(\alpha\right),\ Y_\omega^-\left(\alpha\right),\ X_\omega^+\left(\alpha\right),\ Y_\omega^+\left(\alpha\right)$ are left and right boundaries of α-level sets of fuzzy variables X_ω and Y_ω.

Definition 4. *A variance of a fuzzy random variable Y is*

$$D\left[Y\right] = cov(YY).$$

The expected value, variance and covariance of fuzzy random variables determined in such way inherit basic properties of the corresponding characteristics of real-valued random variables.

LR-type distributions are often used to model fuzzy numbers [7].

Definition 5. *$Z(\gamma)$ is called an LR-type fuzzy variable, if its distribution function has the form*

$$\mu_Z(t) = \begin{cases} L\left(\frac{\underline{m}-t}{\underline{d}}\right), & for\, t < \underline{m}, \\ 1, & for\, \underline{m} \le t \le \overline{m}, \\ R\left(\frac{t-\overline{m}}{\overline{d}}\right), & for\, t > \overline{m}. \end{cases}$$

where $L(t)$ and $R(t)$ are shape functions.

In this case $Z(\gamma)$ is written, as a rule, in the form $Z = \left[\underline{m}, \overline{m}, \underline{d}, \overline{d}\right]_{LR}$, where $\underline{m} \le \overline{m}$ are left and right limits of tolerance (modal) interval, $\underline{d} > 0, \overline{d} > 0$ – coefficients of fuzziness.

We use triangular norms and conorms (t-norms and t-conorms) as an instrument for aggregation of fuzzy information that extends min and max operations, laid in actions on fuzzy sets and fuzzy variables [8,9].

In particular, in this work we consider two extremal t-norms: $T_M(x, y) = \min(x, y)$ and $T_W(x, y) = \begin{cases} \min\{x, y\}, & if\, \max\{x, y\} = 1, \\ 0, & otherwise. \end{cases}$

T_M is called the strongest t-norm and T_W – the weakest t-norm, since for any arbitrary t-norm T and $\forall x, y \in [0, 1]$, the inequality holds:

$$T_W(x, y) \le T(x, y) \le T_M(x, y).$$

One of the main properties of t-norms is their ability to control uncertainty ("fuzziness") growth, which is obvious, for example, when performing arithmetic operations on fuzzy numbers: when adding two fuzzy numbers of LR-type using the strongest t-norm T_M corresponding coefficients of fuzziness are summed, therefore uncertainty is growing. With the help of t-norms other than T_M we can have slower growth of fuzziness. The extreme cases of triangular norms which are considered in the work give us boundaries for control of fuzziness in our minimum risk portfolio model. Following [10], we introduce the notion of mutual t-relatedness of fuzzy sets and fuzzy variables. It is used as an instrument for constructing joint possibility distribution functions.

Definition 6. *Fuzzy sets $A_1, \ldots, A_n \in P(\Gamma)$ are called mutually T-related, if for any index set $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$, $k = 1, \ldots, n$, we have*

$$\pi(A_{i_1} \cap \ldots \cap A_{i_k}) = T(\pi(A_{i_1}), \ldots, \pi(A_{i_k})),$$

where

$$T(\pi(A_{i_1}), ..., \pi(A_{i_k})) = T(T(...T(T(\pi(A_{i_1}), \pi(A_{i_2})), \pi(A_{i_3})), ...), \pi(A_{i_k})).$$

We can transfer the notion of mutual T-relatedness of fuzzy sets on fuzzy variables.

Definition 7. *Fuzzy variables* $Z_1(\gamma), \ldots, Z_n(\gamma)$ *are called mutually* T*-related, if for any index set* $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$, $k = 1, \ldots, n$, *we have*

$$\mu_{Z_{i_1},\ldots,Z_{i_k}}\left(t_{i_1}, \ldots, t_{i_k}\right) = \pi\left\{\gamma \in \Gamma\, : Z_{i_1}\left(\gamma\right) = t_{i_1}, \ldots, Z_{i_k}\left(\gamma\right) = t_{i_k}\right\} =$$

$$\pi\left\{Z_{i_1}^{-1}\left\{t_{i_1}\right\} \cap \ldots \cap Z_{i_k}^{-1}\left\{t_{i_k}\right\}\right\} = T\left\{\pi\left(Z_{i_1}^{-1}\left\{t_{i_1}\right\}\right), \ldots, \pi\left(Z_{i_k}^{-1}\left\{t_{i_k}\right\}\right)\right\}, t_{i_j} \in E^1.$$

3 Expected Portfolio Return Under Conditions of Hybrid Uncertainty

In conditions of hybrid uncertainty of possibilistic-probabilistic type, the return on an investment portfolio can be represented as a fuzzy random function

$$R_P\left(w, \omega, \gamma\right) = \sum_{i=1}^{n} R_i\left(\omega, \gamma\right) w_i, \tag{2}$$

that is, it is a linear function of portfolio's capital shares w_i. Here $R_i\left(\omega, \gamma\right)$ are fuzzy random variables with a shift-scale representation [5]:

$$R_i\left(\omega, \gamma\right) = a_i\left(\omega\right) + \sigma_i\left(\omega\right) Z_i(\gamma), \tag{3}$$

which model the profitability of individual financial assets.

Further in this work we assume that in the representation (3) fuzzy variables $Z_i(\gamma)$ are mutually T_W-related, and $a_i\left(\omega\right)$, $\sigma_i(\omega)$ are random offset and scale factors.

For a better intuitive understanding of such a representation of a fuzzy random variable, imagine a situation where some financial expert is asked to estimate return of a certain financial asset. Both the return and its estimation by the expert are uncertain quantities. We will assume that the uncertainty, determined by market conditions, has a probabilistic nature. On the other hand, uncertainty of expert's estimation is described by some possibility distribution. This model seems quite plausible if we assume that the degree of fuzziness of the expert depends mainly not on the true value of the estimated variable, but on the scale of its variability.

Under the assumptions made, the possibility distribution of portfolio return (2) takes the following form [9,11]

$$R_p\left(w, \omega, \gamma\right) = \left[\underline{m}_{R_p}\left(w, \omega\right), \overline{m}_{R_p}\left(w, \omega\right), \underline{d}_{R_p}\left(w, \omega\right), \overline{d}_{R_p}\left(w, \omega\right)\right]_{LR}, \tag{4}$$

where

$$\underline{m}_{R_p}\left(w, \omega\right) = \sum\nolimits_{i=1}^{n}\left(a_i(\omega) + \sigma_i(\omega), \underline{m}_i\right) w_i,$$

$$\overline{m}_{R_p}\left(w, \omega\right) = \sum\nolimits_{i=1}^{n}\left(a_i(\omega) + \sigma_i(\omega)\overline{m}_i\right) w_i,$$

$$\underline{d}_{R_p}(w,\omega) = \max_{i=1,\ldots,n}\{\sigma_i(\omega)\underline{d}_i w_i\},\ \overline{d}_{R_p}(w,\omega) = \max_{i=1,\ldots,n}\{\sigma_i(\omega)\overline{d}_i w_i\}.$$

To remove the uncertainty of probabilistic type, in accordance with our selected approach, it is necessary to identify the possibility distribution of the mathematical expectation of the function $R_p(w,\omega,\gamma)$, that is, to calculate its parameters. Therefore, the expected portfolio return for a fixed w would be a fuzzy variable. This is shown by the following theorem.

Theorem 1. *Let the conditions (2)–(4) are fulfilled, then the expected portfolio return is characterized by the following possibility distribution:*

$$\hat{R}_p(w,\gamma) = E[R_p(w,\omega,\gamma)] = \left[E\left[\underline{m}_{R_p}(w,\omega)\right], E\left[\overline{m}_{R_p}(w,\omega)\right], E\left[\underline{d}_{R_p}(w,\omega)\right], E\left[\overline{d}_{R_p}(w,\omega)\right]\right]_{LR} = [\underline{m}_{\hat{R}_p}(w), \overline{m}_{\hat{R}_p}(w), \underline{d}_{\hat{R}_p}(w), \overline{d}_{\hat{R}_p}(w)],$$

where $\underline{m}_{\hat{R}_p}(w) = \sum_{i=1}^n(\hat{a}_i + \hat{\sigma}_i\underline{m}_i)w_i,\ \overline{m}_{\hat{R}_p}(w) = \sum_{i=1}^n(\hat{a}_i + \hat{\sigma}_i\overline{m}_i)w_i,$

$$\underline{d}_{\hat{R}_p}(w) = E\left[\max_{i=1,\ldots,n}\{\sigma_i(\omega)\underline{d}_i w_i\}\right],\ \overline{d}_{\hat{R}_p}(w) = E\left[\max_{i=1,\ldots,n}\{\sigma_i(\omega)\overline{d}_i w_i\}\right],$$

$$\hat{a}_i = E[a_i(\omega)],\ \hat{\sigma}_i = E[\sigma_i(\omega)].$$

Proof of Theorem 1 follows from the results obtained in [5,9,11]. Note that functions $\underline{d}_{\hat{R}_p}(w)$ and $\overline{d}_{\hat{R}_p}(w)$ can be calculated in explicit form only for simple probability distributions of random components $a_i(\omega)$ and $\sigma_i(\omega)$ [5] and this is due to the large amount of computation. To reduce the amount of computation and solve problems of this type, we can use stochastic programming methods, in particular, the stochastic quasigradient method.

In accordance with the classical Markowitz approach [12] and with the help of [6], we can construct a risk function for the portfolio and include its expected return in the system of constraints. Since the expected portfolio return in the case of fuzzy random data is fuzzy, the uncertainty of possibilistic type can be removed by imposing requirements on the possibility/necessity of fulfilling an investor's constraints on an acceptable level of expected portfolio return. In this case, the model of feasible portfolios by Markowitz in case of hybrid uncertainty can be represented in the following form

$$F_p = \begin{cases} \tau\left\{\hat{R}_p(w,\gamma)\ \mathcal{R}\ m_d\right\} \geq \alpha, \\ \sum w_i = 1, \\ w \in E_+^n, \end{cases}$$

where $E_+^n = \{w \in E^n : w_i \geq 0\}$, $\hat{R}_p(w,\gamma)$ – expected portfolio return, $\mathcal{R}$ – crisp relation $\{\geq, =\}$; $\alpha \in (0,1], m_d$ – acceptable level of expected portfolio return allowable by an investor. In general, relation $\mathcal{R}$ can be fuzzy, too [13].

The following theorem makes it possible to construct an equivalent crisp analogue of the feasible portfolios model F_P.

Theorem 2. *Let in the constraint model F_p $\tau = \pi$, $\mathcal{R}$ is '$=$'. Then, with possibility not less than α, the level of the expected portfolio return m_d admits the following estimate:*

$$\sum\nolimits_{i=1}^{n} \left(\hat{a}_i + \hat{\sigma}_i \underline{m}_i\right) w_i - E\left[\max_{i=1,\ldots,n} \left\{\sigma_i\left(\omega\right) \underline{d}_i w_i\right\}\right] * L^{-1}\left(\alpha\right) \le m_d$$
$$\le \sum\nolimits_{i=1}^{n} \left(\hat{a}_i + \hat{\sigma}_i \overline{m}_i\right) w_i + E\left[\max_{i=1,\ldots,n} \left\{\sigma_i\left(\omega\right) \overline{d}_i w_i\right\}\right] * R^{-1}\left(\alpha\right).$$

Proof. In accordance with the definition of LR-type fuzzy variable distribution of portfolio's expected return has the following form:

$$\mu_{\hat{R}_p}\left(t\right) = \begin{cases} L\left(\frac{\underline{m}_{\hat{R}_p}(w) - t}{\underline{d}_{\hat{R}_p}(w)}\right), \; t < \underline{m}_{\hat{R}_p}\left(w\right), \\ 1, \qquad \underline{m}_{\hat{R}_p}\left(w\right) \le t \le \overline{m}_{\hat{R}_p}\left(w\right), \\ R\left(\frac{t - \overline{m}_{\hat{R}_p}(w)}{\overline{d}_{\hat{R}_p}(w)}\right), \; t > \overline{m}_{\hat{R}_p}\left(w\right). \end{cases}$$

The left boundary of its α-level set can be found from the relation:

$$L\left(\frac{\underline{m}_{\hat{R}_p}\left(w\right) - t}{\underline{d}_{\hat{R}_p}\left(w\right)}\right) = \alpha,$$

from which, after simple transformations, we have

$$\hat{R}_p^-\left(w, \alpha\right) = \underline{m}_{\hat{R}_p}\left(w\right) - \underline{d}_{\hat{R}_p}\left(w\right) * L^{-1}\left(\alpha\right).$$

Similarly, we find the right boundary $\hat{R}_p^+\left(w, \alpha\right)$. Suppose that the joint distribution $\hat{R}_p\left(w, \gamma\right)$ and m_d is defined by the strongest t-norm T_M. Then the possibilistic inequality $\tau\left\{\hat{R}_p\left(w, \gamma\right) = m_d\right\} \ge \alpha$ is equivalent to the following system of deterministic inequalities [14–16]:

$$\hat{R}_p^-\left(w, \alpha\right) \le m_d \le \hat{R}_p^+\left(w, \alpha\right).$$

After expanding $\hat{R}_p^-\left(w, \alpha\right)$ and $\hat{R}_p^+\left(w, \alpha\right)$ we obtain condition of the Theorem.

Note 1. It is easy to build an equivalent crisp analogue under the assumption that the joint distribution $\hat{R}_p\left(w, \gamma\right)$ and m_d is defined by the weakest t-norm T_W. Although in this case the form of the obtained result will be more cumbersome and thus we will omit it in the present work.

Corollary 1. *If in the condition of Theorem 1 the relation $\mathcal{R}$ is '$\ge$', then we get the following estimate of the expected portfolio return level:*

$$\sum\nolimits_{i=1}^{n} \left(\hat{a}_i + \hat{\sigma}_i \overline{m}_i\right) w_i + E\left[\max_{i=1,\ldots,n} \left\{\sigma_i\left(\omega\right) \overline{d}_i w_i\right\}\right] * R^{-1}\left(\alpha\right) \ge m_d.$$

In the classical portfolio theory by Markowitz, the expected return on the portfolio is limited by the minimum and maximum of the expected returns of its individual assets. It is possible to obtain a relative analogue of a similar estimate for our case (with the possibility of at least α):

Corollary 2. *Under the condition of Theorem 1 the following inequalities for the expected portfolio return level hold:*

$$\min_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\underline{m}_i\right)-E\left[\max_{i=1,\ldots,n}\left\{\sigma_i\left(\omega\right)\underline{d}_i w_i\right\}\right]*L^{-1}\left(\alpha\right)\leq m_d$$
$$\leq\max_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\overline{m}_i\right)+E\left[\max_{i=1,\ldots,n}\left\{\sigma_i\left(\omega\right)\overline{d}_i w_i\right\}\right]*R^{-1}\left(\alpha\right).$$

Proof. This automatically follows from the fact that

$$\sum\nolimits_{i=1}^{n}\left(\hat{a}_i+\hat{\sigma}_i\underline{m}_i\right)w_i\geq\min_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\underline{m}_i\right)\sum\nolimits_{i=1}^{n}w_i=\min_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\underline{m}_i\right),$$

$$\sum\nolimits_{i=1}^{n}\left(\hat{a}_i+\hat{\sigma}_i\overline{m}_i\right)w_i\leq\max_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\overline{m}_i\right)\sum\nolimits_{i=1}^{n}w_i=\max_{i=1,\ldots,n}\left(\hat{a}_i+\hat{\sigma}_i\overline{m}_i\right).$$

4 Portfolio Risk Assessment Under Conditions of Hybrid Uncertainty

In this work we define the variance of a fuzzy random variable in the context of t-norm T_W that describes the interaction of fuzzy factors. For this, we use formula (1) from [6] which takes the following form

$$D_p\left(w\right)=\frac{1}{2}\int_0^1\left(D\left[R_p^-\left(w,\omega,\alpha\right)\right]+D\left[R_p^+\left(w,\omega,\alpha\right)\right]\right)d\alpha,\tag{5}$$

where $R_p^-\left(w,\omega,\alpha\right)$ and $R_p^+\left(w,\omega,\alpha\right)$, are the left and right boundaries, respectively, of the α-level set of portfolio return $R_p(w,\omega,\gamma)$:

$$R_p^-\left(w,\omega,\alpha\right)=\sum\nolimits_{i=1}^{n}\left(a_i(\omega)+\sigma_i(\omega)\underline{m}_i\right)w_i-\max_{i=1,\ldots,n}\left\{\sigma_i\left(\omega\right)\underline{d}_i w_i\right\}*L^{-1}\left(\alpha\right),$$

$$R_p^+\left(w,\omega,\alpha\right)=\sum\nolimits_{i=1}^{n}\left(a_i(\omega)+\sigma_i(\omega)\overline{m}_i\right)w_i+\max_{i=1,\ldots,n}\left\{\sigma_i\left(\omega\right)\overline{d}_i w_i\right\}*R^{-1}\left(\alpha\right).$$

We have:

$$D\left[R_p^-\left(w,\omega,\alpha\right)\right]=\sum_{i=1}^{n}w_i^2D\left[a_i(\omega)+\sigma_i(\omega)\underline{m}_i\right]+D\left[\max_{j=1,\ldots,n}\left\{\sigma_j\left(\omega\right)\underline{d}_j w_j\right\}\right]$$
$$\times\left(L^{-1}\left(\alpha\right)\right)^2+2\sum_{1\leq i<j\leq n}w_iw_j cov\left(\left(a_i\left(\omega\right)+\sigma_i\left(\omega\right)\underline{m}_i\right),\left(a_j\left(\omega\right)+\sigma_j\left(\omega\right)\underline{m}_j\right)\right)$$
$$-2L^{-1}\left(\alpha\right)\sum_{i=1}^{n}w_i cov\left(\left(a_i\left(\omega\right)+\sigma_i\left(\omega\right)\underline{m}_i\right),\max_{j=1,\ldots,n}\left\{\sigma_j\left(\omega\right)\underline{d}_j w_j\right\}\right),$$

$$D\left[R_p^+(w,\omega,\alpha)\right] = \sum_{i=1}^{n} w_i^2 D\left[a_i(\omega)+\sigma_i(\omega)\overline{m}_i\right] + D\left[\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\overline{d}_j w_j\right\}\right]$$
$$\times\left(R^{-1}(\alpha)\right)^2 + 2\sum_{1\le i<j\le n} w_i w_j cov\left((a_i(\omega)+\sigma_i(\omega)\overline{m}_i),(a_j(\omega)+\sigma_j(\omega)\overline{m}_j)\right)$$
$$+2R^{-1}(\alpha)\sum_{i=1}^{n} w_i cov\left((a_i(\omega)+\sigma_i(\omega)\overline{m}_i),\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\overline{d}_j w_j\right\}\right).$$

After substituting the found variances of the left and right boundaries into (5), we obtain:

$$D_p(w) = \frac{1}{2}\sum_{i=1}^{n} w_i^2\left(D\left[a_i(\omega)+\sigma_i(\omega)\underline{m}_i\right] + D\left[a_i(\omega)+\sigma_i(\omega)\overline{m}_i\right]\right)$$
$$+\frac{1}{2}D\left[\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\overline{d}_j w_j\right\}\right]\int_0^1\left(R^{-1}(\alpha)\right)^2 d\alpha$$
$$+\frac{1}{2}D\left[\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\underline{d}_j w_j\right\}\right]\int_0^1\left(L^{-1}(\alpha)\right)^2 d\alpha$$
$$+\sum_{1\le i<j\le n} w_i w_j\left(cov\left((a_i(\omega)+\sigma_i(\omega)\underline{m}_i),(a_j(\omega)+\sigma_j(\omega)\underline{m}_j)\right)\right.$$
$$\left.+cov\left((a_i(\omega)+\sigma_i(\omega)\overline{m}_i),(a_j(\omega)+\sigma_j(\omega)\overline{m}_j)\right)\right)$$
$$+\sum_{i=1}^{n} w_i\left(\int_0^1 R^{-1}(\alpha)\,d\alpha\,cov\left((a_i(\omega)+\sigma_i(\omega)\overline{m}_i),\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\overline{d}_j w_j\right\}\right)\right.$$
$$\left.-\int_0^1 L^{-1}(\alpha)\,d\alpha\,cov\left((a_i(\omega)+\sigma_i(\omega)\underline{m}_i),\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,\underline{d}_j w_j\right\}\right)\right).$$

If in fuzzy random variables all random parameters of distributions are independent and fuzzy components are represented by symmetrical fuzzy numbers of LR-type, i.e. $S(t) = L(t) = R(t), \forall t \ge 0$ and $m_i = \overline{m}_i = \underline{m}_i$, $d_i = \overline{d}_i = \underline{d}_i$, $i = 1,\dots,n$, then the variance formula has the form:

$$\sum_{i=1}^{n} w_i^2\left(D\left[a_i(\omega)\right] + D\left[\sigma_i(\omega)\right]m_i^2\right) + D\left[\max_{j=1,\dots,n}\left\{\sigma_j(\omega)\,d_j w_j\right\}\right]\int_0^1\left(S^{-1}(\alpha)\right)^2 d\alpha. \tag{6}$$

Let us obtain as an example a formula for variance when the shift and scale coefficients $a_i(\omega)$ and $\sigma_j(\omega)$ have specific distribution – uniform on [0,1]. We use the variance property, according to which it is equal to the difference between the second moment and the square of the first moment:

$$D[X] = E\left[X^2\right] - (E[X])^2. \tag{7}$$

In accordance with [5], we have for $E\left[\max_i\left\{\sigma_j(\omega)\,\overline{d}_j w_j\right\}\right]$ (we denote it $EMax$):

$$EMax\left(\overline{d}w\right) = \sum_{i=1}^{n}\frac{\left(\overline{d}w\right)_{(i)}^{n-i+1}}{(n-i+1)(n-i+2)\left(\overline{d}w\right)_{(i+1)}\dots\left(\overline{d}w\right)_{(n)}},$$

where $\left(\overline{d}w\right)_{(1)}\left(\overline{d}w\right)_{(2)},\ldots,\left(\overline{d}w\right)_{(n)}$ is an ordered ascending permutation of elements $\left\{\overline{d}_1 w_1, \overline{d}_2 w_2 \ldots, \overline{d}_n w_n\right\}$. It is easy to show in a similar way that the second moment $E\left[\left(\max_i\left\{\sigma_j\left(\omega\right)\overline{d}_j w_j\right\}\right)^2\right]$ is given by the following formula:

$$EMax2\left(\overline{d}w\right)=\sum\nolimits_{i=1}^{n}\frac{2\left(\overline{d}w\right)_{(i)}^{n-i+2}}{\left(n-i+2\right)\left(n-i+3\right)\left(\overline{d}w\right)_{(i+1)}\cdots\left(\overline{d}w\right)_{(n)}}.$$

Substituting everything in (7), we obtain the variance. Similarly, we can construct $D\left[\max_i\left\{\sigma_j\left(\omega\right)\underline{d}_j w_j\right\}\right]$.

If we denote

$$ESMax(dw)=E\left[\sigma_i\left(\omega\right)\max_{j=1,\ldots,n}\left\{\sigma_j\left(\omega\right)\underline{d}_j w_j\right\}\right],$$

variance formula finally becomes:

$$\begin{aligned}D_p\left(w\right)&=\frac{1}{2}\sum_{i=1}^{n}w_i^2\left(\frac{1}{12}\left(\underline{m}_i^2+\overline{m}_i^2\right)+\frac{1}{6}\right)\\&+\frac{1}{2}\left(EMax2\left(\overline{d}w\right)-\left(EMax\left(\overline{d}w\right)\right)^2\right)\int_0^1\left(R^{-1}\left(\alpha\right)\right)^2d\alpha\\&+\frac{1}{2}\left(EMax2\left(\underline{d}w\right)-\left(EMax\left(\underline{d}w\right)\right)^2\right)\int_0^1\left(L^{-1}\left(\alpha\right)\right)^2d\alpha\\&+\sum\nolimits_{i=1}^{n}w_i\left(\int_0^1R^{-1}\left(\alpha\right)d\alpha\,\overline{m}_i\left(ESMax(\overline{d}w)-\frac{1}{2}EMax\left(\overline{d}w\right)\right)\right.\\&\left.-\int_0^1L^{-1}\left(\alpha\right)d\alpha\,\underline{m}_i\left(ESMax(\underline{d}w)-\frac{1}{2}EMax\left(\underline{d}w\right)\right)\right)\end{aligned}$$

and formula (6) with its corresponding assumptions:

$$\frac{1}{12}\sum_{i=1}^{n}w_i^2\left(m_i^2+1\right)+\left(EMax2\left(dw\right)-\left(EMax\left(dw\right)\right)^2\right)\int_0^1\left(S^{-1}\left(\alpha\right)\right)^2d\alpha.$$

5 An Example of Minimum Risk Portfolio and Model Calculations

The model of minimum risk portfolio under conditions of a hybrid uncertainty of a possibilistic-probabilistic type has the form:

$$D_p(w)\rightarrow min, \tag{8}$$

$$w\in F_p. \tag{9}$$

We will carry out numerical calculations and compare the results in case of the strongest and weakest t-norms in a possibility-necessity context. With the assumptions made earlier on the elements of the distributions and with all fuzzy variables having symmetrical triangular forms with $S(t) = \max\{0, 1 - t\}$, the equivalent crisp analogues of the minimum risk portfolio (8)–(9) take the following forms (depending on the context) when all fuzzy parameters of distributions are mutually T_W-related:

$$\frac{1}{12}\sum_{i=1}^{n} w_i^2 \left(m_i^2 + 1\right) + \frac{1}{3}\left(EMax2\left(dw\right) - \left(EMax\left(dw\right)\right)^2\right) \rightarrow min, \tag{10}$$

$$F_p = \begin{cases} \frac{1}{2}\sum_{i=1}^{n}\left(m_i+1\right)w_i + EMax\left(\mathrm{d}w\right) * \left(1-\alpha\right) \geq m_d, & \text{(in possibility context)} \\ \frac{1}{2}\sum_{i=1}^{n}\left(\mathrm{m}_i+1\right)w_i - EMax\left(\mathrm{d}w\right) * \alpha \geq m_d, & \text{(in necessity context)} \\ \sum w_i = 1, & \\ w \in E_+^n. & \end{cases} \tag{11}$$

In the case when all fuzzy parameters are mutually min-related, we have the following equivalent crisp analogues of the model (8)–(9):

$$\frac{1}{12}\sum_{i=1}^{n} w_i^2 \left(m_i^2 + 1 + \frac{1}{3}d_i^2\right) \rightarrow min, \tag{12}$$

$$F_p = \begin{cases} \frac{1}{2}\sum_{i=1}^{n}\left(m_i + 1 + d_i(1-\alpha)\right)w_i \geq m_d, & \text{(in the possibility context)} \\ \frac{1}{2}\sum_{i=1}^{n}\left(m_i + 1 - d_i\alpha\right)w_i \geq m_d, & \text{(in the necessity context)} \\ \sum w_i = 1, & \\ w \in E_+^n. & \end{cases} \tag{13}$$

Consider a model example for n = 2. Let $Z_1 = [0.3; 0.3; 3.5; 3.5]_{SS}$, $Z_2 = [2.8; 2.8; 1.5; 1.5]_{SS}$, $\alpha = 0.65$. Figure 1 illustrates the sets of investment opportunities calculated from (10)–(13).

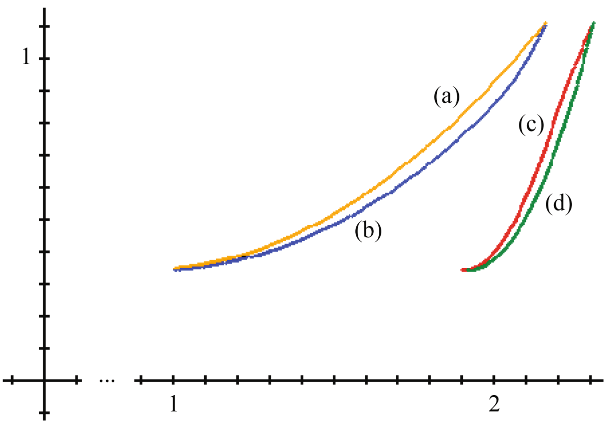


Fig. 1. Dependence of the minimum risk of portfolio on its expected return for min-related parameters in necessity context (a) and possibility context (d), for T_W-related parameters in necessity context (b) and possibility context (c).

Curves (b) and (c) on Fig. 1 show solutions of the problem (10)–(11) in necessity and possibility contexts, respectively, and curves (a) and (d) – solutions of the problem (12)–(13) in necessity and possibility contexts, respectively.

On the figure, the x-axis is the expected portfolio return, and the y-axis is the risk of the portfolio. As one can see from Fig. 1 T_W-related regions are narrower thus less "fuzzy" which is consistent with properties of T_W-related calculus of fuzzy variables.

Figures 2 and 3 show behavior of investment opportunities curves from Fig. 1 depending on α-level. Two extreme cases are shown: when α is almost zero (Fig. 2) we see that necessity curves for T_M and T_W start to coincide and when $\alpha = 1$ (Fig. 3) than the same behavior transitions to possibilities curves.

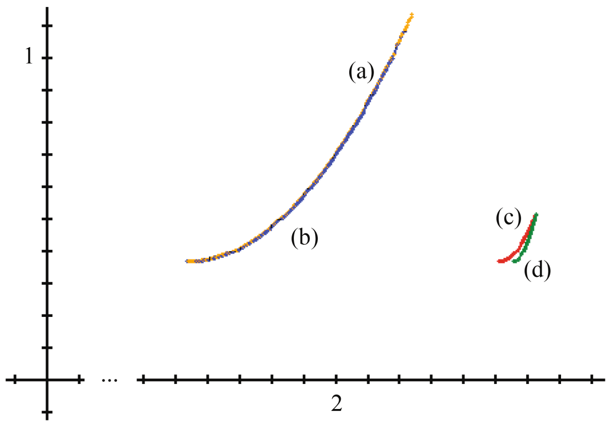


Fig. 2. Behavior of investment opportunities curves when $\alpha = 0.01$.

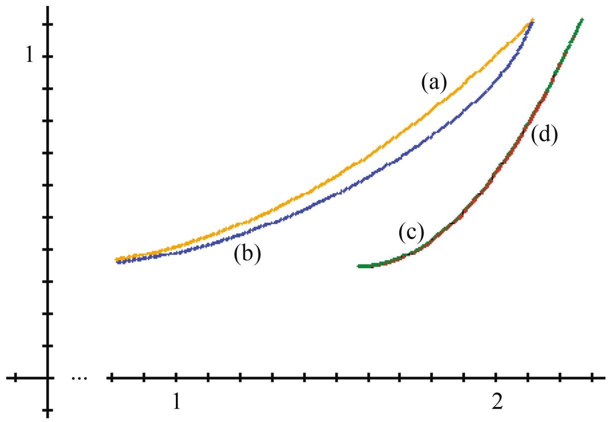


Fig. 3. Behavior of investment opportunities curves when $\alpha = 1$.

6 Conclusion

The following generalizations and amplifications obtained results [1,2] are described: a mathematical model of the minimum risk portfolio was constructed under the conditions of hybrid uncertainty of possibilistic-probabilistic type in the possibility-necessity context; a formula was obtained for the portfolio risk estimation. For this purpose, we derived a formula for the variance of the portfolio in the context of t-norm T_W describing the interaction of fuzzy factors. An equivalent crisp analog of the minimum risk portfolio model is constructed. The approach is demonstrated on a model example for the case when its probabilistic parameters are independent and distributed according to a uniform law on [0,1] and fuzzy factors are symmetrical fuzzy numbers.

Results of methods for finding quasi-efficient portfolios for different t-norms in the possibility-necessity context were compared on a numerical example.

References

1. Yazenin, A.V.: Possibilistic-probabilistic models and methods of portfolio optimization. In: Batyrshin, I., Kacprzyk, J. et al. (eds.) Studies in Computational Intelligence, vol. 36, pp. 241–259. Springer, Heidelberg (2007). doi:10.1007/978-3-540-36247-0_9
2. Xu, J., Zhou, X.: Fuzzy-Like Multiple Objective Decision Making. Studies in Fuzziness and Soft Computing, vol. 263. Springer, Berlin (2011). doi:10.1007/978-3-642-16895-6
3. Nahmias, S.: Fuzzy variables in a random environment. In: Gupta, M.M., Ragade, R.K., Yager, R.R. (eds.) Advances in Fuzzy Sets Theory and Applications, pp. 165–180. NHCP, Amsterdam (1979)
4. Puri, M.L., Ralescu, D.A.: Fuzzy Random Variables. J. Math. Anal. Appl. **114**, 409–422 (1986). doi:10.1016/0022-247X(86)90093-4
5. Yazenin, A.V.: Basic Concepts of Possibility Theory: A Mathematical Apparatus for Decision-making Under Hybrid Uncertainty Conditions. Fizmatlit Publ. Moscow (2016). (in Russian)
6. Feng, Y., Hu, L., Shu, H.: The variance and covariance of fuzzy random variables and their applications. Fuzzy Sets Syst. **120**, 487–497 (2001). doi:10.1016/S0165-0114(99)00060-3
7. Dubois, D., Prade, H.: Théorie des Possibilités. Application à la Représentation des Connaissances en Informatique. Masson, Paris (1988)
8. Nguyen, H.T., Walker, E.A.: A First Course in Fuzzy Logic. CRC Press, Boca Raton (1997)
9. Mesiar, R.: Triangular-norm-based addition of fuzzy intervals. Fuzzy Sets Syst. **91**, 231–237 (1997). doi:10.1016/S0165-0114(97)00143-7
10. Hong, D.H.: Parameter estimations of mutually T-related fuzzy variables. Fuzzy Sets Syst. **123**, 63–71 (2001). doi:10.1016/S0165-0114(00)00113-5
11. Yazenin, A.V., Soldatenko, I.S.: Possibilistic optimization tasks with mutually t-related parameters: solution methods and comparative analysis. In: Lodwick, W.A., Kacprzyk, J. (eds.) Studies in Fuzziness and Soft Computing, vol. 254, pp. 163-192. Springer (2010). doi:10.1007/978-3-642-13935-2_8

12. Markowitz, H.M.: Portfolio selection. J. Finance **7**(1), 77–91 (1952). doi:10.2307/2975974
13. Gordeev, R.N., Yazenin, A.V.: A method for solving a problem of possibilistic programming. J. Comput. Syst. Sci. Int. **45**(3), 442–449 (2006). doi:10.1134/S1064230706030105
14. Yazenin, A.V.: Fuzzy and stochastic programming. Fuzzy Sets Syst. **22**(1–2), 171–180 (1987). doi:10.1016/0165-0114(87)90014-5
15. Yazenin, A., Wagenknecht, M.: Possibilistic Optimization. Brandenburgische Technische Universitat, Cottbus (1996)
16. Yazenin, A.V.: On the problem of possibilistic optimization. Fuzzy Sets Syst. **81**(1), 133–140 (1996). doi:10.1016/0165-0114(95)00245-6

Discrete Wavelet Transform and Fuzzy Logic Algorithm for Classification of Fault Type in Underground Cable

Suntiti Yoomak, Chaichan Pothisarn, Chaiyan Jettanasen, and Atthapol Ngaopitakkul(✉)

Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Chalongkrung Road, Ladkrabang, Bangkok, Thailand
atthapol.ng@kmitl.ac.th

Abstract. This paper proposes the combination of discrete wavelet transform (DWT) and fuzzy logic to classify the fault type in underground distribution cable. The DWT is employed to decompose high frequency component from fault signal with the mother wavelet daubechies4 (db4). The maximum coefficients detail of DWT from phase A, B, C and zero sequence for post-fault current waveforms are considered as an input pattern of decision algorithm. Triangle-shaped S-shaped and Z-shaped membership function with maximum, medium, minimum, and zero are used to create a function for the input variable. Output variable of fuzzy are designated as values range 1 to 10 which corresponding with type of fault. The obtained average accuracy results shown that the proposed decision algorithm is able to classify the fault type with satisfactory accuracy.

Keywords: Fuzzy logic · Fault type · Underground distribution system

1 Introduction

In the literature for transmission line and underground cable protection [1–28], based on the transient-based techniques, the application of wavelet transform is used [2, 3, 8, 9]. In previous research works [8], the division algorithm between the maximum coefficients of DWT at ¼ cycle of phase A, B, C is performed. The obtained results were compared to identifying the phase with fault appearance so that the types of fault can be analysed.

In recent years, the artificial intelligence has been often employed for fault diagnosis for power system [10–16] due that the algorithm can give precise results. Based on studied research papers related to fault classification in transmission and distribution system in the literature, the fuzzy logic is one of algorithm that is interested in previous decade [18–27]. In [28], the decision algorithm using the combination of DWT and fuzzy logic to identify of fault type on single circuit transmission line was presented. The combination of DWT and back-propagation neural network have been compared with the proposed decision algorithm in [28]. The obtained results seen that the proposed decision algorithm can be identified the fault type with highly satisfactory more than the other. Paper [29] presents the proper input pattern of fuzzy logic algorithm for fault type classification in underground cable.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_52

Paper [30] is checked for all types of fault taking into account the fundamental components of voltage and current collected from the sending end. In [31], the results of investigation into a new fault classification and location technique are presented by using EMTP software. Paper [32] presents a new algorithm for fault detection and classification using discrete wavelet transform (DWT) and back-propagation neural network (BPNN) based on Clarke's transformation on parallel transmission. In [33], a novel scheme using wavelet technique for classification of faults in TS is proposed. However, all of the above paper, researchers mostly have used several algorithms to classifying the fault type in overhead transmission line but not for classifying in underground distribution system.

The goal of this paper is to classify the fault type in underground distribution cable using the combination of the DWT and fuzzy logic. The system under this study is from Metropolitan Electricity Authority (MEA) that is a part of Thailand's 115 kV underground distribution systems. The locations of fault including the fault type and fault inception angles are varied to study the behavior of the fault. The ATP/EMTP is employed to simulate the fault signals. In addition, the results from the proposed algorithm are compared with the comparison coefficient technique and probabilistic neural network.

2 Simulations

The ATP/EMTP [8] is employed to simulate fault signals, at a sampling rate of 200 kHz. The system employed in case studies are chosen based on the underground distribution system as illustrated in Fig. 1. In addition, a cross-sectional view of a cable is shown in Fig. 2. To avoid complexity, the fault resistance is assumed to be 10 Ω. Fault patterns in the simulations are performed with various changes of system parameters as follows:

- Fault types are single line to ground, double lines to ground, line to line, and three-phase fault.
- Fault locations are from 1 to 5 km (each step = 1 km) of the underground cable length measured from the sending end
- Fault inception angles on the phase A voltage waveform are varied from 0° to 150° with a step of 30°

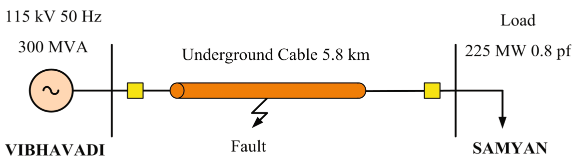


Fig. 1. The system used in simulation studies [8].

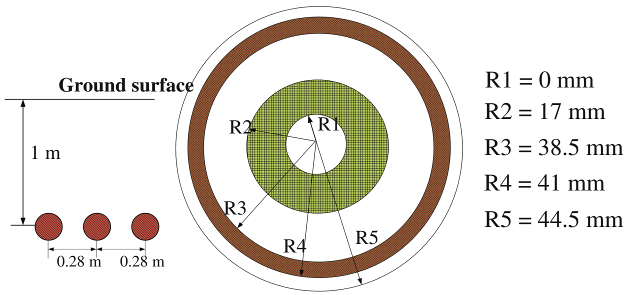


Fig. 2. The configuration of cable in simulation studies [8]

3 Decision Algorithm

The proposed decision algorithm in this paper can be divided into 2 processes. First process, the fault detection decision algorithm must be detected using the discrete wavelet transform to classifying between fault condition and normal condition. For next process, after fault condition can be detected, the fault type decision algorithm will be classifying the type of fault using the fuzzy logic.

With several trial and error processes, the fault detection decision algorithm [8] on the basis of computer programming technique is constructed as shown in Fig. 3. Fault detection decision algorithm is processed using positive sequence current signal. By considering the Fig. 3, fault signals are imported to analyse the high frequency transient components using mother wavelet daubechies4 (db4) in the wavelet toolbox. The Clark's transformation matrix is employed for calculating the positive sequence and zero sequence of currents. After applying the DWT to the positive sequence currents, the comparison of the coefficients from each scale is under investigation. Coefficients obtained using DWT of signals are squared so that the abrupt change in the spectra can be clearly found. This sudden change is used as an index for the occurrence of faults. The fault detection decision algorithm has been proposed that if coefficients of any scales are change around five times before an occurrence of the faults, there are faults occurring in underground cable.

After the fault detection process, the comparison of the coefficients from first scale that can detect fault is considered as input variables for the next process in order to classify the types of fault. Before the fault type decision algorithm process, a structure of the fuzzy logic consists of 4 inputs and 1 output as illustrated in Fig. 4. By observing Fig. 4, it can be observed that, in first stage, the maximum coefficients detail (phase A, B, C and zero sequence of post-fault current signals) of DWT at the first peak time that can detect fault, after that define value range 0 to 1 by normalized, is performed as input variables.

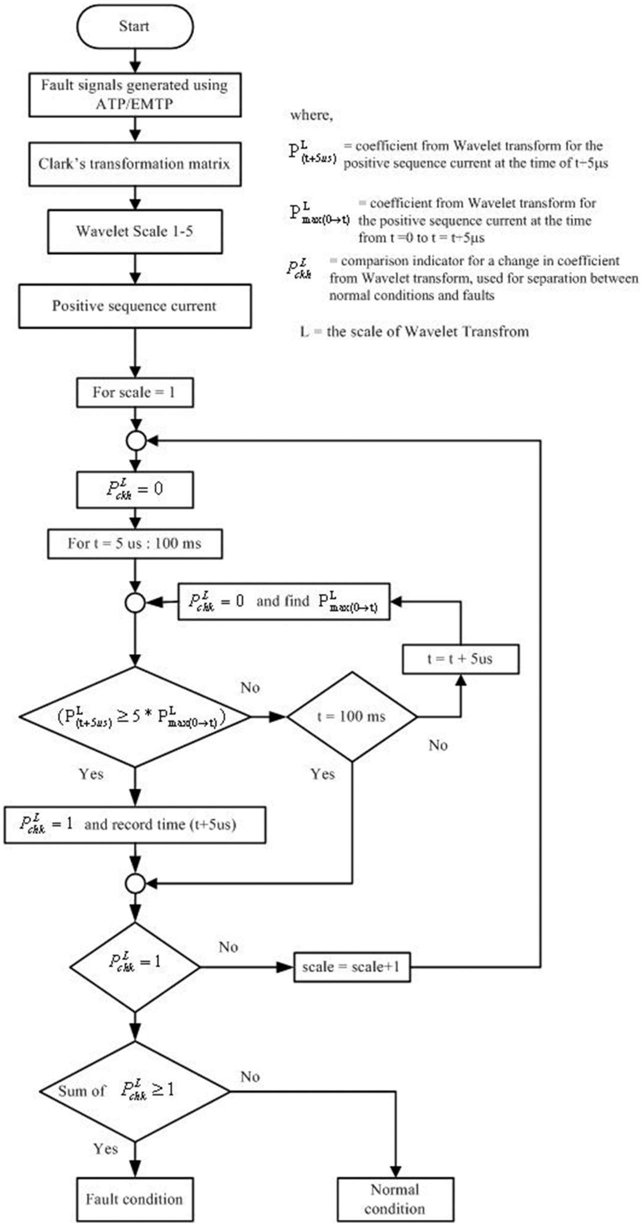


Fig. 3. Flowchart for fault detection [8].

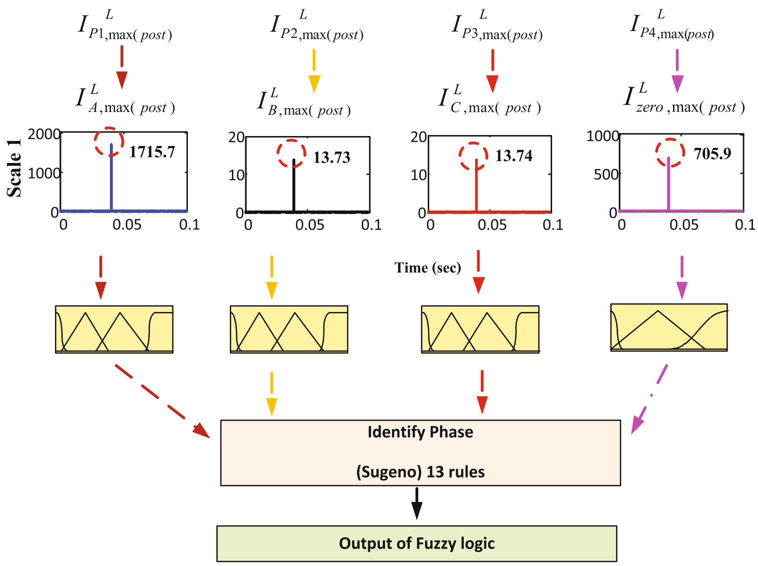


Fig. 4. Structure of fuzzy logic for classifying the fault type in case of maximum coefficients detail of DWT (Case 1).

For the next stage, triangle-shaped S-shaped and Z-shaped membership functions are used to create a function for the input variable. In addition, a term of the number of fuzzy sets has 4 terms and linguistic variable has 4 levels such as maximum, medium, minimum, and zero are designed by Z-shaped, triangle-shaped, triangle-shaped, and S-shaped respectively as shown in Fig. 5. Output variable of fuzzy are designated as values range 1 to 10 which corresponding with type of fault as shown in Table 1.

Finally, for this paper, the fuzzy inference rules are based on the principles of fuzzy logic in the form of IF-THEN. The IF statement is called the condition (antecedent), which is the input of fuzzy while the THEN is called the consequent, which is the output of fuzzy.

According to the fault type can be calculated using proposed decision algorithm, the various case studies were performed with various types of faults at each location on the distribution underground cable including the variation of fault inception angles and locations on each distribution cable in order to verify the decision algorithm capability. The total number of the case studies is 200 sets. The average accuracy of proposed decision algorithm is shown in Table 2. By considering the data in Table 2, it can be seen that the proposed algorithm gives results with satisfactory accuracy.

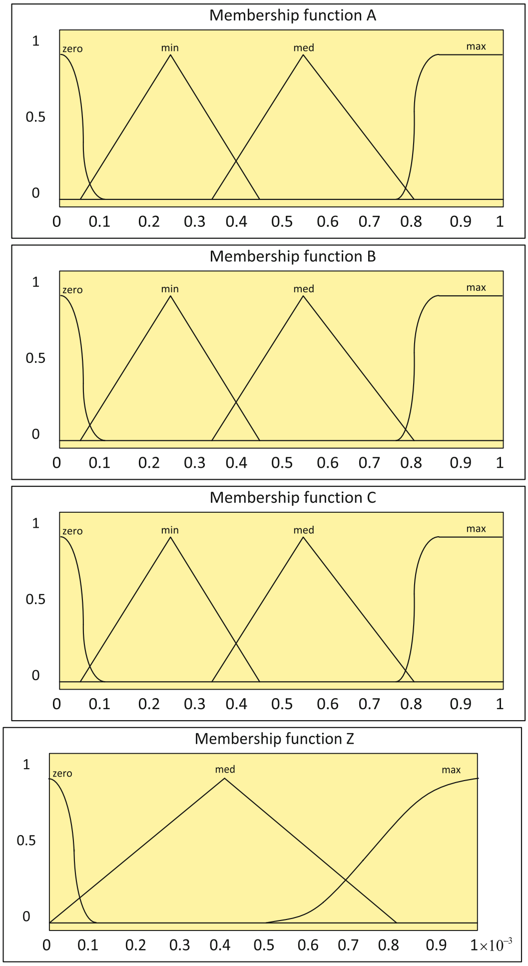


Fig. 5. Membership functions of input variable for classifying the fault types in case of maximum coefficients detail of DWT (Case 1).

Table 1. Rules of fuzzy logic

Output of fuzzy logic	Rules of the module				Classification of fault type	Types of fault
	$I^{L}_{P1,\max(\text{post})}$	$I^{L}_{P2,\max(\text{post})}$	$I^{L}_{P3,\max(\text{post})}$	$I^{L}_{P4,\max(\text{post})}$		
1	MAX	ZERO	ZERO	MAX	Phase A to ground fault	AG
2	ZERO	MAX	ZERO	MAX	Phase B to ground fault	BG
3	ZERO	ZERO	MAX	MAX	Phase C to ground fault	CG
4	NOT ZERO	NOT ZERO	ZERO	MAX	Phase A and phase B to ground fault	ABG
5	MED	MED	ZERO	MED	Phase A to phase B fault	AB
5	MED	MED	MIN	MED	Phase A to phase B fault	AB
6	NOT ZERO	ZERO	NOT ZERO	MAX	Phase C and phase A to ground fault	CAG
7	MED	ZERO	MED	MED	Phase C to phase A fault	CA
7	MED	MIN	MED	MED	Phase C to phase A fault	CA
8	ZERO	NOT ZERO	NOT ZERO	MAX	Phase B and phase C to ground fault	BCG
9	ZERO	MED	MED	MED	Phase B to phase C fault	BC
9	MIN	MED	MED	MED	Phase B to phase C fault	BC
10	NOT ZERO	NOT ZERO	NOT ZERO	ZERO	Three phase fault	ABC

Table 2. Percentage of average accuracy for fault types

Classification of the fault types	Number of case studies	Proposal algorithm
AG	20	85%
BG	20	80%
CG	20	85%
ABG	20	80%
CAG	20	80%
BCG	20	80%
AB	20	100%
CA	20	80%
BC	20	100%
ABC	20	100%
Average		88%

4 Conclusion

This paper proposed the combination of the discrete wavelet transform and fuzzy logic algorithm to classify the fault type in underground distribution cable. The proposed decision algorithm in this paper can be divided into 2 processes. First process, the fault detection decision algorithm must be detected using the DWT to classifying between fault condition and normal condition. Fault detection decision algorithm is processed using positive sequence current signal. The DWT with Daubechies4 (db4) is employed to decompose high frequency components from fault signals. The obtained results can be summarized that the fault detection decision algorithm can be detected the fault condition with average accuracy 100%. For next process, the fault type decision algorithm can be classifying the type of fault using the fuzzy logic. The maximum coefficients detail of DWT is considered as the input variables in constructing a fuzzy logic decision algorithm. By performing many simulations, the obtained results shown that the proposed decision algorithm can classify the fault type with average accuracy 88.00%.

Acknowledgments. This authors wish to gratefully acknowledge financial support for this research (No. KREF025606) from the King Mongkut's Institute of Technology Ladkrabang Research fund, Thailand. They also would like to thank the Srinakharinwirot University and PEA for the DigSILENT presented in this article.

References

1. Sharafi, A., Sanaye-Pasand, M., Jafarian, P.: Ultra-high-speed protection of parallel transmission lines using current travelling waves. IET Gener. Transm. Distrib. **5**(6), 656–666 (2011)
2. Makming, P., Bunjongjit, S., Kunakorn, A., Jiriwibhakorn, S., Kando, M.: Fault diagnosis in transmission lines using wavelet transforms. In: Proceedings of the IEEE Transmission and Distribution Conference, Yokohama, Japan, pp. 2246–2250 (2002)
3. Perez, F.E., Orduna, E., Guidi, G.: Adaptive wavelets applied to fault classification on transmission lines. IET Gener. Transm. Distrib. **5**(7), 694–702 (2011)
4. Yusuff, A.A., Jimoh, A.A., Munda, J.L.: Determinant-based feature extraction for fault detection and classification for power transmission lines. IET Gener. Transm. Distrib. **5**(12), 1259–1267 (2011)
5. Sidhu, T.S., Xu, Z.: Detection of incipient faults in distribution underground cables. IEEE Trans. Power Deliv. **25**(3), 1363–1371 (2010)
6. El Din, E.S.T., Gilany, M., Aziz, M.M.A., Ibrahim, D.K.: A wavelet based fault location technique for aged power cables. IEEE Power Eng. Soc. **3**, 2485–2491 (2005)
7. Han, J., Kim, W.-K., Lee, J.-W., Kim, C.-H.: Fault type classification in transmission line using STFT. In: Proceedings of the 11th International Conference on Developments in Power Systems Protection, pp. 1–5 (2012)
8. Apisit, C., Ngaopitakkul, A.: Identification of fault types for underground cable using discrete wavelet transform. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2010, Hongkong, pp. 1262–1266 (2010)

9. Pandey, A., Younan, N.: Wavelet as a diagnostic tool for fault classification and identification in underground power cable. In: Proceedings of the Annual Report Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), pp. 76–79 (2011)
10. Ngaopitakkul, A., Pothisarn, C.: Discrete Wavelet transform and back-propagation neural networks algorithm for fault location on single-circuit transmission line. In: Proceedings of the International Conference on Robotics and Biomimetics (ROBIO2008), pp. 1613–1618 (2008)
11. Ke, W., Weirong, C., Qi, L.: Power system fault identification method based on multi-wavelet packet and artificial neural network. In: Proceedings of the International Conference on Intelligent Systems Design and Engineering Application (ISDEA), pp. 1457–1462 (2012)
12. Dubey, H.C., Tiwari, A.K., Ray, P.K., Mohanty, S.R., Kishor, N.: A novel fault classification scheme based on least square SVM. In: Proceedings of the Students Conference on Engineering and Systems (SCES), pp. 1–5 (2012)
13. Seyedtabaii, S.: Improvement in the performance of neural network based power transmission line fault classifiers. IET Gener. Transm. Distrib. **6**(8), 731–737 (2012)
14. Livani, H., Evrenosoglu, C.Y.: A fault classification method in power systems using DWT and SVM classifer. In: Proceedings of the IEEE PES Transmission and Distribution Conference and Exposition (T&D), pp. 1–5 (2012)
15. Chen, J., Aggarwal, R.K.: A new approach to EHV transmission line fault classification and fault detection based on the wavelet transform and artificial intelligence. In: Proceedings of the IEEE Power and Energy Society General Meeting, pp. 1–8 (2012)
16. Lout, K., Aggarwal, R.K.: A feedforward artificial neural network approach to fault classification and location on a 132 kV transmission line using current signals only. In: Proceedings of the 47th International Universities Power Engineering Conference (UPEC), pp. 1–6 (2012)
17. Jafarian, P., Sanaye-Pasand, M.: High-frequency transients-based protection of multiterminal transmission lines using the SVM technique. IEEE Trans. Power Deliv. **28**(1), 188–196 (2013)
18. Ferrero, A., Sangiovanni, S., Zappitelli, E.: A Fuzzy-set approach to fault-type identification in digital relaying. IEEE Trans. Power Deliv. **10**(1), 169–175 (1995)
19. Nguyen, T., Liao, Y.: Transmission line fault type classification based on novel features and neuro-fuzzy system. Electr. Power Compon. Syst. **38**(6), 695–709 (2010)
20. Youssef, O.A.S.: Combined fuzzy-logic wavelet-based fault classification technique for power system relaying. IEEE Trans. Power Deliv. **19**(2), 582–589 (2004)
21. Pradhan, A.K., Routray, A., Pati, S., Pradhan, D.K.: Wavelet fuzzy combined approach for fault classification of a series-compensated transmission line. IEEE Trans. Power Deliv. **19** (4), 1612–1618 (2004)
22. Chiradeja, P., Pothisarn, C.: Discrete wavelet transform and fuzzy logic algorithm for identification of fault types on transmission line. In: Proceedings of the 8th IET International Conference on Advances in Power System Control, Operation and Management (APCOM 2009), Hongkong (2008)
23. Khalid, H.M., Khoukhi, A., Al-Sunni, F.M.: Fault detection and classification using kalman filter and genetic neuro-fuzzy systems. In: Proceedings of Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), pp. 1–8 (2011)
24. Rebizant, W., Solak, K.: Transmissin line differential protection with fuzzy signal processing support. In: Proceedings of 7th International Conference on Electrical and Electronics Engineering (ELECO), pp. 162–166 (2011)

25. Moshtagh, J., Rafinia, A.: A new approach to high impedance fault location in three-phase underground distribution system using combination of fuzzy logic & wavelet analysis. In: Proceedings of 11th International Conference on Environment and Electrical Engineering (EEEIC), pp. 90–97 (2012)
26. Cecati, C., Razi, K.: Fuzzy-logic-based high accurate fault classification of single and double-circuit power transmission lines. In: Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion, pp. 883–889 (2012)
27. Long, Z., Younan, N.H., Bialek, T.O.: Underground power cable fault detection using complex wavelet analysis. In: Proceedings of International Conference on High Voltage Engineering and Application, pp. 59–62 (2012)
28. Chiradejal, P., Pothisarn, C.: Discrete wavelet transform and fuzzy logic algorithm for identification of fault type on transmission line. In: Proceedings of 8th IET International Conference on Advances in Power System Control, Operation and Management (APSCOM2009), pp. 1–6 (2009)
29. Klomjit, J., Ngaopitakkul, A.: Selection of proper input pattern in fuzzy logic algorithm for classifying the fault type in underground distribution system. In: 2016 IEEE Region 10 Conference (TENCON), Singapore, pp. 2650–2655 (2016)
30. Nag, A., Yadav, A.: Fault classification using Artificial Neural Network in combined underground cable and overhead line. In: 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, pp. 1–4 (2016)
31. Rafinia, A., Moshtagh, J.: A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ANN and FLS. Int. J. Electr. Energy Syst. **55**, 261–274 (2014). ISSN: 0142-0615
32. Zin, A.A.M., Saini, M., Mustafa, M.W., Sultan, A.R., Rahimuddin: New algorithm for detection and fault classification on parallel transmission line using DWT and BPNN based on Clarke's transformation. Neurocomputing **168**, 983–993 (2015). ISSN 0925-2312
33. Prasad, A., Edward, J.B.: Application of wavelet technique for fault classification in transmission systems. Proc. Comput. Sci. **92**, 78–83 (2016). ISSN 1877-0509

Investigation and Reduction of Effects of Transient Signals for Switching Capacitor into a Power System by Using an Experimental Test Set

Suntiti Yoomak(✉), Chaichan Pothisarn, Chaiyan Jettanasen, and Atthapol Ngaopitakkul

Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Chalongkrung Road, Ladkrabang, Bangkok, Thailand
suntitiatim@hotmail.com, atthapol.ng@kmitl.ac.th

Abstract. This paper aims to investigate switching capacitor bank of the 115 kV Nong Chok substation under the Electricity Generating Authority of Thailand (EGAT). The substation comprises of 3 steps capacitor banks with reactive power of 48 Mvar in each step. In case study, the substation is downscaled to be an experimental unit with 415 V and 5 Mvar in each step in laboratory. Inrush currents, the behavior of transient signals, that occurs when capacitors are switched into the system are studied and analyzed. To reduce the effect of switching capacitors, current limiting reactors connected in series with the capacitors are proposed. In addition, a zero-crossing circuit is designed to control switching angle of the capacitors, since it has a significant effect on the inrush currents. The results of experiment are compared with two case studies: switching capacitors without integrated 7% of reactors and switching capacitors with integrated 7%. It can be summarized that the switching capacitor without integrated reactors has inrush currents change based on the switched angles of the capacitors. However, the switching capacitor with integrated reactors gives inrush current values are almost approximate in each angles of switching and they are lower the case of the switching capacitor without integrated reactors. Nevertheless, reactor integration into the system leads to high current values at the steady states.

Keywords: Capacitor bank · Inrush current · Reactor · Switching

1 Introduction

Nowadays, we need to consume a lot of electrical energy for daily activity but some of the power systems, distributing and transmitting the electrical energy to the consumers, have unsuitable power factors resulting in low efficiency. Thus, a capacitor bank is one of the systems used to enhance the power system efficiency. Many research articles present that capacitor bank installation to an electrical system provides many benefits, for instance power factor correction, voltage support, reduction of harmonic distortion effect in transmission systems and increase of active power transfer capacity [1–3].

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_53

The investigation of efficiency improvement of power systems by integrating different types of capacitor bank units is presented in K. Tilakul [4] and C. Rivera [5] works. A capacitor bank comprises of capacitor units connected in series and parallel inside an enclosed bank. However, capacitor bank switching has a negative effect on power system operation which is inrush currents, overvoltage transience and harmonic problems, hence electrical equipment damage.

The inrush current from switching can be very high during short periods, which leads to the failure of electrical equipment operation. To limit the inrush current, J. C. Das [6] and Mirza Softić [7] present a power factor controller. Thyristors are used to control the power factor of the capacitor banks that are connected in series. Results indicate that this technique is effective and reliable. The installation of an electromagnetic relay with solid-state transient limiter into a capacitor bank is proposed [8–10]. There are two operation modes by using a thyristor to address any transient overvoltage and inrush current problems: limiting mode and bypass mode, which have a very simple structure and reliable performance [8]. Due to an increased voltage stress of a thyristor switch, paper [9] uses IGBT to turn-on and turn-off a switch without an inrush current at the voltage zero-crossing. IEEE Std C37 [11] proposes guidelines for shunt power capacitor bank and filter capacitor bank protection, aiming for many shunt capacitor installations and designs. Software simulation is used to investigate inrush current for single and back-to-back capacitor banks, revealing that the pre-insertion resistor technique can significantly reduce transient [12].

In this paper, switching capacitor bank of the 115 kV Nong Chok substation under the Electricity Generating Authority of Thailand (EGAT) is studied. The substation is downscaled to be an experimental unit in laboratory to study and analyze inrush currents. After that, current limiting reactors connected in series with the capacitors are installed to reduce the effect of switching capacitors.

2 Experimental Setup

The experimental model is set to simulate the 115 kV Nong Chok substation under the Electricity Generating Authority of Thailand (EGAT). A single line diagram of the capacitor bank system is shown in Fig. 1. The substation comprises of 3 steps capacitor banks with reactive power of 48 MVar in each step. The experimental test set is downscaled from 115 kV to 415 V in laboratory. In addition, the size of capacitors and other equipment is determined by using per unit calculation. Hence, the total capacity power is 15 MVar, 5 MVar in each step, with ungrounded wye for internal connection.

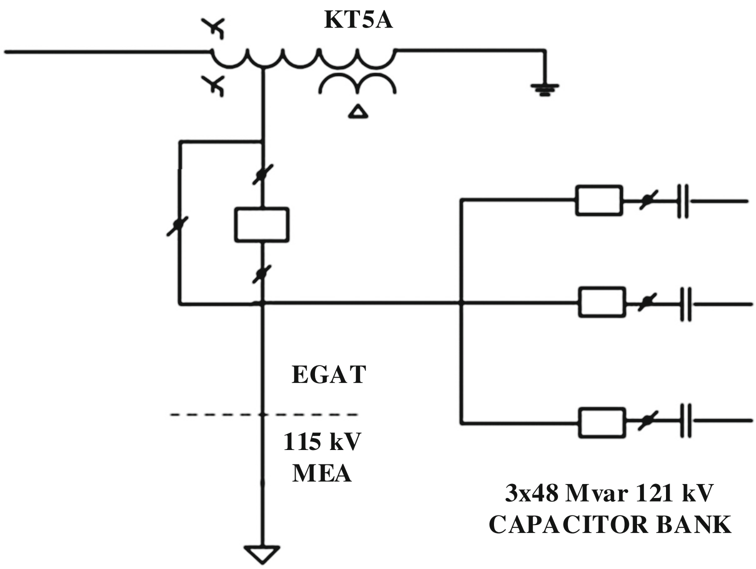
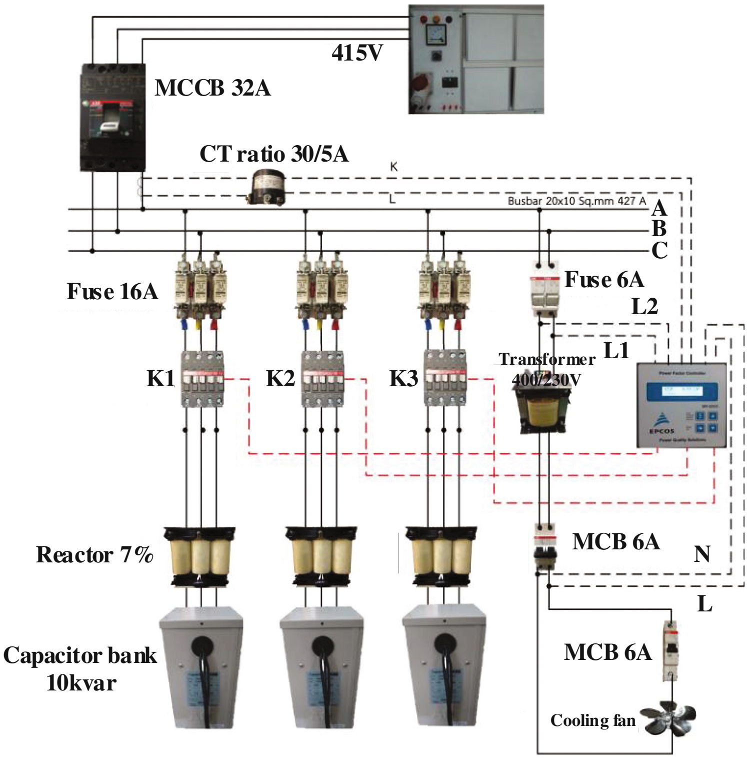


Fig. 1. The single line diagram of the capacitor bank system in the Nong Chok substation

The experimental setup receives a voltage of 415 V from a variable voltage transformer used to step up the voltage of the power supply in the laboratory. The voltage and current measured at phase A by a power factor controller are evaluated and processed, after that it sends a signal to the switching capacitors with the magnetic contactors K1, K2, and K3. The operation of the experimental test set is divided into two parts: a power circuit and a control circuit.

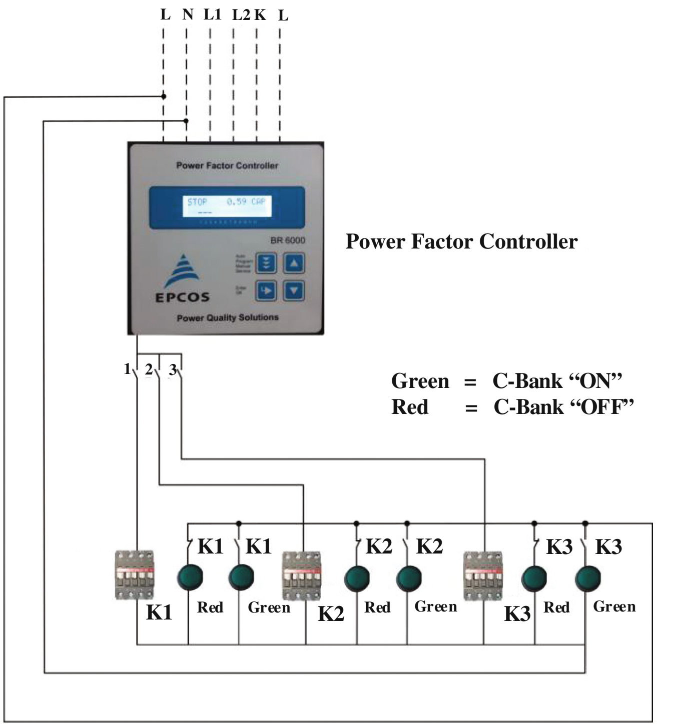
Figure 2(a) shows a schematic diagram of the power circuit of the experimental test set. The busbars of the power circuit receive a 3-phase voltage of 415 V three phase via a main circuit breaker. There are two main circuits connected with them. Firstly, the 415-line voltage of phase A is changed to 230 V by a step down voltage transformer in order to supply the control circuit (power factor controller) and a thermal ventilation system (a cooling fan). Next, the switching of the capacitor bank in each step consists of a HRC fuse, a magnetic contactor, a 7% reactor and a capacitor.

A schematic diagram of the control circuit of the capacitor bank system in the experimental test set is shown in Fig. 2(b) The control system comprises the power factor controller used to control the switching capacitor in each step. The power factor controller obtains current and voltage values from the current transformer (CT) and potential transformer (PT). The power factor of the system is calculated from these values to compare with the power factor value set inside the program. After that, the capacitor in each step is closed or opened with the magnetic contactor by using on/off signals from the power factor controller.



(a) a schematic diagram of the power circuit of the experimental test set

Fig. 2. The operation of the experimental test set (a) a schematic diagram of the power circuit of the experimental test set (b) a schematic diagram of the control circuit of the capacitor bank system



(b) a schematic diagram of the control circuit of the capacitor bank system

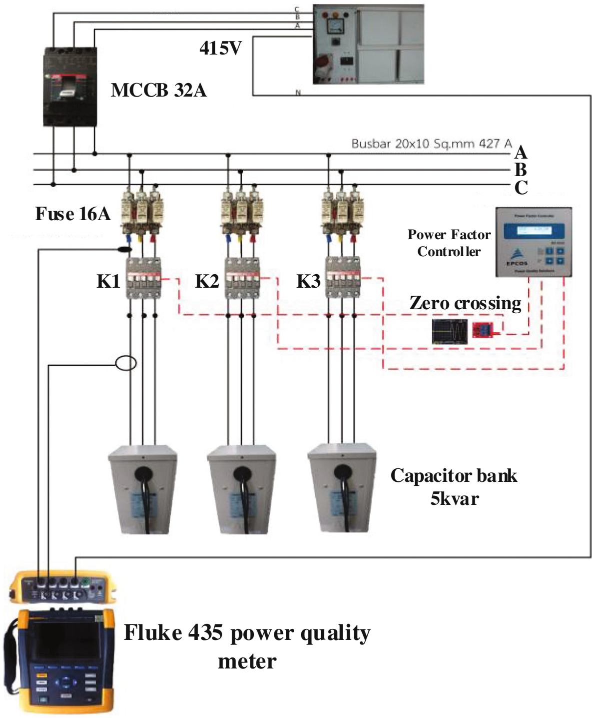
Fig. 2. (*continued*)

3 Experimental Results

The inrush currents from switching capacitors into the experimental setup are measured by using a Fluke 435 power quality meter. There are two case studies: switching capacitors without integrated 7% of reactors and switching capacitors with integrated 7%. In these studies, a zero-crossing detection is installed to control the angle of switching capacitors.

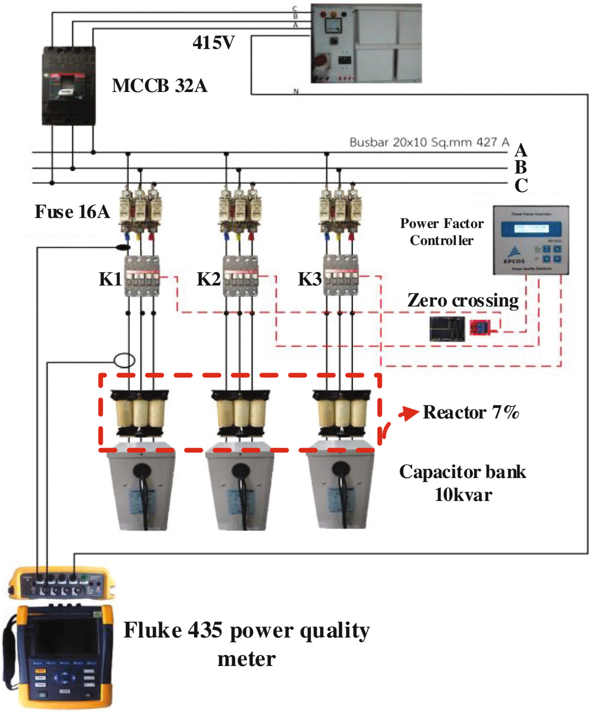
Figure 3(a) gives the schematic diagram of switching capacitors without integrated 7% of reactors. For inrush current measurement, current clamps are installed between magnetic contactors (K1, K2, and K3) and capacitors, while voltage probe installation is set between HRC fuses and the magnetic contactors. For experiment, the test set is supplied from the three-phase voltage of 415 V in the laboratory via a main circuit breaker (32 A, MCCA). The capacitor in each step (1st, 2nd, and 3rd steps respectively) is switched by using the power factor controller to deliver signals to the magnetic contactors K1, K2, and K3. Electrical parameters including the inrush currents are recorded by the power quality meter and results are presented in Table 1. Figure 4 depicts the current and voltage waveforms of switching capacitor bank without integrated 7% of reactors at the phase angles of 0° and 90°.

For the schematic diagram of switching capacitors with integrated 7% of reactors as depicted in Fig. 3 (b), the measurement equipment setting and measurement remain the same case of the switching capacitors without integrated 7% of reactors. In this model, the 7% of reactors are integrated between the magnetic contactors and capacitors in order to reduce effects of the inrush currents; the results are shown in Table 2 and the current and voltage waveforms of the switching capacitors are shown in Fig. 5.



(a) switching capacitors without integrated 7% of reactors

Fig. 3. The schematic diagram of switching capacitors (a) switching capacitors without integrated 7% of reactors (b) switching capacitors with integrated 7% of reactors

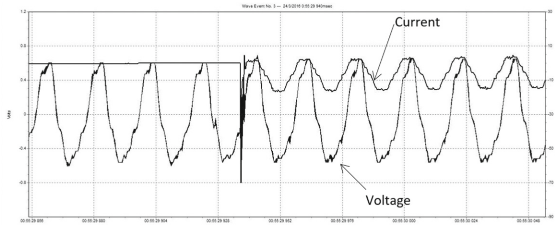


(b) switching capacitors with integrated 7% of reactors

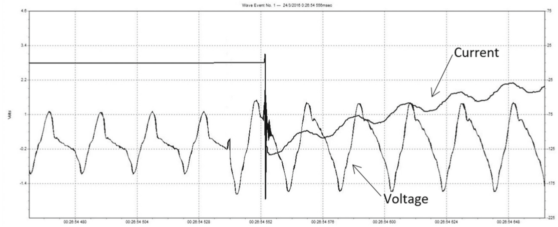
Fig. 3. (*continued*)

Table 1. Results of inrush current in the case of switching capacitors without integrated 7% of reactors

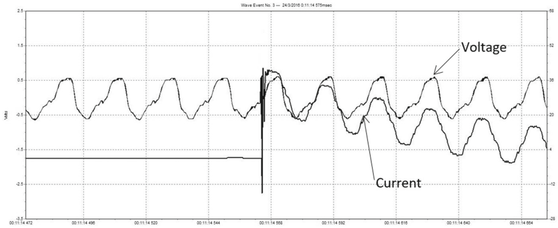
Phase angle (degree)	1st step			2nd step			3rd step		
	Inrush current (A)	Steady state period (ms)	Current (A)	Inrush current (A)	Steady state period (ms)	Current (A)	Inrush current (A)	Steady state period (ms)	Current (A)
0	−69.98	179	8.12	41.56	161	8.93	−21.37	127	9.23
45	−85.06	336	8.35	52.98	225	8.95	37.60	190	9.09
90	−197.60	712	8.57	−176.52	554	8.97	132.31	595	9.21
135	68.28	350	8.74	52.30	204	9.00	34.59	189	9.05
180	75.21	204	8.21	53.28	194	8.97	22.39	154	9.12
225	−82.92	377	8.32	56.01	194	9.07	33.91	168	9.14
270	197.66	698	8.78	182.76	688	8.88	146.67	611	9.25
315	−82.95	307	8.67	54.38	270	8.98	−36.05	242	9.32



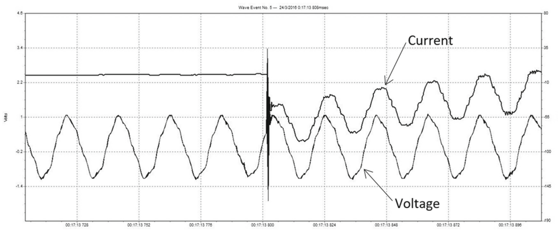
(a) switching capacitors at 0° phase angle in step 1



(b) switching capacitors at 90° phase angle in step 1

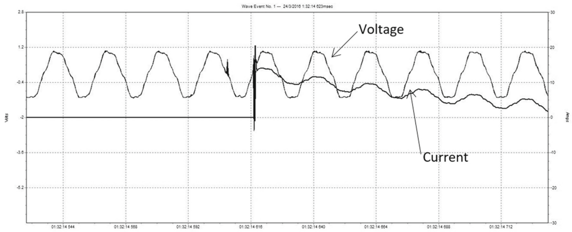


(c) switching capacitors at 0° phase angle in step 2

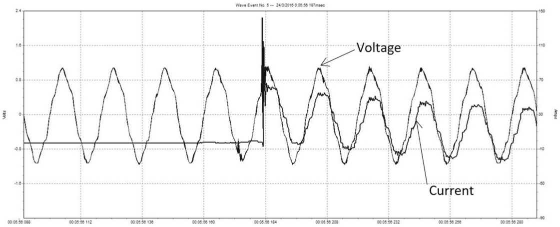


(d) switching capacitors at 90° phase angle in step 2

Fig. 4. Current and voltage waveforms of switching capacitor bank without integrated 7% of reactors (a) switching capacitors at 0° phase angle in step 1(b) switching capacitors at 90° phase angle in step 1(c) switching capacitors at 0° phase angle in step 2(d) switching capacitors at 90° phase angle in step 2(e) switching capacitors at 0° phase angle in step 3(f) switching capacitors at 90° phase angle in step 3



(e) switching capacitors at 0° phase angle in step 3

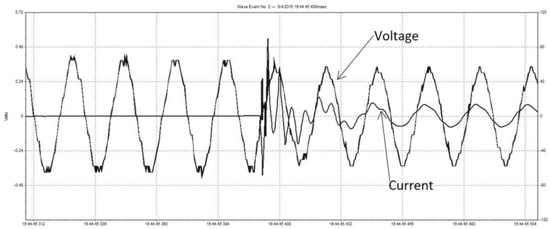


(f) switching capacitors at 90° phase angle in step 3

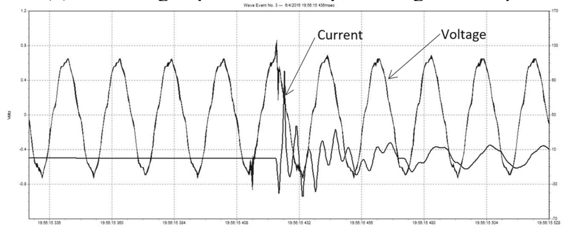
Fig. 4. (*continued*)

Table 2. Results of inrush current in the case of switching capacitors with integrated 7% of reactors

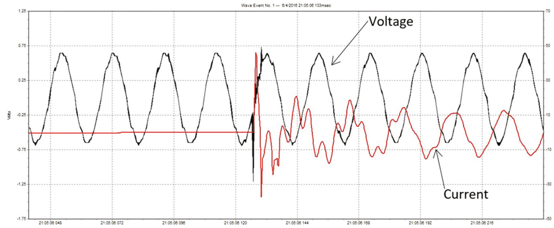
Phase angle (degree)	1st step			2nd step			3rd step		
	Inrush current (A)	Steady state period (ms)	Current (A)	Inrush current (A)	Steady state period (ms)	Current (A)	Inrush current (A)	Steady state period (ms)	Current (A)
0	95.87	79	12.56	91.08	48	14.58	−90.45	39	15.92
45	−95.68	74	12.56	−91.98	53	14.40	−90.92	48	15.90
90	97.98	84	12.45	94.56	68	14.38	92.86	63	15.89
135	−96.45	71	12.57	−91.25	58	14.39	−92.14	52	15.75
180	96.05	64	12.61	−91.09	51	14.39	−91.00	49	15.84
225	97.66	81	12.49	93.87	65	14.40	91.85	55	15.90
270	96.78	73	12.56	−93.56	61	14.14	−91.94	56	15.86
315	95.98	68	12.60	92.87	60	14.19	91.00	51	15.94



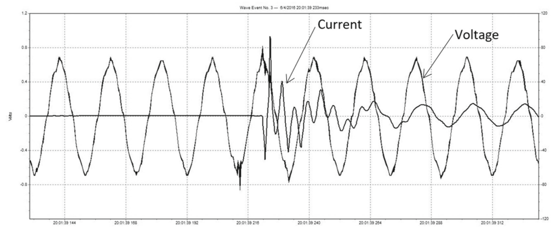
(a) switching capacitors at 0° phase angle in Step 1



(b) switching capacitors at 90° phase angle in Step 1

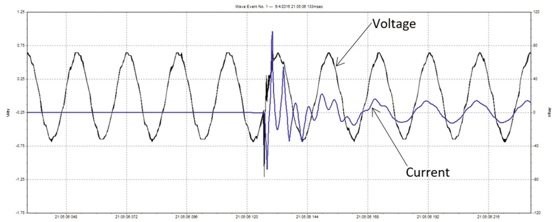


(c) switching capacitors at 0° phase angle in Step 2

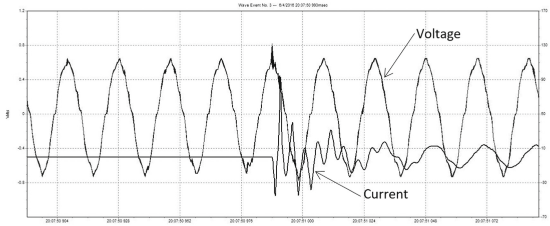


(d) switching capacitors at 90° phase angle in Step 2

Fig. 5. Current and voltage waveforms of switching capacitor bank with integrated 7% of reactors (a) switching capacitors at 0° phase angle in Step 1 (b) switching capacitors at 90° phase angle in Step 1 (c) switching capacitors at 0° phase angle in Step 2 (d) switching capacitors at 90° phase angle in Step 2 (e) switching capacitors at 0° phase angle in Step 3 (f) switching capacitors at 90° phase angle in Step 3



(e) switching capacitors at 0° phase angle in Step 3



(f) switching capacitors at 90° phase angle in Step 3

Fig. 5. (*continued*)

4 Conclusions

This paper aims to study capacitor bank switching transients in a substation by using an experimental test unit. To reduce the effect of switching capacitors, current limiting reactors connected in series with the capacitors are proposed. The results of the switching capacitors without integrated 7% of reactors (see Table 1) show that the inrush current values obtained from the experimental setup have change based on the switched angles of the capacitors. Switching of angles at 90 and 270 degrees causes the maximum inrush current value when comparing with another angles. By contrast, the angles of 0 and 180 degrees give the minimum inrush current value.

For the results of the switching capacitors with integrated 7% of reactors (see Table 2), it can be noticed that inrush current values are almost approximate in each angles of switching capacitors and lower than the inrush currents of the former case. The 7% reactors connected in series with the capacitors increase in an inductance value per a phase, resulting in decreased inrush current values and short periods of time for steady states. However, 7% reactor integration into the experimental setup allows for high current values at the steady states.

Acknowledgments. This authors wish to gratefully acknowledge financial support for this research (No. KREF045507) from the King Mongkut's Institute of Technology Ladkrabang Research fund, Thailand.

References

1. Grünbaum, R., Ingeström, G., Ekehov, B., Marais, R.: 765 kV series capacitors for increasing power transmission capacity to the Cape Region. In: IEEE Power and Energy Society Conference and Exposition in Africa: Intelligent Grid Integration of Renewable Energy Resources (PowerAfrica), Johannesburg, pp 1–8 (2012). doi:10.1109/PowerAfrica.2012.6498623
2. Aman, M.M., Jasmon, G.B., Bakar, A.H.A., Mokhlis, H., Karimi, M.: Optimum shunt capacitor placement in distribution system: a review and comparative study. Renew. Sustain. Energy Rev. **30**, 429–439 (2014). doi:10.1016/j.rser.2013.10.002
3. Issouribehere, P.E., Issouribehere, F., Barbera, G.A., Gomez, D.: Measurements and studies of harmonics and switching, transients in large HV shunt capacitor banks. In: 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, pp. 1–8 (2007). doi:10.1109/PES.2007.385670
4. Tilakul, K., Buasri, P., Kaewrawang, A., Triyangkulsri, J.: Capacitor location and size determination to reduce power losses of a distribution feeder in Lao PDR. Int. J. Comput. Electr. Eng. **4**(1), 32 (2012)
5. Rivera, C., Baretty, R.: Energization Study of 10.4MVAR Capacitor Bank in 38KV System. University of Kentucky Master's Theses (2011)
6. Das, J.C.: Analysis and control of large-shunt-capacitor-bank switching transients. IEEE Trans. Ind. Appl. **41**(6), 1444–1451 (2005). doi:10.1109/TIA.2005.858279
7. Softić, M., Tokić, A., Uglešić, I.: Measurement, modeling and simulation of capacitor bank switching transients. IFAC Proc. Vol. **45**(2), 1254–1259 (2012). doi:10.3182/20120215-3-AT-3016.00222
8. Ghanbari, T., Farjah, E., Zandnia, A.: Solid-state transient limiter for capacitor bank switching transients. IET Gener. Transm. Distrib. **7**(11), 1272–1277 (2013). doi:10.1049/iet-gtd.2012.0495
9. Ruszczyk, A., Kóska, K., Janisz, K.: Solid-state switch for capacitors bank used in reactive power compensation. In: 2015 International School on Nonsinusoidal Currents and Compensation (ISNCC), Lagow, pp. 1–5 (2015). doi:10.1109/ISNCC.2015.7174699
10. Huhman, B.M., Neri, J.M.: Effect of fuse resistance on EML capacitor bank solid-state switching. In: 2008 IEEE International Power Modulators and High-Voltage Conference, Las Vegas, NE, pp. 89–92 (2008). doi:10.1109/IPMC.2008.4743585
11. IEEE Guide for the Protection of Shunt Capacitor Banks. IEEE Std C37.99-2012 (Revision of IEEE Std C37.99-2000). doi:10.1109/IEEESTD.2013.6482148
12. Beanland, M., Speas, T., Rostron, J.: Pre-insertion resistors in high voltage capacitor bank switching. In: Western Protective Relay Conference, pp. 1–14 (2004)

Practical Notes on Applying Generalised Stochastic Orderings to the Study of Performance of Classification Algorithms for Low Quality Data

Patryk Żywica[1(✉)], Katarzyna Basiukajc[1], and Inés Couso[2]

[1] Faculty of Mathematics and Computer Science, Adam Mickiewicz University in Poznań, Umultowska 87, 61-614 Poznań, Poland
bikol@amu.edu.pl

[2] Department of Statistics and Operational Research, University of Oviedo, C/ Luis Berrocal s/n, Despacho 4.1.12, 33071 Gijón, Spain
couso@uniovi.es

Abstract. This paper presents an approach to applying stochastic orderings to evaluate classification algorithms for low quality data. It discusses some known stochastic orderings along with practical notes about their application to classifier evaluation. Finally, a new approach based on fuzzy cost function is presented. The new method allows comparing any two classifiers, but does not require a precise definition of the cost function. All proposed methods were evaluated on real life medical data. The obtained results are very similar to those previously reported but comparatively much weaker assumptions about costs values are adopted.

Keywords: Classification · Loss function · Stochastic ordering · Low quality data · Fuzzy random variable

1 Introduction

As long as machine learning algorithms are becoming more and more popular and their area of application is simultaneously expanding, we are facing a wide range of newly arising problems. One of them is the evaluation of algorithms concerning real-life conditions with some unusual restrictions.

A typical binary classification problem's goal is to find a model $f : \mathcal{X} \to \mathcal{Y}$ assigning the categories from $\mathcal{Y} = \{0, 1\}$ to lists of attributes mapped by $\mathbb{Y} : \Omega \to \mathcal{Y}$ and $\mathbb{X} : \Omega \to \mathcal{X}$, respectively. These kinds of models can be evaluated by widely known evaluation functions such as: accuracy, precision, recall as well as F1-score.

This work was partially supported by the Polish National Science Centre grant number 2016/21/N/ST6/00316.

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7_54

Unfortunately, not every classification problem matches the definition above [1–4]. For instance, in some medical diagnostic problems, the physician is not always able to collect all the data needed for the diagnosis due to the time and money investment this constitutes [5,6]. In such situations we may want to evaluate classifiers similarly to the way real-life doctor's decisions are evaluated. We thus, take into account the data quality, the level of uncertainty, and allow the possibility of receiving a "not available" (NA) value as the output of the classification model. In turn, this prevent the use of mentioned evaluation functions so another solution must be found.

One of the proposed approaches is to introduce the cost function matching model outcomes (projected as *true positive*, *true negative*, etc.) with the cost of the real-life consequences they can contribute to [7]. Although this solution works, the selection of cost matrix values is subjective with possibly divergent opinions held by experts. Furthermore, some small changes in the cost matrix values may cause significant changes in the final classification result (lack of robustness). Another problem related especially to medical decision evaluation is that costs of individual decisions may not be known. For example, for each patient the cost of *false–negative* may be different depending on his or her other medical conditions [8].

Another idea to study the performance of such classification algorithms is to apply stochastic orderings. This method was proposed in [9,10] and it fits into the situation presented. It will be presented and extended in following sections.

The remainder of the paper is organised as follows. In Sect. 2 we present basic notions regarding cost function and stochastic orders. Section 3 describes an evaluated dataset as well as results for three stochastic order based classifier evaluation methods. In Sect. 4 we present details of our proposed approach as well as some analysis of obtained results. Conclusions and further work appear in Sect. 5.

2 Basic Notions

Let $\mathcal{Y}$ be the output space of the models and let $\Delta : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ be the *loss (cost) function* used to find the best classifier. The main goal of the loss function is to penalise wrong outcomes in order to enable finding the best model as the model with the minimal loss. It means, that loss function values usually become positive when the predictions do not match reality.

There also exists a dual definition of the *loss function* called *reward function* which we will use in this paper to match the common stochastic orderings literature. It's a simple opposite to the loss function which means the greater value of the reward function is, the more relevant outcomes are. Therefore, the best model can be found at its maximum. We can associate the random variable *reward* $U_{\Delta,f}^{(\mathbb{X},\mathbb{Y})} : \Omega \to \mathbb{R}$ with the model f using the definition as follows:

$$U_{\Delta,f}^{(\mathbb{X},\mathbb{Y})} = -\Delta(\mathbb{Y}(\omega), f(\mathbb{X}(\omega)) \qquad \forall \omega \in \Omega. \tag{1}$$

Table 1. Cost function (matrix) for example binary classification (a) and its extension to uncertain classification (b).

(a)

		predicted	
		benign	malignant
actual	benign	TN: 0	FP: 2.5
	malignant	FN: 5	TP: 0

(b)

	predicted		
	benign	malignant	NA
benign	TN: 0	FP: 2.5	N0: 1
malignant	FN: 5	TP: 0	N1: 2

Let $X, Y : \Omega \to \mathbb{R}$ be two random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$. From many available kinds of stochastic orderings, of most interest from classification performance evaluation point of view are [11]:

1. Dominance in the sense of expected utility [12]. Given an increasing function $u : \mathbb{R} \to \mathbb{R}$, X dominates Y wrt u (denoted $X \succeq_u Y$) if

$$E_P(u(X)) \geq E_P(u(Y))\,. \tag{2}$$

2. First order stochastic dominance [13]. X dominates Y (denoted $X \succeq_{1st} Y$) if

$$\forall_{x\in\mathbb{R}} \quad P(X > x) \geq P(Y > x)\,. \tag{3}$$

   It is well known that $X \preceq_{1st} Y$ if and only if $X \preceq_u Y$, for all increasing utility functions $u : \mathbb{R} \to \mathbb{R}$.
3. Statistical preference [14]. X is statistically preferred to Y (denoted $X \succeq_{sp} Y$) if

$$P(X > Y) \geq P(Y > X)\,. \tag{4}$$

Based on this we can present the notion of the $(\succeq, \Delta)$–*domination* proposed by Couso and Sánchez [9].

Definition 1. *Let $f_1 : \mathcal{X} \to \mathcal{Y}$ and $f_2 : \mathcal{X} \to \mathcal{Y}$ be the classification models and $\succeq$ be any stochastic ordering. f_1 $(\succeq, \Delta)$–dominates f_2 if*

$$U_{\Delta,f_1}^{(\mathbb{X},\mathbb{Y})} \succeq U_{\Delta,f_2}^{(\mathbb{X},\mathbb{Y})}\,. \tag{5}$$

Example 1. *Let us consider the binary classification problem which refers to determination whether the tumor is malignant (M) or benign (B). Let $\mathbb{X}$ be the random vector of attributes and $\mathbb{Y}$ – the outcome. Let the cost matrix be given as in Table 1a.*

According to the definition of the reward function we have:

$$P(U_{\Delta,f}^{(\mathbb{X},\mathbb{Y})} > c) = \begin{cases} 1, & \text{if } c < -5, \\ 1 - P(\mathbb{Y} = M, f(\mathbb{X}) = B), & \text{if} -5 \leq c < -2.5, \\ P(\mathbb{Y} = f(\mathbb{X})), & \text{if} -2.5 \leq c < 0, \\ 0, & \text{if } c \geq 0. \end{cases} \tag{6}$$

The equation above together with Definition 1 lead to the conclusion that f_1 $(\succeq_{1st}, \Delta)$*–dominates* f_2 *if and only if:*

$$P(\mathbb{Y} = M, f_1(\mathbb{X}) = B) \le P(\mathbb{Y} = M, f_2(\mathbb{X}) = B), \tag{7}$$

$$P(\mathbb{Y} = f_1(\mathbb{X})) \ge P(\mathbb{Y} = f_2(\mathbb{X})). \tag{8}$$

Moreover, this stays true for any cost matrix for which

$$\Delta(\mathbb{Y} = M, f(\mathbb{X}) = B) \ge \Delta(\mathbb{Y} = B, f(\mathbb{X}) = M) \ge \Delta(\mathbb{Y} = f(\mathbb{X}))\,. \tag{9}$$

This means that $(\succeq_{1st}, \Delta)$*–dominance does not depend on actual cost values but only on their order.*

Example 2. *Let us consider again the binary classification problem from Example 1 with the same cost matrix. We can distinguish following cases:*

$$P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_2}) = P(f_1(\mathbb{X}) = B, f_2(\mathbb{X}) = M), \quad if \quad \mathbb{Y} = B, \tag{10}$$

$$P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_2}) = P(f_1(\mathbb{X}) = M, f_2(\mathbb{X}) = B), \quad if \quad \mathbb{Y} = M. \tag{11}$$

According to the (10) and (11) and Definition 1, f_1 $(\succeq_{sp}, \Delta)$*–dominates* f_2 *if and only if:*

$$P(\mathbb{Y} = f_1(\mathbb{X}), \mathbb{Y} \ne f_2(\mathbb{X})) \ge P(\mathbb{Y} = f_2(\mathbb{X}), \mathbb{Y} \ne f_1(\mathbb{X})). \tag{12}$$

Example 3. *Let consider medical classification problem for* 10 *patients, cost matrix from Example 1 and three diagnostic models* f_1, f_2 *and* f_3*. Actual diagnoses and predictions are given in Table 2. We can easily calculate that*

$$P(\mathbb{Y} = f_1(\mathbb{X})) = 0.8 \qquad P(\mathbb{Y} = M, f_1(\mathbb{X}) = B) = 0.1 \tag{13}$$

$$P(\mathbb{Y} = f_2(\mathbb{X})) = 0.6 \qquad P(\mathbb{Y} = M, f_2(\mathbb{X}) = B) = 0 \tag{14}$$

$$P(\mathbb{Y} = f_3(\mathbb{X})) = 0.7 \qquad P(\mathbb{Y} = M, f_3(\mathbb{X}) = B) = 0.2 \tag{15}$$

According to Example 1, f_1 $(\succeq_{1st}, \Delta)$*–dominates* f_3*. Unfortunately, models* f_1 *and* f_2 *are incomparable with the respect to* $(\succeq_{1st}, \Delta)$ *criterion.*

Moreover, according to Example 2,

$$P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_2}) = 0.3 \qquad P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_2} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1}) = 0.1 \tag{16}$$

$$P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_3}) = 0.1 \qquad P(U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_3} > U^{(\mathbb{X},\mathbb{Y})}_{\Delta,f_1}) = 0 \tag{17}$$

thus f_1 $(\succeq_{sp}, \Delta)$*–dominates* f_2 *as well as* f_1 $(\succeq_{sp}, \Delta)$*–dominates* f_3*.*

3 Application of Stochastic Orderings to Low Quality Data Classification Performance Evaluation

Definitions and examples presented in Sect. 2 referred to binary classification problem. However, as it was mentioned in the Introduction, there exist real-life problems, where the classification can be uncertain and the outcome may take the *NA* value. In this section we are going to apply concepts from Sect. 2 to this particular case.

Table 2. Diagnoses for patients from Example 3.

Diagnosis	Patients									
	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
actual $\mathbb{Y}(\omega)$	B	B	B	B	B	M	M	M	M	M
model $f_1(\mathbb{X}(\omega))$	B	B	B	B	M	M	M	M	M	B
model $f_2(\mathbb{X}(\omega))$	B	M	M	M	M	M	M	M	M	M
model $f_3(\mathbb{X}(\omega))$	B	B	B	B	M	M	M	M	B	B

3.1 Medical Data

We base our evaluation on test dataset from recent research on application of aggregation operators to incomplete data classification [2]. Original study group consists of 388 patients diagnosed and treated for ovarian tumor in the Division of Gynecological Surgery, Poznan University of Medical Sciences, between 2005 and 2015. Among them, 61% were diagnosed with a benign tumor and 39% with a malignant one. Moreover, 56% of the patients had no missing values in the attributes required by diagnostic models, 40% had a percentage of missing values in the range $(0\%, 50\%]$, and the remainder had more than 50% missing values. The test set consists of patients with real missing data and some proportion of patients with a complete set of features. As a result, the test set consisted of 175 patients. Patients with more than 50% missing values were excluded from the study. The dataset partition is presented visually in Fig. 1.

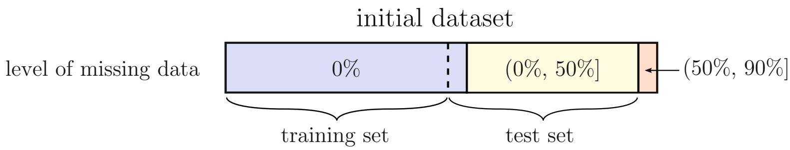


Fig. 1. The division of the dataset. Patients with more than 50% missing values were not included in the experiment. Source [2].

During the research over 4000 different classification strategies were evaluated. Among them 130 were selected into test phase. Our evaluation was performed on outcomes returned by those classifiers on real life test set. For all classifiers it was assumed that no diagnosis may be returned. For more information regarding dataset we refer the reader to original paper [2].

3.2 Expected Utility

One of the methods of dealing with evaluation of the algorithms based on incomplete data is to insert an additional column to the cost matrix as it was proposed in [2]. Then, the loss function may be defined as a sum of costs of all outcomes given by the algorithm. Unfortunately, it's hard to assume that one type of mistake is a certain number of times worse than the other. We also can't say, that every patient has the same loss for every mistake. These make the costs only intuitive and as long as their small changes lead to different final results, we can't be sure that these final results are the best solutions for the patients.

Table 6 presents some selected best classifiers based on this criterion from the dataset described in previous subsection. As can be seen this comparison method is very useful and straightforward. It offers easy to interpret linear order that facilitates selection of the best classifier. The main drawback of this approach concerns the uncertain and subjective selection of cost function. It possible that small change to cost matrix causes significant changes to classifier order and this kind of behaviour is not desired.

3.3 First Stochastic Dominance

To define first of discussed relations, let f be the classifier and Δ the loss function which can be described by the cost matrix similar to one from Table 1b. As the precise cost values do not matter as long as the ordering is saved, let's only assume that the Δ always fulfil (18–21).

$$\Delta(\mathbb{Y} = M, f(\mathbb{X}) = B) \geq \Delta(\mathbb{Y} = B, f(\mathbb{X}) = M) \tag{18}$$

$$\Delta(\mathbb{Y} = B, f(\mathbb{X}) = M) \geq \Delta(\mathbb{Y} = M, f(\mathbb{X}) = NA) \tag{19}$$

$$\Delta(\mathbb{Y} = M, f(\mathbb{X}) = NA) \geq \Delta(\mathbb{Y} = B, f(\mathbb{X}) = NA) \tag{20}$$

$$\Delta(\mathbb{Y} = B, f(\mathbb{X}) = NA) \geq \Delta(\mathbb{Y} = f(\mathbb{X})) \tag{21}$$

Then, analogously to the Example 1, we can conclude, that classifier f_1 $(\succeq_{1st}, \Delta)$–dominates f_2 if and only if:

$$P_{f_1}(TP) + P_{f_1}(TN) \geq P_{f_2}(TP) + P_{f_2}(TN), \tag{22}$$

$$P_{f_1}(TP) + P_{f_1}(TN) + P_{f_1}(N0) \geq P_{f_2}(TP) + P_{f_2}(TN) + P_{f_2}(N0), \tag{23}$$

$$1 - P_{f_1}(FN) - P_{f_1}(FP) \geq 1 - P_{f_2}(FN) - P_{f_2}(FP), \tag{24}$$

$$1 - P_{f_1}(FN) \geq 1 - P_{f_2}(FN). \tag{25}$$

where TP, TN, $N0$, FP, FN are clarified in Table 1b.

Let's take the dataset described in Sect. 3.1 with $(\succeq_{1st}, \Delta)$–dominance relation. Based on the conditions (22–25), we can define a stochastic ordering inside this set. The only change, applied to make this relation irreflexive as well as to avoid cycles, is that the condition *greater than* instead of *greater or equal to* must be fulfilled in at least one of (22–25).

Then, we get the strict partial order which allows us to find maximal elements in the set and as we know they are always better than dominated ones, they can be use as an output to further considerations.

Figure 2a and Table 6 shows the maximal elements from the medical classifiers set along with their costs calculated according to the cost matrix from Table 1b and the information if they are the only ones in the chains they are included or not.

Unfortunately, the number of maximal elements is about one-quarter of all models (33 of 130) so this method couldn't help in determining the best classifier but still it can be used as a very effective process of pre-selection.

3.4 Statistical Preference Stochastic Dominance

In traditional binary classification, Definition 1 leads to the conclusion that $(\succeq_{sp}, \Delta)$–domination depends only on the bigger number of true outcomes given by one of the classifiers. In uncertain classification with possibility of NA the problem becomes more complicated. For example, for $\mathbb{Y} = M$ there are three cases when $P(U_{\Delta,f_1}^{(\mathbb{X},\mathbb{Y})} > U_{\Delta,f_2}^{(\mathbb{X},\mathbb{Y})})$:

- $f_1(\mathbb{X}) = M$ and $f_2(\mathbb{X}) = NA$,
- $f_1(\mathbb{X}) = M$ and $f_2(\mathbb{X}) = B$,
- $f_1(\mathbb{X}) = NA$ and $f_2(\mathbb{X}) = B$.

Similarly, for $\mathbb{Y} = B$, $P(U_{\Delta,f_1}^{(\mathbb{X},\mathbb{Y})} > U_{\Delta,f_2}^{(\mathbb{X},\mathbb{Y})})$ when:

- $f_1(\mathbb{X}) = B$ and $f_2(\mathbb{X}) = NA$,
- $f_1(\mathbb{X}) = B$ and $f_2(\mathbb{X}) = M$,
- $f_1(\mathbb{X}) = NA$ and $f_2(\mathbb{X}) = M$.

Summarising the cases above, we can say that in uncertain classification f_1 $(\succeq_{sp}, \Delta)$–dominates f_2 if the number of times when f_1 gives proper output while f_2 doesn't or f_1 gives NA while f_2 is wrong is bigger than the number of opposite situations.

Table 3. Diagnoses for patients from Example 4.

Diagnosis	Patients				
	ω_1	ω_2	ω_3	ω_4	ω_5
actual	M	M	M	M	M
model f_1	M	M	*NA*	B	B
model f_2	B	B	M	*NA*	*NA*
model f_3	*NA*	*NA*	B	M	M

Example 4. *Let's consider medical classification problem for five patients and three diagnostic models f_1, f_2 and f_3 with actual diagnoses and predictions given in Table 3. We can easily notice, that f_2 has three better predictions than f_1, f_3 has three better predictions than f_2 and finally f_1 has three also better predictions than f_3. It means, that f_2 $(\succeq_{sp}, \Delta)$–dominates f_1, f_3 $(\succeq_{sp}, \Delta)$–dominates f_2 as well as f_1 $(\succeq_{sp}, \Delta)$–dominates f_3.*

The previous example shows that in this case statistical preference based dominance relation makes cycles possible to appear while comparing models. It means, that we can compare each pair of classifiers, but it can be impossible to find the maximal elements in the set with $(\succeq_{sp}, \Delta)$–dominance relation, so the another approach to evaluate the models using this dominance should be defined.

To evaluate models in this case we propose method based on PageRank algorithm [15]. This algorithm is generally used to rate values of the websites by looking how many other sites have reference links to them and how high are the rates of these linking sites. The more links from sites with high rate, the better. All computations are performed on matrix representing graph, where sites are vertices and links are the directed edges pointing to the linked sites. In our situation, models are treated as vertices, the directed edge points to the dominating model and all computations are preserved.

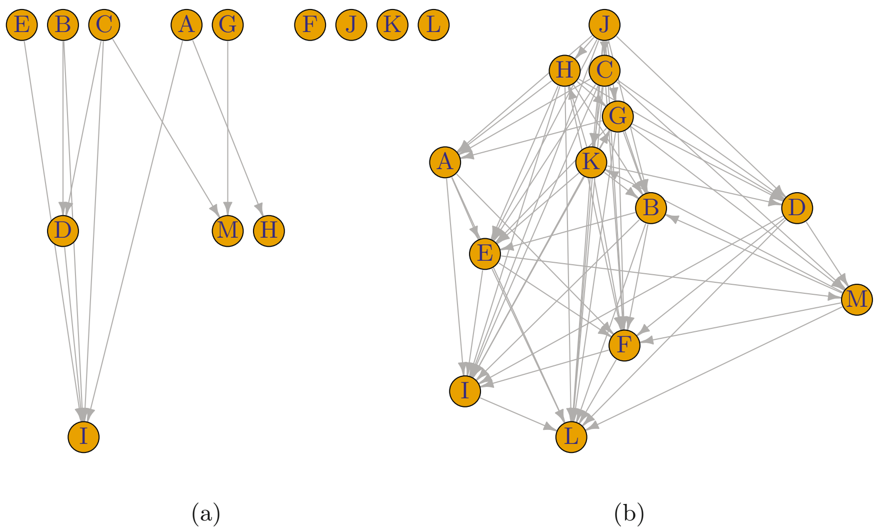


Fig. 2. Graphs showing domination relation for First Stochastic Dominance (a) and Statistical Preference (b) for selected classifiers.

The final score presented in Table 6 is the percentage of time spent in particular classifier vertex while making random PageRank walk. Figure 2 shows the original statistical preference graph for selected best classifiers along with their costs calculated according to the cost matrix from Table 1b.

4 Proposed Approach

4.1 Idea

As can be seen from previous section, each presented method has it own strengths and weaknesses. On the one hand, total cost method gives linear order between all classifiers at the expense of the need to provide concrete numerical cost values. On the other hand, First Stochastic Dominance requires only to know whether one classification outcome is better than other. But this leads to a situation where there are many models that cannot be compared. Application of Statistical Preference results with hard to interpret structure. Although application of Page Rank algorithm gives linear order it is still hard to justify such approach and interpret particular values.

Our aim is to propose a method that retains the ability to compare nearly all classifiers, while imposing the least restrictions on the cost of particular, possibly uncertain, decision. As a starting point we chose the First Stochastic Dominance comparison method, which is highly intuitive and easy to interpret. It can be viewed as a total cost method applied for all possible cost functions [16]. Since, experts are often unable to give precise numerical costs, we propose to model them as fuzzy numbers interpreted, in epistemic way (see [17]), as family of nested confidence sets

$$\widetilde{\Delta} : \mathcal{Y} \times \mathcal{Y} \to \mathcal{FN}(\mathbb{R})\,. \tag{26}$$

As will be shown further in this Section, this will enable comparison of all classifiers with respect to any stochastic dominance.

This approach has one additional benefit. Previously cost values were independent of particular patient and were based only on actual and predicted diagnosis. In real life medical scenario this is not always true. For some patients even proper diagnosis may lead to bad outcome and vice versa. Thanks to this approach actual cost corresponding to diagnosis may vary depending on particular patient conditions as we interpret fuzzy number in epistemic way.

4.2 Definitions

Similarly as in previous sections, for any classification model f we can define *reward fuzzy random variable*

$$\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f} : \Omega \to \mathcal{FN}(\mathbb{R}) \tag{27}$$

as a opposite of cost value:

$$\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f} = -\widetilde{\Delta}(\mathbb{Y}(\omega), f(\mathbb{X}(\omega))) \qquad \forall \omega \in \Omega\,. \tag{28}$$

According to the epistemic interpretation, the reward fuzzy random variable should be also understood in terms of confidence sets.

We will use the Extension Principle based stochastic order proposed by Couso and Dubois [16] for comparing fuzzy sets of random variables. Let $\pi_{\widetilde{X}}(X)$ be the degree of possibility that X is the random variable underlain by the fuzzy random variable $\widetilde{X}$

$$\pi_{\widetilde{X}}(X) = \inf_{\omega \in \Omega} \mu_{\widetilde{X}(\omega)}(X(\omega))\,. \tag{29}$$

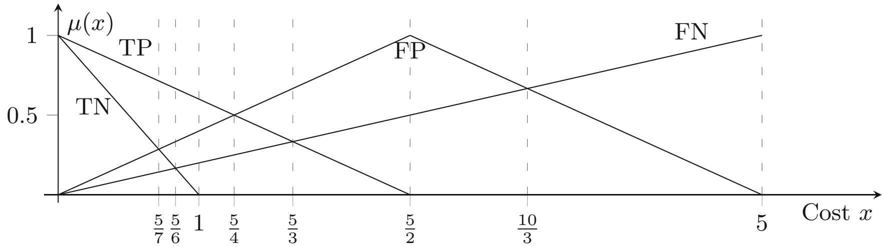


Fig. 3. Fuzzy cost values from Example 5.

Then for any stochastic order $\succeq$ the degree of possibility of dominance between fuzzy random variables $\widetilde{X}$ and $\widetilde{Y}$ can be defined as:

$$\Pi(\widetilde{X} \succeq \widetilde{Y}) = \sup_{X,Y:\, X \succeq Y} \min(\pi_{\widetilde{X}}(X), \pi_{\widetilde{Y}}(Y)) \,. \tag{30}$$

Now we are ready to define our proposed approach to classification model comparison.

Definition 2. *Let $f_1 : \mathcal{X} \to \mathcal{Y}$ and $f_2 : \mathcal{X} \to \mathcal{Y}$ be the classification models. The degree in which f_1 dominates f_2 with respect to stochastic order $\succeq$ and fuzzy cost function $\widetilde{\Delta}$ is defined as*

$$[\![f_1 \succeq f_2]\!]_{\widetilde{\Delta}} = \frac{\Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2})}{\Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2}) + \Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1})} . \tag{31}$$

Such definition, in contrast to simple $[\![f_1 \succeq f_2]\!] = \Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2})$, ensures some desired properties such as $[\![f_1 \succeq f_2]\!] + [\![f_2 \succeq f_1]\!] = 1$ or $[\![f \succeq f]\!] = 0.5$. Moreover, normalisation allows to limit the impact of incomparable random variables when stochastic ordering is a partial preorder. If there are more than two classifiers, we can order them according to maximal degree of being dominated by any other classifier defined for each f_i:

$$p_{\widetilde{\Delta},\succeq}(f_i) = \max_{\substack{\text{all classifiers } f \\ f \neq f_i}} [\![f \succeq f_i]\!]_{\widetilde{\Delta}} \tag{32}$$

When applied stochastic ordering $\succeq$ is a total preorder, then this criterion coincides with the selection of the model f that minimises $\Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_i})$ – the possibility of being dominated by some (arbitrary) model f_i. Thus (31–32) can be seen as a generalisation of that criterion to partial preorders for which not necessarily $\max(\Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2}), \Pi(\widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_2} \succeq \widetilde{U}^{(\mathbb{X},\mathbb{Y})}_{\widetilde{\Delta},f_1})) = 1$.

Example 5. *Let's try to examine the situation from Example 3 using the proposed approach. In the example we will use fuzzy cost function $\widetilde{\Delta}$ with costs* *defined on Fig. 3. The kernels of fuzzy cost values are the same as costs from Example 1.*

Domination degrees are presented in Table 4. Using criterion from (32) we can obtain the following order: f_1 *(*0.3*),* f_2 *(*0.7*) and* f_3 *(*0.75*). Hence,* f_1 *is definitely the best model for given problem.*

Let us now look in more detail at the situation of f_1 *and* f_2 *models. They were incomparable according to classical First Stochastic Dominance order. Thanks to proposed approach, we still are able to find out which one is better. According to (30) we need to find random variables* X *and* Y *that maximise given formula and for which* $X \succeq_{1st} Y$ *holds. Optimal random variables are given in Table 5. It is easy to observe that for patients for which* f_1 *outcome was worse then that of* f_2 *costs are swapped to keep the* $X \succeq_{1st} Y$ *property.*

4.3 Evaluation

We evaluated this approach on the same medical data set as the original classifier comparison strategies. The procedure was following:

1. Extend fuzzy cost function from Example 5 to cover "NA" cases
2. For each pair of classifiers (f_i, f_j):

   (a) Test whether f_i dominates f_j, if so, return 1 (full dominance)
   (b) Try to solve the problem numerically using Nelder and Mead and BFGS methods [18]
   (c) Return the highest value found

3. Normalise the dominance degrees according to (31)
4. For each classifier f_i calculate value of $p_{\widetilde{\Delta},\succeq}(f_i)$

More details on evaluation procedure including data pre– and post–processing can be found on GitHub repository.[1]

Table 4. Degree of domination for classification models from Example 5.

	f_1	f_2	f_3
f_1	0.5	0.7	0.75
f_2	0.3	0.5	0.34
f_3	0.25	0.66	0.5

Graph on Fig. 4 presents selected best classifiers with the lowest $p_{\widetilde{\Delta},\succeq}(f_i)$ value. Significant domination degrees (≥ 0.5) are depicted as arrows pointing from dominating to dominated element. There are 3 classifiers that are not

[1] https://github.com/bikol/stochastic-orders-evaluation.

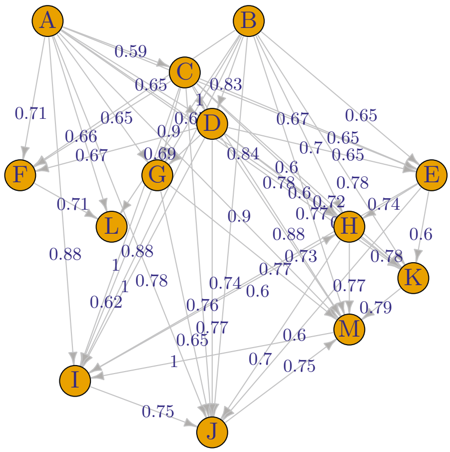


Fig. 4. Graph showing domination degrees ($[\![f_1 \succeq_{1st} f_2]\!]$) obtained with proposed approach for selected classifiers.

Table 5. Costs that maximise degree of domination $[\![f_1 \succeq f_2]\!]_{\widetilde{\Delta}}$.

Cost	Patients									
	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
actual $\mathbb{Y}(\omega)$	B	B	B	B	B	M	M	M	M	M
model f_1	0 (B)	0 (B)	0 (B)	0 (B)	2.5 (M)	0 (M)	0 (M)	0 (M)	0 (M)	3.33 (B)
model f_2	0 (B)	3.33 (M)	2.5 (M)	2.5 (M)	2.5 (M)	0 (M)	0 (M)	0 (M)	0 (M)	0 (M)

Table 6. Summary of various evaluation methods. Shortcuts in header stand for: DEC – Decisiveness, ACC – Accuracy, SEN – Sensitivity, SPC – Specificity.

Model	DEC	ACC	SEN	SPC	Cost	$\succeq_{1st}$	$\succeq_{sp}$	$p_{\widetilde{\Delta},\succeq_{1st}}$
A	0.949	0.886	0.902	0.878	70	0	0.961	0.567
B	0.966	0.876	0.902	0.864	72	0	1.051	0.571
C	0.971	0.876	0.900	0.867	72	0	5.401	0.592
D	0.971	0.871	0.900	0.858	74.5	1	2.293	1.000
E	0.971	0.865	0.918	0.843	75.5	0	2.176	0.675
F	0.931	0.877	0.917	0.861	76	0	0.773	0.713
G	1.000	0.857	0.885	0.846	77.5	0	3.159	0.904
H	0.943	0.885	0.857	0.897	78	1	2.796	0.835
I	0.971	0.859	0.900	0.842	79.5	1	0.727	1.000
J	0.920	0.901	0.826	0.930	80	0	7.462	0.815
K	0.920	0.894	0.848	0.913	80	0	3.186	0.814
L	0.874	0.895	0.909	0.890	80	0	0.583	0.706
M	1.000	0.851	0.731	0.902	100	1	2.739	0.900

significantly dominated. Those models should be considered as potential candidates for choosing. In Table 6 one can see that those models have the lowest and very similar values of $p_{\widetilde{\Delta},\succeq}(f_i)$ which should be used as a final criteria for model selection.

One can see that all comparison methods gave similar results. However, there are few interesting cases which will be discussed here. First one is model J, the best model according to statistical preference. However, this is not confirmed by other methods. The reason of such behaviour is that statistical preference based method, in contrast to other ones, does not take into account the difference in weight between false negatives and false positives. Therefore its results are more similar to those obtained with accuracy.

Second interesting case concerns classifiers D and E. D is being dominated according to First Stochastic Dominance while E is not. This contrasts with the fact that D performs better on other performance measures listed in Table 6 (except sensitivity). Such situation occurred because model C classifies all patients exactly the same as D, except one for which it gives better response (C dominates D, see Fig. 2a). One may say that C is a strictly better version of D. Therefore D should not be chosen as the best classifier. However, this is not true for model E so it still may be considered as a the candidate.

5 Discussion and Further Work

This paper presents an approach to applying stochastic orderings to evaluate classification algorithms for low quality data. We discussed some known stochastic orderings along with practical notes about their application to medical diagnosis support problem. The difficulties that have arisen were our motivation to propose new approach based on fuzzy cost function. The new method allows to compare any two classifiers, but does not require precise definition of the cost function.

All proposed methods were evaluated on real life medical data that comes from recent study on application of aggregation operators to supporting ovarian tumor diagnosis [2]. We were able to obtain results very similar to those previously reported but adopting much weaker assumptions about costs values. This is especially important in this specific problem because as there are still no reliable information on how to estimate costs in medical diagnostics.

Our proposed approach allows to associate numerical metric to each classifier (similarly as in total cost method). This is very useful as it enables the use of this method in more complex evaluation and learning procedures such as cross validation.

As future research we want to evaluate the stability of domination degrees while we slightly change fuzzy cost values. Such stability is very problematic in classical total cost method, where even small changes in costs may lead to big changes in obtained classifier order. As a second line of further research we want to investigate other approaches to fuzzify First Stochastic Dominance based classifier evaluation method such as application of linguistic quantification.

References

1. Żywica, P., Wójtowicz, A., et al.: Improving medical decisions under incomplete data using interval–valued fuzzy aggregation. In: Proceedings of 9th European Society for Fuzzy Logic and Technology (EUSFLAT), Gijón, Spain, pp. 577–584 (2015)
2. Wójtowicz, A., Żywica, P., et al.: Solving the problem of incomplete data in medical diagnosis via interval modeling. Appl. Soft Comput. **47**, 424–437 (2016)
3. Diering, M., Dyczkowski, K., Hamrol, A.: New method for assessment of raters agreement based on fuzzy similarity. In: Proceedings of the 10th International Conference on Soft Computing Models in Industrial and Environmental Applications, pp. 415–425. Springer (2015)
4. Stachowiak, A., Dyczkowski, K.: A similarity measure with uncertainty for incompletely known fuzzy sets. In: 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), pp. 390–394. IEEE (2013)
5. Stukan, M., Dudziak, M., et al.: Usefulness of diagnostic indices comprising clinical, sonographic, and biomarker data for discriminating benign from malignant ovarian masses. J. Ultrasound Med. **34**(2), 207–217 (2015)
6. Moszyński, R., Żywica, P., et al.: Menopausal status strongly influences the utility of predictive models in differential diagnosis of ovarian tumors: an external validation of selected diagnostic tools. Ginekol. Pol. **85**(12), 892–899 (2014)
7. Japkowicz, N., Shah, M.: Evaluating Learning Algorithms: A Classification Perspective. Cambridge University Press, New York (2011)
8. Hatch, S.: Snowball in a Blizzard: A Physician's Notes on Uncertainty in Medicine. Basic Books, New York (2016)
9. Couso, I., Sánchez, L.: Generalized stochastic orderings applied to the study of performance of machine learning algorithms for low quality data. In: Proceedings of 16th International Fuzzy Systems Association World Congress and 9th Conference of the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT 2015), pp. 1534–1541. Atlantis Press (2015)
10. Couso, I., Sánchez, L.: Machine learning models, epistemic set-valued data and generalized loss functions: an encompassing approach. Inf. Sci. **358**, 129–150 (2016)
11. Shaked, M., Shanthikumar, G.: Stochastic Orders. Springer Science & Business Media, New York (2007)
12. Savage, L.J.: The Foundations of Statistics. Courier Corporation, New York (1972)
13. Hadar, J., Russell, W.R.: Rules for ordering uncertain prospects. Am. Econ. Rev. **59**(1), 25–34 (1969)
14. David, H.A.: The Method of Paired Comparisons, vol. 12. Charles Griffin & D. Ltd., London (1963)
15. Page, L., Brin, S., et al.: The PageRank citation ranking: bringing order to the web. Technical report, Stanford InfoLab (1999)
16. Couso, I., Dubois, D.: A perspective on the extension of stochastic orderings to fuzzy random variables. In: Proceedings of 16th International Fuzzy Systems Association World Congress and 9th Conference of the European Society for Fuzzy Logic and Technology, IFSA-EUSFLAT 2015, pp. 1486–1492. Atlantis Press (2015)
17. Dubois, D., Prade, H.: Gradualness, uncertainty and bipolarity: making sense of fuzzy sets. Fuzzy Sets Syst. **192**, 3–24 (2012)
18. Wright, S., Nocedal, J.: Numerical Optimization, vol. 35. Springer Science, New York (1999). pp. 67–68

Author Index

J. Kacprzyk et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*,
Advances in Intelligent Systems and Computing 643,
DOI 10.1007/978-3-319-66827-7