

New Dimensions in Fuzzy Logic and Related Technologies

Proceedings
of the 5th EUSFLAT Conference,
Ostrava, Czech Republic,
September 11–14, 2007

Volume II





Edited by
Martin Štěpnička
Vilém Novák
Ulrich Bodenhofer





Title: **New Dimensions in Fuzzy Logic and Related Technologies. Volume II**
Proceedings of the 5th EUSFLAT Conference, Ostrava, Czech Republic, September 11-14, 2007

Editors:	Mgr. Martin Štěpnička, Prof. Ing. Vilém Novák, DrSc. Univ. Doz. Dipl.-Ing. Dr. Ulrich Bodenhofer
Publisher:	University of Ostrava
Number of Copies:	170
Pages:	488
Edition:	First, 2007
Printing:	Tribun

Typesetting: by authors using LaTeX or MS Word
completion by Viktor Pavliska

Cover design: Robert Rytina, VR Ateliér

© University of Ostrava
© Mgr. Martin Štěpnička, Prof. Ing. Vilém Novák, DrSc., Univ. Doz. Dipl.-Ing. Dr. Ulrich Bodenhofer

ISBN 978-80-7368-387-0

Contents

Regular Sessions **9**

1 Fuzzy Logic 1 **11**

A Fuzzy Formal Logic for Interval-valued Residuated Lattices

Van Gasse, B., Cornelis, C., Deschrijver, G., Kerre, E. E. 13

On Completeness Results for the Expansions with Truth-constants of Some Predicate Fuzzy Logics

Esteva, F., Godo, L., Noguera, C. 21

Resolution Strategies for Fuzzy Description Logic

Habiballa, H. 27

Left-continuous t-norms as Functional Algebras

Vetterlein, T. 37

A Logic Determined by Commutative Residuated Lattices

Kondo, M., Watari, O., Kawaguchi, M. F., Miyakoshi, M. 45

On the First Place Antitonicity in QL-implications

Shi, Y., Ruan, D., Kerre, E. E. 49

2 Applications 1 **57**

On Cluster Analysis Based on Fuzzy Relations Between Spatial Data

Nasibov, E., Ulutagay, G. 59

Using Supervised Fuzzy Clustering and CWT for Ventricular Late Potentials Detection in HRECG Signal

Salimi Khorshidi, G., Jaafari, A., Pourmorteza, A., Motie Nasrabadi, A., Hashemi Golpayegani, M. 63

Distributed Wireless Sensor Network Architecture: Fuzzy Logic Based Sensor Fusion

Murthy, G. R., Iyer, V. 71

Soft Computing Robot Navigation Case Study

Lakov, D. V., Saralieva, M. R. 79

3 Fuzzy Rules **85**

How to Use Controller with Conditionally Firing Rules

Navara, M., Petrík, M. 87

On Additive and Multiplicative Fuzzy Models

Štěpnička, M., De Baets, B., Nosková, L. 95

Automation of Human Reasoning in Economical Analysis

Dvořák, A., Novák, V. 103

Advanced Inference in Fuzzy Systems by Rule Base Compression

Gegov, A., Gobalakrishnan, N. 111

Using Generalized Constraints and Protoforms to Deal with Adverbs

Soto, A., Olivas, J. A., Prieto, M. E. 119

4 Applications 2 **127**

A Less Cumulative Algorithm of Mining Linguistic Browsing Patterns in the World Wide Web

Dyczkowski, K. 129

Skew-Tree Based Multistage Fuzzy Controller for Nonlinear Systems

Khosravi, A., Lu, J., Zheng, X., Barzamini, R. 137

Modifying the Classic Peak Picking Technique Using a Fuzzy Multi Agent to Have an Accurate P300-based BCI

Salimi Khorshidi, G., Jaafari, A., Motie Nasrabadi, A., Hashemi Golpayegani, M. 143

Toward a Solution for "Genetic Algorithm"-"Support Vector Machine" Combination to Have a Reliable P300-based BCI

Salimi Khorshidi, G., Jaafari, A., Motie Nasrabadi, A., Hashemi Golpayegani, M., Lucas, C. 149

5 Measures of Vagueness **155**

A Characterization of the Ambiguity and Fuzziness by Means of Activity Orders on the Closed Interval in $[0, 1]$

Alcalde, C., Burusco, A., Fuentes-González, R. 157

Extended Distances Between Fuzzy Points

Casasnovas, J., Riera, J. V. 163

Conditional Entropy for the Union of Fuzzy and Crisp Partitions

Vivona, D., Divari, M. 169

A Construction of a Fuzzy Valued Measure Based on Minimum t-norm

Ruzha, V., Asmuss, S. 175

Compenzational Vagueness

Mareš, M. 179

On Monotonic Operations which Are Locally Internal on Some Subset of Their Domain

Drygaś, P. 185

6 Fuzzy Logic 2 **193**

The Logic of Perfect MV-algebras

Belluce, L. P., Di Nola, A., Gerla, B. 195

An Algebraic Approach to States on MV-algebras

Flaminio, T., Montagna, F. 201

Free MV-algebras

Di Nola, A., Grigolia, R. 207

Concept Lattices under Non-commutative Conjunctors Are Generalized Concept Lattices

Medina, J., Ojeda-Aciego, M., Ruiz-Calviño, J. 209

Mixed Integer Programming, General Concept Inclusions and Fuzzy Description Logics

Bobillo, F., Straccia, U. 213

From Fuzzy- to Bipolar- Datalog

Achs, Á. 221

7 Neuro-fuzzy, Predictions and Evolutionary Approach 229

Training Neurofuzzy Networks with Participatory Learning

Hell, M., Ballini, R., Costa, P., Gomide, F. 231

A New Preprocessing Approach to Preparation of Binary Patterns for FAM Neural Networks

Chitsaz, M., Sadati, N., Barzamini, R., Jouzdani, J., Khosravi, A. 237

Time Series Prediction by Perturbed Fuzzy Model

Salgado, P., Igrejas, G., Gouveia, F. 243

Competitive Self-adaptation in Evolutionary Algorithms

Tvrdík, J., Křivý, I. 251

Differential Evolution Methods for LU-fuzzy Arithmetic

Stefanini, L. 259

Selection of Optimal Set of Diagnostic Tests with Use of Evolutionary Approach in Intelligent Systems

Yankovskaya, A. E., Tsoy, Y. R. 267

8 Fuzzy Numbers 271

Object Oriented Approach in Fuzzy Arithmetic Using Parametric Representation of Fuzzy Numbers

Huňka, F. 273

Inverse Arithmetic Operators for Fuzzy Intervals

Boukezzoula, R., Galichet, S., Foulloy, L. 279

Representing Incomplete Knowledge about Fuzzy Sets

Wygralak, M. 287

9 Decision Making 1 293

Two Consensus Protocols Based on an Acceptance Threshold in Group Decision Making

Labreuche, Ch. 295

Modeling the Context in Multi-Criteria Decision Analysis

Labreuche, Ch. 303

A Fuzzy Representation for Weights of Alternatives in AHP

Ohnishi, S., Yamanoi, T., Imai, H. 311

Dealing with Incompleteness of Preferences in Group Decision Making Problems

Pankowska, A. 317

10 Fuzzy Relations **323**

Algorithms for the Construction of Digital Convex Fuzzy Hulls

Ralević, N. M., Čomić, L. 325

Lukasiewicz Filters as Convex Combinations of Crisp Filters

Havranová, Z., Kalina, M. 331

Fuzzy Preference Relations and Lukasiewicz Filters

Havranová, Z., Kalina, M. 337

On a Proximity-based Tolerant Inclusion

Bosc, P., Hadjali, A., Pivert, O. 343

Finding Close T-indistinguishability Operators to a Given Proximity

Garmendia, L., Recasens, J. 351

Preservation of Graded Properties of Fuzzy Relations by Aggregation Functions

Dudziak, U. 359

11 Uncertainty Models **365**

Transforming Probability Intervals into Other Uncertainty Models

Destercke, S., Dubois, D., Chojnacki, E. 367

Lattice-Valued Possibilistic Entropy Functions

Kramosil, I. 375

Revisiting Bobylev's Definitions

Rodríguez-Muñiz, L. J. 381

12 Categories and Topology **385**

On L-valued Equalites on the L-powersets of Many-valued Sets

Uljane, I., Šostak, A. 387

On Fuzzification of Algebraic and Topological Structures

Solovjovs, S. 395

Fuzzy Order Relation and Fuzzy Ordered Set Category

Lebedeva, O. 403

Monadic L-fuzzy Quantifiers of the Type $\langle 1^n, 1\rangle$

Holčapek, M. 409

Homomorphisms of Fuzzy Logic Models Based on Sets With Similarities

Močkoř, J. 417

13 Decision Making 2 **423**

Fault Diagnosis with Progressive Symptoms Based on Multi-Agent Approach

Sokolov, O., Wagenknecht, M., Gocht, U. 425

Similarity of Fuzzy Preference Structures Based on Metrics

Hliněná, D., Kráľ, P. 431

Supporting Consumer Decision by Fuzzy Revealed Preference in Online Sales

Georgescu, I., Qiu, X., Lin, R. 437

Revealed Preference in a Fuzzy Framework

Georgescu, I. 445

14 Aggregation Operators in Decision Making **453**

Applications of the Fuzzy Weighted Average of Fuzzy Numbers in Decision Making Models

Pavlačka, O., Talašová, J. 455

The Induced Generalized OWA Operator

Merigó, J. M., Gil-Lafuente, A. M. 463

Mixture Operators and Additional Insurance

Špirková, J., Kráľ, P. 471

Using Quasiarithmetic Means in a Sequential Decision Procedure

Ballester, M. A., García-Lapresta, J. L. 479

Regular Sessions

Session 1

Fuzzy Logic 1

A Fuzzy Formal Logic for Interval-valued Residuated Lattices

B. Van Gasse
Ghent University
Krijgslaan 281 (S9)
9000 Gent, Belgium
Bart.VanGasse@UGent.be

C. Cornelis
Ghent University
Krijgslaan 281 (S9)
9000 Gent, Belgium
Chris.Cornelis@UGent.be

G. Deschrijver
Ghent University
Krijgslaan 281 (S9)
9000 Gent, Belgium
Glad.Deschrijver@UGent.be

E.E. Kerre
Ghent University
Krijgslaan 281 (S9)
9000 Gent, Belgium
Etienne.Kerre@UGent.be

Abstract

Fuzzy formal logics were introduced in order to handle graded truth values instead of only 'true' and 'false'. A wide range of such logics were introduced successfully, like Monoidal T-norm based Logic, Basic Logic, Gödel Logic, Łukasiewicz Logic etc. However, in general, fuzzy set theory is not only concerned with vagueness, but also with uncertainty. A possible solution is to use intervals instead of real numbers as membership values. In this paper, we present an approach with triangle algebras, which are algebraic characterizations of interval-valued residuated lattices. The variety of these structures corresponds in a sound and complete way to a logic that we introduce, called Triangle Logic (in the same way as, e.g., BL-algebras and Basic Logic). We will show that this truth-functional approach, along with the residuation principle, has some consequences that seem to obstruct an easy and proper interpretation for the semantics of Triangle Logic.

Keywords: Fuzzy logic, Interval-valued fuzzy set theory, Residuated lattices.

1 Introduction and Preliminaries

1.1 Fuzzy formal logics

In classical logic[1], a formula is provable if, and only if, its value (under every evaluation) is 1 in every Boolean algebra (this is called soundness and completeness). Moreover, it suffices to consider the only linear Boolean algebra, which has two elements (0 and 1). We can therefore say that classical logic is a two-valued logic. The value 0 stands for 'false', 1 for 'true'. However, in many cases it may be preferable to work with graded truth values, e.g. for propositions like 'a cat is a small animal', 'Peter has dark hair' or 'the weather is bad today'. This is the reason why fuzzy (formal) logics, which are generalizations of classical logic, were introduced. One of the most general examples is Höhle's Monoidal Logic (ML) [9]. In this logic, a formula is provable iff its value is 1 in every residuated lattice. Recall that a residuated lattice is a structure $\mathcal{L} = (L, \sqcap, \sqcup, *, \Rightarrow, 0, 1)$ in which $\sqcap, \sqcup, *$ and $\Rightarrow$ are binary operators on L and

- $(L, \sqcap, \sqcup)$ is a bounded lattice with 0 as smallest and 1 as greatest element,
- $*$ is commutative and associative, with 1 as neutral element, and
- $x * y \leq z$ iff $x \leq y \Rightarrow z$ for all x, y and z in L (residuation principle).

This allows one to assign graded truth values to formulas. Of course, not every formula that is provable in classical logic, is provable in ML too. However, we can add axioms to obtain intermediary logics. An interesting example is Esteva and Godo's Monoidal T-norm based Logic (MTL) [7]. For this fuzzy logic, the so-called prelinearity axiom $((\phi \to \psi) \vee (\psi \to \phi))$ is added to the axioms of ML. As for all axiomatic extensions of ML, MTL is sound and complete w.r.t. the associated variety of algebraic structures (in this case: MTL-algebras, which are residuated lattices in which $(x \Rightarrow y) \sqcup (y \Rightarrow x) = 1$ holds for all x and y). Moreover, because of the prelinearity MTL is also sound and complete w.r.t. linear MTL-algebras[2] [7]. The latter property is preserved for axiomatic extensions of MTL. For example, Basic Logic (BL) [8] is sound and complete w.r.t. linear BL-algebras, Łukasiewicz logic [11] is sound and complete w.r.t. Wajsberg algebras,

[1] In this paper we only deal with propositional logics. Therefore we will not write this explicitly.

[2] Recall that linear MTL-algebra is the same as linear residuated lattices, because linearity is a stronger property than prelinearity in residuated lattices.

and classical logic is sound and complete w.r.t. the linear Boole algebra. In some of these cases, it is even possible to further improve this result in the sense that one only has to take into account the algebraic structures on the unit interval. Two examples are MTL and BL, which are sound and complete w.r.t. MTL-algebras on $[0,1]$ and BL-algebras on $[0,1]$, respectively [1, 10]. These are exactly the structures induced by left-continuous t-norms and continuous t-norms, respectively. So we can speak of BL as the logic of continuous t-norms on $[0,1]$.

1.2 Intervals as Truth Values

In order to handle also uncertainty together with vagueness, it seems a good idea to use closed subintervals of $[0,1]$ instead of real numbers in $[0,1]$. Interval-valued truth degrees have been widely adopted in knowledge-based systems [6]. This is due to the relative efficiency of operations defined on them, as well as to the fact that they carry an attractive and straightforward semantical interpretation as partial, or incomplete, truth values, i.e., they exhibit a lack of knowledge about a formula's exact truth value; the wider the interval, the greater the uncertainty.

The lattice $\mathcal{L}^I = (L^I, \sqcap, \sqcup)$ that contains these closed subintervals is shown graphically in Figure 1 and defined by

$$L^I = \{[x_1, x_2] \mid (x_1, x_2) \in [0,1]^2 \text{ and } x_1 \leq x_2\},$$
$$[x_1, x_2] \sqcap [y_1, y_2] = [\min(x_1, y_1), \min(x_2, y_2)],$$
$$[x_1, x_2] \sqcup [y_1, y_2] = [\max(x_1, y_1), \max(x_2, y_2)].$$

Its partial ordering $\leq_{L^I}$ is given by componentwise extension of $\leq$,

$$[x_1, x_2] \leq_{L^I} [y_1, y_2] \iff x_1 \leq y_1 \text{ and } x_2 \leq y_2.$$

We use the 'componentwise' ordering, which is in our opinion the most natural one. Indeed, it satisfies $[x_1, x_2] \leq_{L^I} [y_1, y_2]$ iff $[x_1, x_2] \sqcap [y_1, y_2] = [x_1, x_2]$ iff $[x_1, x_2] \sqcup [y_1, y_2] = [y_1, y_2]$, and $[x_1, x_2] \sqcap [y_1, y_2] = \{\min(x, y) \mid (x, y) \in [x_1, x_2] \times [y_1, y_2]\}$ and $[x_1, x_2] \sqcup [y_1, y_2] = \{\max(x, y) \mid (x, y) \in [x_1, x_2] \times [y_1, y_2]\}$.

One way to interpret these intervals is to view them as approximations of truth values in $[0,1]$. For example, a truth value $[a, b]$ might mean that the exact, but unknown, truth value is definitely greater than or equal to a and smaller than or equal to b.

Research on fuzzy formal logics has centered on prelinear residuated structures. However, while prelinearity holds in every residuated lattice $([0,1], \min, \max, *, \Rightarrow, 0, 1)$, it is not necessarily preserved for closed intervals of a bounded lattice $\mathcal{L}$; for example, it was shown in [2] that no MTL-algebra exists on the lattice $\mathcal{L}^I$.

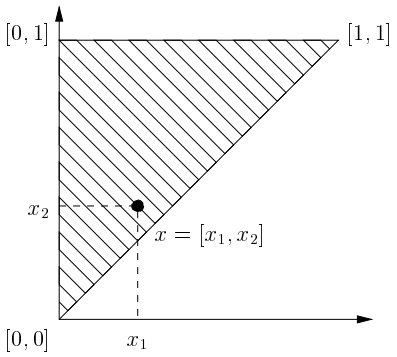


Figure 1: The lattice $\mathcal{L}^I$

The goal of this paper is to develop a logic that formally characterizes tautologies (true formulas) in interval-valued residuated lattices (IVRLs). Monoidal Logic, which corresponds to the complete class of residuated lattices, is too general for our purposes, and we need to extend it with suitable axioms to replace prelinearity. To achieve this, we propose the use of modal-like operators.

In Section 2 we first recall triangle algebras and their relationship with residuated lattices on interval-valued algebraic structures [12]. In Section 3 we will introduce Triangle Logic and show soundness and completeness w.r.t. triangle algebras. We round up the paper with some remarks about the suitability of TL for modelling reasoning with uncertain propositions, and a conclusion.

2 Triangle algebras

In view of the considerations in Section 1, we defined [12] algebraic structures that should satisfy as many properties as possible of residuated lattices on L^I, like MTL-algebras are the algebraic structures that satisfy as many properties as possible of residuated lattices on $[0,1]$. These structures are called triangle algebras. Before we give their definition and discuss triangle algebras, we give some additional definitions:

Definition 1 Given a lattice $\mathcal{L} = (L, \sqcap, \sqcup)$, its triangularization $\mathbb{T}(\mathcal{L})$ is the structure $\mathbb{T}(\mathcal{L}) = (Int(\mathcal{L}), \bigsqcap, \bigsqcup)$ defined by

- $Int(\mathcal{L}) = \{[x_1, x_2] \mid (x_1, x_2) \in L^2 \text{ and } x_1 \leq x_2\}$
- $[x_1, x_2] \bigsqcap [y_1, y_2] = [x_1 \sqcap y_1, x_2 \sqcap y_2]$
- $[x_1, x_2] \bigsqcup [y_1, y_2] = [x_1 \sqcup y_1, x_2 \sqcup y_2]$

The set $D_{\mathcal{L}} = \{[x, x] \mid x \in L\}$ is called the diagonal of $\mathbb{T}(\mathcal{L})$.

The first and the second projection pr_1 and pr_2 are the mappings from $Int(\mathcal{L})$ to L, defined by $pr_1([x_1, x_2]) = x_1$ and $pr_2([x_1, x_2]) = x_2$, for all $[x_1, x_2]$ in $Int(\mathcal{L})$. We also define the mappings p_v and p_h from $Int(\mathcal{L})$ to $D_{\mathcal{L}}$ as $p_v([x_1, x_2]) = [x_1, x_1]$ and $p_h([x_1, x_2]) = [x_2, x_2]$, for all $[x_1, x_2]$ in $Int(\mathcal{L})$.
A t-norm $\mathcal{T}$ on a triangularization of a bounded lattice is called t-representable if $pr_1(\mathcal{T}([x_1, x_2], [y_1, y_2]))$ is independent of x_2 and y_2, and $pr_2(\mathcal{T}([x_1, x_2], [y_1, y_2]))$ is independent of x_1 and y_1.
An interval-valued residuated lattice (IVRL) is a residuated lattice $(Int(\mathcal{L}), \bigsqcap, \bigsqcup, \odot, \Rightarrow_{\odot}, [0, 0], [1, 1])$ on the triangularization $\mathbb{T}(\mathcal{L})$ of a bounded lattice $\mathcal{L}$, in which the diagonal $D_{\mathcal{L}}$ is closed under $\odot$ and $\Rightarrow_{\odot}$, i.e., $[x_1, x_1] \odot [y_1, y_1] \in D_{\mathcal{L}}$ and $[x_1, x_1] \Rightarrow_{\odot} [y_1, y_1] \in D_{\mathcal{L}}$ for x_1, y_1 in L.

Note for example that $\mathcal{L}^I$ is the triangularization of $([0, 1], \min, \max)$.

Example 1 *It was shown in [3] that if T is a left-continuous t-norm on $[0, 1]$, then for each α in $[0, 1]$, the mapping $\mathcal{T}_{T,\alpha}$ defined by, for $x = [x_1, x_2]$ and $y = [y_1, y_2]$ in L^I, $\mathcal{T}_{T,\alpha}(x, y) =$*

$$[T(x_1, y_1), \max(T(\alpha, T(x_2, y_2)), T(x_1, y_2), T(x_2, y_1))],$$

is a t-norm on L^I, with residual implicator[3] $\mathcal{I}_{\mathcal{T}_{T,\alpha}}(x, y)$

$$= \quad [\min(I_T(x_1, y_1), I_T(x_2, y_2)), \\ \min(I_T(T(x_2, \alpha), y_2), I_T(x_1, y_2))],$$

that induces a residuated lattice on $\mathcal{L}^I$. Because also $\mathcal{T}_{T,\alpha}([x, x], [y, y]) = [T(x, y), T(x, y)]$ and $\mathcal{I}_{\mathcal{T}_{T,\alpha}}([x, x], [y, y]) = [I_T(x, y)), I_T(x, y)]$ for all x and y in $[0, 1]$, $(L^I, \sqcap, \sqcup, \mathcal{T}_{T,\alpha}, \mathcal{I}_{\mathcal{T}_{T,\alpha}}, [0, 0], [1, 1])$ is an IVRL.

Two important values of α can be distinguished:

- *If $\alpha = 1$, we obtain t-representable t-norms on $\mathcal{L}^I$: $\mathcal{T}_{T,1}(x, y) = [T(x_1, y_1), T(x_2, y_2)]$, which can be seen as the straightforward (and most commonly used) extension of T to $\mathcal{L}^I$. These t-norms on $\mathcal{L}^I$ are characterized by the property $p_h(\mathcal{T}(x, y)) = \mathcal{T}(p_h(x), p_h(y))$.*
- *If $\alpha = 0$, we obtain pseudo t-representable t-norms on $\mathcal{L}^I$:*
  $\mathcal{T}_{T,0}(x, y) = [T(x_1, y_1), \max(T(x_1, y_2), T(x_2, y_1))]$.
  These t-norms are inherently more complex than their t-representable counterparts, but, as we shall see at the end of this section, satisfy more relevant properties.

[3] If T is a t-norm on a bounded lattice $(L, \sqcap, \sqcup)$, its residual implicator I_T is defined as $I_T(x, y) = \sup\{z \in L \mid T(x, z) \leq y\}$ for all x and y in L, if the supremum exists. It always exists in complete lattices and residuated lattices (provided that $*$ is the t-norm).

Definition 2 *[12]* A triangle algebra is a structure $\mathcal{A} = (A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)$, in which $(A, \sqcap, \sqcup, *, \Rightarrow, 0, 1)$ is a residuated lattice, in which ν and μ are binary operators and u a constant, and in which the following conditions hold:

$T.1$	$\nu x \leq x$,	$T.1'$	$x \leq \mu x$,
$T.2$	$\nu x \leq \nu\nu x$,	$T.2'$	$\mu\mu x \leq \mu x$,
$T.3$	$\nu(x \sqcap y) = \nu x \sqcap \nu y$,	$T.3'$	$\mu(x \sqcap y) = \mu x \sqcap \mu y$,
$T.4$	$\nu(x \sqcup y) = \nu x \sqcup \nu y$,	$T.4'$	$\mu(x \sqcup y) = \mu x \sqcup \mu y$,
$T.5$	$\nu 1 = 1$,	$T.5'$	$\mu 0 = 0$,
$T.6$	$\nu u = 0$,	$T.6'$	$\mu u = 1$,
$T.7$	$\nu\mu x = \mu x$,	$T.7'$	$\mu\nu x = \nu x$,

$T.8$	$\nu(x \Rightarrow y) \leq \nu x \Rightarrow \nu y$,
$T.9$	$(\nu x \Leftrightarrow \nu y) * (\mu x \Leftrightarrow \mu y) \leq (x \Leftrightarrow y)$,
$T.10$	$\nu x \Rightarrow \nu y \leq \nu(\nu x \Rightarrow \nu y)$.

The set $\{x \in A \mid \nu x = x\}$ is called the set of exact elements $E(\mathcal{A})$ of the triangle algebra $\mathcal{A}$.

We use the modal-like operators ν and μ in order to represent the lower and upper bound of an interval. The set of exact elements corresponds to the diagonal of an IVRL.

Proposition 1 *[12] Let $\mathcal{A} = (A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)$ be a triangle algebra. Then $E(\mathcal{A})$ is the direct image of A under ν, as well as under μ. Moreover, this set is invariant under ν and μ, and contains 0 and 1, but not u (unless in the trivial case when $|A| = 1$). It is closed under $\sqcap$, $\sqcup$, $*$ and $\Rightarrow$.*

The next theorem [12] establishes the equivalence between IVRLs and triangle algebras:

Theorem 1 *Every triangle algebra $\mathcal{A} = (A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)$ is isomorphic to a triangle algebra $(Int(\mathcal{E}(\mathcal{A})), \bigsqcap, \bigsqcup, \odot, \Rightarrow_{\odot}, p_v, p_h, [0, 0], [0, 1], [1, 1])$ where $(Int(\mathcal{E}(\mathcal{A})), \bigsqcap, \bigsqcup, \odot, \Rightarrow_{\odot}, [0, 0], [1, 1])$ is an IVRL.*
Conversely, if $(A, \sqcap, \sqcup, \odot, \Rightarrow_{\odot}, [0, 0], [1, 1])$ is an IVRL and ν and μ are defined by $\nu[x_1, x_2] = [x_1, x_1]$ and $\mu[x_1, x_2] = [x_2, x_2]$, then $(A, \sqcap, \sqcup, \odot, \Rightarrow_{\odot}, \nu, \mu, [0, 0], [0, 1], [1, 1])$ is a triangle algebra.

This theorem shows that triangle algebras are a good characterization for IVRLs. However, IVRLs are still more general than residuated lattices on L^I. For example, the diagonal of an IVRL needs not be linear. We can impose the property that the diagonal should be prelinear, but it is currently unknown if this is enough to guarantee that all identities that are true in IVRLs with linear diagonal are also true in every IVRL with prelinear diagonal. This property would be comparable to the fact that all the identities that are true in linear residuated lattices are also true in prelinear residuated lattices [7].

3 Triangle Logic

In this section we translate the defining properties of triangle algebras into logical axioms, and show that the resulting logic TL is sound and complete w.r.t. the variety of triangle algebras.

The language of TL consists of countably many proposition variables (p_1, p_2,...), the constants $\overline{0}$ ('falsity') and $\bot$ ('uncertainty'), the unary operators $\Box$ ('necessity'), $\Diamond$ ('possibility'), the binary operators $\wedge$ ('weak conjunction'), $\vee$ ('disjunction'), & ('strong conjunction'), $\rightarrow$ ('implication'), and finally the auxiliary symbols '(' and ')'. Formulas are defined inductively: proposition variables, $\overline{0}$ and $\bot$ are formulas; if ϕ and ψ are formulas, then so are $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \& \psi)$, $(\phi \rightarrow \psi)$, $(\Box\psi)$ and $(\Diamond\psi)$.

In order to avoid unnecessary brackets, we agree on the following priority rules:

- unary operators always take precedence over binary ones,
- among the binary operators, & has the highest priority; furthermore $\wedge$ and $\vee$ take precedence over $\rightarrow$,
- the outermost brackets are not written.

We also introduce some useful shorthand notations: $\overline{1}$ for $\overline{0} \rightarrow \overline{0}$, $\neg\phi$ for $\phi \rightarrow \overline{0}$ and $\phi \leftrightarrow \psi$ for $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$ for formulas ϕ and ψ.

The axioms of TL are those of ML (Monoidal Logic)

$ML.1$	$(\phi \rightarrow \psi) \rightarrow ((\psi \rightarrow \chi) \rightarrow (\phi \rightarrow \chi))$,
$ML.2$	$\phi \rightarrow (\phi \vee \psi)$,
$ML.3$	$\psi \rightarrow (\phi \vee \psi)$,
$ML.4$	$(\phi \rightarrow \chi) \rightarrow ((\psi \rightarrow \chi) \rightarrow ((\phi \vee \psi) \rightarrow \chi))$,
$ML.5$	$(\phi \wedge \psi) \rightarrow \phi$,
$ML.6$	$(\phi \wedge \psi) \rightarrow \psi$,
$ML.7$	$(\phi \& \psi) \rightarrow \phi$,
$ML.8$	$(\phi \& \psi) \rightarrow (\psi \& \phi)$,
$ML.9$	$(\phi \rightarrow \psi) \rightarrow ((\phi \rightarrow \chi) \rightarrow (\phi \rightarrow (\psi \wedge \chi)))$,
$ML.10$	$(\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \& \psi) \rightarrow \chi)$,
$ML.11$	$((\phi \& \psi) \rightarrow \chi) \rightarrow (\phi \rightarrow (\psi \rightarrow \chi))$,
$ML.12$	$\overline{0} \rightarrow \phi$,

complemented with axioms corresponding to T.1–T.10 and T.1'–T.7':

$TL.1$	$\Box\phi \rightarrow \phi$,
$TL.1'$	$\phi \rightarrow \Diamond\phi$,
$TL.2$	$\Box\phi \rightarrow \Box\Box\phi$,
$TL.2'$	$\Diamond\Diamond\phi \rightarrow \Diamond\phi$,
$TL.3$	$(\Box\phi \wedge \Box\psi) \rightarrow \Box(\phi \wedge \psi)$,
$TL.3'$	$(\Diamond\phi \wedge \Diamond\psi) \rightarrow \Diamond(\phi \wedge \psi)$,
$TL.4$	$\Box(\phi \vee \psi) \rightarrow (\Box\phi \vee \Box\psi)$,
$TL.4'$	$\Diamond(\phi \vee \psi) \rightarrow (\Diamond\phi \vee \Diamond\psi)$,
$TL.5$	$\Box\overline{1}$,
$TL.5'$	$\neg\Diamond\overline{0}$,
$TL.6$	$\neg\Box\bot$,
$TL.6'$	$\Diamond\bot$,
$TL.7$	$\Diamond\phi \rightarrow \Box\Diamond\phi$,
$TL.7'$	$\Diamond\Box\phi \rightarrow \Box\phi$,
$TL.8$	$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$,
$TL.9$	$(\Box\phi \leftrightarrow \Box\psi) \& (\Diamond\phi \leftrightarrow \Diamond\psi) \rightarrow (\phi \leftrightarrow \psi)$,
$TL.10$	$(\Box x \rightarrow \Box y) \rightarrow \Box(\Box x \rightarrow \Box y)$.

The deduction rules are modus ponens (MP, from ϕ and $\phi \rightarrow \psi$ infer ψ), generalization (G, from ϕ infer $\Box\phi$) and monotonicity of $\Diamond$ (M$\Diamond$, from $\phi \rightarrow \psi$ infer $\Diamond\phi \rightarrow \Diamond\psi$).

The consequence relation $\vdash$ is defined as follows, in the usual way. Let V be a theory, i.e., a set of formulas in TL. A (formal) proof of a formula ϕ in V is a finite sequence of formulas with ϕ at its end, such that every formula in the sequence is either an axiom of TL, a formula of V, or the result of an application of an inference rule to previous formulas in the sequence. If a proof for ϕ exists in V, we say that ϕ can be deduced from V and we denote this by $V \vdash \phi$.

Note that TL.5 is in fact superfluous, as it immediately follows from $\emptyset \vdash \overline{1}$ and generalization; we include it here to obtain full correspondence with Definition 2.

To show that this logic is sound and complete w.r.t. the variety of triangle algebras, we use the same approach as for, e.g. ML [9], MTL [7] and BL [8]. We first establish some intermediate results. For a theory V, and formulas ϕ and ψ in TL, denote $\phi \sim_V \psi$ iff $V \vdash \phi \rightarrow \psi$ and $V \vdash \psi \rightarrow \phi$ (this is also equivalent with $V \vdash \phi \leftrightarrow \psi$). Note that $\sim_V$ is an equivalence relation on the set of formulas. Moreover, it is a congruence w.r.t. $\Box$, $\Diamond$, $\wedge$, $\vee$, & and $\rightarrow$; this means that the results of the application of these connectives are equivalent whenever the arguments are equivalent.As a consequence, we can meaningfully consider the structure $(A_V, \wedge_V, \vee_V, \&_V, \rightarrow_V, \Box_V, \Diamond_V, [\overline{0}]_V, [\bot]_V, [\overline{1}]_V)$, in which

- A_V is the set of equivalence classes of $\sim_V$, i.e.

$A/_{\sim_V}$,

- $\wedge_V$ is the binary operation on A_V that maps $([\phi]_V, [\psi]_V)$ to $[\phi \wedge \psi]_V$,
- $\vee_V$ is the binary operation on A_V that maps $([\phi]_V, [\psi]_V)$ to $[\phi \vee \psi]_V$,
- $\&_V$ is the binary operation on A_V that maps $([\phi]_V, [\psi]_V)$ to $[\phi \& \psi]_V$,
- $\rightarrow_V$ is the binary operation on A_V that maps $([\phi]_V, [\psi]_V)$ to $[\phi \rightarrow \psi]_V$,
- $\Box_V$ is the unary operation on A_V that maps $[\phi]_V$ to $[\Box\phi]_V$,
- $\Diamond_V$ is the unary operation on A_V that maps $[\phi]_V$ to $[\Diamond\phi]_V$,
- $[\overline{0}]_V$, $[\bot]_V$, $[\overline{1}]_V$ are the elements of A_V that contain $\overline{0}$, $\bot$ and $\overline{1}$ resp.

Proposition 2 The structure $(A_V, \wedge_V, \vee_V, \&_V, \rightarrow_V, \Box_V, \Diamond_V, [\overline{0}]_V, [\bot]_V, [\overline{1}]_V)$ is a triangle algebra.

If $V = \emptyset$, this structure is called the Lindenbaum-algebra of TL.

Definition 3 *Let $\mathcal{A} = (A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)$ be a triangle algebra and V a theory. An $\mathcal{A}$-evaluation is a mapping e from the set of formulas of TL to A that satisfies, for each two formulas ϕ and ψ:*

- $e(\phi \wedge \psi) = e(\phi) \sqcap e(\psi)$,
- $e(\phi \vee \psi) = e(\phi) \sqcup e(\psi)$,
- $e(\phi \& \psi) = e(\phi) * e(\psi)$,
- $e(\phi \rightarrow \psi) = e(\phi) \Rightarrow e(\psi)$,
- $e(\Box\phi) = \nu e(\phi)$,
- $e(\Diamond\phi) = \mu e(\phi)$,
- $e(\overline{0}) = 0$ *and*
- $e(\bot) = u$.

If an $\mathcal{A}$-evaluation e satisfies $e(\chi) = 1$ for every χ in V, it is called an $\mathcal{A}$-model for V.

Theorem 2 (**Soundness and completeness of TL**) A formula ϕ can be deduced from a theory V in TL iff for every triangle algebra $\mathcal{A}$ and for every $\mathcal{A}$-model e of V, $e(\phi) = 1$.

Theorem 2 implies similar results for more specific logics.

- For example, if we add $\mu(x * y) = \mu x * \mu y$ to the conditions of a triangle algebra and $\Diamond\phi \& \Diamond\psi \rightarrow \Diamond(\phi \& \psi)$ to the axioms of TL, then we can obtain a valid theorem by replacing 'triangle algebra' and 'Triangle Logic' in the formulation of Theorem 2 by the new algebra and logic. This property implies that (in terms of IVRL) the second component of $[x_1, x_2] * [y_1, y_2]$ is independent of x_1 and y_1. This means that we can use this property to characterize IVRLs with t-representable t-norms by triangle algebras satisfying $\mu(x * y) = \mu x * \mu y$.
- Another interesting example is $x = \neg\neg x$, connected to the axiom $\neg\neg\phi \rightarrow \phi$. The only involutive t-norms of the form $\mathcal{T}_{T,\alpha}$ are the pseudo t-representable ones [3]. More generally, in an involutive triangle algebra, it can be verified that $u * u = 0$.
- As a final example, we can add $(\nu x \Rightarrow \nu y) \sqcup (\nu y \Rightarrow \nu x) = 1$ to the conditions of a triangle algebra (remark that this property is always satisfied for triangle algebras on $\mathcal{L}^I$, because its diagonal is linearly ordered). If also we add $(\Box\phi \rightarrow \Box\psi) \vee (\Box\psi \rightarrow \Box\phi)$ to the axioms of TL, then again we obtain a valid theorem by replacing 'triangle algebra' and 'Triangle Logic' in Theorem 2 by the new algebra and logic. In this case $(E(\mathcal{A}), \sqcap, \sqcup, *, \Rightarrow, 0, 1)$ is an MTL-algebra (prelinear residuated lattice). This means that it is a subalgebra of the direct product of a system of linearly ordered residuated lattices [9]. Using this property, a stronger form of completeness, called chain completeness, can be proven for MTL: a formula ϕ can be deduced from a theory V in MTL iff for every linearly ordered MTL-algebra $\mathcal{A}$ and for every $\mathcal{A}$-model e of V, $e(\phi) = 1$. Similar results hold for subvarieties of the variety of MTL-algebras and their corresponding logics (e.g. BL and Ł). We would like to find analogous theorems for triangle algebras (and subvarieties) too, but at this moment it is still an open question if every triangle algebra satisfying $(\nu x \Rightarrow \nu y) \sqcup (\nu y \Rightarrow \nu x) = 1$ is a subalgebra of the direct product of a system of triangle algebras with linearly ordered diagonal.

Note 1 Triangle Logic is a truth-functional logic: the truth degree of a compound proposition is determined by the truth degree of its parts. This causes some counterintuitive results, if we want to interpret the element $[0, 1]$ of an IVRL as uncertainty. For example: suppose we don't know anything of the truth value of propositions p and q, i.e., $v(p) = v(q) = [0, 1]$. Then yet the implication $p \rightarrow q$ is definitely valid: $v(p \rightarrow q) = v(p) \Rightarrow v(q) = [1, 1]$. However, if

$\neg[0,1] = [0,1]$ [4] (which is intuitively preferable, since the negation of an uncertain proposition is still uncertain), then we can take $q = \neg p$, and obtain that $p \rightarrow \neg p$ is true. Or, equivalently (using the residuation principle), that $p\&p$ is false. This does not seem intuitive, as one would rather expect $p\&p$ to be uncertain if p is uncertain.
Another consequence of $[0,1] \Rightarrow [0,1] = [1,1]$ is that it is impossible to interpret the intervals as a set in which the 'real' (unknown) truth value is contained, and $X \Rightarrow Y$ as the smallest closed interval containing every $x \Rightarrow y$, with x in X and y in Y (as in [6]). Indeed: $1 \in [0,1]$ and $0 \in [0,1]$, but $1 \Rightarrow 0 = 0 \notin [1,1]$.
On the other hand, for t-norms it is possible that $X * Y$ is the smallest closed interval containing every $x * y$, with x in X and y in Y, but only if they are t-representable (described by the axiom $\mu(x * y) = \mu x * \mu y$). However, in this case $\neg[0,1] = [0,0]$, which does not seem intuitive ('the negation of an uncertain proposition is absolutely false').
These considerations seem to suggest that Triangle Logic is not suitable to reason with uncertainty. This does not mean that intervals are not a good way for representing degrees of imprecise knowledge, only that they are not suitable as truth values in a truth functional logical calculus when we interpret them as expressing imprecision. It might even be impossible to model uncertainty as a truth value in any truth-functional logic. This question is discussed in [4, 5]. However, nothing prevents the intervals in Triangle Logic from having more adequate interpretations.

4 Conclusion

In this paper we explained why we wanted to construct a (family of) fuzzy formal logic(s) with intervals as truth values. We showed how triangle algebras can be used to achieve this goal, as they characterize interval-valued residuated lattices. We introduced Triangle Logic and proved soundness and completeness w.r.t. triangle algebras, and thus w.r.t. interval-valued residuated lattices. Finally, we made some observations which seem to limit the usefulness of Triangle Logic, and indeed of any truth-functional logic, to model reasoning about uncertain propositions.

Acknowledgement

Bart Van Gasse and Chris Cornelis would like to thank the Research Foundation–Flanders for funding their research.

[4] This is for example the case if $\neg$ is involutive.

References

[1] R. Cignoli, F. Esteva, L. Godo, A. Torrens, Basic fuzzy logic is the logic of continuous t-norms and their residua, Soft Computing 4, 2000, pp. 106–112.

[2] C. Cornelis, G. Deschrijver and E.E. Kerre, Advances and Challenges in Interval-Valued Fuzzy Logic, Fuzzy Sets and Systems 157(5), 2006, pp. 622–627.

[3] G. Deschrijver and E.E. Kerre, Classes of Intuitionistic Fuzzy t-norms Satisfying the Residuation Principle, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 11, 2003, pp. 691–709.

[4] D. Dubois and H. Prade, Can we enforce full compositionality in uncertainty calculi?, in: *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI'94)*, Seattle, Washington, 1994, pp. 149–154.

[5] D. Dubois, On ignorance and contradiction considered as truth values, personal communication.

[6] F. Esteva, P. Garcia-Calvés, L. Godo, Enriched interval bilattices and partial many-valued logics: an approach to deal with graded truth and imprecision, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 2(1), 1994, pp. 37–54.

[7] F. Esteva, L. Godo, Monoidal t-norm based logic: towards a logic for left-continuous t-norms, Fuzzy Sets and Systems 124, 2001, pp. 271–288.

[8] P. Hájek, Metamathematics of fuzzy logic, Trends in Logic—Studia Logica Library, Kluwer Academic Publishers, 1998.

[9] U. Höhle, Commutative, residuated l-monoids, in: U. Höhle and E.P. Klement (Eds.), *Non-classical Logics and their Applications to Fuzzy Subsets: a Handbook of the Mathematical Foundations of Fuzzy Set Theory*, Kluwer Academic Publishers, 1995, pp. 53–106.

[10] S. Jenei, F. Montagna, A proof of standard completeness for Esteva and Godo's logic MTL, Studia Logica 70, 2002, pp. 1–10.

[11] J. Łukasiewicz, A. Tarski, Untersuchungen über den Aussagenkalkül, Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, 1930, pp. 1–21.

[12] B. Van Gasse, C. Cornelis, G. Deschrijver, E.E. Kerre, Triangle algebras: towards an axiomatization of interval-valued residuated lattices, in: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H. S. Nguyen, R. Slowinski (Eds.), *Lecture notes in computer sciences, Proceedings of RSCTC2006*, Springer, 2006.

On completeness results for the expansions with truth-constants of some predicate fuzzy logics

Francesc Esteva
AI Research Institute (IIIA)
Spanish Research Council (CSIC)
08193 Bellaterra, Catalonia, Spain
esteva@iiia.csic.es

Lluís Godo
AI Research Institute (IIIA)
Spanish Research Council (CSIC)
08193 Bellaterra, Catalonia, Spain
godo@iiia.csic.es

Carles Noguera
Dept. Comp. Science
University of Lleida
25001 Lleida, Catalonia, Spain
cnoguera@diei.udl.cat

Abstract

In this paper we study generic expansions of predicate logics of some left-continuous t-norms (mainly Gödel and Nilpotent Minimum predicate logics) with a countable set of truth-constants. Using known results on t-norm based predicate fuzzy logics we obtain results on the conservativeness and completeness for the expansions of some predicate fuzzy logics. We describe the problem for the cases of Łukasiewicz and Product predicate logics and prove that the expansions of Gödel and Nilpotent Minimum predicate logics are canonical complete for tautologies, and strong standard complete for deduction upon any set of premisses.

Keywords: Monoidal T-norm based Logic MTL, Gödel, Łukasiewicz, Product and Nilpotent Minimum propositional and predicate logics, t-norm-based logic, Rational Pavelka Logic, Gödel, Product and Nilpotent Minimum logics with truth-constants, completeness results.

1 Introduction

In the context of expansions of propositional t-norm based fuzzy logics with truth-constants, an algebraic analysis has been recently used to establish different completeness results (with respect to a finitary notion of deduction) for a number of propositional logics, among them Gödel and Nilpotent Minimum logics [5], Product logic [17], logics a continuous t-norm [4] and logics of Weak Nilpotent Minimum t-norms [6].[1]

This approach contrasts with the one related to Łukasiewicz logic, initiated by Pavelka [16] for the propositional case, and extended by Nóvak [14, 15] for the first-order case. This latter approach, based on a infinitary notion of provability, strongly relies on the continuity of the truth functions of Łukasiewicz logic (and hence not applicable to other t-norm based logics), and it was simplified by Hájek in [7], both for the propositional and first-order cases. In particular, Hájek defines what he calls *Rational Pavelka Predicate Logic*, RPL$\forall$ for short, as the expansion of Łukasiewicz predicate logic Ł$\forall$ by introducing in the language a truth-constant $\overline{r}$ for each rational r of $[0,1]$ and by adding the well-known book-keeping axioms

$$\overline{r}\&\overline{s} \leftrightarrow \overline{\max(0, r+s-1)}$$
$$(\overline{r} \rightarrow \overline{s}) \leftrightarrow \overline{\min(1, 1-r+s)}$$

Hájek shows that RPL$\forall$ enjoys the so-called *Pavelka style completeness*, which means that for any theory T and formula φ, one has

$$\|\varphi\|_T \;=\; |\varphi|_T,$$

where $\|\varphi\|_T = \inf\{\|\varphi\|_{\mathbf{M}} \mid \mathbf{M} \text{ model of } T\}$ is the truth degree of φ in T and $|\varphi|_T = \sup\{r \mid T \vdash_{\mathrm{RPL}\forall} \overline{r} \rightarrow \varphi\}$ is the provability degree of φ from T.

In this paper, following the algebraic approach mentioned above, we consider the expansions with truth-constants of the corresponding predicate logics, with special attention to the cases of Gödel and Nilpotent Minimum. In fact, to the best of our knowledge, until now only the expansion of Łukasiewicz predicate logic with truth-constants had been considered in the literature. A nice and deep result contained in [10] proves that Rational Pavelka Predicate logic RPL$\forall$ is a conservative expansion of Łukasiewicz predicate logic Ł$\forall$ and, since Ł$\forall$ is not recursively axiomatizable with respect to the standard semantics, so neither is RPL$\forall$. This is a negative result. In this paper we show other negative results, but also two positive new results. Namely, after some preliminary definitions and results in next section, we first prove that the expansions of predicate Product, Gödel, Nilpotent Minimum logics

[1] For a number these logics their complexity issues have been recently studied in [8].

(and more generally any pseudo-complemented t-norm based logic) are conservative expansions of their corresponding predicate logics. Moreover, we prove (canonical) standard completeness results for the expansions with truth-constants of Gödel and Nilpotent Minimum predicate logics. The paper ends with some conclusions and remarks.

2 Preliminaries

2.1 Propositional expansions with truth-constants

The basic logic we will use is the Monoidal t-norm based logic MTL introduced in [3] and proved to be the logic of left-continuous t-norms and their residua in [11]. In this setting, given a left-continuous t-norm $*$ we will denote by $[0,1]_*$ the standard MTL-chain given by the left-continuous t-norm $*$ and its residuum $\Rightarrow$, i. e. $[0,1]_* = \langle [0,1], *, \Rightarrow, \min, \max, 0, 1\rangle$, and by L_* the axiomatic extension of MTL whose equivalent algebraic semantics is $\mathbf{V}([0,1]_*)$, i. e. the variety generated by $[0,1]_*$. Well-known examples of these logics are the cases when $*$ is the minimum t-norm ($\mathrm{L}_*=$ G), the Łukasiewicz t-norm ($\mathrm{L}_* =$ Ł), the product t-norm ($\mathrm{L}_* = \Pi$) or the nilpotent minimum t-norm (L_* = NM).

Now, given a left-continuous t-norm $*$ and its corresponding logic L_*, let $\mathcal{C} = \langle C, *, \Rightarrow, \min, \max, 0, 1\rangle$ be a *countable* subalgebra of the standard L_*-algebra $[0,1]_*$. Then, the logic $\mathrm{L}_*(\mathcal{C})$ is defined as follows:

(i) the language of $\mathrm{L}_*(\mathcal{C})$ is the one of L_* expanded with a new propositional variable $\overline{r}$ for each $r \in C \setminus \{0,1\}$,

(ii) the axioms of $\mathrm{L}_*(\mathcal{C})$ are those of L_* plus the book-keeping axioms:

$$\overline{r}\&\overline{s} \leftrightarrow \overline{r * s}$$
$$\overline{r} \rightarrow \overline{s} \leftrightarrow \overline{r \Rightarrow_* s}$$

for each $r, s \in C$. The algebraic counterpart of the $\mathrm{L}_*(\mathcal{C})$ logic consists of the class of $\mathrm{L}_*(\mathcal{C})$-algebras, defined as structures

$$\mathcal{A} = \langle A, \&, \rightarrow, \wedge, \vee, \{\overline{r}^{\mathcal{A}} : r \in C\}\rangle$$

such that:

1. $\langle A, \&, \rightarrow, \wedge, \vee, \overline{0}^{\mathcal{A}}, \overline{1}^{\mathcal{A}}\rangle$ is an L_*-algebra, and
2. for every $r, s \in C$ the following identities hold:

$$\overline{r}^{\mathcal{A}}\&\overline{s}^{\mathcal{A}} = \overline{r * s}^{\mathcal{A}}$$
$$\overline{r}^{\mathcal{A}} \rightarrow \overline{s}^{\mathcal{A}} = \overline{r \Rightarrow s}^{\mathcal{A}}.$$

The $\mathrm{L}_*(\mathcal{C})$-chains defined over the real unit interval $[0,1]$ are called *standard*. Among them there is one which reflects the intended semantics, the so-called *canonical standard* $\mathrm{L}_*(\mathcal{C})$-chain

$$[0,1]_{\mathrm{L}_*(\mathcal{C})} = \langle [0,1], *, \Rightarrow, \min, \max, \{r : r \in C\}\rangle,$$

i. e. the standard chain where the truth-constants are interpreted by themselves. Note that, for a logic $\mathrm{L}_*(\mathcal{C})$ there may exist multiple standard chains as soon as there exist different ways of interpreting the truth-constants on $[0,1]$ respecting the book-keeping axioms. For instance, let $C = [0,1] \cap \mathbb{Q}$ and let $*$ be a *pseudo-complemented t-norm*, i.e. a left-continuous t-norm $*$ whose definable negation $\neg x = x \Rightarrow 0$ is the so-called Gödel negation, i.e. $\neg x = 0$ for all $x \neq 0$ and $\neg 0 = 1$. In such a case, if $*$ is closed on C, it is easy to check that the algebra $\mathcal{A} = \langle [0,1], *, \Rightarrow, \wedge, \vee, \{\overline{r}^{\mathcal{A}} : r \in C\}\rangle$ where

$$\overline{r}^{\mathcal{A}} = \begin{cases} 1, & \text{if } r > 0 \\ 0, & \text{otherwise} \end{cases}$$

is always a standard $\mathrm{L}_*(\mathcal{C})$ algebra. This is the case e.g. of minimum and product t-norms. Furthermore, in the particular case of $* = \min$, for any $\alpha > 0$, the algebra $\mathcal{A} = \langle [0,1], *, \rightarrow, \wedge, \vee, \{\overline{r}^{\mathcal{A}} : r \in C\}\rangle$ where

$$\overline{r}^{\mathcal{A}} = \begin{cases} 1, & \text{if } r \geq \alpha \\ 0, & \text{otherwise} \end{cases}$$

is also standard $\mathrm{G}_*(\mathcal{C})$.

Since the additional symbols added to the language are 0-ary, $\mathrm{L}_*(\mathcal{C})$ is also an algebraizable logic and its equivalent algebraic semantics is the variety of $\mathrm{L}_*(\mathcal{C})$-algebras. This, together with the fact that $\mathrm{L}_*(\mathcal{C})$-algebras are representable as a subdirect product of $\mathrm{L}_*(\mathcal{C})$-chains, leads to the following general completeness result of $\mathrm{L}_*(\mathcal{C})$ with respect to the class of $\mathrm{L}_*(\mathcal{C})$-chains. In the following, for any set $\Gamma \cup \{\varphi\}$ of $\mathrm{L}_*(\mathcal{C})$-formulae and any class $\mathbb{K}$ of $\mathrm{L}_*(\mathcal{C})$-chains, we write $\Gamma \models_{\mathbb{K}} \varphi$ to denote that, for each $\mathcal{A} \in \mathbb{K}$, $e(\varphi) = \overline{1}^{\mathcal{A}}$ for all $\mathcal{A}$-evaluation e model of Γ

Theorem 1. *For any set $\Gamma \cup \{\varphi\}$ of $\mathrm{L}_*(\mathcal{C})$-formulae, it holds that $\Gamma \vdash_{\mathrm{L}_*(\mathcal{C})} \varphi$ if, and only if, $\Gamma \models_{\{\mathrm{L}_*(\mathcal{C})-chains\}} \varphi$.*

The issue of studying when a logic $\mathrm{L}_*(\mathcal{C})$ is also complete with respect to the class of standard $\mathrm{L}_*(\mathcal{C})$-chains (called *standard completeness* property) or with respect to the canonical standard $\mathrm{L}_*(\mathcal{C})$-chain (called *canonical standard completeness* property) has been addressed in the literature for some logics L_*. Hájek already proved in [7] the canonical completeness of the expansion of Łukasiewicz logic with rational truth-constants for finite theories. More recently, the expansions of Gödel (and of some t-norm based logic related to the nilpotent minimum t-norm) and of Product logic with countable sets of truth-constants have been proved to enjoy the canonical standard completeness for theorems in [5] and in [17] respectively. A

rather exhaustive description of completeness results for the logics $\mathrm{L}_*(\mathcal{C})$ can be found in [4] and in [6].

2.2 Core predicate fuzzy logics

Predicate versions of the propositional t-norm based logics described above have also been defined and studied in the literature. Following [9] we give below a general definition of the predicate logic L$\forall$ for any (propositional) *core fuzzy logic* L. A finitary logic L in a countable language is a *core fuzzy logic* [1] if:

(i) L expands MTL;

(ii) L satisfies the congruence condition: $\varphi \leftrightarrow \psi \vdash_{\mathrm{L}} \chi(\varphi) \leftrightarrow \chi(\psi)$, for every φ, ψ, χ;

(iii) L satisfies the following local deduction theorem: $\Gamma, \varphi \vdash_{\mathrm{L}} \psi$ iff there a is natural number n such that $\Gamma \vdash_{\mathrm{L}} \varphi \& \overset{n}{\ldots} \& \varphi \rightarrow \psi$.

Note that the logics $\mathrm{L}_*(\mathcal{C})$ introduced above are core fuzzy logics, so what follows also applies to them.

Given a core fuzzy logic L, the language $\mathcal{PL}$ of L$\forall$ is built from the propositional language $\mathcal{L}$ of L by enlarging it with a set of predicates $Pred$, a set of object variables Var and a set of object constants $Const$, together with the two classical quantifiers $\forall$ and $\exists$. The notion of formula trivially generalizes taking into account that now, if φ is a formula and x is an object variable, then $(\forall x)\varphi$ and $(\exists x)\varphi$ are formulae as well.

In first-order fuzzy logics it is usual to restrict the semantics to L-chains only. For each L-chain $\mathcal{A}$ an L-interpretation for a predicate language $\mathcal{PL} = (Pred, Const)$ of L$\forall$ is a structure

$$\mathbf{M} = (M, (r_P)_{P \in Pred}, (m_c)_{c \in Const})$$

where $M \neq \emptyset$, $r_P : M^{ar(P)} \rightarrow A$ and $m_c \in M$ for each $P \in Pred$ and $c \in Const$. For each evaluation of variables $v : Var \rightarrow M$, the truth-value $\|\varphi\|^{\mathcal{A}}_{\mathbf{M},v}$ of a formula (where $v(x) \in M$ for each variable x) is defined inductively from

$$\|P(x, \cdots, c, \cdots)\|^{\mathcal{A}}_{\mathbf{M},v} = r_P(v(x), \cdots, m_c \cdots),$$

taking into account that the value commutes with connectives, and defining

$$\begin{aligned} \|(\forall x)\varphi\|^{\mathcal{A}}_{\mathbf{M},v} &= \inf\{\|\varphi\|^{\mathcal{A}}_{\mathbf{M},v'} \mid v(y) = v'(y) \text{ for all variables } y, \text{ except } x\} \\ \|(\exists x)\varphi\|^{\mathcal{A}}_{\mathbf{M},v} &= \sup\{\|\varphi\|^{\mathcal{A}}_{\mathbf{M},v'} \mid v(y) = v'(y) \text{ for all variables } y, \text{ except } x\} \end{aligned}$$

if the infimum and supremum exist in $\mathcal{A}$, otherwise the truth-value(s) remain undefined. An structure $\mathbf{M}$ is called $\mathcal{A}$-safe if all infs and sups needed for definition of the truth-value of any formula exist in $\mathcal{A}$. Then, the truth-value of a formula φ in a safe $\mathcal{A}$-structure $\mathbf{M}$ is just

$$\|\varphi\|^{\mathcal{A}}_{\mathbf{M}} = \inf\{\|\varphi\|^{\mathcal{A}}_{\mathbf{M},v} \mid v : Var \rightarrow M\}.$$

When $\|\varphi\|^{\mathcal{A}}_{\mathbf{M}} = 1$ for a $\mathcal{A}$-safe structure $\mathbf{M}$, the pair $(\mathbf{M}, \mathcal{A})$ is said to be a model for φ, written $(\mathbf{M}, \mathcal{A}) \models \varphi$. Sometimes we will call the pair $(\mathbf{M}, \mathcal{A})$ an L$\forall$-structure.

The axioms for L$\forall$ are the axioms resulting from those of L by substitution of propositional variables with formulae of $\mathcal{PL}$ plus the following axioms on quantifiers (the same used in [7] when defining BL$\forall$):

($\forall$1)	$(\forall x)\varphi(x) \rightarrow \varphi(t)$ (t substitutable for x in $\varphi(x)$)
($\exists$1)	$\varphi(t) \rightarrow (\exists x)\varphi(x)$ (t substitutable for x in $\varphi(x)$)
($\forall$2)	$(\forall x)(\nu \rightarrow \varphi) \rightarrow (\nu \rightarrow (\forall x)\varphi)$ (x not free in ν)
($\exists$2)	$(\forall x)(\varphi \rightarrow \nu) \rightarrow ((\exists x)\varphi \rightarrow \nu)$ (x not free in ν)
($\forall$3)	$(\forall x)(\varphi \vee \nu) \rightarrow ((\forall x)\varphi \vee \nu)$ (x not free in ν)

The rules of inference of L$\forall$ are modus ponens and generalization: from φ infer $(\forall x)\varphi$.

A completeness theorem for first-order BL was proven in [7] and the completeness theorems of other predicate fuzzy logics defined in the literature have been proven in the corresponding papers where the propositional logics are introduced. The following general formulation of completeness for predicate core fuzzy logics is from the paper [9].

Theorem 2. *For any core fuzzy logic* L *over a predicate language $\mathcal{PL}$, it holds that*

$$T \vdash_{\mathrm{L}\forall} \varphi \text{ iff } (\mathbf{M}, \mathcal{A}) \models \varphi \text{ for each model } (\mathbf{M}, \mathcal{A}) \text{ of } T,$$

for any set of sentences T and formula φ of the predicate language $\mathcal{PL}$.

3 Types of completeness properties and their relationships

We will use the following terminology and notation to refer to the usual three notions of completeness for core fuzzy logics.

Definition 3. *Let* L *be a core fuzzy and let* $\mathbb{K}$ *be a class of* L*-algebras. We define:*

- L *has the property of* strong $\mathbb{K}$-completeness, S$\mathbb{K}$C *for short, when for every set of* L*-formulae* Γ *and every* L*-formula* φ, $\Gamma \vdash_{\mathrm{L}} \varphi$ *iff* $\Gamma \models_{\mathbb{K}} \varphi$.
- L *has the property of* finite strong $\mathbb{K}$-completeness, FS$\mathbb{K}$C *for short, when for every* finite *set of* L*-formulae* Γ *and every* L*-formula* φ, $\Gamma \vdash_{\mathrm{L}} \varphi$ *iff* $\Gamma \models_{\mathbb{K}} \varphi$.

- L *has the property of* $\mathbb{K}$-completeness, $\mathbb{K}$C *for short, when for every* L*-formula* φ, $\vdash_{\mathrm{L}} \varphi$ *iff* $\models_{\mathbb{K}} \varphi$.

They are analogously defined for the predicate logics.

Definition 4. *Let* L *be a core fuzzy logic. We say that* L$\forall$ *has the* S$\mathbb{K}$C *if for each language* $\mathbf{\Gamma}$*, theory* T*, and formula* φ *the following are equivalent:*

- $T \vdash_{\mathrm{L}\forall} \varphi$.
- $(\mathbf{M}, \mathcal{A}) \models \varphi$ *for each* $\mathcal{A} \in \mathbb{K}$ *and each model* $(\mathbf{M}, \mathcal{A})$ *of the theory* T.

We say that L$\forall$ *has the* FS$\mathbb{K}$C *if the above condition holds for* finite *theories. Finally, we say that* L$\forall$ *has the* $\mathbb{K}$C *if the above condition holds for the empty theory.*

When $\mathbb{K}$ is the class of standard algebras in the variety of L-algebras, then instead of $\mathbb{K}$*-completeness* properties we talk about *standard completeness* properties and we use the notation $\mathcal{R}$C instead of $\mathbb{K}$C (to stress that it is a completeness with respect to algebras defined of the real unit interval). Moreover, as mentioned above, when the considered core fuzzy logic is of the form $\mathrm{L}_*(\mathcal{C})$ we can think of further restricting the semantics to the canonical standard algebra. Thus, we also consider the three *canonical standard completeness* properties for these logics both in the propositional and in the first order case.

The completeness properties, their relations and algebraic equivalent (or sufficient) conditions have been studied in [2]. In particular, the following results for the S$\mathbb{K}$C have been proved.

Theorem 5 ([2])**.** *Let* L *be a core fuzzy logic and let* $\mathbb{K}$ *be a class of* L*-algebras. Then:*

1. L *has the* S$\mathbb{K}$C *if, and only if, every countable* L*-chain can be embedded into some chain from* $\mathbb{K}$.
2. L$\forall$ *has the* S$\mathbb{K}$C *if every countable* L*-chain can be* σ*-embedded (i.e. with an embedding which preserves existing suprema and infima) into some chain from* $\mathbb{K}$.

Now we recall a relation between completeness of a propositional core fuzzy logic L and completeness of its corresponding predicate logic L$\forall$.

Proposition 6 (cf. [2])**.** *If for some family* $\mathbb{K}$ *of* L*-chains* L$\forall$ *enjoys the* $\mathbb{K}$C *(*FS$\mathbb{K}$C*,* S$\mathbb{K}$C *resp.), then* L *enjoys the* $\mathbb{K}$C *(*FS$\mathbb{K}$C*,* S$\mathbb{K}$C *resp.) as well.*

This proposition yields a necessary condition for the completeness properties of predicate fuzzy logics that will be useful to refute some completeness results in the next section. In a similar way we will also use the following proposition relating completeness of two predicate logics when one is a conservative expansion of the other one.

Proposition 7. *Let* L *and* L$'$ *be two core fuzzy logics such that* L$'\forall$ *is a conservative expansion of* L$\forall$*. Let* $\mathbb{K}'$ *be a class of* L$'$*-chains and let* $\mathbb{K}$ *be the class of their* L*-reducts. If* L$'\forall$ *enjoys the* $\mathbb{K}'$C *(*FS$\mathbb{K}'$C*,* S$\mathbb{K}'$C *resp.), then* L$\forall$ *enjoys* $\mathbb{K}$C *(*FS$\mathbb{K}$C*,* S$\mathbb{K}$C *resp.) as well.*

Proof: Assume that L$'\forall$ enjoys the $\mathbb{K}$C and we prove that L$\forall$ also enjoys it. Suppose that $\nvdash_{\mathrm{L}\forall} \varphi$ for some formula φ in language of L$\forall$. Then, since L$'\forall$ is a conservative expansion of L$\forall$ we also have $\nvdash_{\mathrm{L}'\forall} \varphi$, hence there is some structure $(\mathbf{M}, \mathcal{A}')$ with $\mathcal{A}' \in \mathbb{K}'$ such that $(\mathbf{M}, \mathcal{A}') \not\models \varphi$. Let $\mathcal{A}$ be the L-reduct of $\mathcal{A}'$. Since φ is an L$\forall$-formula, it is clear that $(\mathbf{M}, \mathcal{A}) \not\models \varphi$. □

4 Completeness results for some $\mathbf{L}_*\forall(\mathcal{C})$ logics

In the following, given a left-continuous t-norm $*$ and its corresponding logic L_*, and a *countable* subalgebra $\mathcal{C}$ of the standard L_*-algebra $[0,1]_*$, we will denote by $\mathrm{L}_*\forall(\mathcal{C})$ the predicate version of the (core fuzzy) logic $\mathrm{L}_*(\mathcal{C})$ according to the definitions in Section 2.2.

Remark about the notation used. In the way we have just defined $\mathrm{L}_*\forall(\mathcal{C})$, we should have rather used the notation $\mathrm{L}_*(\mathcal{C})\forall$, since we have started by the expanded logic $\mathrm{L}_*(\mathcal{C})$ and then we have defined the predicate variant over it. But in fact, starting with the $\mathrm{L}_*\forall$ logic and then expanding it with the truth constants from $\mathcal{C}$ leads exactly to the same predicate logic and thus we will keep using $\mathrm{L}_*\forall(\mathcal{C})$. Moreover, we will make also use of the notations G$\forall$, Ł$\forall$, $\Pi\forall$ and NM$\forall$ when referring to $\mathrm{L}_*\forall$ logic when the t-norm $*$ is the minimum, Łukasiewicz, product and nilpotent minimum t-norm respectively.

In the case of expansions of $\mathrm{L}_*\forall$ logics with truth-constants, it was already proved by Hájek *et al.* in [10] that RPL$\forall$ (Rational Pavelka predicate logic[2]) is a conservative expansion of Ł$\forall$. Next theorem proves the analogous result for other logics.

Proposition 8. *If* $*$ *is a pseudo-complemented t-norm or the nilpotent minimum t-norm, then* $\mathrm{L}_*\forall(\mathcal{C})$ *is a conservative expansion of* $\mathrm{L}_*\forall$.

Proof: Let φ be an $\mathrm{L}_*\forall$-formula such that $\nvdash_{\mathrm{L}_*\forall} \varphi$. We must show that $\nvdash_{\mathrm{L}_*\forall(\mathcal{C})} \varphi$. By hypothesis, there is some $\mathrm{L}_*\forall$-structure $(\mathbf{M}, \mathcal{A})$ such that $(\mathbf{M}, \mathcal{A}) \not\models \varphi$. It is enough to show that $\mathcal{A}$ can be expanded to an

[2] In our notation RPL$\forall$ corresponds to Ł$\forall(\mathcal{C})$ when $C = [0,1] \cap \mathbb{Q}$.

$\mathrm{L}_*(\mathcal{C})$-chain. If $*$ is a pseudo-complemented t-norm we can define the interpretation of every truth-constant $\overline{r}$ such that $r \neq 0$ as $\overline{1}^{\mathcal{A}}$ (see Section 2.1). Assume now that $*$ is the nilpotent minimum t-norm. If $\mathcal{C}$ has no negation fixpoint, we define the interpretation of a truth-constant $\overline{r}$ as $\overline{1}^{\mathcal{A}}$ when $\neg r < r$, and we define it as $\overline{0}^{\mathcal{A}}$ when $\neg r > r$. If $\mathcal{C}$ has the negation fixpoint $\frac{1}{2}$, we can suppose that $\mathcal{A}$ also has a negation fixpoint a (otherwise it could be added as shown in [13]), and then we interpret $\overline{\frac{1}{2}}$ and the rest of the constants as in the previous case. □

This result, together with the one above mentioned by Hájek *et al.*, shows that when $*$ is one of the three basic continuous t-norms (Łukasiewicz, product and minimum), $\mathrm{L}_*\forall(\mathcal{C})$ is a conservative expansion of $\mathrm{L}_*\forall$ for every subalgebra $\mathcal{C}$ of truth-constants, except for the case of Łukasiewicz t-norm where the result has only been proved for $C = [0,1] \cap \mathbb{Q}$.

Now we are prepared to deal with the standard completeness properties of predicate logics with truth-constants. It is well known that Product and Łukasiewicz predicate logics do not enjoy the standard completeness. Therefore, by Propositions 7 and 8, $\mathrm{Ł}\forall(\mathcal{C})$ and $\Pi\forall(\mathcal{C})$ do not have the $\mathbb{K}\mathrm{C}$ when $\mathbb{K}$ is the class of all standard $\mathrm{Ł}\forall(\mathcal{C})$-chains and the class of all standard $\Pi\forall(\mathcal{C})$-chains, respectively; hence these logics do not enjoy the canonical standard completeness neither. The same reasoning would also hold for every logic based on a pseudo-complemented t-norm $*$ for which we know that $\mathrm{L}_*\forall$ fails to enjoy the standard completeness.[3] This is not the case for Gödel and Nilpotent Minimum predicate logics which, in fact, satisfy the strong standard completeness. Next we show that these completeness properties extend to their expansions with truth-constants.

Theorem 9. *The logics* $\mathrm{G}\forall(\mathcal{C})$ *and* $\mathrm{NM}\forall(\mathcal{C})$ *enjoy the* $\mathrm{S}\mathcal{R}\mathrm{C}$.

Proof: As stated in the preliminaries, the strong standard completeness is equivalent to the property of σ-embeddability. Since Gödel logic already satisfies the strong standard completeness, we know that any countable G-chain is σ-embeddable into $[0,1]_{\mathrm{G}}$, thus every countable $\mathrm{G}(\mathcal{C})$-chain is also σ-embeddable in a standard $\mathrm{G}(\mathcal{C})$-chain. Indeed, given a countable $\mathrm{G}(\mathcal{C})$-chain $\mathcal{A}$ let f be the σ-embedding of its G-reduct into $[0,1]_G$. Then $\mathcal{A}$, as $\mathrm{G}(\mathcal{C})$-chain, is also σ-embeddable into the standard $\mathrm{G}(\mathcal{C})$-chain defined over $[0,1]_{\mathrm{G}}$ interpreting each truth-constant $\overline{r}$ as $f(\overline{r}^{\mathcal{A}})$. The proof for $\mathrm{NM}\forall(\mathcal{C})$ is completely analogous. □

[3] This is the case of all pseudo-complemented continuous t-norms except for $[0,1]_{\mathrm{G}}$, which enjoys the standard completeness (see [12]).

A natural question here is whether these completeness results can be improved by restricting the semantics to the canonical standard algebra. As a matter of fact, the canonical $\mathrm{FS}\mathcal{R}\mathrm{C}$ fails for the logics $\mathrm{G}(\mathcal{C})$ and $\mathrm{NM}(\mathcal{C})$, as shown in [4, 6]. Therefore, by Proposition 6, this completeness property also fails for their predicate versions. Nevertheless, we can still prove the canonical standard completeness for these logics.

Theorem 10. *The logics* $\mathrm{G}\forall(\mathcal{C})$ *and* $\mathrm{NM}\forall(\mathcal{C})$ *enjoy the canonical* $\mathcal{R}\mathrm{C}$*, i.e. the provable formulae coincide with the 1-tautologies of the canonical standard chain* $[0,1]_{\mathrm{G}(\mathcal{C})}$ *and of* $[0,1]_{\mathrm{NM}(\mathcal{C})}$ *respectively.*

Proof: We only give a sketch of the proof for $\mathrm{G}\forall(\mathcal{C})$ (the proof for $\mathrm{NM}\forall(\mathcal{C})$ is similar with the obvious changes). Soudness is obvious as usual. For the converse direction we will argue by contraposition, i.e. we will prove that if $\nvdash_{\mathrm{G}\forall(\mathcal{C})} \varphi$ for some formula φ, then there is a $\mathrm{G}\forall(\mathcal{C})$-structure $(\mathbf{M}, [0,1]_{\mathrm{G}(\mathcal{C})})$ such that $(\mathbf{M}, [0,1]_{\mathrm{G}(\mathcal{C})}) \not\models \varphi$.

If $\nvdash_{\mathrm{G}\forall(\mathcal{C})} \varphi$, then there exists a $\mathrm{G}\forall(\mathcal{C})$-structure $(\mathbf{M}, \mathcal{A})$ over a countable G-chain $\mathcal{A}$ and an evaluation v such that $\|\varphi\|^{\mathcal{A}}_{\mathbf{M},v} < \overline{1}^{\mathcal{A}}$. Take $s = \min\{r \in C \mid \overline{r}^{A} = \overline{1}^{\mathcal{A}}, r \text{ appears in } \varphi\}$ and define $g : A \to [0,1]$ by taking $g(\overline{1}^{\mathcal{A}}) = 1$ and $g\,|_{A\setminus\{\overline{1}^{\mathcal{A}}\}}$ a bijection of $A \setminus \{\overline{1}^{\mathcal{A}}\}$ into $[0,s)$ preserving existing suprema and infima, and such that $g(\overline{r}^{\mathcal{A}}) = r$ for every truth-constant appearing in φ such that $\overline{r}^{\mathcal{A}} \neq \overline{1}^{\mathcal{A}}$. Using the mapping g we can produce an structure $(\mathbf{M}', [0,1]_{\mathrm{G}(\mathcal{C})})$ in such a way that for every evaluation e on $(\mathbf{M}, \mathcal{A})$ and every evaluation e' on $(\mathbf{M}', [0,1]_{\mathrm{G}(\mathcal{C})})$ it holds that

$$\|P(t_1, t_2, .., t_n)\|^{[0,1]_{\mathrm{G}(\mathcal{C})}}_{M',e'} = g(\|P(t_1, t_2, .., t_n)\|^{A}_{M,e})$$

for each predicate symbol P.

Now by induction we can prove that given any $\mathbf{M}$ and e and their associated $\mathbf{M}'$ and e', the following statement is true for all ψ subformula of φ, then:
a) If $\|\varphi\|^{A}_{M,e} = \overline{1}^{A}$, then $\|\varphi\|^{[0,1]_{\mathrm{G}(\mathcal{C})}}_{M',e'} \geq s$,
b) If $\|\varphi\|^{A}_{M,e} \neq \overline{1}^{A}$, then $\|\varphi\|^{[0,1]_{\mathrm{G}(\mathcal{C})}}_{M',e'} = g(\|\varphi\|^{A}_{M,e})$.
From this result the theorem is easily proved. □

5 Conclusions

In this short paper we have considered the (canonical) standard completeness properties for several prominent predicate fuzzy logics enriched with constants for intermediate truth-values. Some of these properties have been denied by showing that the standard completeness already fails for the corresponding logic without additional truth-constants, while in some other cases the answer has turned out to be positive by some ad hoc proofs. The following tables collect these results.

Logic	$\mathcal{R}$C	FS$\mathcal{R}$C	S$\mathcal{R}$C
Ł$\forall(\mathcal{C})$	No	No	No
$\Pi\forall(\mathcal{C})$	No	No	No
G$\forall(\mathcal{C})$	Yes	Yes	Yes
NM$\forall(\mathcal{C})$	Yes	Yes	Yes

Logic	Can. $\mathcal{R}$C	Can. FS$\mathcal{R}$C	Can. S$\mathcal{R}$C
Ł$\forall(\mathcal{C})$	No	No	No
$\Pi\forall(\mathcal{C})$	No	No	No
G$\forall(\mathcal{C})$	Yes	No	No
NM$\forall(\mathcal{C})$	Yes	No	No

As open problems that we plan to address in a forthcoming paper we can mention the following:

1. For which left-continuous t-norms $*$ is $\mathrm{L}_*\forall(\mathcal{C})$ a conservative extension of $\mathrm{L}_*\forall$?
2. Is it possible to prove the result of Proposition 8 for Łukasiewicz logic when the language is expanded with truth-constants for irrational values?
3. Is the canonical FS$\mathcal{R}$C true for G$\forall(\mathcal{C})$ and NM$\forall(\mathcal{C})$ when the language is restricted to evaluated formulae (i.e. formulae of the kind $\overline{r} \to \varphi$, where there are no new truth-constants in φ)?
4. Investigate completeness results for the expansions of the logics $\mathrm{L}_*\forall(\mathcal{C})$ with the projection connective Δ.

Acknowledgement

The authors acknowledge partial support of the Spanish project MULOG TIN2004-07933-C03-01.

References

[1] P. Cintula. Weakly Implicative (Fuzzy) Logics (I): Basic Properties, *Archive for Mathematical Logic*, 45: 673–704, 2006.

[2] P. Cintula, F. Esteva, J. Gispert, L. Godo, F. Montagna, C. Noguera, Distinguished algebraic semantics for t-norm based fuzzy logics: methods and algebraic equivalencies, Submitted, 2007.

[3] F. Esteva and L. Godo. Monoidal t-norm based logic: Towards a logic for left-continuous t-norms. *Fuzzy Sets and Systems*, 124: 271–288, 2001.

[4] F. Esteva, J. Gispert, L. Godo, and C. Noguera. Adding truth-constants to logics of a continuous t-norm: axiomatization and completeness results. *Fuzzy Sets and Systems* 158: 597–618, 2007.

[5] F. Esteva, L. Godo and C. Noguera. On rational weak nilpotent minimum logics, *Journal of Multiple-Valued Logic and Soft Computing* 12, Number 1-2: 9–32, 2006.

[6] F. Esteva, L. Godo, and C. Noguera. On expansions of t-norm based logics with truth-constants. To appear in the book *Fuzzy Logics and Related Structures* (S. Gottwald, P. Hájek, U. Höhle and E.P. Klement eds.), 2007.

[7] P. Hájek. *Metamathematics of Fuzzy Logic*, volume 4 of *Trends in Logic-Studia Logica Library*. Dordrecht/Boston/London, 1998.

[8] P. Hájek. Computational complexity of t-norm based propositional fuzzy logics with rational truth constants. *Fuzzy Sets and Systems* 157: 677–682, 2006.

[9] P. Hájek and P. Cintula. Triangular Norm Predicate Fuzzy Logics, To appear in the book *Fuzzy Logics and Related Structures* (S. Gottwald, P. Hájek, U. Höhle and E.P. Klement eds.), 2007.

[10] P. Hájek, J. Paris, and J. Shepherdson. Rational Pavelka predicate logic is a conservative extension of Łukasiewicz predicate logic. *Journal of Symbolic Logic*, 65(2): 669–682, 2000.

[11] S. Jenei and F. Montagna. A proof of standard completeness for Esteva and Godo's logic MTL. *Studia Logica*, 70: 183–192, 2002.

[12] F. Montagna. On the predicate logics of continuous t-norm BL-algebras. *Archive for Mathematical logic*, 44: 97–114, 2005.

[13] C. Noguera, F. Esteva, and J. Gispert. On some varieties of MTL-algebras, *Logic Journal of the IGPL*, 13: 443–466, 2005.

[14] V. Novák. On the syntactico-semantical completeness of first-order fuzzy logic. Part i. Syntax and Semantics. *Kybernetika*, 26: 47–66, 1990.

[15] V. Novák. On the syntactico-semantical completeness of first-order fuzzy logic. Part II. Main results. *Kybernetika*, 26: 134–154, 1990.

[16] J. Pavelka. On Fuzzy Logic I, II, III. *Zeitschrift fur Math. Logik und Grundlagen der Math*, 25: 45–52, 119–134, 447–464, 1979.

[17] P. Savický, R. Cignoli, F. Esteva, L. Godo, C. Noguera. On product logic with truth-constants. *Journal of Logic and Computation* 16: 205-225, 2006.

Resolution Strategies for Fuzzy Description Logic

Hashim Habiballa
Institute for Research and Applications of Fuzzy Modeling
University of Ostrava, 30. dubna 22, 701 03 Ostrava 1, Czech Republic
hashim.habiballa@osu.cz

Abstract

The paper presents experimental comparison of several resolution strategies for reasoning in Fuzzy Description Logic based on Fuzzy Predicate Logic with evaluated syntax. Resolution-based reasoning is established on previous works concerning non-clausal resolution principle both theoretical and application-oriented (FPLGERDS inference engine).

Keywords: Fuzzy inference, Fuzzy Description Logic, Resolution Strategies.

1 Introduction

Description logic (DL) has been intensively studied and it forms an interesting formalism for knowledge representation [3]. Its expressive power is sufficient for applications like semantic web. On the other hand DL is relatively simple in contrast to the full First-order Logic (FOL) bringing some significant properties like decidability. Fuzzy Predicate Logic with Evaluated Syntax (FPL) [12] is a well-studied and wide-used logic capable of expressing vagueness. It has a lot of applications based on robust theoretical background. The knowledge representation itself doesn't lead to full applicable deductive system. It also requires an efficient formal proof theory. Since DL is semantically a subclass of FOL it may use various proved techniques like tableaux algorithm. However the most widely applied resolution principle [4] brings syntactically several obstacles mainly arising from normal form transformation. FPL is associating with even harder problems when trying to use the resolution principle. The solutions to these obstacles based on the non-clausal resolution [2] were already proposed in [8] and [7].

In this article we would like to present a natural integration of these two formal logical systems into Fuzzy Description Logic (FDL). It leads to the refutational resolution theorem prover for FDL ($RRTP_{FDL}$). FDL has been presented in a simple form in [13] and in more general form in [11] with strong computability and time complexity results. The article refers to some definitions from cited refutational resolution theorem provers for DL ($RRTP_{DL}$) and FPL ($RRTP_{FPL}$).

Another issue addressed in the paper concerns to the efficiency of presented inference strategies developed originally for the proving system. We show their perspectives in combination with standard proof-search strategies. The main problem for the fuzzy logic theorem proving lies in the large amount of possible proofs with different degrees and there is presented an algorithm (Detection of Consequent Formulas - DCF) solving this problem. The algorithm is based on detection of such redundant formulas (proofs) with different degrees.

2 General resolution and unification extensions for existentiality

For the purposes of ($RRTP_{FPL}$) we will use generalized principle of resolution, which is defined in the research report [1]. There is a propositional form of the rule defined at first and further it is lifted into first-order logic. We will introduce the propositional form of the general resolution.

General resolution - propositional version

$$\frac{F[G] \quad F'[G]}{F[G/\bot] \vee F'[G/\top]} \tag{1}$$

where the propositional logic formulas F and F' are the premises of inference and G is an occurrence of a subformula of both F and F'. The expression $F[G/\bot] \vee F'[G/\top]$ is the resolvent of the premises on G. Every occurrence of G is replaced by false in the first formula and by true in the second one. It is also called F the positive, F' the negative premise, and G the resolved subformula.

The proof of the soundness of the rule is similar to clausal resolution rule proof. Suppose the Interpretation I in which both premises are valid. In I, G is either true or false. If G ($\neg G$) is true in I, so is $F'[G/\top]$ ($F[G/\bot]$).

Revised version of the paper which forms the core of the handbook [2] is closely related with notion of selection functions and ordering constraints. By a selection functions we mean a mapping S that assigns to each clause C a (possibly empty) multiset $S(C)$ of negative literals in C. In other words, the function S selects (a possibly empty) negative subclause of C. We say that an atom A, or a literal $\neg A$, is selected by S if $\neg A$ occurs in $S(C)$. (There are no selected atoms or literals if $S(C)$ is empty.) As an usual ordering can be used lexicographic path ordering over a total precedence. But in this case the ordering is admissible if predicate symbols have higher precedence than logical symbols and the constants $\top$ and $\bot$ are smaller than the other logical symbols. It means the ordering is following $A \succ \equiv \succ \supset \succ \neg \succ \vee \succ \wedge \succ \top \succ \bot$. The handbook also addresses another key issues for automated theorem proving - the efficiency of the proof search. This efficiency is closely related with the notion of *redundancy*.

If we want to generalize the notion of resolution and lift it into first-order case we have to define first the notion of selection function for general clauses. General clauses are multisets of arbitrary quantifier-free formulas, denoting the disjunction of their elements. Note, that we can also work with a special case of such a general clauses with one element, which yields to a standard quantifier-free formula of first-order logic. A (general selection) function is a mapping S that assigns to each general clause C a (possibly empty) set C of non-empty sequences of (distinct) atoms in C such that either $S(C)$ is empty or else, for all interpretations I in which C is false, there exists a sequence $A_1, ..., A_k$ in $S(C)$, all atoms of which are true in I. A sequence $A_1, ..., A_k$ in $S(C)$ is said to be *selected* (by S).

We have to define the notion of polarity for these reasons according to the handbook [2]. It is based on the following assumption that a subformula F' in $E[F']$ is *positive* (resp. *negative*), if $E[F'/\top]$ (resp. $E[F'/\bot]$) is a tautology. Thus, if F' is *positive* (resp. *negative*) in E, F' (resp. $\neg F'$) logically implies E. Even it should seem that determining of the polarity of any subformula is NP-complete (hard) problem, we can use syntactic criteria for this computation. In this case the complexity of the algorithm is linear (note that we base our theory on similar syntactic criteria below - structural notions definition).

When trying to refine the general resolution rule for fuzzy predicate logic and description logic, it is important to devise a sound and complete unification algorithm. Standard unification algorithms require variables to be treated only as universally quantified ones. We will present a more general unification algorithm, which can deal with existentially quantified variables without the need for those variables be eliminated by skolemization. It should be stated that the following unification process doesn't allow an occurrence of the equivalence connective. It is needed to remove equivalence by rewrite rule: $A \leftrightarrow B \Leftrightarrow [A \rightarrow B] \wedge [B \rightarrow A]$.

We assume that the language and semantics of FOL is standard. We use terms - individuals $(a, b, c, ...)$, functions (with n arguments) $(f, g, h, ...)$, variables $(X, Y, Z, ...)$, predicates(with n arguments) $(p, q, r, ...)$, logical connectives $(\wedge, \vee, \rightarrow, \neg)$, quantifiers $(\exists, \forall)$ and logical constants $(\bot, \top)$. We also work with standard notions of logical and special axioms (sets LAx, SAx), logical consequence, consistency etc. as they are used in mathematical logic.

Definition 1 ***Structural notions of a FOL formula***
Let F be a formula of FOL then the structural mappings Sub (subformula), Sup (superformula), Pol (polarity) and Lev (level) are defined as follows:

$F = G \wedge H$ *or* $F = G \vee H$	$Sub(F) = \{G, H\}$, $Sup(G) = Sup(H) = F$ $Pol(G) = Pol(F) = Pol(H)$
$F = G \rightarrow H$	$Sub(F) = \{G, H\}$, $Sup(G) = Sup(H) = F$ $Pol(G) = -Pol(F)$, $Pol(H) = Pol(F)$
$F = \neg G$	$Sub(F) = \{G\}$, $Sup(G) = F$ $Pol(G) = -Pol(F)$
$F = \exists \alpha G$ *or* $F = \forall \alpha G$ *(α is a variable)*	$Sub(F) = \{G\}$, $Sup(G) = F$ $Pol(G) = Pol(F)$

$Sup(F) = \emptyset \Rightarrow Lev(F) = 0, Pol(F) = 1,$
$Sup(F) \neq \emptyset \Rightarrow Lev(F) = Lev(Sup(F)) + 1$
For mappings Sub and Sup reflexive and transitive closures Sub^* and Sup^* are defined recursively as follows:
1. $Sub^*(F) \supseteq \{F\}$, $Sup^*(F) \supseteq \{F\}$
2. $Sub^*(F) \supseteq \{H | G \in Sub^*(F) \wedge H \in Sub(G)\}$, $Sup^*(F) \supseteq \{H | G \in Sup^*(F) \wedge H \in Sup(G)\}$

Example: $A \rightarrow B$ - $Pol(A) = -1$, $Pol(B) = 1$, $Lev(A) = 1$

These structural mappings provide framework for assignment of quantifiers to variable occurrences. It is needed for the correct simulation of skolemization (the information about a variable quantification in the prenex form). Subformula and superformula mappings and its closures encapsulate essential hierarchical information of a formula structure. Level gives the ordering with respect to the scope of variables (which is also essential for skolemization simulation - unification is restricted for existential variables). Polarity enables to decide the global meaning of a variable (e.g. globally an existential variable is universal if its quantification subformula has negative polarity). Sound unification requires further definitions on variable quantification. We will introduce notions of the corresponding quantifier for a variable occurrence, substitution mapping and significance mapping (we have to distinguish between original variables occurring in special axioms and newly introduced ones in the proof sequence).

Definition 2 ***Variable assignment, substitution and significance***
Let F be a formula of FOL, $G = p(t_1, ..., t_n) \in Sub^*(F)$ atom in F and α a variable occurring in t_i. Variable mappings Qnt(quantifier assignment), Sbt (variable substitution) and Sig(significance) are defined as follows:

$Qnt(\alpha) = Q\alpha H, where Q = \exists \vee Q = \forall, H, I \in Sub^*(F), Q\alpha H \in Sup^*(G)$,
$\forall Q\alpha I \in Sup^*(G) \Rightarrow Lev(Q\alpha I) < Lev(Q\alpha H)$.
$F[\alpha/t']$ is a substitution of term t' into α in $F \Rightarrow Sbt(\alpha) = t'$.
A variable α occurring in $F \in LAx \cup SAx$ is significant w.r.t. existential substitution, $Sig(\alpha) = 1$ iff variable is significant, $Sig(\alpha) = 0$ otherwise.

Example: $\forall x(\forall x A(x) \rightarrow B(x))$ - $Qnt(x) = \forall x A(x)$, for x in $A(x)$ and $Qnt(x) = \forall x(\forall x A(x) \rightarrow B(x))$, for x in $B(x)$.

Note that with Qnt mapping (assignment of first name matching quantifier variable in a formula hierarchy from bottom) we are able to distinguish between variables of the same name and there is no need to rename any variable. Sbt mapping holds substituted terms in a quantifier and there is no need to rewrite all occurrences of a variable when working with this mapping within unification. It is also clear that if $Qnt(\alpha) = \emptyset$ then α is a free variable. These variables could be simply avoided by introducing new universal quantifiers to F. Significance mapping is important for differentiating between original formula universal variables and newly introduced ones during proof search (an existential variable can't be bounded with it).

Before we can introduce the standard unification algorithm, we should formulate the notion of global universal and global existential variable (it simulates conversion into prenex normal form).

Definition 3 ***Global quantification***
Let F be a formula without free variables and α be a variable occurrence in a term of F.

1. *α is a global universal variable ($\alpha \in Var_\forall(F)$) iff ($Qnt(\alpha) = \forall \alpha H \wedge Pol(Qnt(\alpha)) = 1$) or ($Qnt(\alpha) = \exists \alpha H \wedge Pol(Qnt(\alpha)) = -1$)*
2. *α is a global existential variable ($\alpha \in Var_\exists(F)$) iff ($Qnt(\alpha) = \exists \alpha H \wedge Pol(Qnt(\alpha)) = 1$) or ($Qnt(\alpha) = \forall \alpha H \wedge Pol(Qnt(\alpha)) = -1$)*

$Var_\forall(F)$ and $Var_\exists(F)$ are sets of global universal and existential variables.

Example: $F = \forall y(\forall x A(x) \rightarrow B(y))$ - x is a global existential variable, y is a global universal variable.

It is clear w.r.t. skolemization technique that an existential variable can be substituted into an universal one only if all global universal variables over the scope of the existential one have been already substituted by a term. Skolem functors function in the same way. Now we can define the most general unification algorithm based on recursive conditions (extended unification in contrast to standard MGU).

Definition 4 ***Most general unifier algorithm***
Let $G = p(t_1, ..., t_n)$ and $G' = r(u_1, ..., u_n)$ be atoms. Most general unifier (substitution mapping) MGU(G, G') = σ is obtained by following atom and term unification steps or the algorithm returns fail-state for unification. For the purposes of the algorithm we define the Variable Unification Restriction (VUR).

Variable Unification Restriction

Let F_1 be a formula and α be a variable occurring in F_1, F_2 be a formula, t be a term occurring in F_2 and β be a variable occurring in F_2. Variable Unification Restriction (VUR) for (α,t) holds if one of the conditions 1. and 2. holds:

1. *α is a global universal variable and $t \neq \beta$, where β is a global existential variable and α not occurring in t (non-existential substitution)*
2. *α is a global universal variable and $t = \beta$, where β is a global existential variable and $\forall F \in Sup^*(Qnt(\beta))$, $F = Q\gamma G$, $Q \in \{\forall, \exists\}$, γ is a global universal variable, $Sig(\gamma) = 1 \Rightarrow (Sbt(\gamma) = r') \in \sigma$, r' is a term (existential substitution).*

Atom unification

1. *if $n = 0$ and $p = r$ then $\sigma = \emptyset$ and the unifier exists (success-state).*
2. *if $n > 0$ and $p = r$ then perform term unification for pairs $(t_1, u_1), \ldots, (t_n, u_n)$; If for every pair unifier exists then $MGU(G, G') = \sigma$ obtained during term unification (success state).*
3. *In any other case unifier doesn't exist (fail-state).*

Term unification (t', u')

1. *if $u' = \alpha$, $t' = \beta$ are variables and $Qnt(\alpha) = Qnt(\beta)$ then unifier exists for (t', u') (success-state) (occurrence of the same variable).*
2. *if $t' = \alpha$ is a variable and $(Sbt(\alpha) = v') \in \sigma$ then perform term unification for (v', u'); The unifier for (t', u') exists iff it exists for (v', u') (success-state for an already substituted variable).*
3. *if $u' = \alpha$ is a variable and $(Sbt(\alpha) = v') \in \sigma$ then perform term unification for (t', v'); The unifier for (t', u') exists iff it exists for (t', v') (success-state for an already substituted variable).*
4. *if $t' = a$, $u' = b$ are individual constants and a = b then for (t',u') unifier exists (success-state).*
5. *if $t' = f(t'_1, ..., t'_m)$, $u' = g(u'_1, ..., u'_n)$ are function symbols with arguments and $f = g$ then unifier for (t', u') exists iff unifier exists for every pair (t'_1, u'_1), ..., (t'_n, u'_n) (success-state).*

6. *if $t' = \alpha$ is a variable and VUR for (t', u') holds then unifier exists for (t', u') holds and $\sigma = \sigma \cup (Sbt(\alpha) = u')$ (success-state).*
7. *if $u' = \alpha$ is a variable and VUR for (u', t') holds then unifier exists for (t', u') holds and $\sigma = \sigma \cup (Sbt(\alpha) = t')$ (success-state).*
8. *In any other case unifier doesn't exist (fail-state).*

$MGU(A) = \sigma$ for a set of atoms $A = \{G_1, \ldots, G_k\}$ is computed by the atom unification for $(G_1, G_i), \sigma_i = MGU(G_1, G_i), \forall i, \sigma_0 = \emptyset$, where before every atom unification (G_1, G_i), σ is set to σ_{i-1}.

With above defined notions it is simple to state the general resolution rule for FOL (without the equivalence connective). It conforms to the definition from [1].

Definition 5 ***General ordered resolution with selection for first-order logic*** *(GR_{FOL})*

$$\frac{F[G_1, , ..., G_k] \quad F'[G'_1, ..., G'_n]}{F\sigma[G/\bot] \vee F'\sigma[G/\top]} \tag{2}$$

where $\sigma = MGU(A)$ is the most general unifier (MGU) of the set of the atoms $A = \{G_1, \ldots, G_k, G'_1, \ldots, G'_n\}$, $G = G_1\sigma$. For every variable α in F or F', $(Sbt(\gamma) = \alpha) \cap \sigma = \emptyset \Rightarrow Sig(\alpha) = 1$ in F or F' iff $Sig(\alpha) = 1$ in $F\sigma[G/\bot] \vee F'\sigma[G/\top]$. F is called positive and F' is called negative premise, G represents an occurrence of an atom. The expression $F\sigma[G/\bot] \vee F'\sigma[G/\top]$ is the resolvent of the premises on G.
and
(i) either G is selected by S in F', or else $S(F')$ is empty, G is maximal in F', (ii) atom G is maximal in F, and (iii) F does not contain a selected atom.

Note that with Qnt mapping we are able to distinguish variables not only by its name (which may not be unique), but also with this mapping (it is unique). Sig property enables to separate variables, which were not originally in the scope of an existential variable. When utilizing the rule it should be set the Sig mapping for every variable in axioms and negated goal to 1. We present a very simple example of existential variable unification before we introduce the refutational theorem prover for FOL.

3 Fuzzy Predicate Logic and refutational proof

The fuzzy predicate logic with evaluated syntax is a flexible and fully complete formalism, which will be used for the below presented extension [12]. In order to use an efficient form of the resolution principle we have to extend the standard notion of a proof (provability value and degree) with the notion of refutational proof (refutation degree). Propositonal version of the fuzzy resolution principle has been already presented in [6]. We suppose that set of truth values is Łukasiewicz algebra. Therefore we assume standard notions of conjunction, disjunction etc. to be bound with Łukasiewicz operators. It is important that we assume that for every subformula $Sub, Sup, Pol, Lev, Qnt, Sbt, Sig$ and other derived properties defined in section 2 hold (where the classical FOL connective is presented the Łukasiewicz one has the same mapping value).

Definition 6 ***Evaluated proof, refutational proof and refutation degree***
An evaluated formal proof of a formula A from the fuzzy set $X \subsetsim F_J$ is a finite sequence of evaluated formulas $w := \ a_0/A_0, \ a_1/A_1, ..., \ a_n/A_n$ such that $A_n := A$ and for each $i \leq n$, either there exists an m-ary inference rule r such that
$a_i/A_i := \ r^{evl}(a_{i_1}, ..., a_{i_m})/r^{syn}(A_{i_1}, ..., A_{i_m})$,
$i_1, ..., i_m < n$ or $a_i/A_i := \ X(A_i)/A_i$.
We will denote the value of the evaluated proof by $Val(w) = a_n$.
An evaluated refutational formal proof of a formula A from X is w, where additionally $a_0/A_0 := \ 1/\neg A$ and $A_n := \bot$. $Val(w) = a_n$ is called refutation degree of A.

Definition 7 ***General ordered resolution with selection for fuzzy predicate logic*** *(GR_{FPL})*

$$r_{GR} : \frac{a/F[G_1, , ..., G_k], \ b/F'[G'_1, ..., G'_n]}{a \otimes b/F\sigma[G/\bot] \nabla F'\sigma[G/\top]} \tag{3}$$

where $\sigma = MGU(A)$ is the most general unifier (MGU) of the set of the atoms $A = \{G_1, \ldots, G_k, G'_1, \ldots, G'_n\}$, $G = G_1\sigma$. For every variable α in F or F', $(Sbt(\gamma) = \alpha) \cap \sigma = \emptyset \Rightarrow Sig(\alpha) = 1$ in

F or F' iff $Sig(\alpha) = 1$ in $F\sigma[G/\bot] \vee F'\sigma[G/\top]$. F is called positive and F' is called negative premise, G represents an occurrence of an atom. The expression $F\sigma[G/\bot] \nabla F'\sigma[G/\top]$ is the resolvent of the premises on G.
and
(i) either G is selected by S in F', or else $S(F')$ is empty, G is maximal in F', (ii) atom G is maximal in F, and (iii) F does not contain a selected atom.

Lemma 1 ***Soundness of*** *r_{GR}*
The inference rule r_{GR} for FPL based on $\mathcal{L}_{\text{Ł}}$ is sound i.e. for every truth valuation $\mathcal{D}$,

$$\mathcal{D}(r^{syn}(A_1, ..., A_n)) \geq r^{evl}(\mathcal{D}(A_1), ..., \mathcal{D}(A_n)) \tag{4}$$

holds true.

Definition 8 ***Refutational resolution theorem prover***
Refutational non-clausal resolution theorem prover for

FPL ($RRTP_{FPL}$) is the inference system with the inference rule GR_{FPL} and sound simplification rules for $\bot$, $\top$ (standard equivalencies for logical constants). A refutational proof by definition 6 represents a proof of a formula G (goal) from the set of special axioms N. It is assumed that $Sig(\alpha) = 1$ for $\forall\alpha$ in $F \in N \cup \neg G$ formula, every formula in a proof has no free variable and has no quantifier for a variable not occurring in the formula.

Definition 9 ***Simplification rules for*** $\nabla, \Rightarrow$

$$r_{s\nabla} : \frac{a/\bot \nabla A}{a/A} \quad and \quad r_{s\Rightarrow} : \frac{a/\top \Rightarrow A}{a/A}$$

Lemma 2 ***Provability and refutation degree for*** GR_{FPL}
$T \vdash_a A$
iff $a = \bigvee\{Val(w)|$ w is a refutational proof of A from LAx$\cup$ SAx$\}$

Theorem 1 ***Completeness for fuzzy logic with*** r_{GR}, $r_{s\nabla}$, $r_{s\Rightarrow}$ ***instead of*** r_{MP}
Formal fuzzy theory, where r_{MP} is replaced with r_{GR}, $r_{s\nabla}$, $r_{s\Rightarrow}$, is complete i.e. for every A from the set of formulas $T \vdash_a A$ iff $T \models_a A$.

4 Resolution principle for Fuzzy Description Logic

In this section we will naturally build up the Refutational Resolution Theorem Prover for Fuzzy Description Logic $RRTP_{FDL}$ based on resolution rule for FPL. Whole the theory is based on mapping FDL formulas into FPL formulas and hence the properties of interpretation and resolution are preserved. The paper works with standard interpretation structures and rules of FPL [12] and also for ALC-like FDL [11] based on Łukasiewicz algebra. We use atomic concepts to be bound with unary predicates $A_1, ..., A_n$ and atomic roles are bound with binary predicates $R_1, ..., R_m$. Terms are constants $a_1, ..., a_l$ or variables $x_1, ..., x_k$. There is also top and bottom concepts $(\top, \bot)$. Concepts may be constructed using negation $\neg$, conjunction $\&$, disjunction ∇, implication $\rightarrow$, restricted quantification $\forall R.C$ bounded with $\forall y(R(x,y) \rightarrow C(y))$, $\exists R.C$ bounded with $\exists y(R(x,y) \,\&\, C(y)))$. Instances of concepts could be written in the form $C(t)$, where t is a term. When speaking about instances of quantified concepts its meaning in FPL is equivalent to replacing the occurrence of x variable by t. For the purposes of the prover we extend the notion of roles. At first we enable to use the atomic role negation $\neg R$, which is equivalent to its FPL representation and in semantic meaning it relates to complement operation on the extent of R. We also enable to use negation, conjunction and disjunction of roles and we introduce notion of full role $(\top_R)$ and empty role $(\bot_R)$. We call formulas of FDL - FDL formulas.

Definition 10 ***Embedding of structural notions into FOL***
Let D be a DL formula. We call a formula F of FOL equivalent embedding formula by the following mapping $Emb(D) = F$ (embedding). All the above defined notions for FOL holds also for $Emb(D)$. Additionally we define a mapping Ext (extension) for D, which returns an explicit form of DL-formula with terms assigned to atomic concepts and roles.

Let C, D be concepts, R, S be roles, t, t_1, t_2 be terms, x, y be variables:
Recursive definiton of Ext:
$Ext(C) = C(x)$, $Ext(R) = R(x,y)$, $Ext(C(t)) = C(t)$, $Ext(R(t_1,t_2)) = R(t_1,t_2)$ *(for atomic concepts and roles)*

For every $\bullet \in \{\sqcap, \sqcup, \rightarrow\}$ *it holds:*
$Ext(C \bullet D) = (C \bullet D)$, $Ext(R \bullet S) = (R \bullet S)$
$Ext(\neg C) = \neg C$, $Ext(\neg R) = \neg R$

For every $Q \in \{\forall, \exists\}$ *it holds:*
$Ext(QR.C) = QR.C$,

Recursive definiton of Emb:
Let C, D be concepts and R, S be roles:
$Emb(C(t)) = C(t)$, $Emb(R(t_1,t_2) = R(t_1,t_2))$, $Emb(A) = A(x)$, $Emb(R) = R(x,y))$ *(atomic concept and role)*

$Emb(C \sqcap D) = C \wedge D$, $Emb(C \sqcup D) = C \vee D$, $Emb(C \rightarrow D) = C \rightarrow D$, $Emb(\neg C) = \neg C$, $Emb(R \sqcap S) = R \wedge S$, $Emb(R \sqcup S) = R \vee S$, $Emb(R \rightarrow S) = R \rightarrow S$, $Emb(\neg R) = \neg R$,

$Emb(\forall R.C) = \forall y(X)$, *where* $Emb(R.C) = X = R(x,y) \rightarrow C(y)$,
$Emb(\exists R.C) = \exists y(X)$, *where* $Emb(R.C) = X = R(x,y) \wedge C(y)$,
$Emb(\top) = \top$, $Emb(\bot) = \bot$, $Emb(\top_R) = \top$, $Emb(\bot_R) = \bot$.

For every DL formula D or DL formula part $R.C$ = D, a variable y occurring in D or $R.C$, where $Emb(D) = F$, and formula mappings
$Map \in \{Sub, Sup, Sub^*, Sup^*, Pol, Lev\}$ *, it holds $Map(D) = Map(F)$. For a variable y in an atomic concept or role (C or R) and F it holds $Qnt(y) = Qnt(\alpha)$, where α is a variable occurring in $Emb(C)$ or $Emb(R)$. Sig and Sbt mappings could be defined subsequently. If not stated otherwise $Sig(\alpha) = 0$. No free variables should occur in $Emb(D)$.*

Example 1 ***Examples of structural mappings***
Let $C \rightarrow \exists R.C'$ be a DL formula D. Then there are some notions (not all possible) holding for D:
$Ext(D) = C(x) \rightarrow \exists R(x,y).C'(y)$ *and*
$Pol(D) = 0$, $Pol(C) = -1$, $Pol(\exists R.C') = 1$, $Pol(C') = 1$, $Lev(D) = 0$, $Lev(\exists R.C') = 1$, $Lev(R.C) = 2$, $Lev(R) = 3$, $Sub(\exists R.C') = R(x,y) \rightarrow C'(y)$, $Sub^*(\exists R(x,y).C'(y)) = \{R(x,y) \rightarrow C'(y), R(x,y), C'(y)\}$,
$Sub(D) = \{C(x), \exists y(R(x,y) \rightarrow C'(y))\}$, $Sup(C'(y)) = R(x,y) \rightarrow C(y)$
For y in $C'(y)$ $Qnt(y) = \exists y(R(x,y) \rightarrow C(y))$

Definition 11 ***General ordered resolution with selection for Fuzzy Description Logic*** (GR_{FDL})
Let D and D' be DL formulas projected by Ext_{Max},

G_i, G'_i be an atomic concept or role (a subformula of D, D′).

$$r_{GR} : \frac{a/D[G_1, , ..., G_k],\ b/D'[G'_1, ..., G'_n]}{a \otimes b/D\sigma[G/\circ]\nabla D'\sigma[G/\bullet]} \qquad (5)$$

where $\sigma = MGU(A)$ is the most general unifier (MGU) of the set of the atoms, $A = \{G_1, \ldots, G_k, G'_1, \ldots, G'_n\}$, $G = G_1\sigma$. For G being an atomic concept $\circ = \bot$, $\bullet = \top$, for G being an atomic role $\circ = \bot_R$, $\bullet = \top_R$. For every variable α in D or D′, $(Sbt(\gamma) = \alpha) \cap \sigma = \emptyset \Rightarrow (Sig(\alpha) = 1$ in D or D′ iff $Sig(\alpha) = 1$ in $D\sigma[G/\circ]\nabla D'\sigma[G/\bullet]$.
and
(i) either G is selected by S in F′, or else $S(F')$ is empty, G is maximal in F′, (ii) atom G is maximal in F, and (iii) F does not contain a selected atom.

Definition 12 ***Refutational resolution theorem prover for FDL***
Refutational non-clausal resolution theorem prover for FDL ($RRTP_{FDL}$) is the inference system with the inference rule GR_{FDL} and sound auxiliary simplification rules. A refutational proof of the goal G from the set of axioms $N = \{A_1, ..., A_m\}$ is a sequence of DL formulas with full terms on atoms (max. atomic extension) $w := a_0/A_0,\ a_1/A_1, ...,\ a_n/A_n,\ a/\bot$, where D_i is an axiom from N, $\neg G$ or a resolvent from premises D_k and D_l $(k, l < i)$; simplification rules apply to the resolvent. It is assumed that $Sig(\alpha) = 1$ $\forall\alpha$ variable in $N \cup \neg G$.

Example 2 ***Proof of child's happiness by*** r_{GR}

Consider the following knowledge (significantly simplified in contrast to the reality) about child's happiness. We suppose that a child is happy in the degree 0.8 if it has mother and father. It can be expressed in compact FDL-formula form:
$0.8/[\exists child.female \wedge \exists child.male] \Rightarrow happy$

Further we suppose that a child is happy in the degree 0.5 if it has a lot of toys (we suppose parents are a bit more important for children). We will present proofs and then we mark the best provability degree from the following axioms. According to the definitions FDL-formulas are projected by Ext_{Max}. It was used the automated theorem prover of the author for classical logic [5]. Xa. steps represent application of simplification rules.

Common proof members (axioms):

1. $0.8/[\exists child(X,Y).female(Y) \wedge \exists child(X,Y).male(Y)] \Rightarrow happy(X)$ *(happy with parents - 0.8)*
2. $0.5/toys(X) \Rightarrow happy(X)$ *(happy with toys - 0.5)*
3. $1/child(johana, hashim)$ *(clear crisp fact)*
4. $1/child(johana, lucie)$ *(clear crisp fact)*
5. $1/male(hashim)$ *(clear crisp fact)*
6. $1/female(lucie)$ *(clear crisp fact)*
7. $0.9/toys(johana)$ *(johana has a lot of toys - 0.9)*
8. $1/\neg happy(johana)$ *(negated goal - is johana happy?)*

Proof 1:

9. $0.9 \otimes 0.5/\bot \vee [\top \Rightarrow happy(johana)]$
9a. $0.4/happy(johana)$ *(r_{GR} on 7.,2., $X = johana$)*
10. $1 \otimes 0.4/\bot \vee \neg\top$
10a. $0.4/\bot$ *(r_{GR} on 9.,8.)*

(happy(johana) is provable in 0.4)

Proof 2:

9. $0.8 \otimes 1/[[\exists child(johana,Y).female(Y) \wedge \exists child(johana,Y).male(Y)] \Rightarrow \bot]\nabla\neg\top$
9a. $0.8/\neg[\exists child(johana,Y).female(Y) \wedge \exists child(johana,Y).male(Y)]$ *(r_{GR} on 1.,8., $X = johana$)*
10. $0.8 \otimes 1/\neg[\exists child(johana, lucie).\top \wedge \exists child(johana,Y).male(Y)]]\nabla\bot$
10a. $0.8/\neg[\exists child(johana, lucie).\top \wedge \exists child(johana,Y).male(Y)]$ *(r_{GR} on 6.,9., $Y = lucie$)*
11. $0.8 \otimes 1/\neg[\exists child(johana, lucie).\top \wedge \exists child(johana, hashim).\top]\nabla\bot$
11a. $0.8/\neg[\exists child(johana, lucie).\top \wedge \neg child(johana, hashim).\top]$ *(r_{GR} on 5.,10., $Y = hashim$)*
12. $0.8 \otimes 1/\neg[\exists\top_R.\top \wedge \exists child(johana, hashim).\top]\nabla\bot$
12a. $0.8/\neg[\exists child(johana, hashim).\top]$ *(r_{GR} on 4.,11.)*
13. $0.8 \otimes 1/\neg\exists\top_R.\top\nabla\bot$
13a. $0.8/\bot$ *(r_{GR} on 3.,12.)*

(happy(johana) is provable in 0.8)

We have stated two different proofs and it is clear that several other proofs could be constructed. Let us note that these proofs either consist of redundant steps or they are variants of Proof 1 and Proof 2, where only the order of resolutions is different. So we can conclude that it is effectively provable that Johana is a happy child in the degree 0.8.

The proof of soundness and completeness of the $RRTP_{DL}$ is straightforward. Since the FPL embedding is semantically equivalent to the interpretation of FDL formulas, all the properties of $RRTP_{FPL}$ holding for FPL should also hold for $RRTP_{FDL}$.

5 Implementation and efficiency

The author also currently implements the non-clausal theorem prover into fuzzy logic and description logic as an extension of previous prover for FOL (GEneralized Resolution Deductive System - GERDS) [5]. Experiments concerning prospective inference strategies can be performed with this extension. The prover called Fuzzy Predicate Logic GEneralized Resolution Deductive System - FPLGERDS provides standard interface for input (knowledge base and goals) and output (proof sequence and results of fuzzy inference, statistics).

There are already several efficient strategies proposed by author (mainly Detection of Consequent Formulas (DCF) adopted for the usage also in FPL). With these strategies the proving engine can be implemented in "real-life"

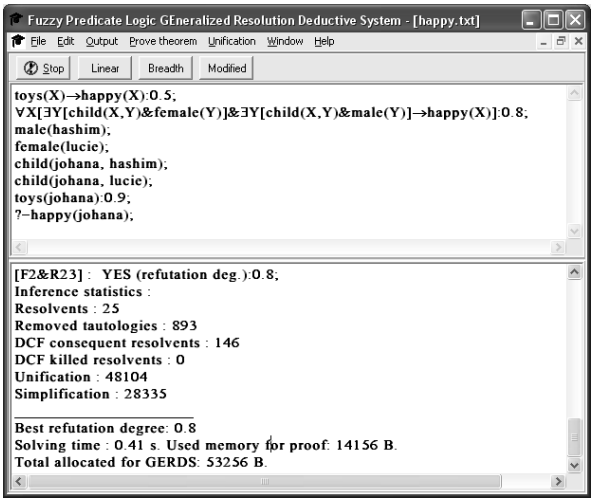


Figure 1: Fuzzy Predicate Logic GEneralized Resolution Deductive System

applications since the complexity of theorem proving in FPL is dimensionally harder than in FOL (the need to search for all possible proofs - we try to find the best refutation degree). The DCF idea is to forbid the addition of a resolvent which is a logical consequence of any previously added resolvent. For refutational theorem proving it is a sound and complete strategy and it is emiprically very effective. Completeness of such a strategy is also straight-forward in FOL:

$$(R_{old} \vdash R_{new}) \wedge (U, R_{new} \vdash \bot) \Rightarrow (U, R_{old} \vdash \bot)$$

Example: $R_{new} = p(a)$, $R_{old} = \forall x(p(x))$, $R_{old} \vdash R_{new}$.

DCF could be implemented by the same procedures like General Resolution (we may utilize self-resolution). Self-resolution has the same positive and negative premise and needs to resolve all possible combinations of an atom. It uses the following scheme:

$$R_{old} \vdash R_{new} \Leftrightarrow \neg(R_{old} \rightarrow R_{new}) \vdash \bot$$

Even the usage of this teachnique is a semidecidable problem, we can use time or step limitation of the algorithm and it will not affect the completeness of the $RRTP_{FOL}$.
Example: $R_{new} = p(a)$, $R_{old} = \forall x(p(x))$, $\neg(\forall x(p(x)) \rightarrow p(a))$
MGU: $Sbt(x) = a$, $Res = \neg(\bot \rightarrow \bot) \vee \neg(\top \rightarrow \top) \Rightarrow \bot$
We have proved that R_{new} is a logical consequence of R_{old}.

In FPL we have to enrich the DCF procedure by the limitation on the provability degree. if $U \vdash_a R_{old} \wedge U \vdash_b R_{new} \wedge b \leq a$ then we can apply DCF. DCF Trivial check performs a symbolic comparison of R_{old} and R_{new} we use the same provability degree condition. In other cases we have to add R_{new} into the set of resolvents and we can apply "DCF Kill" procedure. DCF Kill searches for every R_{old} being a logical consequence of R_{new} and if $U \vdash_a R_{old} \wedge U \vdash_b R_{new} \wedge b \geq a$ then Kill R_{old} (resolvent is removed).

We will now show some efficiency results concerning many-valued logic both for Fuzzy Predicate Logic and Fuzzy Description Logic. We have used the above mentioned application FPLGERDS and originally developed DCF strategy for FPL. It is clear that inference in $RRTP_{FPL}$ and $RRTP_{FDL}$ on general knowledge bases is a problem solved in exponential time. Nevertheless as we would like to demonstrate the need to search for every possible proof (in contrast to the two-valued logic) will not necessarily in particular cases lead to the inefficient theory. We have devised knowledge bases (KB) on the following typical problems related to the use of fuzzy logic.

We have performed experimental measurements concerning efficiency of the presented non-clausal resolution principle and also DCF technique. This measurements were done using the FPLGERDS application [9]. Special testing knowledge bases were prepared and several types of inference were tested on a PC with standard Intel Pentium 4 processor as described below.

Fuzzy Description Logic redundancy-based inefficient knowledge bases

As it was shown above in the theorem proving example the problem of proof search is quite different in FPL and FDL in comparison with the two-valued logic. We have to search for the best refutation degree using refutational theorem proving in order to make sensible conclusions from the inference process. It means we cannot accept the **"first successful"** proof, but we have to check **"all possible proofs"** or we have to be sure that every omitted proof is **"worse"** that some another one. The presented DCF and DCF Kill technique belong to the third sort of proof search strategies, i.e. they omit proofs that are really "worse" than some another (see the explication above). Proofs and formulas causing this could be called redundant proofs and redundant formulas. Fuzzy logic makes this redundancy dimensionally harder since we could produce not only equivalent formulas but also equivalent formulas of different evaluation degree.

Example 3 ***Redundant knowledge base***
Consider the following knowledge base (fragment):

...,
$0.51/a \wedge b_1 \Rightarrow z, 0.61/a \wedge b_1 \wedge b_1 \Rightarrow z,$
$0.71/a \wedge b_1 \wedge b_1 \wedge b_1 \Rightarrow z, 0.81/a \wedge b_1 \wedge b_1 \wedge b_1 \wedge b_1 \Rightarrow z,$
$0.91/a \wedge b_1 \wedge b_1 \wedge b_1 \wedge b_1 \wedge b_1 \Rightarrow z, 1/b_1,$
...,
$0.52/a \wedge b_2 \Rightarrow z, 0.62/a \wedge b_2 \wedge b_2 \Rightarrow z,$
$0.72/a \wedge b_2 \wedge b_2 \wedge b_2 \Rightarrow z, 0.82/a \wedge b_2 \wedge b_2 \wedge b_2 \wedge b_2 \Rightarrow z,$
$0.92/a \wedge b_2 \wedge b_2 \wedge b_2 \wedge b_2 \wedge b_2 \Rightarrow z, 1/b_2,$
...,

Goal: $? - a \Rightarrow z$

Searching for the best proof of a goal will produce a lot of logically equivalent formulas with different degrees. These resolvents make the inference process inefficient and one of the essential demands to the presented refutational theorem prover is a reasonable inference strategy with acceptable time complexity.

We have compared efficiency of the standard **breadth-first search**, **linear search** and **modified linear search** (starting from every formula in knowledge base) and also combinations with DCF and DCF-kill technique [9]. We have prepared knowledge bases of the size 120, 240, 360, 480 and 600 formulas. It has been compared the time and space efficiency on the criterion of 2 redundancy levels. This level represents the number of redundant formulas to which the formula is equivalent (including the original formula). For example the level 5 means the knowledge base contain 5 equivalent redundant formulas for every formula (including the formula itself). The basic possible state space search techniques and DCF heuristics and their combinations are presented in the following tables.

Table 1: Proof search algorithms

Search method		**Description**
Breadth	B	Level order generation, start - special axioms + goal
Linear	L	Resolvent ⇒ premise, start - goal
Modified-Linear	M	Resolvent ⇒ premise, start - goal + special axioms

We use standard state space search algorithms in the FPLGERDS application - Breadth-first and Linear search. Breadth-first method searches for every possible resolvent from the formulas of the level 0 (goal and special axioms). These resolvents form formulas of the level 1 and we try to combine them with all formulas of the same and lower level and continue by the same procedure until no other non-redundant resolvent could be found. Linear search performs depth-first search procedure, where every produced resolvent is used as one of the premises in succeeding step of inference. The first produced resolvents arises from the goal formula. Modified linear search method posses the same procedure as linear one, but it starts from goal and also from all the special axioms.

Table 2: DCF heuristics

DCF Method		**Description**
Trivial	T	Exact symbolic comparison
DCF	DC	Potential resolvent is consequent (no addition)
DCF Kill	DK	DCF + remove all consequent resolvents

DCF methods for reduction of resolvent space are basically three. The simplest is trivial DCF method, which detects redundant resolvent only by its exact symbolic comparison, i.e. formulas are equivalent only if the are syntactically the same. Even it is a very rough method, it is computationally very simple and forms necessary essential restriction for possibly infinite inference process. The next method of DCF technique enables do detect the equivalency of a formula (potential new resolvent) by the means described above. DCF Kill technique additionally tries to remove every redundant resolvent from the set of resolvents.

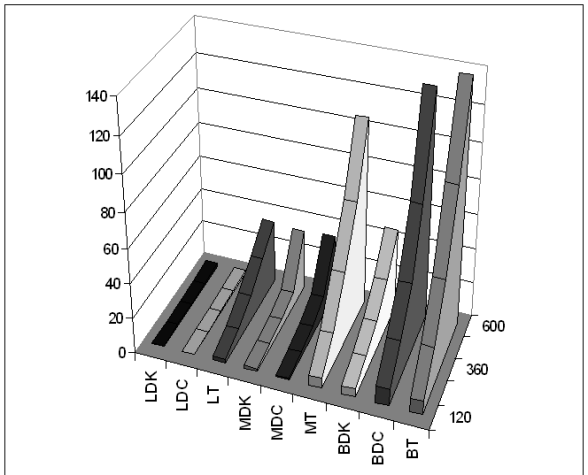
Table 3: Inference strategies

Search	**DCF**	Code	**Description**
Breadth	Trivial	BT	Complete
Breadth	DCF	BDC	Complete
Breadth	DCF Kill	BDK	Complete
Mod. Linear	Trivial	MT	Incomplete (+)
Mod. Linear	DCF	MDC	Incomplete (+)
Mod. Linear	DCF Kill	MDK	Incomplete (+)
Linear	Trivial	LT	Incomplete
Linear	DCF	LDC	Incomplete
Linear	DCF Kill	LDK	Incomplete

We have built-up 9 combinations of inference strategies from the mentioned proof search and DCF heuristics. They have different computational strength, i.e. their completeness is different for various classes of formulas. Fully complete (as described above) for general formulas of FPL and FDL are only breadth-first search combinations. Linear search strategies are not complete even for two-valued logic and horn clauses. Modified linear search has generally bad completeness results when an infinite loop is present in proofs, but for guarded knowledge bases it can assure completeness preserving better space efficiency than breadth-first search.

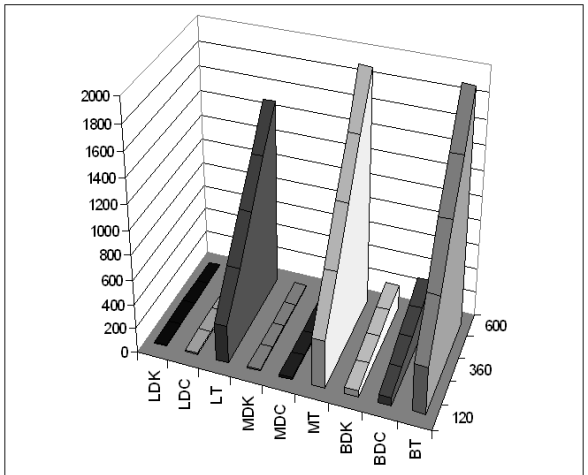
We have tested presented inference strategies on the sample knowledge bases with redundancy level 5 with 20, 40, 60, 80 and 100 groups of mutually redundant formulas (the total number of formulas in knowledge base is 120, 240, 360, 480 and 600). At first we have tested their time efficiency for inference process. As it could be observed from figure 2, the best results have **LDK and LDC** strategies. For simple guarded knowledge bases (not leading to an infinite loop in proof search and where the goal itself assures the best refutation degree) these two methods are **very efficient**. DCF strategies significantly reduces the proof search even in comparison with LT strategy (standard), therefore the usage of any non-trivial DCF heuristics is significant. Next important result concludes from the comparison of BDK and MDK, MDC strategies. We can conclude that MDK and MDC strategies are relatively comparable to BDK and moreover BDK preserves completeness for general knowledge bases.

Space complexity is even more significantly affected by the DCF heuristics. There is an interesting comparison of trivial and non-trivial DCF heuristics in figure 3. Even BDK strategy brings significant reduction of resolvents amount, while LDK, LDC, MDK, MDC strategies have minimal necessary amount of kept resolvents during inference process.



	120	240	360	480	600
BT	7,44	25,09	50,55	90,91	138,67
BDC	9,52	23,73	43,76	76,63	129,94
BDK	4,68	10,1	17,64	29,54	47,31
MT	5,86	18,85	41,46	70,01	106,68
MDC	0,93	3,09	5,14	10,54	35,14
MDK	1,45	3,2	5,03	8,84	34,32
LT	2,17	7,24	14,76	24,98	35,77
LDC	0,26	0,52	0,57	0,8	1,96
LDK	0,28	0,5	0,57	0,78	1,99

Figure 2: Time complexity for redundancy level 5 (seconds)



	120	240	360	480	600
BT	385	765	1185	1525	1905
BDC	65	125	185	245	306
BDK	46	86	126	166	205
MT	385	765	1145	1525	1905
MDC	14	14	14	14	12
MDK	14	14	14	14	12
LT	304	604	904	1204	1504
LDC	12	12	12	12	10
LDK	12	12	12	12	10

Figure 3: Space complexity for redundancy level 5 (resolvents)

Performed experiments shows the significance of originally developed DCF strategies in combination with standard breadth-first search (important for general knowledge bases - **BDK**). We also outlined high efficiency for linear search based strategies (mainly **LDK**). Even this strategy is not fully complete and could be used only for guarded fragment of FDL, this problem is already known in classical (two-valued) logic programming and automated theorem proving. We also use these highly efficient linear search strategies, even they are not complete.

6 Conclusion, further research and applications

The *Non-clausal Refutational Resolution Theorem Prover* forms a powerful inference system for automated theorem proving in fuzzy description logic. The main advantage in contrast with other inference systems lies in the possibility to utilize various inference strategies for effective reasoning. Therefore it is essential for practically successful theorem proving. The recent idea of fuzzy description logic is suitable for further extensions with the presented inference rules [11] and also reflects real situations as it could be observed from the example.

The Detection of Consequent Formulas algorithms family brings significant improvements in time and space efficiency for the best proof search. We have shown results indicating specific behavior of some combinations of the DCF and standard proof search (breadth-first and linear search). DCF strategies (BDC, BDK) have interesting results even for fully general fuzzy predicate logic with evaluated syntax, where the strategy makes the inference process practically manageable (in contrast to unrestricted "blind" proof-search). However it seems to be more promising for practical applications to utilize incomplete strategies with high time efficiency like LDK (even for large knowledge bases it has very short solving times). It conforms to another successful practical applications in two-valued logic like logic programming or deductive databases where we also use efficient incomplete strategies for fragments of fully general logics.

We have briefly shown some efficiency results for the presented automated theorem prover and inference strategies. They show the significant reduction of time and space complexity for the DCF technique. There is also prepared the experimental application FPLGERDS, which is obtainable from URL: *http://www1.osu.cz/home/habibal/files/gerds.zip*. The package contains current version of the application, source

codes, examples and documentation.

Acknowledgement

This work was supported by research plan of MSMT CR - MSM 6198898701.

References

[1] Bachmair, L., Ganzinger, H. A theory of resolution. Technical report: Max-Planck-Institut, 1997.

[2] Bachmair, L., Ganzinger, H. Resolution theorem proving. In Handbook of Automated Reasoning, MIT Press, 2001.

[3] Baader et col. (eds.). The Description Logic Handbook - Theory, Interpretation and Applications. Cambridge Univ. Press, 2003.

[4] Dukić, N., Avdagić, Z. Fuzzy Functional Dependency and the Resolution Principle. In Informatica, Vilnius: Lith. Acad. Sci. (IOSPRESS), 2005, Vol.16, No. 1, pp. 45 - 60, 2005.

[5] Habiballa, H. Non-clausal resolution - theory and practice. Research report: University of Ostrava, 2000, http://www.volny.cz/habiballa/files/gerds.pdf

[6] Habiballa, H., Novák, V. Fuzzy General Resolution. In Proc. of Intl. Conf. Aplimat 2002. Bratislava, Slovak Technical University, 2002. pp. 199-206, also available as research rep. at http://ac030.osu.cz/irafm/ps/rep47.ps

[7] Habiballa, H. Non-clausal Resolution in Fuzzy Predicate Logic with Evaluated Syntax (background and implementation). In Proc. of Intl. Conf. The Logic of Soft Computing IV, Ostrava, pp. 51 - 54, 2005, also available as research rep. at http://ac030.osu.cz/irafm/ps/rep70.ps.gz

[8] Habiballa, H. Resolution Based Reasoning in Description Logic. In Proc. of Intl. Conf. ZNALOSTI 2006, Univ. of Hradec Kralove, 2006, also available as research rep. at http://ac030.osu.cz/irafm/ps/rep66.ps.gz.

[9] Habiballa, H. Fuzzy Predicate Logic Generalized Resolution Deductive System. Technical Report, Institute for Research and Application of Fuzzy Modeling, University of Ostrava, 2006.

[10] Hájek, P. Metamathematics of fuzzy logic. Kluwer Academic Publishers - Dordrecht, 2000.

[11] Hájek, P. Making fuzzy description logic more general. Fuzzy Sets and Systems 154(2005),pp. 1-15.

[12] Novák, V., Perfilieva, I., Močkoř, J. Mathematical principles of fuzzy logic. Kluwer, 1999.

[13] Straccia, U. Reasoning within Fuzzy Description Logics. J. of AI Research 14 (2001), pp. 137-166.

Left-continuous t-norms as functional algebras

Thomas Vetterlein
European Centre for Soft Computing
Edificio Científico-Tecnológico, C/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Spain
Thomas.Vetterlein@softcomputing.es

Abstract

With a left-continuous t-norm $\odot$, we may associate the set of its vertical cuts, namely, the set F of functions $f_a : [0,1] \to [0,1],\ x \mapsto x \odot a$. Endowed with the pointwise order, with the functional composition, with the constant 0 function and with the identity function, F is an algebra which is isomorphic to $([0,1]; \leq, \odot, 0, 1)$.

We characterize the functional algebras arising in this way from left-continuous t-norms; the key property is that every two functions commute. On the basis of this approach, we describe a subclass of the left-continuous t-norms in a unified way. This subclass comprises most left-continuous t-norms discussed in the literature.

Keywords: Left-continuous t-norms, functional algebra.

1 Introduction

When compiling the most basic properties which a function interpreting the conjunction in fuzzy logics should fulfill, one arrives at the notion of a left-continuous t-norm; see e.g. [7]. The logic which is based on all left-continuous t-norms together with their respective residua, is MTL [2].

It has been an open problem for many years how to describe the structure of left-continuous t-norms in general. A general approach being missing, a great deal of work has been done to define various ways to construct specific such two-placed functions. See [5] for an overview, and see also [10].

This paper is meant as a contribution to the difficult problem how to systematize left-continuous t-norms. We have developed a simple idea with the aim to bring order into the picture, which presently causes the impression that just a little bit of creativity is needed to find once again a new left-continuous t-norm, seemingly unrelated to earlier defined ones.

Still, what we offer is a way to construct left-continuous t-norms, and not all of them are included. However, our approach differs from previous ones. In the works of S. Jenei and others (see [5]), the geometrical viewpoint is frequently stressed, and the t-norm is visualized by its three-dimensional graph. In our own previous work [9, 10], the relationship between MTL-algebras and partially ordered groups is examined.

The present work stresses again the geometric viewpoint; but no three-dimensional objects are considered, we restrict to two dimensions. We simply work with the collection of vertical cuts through the three-dimensional graph. In this way, it is possible to visualize in a pleasant way our object of study. We are given a set of functions from the real unit interval into itself, and the basic property of this set is that every two functions commute. Indeed, the commutativity of the functional composition corresponds to both the associativity and the commutativity of the underlying t-norm.

When taking into account the typical examples of left-continuous t-norms, we see that the associated functional algebras show certain regularities. We have collected these regularities and derived from them a general way to construct left-continuous t-norms. In this way, we systematize a class of t-norms which have a particularly simple structure, but is not too restrictive either.

In this paper, we present our idea and its mathematical specification, together with a demonstration how the method works in known cases. The further elaboration is the subject of a forthcoming paper.

2 Preliminaries

The basic notion of this paper is the following one.

Definition 2.1 An operation $\odot : [0,1]^2 \to [0,1]$ is called a *t-norm* if, for all $a, b, c \in [0,1]$, (i) $(a \odot b) \odot c = a \odot (b \odot c)$, (ii) $a \odot b = b \odot a$, (iii) $a \odot 1 = a$, and (iv) $a \leq b$ implies $a \odot c \leq b \odot c$. A t-norm $\odot$ is called *left-continuous*, or *l.c.* for short, if for each $a \in [0,1]$, the function $(0,1] \to [0,1]$, $x \mapsto x \odot a$ is left-continuous.

Let $\odot$ a l.c. t-norm. Then we call $([0,1]; \leq, \odot, 0, 1)$ the *t-norm monoid* based on $\odot$.

Note that usually, we associate with a l.c. t-norm the MTL-algebra $([0,1]; \wedge, \vee, \odot, \to, 0, 1)$, where $\to$ is the residuum belonging to $\odot$. Classifying the t-norm algebras is, however, the same as to classify the t-norm monoids, as the operations $\odot$ and $\to$ are mutually definable. We will deal here only with the monoidal operation $\odot$.

The article is based on a simple observation, subject of the subsequent Theorem 2.3.

Definition 2.2 Let $\odot$ be a l.c. t-norm. For any $a \in [0,1]$, call

$$f_a^{\odot} : [0,1] \to [0,1], \quad x \mapsto x \odot a$$

the *vertical cut*, or *cut* for short, of $\odot$ at the point a. Moreover, call

$$F^{\odot} = \{f_a^{\odot} : a \in [0,1]\}$$

cut set belonging to $\odot$.

We will denote the cut of a l.c. t-norm at a simply by f_a and the cut set by F, since the reference to a specific t-norm will always be clear.

Theorem 2.3 *Let $\odot$ be a l.c. t-norm. Then the cut set F belonging to $\odot$ is a set of functions from $[0,1]$ to $[0,1]$ with the following properties:*

(T1) *Every $f \in F$ is increasing.*

(T2) *Every two functions in F commute, that is, $f \circ g = g \circ f$ for any $f, g \in F$.*

(T3) *For every $a \in [0,1]$, there is exactly one $f \in F$ such that $f(1) = a$.*

(T4) *Every $f \in F$ is on $(0,1]$ left-continuous.*

Conversely, let F be a set of functions from $[0,1]$ to $[0,1]$ fulfilling (T1)–(T4)*. Then F is the cut set of a l.c. t-norm $\odot$, and $\odot$ is uniquely determined by*

$$a \odot b = f(a), \text{ where } f \in F \text{ is such that } f(1) = b. \quad (1)$$

Proof. Clearly, the cut set of a l.c. t-norm fulfills the conditions (T1)–(T4).

Conversely, let F be a set of functions from $[0,1]$ to itself such that (T1)–(T4) hold. Using (T3), denote by f_a, where $a \in [0,1]$, the unique element of F such that $f_a(1) = a$. We may then define $\odot$ by (1).

From (T2), we conclude $a \odot b = f_b(a) = f_b(f_a(1)) = f_a(f_b(1)) = f_a(b) = b \odot a$. Again by (T2), we have $(a \odot b) \odot c = f_c(a \odot b) = f_c(f_b(a)) = f_b(f_c(a)) = (a \odot c) \odot b$, and since $\odot$ is proved to be commutative, $\odot$ is associative. Moreover, $a \odot 1 = 1 \odot a = f_a(1) = a$. Finally, $a \leq b$ implies $a \odot c = f_c(a) \leq f_c(b) = b \odot c$ by (T1). So $\odot$ is a t-norm, which is left-continuous by (T4).

Obviously, F is the cut set belonging to $\odot$. In particular, f_a is the cut of $\odot$ at a. □

We shall endow the cut set belonging to a t-norm with an algebraic structure. For pairs of functions $f, g : [0,1] \to [0,1]$, we denote the pointwise order by $\leq$. Furthermore, with each pair $f, g : [0,1] \to [0,1]$, we associate their composition $f \circ g$. The basic properties of the order $\leq$ and of the operation $\circ$ on F are as follows.

Lemma 2.4 *Let F be a cut set belonging to some l.c. t-norm. Then:*

(T5) *The pointwise defined partial order $\leq$ on F is a total order which is moreover dense, separable, complete, and bounded. The lower bound is*

$$\bar{0} : [0,1] \to [0,1], \quad x \mapsto 0,$$

and the upper bound is

$$id : [0,1] \to [0,1], \quad x \mapsto x.$$

Moreover, all suprema are calculated pointwise.

(T6) *F is closed under functional composition, that is, $f \circ g \in F$ for any $f, g \in F$.*

We note that, by Theorem 2.3, the properties (T5)–(T6) are consequences of (T1)–(T4).

We summarize what we have shown.

Definition 2.5 Let F be a cut set belonging to a l.c. t-norm $\odot$. Endow F with the pointwise order $\leq$, with the composition of functions $\circ$, and with the constants $\bar{0}$ and id. Then we call $(F; \leq, \circ, \bar{0}, id)$ the functional algebra belonging to $\odot$.

Theorem 2.6 *A l.c. t-norm algebra $([0,1]; \leq, \odot, 0, 1)$ and the functional algebra $(F; \leq, \circ, \bar{0}, id)$ belonging to*

$\odot$ are isomorphic. The isomorphism is given by $\Phi(a) = f_a$, where $a \in [0,1]$ and f_a is the unique element of F such that $f_a(1) = a$.

In this way, a one-to-one correspondence is defined between the l.c. t-norms and the algebras of functions from $[0,1]$ to $[0,1]$ fulfilling the properties (T1)–(T4).

Note that this correspondence can be extended to all t-norms, by dropping the condition (T4).

3 Examples

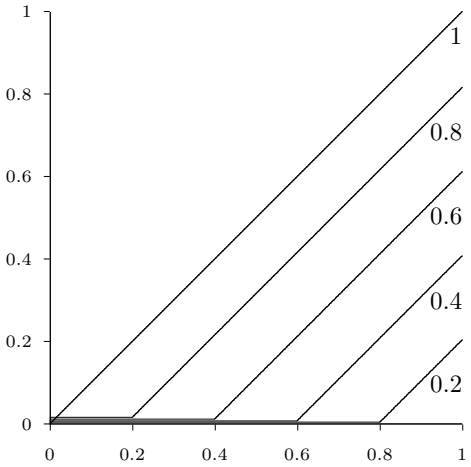
Let us have a look at the cut sets of some frequently encountered l.c. t-norms.

We consider first the continuous t-norms. Let $\odot_L$ be the Łukasiewicz t-norm: let

$$a \odot_L b = (a + b - 1) \vee 0$$

for $a, b \in [0,1]$. We plot the cuts of $\odot_L$ at the points $0.2, 0.4, ..., 1$:

The Łukasiewicz t-norm



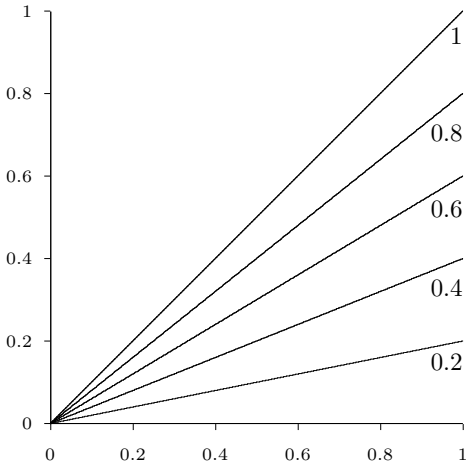
Note that our figures have only a schematic character; the cuts are drawn inaccurate in case that two of them overlap.

Next, let $\odot_P$ be the product t-norm: let

$$a \odot_P b = a \cdot b$$

for $a, b \in [0,1]$, where $\cdot$ denotes the multiplication of reals.

The product t-norm

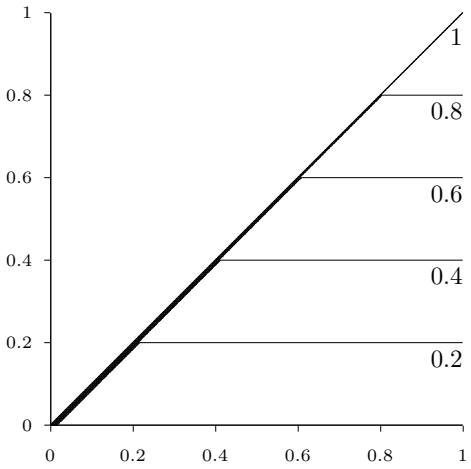


Finally, let $\odot_G$ be the Gödel t-norm: let

$$a \odot_G b = a \wedge b$$

for $a, b \in [0,1]$.

The Gödel t-norm



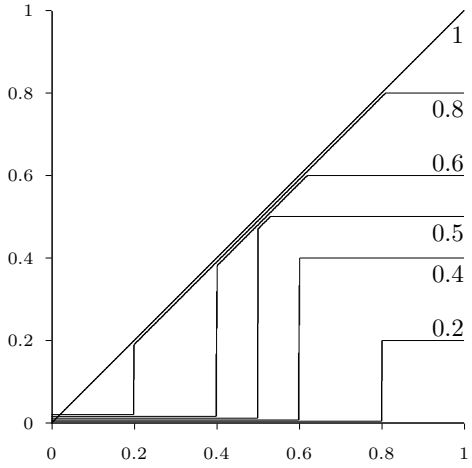
Any continuous t-norm arises from the three mentioned ones by means of an ordinal sum. So with these three examples, it is easy to see how the cut set of an arbitrary continuous t-norm looks like.

We now turn to examples of non-continuous l.c. t-norms. These are taken from the survey article [5], where further explanations and also the appropriate references can be found. Let $\odot_n$ be the nilpotent minimum t-norm: let

$$a \odot_n b = \begin{cases} a \wedge b & \text{if } a + b > 1, \\ 0 & \text{else} \end{cases}$$

for $a, b \in [0,1]$. Discontinuity points are marked in the subsequent figures by vertical lines inserted into the graph, whose lower edge indicates the respective function value.

The nilpotent minimum t-norm

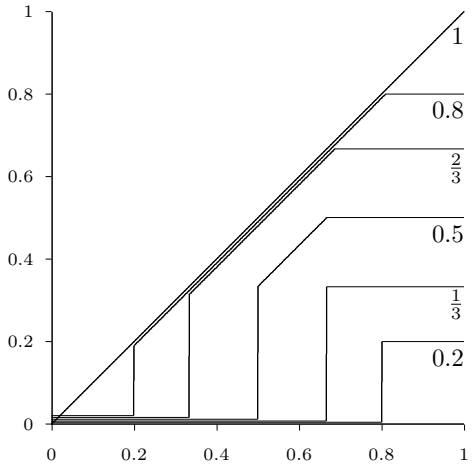


Next, we consider the t-norm $\odot_J$: let

$$a \odot_J b = \begin{cases} a \wedge b & \text{if } a+b>1 \text{ and } a \le \frac{1}{3} \text{ or } a > \frac{2}{3}, \\ a+b-\frac{2}{3} & \text{if } a+b>1 \text{ and } \frac{1}{3} < a,b \le \frac{2}{3}, \\ 0 & \text{if } a+b \le 1 \end{cases}$$

for $a, b \in [0,1]$.

The t-norm J

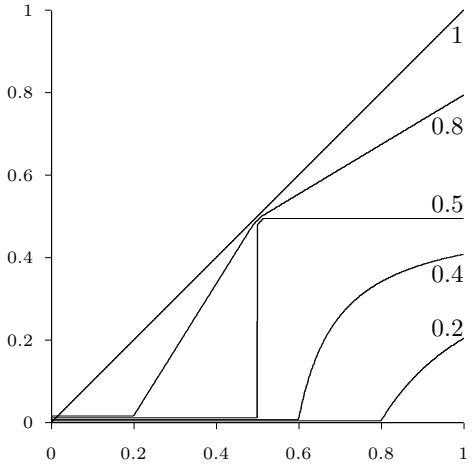


Furthermore, let $\odot_{rP}$ the rotated product t-norm: let

$$a \odot_{rP} b = \begin{cases} 2ab-a-b+1 & \text{if } a,b > \frac{1}{2}, \\ \frac{a+b-1}{2a-1} & \text{if } a > \frac{1}{2},\ b \le \frac{1}{2}, \\ & \text{and } a+b>1, \\ 0 & \text{if } a+b \le 1. \end{cases}$$

for $a, b \in [0,1]$.

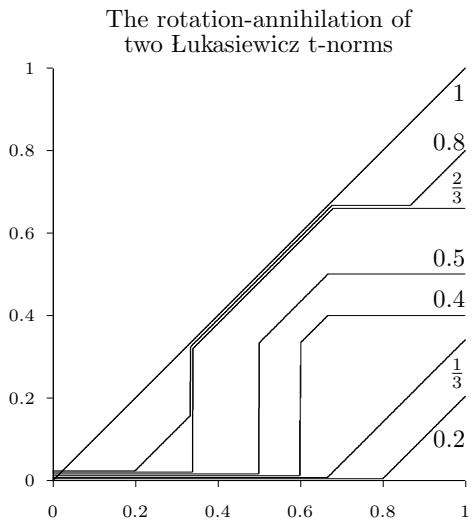
The rotated product t-norm



Finally, let $\odot_{raLL}$ be the rotation-annihilation of two Łukasiewicz t-norms: let

$$a \odot_{raLL} b = \begin{cases} a+b-1 & \text{if } a,b > \frac{2}{3} \text{ and } a+b > \frac{5}{3} \text{ or} \\ & a \le \frac{1}{3},\ b > \frac{2}{3} \text{ and } a+b>1, \\ \frac{2}{3} & \text{if } a,b > \frac{2}{3} \text{ and } a+b \le \frac{5}{3}, \\ a+b-\frac{2}{3} & \text{if } \frac{1}{3} < a,b \le \frac{2}{3} \text{ and } a+b>1, \\ a & \text{if } \frac{1}{3} < a \le \frac{2}{3},\ b > \frac{2}{3}, \\ 0 & \text{if } a+b \le 1. \end{cases}$$

for $a, b \in [0,1]$.



4 The regular left-continuous t-norms

We will single out a class of left-continuous t-norms which allow a particularly easy description. All examples of the previous section belong to this class.

As the first step, we restrict the number of discontinuity points. Let $\odot$ be a l.c. t-norm; we introduce the following condition.

(R1) There is a an $n < \omega$ such that each cut of $\odot$ has at most n points of discontinuity.

From the universal-algebraic point of view, the class of l.c. t-norms fulfilling only this condition (R1) can be considered quite large. Namely, the algebras $([0,1]; \wedge, \vee, \odot, \to, 0, 1)$, where $\odot$ is a l.c. t-norm fulfilling (R1), generate the variety of MTL-algebras. Indeed, this variety is generated by its finite members [1], and every finite member can be isomorphically embedded into a t-norm algebra as explained in [6], and this t-norm algebra fulfills (R1).

Among the l.c. t-norm ruled out by (R1), we may mention the t-norm defined by Hájek in [4] and the t-norm defined by Hliněná in [8].

Our second condition refers directly to the representation of a l.c. t-norm $\odot$ by the corresponding cut set F. Recall that the cut of $\odot$ at 1 is the identity, the largest function in F; the smaller $a \in [0,1]$ gets, the smaller is the cut at a; finally, the cut at 0 is the zero function, the smallest function in F. In particular, the set of points which belong both to the graph of the identity function and to the closure of the graph of an $f_a \in F$ gets smaller if we make a smaller. It is this dependence which we assume to be of an easy kind.

Recall that for a left-continuous function $f\colon [0,1] \to [0,1]$, the right-side limit of f at a point x is denoted by $f^+(x)$.

Definition 4.1 For each $f \in F$, let

$$Q(f) = \{x \in [0,1]\colon f^+(x) = x\},$$

and put $q(a) = Q(f_a)$ for each $a \in [0,1]$. Moreover, let $\mathcal{C}([0,1])$ be the set of closed subsets of $[0,1]$ containing 0, and partially order $\mathcal{C}([0,1])$ by the set-theoretic inclusion.

Note that q is a monotonously increasing function from $[0,1]$ to $\mathcal{C}([0,1])$. The dependence of $q(a)$ from a is in the cases of commonly known l.c. t-norm simple, but can in general be arbitrarily complicated. This motivates us to introduce the following further condition on $\odot$. The power set of a set X will be denoted by $\mathcal{P}(X)$.

(R2) There are $0 = x_0 < x_1 < \ldots < x_k = 1$, $k \geq 1$, and $0 = a_0 < a_1 < \ldots < a_l = 1$, $l \geq 1$, such that, for every $1 \leq i < k$ and $1 \leq j < l$, the function

$$\begin{aligned} q_i^j\colon \quad & [a_{j-1}, a_j) \ \to \ \mathcal{P}((x_{i-1}, x_i)), \\ & b \ \mapsto \ q(b) \cap (x_{i-1}, x_i) \end{aligned}$$

is either constant $\emptyset$, or constant (x_{i-1}, x_i), or of the form $q(b) = (x_{i-1}, r(b))$ for some order-preserving bijection $r : [a_{j-1}, a_j) \to [x_{i-1}, x_i)$, or of the form $q(b) = (l(b), x_i)$ for some order-reversing bijection $l\colon [a_{j-1}, a_j) \to (x_{i-1}, x_i]$.

The condition (R2) is more restrictive than (R1). Let us call a t-norm *regular* if it fulfills (R1) and (R2). We have:

- All continuous t-norms are regular.
- A finite ordinal sum of regular l.c. t-norms is again regular.
- The t-norms constructed from regular l.c. t-norms by means of annihilation, rotation, or rotation-annihilation, as explained in [5], are again regular.

On the other side, an ordinal sum of countably many regular l.c. t-norms such that the number of discontinuity points of the cuts is not globally bounded, does not fulfill (R1). Furthermore, let M be the finite totally ordered monoid given in [3, Chapter 5] as an example for the property of not being formally integral. The dual of M expands to a finite MTL-algebra which, embedded into a t-norm algebra according to [6], results in a l.c. t-norm fulfilling (R1), but not (R2).

5 The description of regular left-continuous t-norms

Regular t-norms are made up from only six different kinds of "building blocks".

The six possible constituents of regular l.c. t-norms can be identified with certain algebras of functions from $[0,1]$ to $[0,1]$.

Definition 5.1 Let C be a set of functions $f_a : [0,1] \to [0,1]$, where $0 \leq a \leq 1$, such that $f_0 = \bar{0}$, $f_1 = id$, and $f_0 \leq f \leq f_1$ for every $f \in C$. Endow C with the order $\leq$, the functional composition $\circ$, and the constant id.

(i) Let $f_a\colon [0,1] \to [0,1]$, $x \mapsto a \cdot x$. Then C is called the *product algebra*.

(ii) Let $f_a\colon [0,1] \to [0,1]$, $x \mapsto (x + a - 1) \vee 0$. Then C is called the *Łukasiewicz algebra*.

(iii) Let $f_a\colon [0,1] \to [0,1]$, $\frac{(a+x-1)\vee 0}{a}$ if $a > 0$. Then C is called the *reversed product algebra*.

(iv) Let $f_a\colon [0,1] \to [0,1]$, $x \mapsto x^{\frac{1}{a}}$ if $a > 0$. Then C is called the *power algebra*.

(v) Let

$$f_a(x) = \begin{cases} x & \text{for } x > a, \\ 0 & \text{for } x \leq a \end{cases}$$

for $x \in [0,1]$. Then C is called the *left-idempotency algebra*.

(vi) Let

$$f_a(x) = \begin{cases} x & \text{for } x \leq a, \\ a & \text{for } x \geq a \end{cases}$$

for $x \in [0,1]$. Then C is called the *right-idempotency algebra*.

In the sequel, we will actually have to do with sets of functions from an interval $(a,b]$, rather than $[0,1]$, to itself. Each of the algebras just defined may be transformed by means of an order-preserving bijection from $[0,1]$ to $(a,b]$, and will then be called a isomorphic to the algebra of the respective type.

Definition 5.2 Let $\odot$ be a l.c. t-norm and $(F; \leq, \circ, \bar{0}, id)$ the functional algebra belonging to $\odot$, and let $0 \leq a < b \leq 1$. Let $F_{(a,b]}$ be the set of functions

$$f_{(a,b]}\colon [a,b] \to [a,b], \quad t \mapsto \begin{cases} f(t) & \text{if } t > a \text{ and } f(t) > a, \\ a & \text{else,} \end{cases}$$

where $f \in F$; and endow $F_{(a,b]}$ with the pointwise order $\leq$, the functional composition $\circ$, and the identity $id|_{[a,b]}$. Then $(F_{(a,b]}; \leq, \circ, id|_{[a,b]})$ is called the *basic algebra* belonging to the interval $(a,b]$.

Moreover, the *parameter set* belonging to $(a,b]$ is the smallest left-open right-closed interval $(u,v]$ such that, for any $f \in F$, $a < f(b) < b$ if and only if $u < f(1) < v$.

So given an interval $(a,b] \subseteq [0,1]$, the basic algebra belonging to $(a,b]$ arises from F by considering the graph of each cut f only inside the triangle with the points (a,a), (b,b), and (b,a). More precisely, we first restrict f to $(a,b]$; we map each $x \in (a,b]$ to $f(x)$ only in case that this value is larger than a, else we map x to a; and we map the remaining boundary point a to a.

The crucial observation is the following.

Theorem 5.3 *Let $\odot$ be a regular l.c. t-norm. Then there are $0 = a_0 < \ldots < a_k = 1$ such that, for each $i = 1, ..., k$, the algebra $F_{(a_{i-1},a_i]}$ is isomorphic to the product, Łukasiewicz, reversed product, power, left-idempotency, or right-idempotency algebra.*

According to this theorem, let us associate with a regular l.c. t-norm its *characteristic data*: (i) the *basic intervals* $(a_0, a_1], ..., (a_{k-1}, a_k]$, (ii) for each $i = 1, ..., k$, the *type* of $F_{(a_{i-1},i]}$, which is one of product, Łukasiewicz, reversed product, power, left-idempotency, or right-idempotency, and (iii) for each $i = 1, ..., k$, the parameter set belonging to $(a_{i-1}, a_i]$. We remark that this data can in general not be chosen unambiguously.

As an example, consider the t-norm $\odot_J$ above. In this case, the basic intervals are $(0, \frac{1}{3}]$, $(\frac{1}{3}, \frac{2}{3}]$, and $(\frac{2}{3}, 1]$. $F_{(0,\frac{1}{3}]}$ is of type left-idempotency, $F_{(\frac{1}{3},\frac{2}{3}]}$ is of type Łukasiewicz, and $F_{(\frac{2}{3},1]}$ is of type right-idempotency. The parameter set belonging to $(0, \frac{1}{3}]$ and $(\frac{2}{3}, 1]$ is $(\frac{2}{3}, 1]$; the parameter set belonging to $(\frac{1}{3}, \frac{2}{3}]$ is $(\frac{1}{3}, \frac{2}{3}]$.

In this case, like in the case of all other examples given in Section 3, this information determines the t-norm, up to isomorphism, uniquely. However, this is not true in the general case.

Definition 5.4 We call a regular l.c. t-norm $\odot$ *locally determined* if any other regular l.c. t-norm with the same characteristic data is isomorphic to $\odot$.

One sufficient condition to ensure that a regular l.c. t-norm is locally determined is as follows. Here, a function f is called idempotent if $f \circ f = f$.

Proposition 5.5 *Let $\odot$ be a regular l.c. t-norm. Let the following condition hold:*

(R3) For each basic interval $(a,b]$ whose parameter is

also $(a, b]$, the cut f_a restricted to $[0, a]$ is idempotent.

Then $\odot$ is locally determined.

6 Conclusion

We have presented an approach to describe the so-called regular l.c. t-norms in a unified way. The main tool is the algebra of commuting functions naturally associated with each t-norm. Not every l.c. t-norm is regular, but as measured by the set of l.c. t-norms which can be found in the literature, the class of regular t-norms is quite comprehensive.

Our method provides also a recipe how to construct new l.c. t-norms. To this end, the characteristic data is to be chosen tentatively in a way that, e.g. using condition (R3), a whole functional algebra is determined by it. It is then straightforward to check whether this functional algebra actually corresponds to a l.c. t-norm.

To characterize the exact rules according to which we have to associate with a number of basic intervals types and paramater sets, so as to get the characteristic data of a left-continuous t-norm, remains as a task for subsequent papers.

References

[1] A. Ciabattoni, G. Metcalfe, F. Montagna, Adding modalities to MTL and its extensions, in: U. Höhle et al. (eds.), Proceedings of the Linz Symposium 2005, to appear.

[2] F. Esteva, L. Godo, Monoidal t-norm based logic: Towards a logic for left-continuous t-norms, *Fuzzy Sets Syst.* **124** (2001), 271 - 288.

[3] K. Evans, M. Konikoff, J. J. Madden, R. Mathis, G. Whipple, Totally ordered commutative monoids, *Semigroup Forum* **62** (2001), 249 - 278.

[4] P. Hájek, Observations on the monoidal t-norm logic, *Fuzzy Sets Syst.* **132** (2003), 107 - 112.

[5] S. Jenei, How to construct left-continuous triangular norms – state of the art, *Fuzzy Sets Syst.* **143** (2004), 27 - 45.

[6] S. Jenei, F. Montagna, A proof of standard completeness for Esteva and Godo's logic MTL, *Stud. Log.* **70** (2002), 183 - 192.

[7] E. P. Klement, R. Mesiar, E. Pap, "Triangular Norms", Kluwer Acad. Publ., Dordrecht 2000.

[8] D. Smutná (Hliněná), On a peculiar t-norm, *Busefal* **75** (1998), 60 - 67.

[9] T. Vetterlein, Weak effect algebras, *Alg. universalis*, to appear.

[10] T. Vetterlein, MTL-algebras arising from ly ordered groups, in: S. Gottwald, U. Höhle, P. Klement, "Fuzzy logics and related structures. Selected papers from the 26th Linz seminar on fuzzy set theory (Linz 2005)"; to appear.

A logic determined by commutative residuated lattices

M.Kondo *
School of Information Environment
Tokyo Denki University, Japan
kondo@sie.dendai.ac.jp

O.Watari
Hokkaido Automotive
Engineering College, Japan
watari@haec.ac.jp

M.F.Kawaguchi
Division of Computer Science
Hokkaido University, Japan
mayuka@main.ist.hokudai.ac.jp

M.Miyakoshi
Division of Computer Science
Hokkaido University, Japan
miyakosi@main.ist.hokudai.ac.jp

Abstract

We give an axiom system of a logic CRL which is characterized by the class of commutative residuated lattices. It has as axiomatic extensions well-known logics, UL by Metcalfe and Montanga (ML by Höhle, MTL by Esteva and L.Godo, BL by Hájek). Moreover we show that in any CRL algebra, the following four conditions (PL), (C_1) + (C_2), (E_1^*) and (E_2^*) are equivalent to each other in $\mathcal{CRL}$. This is a generalization of the results proved in [5] and [6].

Keywords: commutative residuated lattice, MTL, BL, UL

1 Introduction

Residuation is a fundamental concept of ordered structures and categories. So far algebraic researches with residuations have done and are doing now for many logics. This is a hot research field in many-valued logic, especially. Many logics are characterized by the class of algebras based on lattices which have residuations. For example, propositional logic, intuitionistic logic, Hájek's BL (basic logic) and Lukasiewizc's MV (many-valued logic) are determined by the class of Boolean algebras, Heyting algebras, BL-algebras and of MV-algebras, respectively. All of these algebras have a common base as algebras, lattices with residuations. Thus, it is very important to develop logics in the view point of residuations. Here we propose an axiom system called CRL which completeness theorem will be proved by the class of *commutative residuated lattices*. The axiom system of the logic CRL has naturally interpreted axioms, that is, all algebraic properties of commutative residuated lattices are reflected directly to axioms of CRL. Thus, it is easy to prove the completeness theorem. Also it gives a uniform treat of many other logics based on commutative residuated lattices, in particular UL by Metcalfe and Montanga ([4]), ML by Höhle ([2]), MTL by Esteva and Godo ([1]), BL by Hájek ([3]), uniform based logic ([7]) and so on.

This work was supported by Tokyo Denki University Science Promotion Fund (Q06J-12)

2 Logic CRL

We define a logic CRL here, which is determined by the class of all commutative residuated lattices. Thus, we name it CRL. The logic has a following language.

Propositional variables : $p_0, p_1, p_2, \cdots$

Constants : $E, \bot$

Logical symbols : $\wedge, \vee, \rightarrow, \circ$

A formula of CRL is defined as follows:

(1) Every propositional variable is a formula;

(2) Each constant is a formula;

(3) If A and B are formulas, then so are $A \wedge B, A \vee B, A \rightarrow B, A \circ B$.

Let $\Phi_0 = \{p_0, p_1, p_2, \cdots\} \cup \{E, \bot\}$ and Φ be the set of all formulas of CRL. A logical system CRL has the following axioms and rules of inference:

Axioms:

(A1) $A \rightarrow A \vee B$

(A2) $A \vee B \rightarrow B \vee A$

(A3) $A \wedge B \rightarrow A$

(A4) $A \wedge B \rightarrow B \wedge A$

(A5) $A \circ B \rightarrow B \circ A$

(A6) $A \circ E \rightarrow A$

(A7) $A \rightarrow A \circ E$

(A8) $\bot \to A$

(A9) $(A \to B) \to ((B \to C) \to (A \to C))$

Rules of inference

$$(R_\vee)\ \frac{A \to C \ \ B \to C}{A \vee B \to C},\quad (R_\wedge)\ \frac{A \to B \ \ A \to C}{A \to B \wedge C}$$

$$(R_{I\circ})\ \frac{A \to (B \to C)}{A \circ B \to C},\ (R_{E\circ})\ \frac{A \circ B \to C}{A \to (B \to C)}$$

$$(\mathrm{MP})\ \frac{A \quad A \to B}{B}$$

A formula A is called *provable* in CRL when there is a finite sequence $A_1, A_2, \cdots, A_n (= A)$ $(n \geq 1)$ of formulas such that, for every i $(1 \leq i \leq n)$,

(1) A_i is an axiom;

(2) A_i is deduced from A_j, A_k $(j, k < i)$ by one of the rules of inference above.

By $\vdash_{CRL} A$, we mean that A is provable in CRL. We note that $A \to (\bot \to \bot)$ is provable for every formula A in CRL. Thus we use a new symbol $\top$ as an abbreviation of the formula $\bot \to \bot$. Thus for every formula A we have $\vdash_{CRL} A \to \top$. It is easy to prove the following.

Proposition 1. *For all $A, B, C \in \Phi$, we have*

(1) $\dfrac{A \to B \quad B \to C}{A \to C}$

(2) $\vdash A \to A$

(3) $\vdash A \circ (A \to B) \to B$

(4) $\dfrac{E \to A}{A}$, $\dfrac{A}{E \to A}$

(5) $\vdash (A \circ B) \circ C \to A \circ (B \circ C)$

(6) $\vdash A \circ (B \circ C) \to (A \circ B) \circ C$

(7) $\dfrac{A \to B}{A \wedge C \to B \wedge C}$

(8) $\dfrac{A \to B}{A \vee C \to B \vee C}$

(9) $\dfrac{A \to B}{A \circ C \to B \circ C}$

(10) $\dfrac{A \to B}{(B \to C) \to (A \to C)}$

(11) $\dfrac{A \to B}{(C \to A) \to (C \to B)}$

3 CRL algebra

We define an algebraic semantics of the logic CRL, which gives a completeness theorem of CRL. By CRL algebra $(X, \wedge, \vee, \cdot, \to, 0, e, 1)$, we mean the following algebra

(1) $(X, \wedge, \vee, 0, 1)$ is a bounded lattice

(2) $(X, \cdot, e)$ is a commutative monoid with unit element e

(3) For all $x, y, z \in X$, $x \cdot y \leq z$ if and only if $x \leq y \to z$.

As to properties of CRL algebras, we have the following.

Proposition 2. *Let X be a CRL algebra. For all $x, y, z \in X$, we have*

(1) $x \leq y \iff e \leq x \to y$

(2) $x \cdot (x \to y) \leq y$

(3) $x \leq y \implies x \cdot z \leq y \cdot z,\ z \to x \leq z \to y,\ y \to z \leq x \to z$

(4) $e \to x = x$

(5) $x \to 1 = 1$

(6) $1 \to x \leq x$

(7) $e \leq 0 \to x$

(8) $1 \cdot 1 = 1$

(9) $(x \vee y) \cdot z = (x \cdot z) \vee (y \cdot z)$

Let X be a CRL algebra. A map $v : \Phi_0 \to X$ is called a *valuation* on X and it can be extended uniquely to the set Φ of all formulas as follows:

(1) $v(A \wedge B) = v(A) \wedge v(B)$

(2) $v(A \vee B) = v(A) \vee v(B)$

(3) $v(A \to B) = v(A) \to v(B)$

(4) $v(A \circ B) = v(A) \cdot v(B)$

We denote the extended valuation v by the same symbol v. It is easy to show that $v(\bot) = 0, v(E) = e$ and $v(\top) = 1$. The following is easy to prove.

Lemma 1 (Soundness Theorem)**.** *For every formula A, if $\vdash_{CRL} A$ then $v(A) \geq e$ for any valuation v on any CRL algebra X.*

We use the well-known method called *Lindenbaum-Tarski* algebra to prove the converse direction (Completeness Theorem) of the above. At first we define a relation $\equiv$ on the set Φ of formulas of CRL : For $A, B \in \Phi$,

$$A \equiv B \iff \vdash_{CRL} A \to B \text{ and } \vdash_{CRL} B \to A$$

As to the relation $\equiv$, we see directly from the proposition 1 that

Lemma 2. *$\equiv$ is a congruence on Φ, that is, it is an equivalence relation and satisfies the substitution property : If $A \equiv B$ and $C \equiv D$, then*

$$A \wedge C \equiv B \wedge D,\ A \vee C \equiv B \vee D,$$
$$A \to C \equiv B \to D,\ A \circ C \equiv B \circ D$$

Since $\equiv$ is the congruence, we can define operations on $\Phi/\equiv$: For $A, B \in \Phi$, we define

$$[A] \sqcap [B] = [A \wedge B],$$
$$[A] \sqcup [B] = [A \vee B],$$
$$[A] \circ [B] = [A \circ B],$$
$$[A] \to [B] = [A \to B],$$
$$\mathbf{0} = [\bot],\ \mathbf{e} = [E],\ \mathbf{1} = [\top].$$

Lemma 3. *$(\Phi/\equiv, \sqcap, \sqcup, \circ, \to, \mathbf{0}, \mathbf{e}, \mathbf{1})$ is a CRL algebra.*

We note that if $\vdash_{CRL} A \to B$ then $[A] \leq [B]$, that is, $\mathbf{e} \leq [A] \to [B]$ in the CRL algebra $\Phi/\equiv$.

We define a valuation $V^* : \Phi_0 \to \Phi/\equiv$ as $V^*(p) = [p]$, then we have $V^*(A) = [A]$ for every formula A of CRL.

Lemma 4. *For any formula $A \in \Phi$,*

$$\vdash_{CRL} A \iff V^*(A) \geq \mathbf{e} \text{ in } \Phi/\equiv$$

Proof. If $\vdash A$, since $\vdash A \circ E \to A$, then we have $\vdash A \to (E \to A)$. It follows from assumption that $\vdash E \to A$. This means that $\mathbf{e} \leq [A]$.

Conversely, $\mathbf{e} \leq [A]$ implies $\vdash E \to A$. It follows from proposition 1 that $\vdash A$. □

From the above, we can prove the next theorem.

Theorem 1 (Completeness Theorem)**.** *Let $A \in \Phi$. Then we have*

> *$\vdash_{CRL} A$ if and only if $v(A) \geq e$ for every valuation $v : \Phi_0 \to X$ on any CRL algebra X.*

Proof. Suppose that $v(A) \geq e$ for every valuation $v : \Phi_0 \to X$ on any CRL algebra X. Thus, in particular, we have $V^*(A) \geq \mathbf{e}$ for the valuation V^* on the CRL algebra $\Phi/\equiv$. From the lemma above it follows that $\vdash A$. □

Next we consider relations between our logic and other familiar logics, MTL, ML, UL and so on. The following is the axiom system of MTL according to [4]:

Axioms:

(MTL1) $(A \to B) \to ((B \to C) \to (A \to C))$

(MTL2) $A \circ B \to A$

(MTL3) $A \circ B \to B \circ A$

(MTL4) $A \wedge B \to A$

(MTL5) $A \wedge B \to B \wedge A$

(MTL6) $A \circ (A \to B) \to A \wedge B$

(MTL7a) $(A \to (B \to C)) \to (A \circ B \to C)$

(MTL7b) $(A \circ B \to C) \to (A \to (B \to C))$

(MTL8) $((A \to B) \to C) \to (((B \to A) \to C) \to C)$

(MTL9) $\bot \to A$

Rules of inference

(MP) $\dfrac{A \quad A \to B}{B}$

It is easy to prove that our logic CRL is equivalent to the system MTL without (I) $A \to (B \to A)$ and $(PL1)$ $(A \to B) \vee (B \to A)$. That is, MTL is an extension of our logic CRL with two axioms (I) and $(PL1)$. More accurately, we take $\mathcal{CRL}+(I)+(PL1)$ as a logic which has extra axioms (I) $A \to (B \to A)$ and (PL1) $(A \to B) \vee (B \to A)$ besides those of CRL. Then we can show without difficulty that MTL is equivalent to the logic CRL + (I) + (PL1), that is,

$$\vdash_{MTL} A \text{ if and only if } \vdash_{CRL+(I)+(PL1)} A.$$

4 Extension of CRL logic

By adding some extra axioms to CRL according to [5] and [6], we can consider interesting subvarieties of the variety of commutative residuated lattices. We list axioms to be added to CRL.

(I) $A \to (B \to A)$

$(PL1)$ $(A \to B) \vee (B \to A)$

$(PL2)$ $(E \wedge (A \to B)) \vee (E \wedge (B \to A))$

(C_1) $E \to (A \to B) \vee (B \to A)$

(C_2) $E \wedge (A \vee B) \to (E \wedge A) \vee (E \wedge B)$

(E_1) $(A \wedge B \to C) \to (A \to C) \vee (B \to C)$

(E_2) $(A \to B \vee C) \to (A \to B) \vee (A \to C)$

(PL) $E \to (E \wedge (A \to B)) \vee (E \wedge (B \to A))$

(E_1^*) $(A \wedge B \to C) \wedge E \to ((A \to C) \wedge E) \vee ((B \to C) \wedge E)$

(E_2^*) $(A \to B \vee C) \wedge E \to ((A \to B) \wedge E) \vee ((A \to C) \wedge E)$

These axioms are correspond to the following conditions, respectively.

(I) e is the greatest element, that is, $e = 1$.

$(PL1)$ $(a \to b) \vee (b \to a) = 1$

$(PL2)$ $(e \wedge (a \to b)) \vee (e \wedge (b \to a)) \geq e$

(C_1) $e \leq (a \to b) \vee (b \to a)$

(C_2) $e \wedge (a \vee b) \leq (e \wedge a) \vee (e \wedge b)$

(E_1) $(a \wedge b \to c) \leq (a \to c) \vee (b \to c)$

(E_2) $(a \to b \vee c) \leq (a \to b) \vee (a \to c)$

(PL) $e \leq (e \wedge (a \to b)) \vee (e \wedge (b \to a))$

(E_1^*) $(a \wedge b \to c) \wedge e \leq ((a \to c) \wedge e) \vee ((b \to c) \wedge e)$

(E_2^*) $(a \to b \vee c) \wedge e \leq ((a \to b) \wedge e) \vee ((a \to c) \wedge e)$

As the similar argument above, we see that the *monoidal logic* ML ([2]) is the CRL logic with the extra axiom (I), the *monoidal t-norm logic* MTL ([1]) is the CRL logic with (I) and $(PL1)$ and the *uninorm logic* UL [4] is one with $(PL2)$. We represent such situation by the following informal form.

$$ML = CRL + (I)$$
$$MTL = CRL + (I) + (PL1)$$
$$UL = CRL + (PL2)$$

As to these axioms we have the following fundamental result.

Proposition 3. *For all formulas A, B, we have $\vdash_{CRL} A \to (B \to A)$ if and only if $\vdash_{CRL} \top \to E$.*

From the above we see that $\vdash_{ML} A$ if and only if A is valid on the class of all CRL algebras with the condition that e is the largest element, that is, $e = 1$ in the lattice. It is similar to the other logics. Thus we have completeness theorems of these other logics.

Theorem 2. *Let $A \in \Phi$ and $\alpha, \cdots, \beta$ be some of conditions $\{(I), (PL1), \cdots, (E_1^*), (E_2^*)\}$ described above. Then we have*

$\vdash_{CRL+\{\alpha,\cdots,\beta\}} A$ if and only if $v(A) \geq e$ for every valuation $v : \Phi_0 \to X$ on any CRL algebra X with the corresponding conditions $\alpha, \cdots, \beta$.

For the case of MTL (or UL) algebra, it is known ([5]) that all subdirectly irreducible MTL (UL)-algebras are totally ordered. Thus it follows that the logic MTL (UL) is characterized by the totally ordered MTL (UL)-algebras, that is,

$\vdash_{MTL} A$ $(\vdash_{UL} A)$ if and only if $v(A) = 1$ $(v(A) \geq e)$ for every valuation $v : \Phi_0 \to X$ on any *totally ordered* MTL (UL) algebras X.

Moreover the class $\mathcal{MTL}$ of all MTL algebras is the least distributive variety containing the class $\mathcal{WL}$ of all CRL algebras with (C_1) $(x \to y) \vee (y \to x) \geq e$. Indeed, let $\mathcal{V}$ be the class of all distributive CRL algebras containing $\mathcal{WL}$. Since any algebra A in $\mathcal{MTL}$ is distributive, it satisfies the condition (C_2) $e \wedge (a \vee b) = (e \wedge a) \vee (e \wedge b)$. Thus A satisfies (C_1) and (C_2). It follows from the result in [5] that $\mathcal{MTL} \subseteq \mathcal{V}$.

As to the other conditions above, it is proved that

• Three conditions (C_1), (E_1) and (E_2) are equivalent to each other in $\mathcal{CRL}$ with (I) ([6]) ;

• Three conditions (C_1), (E_1) and (E_2) are equivalent to each other in $\mathcal{CRL}$ with (C_2) ([5]).

We now show that

Theorem 3. *The following four conditions (PL), (C_1) + (C_2), (E_1^*) and (E_2^*) are equivalent to each other in $\mathcal{CRL}$.*

Proof. We only show that (PL) implies (E_1^*). Since the class of all CRL algebras with (PL) is the variety whose subdirectly irreducible members are just totally ordered CRL algebras ([5]), it is enough to show that any totally ordered CRL algebra satisfies the condition (E_1^*). Simple calculation yields to the desired result. □

References

[1] F. Esteva and L. Godo, Monoidal t-norm based logic : towards a logic for left-continuous t-norm, Fuzzy sets and systems, vol.124 (2001), 271-288

[2] U. Höhle, Commutative, residuated l-monoids, in : U. Höhle, E.P.Klements (Eds.), Non-classical logics and their applications to the fuzzy subsets, Kluwer Acad. Publ., Dordrecht 1995, 53-106

[3] P. Hájek, Metamathematics of fuzzy logic, Kluwer Academic Publishers, 1998

[4] G. Metcalfe and F. Montanga, Substructural fuzzy logics, to appear in Journal of Symbolic Logic.

[5] J.B. Hart, L. Rafter and C. Tsinakis, The structure of commutative residuated lattices, International Journal of Algebra and Computation, Vol.12 (2002), 509-524.

[6] M. Ward and R.P. Dilworth, Residuated lattices, Trans. of the AMS, vol.45 (1939), 335-354

[7] O. Watari, M.F. Kawaguchi and M. Miyakoshi, Uninorm based logic as an extension of substructural logics FLe, Proceedings of 11th International Conference of Information processing and Management of Uncertainty in Knowledge-based Systems (IPMU2006), Paris (2006), 460-465

On the first place antitonicity in QL-implications

Y. Shi
Ghent University,
Krijgslaan 281 (S9),
9000 Gent, Belgium
Yun.Shi@UGent.be

D. Ruan
Belgian Nuclear Research Centre
(SCK•CEN), 2400 Mol, &
Ghent University,
Krijgslaan 281 (S9),
9000 Gent, Belgium
druan@SCKCEN.BE

E.E. Kerre
Ghent University,
Krijgslaan 281 (S9),
9000 Gent, Belgium
Etienne.Kerre@UGent.be

Abstract

In order to obtain a demanded fuzzy implication, a number of properties have been proposed, among which the first place antitonicity, the second place monotonicity and the boundary conditions are the most important ones. The three classes of fuzzy implications derived from the implication in binary logic, S-, R- and QL-implications all satisfy the second place monotonicity and the boundary conditions. However, not all the QL-implications satisfy the first place antitonicity as S- and R- implications do. In this paper we study the QL-implications satisfying the first place antitonicity. First we study the relationship between the first place antitonicity and other required properties of QL-implications. And then we work on the conditions under which a QL-implication generated by different combinations of a t-conorm S, a t-norm T and a strong fuzzy negation n will satisfy the first place antitonicity, especially on the cases that both S and T are continuous. We also investigate the interrelationships between S- and R-implications on one hand and QL-implications satisfying the first place antitonicity on the other.
Keywords: Fuzzy implication, QL-implication, the First place antitonicity.

1 Introduction

A fuzzy implication is a fuzzy connective that has played important roles in different fuzzy domains [5, 6, 7, 13, 14]. There are several different definitions of a fuzzy implication, e.g., [1, 3, 4, 9]. In this paper we define a fuzzy implication as a $[0,1]^2 \to [0,1]$ mapping that satisfies the boundary conditions:

I0. I(0,0)=I(0,1)=I(1,1)=1, I(1,0)=0

One of the fuzzy inference methods is the generalized modus ponens. In fuzzy logic, the generalized modus ponens is realized through IF-THEN rules. Let X and Y be two linguistic variables on the universe of discourses U and V respectively. Moreover, let A and $A^{'}$ be two fuzzy sets on U and let B and $B^{'}$ be two fuzzy sets on V. A, $A^{'}$, B and $B^{'}$ may refer to linguistic concepts. An IF-THEN rule is expressed by:
IF X is A, THEN Y is B
X is $A^{'}$

Y is $B^{'}$
where $B^{'}$ is obtained through Zadeh's compositional rule of inference:

$$B^{'}(v) = \sup_{u \in U} T(A^{'}(u), I(A(u), B(v)))$$

In this formula, T is a t-norm and I is a fuzzy implication. In order to obtain a suitable conclusion of the fuzzy inference, a number of properties have been proposed for the fuzzy implication I [4, 10, 11, 15, 17], among which the most important ones are:

I1. the first place antitonicity:
$x_1 < x_2 \Rightarrow I(x_1, y) \geq I(x_2, y)$,
for all x_1, x_2, $y \in [0,1]$;

I2. the second place monotonicity:
$y_1 < y_2 \Rightarrow I(x, y_1) \leq I(x, y_2)$,
for all x, y_1, $y_2 \in [0,1]$;

I3. the neutrality property:
$I(1, x) = x$, for all $x \in [0,1]$;

I4. the exchange principle:
$I(x, I(y, z)) = I(y, I(x, z))$, for all x, y, $z \in [0,1]$;

I5. the ordering property:
$x \leq y \Leftrightarrow I(x, y) = 1$, for all x, $y \in [0,1]$;

I6. the contrapositive principle:
$I(x, y) = I(n(y), n(x))$, for all x, $y \in [0,1]$, w.r.t. a strong fuzzy negation n;

I7. the continuity.

I1 and I2 are the most important properties. Some authors even define a fuzzy implication I as a $[0,1]^2 \to [0,1]$ mapping that satisfies I1, I2 and

I9 $I(0,x) = 1$, for all $x \in [0,1]$;

I10 $I(x,1) = 1$, for all $x \in [0,1]$;

I11 $I(1,0) = 0$.

Notice that if I satisfies I0, I1 and I2, then I9, I10 and I11 will be satisfied immediately.
There are three important classes of fuzzy implications derived from the implication in binary logic ([11], Chapter 11):

1. Strong implication (S-implication), $I(x,y) = S(n(x),y)$, where S is a t-conorm and n is a strong fuzzy negation;
2. Residuated implication (R-implication), $I(x,y) = \sup\{t \in [0,1] | T(x,t) \le y\}$, where T is a t-norm;
3. Quantum logic implication (QL-implication), $I(x,y) = S(n(x), T(x,y))$, where S is a t-conorm, T is a t-norm and n is a strong fuzzy negation.

All these three classes of fuzzy implications satisfy I0 and I2. However, although all S-implications and R-implications satisfy I1 ([10], Definition 1.15), not all the QL-implications do.
Some work on whether a QL-implication satisfies I1 or not has been done in [8], [16] and [12]. In [8], the conditions under which a QL-implication I_{QL} and a t-norm T_* satisfy the residuation property: $T_*(x,z) \le y \Leftrightarrow z \le I_{QL}(x,y)$, for all x, y, $z \in [0,1]$ are found. This means that I_{QL} is an R-implication as well ([8], Example 4.5). Hence I_{QL} satisfies I1 provided these conditions are fulfilled. However, being an R-implication is sufficient but not necessary for a QL-implication to satisfy I1 (see Remark 4, 5, 6 in this paper). In [16], the authors have worked out how a QL-implication satisfies I1 as well as I4 ([16], Definition 1, Theorem 7, Theorem 11). It is proved that such a QL-implication is an S-implication as well. Again, being an S-implication is sufficient but not necessary for a QL-implication to satisfy I1 (see Remark 5, 6 in this paper). And in [12], the authors have worked out for a group of QL-implications the conditions under which they satisfy I1. They restrict the relationship between the t-conorm and the strong fuzzy negation which construct the QL-implications ([12], Proposition 9).
In this paper, we study the QL-implications generated by a t-conorm S, a t-norm T and a strong fuzzy negation n that satisfy I1, especially for the cases that both S and T are continuous. First the relationship between I1 and the other properties of QL-implications is studied in Section 3.1. And then the conditions under which a QL-implication satisfies I1 are obtained in Section 3.2. Some QL-implications that satisfy I1 are equivalent to S-implications or R-implications while others are not. We denote these facts in Remark 4 to 7. Preliminaries are given in Section 2 and conclusions are given in Section 4 respectively.

2 Preliminaries

Definition 1. *An automorphism of the interval $[a,b] \subset \boldsymbol{R}$ is a continuous, strictly increasing mapping φ from $[a,b]$ to $[a,b]$ with boundary conditions $\varphi(a) = a$ and $\varphi(b) = b$ ([4], Definition 0).*

Lemma 1. *If ϕ is an automorphism of the unit interval, then ϕ^{-1} is also an automorphism of the unit interval.*

Lemma 2 (The chain rule)**.** *The composition of two automorphisms of the unit interval is again an automorphism.*

Definition 2. *Two mappings F, G: $[0,1]^n \to [0,1]$ are conjugated to each other, if there exists an automorphism ϕ of the unit interval such that $G = F_\phi$, where $F_\phi(x_1, x_2, \cdots, x_n)$ $= \phi^{-1}(F(\phi(x_1), \phi(x_2), \cdots, \phi(x_n)))$, $x_1, x_2, \cdots, x_n \in [0,1]$ ([2], Definition 2).*

It is easy to see that $G = F_\phi \Leftrightarrow F = G_{\phi^{-1}}$.

Definition 3. *A mapping n: $[0,1] \to [0,1]$ is a fuzzy negation if it is decreasing and satisfies: $n(0) = 1$, $n(1) = 0$.*

Definition 4. *A fuzzy negation that satisfies $n(n(x)) = x$, for all $x \in [0,1]$ is called a strong fuzzy negation.*

We denote the standard strong fuzzy negation as n_0, i.e., $n_0(x) = 1 - x$, for all $x \in [0,1]$.

Definition 5. *Let ϕ be an automorphism of the unit interval. Then n_ϕ denotes the strong fuzzy negation that is conjugated to n_0, i.e., $n_\phi(x) = \phi^{-1}(1 - \phi(x))$, for all $x \in [0,1]$.*

Definition 6. *A mapping T: $[0,1]^2 \to [0,1]$ is a triangular norm (t-norm for short) if for all $x, y, z \in [0,1]$ it satisfies:*

T1. boundary condition: $T(x,1) = x$;

T2. monotonicity: $y \le z \Rightarrow T(x,y) \le T(x,z)$;

T3. commutativity: $T(x,y) = T(y,x)$;

T4. associativity: $T(x,T(y,z)) = T(T(x,y),z)$.

Three important continuous t-norms are:

1. $T_M(x,y) = \min(x,y)$, (minimum)
2. $T_P(x,y) = xy$, (product)
3. $T_{\text{Ł}}(x,y) = \max(x+y-1,0)$, (bounded product)

Definition 7. *Let ϕ be an automorphism of the unit interval and T be a t-norm. Then T_ϕ denotes the t-norm that is conjugated to T, i.e.,* $T_\phi(x,y) = \phi^{-1}(T(\phi(x),\phi(y)))$, *for all* $(x,y) \in [0,1]^2$.

T_M is conjugated to itself, i.e., if ϕ is an automorphism of the unit interval, then $T_M = T_{M\phi}$.

Definition 8. *Let $\{[a_m,b_m]\}$ be a non-empty family of non-overlapping, closed, proper subintervals of $[0,1]$ and $\{\phi_m\}$ be a family of automorphisms of the unit interval. Then a continuous t-norm T_o is called an ordinal sum of $\{[a_m,b_m],T_m\}$, where $T_m = T_{P\phi m}$ or $T_m = T_{\text{Ł}\phi m}$, if $T_o(x,y)$ is equal to:*

$$\begin{cases} a_m + (b_m - a_m)T_m(\frac{x-a_m}{b_m-a_m}, \frac{y-a_m}{b_m-a_m}), (x,y) \in [a_m,b_m]^2 \\ T_M(x,y), \quad otherwise \end{cases}$$

If there exists only one subinterval $[a_1,b_1]$ of $[0,1]$ with $a_1 = 0$, $b_1 = 1$ and ϕ_1 being the automorphism of the unit interval, then $T_o = T_{P\phi 1}$ or $T_o = T_{\text{Ł}\phi 1}$. In this paper, as to 'an ordinal sum of $\{[a_m,b_m],T_m\}$', we mean that there exists at least one subinterval $[a_k,b_k]$ such that $a_k \neq 0$ or $b_k \neq 1$.
It is stated in ([10], Section 1.3.4) that a continuous t-norm is either T_M, or conjugated to T_P, or conjugated to $T_{\text{Ł}}$, or an ordinal sum of the non-empty family $\{[a_m,b_m],T_m\}$ with T_m being conjugated to T_P or $T_{\text{Ł}}$.

Definition 9. *A mapping S: $[0,1]^2 \to [0,1]$ is a triangular conorm (t-conorm for short) if for all $x,y,z \in [0,1]$ it satisfies:*

S1. boundary condition: $S(x,0) = x$;

S2. monotonicity: $y \leq z \Rightarrow S(x,y) \leq S(x,z)$;

S3. commutativity: $S(x,y) = S(y,x)$;

S4. associativity: $S(x,S(y,z)) = S(S(x,y),z)$.

Three important continuous t-conorms are:

1. $S_M(x,y) = \max(x,y)$, (maximum)
2. $S_P(x,y) = x+y-xy$, (probabilistic sum)
3. $S_{\text{Ł}}(x,y) = \min(x+y,1)$, (bounded sum)

Definition 10. *Let ϕ be an automorphism of the unit interval and S be a t-conorm. Then S_ϕ denotes the t-conorm that is conjugated to S, i.e.,* $S_\phi(x,y) = \phi^{-1}(S(\phi(x),\phi(y)))$, *for all* $(x,y) \in [0,1]^2$.

Definition 11. *Let ϕ be an automorphism of the unit interval and I be a fuzzy implication. Then I_ϕ denotes the fuzzy implication that is conjugated to I, i.e.,* $I_\phi(x,y) = \phi^{-1}(I(\phi(x),\phi(y)))$, *for all* $(x,y) \in [0,1]^2$.

3 QL-implications and the first place antitonicity

First we give two propositions and three lemmas that will play important roles in this section.

Proposition 1. *A necessary condition for a QL-implication generated by a t-conorm S, a t-norm T and a strong fuzzy negation n to satisfy I1, I4, I5 or I6 is $S(n(x),x) = 1$, for all $x \in [0,1]$.*

For the case that the t-conorm S is continuous, Proposition 1 can be further expressed by next proposition, according to [4] and [12].

Proposition 2. *A necessary condition for a QL-implication generated by a continuous t-conorm S, a t-norm T and a strong fuzzy negation n to satisfy I1, I4, I5 or I6 is that there exists an automorphism ϕ of the unit interval such that $S = S_{\text{Ł}^\phi}$ and n satisfies $n(x) \geq n_\phi(x)$, for all $x \in [0,1]$.*

Lemma 3. *([10], Theorem 1.13) A fuzzy implication is an S-implication if and only if it satisfies I3, I4 and I6.*

Lemma 4. *([10], Theorem 1.14) A fuzzy implication is an R-implication if and only if it satisfies I2, I4 and I5.*

Lemma 5. *Let ϕ be an automorphism of the unit interval. Then a QL-implication I_{QL} satisfies I1 iff $I_{QL\phi}$ satisfies I1.*

3.1 Relationship between the first place antitonicity and the other potential properties of QL-implications

As stated in the Introduction, all QL-implications satisfy I0 and I2. It is also easy to see that each QL-implication satisfies I3. Moreover, a QL-implication generated by a t-conorm S, a t-norm T and a strong fuzzy negation is continuous if both S and T are continuous. Thus we will only consider how a QL-implication I_{QL} satisfies I1, I4, I5 or I6 and the interrelationship between the I_{QL}'s satisfying them.

Theorem 1. *A QL-implication I_{QL} satisfies I4 iff I_{QL} is also an S-implication.*

Proof. $\Longleftarrow$: Directly from Lemma 3.
$\Longrightarrow$: According to ([12], Remark 2), if I_{QL} satisfies I4, then I_{QL} also satisfies I6. Since I_{QL} always satisfies I3, according to Lemma 3, I_{QL} is also an S-implication. □

Remark 1. As stated in the Introduction, an S-implication always satisfies I1. Thus if a QL-implication satisfies I4, then it also satisfies I1.

Next theorem is for the case that the t-conorm S which constructs the QL-implication is continuous, i.e, according to Proposition 2, there exists an automorphism ϕ of the unit interval such that $S = S_{\mathbf{L}^\phi}$.

Theorem 2. *Let ϕ be an automorphism of the unit interval. A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}^\phi}$, a t-norm T and the strong fuzzy negation n_ϕ satisfies I4 iff there exists $s \in [0, +\infty]$ such that $T = T^s_\phi$, where T^s is a Frank t-norm, defined as:*

$$T^s(x,y) = \begin{cases} T_{\mathbf{L}}(x,y), & s = 0 \\ T_P(x,y), & s = 1 \\ T_M(x,y), & s = +\infty \\ log_s(1 + \frac{(s^x-1)(s^y-1)}{s-1}), & otherwise \end{cases} \tag{1}$$

Proof. $\Longleftarrow$: According to the proof of ([12], Corollary 1), such a QL-implication I_{QL} is also an S-implication. Thus according to Lemma 3, I_{QL} satisfies I4.
$\Longrightarrow$: According to ([12], Remark 2), if I_{QL} satisfies I4, then I_{QL} also satisfies I6. Moreover, since I_{QL} satisfies I2, according to ([4], Lemma 1 (ii)), I_{QL} satisfies I1. Thus according to ([12], Proposition 9) and ([12], Corollary 1), $T_{\phi^{-1}}$ is a Frank t-norm, i.e., $T = T^s_\phi$. □

A QL-implication satisfies I4 implies that it satisfies I6, but not the reverse. Comparing next theorem and Theorem 2, we can see that there exist QL-implications that satisfy I6 but not I4.

Theorem 3. *([12], Proposition 11) Let ϕ be an automorphism of the unit interval. A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}^\phi}$, a t-norm T and the strong fuzzy negation n_ϕ satisfies I6 iff $T_{\phi^{-1}}$ is an ordinal sum of the non-empty family $\{[a_m, b_m], T_m\}$, where T_m are Frank t-norms defined in (1) with the parameter $s \in [0, +\infty[$.*

Remark 2. Since a QL-implication I_{QL} always satisfies I2, according to ([4], Lemma 1 (ii)), if I_{QL} satisfies I6, then it also satisfies I1.

Now we consider the conditions under which a QL-implication satisfies I5. Next theorem is for the case that both the t-conorm and the t-norm which construct the QL-implication are continuous.

Theorem 4. *Let ϕ be an automorphism of the unit interval. A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}^\phi}$, a continuous t-norm T and a strong fuzzy negation n satisfies I5 iff for all $y \in [0,1]$:*

i) $T(y,y) = n_\phi(n(y))$ and

ii) $T(x,y) \geq n_\phi(n(x))$, for all $x \in [0,y]$ and

iii) $T(x,y) < n_\phi(n(x))$, for all $x \in [y,1]$.

Proof. I_{QL} satisfies I5 iff
$I_{QL}(x,y) = \phi^{-1}(\min(\phi(n(x)) + \phi(T(x,y)), 1)) = 1 \Leftrightarrow x \leq y$, which means $\phi(n(x)) + \phi(T(x,y)) \geq 1 \Leftrightarrow x \leq y$.
Define for all $y \in [0,1]$, $F_y(x) = \phi(n(x)) + \phi(T(x,y))$, for all $x \in [0,1]$. Then $F_y(x) \geq 1$, i.e.,
$T(x,y) \geq n_\phi(n(x))$ iff $x \in [0,y]$ and $F_y(x) < 1$, i.e.,
$T(x,y) < n_\phi(n(x))$ iff $x \in]y,1]$. Moreover, since T, n and ϕ are all continuous, F_y is continuous. Thus $F_y(y) = 1$, i.e., $T(y,y) = n_\phi(n(y))$. □

Example 1. Let I_{QL} be the QL-implication defined in Theorem 4 with $T(y,y) = n_\phi(n(y))$, for all $y \in [0,1]$. Since T_M is the one and only the one t-norm that satisfies $T(y,y) = y$, for all $y \in [0,1]$, we have that if $T = T_M$, then $n = n_\phi$ and that if $n = n_\phi$, then $T = T_M$. Actually I_{QL} generated by $S_{\mathbf{L}^\phi}$, T_M and n_ϕ is an R-implication, i.e., $I_{QL}(x,y) = \sup\{t \in [0,1] | T_{\mathbf{L}^\phi}(x,t) \leq y\}$. According to Lemma 4, I_{QL} satisfies I4 and I5. Moreover, as stated in the Introduction, I_{QL} satisfies I1.

Remark 3. The QL-implication I_{QL} defined in Example 1 is also an S-implication, i.e.,
$I_{QL}(x,y) = S_{\mathbf{L}^\phi}(n_\phi(x), y)$. Thus according to Lemma 3, I_{QL} also satisfies I6.

3.2 QL-implications that satisfy the first place antitonicity

In this section we will focus on the characterizations of QL-implications satisfying I1. We mainly focus on the continuous cases. We will also indicate whether a QL-implication satisfying I1 is also an S-implication or an R-implication.

Theorem 5. *A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}}$, the t-norm T_M and a strong fuzzy negation n satisfies I1 iff $n(x) \geq n_0(x)$, for all $x \in [0,1]$.*

Proof. $\Longrightarrow$: Straightforward from Proposition 2.
$\Longleftarrow$: For all x_1, x_2 and $y \in [0,1]$, assume $x_1 < x_2$. If $x_1 \leq y$, then $I_{QL}(x_1,y) = S_{\mathbf{L}}(n(x_1), \min(x_1,y))$
$= S_{\mathbf{L}}(n(x_1), x_1)$. Since $n(x_1) \geq n_0(x_1)$,
$I_{QL}(x_1,y) = 1 \geq I_{QL}(x_2,y)$. Thus we need only consider the situation that $y < x_1 < x_2$. In this case,

$I_{QL}(x_1, y) = S_{\mathbf{L}}(n(x_1), \min(x_1, y)) = S_{\mathbf{L}}(n(x_1), y)$ and $I_{QL}(x_2, y) = S_{\mathbf{L}}(n(x_2), y)$. Since $S_{\mathbf{L}}(\cdot, y)$ is increasing and n is decreasing, we have $I_{QL}(x_1, y) \geq I_{QL}(x_2, y)$. Thus for all x_1, x_2 and $y \in [0, 1]$, $x_1 < x_2$ implies: $I_{QL}(x_1, y) \geq I_{QL}(x_2, y)$, i.e., I_{QL} satisfies I1. □

Corollary 1. *Let ϕ and φ denote two automorphisms of the unit interval. Then a QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}\phi}$, the t-norm T_M and a strong fuzzy negation n_φ satisfies I1 iff $n_\varphi(x) \geq n_\phi(x)$, for all $x \in [0, 1]$.*

Proof. $I_{QL}(x, y) = S_{\mathbf{L}\phi}(n_\varphi(x), T_M(x, y))$
$= \phi^{-1}(S_{\mathbf{L}}(\phi(n_\varphi(x)), T_M(\phi(x), \phi(y))))$.
Putting $\gamma = \varphi \circ \phi^{-1}$, then $\phi(n_\varphi(x)) = n_\gamma(\phi(x))$.
According to Lemma 1 and Lemma 2, γ is also an automorphism of the unit interval. So n_γ is a strong fuzzy negation. Thus
$I_{QL}(x, y) = \phi^{-1}(S_{\mathbf{L}}(n_\gamma(\phi(x)), T_M(\phi(x), \phi(y))))$
$= \phi^{-1}(I'_{QL}(\phi(x), \phi(y)))$, where
$I'_{QL}(x, y) = S_{\mathbf{L}}(n_\gamma(x), T_M(x, y))$. According to Theorem 5, I'_{QL} satisfies I1 iff $n_\gamma(x) \geq n_0(x)$, for all $x \in [0, 1]$. And according to Lemma 5, I_{QL} satisfies I1 iff I'_{QL} satisfies I1. Thus I_{QL} satisfies I1 iff $n_\gamma(x) \geq n_0(x)$, which leads to $n_\gamma(\phi(x)) \geq 1 - \phi(x)$, which means $\phi(n_\varphi(x)) \geq 1 - \phi(x)$, i.e.,
$n_\varphi(x) \geq \phi^{-1}(1 - \phi(x)) = n_\phi(x)$, for all $x \in [0, 1]$. □

Remark 4. According to Example 1 and Remark 3, for the QL-implication I_{QL} defined in Theorem 5 with $n \geq n_0$, if $n = n_0$, then I_{QL} is equivalent to both an S-implication and an R-implication.
On the contrary, we suppose that there exists $x_0 \in]0, 1[$ such that $n(x_0) > n_0(x_0)$. Then there exists y_0 such that $x_0 > y_0 \geq 1 - n(x_0)$, which leads to $n(x_0) + y_0 \geq 1$, which means $I_{QL}(x_0, y_0) = 1$ provided $x_0 > y_0$. Thus I_{QL} does not satisfy I5. Therefore according to Lemma 4, I_{QL} is not an R-implication. But I_{QL} is an S-implication, i.e., $I_{QL}(x, y) = S_{\mathbf{L}}(n(x), y)$.
Similarly, for the QL-implication I_{QL} defined in Corollary 1 with $n_\varphi \geq n_\phi$, if $n_\varphi = n_\phi$, then I_{QL} is equivalent to both an S-implication and an R-implication. If on the contrary $n_\varphi \neq n_\phi$, then I_{QL} is not an R-implication but an S-implication, i.e., $I_{QL}(x, y) = S_{\mathbf{L}\phi}(n_\varphi(x), y)$.

According to Proposition 2, a necessary condition for a QL-implication I_{QL} generated by a continuous t-conorm S, a t-norm T and a strong fuzzy negation n to satisfy I1 is that there exists an automorphism ϕ of the unit interval such that $S = S_{\mathbf{L}\phi}$ and $n \geq n_\phi$. The authors of [12] have done the work for the special case that $n = n_\phi$. Next theorem gives the sufficient and necessary condition for I_{QL} of such case to satisfy I1.

Theorem 6. *([12], Proposition 9) Let ϕ be an automorphism of the unit interval. A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}\phi}$, a t-norm T and the strong fuzzy negation n_ϕ satisfies I1 iff $T_{\phi^{-1}}$ satisfies the Lipschitz condition, i.e., for all $x_1, x_2, y \in [0, 1]$,*

$$x_1 \leq x_2 \Rightarrow T_{\phi^{-1}}(x_2, y) - T_{\phi^{-1}}(x_1, y) \leq x_2 - x_1, \quad (2)$$

There are t-norms sufficient to fulfill the Lipschitz condition (2), here we give examples:

Example 2 According to (1), a Frank t-norm $T^s = T_P$ if $s = 1$ and $T^s = T_{\mathbf{L}}$ if $s = 0$. It has been stated in ([12], Remark 4) that a t-norm T which is a Frank t-norm or an ordinal sum of the non-empty family $\{[a_m, b_m], T_m\}$, where T_m are Frank t-norms, always satisfies the Lipschitz condition (2). Thus T_P, $T_{\mathbf{L}}$ and T_o which is an ordinal sum of the non-empty family $\{[a_m, b_m], T_m\}$, where $T_m = T_P$ or $T_m = T_{\mathbf{L}}$ all satisfy the Lipschitz condition (2). Hence according to Theorem 6, a QL-implication generated by the t-conorm $S_{\mathbf{L}\phi}$, the t-norm $T_{P\phi}$, $T_{\mathbf{L}\phi}$ or $T_{o\phi}$ and the strong fuzzy negation n_ϕ satisfies I1.

Remark 5 Let I_{QL} be a QL-implication generated by the t-conorm $S_{\mathbf{L}\phi}$, a t-norm T that $T_{\phi^{-1}}$ satisfies the Lipschitz condition (2) and the strong fuzzy negation n_ϕ. Then according to ([12], Corollary 1), I_{QL} satisfies I4 iff $T_{\phi^{-1}}$ is a Frank t-norm defined in (1). Thus according to Theorem 1, I_{QL} is also an S-implication as soon as $T_{\phi^{-1}}$ is a Frank t-norm. Moreover, if $T_{\phi^{-1}}$ is not a Frank t-norm, eg., an ordinal sum of the non-empty family $\{[a_m, b_m], T_m\}$, where $T_m = T_P$ or $T_m = T_{\mathbf{L}}$, then I_{QL} does not satisfy I4. Thus according to Lemma 3 and Lemma 4, it is neither an S-implication nor an R-implication.

Besides the QL-implications generated by the t-conorm $S_{\mathbf{L}\phi}$, a t-norm T and the strong fuzzy negation n_ϕ, which we discussed above, there exist other combinations of a t-conorm S, a t-norm T and a strong fuzzy negation n to generate a QL-implication I_{QL} which satisfies I1. It is sufficient but not necessary for n to be n_ϕ while $S = S_{\mathbf{L}\phi}$. Next we discuss the cases that provided both S and T are continuous, what conditions should n fulfill to make I_{QL} satisfy I1. Since T is either T_M, or conjugated to T_P or conjugated to $T_{\mathbf{L}}$, or an ordinal sum the non-empty family $\{[a_m, b_m], T_m\}$ with T_m being conjugated to T_P or $T_{\mathbf{L}}$, we have the next theorems and corollaries. First we consider the cases that $T = T_P$ or T is conjugated to T_P.

Theorem 7. *Let n be a strong fuzzy negation and define a mapping f as $f(x) = \frac{1-n(x)}{x}$, for all $x \in]0,1]$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}}$, the t-norm T_P and n satisfies I1 iff f is increasing.*

Proof. For all x_1, x_2 and $y \in [0,1]$, assume $x_1 < x_2$.
$\Longrightarrow$: In order for I_{QL} to satisfy I1, it is necessary that $I_{QL}(x_2, y) = 1$ implies $I_{QL}(x_1, y) = 1$, namely, $y \geq \frac{1-n(x_2)}{x_2}$ implies $y \geq \frac{1-n(x_1)}{x_1}$. Thus
$f(x_2) = \frac{1-n(x_2)}{x_2} \geq \frac{1-n(x_1)}{x_1} = f(x_1)$, for all $x_1 < x_2$, i.e., f is increasing.
$\Longleftarrow$: If $I_{QL}(x_1, y) = 1$, then it is always greater than $I_{QL}(x_2, y)$. If $I_{QL}(x_2, y) = 1$, then since f is increasing, according to the proof above, $I_{QL}(x_1, y) = 1 = I_{QL}(x_2, y)$. Thus we need only consider the situation that $I_{QL}(x_1, y) = n(x_1) + x_1 y < 1$ and
$I_{QL}(x_2, y) = n(x_2) + x_2 y < 1$, i.e., $x_1 > 0$, $x_2 > 0$, $y < \frac{1-n(x_1)}{x_1}$ and $y < \frac{1-n(x_2)}{x_2}$. Since f is increasing, $\frac{1-n(x_2)}{x_2} \geq \frac{1-n(x_1)}{x_1}$. Thus:
$n(x_1)x_1 - n(x_2)x_1 \geq x_2 - x_2 n(x_1) - x_1 + n(x_1)x_1$,
which leads to $\frac{n(x_1)-n(x_2)}{x_2-x_1} \geq \frac{1-n(x_1)}{x_1} > y$. Therefore $n(x_1) + x_1 y > n(x_2) + x_2 y$, i.e., $I_{QL}(x_1, y) \geq I_{QL}(x_2, y)$. Hence I_{QL} satisfies I1. □

Example 3. Let $\phi(x) = x^2$. Then $n(x) = \sqrt{1-x^2}$ and $f(x) = \frac{1-n(x)}{x} = \frac{1-\sqrt{1-x^2}}{x}$, for all $x \in]0,1]$. Since $\frac{df(x)}{dx} = \frac{1-\sqrt{1-x^2}}{\sqrt{1-x^2}x^2} \geq 0$, f is increasing. Thus I_{QL} defined by $I_{QL}(x, y) = S_{\mathbf{L}}(\sqrt{1-x^2}, xy)$ satisfies I1.

Corollary 2. *Let ϕ, φ and γ be three automorphisms of the unit interval, where $\gamma = \varphi \circ \phi^{-1}$. And define a mapping f as $f(x) = \frac{1-n_\gamma(x)}{x}$, for all $x \in]0,1]$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}^\phi}$, the t-norm $T_{P\phi}$ and n_φ satisfies I1 iff f is increasing.*

Proof. $I_{QL}(x, y) = S_{\mathbf{L}^\phi}(n_\varphi(x), T_{P\phi}(x, y))$
$= \phi^{-1}(S_{\mathbf{L}}(\phi(n_\varphi(x)), T_P(\phi(x), \phi(y))))$.
Since $\gamma = \varphi \circ \phi^{-1}$,
$I_{QL}(x, y) = \phi^{-1}(S_{\mathbf{L}}(n_\gamma(\phi(x)), T_P(\phi(x), \phi(y))))$.
According to Lemma 1 and Lemma 2, γ is also an automorphism of the unit interval. So n_γ is a strong fuzzy negation. Thus
$I_{QL}(x, y) = \phi^{-1}(I'_{QL}(\phi(x), \phi(y)))$, where
$I'_{QL}(x, y) = S_{\mathbf{L}}(n_\gamma(x), T_P(x, y))$. According to Theorem 7, I'_{QL} satisfies I1 iff f is increasing. And according to Lemma 5, I_{QL} satisfies I1 iff I'_{QL} satisfies I1. Thus I_{QL} satisfies I1 iff f is increasing. □

Remark 6. Let I_{QL} be the QL-implication and f be the mapping defined in Theorem 7 with f being increasing. Then $f(x) \leq f(1) = 1$, for all $x \in]0,1]$, which leads to $n(x) \geq n_0(x)$, for all $x \in]0,1]$. Since $n(0) = n_0(0)$, we have $n(x) \geq n_0(x)$, for all $x \in [0,1]$. If $n = n_0$, then according to Example 2 and Remark 5, I_{QL} is also an S-implication. If on the contrary $n \neq n_0$, then consider:
$I_{QL}(x, I_{QL}(y, z))$
$= \min(n(x) + x \cdot \min(n(y) + yz, 1), 1)$,
which is equivalent to:

i) $n(x) + x(n(y) + yz)$, if $n(y) + yz < 1$ and $n(x) + x(n(y) + yz) < 1$;
ii) 1 otherwise,

and $I_{QL}(y, I_{QL}(x, z))$
$= \min(n(y) + y \cdot \min(n(x) + xz, 1), 1)$,
which is equivalent to:

i) $n(y) + y(n(x) + xz)$, if $n(x) + xz < 1$ and $n(y) + y(n(x) + xz) < 1$;
ii) 1 otherwise.

Since n is continuous and since $g(x) = \frac{1-x}{n(x)}$, $x \in [0,1[$ cannot be constant, there exist x_0, $y_0 \in [0,1[$ such that $x_0 \neq y_0$ and $g(x_0) \neq g(y_0)$. Let $g(x_0) > g(y_0)$. Then we have $\frac{\frac{1-n(y_0)}{y_0} - n(x_0)}{x_0} < \frac{\frac{1-n(x_0)}{x_0} - n(y_0)}{y_0}$. Thus there exists z_0 such that
$\frac{\frac{1-n(y_0)}{y_0} - n(x_0)}{x_0} \leq z_0 < \frac{\frac{1-n(x_0)}{x_0} - n(y_0)}{y_0} \leq \frac{1-n(y_0)}{y_0}$.
Therefore x_0, y_0, z_0 satisfy $n(y_0) + y_0 z_0 < 1$ and $n(x_0) + x_0(n(y_0) + y_0 z_0) < 1$ and
$n(y_0) + y_0(n(x_0) + x_0 z_0) \geq 1$, which means $I_{QL}(y_0, I_{QL}(x_0, z_0)) = 1$ while $I_{QL}(x_0, I_{QL}(y_0, z_0)) < 1$. Thus I_{QL} does not satisfy I4. According to Lemma 3 and Lemma 4, I_{QL} is neither an S-implication nor an R-implication.
Similarly, let I_{QL} be the QL-implication and f be the mapping defined in Corollary 2 with f being increasing, we have $n_\varphi(x) \geq n_\phi(x)$, for all $x \in [0,1]$. If $n_\varphi = n_\phi$, then I_{QL} is also an S-implication. If on the contrary $n_\varphi \neq n_\phi$, then I_{QL} is neither an S-implication nor an R-implication.

Theorem 8. *A QL-implication I_{QL} generated by the t-conorm $S_{\mathbf{L}}$, the t-norm $T_{\mathbf{L}}$ and a strong fuzzy negation n satisfies I1 iff $n = n_0$.*

Proof. $\Longleftarrow$: Straightforward from Theorem 6 and Example 2.
$\Longrightarrow$: Take $0 < x_1 < x_2 < 1$. Since $1 - x_2 < 1 - x_1 < 2 - n(x_1) - x_1$, for all $x_1 \in]0,1[$, there exists y_0 such that $1 - x_1 < y_0 < 2 - n(x_1) - x_1$ and $1 - x_2 < y_0$. Thus $I_{QL}(x_1, y_0) = n(x_1) + x_1 + y_0 - 1 < 1$ and
$I_{QL}(x_2, y_0) = \min(n(x_2) + x_2 + y_0 - 1, 1)$. If I_{QL} satisfies I1, then it is necessary that $I_{QL}(x_2, y_0) < 1$, i.e.,

$n(x_2) + x_2 + y_0 - 1 < 1$. Namely, $y_0 < 2 - n(x_1) - x_1$ implies $y_0 < 2 - n(x_2) - x_2$. Thus define f as $f(x) = n(x) + x$, f must be decreasing for all $x \in]0,1[$. Since n is continuous, f is continuous. Thus f must be decreasing for all $x \in [0,1]$. Moreover, we have $f(0) = f(1) = 1$. Therefore $f(x) = 1$, for all $x \in [0,1]$, which means $n(x) = 1 - x$, for all $x \in [0,1]$, i.e., $n = n_0$. □

Corollary 3. *Let ϕ and φ be two automorphisms of the unit interval. Then a QL-implication I_{QL} generated by the t-conorm $S_{\L^\phi}$, the t-norm $T_{\L^\phi}$ and a strong fuzzy negation n_φ satisfies I1 iff $n_\varphi = n_\phi$.*

Proof. $I_{QL}(x,y) = S_{\L^\phi}(n_\varphi(x), T_{\L^\phi}(x,y))$
$= \phi^{-1}(S_{\L}(\phi(n_\varphi(x)), T_{\L}(\phi(x), \phi(y))))$.
Putting $\gamma = \varphi \circ \phi^{-1}$, then $\phi(n_\varphi(x)) = n_\gamma(\phi(x))$.
According to Lemma 1 and Lemma 2, γ is also an automorphism of the unit interval. So n_γ is a strong fuzzy negation. Thus $I_{QL}(x,y)$
$= \phi^{-1}(S_{\L}(n_\gamma(\phi(x)), T_{\L}(\phi(x), \phi(y))))$
$= \phi^{-1}(I'_{QL}(\phi(x), \phi(y)))$,
where $I'_{QL}(x,y) = S_{\L}(n_\gamma(x), T_{\L}(x,y))$.
According to Theorem 8, I'_{QL} satisfies I1 iff $n_\gamma = n_0$. And according to Lemma 5, I_{QL} satisfies I1 iff I'_{QL} satisfies I1. Thus I_{QL} satisfies I1 iff $n_\gamma = n_0$. So $n_\gamma(x) = 1 - x$, $n_\gamma(\phi(x)) = 1 - \phi(x)$, being equivalent to $\phi(n_\varphi(x)) = 1 - \phi(x)$, which means $n_\varphi(x) = \phi^{-1}(1 - \phi(x))$, for all $x \in [0,1]$. Hence $n_\varphi = n_\phi$. □

Remark 7. According to Example 2 and Remark 5, the QL-implication defined in Theorem 8 with $n = n_0$ and the QL-implication defined in Corollary 3 with $n_\varphi = n_\phi$ are equivalent to S-implications.

Next we consider the t-norm T which constructs the QL-implication to be an ordinal sum of the non-empty family $\{[a_m, b_m], T_m\}$, where $T_m = T_P$ or $T_m = T_{\L}$. Because of space limitation, we omit the proofs of the theorems and corollaries below.

Theorem 9. *Let $\{[a_m, b_m]\}$ be a non-empty family of non-overlapping, closed, proper subintervals of $[0,1]$ and T_o be an ordinal sum of $\{[a_m, b_m], T_m\}$, where $T_m = T_P$. Moreover, let n be a strong fuzzy negation and define $f_m(x) = \frac{1-n(x)-a_m}{x-a_m}$, for all m and $x \in]a_m, b_m]$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\L}$, T_o and n satisfies I1 iff $n(x) \geq n_0(x)$, for all $x \in [0,1]$ and f_m is increasing, for all m.*

Theorem 10. *Let $\{[a_m, b_m]\}$ be a non-empty family of non-overlapping, closed, proper subintervals of $[0,1]$ and T_o be an ordinal sum of $\{[a_m, b_m], T_m\}$, where $T_m = T_{\L}$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\L}$, T_o and a strong fuzzy negation n satisfies I1 iff*

i) $n(x) = n_0(x)$, if $x \in [a_m, b_m]$ and

ii) $n(x) \geq n_0(x)$, otherwise.

Synthesize Theorem 9 and Theorem 10, we have the next corollary.

Corollary 4. *Let $\{[a_i, b_i]\}$ and $\{[a_j, b_j]\}$ be two non-empty families of non-overlapping, closed, proper subintervals of $[0,1]$ and $\{[a_m, b_m]\} = \{[a_i, b_i]\} \cup \{[a_j, b_j]\}$. T_o is an ordinal sum of $\{[a_m, b_m], T_m\}$, which is defined as:*

$$\begin{cases} a_i + (b_i - a_i) T_P(\frac{x-a_i}{b_i-a_i}, \frac{y-a_i}{b_i-a_i}), & (x,y) \in [a_i, b_i]^2 \\ a_j + (b_j - a_j) T_{\L}(\frac{x-a_j}{b_j-a_j}, \frac{y-a_j}{b_j-a_j}), & (x,y) \in [a_j, b_j]^2 \\ T_M(x,y), & \text{otherwise} \end{cases} \tag{3}$$

Moreover, let n be a strong fuzzy negation and define $f_i(x) = \frac{1-n(x)-a_i}{x-a_i}$, for all i and $x \in]a_i, b_i]$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\L}$, T_o and n satisfies I1 iff

i) f_i is increasing, for all i, and

ii) $n(x) = n_0(x)$, for all $x \in [a_j, b_j]$, and

iii) $n(x) \geq n_0(x)$, for all $x \notin [a_i, b_i]$ and all $x \notin [a_j, b_j]$.

For the QL-implication being conjugated to the one defined in Corollary 4, we have the next corollary.

Corollary 5. *Let T_o be a t-norm defined by (3), n be a strong fuzzy negation and ϕ be an automorphism of the unit interval. Define $f_i(x) = \frac{1-\phi(n(\phi^{-1}(x)))-a_i}{x-a_i}$, for all i and $x \in]a_i, b_i]$. Then a QL-implication I_{QL} generated by the t-conorm $S_{\L^\phi}$, $T_{o\phi}$ and n satisfies I1 iff*

i) f_i is increasing, for all i, and

ii) $n(x) = n_\phi(x)$, for all $x \in [a_j, b_j]$, and

iii) $n(x) \geq n_\phi(x)$, for all $x \notin [a_i, b_i]$ and all $x \notin [a_j, b_j]$.

Proof. Suppose $n'(x) = \phi(n(\phi^{-1}(x)))$. According to Lemma 5, I_{QL} satisfies I1 iff $I_{QL\phi^{-1}}$, which is expressed as $I_{QL\phi^{-1}}(x,y) = S_{\L}(n'(x), T_o(x,y))$, satisfies I1. According to Corollary 5, $I_{QL\phi^{-1}}$ satisfies I1 iff

i) f'_i is increasing, for all i, and

ii) $n'(x) = n_0(x)$, for all $x \in [a_j, b_j]$, and

iii) $n^{'}(x) \geq n_0(x)$, for all $x \notin [a_i, b_i]$ and all $x \notin [a_j, b_j]$,

where $f_i^{'}(x) = \frac{1-n^{'}(x)-a_i}{x-a_i}$, for all $x \in]a_i, b_i]$. The three conditions are equivalent to

i) f_i is increasing, for all i, and

ii) $n(x) = n_\phi(x)$, for all $x \in [a_j, b_j]$, and

iii) $n(x) \geq n_\phi(x)$, for all $x \notin [a_i, b_i]$ and all $x \notin [a_j, b_j]$.

□

4 Conclusions

In this paper, we have studied the QL-implications generated by a t-conorm, a t-norm and a strong fuzzy negation that satisfy the properties which are required to obtain a suitable conclusion in fuzzy inference. Especially the first place antitonicity (property I1 in the Introduction) of QL-implications has been studied. Theorem 1 to Theorem 4 state the general relationship between a QL-implication satisfying the first place antitonicity and the other properties. Moreover, Theorem 5 to Theorem 10 together with Corollary 1 to Corollary 5 state sufficient and necessary conditions for the QL-implications generated by different combinations of a t-conorm, a t-norm and a strong fuzzy negation to satisfy the first place antitonicity. Whether the QL-implications which satisfy the first place antitonicity are equivalent to S- or R- implications have been illustrated in Remarks 4 to 7.

References

[1] M. Baczynski J. Drewniak, Conjugacy classes of fuzzy implication, in: B. Reusch (Eds.), *Computational intelligence: Theory and Applications*. Springer-Verlag, Berlin (Lecture Notes in Computer Science, Vol. 1625), 1999, pp. 287-298.

[2] M. Baczynski, Residual implications revisited. Notes on the Smets-Magrez Theorem, Fuzzy Sets and Systems 145 (2004) 267-277.

[3] M. Baczynski, J. Balasubramaniam, On the characterizations of (S, N)-implications, Fuzzy Sets and Systems, to appear.

[4] H. Bustince, P. Burillo, F. Soria, Automorphisms, negations and implication operators, Fuzzy Sets and Systems 134 (2003) 209-229.

[5] D. Dubois, H Prade, Fuzzy sets in approximate reasoning, part 1: inference with possibility distributions, Fuzzy Sets and Systems 40 (1991) 143-202.

[6] D. Dubois, H Prade, Fuzzy sets in approximate reasoning, part 2: logical approaches, Fuzzy Sets and Systems 40 (1991) 202-244.

[7] D. Dubois, H Prade, Gradual inference rules in approximate reasoning, Information Sciences 61 (1992) 103-122.

[8] J.C. Fodor, On fuzzy implication operators, Fuzzy Sets and Systems 42 (1991) 293-300.

[9] J.C. Fodor, Contrapositive symmetry of fuzzy implications, Fuzzy Sets and Systems 69 (1995) 141-156.

[10] J.C. Fodor, M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support, Kluwer Academic Publishers, Netherlands, 1994.

[11] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic, Theory and Applications, Prentice Hall, USA, 1995.

[12] M. Mas, M. Monserrat, J. Torrens, QL-implications versus D-implications, Kybernetika 42 (2006) 351-366.

[13] D. Ruan, E.E. Kerre, Fuzzy implication operators and generalized fuzzy method of cases, Fuzzy Sets and Systems 54 (1993) 23-37.

[14] D. Ruan, E.E. Kerre, Fuzzy IF-THEN Rules in Computational Intelligence: Theory and Applications, Kluwer Academic Publishers, Boston, 2000.

[15] P. Smets, P. Magrez, Implication in fuzzy logic, Internat. J. Approx. Reason. 1 (1987) 327-347.

[16] E. Trillas, C.D. Campo, S. Cubillo, When QM-operators are implication functions and conditional fuzzy relations, International Journal of Intelligent Systems 15 (2000) 647-655.

[17] R.R. Yager, On global requirements for implication operators in fuzzy modus ponens, Fuzzy Sets and Systems 106 (1999) 3-10.

Session 2

Applications 1

ON CLUSTER ANALYSIS BASED ON FUZZY RELATIONS BETWEEN SPATIAL DATA

Efendi NASIBOV
Dept. of Statistics, Dokuz Eylul University,
Buca 35160 Izmir, TURKEY
efendi.nasibov@deu.edu.tr

Gozde ULUTAGAY
Dept. of Statistics, Dokuz Eylul University,
Buca 35160 Izmir, TURKEY
gozde.ulutagay@ogr.deu.edu.tr

Abstract

Methods like DBSCAN are widely used in the analysis of spatial data. These methods are based on the neighborhood relations which use distance between points. However, these neighborhood relations consider to have at least a certain number of neighbors within a definite boundary. In this proposed work such a neighborhood analysis is done by using the benefits of fuzzy sets theory. Usage of fuzzy logic gives more sensitive and realistic results. In this paper, Fuzzy Joint Points (FJP) based on this theory is handled and some theoretical properties used in neighborhood analysis are investigated.

Keywords: Cluster analysis, fuzzy neighborhood relation, Fuzzy Joint Points (FJP).

1 Introduction

Cluster analysis has an important role in statistical data analysis. The main objective of clustering is to facilitate the analysis process by constructing similar objects in a cluster. Clustering methods can be divided into two groups such as hierarchical and prototype-based [3]. In hierarchical clustering, the remoteness of elements is the cornerstone. First of all, closer elements are put into same cluster, and in the next step, elements a little bit far away from the previous ones are put in another cluster, etc. In prototype-based methods, however, prototypes which have common features of some certain classes are formed and then the elements are taken into these classes with respect to the proximity degrees to the prototypes. Namely, in such a situation, not the remoteness of above-mentioned elements from each other, but their remoteness from the prototypes is considered. Some examples of these methods are k-means, k-medoids and FCM [1, 3, 4]. Besides, single-linkage (SLINK), complete-linkage (CLINK), DBSCAN, OPTICS and FJP are some examples of hierarchical methods [2, 3, 5, 6, 7, 8]. In DBSCAN-like methods, in order to determine the core points of clusters or noise points, classical neighborhood density analysis is done. Thus, a point is conceived as a core point, if the number of points in a certain radius is larger than a specified threshold. However, in FJP-based methods, fuzzy neighborhood analysis that provides more sensitive results is used [7].

In this study, some researches are executed on fuzzy neighborhood analysis based on distance between points and it is showed that such a method can provide more sensitive and realistic results in comparison with non-fuzzy methods.

2 Fuzzy Neighborhood Based on Distance Between Points

Let us consider the data set $X = \{x_1, x_2, \ldots, x_n\}$. The purpose of clustering is to determine sets $X^1, X^2, \ldots, X^k$ that constitute X which provides $\forall i, j : i \neq j \Rightarrow X^i \cap X^j = \emptyset$ and $\bigcup_{i=1}^{k} X^i = X$.

Let $F(E^p)$ indicate the set of all p-dimensional fuzzy sets in the space E^p. Let $\mu_A : E^p \to [0,1]$ represent the membership function of the fuzzy set $A \in F(E^p)$.

Definition 1. A conical fuzzy point $A = (a, R) \in F(E^p)$ of the space E^p is a fuzzy set with membership function

$$\mu_A(x) = \begin{cases} 1 - \frac{d(x,a)}{R} & \text{if } d(x,a) \leq R \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

where $a \in E^p$ is the center of fuzzy point A, and $R \in E^1$ is the radius of its support, *supp* A.

α-level set of conical fuzzy point $A = (a, R)$ is calculated as follows:

$$\begin{aligned} A_\alpha &= \{x \in E^p \,|\, \mu_A(x) \geq \alpha\} \\ &= \{x \in E^p \,|\, d(x,a) \leq R \cdot (1 - \alpha)\}. \end{aligned}$$

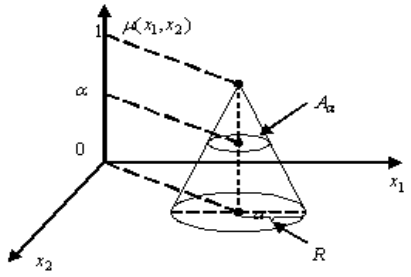


Figure 1: $A = (a, R) \in F(E^2)$ conical fuzzy point.

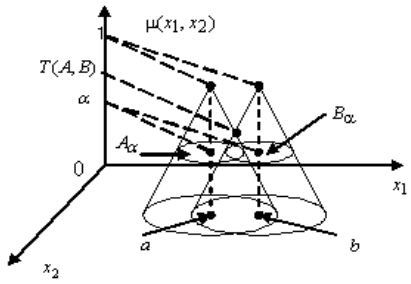


Figure 2: α-neighbor points $A = (a, R)$ and $B = (b, R)$ in fuzzy space E^2.

The equivalent of the fuzzy conical point $A = (a, R) \in F(E^1)$ in the space E^1 is triangular and symmetric fuzzy number $A = (a, R, R)$.

Let $A = (a, R)$ and $B = (b, R)$, be fuzzy sets from the set $X \subset F(E^p)$. Let's denote a fuzzy similarity relation $T : X \times X \to [0, 1]$ on the set X as follows:

$$T(A, B) = 1 - \frac{d(a, b)}{2R}, \tag{2}$$

where $a \in E^p$ and $b \in E^p$ are the centers of the fuzzy points A and B respectively.

(2) can can be written as follows:

$$d(a, b) = 2R \cdot (1 - T(A, B)). \tag{3}$$

It is obvious that the relation T is reflexive, thus $\forall A \in X$, $T(A, A) = 1$ is satisfied.

Lemma 2. *The fuzzy points $A = (a, R)$ and $B = (b, R)$ are called fuzzy α-neighbor points if and only if*

$$d(a, b) \leq 2R \cdot (1 - \alpha) \tag{4}$$

is satisfied.

Definition 3. Let A and B be fuzzy points on the set $X \subset F(E^p)$. If

$$T(A, B) \geq \alpha, \tag{5}$$

is satisfied for fixed $\alpha \in (0, 1]$, then the points A and B are called the α-neighbor fuzzy points and it is denoted by $A \sim_\alpha B$ (Fig.2).

Definition 4. If there is a sequence of α-neighbor fuzzy points $C^1, \ldots, C^k,, k \geq 0$, for fixed $\alpha \in (0, 1]$, between the points A and B, i.e.

$$A \sim_\alpha C^1, C^1 \sim_\alpha C^2, \ldots, C^{k-1} \sim_\alpha C^k \quad and \quad C^k \sim_\alpha B,$$

then the fuzzy points A and B are called α-joint fuzzy points.

Definition 5. Let $X \subset F(E^p)$ be a set of fuzzy points. If the fuzzy points A and B are α-joint for $\alpha \in (0, 1]$ and $\forall A, B \in X$, then the set X is called fuzzy α-joint set.

Suppose that the distance between the level sets A_α and B_α is defined as follows:

$$d(A_\alpha, B_\alpha) = min\{d(x, y) \mid x \in A_\alpha, y \in B_\alpha\}.$$

Let the relation $\hat{T} : X \times X \to [0, 1]$ be the transitive closure of relation $T : X \times X \to [0, 1]$ by using max-min composition.

Lemma 6. *The fuzzy points A and B are called α-neighbor points if and only if*

$$A_\alpha \cap B_\alpha \neq \emptyset \tag{6}$$

is satisfied.

3 Fuzzy Neighborhood Membership Functions

Let us deal with two sets given in Fig. 3 within the same radius ε_1, with equal number of elements, but with different densities. A method like DBSCAN that uses classical neighborhood relation, will consider the points x_1 and x_2 as core points and the cardinality of these sets will be equal within radius ε_1.

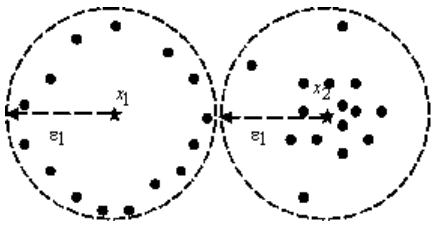


Figure 3: Points x_1 and x_2 are similar according to crisp neighborhood cardinality, but dissimilar according to fuzzy neighborhood cardinality.

Nevertheless, if these points are evaluated according to fuzzy neighborhood cardinality as done in FJP

method, since the proximity of the points around them will be taken into account, it is obvious that they will be different from each other. It can be said that analysis realized with such a point of view will produce more sensitive and realistic results.

As seen in Fig.4, since the classical neighborhood relation is used in DBSCAN method, although the point y_1 is closer to the point x, the neighborhood degree of the points y_1 and y_2 within radius ε_1 are equal with respect to the point x.

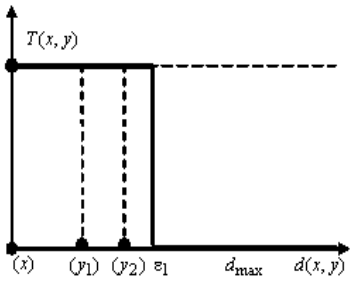


Figure 4: Neighborhood relation used in DBSCAN method.

However in FJP method, since the fuzzy neighborhood relation is used by taking into account the proximity of the points y_1 and y_2 to the point x is considered within radius ε_1, a distinct difference is seen between the neighborhood degrees α_1 and α_2 of the points y_1 and y_2 respectively to the point x (Fig.5).

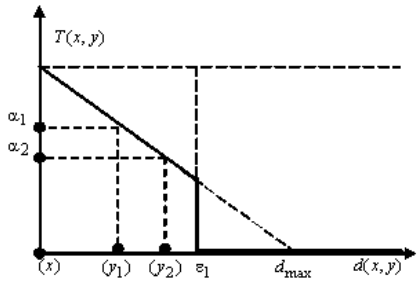


Figure 5: Neighborhood relation used in FJP method.

When dealing with fuzzy neighborhood analysis, different membership functions can be used in order to state neighborhood degrees. Also, the sensitivity of analysis can be increased by using different functions in calculation of the relation T (Fig. 5-6).

As seen from the figures, different sensitivity can be reached by using different membership functions in neighborhood analysis. For instance, the membership function in Fig. 4 handles the points within radius ε_1 as identical whereas there is an obvious distinction in Fig. 5. The difference becomes more evident in Fig.

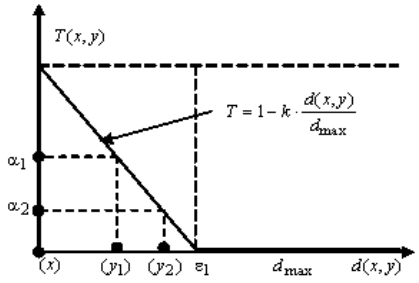


Figure 6: Linear neighborhood relation within an ε_1 radius.

6. However, the sensitivity is the same on near and far distance of the reference point. But in Fig. 7, the sensitivity differs exponentially.

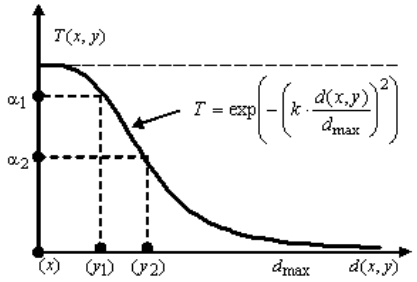


Figure 7: The effect of using different membership functions for a fuzzy point (exponential).

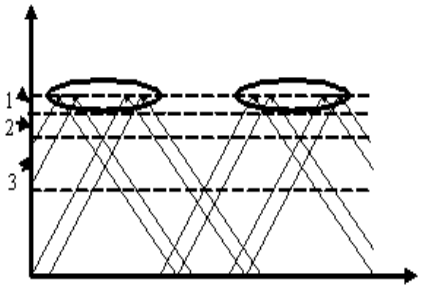


Figure 8: The effect of using different membership functions for a fuzzy point (linear).

Note that, in definition of the fuzzy point, if different nonlinear functions are used, different results can be found by the FJP algorithm. For example, consider a data set with 8 elements. If the membership function is linear as in Fig. 8, the widest change interval in which the α-parameter does not affect the number of clusters is found as the interval number 3 and according to the working principle of the algorithm, such a situation is appropriate for two clusters.

However, when the membership function of the fuzzy points looks like bell-shaped as in Fig. 9, the widest change interval found as number 2 and such a situation is suitable for four clusters. Thus, FJP algorithm finds two clusters if a membership function of a fuzzy point as in Fig. 8 is used whereas it detects four clusters if a membership function of a fuzzy point as in Fig. 9 is used.

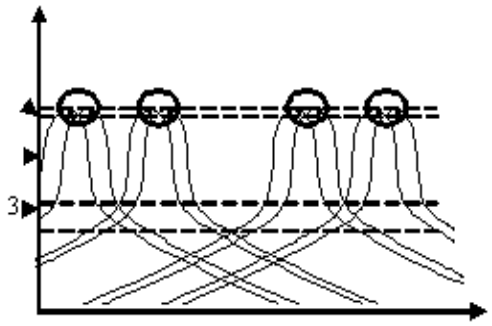


Figure 9: The effect of using different membership functions for a fuzzy point (bell-shaped).

Note that in FJP algorithm, the lowest layer that is suitable for one cluster is not taken into consideration.

4 Conclusion

In this study, the Fuzzy Joint Points (FJP) method is handled and fuzzy neighbor relations between points are investigated. FJP method may be more effectively used in the analysis of statistical databases with spatial data. As a result of our researches, since FJP method uses fuzzy logic, it produces more sensitive realistic results in comparison of methods based on classical neighborhood relation.

Also in our study, some theoretical results are given for the analysis of the cluster structure. Moreover, it is mentioned that using different membership functions for the neighborhood relation results in different partitions.

We think that FJP method may be more effective in areas such as mapping of national or regional geographical information, medical tomography, ultrasound or X-ray analysis, etc.

Acknowledgement

This study is supported as a research project No.106T312 by Scientific and Technological Research Institute of Turkey (TUBITAK).

References

[1] J. C. Bezdek, Fuzzy Mathematics in Pattern Classification, PhD Thesis, Cornell Univ., New York, 1973.

[2] M. Ester, H. Kriegel, J. Sander, X. Xu, A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, In Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining (1996) pp.226–231.

[3] J. Han, M. Kamber, Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, 2001.

[4] J. MacQueen, Some Methods for Classification and Analysis of Multivariate Observations, In Proc. 5th Berkeley Symp. Math. Statist. Prob. 1 (1967) pp.281–297.

[5] E. N. Nasibov, An Alternative Fuzzy-Hierarchical Approach to Cluster Analysis, In Proc. 7th Int. Conf. on Application of Fuzzy Systems and Soft Computing, Germany, (2006) pp.113–123.

[6] E. N. Nasibov, G. Ulutagay, A New Approach to Clustering Problem Using the Fuzzy Joint Points Method, Automatic Control and Computer Sciences,39 (2005) No.6, pp.8-17.

[7] E. N. Nasibov, G. Ulutagay, On the Fuzzy Joint Points Method for Fuzzy Clustering Problem, Automatic Control and Computer Sciences, 40 (2006) No.5, pp.33-44.

[8] E. N. Nasibov, G. Ulutagay, FJP: A New Hierarchical Method for Fuzzy Clustering, In Proc. 3rd Int. Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control (2005) pp.212-220.

[9] W. Pedrycz, F. Gomide, An Introduction to Fuzzy Sets, Massachusetts Institute, 1998.

[10] J. Sander, M. Ester, H. P. Kriegel, X. Xu, Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications, Data Mining and Knowledge Discovery 2 (1998), pp.169-194.

Using Supervised Fuzzy Clustering and CWT for Ventricular Late Potentials (VLPs) Detection in High-Resolution ECG Signal

Gholamreza Salimi Khorshidi
School of cognitive sciences, Institute for studies in theoretical physics and mathematics, Tehran, Iran
gsalimi@bme.aut.ac.ir

Ayyoub Jaafari
Faculty of biomedical engineering, Amirkabir University of technology (Tehran Polytechnic), Tehran, Iran
ajafari@bme.aut.ac.ir

Amir Pourmorteza
Control And Intelligent Processing Center Of Excellence, Electrical And Computer Engineering Department, University Of Tehran, Tehran, Iran
amir.pourmorteza@gmail.com

Ali Motie Nasrabadi
Group of biomedical engineering, Faculty of Engineering, Shahed University, Tehran, Iran
nasrabadi@shahed.ac.ir

Mohammadreza Hashemi Golpayegani
Faculty of biomedical engineering, Amirkabir University of technology, Tehran, Iran
hashemi_g@morva.net

Abstract

Ventricular Late Potentials (VLPs) are low-amplitude, high-frequency components that appear at the end of the QRS complex of a High-Resolution ECG (HRECG) record. VLPs are clinically useful for identifying post-MI (Myocardial Infarction) patients prone to Ventricular Tachycardia (VT) and Sudden Cardiac Death (SCD). In this paper, the Continuous Wavelet Transform (CWT) and a supervised fuzzy clustering algorithm are used together to detect VLP. The terminal part of the QRS complex in the Vector Magnitude (VM) waveform is processed with the CWT to extract a feature vector. Resulting time-frequency representation is subdivided into several sub bands, and the sum of the squared decomposition coefficients is computed in each region. Finally, a supervised Fuzzy clustering method, trained by an appropriate set of these feature vectors, is applied to this data in order to identify VLP.

Keywords: ECG, VLP, Fuzzy Clustering, CWT

1 Introduction

Most of the Sudden Cardiac Deaths (SCDs) due to cardiac diseases are thought to be initiated by Ventricular Tachycardia (VT) which is one of the most serious types of cardiac arrhythmia [11]. As the appearance of Ventricular Late Potentials (VLPs) is associated with VT, there is a clinical interest in detection of these signals as a non-invasive diagnosis for post-MI (Myocardial Infarction) patients prone to VT. VLPs are low-amplitude, high-frequency signals which appear at the end of the QRS complex and arise as a result of the late depolarization of damaged myocardium. Because of their very low amplitudes and the noise overlay, VLPs are obscure in a standard electrocardiogram (ECG) but they can be detected by a High-Resolution ECG (HRECG) record acquired using three orthogonal XYZ leads with a minimum sampling frequency of 1000Hz and a resolution of 12 bits [1,9,11].

There are various techniques to improve signal-to-noise ratio (SNR) in VLP analysis. Typically, several heart beats (200-300) are averaged to suppress the background noise and form the Signal Averaged ECG (SAECG). In this paper, a new fuzzy model structure is proposed for the VLP classification problem in which each rule can represent more than one class with different probabilities. Typical fuzzy classifiers consist of interpretable if-then rules with fuzzy antecedents and class labels in the consequent part. The antecedents (if-parts) of the rules partitions the input space into a number of fuzzy regions by means of fuzzy sets, while the consequents (then-parts) describe the output of the classifier in these regions.

Fuzzy logic improves rule-based classifiers by allowing the use of overlapping class definitions and improves the interpretability of the results by providing more insights into the decision making process. The automatic determination of compact fuzzy classifiers rules from data has been approached by several different techniques like; neuro-fuzzy methods, genetic-algorithm (GA) based rule selection, and fuzzy clustering in combination with GA-optimization. Generally, the bottleneck of the data-driven identification of fuzzy systems is the structure identification that requires nonlinear optimization. Thus for high-dimensional problems, the initialization of the fuzzy model becomes very significant.

Common initialization methods such as grid-type partitioning [26] and rule generation on extreme initialization, result in complex and non-interpretable initial models and the rule-base simplification and reduction steps become computationally demanding. To avoid these problems, fuzzy clustering algorithms were proposed [37]. However, the obtained membership values have to be projected onto the input variables and approximated by parameterized membership functions that deteriorate the performance of the classifier. This decomposition error can be reduced by using eigenvector projection [28], but the obtained linearly transformed input variables do not allow the interpretation of the model. To avoid the projection error and maintain the interpretability of the model, the proposed approach is based on the Gath-Geva (GG) clustering algorithm [22] instead of the widely used Gustafson-Kessel (GK) algorithm [38], because the simplified version of GG clustering allows the direct identification of fuzzy models with exponential membership functions [25]. Neither GG nor GK algorithm utilizes the class labels. Hence, they give suboptimal result if the obtained clusters are directly used to formulate a classical fuzzy classifier.

Hence, there is a need for fine-tuning of the model. This GA or gradient-based fine-tuning, however, can result in overfitting and thus poor generalization of the identified model. Unfortunately, the severe computational requirements of these approaches limit their applicability as a rapid model-development tool. This paper focuses on the design of interpretable fuzzy rule-based classifiers from data with low-human intervention and low-computational complexity. Hence, a new modeling scheme is introduced based only on fuzzy clustering. The proposed algorithm is similar to the Multi-Prototype Classifier technique [18, 33]. The main difference of this approach is that each cluster represents different classes, and the number of clusters used to approximate a given class has to be determined manually, while the proposed approach does not suffer from such problems. Generally, there is a very large set of possible features to compose feature vectors of classifiers. As it is ideal that the training dataset's size should increase exponentially with the feature vector size, it is desired to choose a minimal subset among it. Some generic tips to choose a good feature set include the facts that they should discriminate as much as possible the pattern classes and they should not be correlated or redundant.

2 Simson Method

The conventional time-domain method of VLP detection, developed by Simson [8], is based on feature extraction from the filtered SAECG [8, 9, 13]. Simson's method employs a high-pass filter (cutoff frequency of 25 or 40Hz) to attenuate low-frequency components of averaged XYZ signals (SAECG). To avoid the filter ringing effect in the terminal parts of QRS complex, Simson proposed a bi-directional four-pole Butterworth high-pass filter [8]. After high-pass filtering of the averaged XYZ signals, these signals are combined into a Vector Magnitude (VM) waveform defined by

$$VM = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

After estimating the onset and offset of the filtered QRS complex (the QRS complex in the VM signal), three conventional time-domain features can be measured to detect VLPs [1, 11, 12, 13]:

- QRS_T: Duration of the filtered QRS complex (from the onset to the offset)
- $D_{40:}$ Low-amplitude signal duration (from the offset backward to the point where VM reaches the 40μV)
- $V_{40:}$Root-mean-square value of the last 40ms of the filtered QRS (showed in Fig1)

The criteria to define a VLP positive test are QRS_T>114ms, D_{40}>38ms and V_{40}<20μV [11]. In Fig.1, a plot of a typical filtered QRS complex and the definition of the conventional time-domain features, introduced above, can be viewed.

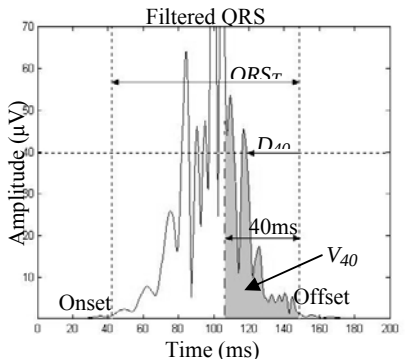


Fig.1 A typical filtered QRS complex and the definition of three features: QRS_T, D_{40} and V_{40}.

3 Feature Extraction

The adopted feature extraction method for VLP consists of five steps:

- Averaging XYZ leads to improve the SNR
- Bidirectional Butterworth filtering of averaged *XYZ* signals consisting of a 4th order high-pass and a 5th order low-pass filters with cutoff frequencies of 40Hz and 250 Hz respectively.
- Combination of the filtered averaged signals into a *VM* waveform using equation (1).
- Applying the CWT to the terminal part of the QRS complex in the *VM* signal.
- Feature extraction from the resulted time-scale plot.

In recent years, the wavelet analysis has been used widely in biomedical researches [2, 5, 7, and 14]. The wavelet transform is a linear time-frequency transform which is based on decomposition of a signal using a set of basis functions. These basis functions are scaled and shifted versions of a prototype mother wavelet [4, 11]. The CWT produces a time-frequency representation of the signal which is a function of time τ and scale a, like CWT (τ, a) named wavelet decomposition coefficient. The scale can be considered as the inverse of the frequency [9].

The smaller scales bring about a higher resolution in time which is useful to detect VLPs as high-frequency, short-duration signals. In this study, the CWT is adopted, using the MATLAB wavelet toolbox, to the last 40ms of the QRS complex in the *VM* signal after estimating the offset point. To improve the robustness of the method to the error in QRS offset detection, a 8ms right shift is considered for the estimated offset [13] so that the CWT is applied to 48ms interval of the *VM* waveform, with scale range of 1-8 and the Morlet wavelet as the mother wavelet [6]. Then, the resulted time-scale representation is subdivided into 81 regions [6] (nine subdivisions on both the time and scale axes). Finally, the sum of the squared wavelet decomposition coefficients is computed in each region to form a feature vector composed of the 81 elements.

4 Supervised Fuzzy Clustering

The objective of clustering is to partition the identification data Z into R clusters. This means, each observation consists of input and output variables grouped into a row vector $Z_k = [x^T_k\ y_k]$, where the k subscript denotes the $k = 1\ldots N$th row of the pattern matrix of Z. The fuzzy partition is represented by the $U= [\mu_{i,k}]_{R*N}$ matrix, where the $\mu_{i,k}$ element of the matrix represents the degree of membership, how the Z_k observation is in the cluster $i = 1\ldots R$. The clustering is based on the minimization of the sum of weighted $D^2_{i,k}$ squared distances between the data points and the η_i cluster prototypes that contains the parameters of the clusters.

$$J(Z,U,\eta)=\sum_{i=1}^{R}\sum_{k=1}^{N}(\mu_{i,k})^m D^2(z_k,r_i) \qquad (3)$$

In which, m is the fuzzy weighting exponent that determines the fuzziness of the resulting clusters. Usually, the chosen value for m will be m = 2. Classical fuzzy clustering algorithms are used to estimate the distribution of the data. Hence, they do not utilize the class label of each data point available for the identification. Furthermore, the obtained clusters cannot be directly used to build the classifier. In the following a new cluster prototype and the related distance measure will be introduced that allows the direct supervised identification of fuzzy classifiers. As the clusters are used to obtain the parameters of the fuzzy classifier, the distance measure is defined similarly to the distance measure of the Bayes classifier:

$$\frac{1}{D^2_{i,k}(z_k,r_i)}=P(r_i)\prod_{j=1}^{n}\exp\left(\frac{1}{2}\frac{(x_{j,k}-v_{i,j})^2}{\sigma^2_{i,j}}\right)P(c_j=y_k|r_i) \qquad (4)$$

This distance measure consists of two terms; the first term is based on the geometrical distance between the v_i cluster centers and the x_k observation vector, while

the second is based on the probability that the r_i-th cluster describes the density of the class of the k-th data, P ($c_j = y_k \mid r_i$) It is interesting to note, that this distance measure only slightly differs from the unsupervised Gath–Geva clustering algorithm which can also be interpreted in a probabilistic framework [21]. However, the novelty of the proposed approach is the second term, which allows the use of class labels. To get a fuzzy partitioning space, the membership values have to satisfy the following conditions:

$$U \in R^{C\times N},\ \mu_{i,k} \in [0,1],\ \forall i,k;$$
$$\sum_{i=1}^{R}\mu_{i,k} = 1,\ \forall k;\ 0 < \sum_{k=1}^{N}\mu_{i,k} < N,\ \forall i \qquad (5)$$

The minimization of the (6) functional represents a non-linear optimization problem that is subject to constraints defined by (5) and can be solved by using a variety of available methods. The most popular method, is the alternating optimization (AO), which consists of the application of Picard iteration through the first-order conditions for the stationary points of (6), which can be found by adjoining the constraints (5) to J by means of LaGrange multipliers [24] and setting the gradients of J to zero, with respect to Z, U, η and λ.

$$\bar{J}(Z,U,\eta,\lambda) =$$
$$\sum_{i=1}^{R}\sum_{k=1}^{N}(\mu_{i,k})^m D^2(z_k,r_i) + \sum_{k=1}^{N}\lambda_k\left(\sum_{j=1}^{R}\mu_{i,k} - 1\right) \qquad (6)$$

Hence, similarly to update the Gath-Geva clustering algorithm equations, the following algorithm will result in a solution that satisfies the (6) constraints.

Initialization: Given a set of data Z specify R, choose a termination tolerance ε>0. Initialize the U= [$\mu_{i,k}$] RεN partition matrix randomly, where $\mu_{i,k}$ denotes the membership that the Z_k data is generated by the i'th cluster. Repeat for l = 1, 2….

Step1: Calculate the parameters of the clusters. Calculate the centers and standard deviation of the Gaussian membership functions (the diagonal elements of the F_i covariance matrices) (7) and estimate the consequent probability parameters (8) and *a priori* probability of the cluster and the weight (impact) of the rules (9).

$$v_i^{(l)} = \frac{\sum_{k=1}^{N}(\mu_{i,k}^{(l-1)})^m x_k}{\sum_{k=1}^{N}(\mu_{i,k}^{(l-1)})^m}$$
$$\sigma_{i,j}^{2(l)} = \frac{\sum_{k=1}^{N}(\mu_{i,k}^{(l-1)})^m (x_{j,k} - v_{j,k})^2}{\sum_{k=1}^{N}(\mu_{i,k}^{(l-1)})^m} \qquad (7)$$

$$p(c_i \mid r_j) = \frac{\sum_{k|y_k=c_i}(\mu_{j,k}^{(l-1)})^m}{\sum_{k=1}^{N}(\mu_{j,k}^{(l-1)})^m}$$
$$1 \le i \le C, 1 \le j \le R \qquad (8)$$

$$P(r_i) = \frac{1}{N}\sum_{k=1}^{N}(\mu_{i,k}^{(l-1)})^m$$
$$w_i = P(r_i)\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi\sigma_{i,j}^2}} \qquad (9)$$

Step2: Compute the distance measure $D^2_{i,k}$ by (4).

Step3: Update the partition matrix

$$\mu_{i,k}^{(l)} = \frac{1}{\sum_{j=1}^{R}(D_{i,k}(z_k,r_i)/D_{j,k}(z_k,r_j))^{2/(m-1)}},$$
$$1 \le i \le R,\ 1 \le k \le N \quad until \left\|U^{(l)} - U^{(l-1)}\right\| < \varepsilon \qquad (10)$$

5 Feature Selection

Using too many input variables may result in difficulties in the interpretability capabilities of the obtained classifier. Hence, selection of the relevant features is usually necessary. In this paper, the modified Fischer interclass separability method is used which is based on statistical properties of the data. The interclass separability criterion is based on the F_B between-class and the F_W within class covariance matrices that sum up to the total covariance of the data $F_T=F_B+F_W$, where:

$$F_W = \sum_{l=1}^{R} P(r_l) F_l \ ,$$

$$F_B = \sum_{l=1}^{R} P(r_l)(v_l - v_0)^T (v_l - v_0),$$

$$v_0 = \sum_{l=1}^{R} P(r_l) v_l \qquad (11)$$

$$J = \frac{\det(F_B)}{\det(F_w)} \qquad (12)$$

The feature interclass seperatibility selection criterion is a trade-off between F_W and F_B (Eq.12). The importance of a feature is measured by leaving out the interested feature and calculating *J* for the reduced covariance matrices. The feature selection is a step-wise procedure, when in every step the least needed feature is deleted from the model [29].

5 Results and Discussion

To evaluate the method used in this study a HRECG database consisting of two groups of signals was selected. The first group contained 50 healthy volunteers' HRECG records acquired by a digital data acquisition system, ML785 PowerLab/8SP, with a sampling frequency of 2000Hz and a 16-bit analog-to-digital converter (ADC). The second group was consisted of semi-simulated HRECG signals with VLPs. In order to simulate each of these signals, three basic simulated waveforms resembling the VLP characteristics were added to *XYZ* leads of a basic HRECG record, a HRECG record without VLP. VLPs are low-amplitude signals (~1-20μV) with short duration (~5-50ms) and broadband spectrum (~40-250Hz) [11]. According to these characteristics, VLPs were simulated as colored Gaussian processes resembling better the real world signals. The basic VLP waveforms were added to the end part of the QRS complex of every heart beat of the *XYZ* leads belonging to the basic HRECG records. The position of the VLPs was varied randomly from beat to beat with respect to the fiducial mark, QRS peak [10, 11]. This HRECG database was divided into a training set, including thirty HRECG signals with VLPs and thirty without, and a test set consisting of 20 records without VLPs and 20 with. For a better training of the neural network and preserving its generalization, the training set was expanded; five sets containing 300 heart beats were selected from every HRECG record of training set randomly. Because of the fact that every HRECG record had at least 350 heart beats, the beat selection was done without replacement for each set. Therefore, an expanded training set consisting of 300 patterns was obtained. The performance of the VLP detection method was measured using conventional criteria i.e. the accuracy *ACC*, sensitivity *SE*, and specificity *SP* defined by

$$ACC = 100 \times (TP + TN) / N$$
$$SE = 100 \times TP / (TP + FN)$$
$$SP = 100 \times TN / (TN + FP) \qquad (13)$$

Where *N, TP, TN, FP*, and *FN* are respectively the total number of patterns, the number of true positive, the number of true negative, the number of false positive, and the number of false negative [6]. Using the expanded training set and the test set, the method based on the CWT and the Fuzzy Supervised Clustering system, introduced in Fig.2, was evaluated that showed good results for the test set. To investigate the performance of the VLP detection method proposed in this work, the conventional time-domain method (Simson's method) was applied to the test set; also, a method based on applying our system to the conventional time-domain features [4,13] was used to detect VLP. Table1 presents the results of the proposed method in comparison with Simson's method and applying a Proposed Fuzzy System to the conventional time-domain features, for the test set. In Simson's method, the balance between *SE* and *SP* can be controlled by choosing one, two, or three of the positive VLP criteria (QRS_T>114ms, D_{40}>38ms, and V_{40}<20μV) at the same time. For example, if higher *SE* is desired, only one of the three criteria is chosen; in contrast, all of the criteria must be satisfied at the same time to achieve higher *SP*. This fact can be seen in Table 1. In most cases, two of the three criteria are used to obtain the balance between *SE* and *SP* [13]. These results obtained with a Fuzzy system with 5 rules.

Table1: VLP detection results. The comparison between the proposed method, Simson's method and applying a Supervised Fuzzy Clustering system to the features (for the test set).

Method		ACC(%)	SE(%)	SP(%)
Simson's Method	One Criterion	80	100	60

Two Criterion	75	80	70
Three Criterion	72	60	85
Applying Neural Network Method to mVM_CWT1 Features	92.5	91	94
Applying Supervised Fuzzy Clustering to the conventional time-domain features	82	81.26	83.66
Applying Supervised Fuzzy Clustering to VM_QRS40 Features	84	73.76	94.13
Applying Supervised Fuzzy Clustering to mVM_QRS40 Features	90	89.88	91.16
Applying Supervised Fuzzy Clustering to mVM_CWT1 Features	92	88.83	95.16

6 Conclusion

The aim of this work was to investigate the capability of a method based on the continuous wavelet transform and a Supervised Fuzzy Clustering, in order to extract features and classify for VLP detection problem in HRECG. The results show good improvements in sensitivity and discriminancy compared to Simson's method and Supervised Fuzzy Clustering to the conventional time-domain features (see Table1). Another possible advantage of the proposed method can be the ability of the VLP detection in patients with bundle branch block, while Simson's method can not be used with these patients because the bundle branch block causes the QRS duration to be extended and consequently increases QRS_T. Simson's method is very sensitive to QRS offset, and even a few errors in estimating of this point can result in wrong diagnostic. However the method used in this study is robust to the error in QRS offset detection. The results represented in Table 1 shows that the CWT based features, used in this study, are more capable than the conventional time-domain features to detect VLPs using Supervised Fuzzy Clustering Systems. According to the above advantages, the proposed method may be an appropriate alternative method for the clinical detection of VLP, if it is evaluated completely. Due to the lack of real HRECG for VLP in this research, the proposed method must be applied to a larger HRECG database consisting of real signals with and without VLPs to complete the evaluation. Form Table1, Neural Network Method is a little better than Our Supervised Clustering method but our system is based on only five fuzzy rules and so the training of system is much faster than a MLP Neural Network and can be used in practical applications.

References

[1] I.C.Baykal and A.Yilmaz, Detection of Late Potentials in Electrocardiogram Signals in both Time and Frequency Domains Using Artificial Neural Networks, *44th IEEE Midwest Symposium on Circuits & Systems*, 2001, pp.576-579.

[2] S.W.Chen, A Wavelet-Based Heart Rate Variability Analysis for the Study of Nonsustained Ventricular Tachycardia, IEEE Transactions on Biomedical Engineering, Vol.49, No.7, 2002, pp.736-742.

[3] A.Cohen , Biomedical Signal Processing, CRC Press, VolumeII, 1986.

[4] A.Mousa and A.Yilmaz, Neural Network Detection of Ventricular Late Potentials in ECG Signals Using Wavelet Transform Extracted Parameters, IEEE Proceedings of the 23rd Annual EMBS International Conference, Turkey, 2001, pp.1668-1671.

[5] A.Mousa and A.Yilmaz, A Method Based on Wavelet Analysis for the Detection of Ventricular Late Potentials in ECG Signals, 44th IEEE Midwest Symposium on Circuits & Systems, 2001, pp.497-500 .

[6] A.Rakotomamonjy, B.Migeon, and P.Marche, Automated Neural Network Detection of Wavelet preprocessed Electrocardiogram Late Potentials, Med.Biol.Eng.Comput.,Vol.36, 1998, pp.346 - 350.

[7] M.B.Simson, Use of Signals in The Terminal QRS Complex to Identify Patients with Ventricular Tachycardia after Myocardial Infarction, Circulation, Vol.64, No.2, 1981, pp. 235-242.

[8] A.Spaargaren and M.J.English, Detecting Ventricular Late Potentials Using the Continuous Wavelet Transform, Proceedings of Computers in Cardiology,IEEE Comput.Soc.,Vol.26, 1999, pp.5-8.

[9] A.Taboada-Crispi, J.V.Lorenzo-Ginori, and D.F.Lovely, Adaptive Line Enhancing Plus Modified Signal Averaging for Ventricular Late Potential Detection, Electronics Letters, Vol.35, No.16, 1999, pp.1293-1295.

[10] A.Taboada-Crispi, Improving Ventricular Late Potentials Detection Effectiveness, Doctoral Thesis, University of New Brunswick,2002.

[11] S.Wu, Y.Qian, Z.Gao, and J.Lin, A Novel Method for Beat-to-Beat Detection of Ventricular

Late Potentials, IEEE Transactions on Biomedical Engineering,Vol.48,No.8,2001, pp.931-935.
[12] Q.Xue and B.R.S.Reddy, Late Potential Recognition by Artificial Neural Networks, IEEE Transactions on Biomedical Engineering, Vol.44, No.2, 1997, pp.132-143.
[13] Janos Abonyi and Ferenc Szeifert , Supervised Fuzzy Clustering for the Identification Of Fuzzy Classifiers , University of Veszprem, Dep. of Process Engineering,
[14] Detection of Ventricular Late Potentials in High-Resolution ECG Signals by a Method Based on the Neural Network and Continuous Wavelet Transform , Ali Shahidi Zandi , Mohammad Hassan Moradi
[15] Baraldi A. and Blonda P. (1999) A survey of fuzzy clustering algorithms for pattern recognition - Part I, IEEE Transaction On Systems, Man and Cybernetics Part B29 (6): 778-785.
[16] Bezdek J.C., Hathaway R.J., Howard R.E., Wilson C.A. and Windham M.P., (1987) Local Convergence Analysis of a Grouped Variable Version of Coordinate Descent. Journal of Optimization Theory and Applications, 71:471–477.
[17] Biem A., Katagiri S., McDermott E., Juang BH. (2001) An application of discriminative feature extraction lo filter-bank-based speech recognition.IEEE Transactions On Speech And Audio Processing 9 (2): 96-110.
[18] Campos T. E., Bloch I., Cesar R. M. Jr. (2001) Feature Selection Based on Fuzzy Distances Between Clusters: First Results on Simulated Data.Lecture Notes in Computer Science, Springer-Verlag press, ICAPR'2001 - International Conference on Advances in Pattern Recognition, Rio de Janeiro, Brazil, May.16
[19] Cios K.J., Pedrycz W., Swiniarski R.W. (1998) Data Mining Methods forKnowledge Discovery. Kluwer Academic Press, Boston.
[20] Corcoran A.L., Sen S. (1994) Using real-valued genetic algorithms to evolve rule sets for classification. In IEEE-CEC, June 27-29, 120–124, Orlando, USA.
[21] Gath I. and Geva. A.B. (1989) Unsupervised Optimal Fuzzy Clustering,IEEE Transactions on Pattern Analysis and Machine Intelligence 7 773–781.
[22] Gustafson, D.E., Kessel, W.C. (1979) Fuzzy Clustering With a Fuzzy Covariance Matrix, In Proc. IEEE CDC, San Diego, USA.
[23] Hathaway R.J. and Bezdek J.C. (1993) Switching Regression Models and Fuzzy Clustering. IEEE Transactions on Fuzzy Systems, 1:195–204.
[24] Hoppner F., Klawonn F., Kruse R. and Runkler T. (1999) Fuzzy Cluster Analysis – Methods for Classification, Data Analysis and Image Recognition,John Wiley and Sons.
[25] Ishibuchi H., Nakashima T., Murata T. (1999) Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems.IEEE Trans. SMC–B 29, 601–618.
[26] Kambhatala N. (1996) Local Models and Gaussian Mixture Models for Statistical Data Processing, Ph.D. Thesis, Oregon Gradual Institute of Science and Technology.
[27] Kim E., Park M., Kim S. and Park M. (1998) A Transformed Input–Domain Approach to Fuzzy Modeling. IEEE Transactions on Fuzzy Systems, 6:596–604.
[28] Loog L.C. M., Duin R.P.W. ,Haeb-Umbach R. (2001) Multiclass linear dimension reduction by weighted pairwise Fisher criteria, IEEE Trans. On PAMI, vol. 23, no. 7, pp. 762-766.
[29] Nauck D. and Kruse R. (1999) Obtaining interpretable fuzzy classification rules from medical data. Artificial Intelligence in Medicine, 16: 149–169.
[30] Pe˜na-Reyes C. A. and Sipper M. (2000) A fuzzy genetic approach to breast cancer diagnosis. Artificial Intelligence in Medicine, 17: 131–155.
[31] Quinlan J. R. (1996) Improved Use of Continuous Attributes in C4.5, Journal of Artificial Intelligence Research, 4: 77–90.
[32] Rahman A.F.R. and Fairhurst M.C. (1997) Multi-prototype classification: improved modelling of the variability of handwritten data using statistical clustering algorithms. Electron. Lett. 33 14, pp. 12081209.17
[33] Roubos J.A., Setnes M. (2000) Compact fuzzy models through complexity reduction and evolutionary optimization. In FUZZ-IEEE, pp762-767, May 7-10, San Antonio, USA.
[34] Roubos J.A., Setnes M. and Abonyi J. (2001) Learning fuzzy classification rules from data. Developments in Soft Computing, eds. John, R. and Birkenhead, R., Springer - Verlag Berlin/Heidelberg, 108-115.
[35] Setiono R. (2000) Generating concise and accurate classification rules for breast cancer diagnosis. Artificial Intelligence in Medicine, 18: 205–219.
[36] Setnes M., Babu˘ska R. (1999) Fuzzy Relational Classifier Trained by Fuzzy Clustering, IEEE Trans. SMC–B, 29, 619–625
[37] Setnes M., Babu˘ska R., Kaymak U., van Nauta Lemke H.R. (1998) Similarity measures in fuzzy rule base simplification. IEEE Trans. SMC–B 28, 376–386.
[38] Takagi T., Sugeno M. (1985) Fuzzy identification of systems and its application to modeling and control. IEEE Trans. SMC 15, 116–132.
[39] Valente de Oliveira J. (1999) Semantic constraints for membership function optimization. IEEE Trans. FS 19, 128–138.

Distributed Wirelss Sensor Network Architecture: Fuzzy Logic based Sensor Fusion

Garimella Rama Murthy
IIIT-Hyderabad, India
rammurthy@iiit.ac.in

Vasanth Iyer
IIIT-Hyderabad, India
vasanth@research.iiit.ac.in

Abstract

In this research paper, power aware distributed wireless sensor network architecture proposed by us (in the context of design of routing algorithm) is summarized. An approach to combined routing and fusion algorithm is discussed. Fuzzy logic based approach to wireless sensor fusion is discussed. The need for modeling sensor readings using fuzzy uncertainty is discussed. Novel overlap function for sensor fusion is presented. Concept of fuzzy overlap function is proposed. Method of distributed computation of means is proposed.

Keywords: Sensor Fusion, Fuzzy Logic, Network Architecture.

1 Introduction

Traditionally sensors were used to monitor natural or artificial phenomena. For instance, in the case of manufacturing process, the temperature and pressure in a set of boilers is monitored by sensors. The sensor readings are communicated to a centralized control center. Thus, conventionally the sensors are connected to one another through wires. These wired sensor based systems were utilized in many practical applications.

Motivation for Wireless Sensors:

Advances in VLSI (Very Large Scale Integration) as well as MEMS (Micro-Electro-Mechanical Systems) facilitated the manufacture of tiny sensors (say sensor motes). These tiny sensors have multiple functionalities such as sensing, processing/computing and communication. Thus, each sensor is typically equipped with a sensor (sensing unit), a processor (computing unit), a transmitter/receiver(transceiver for communication), memory, small battery etc. It is demonstrated by companies like Crossbow that these tiny sensors are commercially viable.

Wireless Sensor Networks:

Advances in sensor technology and computer networks have enabled distributed sensor networks (DSNs) to evolve from small clusters of large sensors to large swarms of micro-sensors, from fixed sensor nodes to mobile nodes, from wired communication to wireless communication, from static network topology to dynamically changing topology [5]. However, these technological advances have also brought new challenges to processing large amount of data in a bandwidth-limited, power constrained, unstable and dynamic environment [5]. To be effective, a wireless sensor network must implement the following procedures:

(a) Neighbor Discovery and Self Organization, (b) Distributed Computation Algorithms, (c) Capacity Optimization, (d) Localization, (e) Routing, (f) Sensor Fusion etc. Details related to these procedures can be found in [2].

Let us consider the routing problem/procedure. It is easy to see that most wireless sensor networks monitoring a phenomenon are data-centric. The goal is to communicate an event from a certain region of sensor field to the base station. In such data-centric networks absolute addressing of sensor nodes is not necessary. Our research group proposed two routing algorithms in wireless sensor networks. Details of these algorithms are described in Section 2. Now let us consider the sensor fusion problem.

Sensor Fusion Problem:

We illustrate this problem using an application. Consider the problem of monitoring fires in a forest. Utilizing a helicopter/air-craft temperature sensors are deployed over a certain terrain. In the deployment phase, some sensors could become faulty. Some more sensors could malfunction due to various reasons after deployment. These faulty as well as functioning wireless sensors monitor the temperature and transmit the information to a base station (using various routing procedures).

At the base station, the temperature information (raw data) coming from various sensors has to be aggregated to detect an event such as fire. Thus we are naturally led to the problem of designing a fusion/aggregation algorithm. Suppose the average value of raw temperature readings is computed it is easy to see that the temperature readings coming from faulty sensors influence the aggregated value and thus leading to a wrong aggregated value. In the worst case, it might so happen that an event such as forest fire is not properly detected. Hence it is necessary to design a robust/fault tolerant sensor fusion algorithms.

In the remaining portion of the research paper, we often refer to a wireless sensor network deployed to monitor the temperature in a forest region and report events such as a forest fire.

2 Power Aware, Distributed Network Architecture

In the case of wireless sensor network, the sensors are deployed over the sensor field. The performance of sensor fusion algorithm depends on the network architecture. Also, sensors should be coordinated using a distributed architecture so that an efficient routing algorithm can be designed. Thus, in the following we summarize the routing algorithms designed by us and the associated network architecture.

The sensor network lifetime is limited due to constrained power requirements of sensor motes. The network lifetime is defined as period of network reliability till the first sensor drains out of power as shown in figures 7, 8, 9&10. To maximize the sensor network life-time the architecture relies on using power-aware routing and data aggregation algorithms. As the need to minimize network traffic only on-demand protocols are considered for power-aware routing. As seen from simulation results the power-aware LEACH algorithm does 40% better than the standard LEACH in terms of number of rounds before the first sensor fails. The remaining Life-time of the sensor network gives the residual power, from the simulation results the standard LEACH increases by 12% when the number of cluster heads is increased gradually up to 50% of the total sensors deployed. The main reason of better efficiency with higher cluster head count is due to evenly distributing the number of sensors per cluster head. In the case of power-aware LEACH the remaining life-time is evenly spread so as to maximize the number of rounds by better balancing the selection of cluster heads. At lower percentage of cluster heads the power balancing algorithm performs best so to not fail a single sensor till the last sensor has no more power. When the percentage of cluster head is gradually increased up to 50% the power balancing algorithm selection is suboptimal due to the increased need to find more optimal cluster heads. Due to this some of the cluster heads have less resource than average remaining residual power causing earlier failure of sensor nodes. The protocols considered can be categorized in the following ways-- flooding, multi-hoping, clustering and our own distributed algorithms like Control Radius Flooding [4], Hierarchical Leveling, Power-aware routing using dual-cluster members.

2.1 Conventional Routing Algorithms

Flooding is a conventional routing algorithm as shown in Figure 1 which allows discovering the topology and the destination path. Due to the nature of the flooding process it generates many duplicates causing collision in the network. The variants of this type of protocol allow a transmission probability to be sent with the message header which can be used to forward or not to forward the current message to its neighbor. Such protocols are classified as gossip and regional flooding. The problem of power-awareness is partially addressed in these types of protocols as these could be deployed on an on-demand or reactive type network without using any prior information of the topology or any other network protocol discovery overheads.

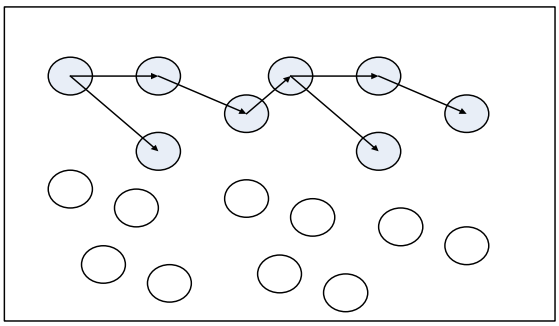


Figure 1: Example of flooding.

Multi-hopping as shown in Figure 2 is an enhancement of the earlier routing protocol by introducing a hop-count in the protocol. This allows forwarding the messages discretely by checking the least hop count to the destination. The overhead in this is to maintain the routing information for the underlying network. The problem of power-awareness is addressed as the total power dissipated by the network is minimized due to multi-hopping. Hence the power dissipated during forwarding a message to its nearest neighbor is the least compared to sending it to the base station directly, enhancing the total life-time of the network.

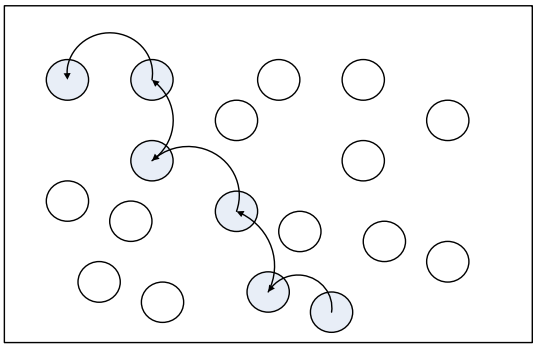


Figure 2: Example of multi-hoping.

2.2 Distributed power-aware architectures

Clustering is a technique in which a cluster head is elected using a power-aware cost factor and its entire neighbors join the cluster dividing the network as shown in Figure 3 into many different regions. This further reduces the power dissipation compared to multi-hopping. The clustering algorithms like LEACH [9] (Low Energy Adaptive Clustering Hierarchy) and HEED (Hybrid, Energy-Efficient, Distributed clustering) help to find suitable cluster heads using localized optimizations or finding optimizations using a distributed parameter. As in the case of sensor networks periodic data is gathered and sent to the base station. Clustering aggregates data to a single cluster head further reducing the energy constraints and enhancing the total lifetime of the network.

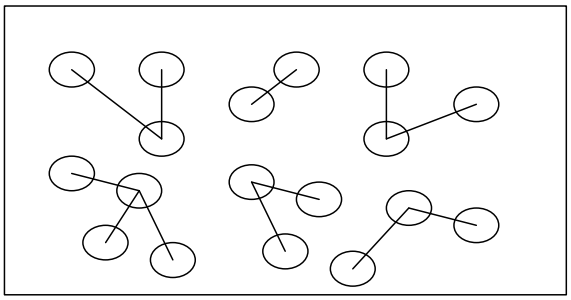


Figure 3: Example of clustering.

2.3 Our Routing approaches

A zone based protocol called CRF (Control Radius/Radio Flooding) is proposed by our group which partitions the sensor network mainly into three hierarchically functional zones as shown in figure 4 using sensor's variable transmit power, the zone farthest from the base station does the data collection and the middle zone does the routing using a route cache and the zone closest to the base station does multi-hopping of data to the base station and enhancing the lifetime of the sensor network. The route cache maintains a use count which is checked against a residual cost factor and if it reaches or exceeds that threshold the cluster head is rotated with one which has a lower use count.

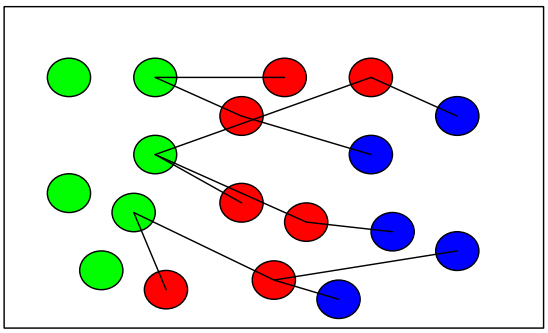


Figure 4: Example of Control Radius/Radio Flooding.

- Leveling Algorithm

Hierarchical Leveling is a frame work as shown in figure 5 in which base station divides the network using proximity of the sensor distance from itself using the power to reach a segment of the network. By this all the nodes have a specific level number and eventually at the end of the division of the network all nodes farthest away from the base station have a large number compared to the nodes closer to the base station making the base station at zero. When an event is detected in the outmost level the message is embedded with its current level and broadcasted, the listening neighbors check if the level is lower than the source of the message if so they are re-broadcasted over the network by which propelling the message to the base station and avoiding collisions generated by the same levels. The problem of power-awareness is addressed by minimizing re-broadcast of unnecessary packets as in the case of conventional flooding.

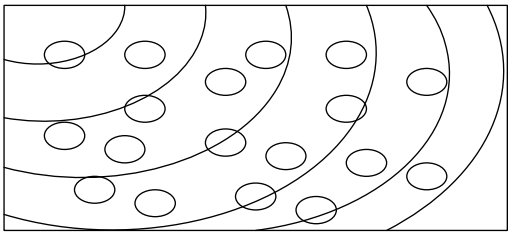


Figure 5: Example of Hierarchical Leveling

- Clustering Algorithm

The new algorithm proposes dual-membership clustering as shown in Figure 6, which is very useful to minimize the use of epidemic algorithms of any sort, thereby saving a lot of resources and also making route discovery more deterministic. This novel technique uses selection of cluster heads and also keeps track of the most recently used path for using the off-line cluster head. The problem of power-awareness is addressed by combining clustering and multi-hopping during transmission and also uses a dual membership of clusters which minimizes the overhead of route discovery.

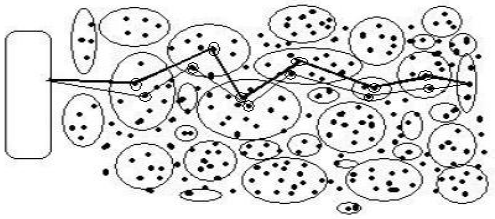


Figure 6: Example of dual membership cluster and recent path usage

2.4 Intuitive betterness of our algorithms

Intuitively, we reason that our routing algorithms are better than various known variants of flooding algorithms from the point of view of minimizing duplicate packets in the network and thus power.

Our first routing algorithm effectively achieves hierarchical partitioning of the network into levels [1]. The data packets are propelled in the direction of base station using a controlled flooding algorithm (also called gossip routing algorithm). The algorithm prevents propagation of packets in the opposite direction of the base station. Thus the algorithm ensures that duplicates packets that are generated (due to broadcast nature of transmission) are prevented from being transmitted repeatedly. The idea of our algorithm has some resemblance to the regional gossip algorithm proposed in [7]. This regional gossip algorithm makes the strong assumption of location awareness.

Traditionally clustering algorithms such as LEACH, HEED are proposed to minimize energy consumption. We proposed a dual membership clustering algorithm [3]. This algorithm also proposed the innovative idea of forming data highways on the sensor field and utilizing them for routing. Thus, in this algorithm, the overhead associated with route discovery is minimized.

• Leveling plus clustering algorithm

We also proposed a hybrid algorithm which combines the advantages of leveling algorithm and clustering algorithm. Detailed performance study is reported in [8]. Thus we expect this hybrid algorithm to be a good power aware algorithm of utility in wireless sensor

networks. In the following section, we propose a combined routing and fusion algorithm.

3. Traditional Wireless Sensor Fusion Approaches

Decision/Information fusion problems have been investigated since few centuries. For instance, the works of Condorcet on democracy models in 1786 and Laplace on composite methods studied the fusion problem. In engineering, the investigations of Von Neumann showed that a reliable system can be built using unreliable components by employing simple majority fusers [5].

Wireless sensor fusion is one among numerous application areas in which information/decision fusion methods have been employed. Sensor fusion techniques must be robust and fault tolerant so that uncertainty and faultiness in sensor readings can be handled. Naturally redundancy in sensor readings are used to provide error tolerance in fusion [5].

- Fault Tolerant Fusion Algorithms: Overlap Functions

An abstract sensor is defined as a sensor that reads a physical parameter and gives out an abstract interval estimate which is bounded and connected subset of the real line. Abstract sensors can be classified into correct sensors and faulty sensors. A correct sensor is an abstract sensor whose interval estimate contains the actual value of the parameter being measured. Otherwise, it is a faulty sensor. A faulty sensor is tamely faulty if it overlaps with a correct sensor, and is wildly faulty if it is does not overlap with any correct sensor [5].

Several overlap functions are defined that process the readouts from abstract sensors. Four such functions are discussed in [5] in the references. For instance Marzullo defined the M function. Let n be the number of sensor readouts, f be the number of faulty sensors. $M([I_1, I_2, ..., I_n])$ is defined to be the smallest interval that contains all the intersections of n-f intervals. It is guaranteed to contain the true value provided the number of faulty sensors is at most f. The disadvantages of such a function are discussed in [5].

- Also, there are efforts to model the sensor readings using probabilistic uncertainty. In the following, we reason that in the case of a wireless sensor network monitoring a phenomenon (such as forest fire), the sensor measurements and/or cluster head measurements should be modeled using fuzzy uncertainty

- Combined Routing and Fusion:

In all wireless sensor networks, the sensors monitoring a phenomenon collect raw data. In pure routing algorithms (epidemic algorithms) such as flooding and its variants (controlled flooding: gossip, regional gossip), the collected raw data is propagated to the base station without any processing.

We propose an approach in which the routing decisions as well as aggregation/fusion decisions are combined. For instance the leveling and clustering algorithm described in Section 2 enables combined routing and fusion.

4. Wireless Sensor Fusion: Fuzzy Logic Based Approach

Limitation of Traditional Fusion Approaches:

In the case of conventional, wireless sensor fusion algorithms, the raw temperature (for example) readings are transmitted to the cluster head. The cluster head can transmit all the raw/unprocessed values to the base station. Alternatively, the cluster head processes the raw values and produces a fused estimate. Typically the fusion measure employed at the cluster head is mean, medium, mode etc. The cluster heads propagate the fused values to the Base Station (in a cooperative manner).

From the point of view of natural/artificial phenomenon being monitored (e.g. Forest Fire), the actual sensor readings (or processed estimate, say mean) are only of limited utility. *To take any control action (say to control forest fire), we want to decide whether the temperature is LOW, MEDIUM, HIGH etc.*

Thus the linguistic variable corresponding to the monitored variable (say temperature) is characterized by fuzzy uncertainty. The linguistic variable (say temperature) assumes values in the fuzzy sets, “LOW”, “MEDIUM”, “HIGH” etc. To emphasize the point, it should be understood that the uncertainty cannot be modeled using probabilistic uncertainty, but requires fuzzy uncertainty

Fuzzy Sensor Fusion:

There are two possible approaches to fusion of measurements collected at sensor nodes

Approach—1:

Crisp measurements at the sensor nodes are aggregated using crisp fusion measures such as median, mode, mean etc (It should be seen that using median leads to a robust estimate as in the case of median filtering).Using the crisp estimates, fuzzification is done at the cluster heads. These fuzzified values at the cluster heads are propagated to the base station. At the base station, different aggregation measures are employed. After proper de-fuzzification, fire control decisions are taken.

Approach—2:

In contrast to approach-1, the monitored variables (temperature) at the sensors are treated as being fuzzy i.e. they assume values in the fuzzy sets LOW, MEDIUM, HIGH (temperature). The crisp measurements are fuzzified. Using fuzzy arithmetic rules, median, mode type measures are computed at the cluster head. The fused values at various cluster heads are composed using FUZZY IF THEN ELSE rules. As in the case of a Fuzzy control system [page 635, 6], fuzzy rule base, fuzzy inference engine, defuzzification interface are utilized to take a control action for monitoring temperature in a forest (Specifically, in case of forest fire event, appropriate control action is taken).

Let temperature at the i-th sensor node/i-th cluster head be denoted by TEMP-i. Typical fuzzy IF, THEN, ELSE rules are:

(1) IF TEMP-1 is LOW, TEMP-2 is LOW, …., TEMP-N is LOW,

THEN FIRE CONTROL IS (NOT REQUIRED)

(2) IF MAJORITY OF (TEMP-1, TEMP-2, …, TEMP-N) is HIGH,

THEN FIRE CONTROL is (REQUIRED).

5. Distributed Sensor Fusion: Mean Computation

In traditional sensor arrays, measurements are simply sent to a central signal processor for computation. In contrast a wireless sensor network should be able to provide some useful inferences based on measurements at the sensors [2]. Anurag et.al proposed distributed computation of maximum and reasoned that the communication complexity is reduced.

We propose distributed computation of mean. For 'n' sensors placed in one dimension, if the average value of measurements is needed, then the method of sending all individual values to a central operator would be extremely inefficient in communication complexity ($0(n^2)$ for 'n' sensors placed in one dimension , with the operator node at one end). Mean can be recursively computed in the network as the sensors exchange their measured values. The distributed computation of mean is illustrated below:

Let the lower case letters denote raw measurements i.e. $\{x_0, x_1, x_2, ...x_n\}$. As shown below, the mean values are computed in a distributed manner:

$X(0) = x_0$, $X(1) = \frac{X(0) + x_1}{2}$, $X(2) = \frac{2X(1) + x_2}{3}$, $X(3) = \frac{3X(2) + x_3}{4}$ and so on. This distributed computation based approach leads to minimal delay in determination of average temperature at the base station.

Now we return to some innovative ideas related to fuzzy sensor fusion.

6. Negative membership function: Modeling need

It is well accepted that a linguistic variable is modeled using fuzzy uncertainty. Now we reason the need for allowing the membership function to assume negative values i.e. membership function assumes values in the set [-1, 1].

Consider a "linguistic variable" such as temperature. Let the associated sets be "COLD", "HOT" and "MEDIUM" (to capture the degree of heat). The

following situation directs our attention towards negative membership function values. Consider a linguistic variable such as "Temperature". Let the temperature be -20°C. The degree of membership of that observation in the fuzzy set "HOT TEMPRATURE" is not zero but SRICTLY NEGATIVE. Suppose these is a common "UNIVERSE OF DISCOURSE" (for the elements of Fuzzy sets). It seems that using the current framework of fuzzy sets, it is possible to make into account the "negative membership" value and the need for it. We are thus BROADENING the scope of fuzzy sets.

It may be possible to convert the above situation in such a way that the traditional fuzzy set idea (with only positive membership function values) is sufficient. But we feel that there is loss of information in such an approach. One might conceive of situations where the membership function assumes complex values and not necessarily real values.

7. Novel Overlap Functions

• In [5], Qi et.al proposed Ω(.) overlap function. We propose a novel overlap function based on the prior Ω(.)

Dual of Ω – Function θ(x): It gives the number of intervals non-overlapping at 'x'. This function is more natural under the assumption that "most of the sensors" are non-faculty.

θ(x) = *N* – Ω(x), where *N* is the total number of sensors.

• θ(x) Associated overlap function β(x), Just as *N* function is associated with θ(x). β(x) gives the interval with the θ(x) Є [0,f]

Novel Idea: Associate a measure with the fused interval specified by the overlap function. This measure quantifies the agreement/disagreement among sensors.

8. Fuzzy overlap function

Suppose we model the sensed value to be a linguistic variable which assumes the values in an interval with associated membership function.

- Arrive at a "fused" interval estimate using one of the various overlap functions.
- Using the membership functions (corresponding to different sensors) over the "fused" interval, associate/compute the resulting membership function as follows:

(i) Compute the Maximum/Minimum of the membership values.

(ii) Compute the average/median other measure of the membership values.

9. Conclusions

In this research paper, design of wireless sensor networks is addressed. Specifically, the problem of wireless sensor fusion is considered.

To address the fusion problem, the distributed network architecture proposed by us is discussed. This architecture was utilized for designing routing algorithms. We propose a combined routing and fusion algorithm. Traditional approach to wireless sensor fusion is summarized. Fuzzy logic based approach to sensor fusion is proposed as a novel solution. The problem of distributed sensor fusion is discussed in the context of computation of average/mean value of the sensor readings.

10. Acknowledgement

Authors are grateful to DST, Govt. of India, for supporting the work with a grant. Also will like to acknowledge the faculty Dr. Rama Murthy department of Communication Research Centre and Dr. Srinivas of VLSI and Embedded systems who had constantly advised on the project since it began a year back.

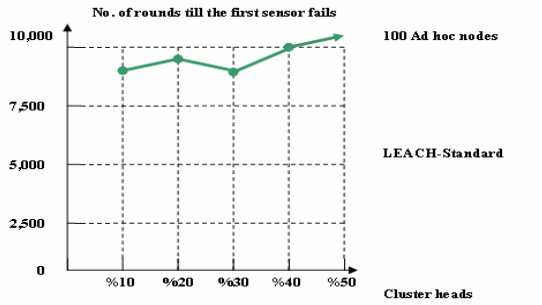


Figure 7: Plot of no. of rounds using LEACH standard with different number of cluster heads.

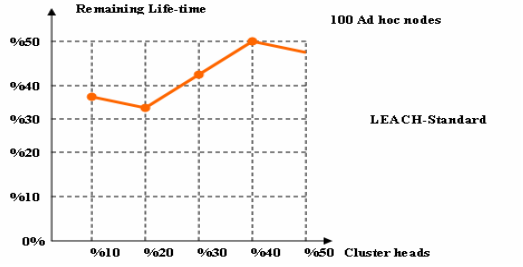


Figure 8: Plot of remaining life-time residual power using LEACH standard with different number of cluster heads.

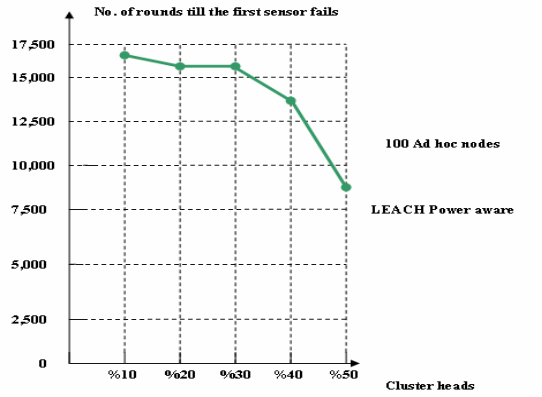


Figure 9: Plot of no. of rounds using LEACH power-aware with different number of cluster heads.

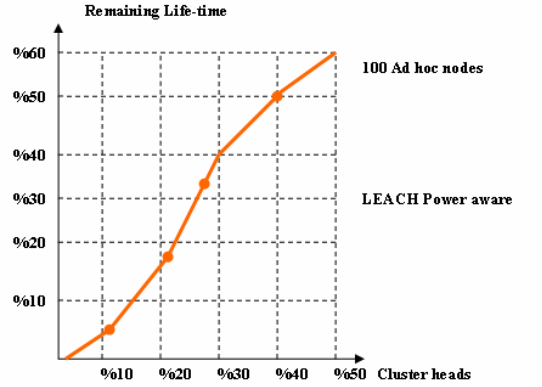


Figure 10: Plot of remaining life-time residual power using LEACH power-aware with different number of cluster heads.

References

[1] S. Arora, S. Chaudhary, M. B. Srinivas and G. Rama Murthy, "Power Aware, Probabilistic and Adaptable Routing Algorithm for Wireless Sensor Networks," Proceedings of 10th National Communications Conference (NCC), January 2004, Organized by IISc, Bangalore. Submitted to SADHANA.

[2] Anurag Kumar, D. Manjunath and Joy Kuri, "Communication Networking: An Analytical Approach," Morgan Kaufmann Publisher, 2005.

[3] N. Bharat Varma, M. Pratyush, G.Rama Murthy and M.B.Srinivas, "Energy-Aware Routing in Sensor Networks using Dual Membership Clusters and Data Highways," Proceedings of IEEE TENCON, Bangkok, 2004. GESTS International Transactions on Computer Science and Engineering, Vol.25, #1, December 2005, Pages. 196-208, ISSN 1738-6438.

[4] Vasanth Iyer, Rama Murthy, M.B. Srinivas "Information Processing of Wireless Sensor Networked Motes by Harvesting Residual Power", The Second IEEE International Workshop on Next Generation Wireless Networks 2006 (IEEE WoNGeN '06).

[5] H. Qi, S.S. Iyengar and K.Chakrabarty, "Distributed Sensor Networks---a Review of Recent Research, " Journal of the Franklin Institute 338 (2001), Pages 655-668.

[6] Satish Kumar, "Neural Networks: A Classroom Approach," Tata McGraw Hill Publishing Company Limited, New Delhi.

[7] Regional Gossip Routing for Wireless Ad Hoc Networks Xiang-Yang Li Kousha Moaveninejad Ophir Frieder.

[8] Amit Kumar, Rumeet Singh, G. Rama Murthy, M.B. Srinivas, "Energy Effcient Layer and cluster formation in Wireless Sensor Networks" B.Tech final year project, IIIT, Hyderabad.

[9] LEACH (Low Energy Adaptive Clustering Hierarchy) Wendi Heinzelman,Hari Balakrishnan,Anantha Chandrakasan.

Soft Computing Robot Navigation Case Study

Dimitar Vasilev Lakov
ICCS – BAS, Acad. G. Bonchev street, Bl.2
Sofia, 1113, Bulgaria
lakov@iccs.bas.bg

Margarita Raykova Saralieva
ICCS – BAS, Acad. G. Bonchev street, Bl.2
Sofia, 1113, Bulgaria
margi@iccs.bas.bgl

Abstract

This contribution presents Soft Computing Robot Navigation for mobile robot SCORCAS, a Case Study. The Navigation uses fuzzy logic experienced on mobile robot systems. Fuzzy logic is applied in analogous way into three sub-systems:

i. Sensor, consisting of five autonomous sensor sub-systems: visual, acoustic, tactile, thermal and inertial;
ii. Sensor fusion that comprise all the robot stimulus in environment image representation;
iii. Path planning – responsible for trajectory generation.

These systems are implemented into experimental prototype. Some results are reported.

Keywords: Soft Computing, Fuzzy System, Robot Navigation.

1 Introduction

This contribution presents Soft Computing Robot Navigation for mobile robot SCORCAS, a Case Study. It is necessary step in our project dedicated to Intelligent Technology Application in Robot Navigation Field. From great variety of soft computing models we chose fuzzy logic representation. The aim of the publication is twofold:

i. To show growing potential of fuzzy logic in real object implementation of robot navigation;
ii. To create the basis and pave path for more profound investigation in the field pertain to the next generation of Mobile Soft Computing Agents dedicated to mobile robot navigation [1].

As far as the Fuzzy Navigation Strategy is based on well known Fuzzy Inference Engine our achievement can be sought more in application aspect of this intelligent technique than in any new sophistication in the field. We show the main stages in realization of multi-sensor navigation system dedicated to mobile robot implementation: some peculiarities of mobile construction, architecture of the control and navigation system, sensor fusion technique, program realization, and some preliminary results of SCORCAS.

Among many navigation tasks of mobile robots such as: optimal trajectory planning; obstacle avoidance; smooth trajectory formation, control with time and spatial constraints, etc., we turn the spotlight on the first, since it comprises in some sense all the others. The main peculiarities of this task can be summarized as follows:

i. Uncertain still dynamically changing working environment that require effective sensor system in combination with some intelligent approach;
ii. Fairy constraint artificial sensor systems of nowadays robots. Although we involve: visual, acoustic, tactile, thermal and inertial sensors, which in some extent mimic human beings: sight, hearing, tactile, thermo, and vestibular sensing, they are incomparable imperfect in comparison with those of human beings;
iii. The same is situation with effectors that in our case are implemented by means of double caterpillar, reversible system and additionally mounted rotating tower that would correspond to human neck functions of human being. Again the human abilities prevail in many times;
iv. Even human being activity is constrained in two dimensional space they are able to assess and perform actions in the third dimension without problems (refer paddling);

The most attractive feature of live systems, unreachable of artificial ones by now, are their ability of emotional actions, which act in emergency situations. It is processed by another informational channel that bypass normal brain controlled activities and is an order faster that the last. Surprisingly enough this

channel is 300 000 years older than consciousness brain activity, but practically proven during many centuries. This is nature gift left from prehistory times, eternal dream for every explorer.

There are many investigations in the field of intelligent navigation technique, concerning one or other of the problems already marked. For examples [2], [3], and [4] present: static fuzzy based model, schema oriented, and intention based navigation systems. Ref. [5], [6], and [7] demonstrate another approach: fuzzy localization on landmarks, genetic based approach, and Kalman filter combining intelligent agents, whereas [8] and [9] stress their attention to intelligent learning: reactive, supervised, and hierarchically, all based oh fuzzy logic. In contrast to all these investigations our system applies five sensing systems mentioned in point ***ii***. The paper is organized as follows: Section two presents the robot structure: mechanics, robot control and navigation system, as well as sensor system. In Section three we describe intelligent control system, whereas Section four refers to a Case Study. Paper ends with discussion about future trends of the project.

2 Presenting the Robot Structure

2.1 Mechanical and Information Structure

As an experimental tool we have chosen a kid toy tank construction consisting of a head (tower), platform, and base unit. The head has two grades of freedom to mimic human neck swivelling and joggling. It carries the all sensor systems. Platform controls and performs communication tasks. Base unit is placed in movement by two independent caterpillar direct current drive units. Mechanical and Information Structure is shown in Figure 1. A combination of velocity and direction of movements of both caterpillars produces controllable smooth rotation over centres placed on cross direction line. Depending on velocity and direction, the centre of rotation is moved from left to right infinities. When the centre of rotation is in infinity, the tank moves straight line in fore or back direction. The smooth movement of the centre from left to right infinity causes left/right rotation with decreasing/increasing radius of rotation. If a radius coincides with the geometrical centre of the tank, the last spins at a place in clock/contra wise directions depending on the directions of both caterpillars. By this way the tank possesses exclusively flexible dynamic features.

2.2 Robot Control Organization

A simplified robot control organization scheme is shown in Figure 1. Navigation strategy is generated in highest level. After generation of appropriate optimal paths they are distributed in two tasks: base unit control and head control. A middle synchronization level is used to coordinate these tasks.

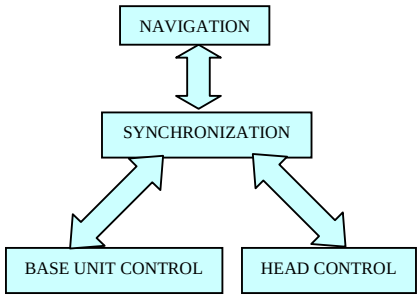


Fig. 1 Robot Control Organization

2.3 Robot Control and Navigation

Robot control system is three layers hierarchical, shown in Figure 2:

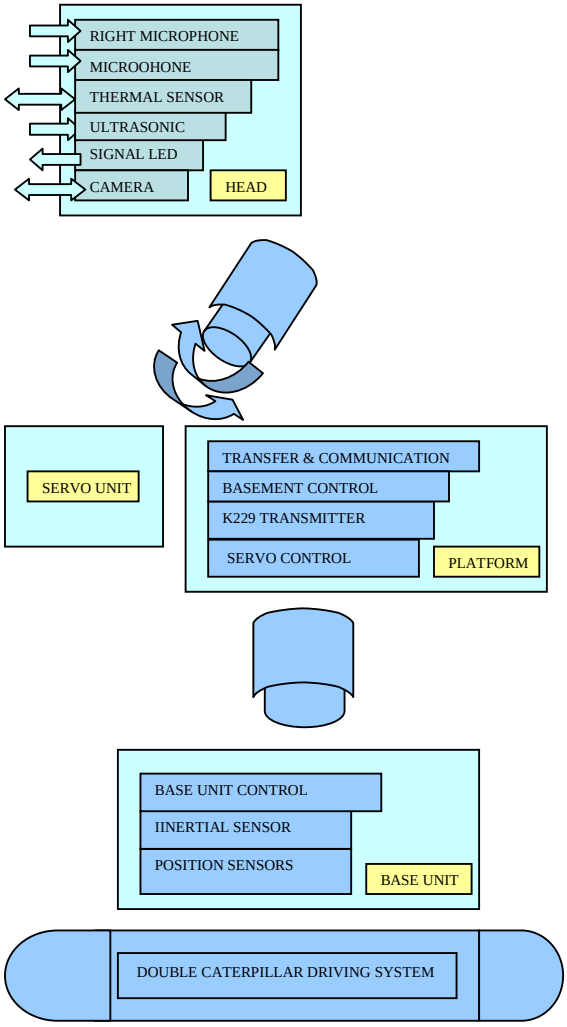


Fig. 2 Robot Control and Navigation System

i. The lowest consists of two independent control units. One for direct drive of caterpillar DC motors, which are precisely control by specialized modulation controller. The other - for control of swivelling and joggling head movements;

ii. The second level is responsible for synchronization of left and right drivers as well as coordination concerning motion control;

iii. The highest is tactic and strategic level defining trajectories. Every trajectory is defined by time dependant steady states. They are transferred by wireless connection. The task of this level is robot navigation. It is implemented via PC using Intelligent Soft Computing Technology. States are transferred to the lowest level of control system of the robot.

2.3 Sensor Processing System

Since every sensor system is autonomous, the whole information is processed in parallel. The raw data are normalized within (0, 1] semi closed interval, then fuzzified using the same granulation of triangular fuzzy sets representation. This way of granulation is chosen for simplicity although some other possible representation could turn out to be more useful. For now it is left for next investigation. A conventional inference mechanism of all the fuzzy systems is chosen. Every fuzzy stimulus system is one in - one out system. It task is to transfer stimulus into its corresponding fuzzy representation. Inputs are three terms granulated stimuli and outputs are five granulated assessment of quality in respect to the defined strategy. The inputs are presented by ‘low‘, ‘middle’, and ‘big’ terms, whereas their counterpart quality assessments are: ‘good’, ‘middle’, and’ ‘bad‘ with two additional intermediate terms; ‘good-middle’, and ‘bad-middle’. All in/out terms have triangular shape definition. An example of visual fuzzy inference system is given in Figure 3. and Figure 4.

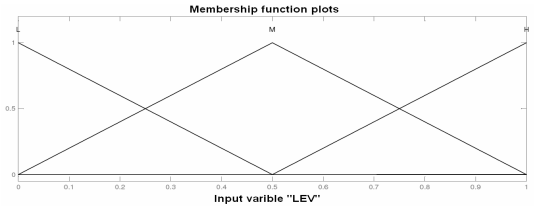


Fig. 3 Fuzzy input visual stimulus

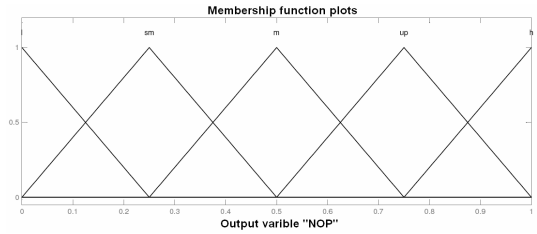


Fig. 4 Fuzzy output visual stimulus

Hence we have five simple inference systems. Everyone is presented by its fuzzy rule base with nine rules - fully exhaustive representation. In such a way we involve qualitative instead of quantitative stimuli assessment. The reason is to enlighten a process of their sequel fusion.

2.4 Sensor Fusion System

The sensor information data are transferred via transmitter to control PC; there an information fusion is performed. Existence of five sensor system requires sophisticated fusion mechanism able to process on line the information. We apply fuzzy information fusion of the pre-processed sensor data, which are outputs of the sensor processing systems. The scheme of information fusion is shown in Figure 7. After quality assessment of every sensor data they are ready for fusion. The fusion is performed using the same fuzzy logic technique. This is five in - two out fuzzy inference machine: five quality assessment of every sensor as inputs, and two decisions for both left and right drivers. In/out granulation is shown in Figures 5. and 6.

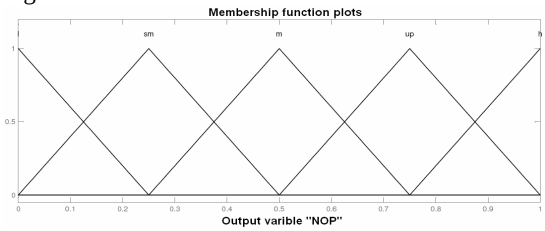


Fig. 5 Quality assessment of input for visual stimuli

Here inputs are presented by five stimulus granules of its quality: ‘good’, ‘good-middle’, ‘middle’, ‘bad-middle’, and’ ‘bad‘.

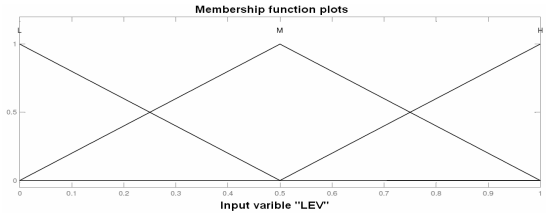


Fig. 6 Fuzzy output of fusion system

Outputs are granulated in three quality assessments for each left and right drivers, i.e.: ‘good’, ‘middle’, and’ ‘bad‘.

3. Intelligent Control System

Control System comprises three sub-systems:

i. Fuzzy Data Representation;

ii. Path Planning.;

iii. Navigation.

All the systems utilize similar fuzzy inference technique presented in Figure 7. On this stage the tower

is controlled independently from caterpillar motions. It needs only synchronization with base unit control system and high priority action in emergency situations.

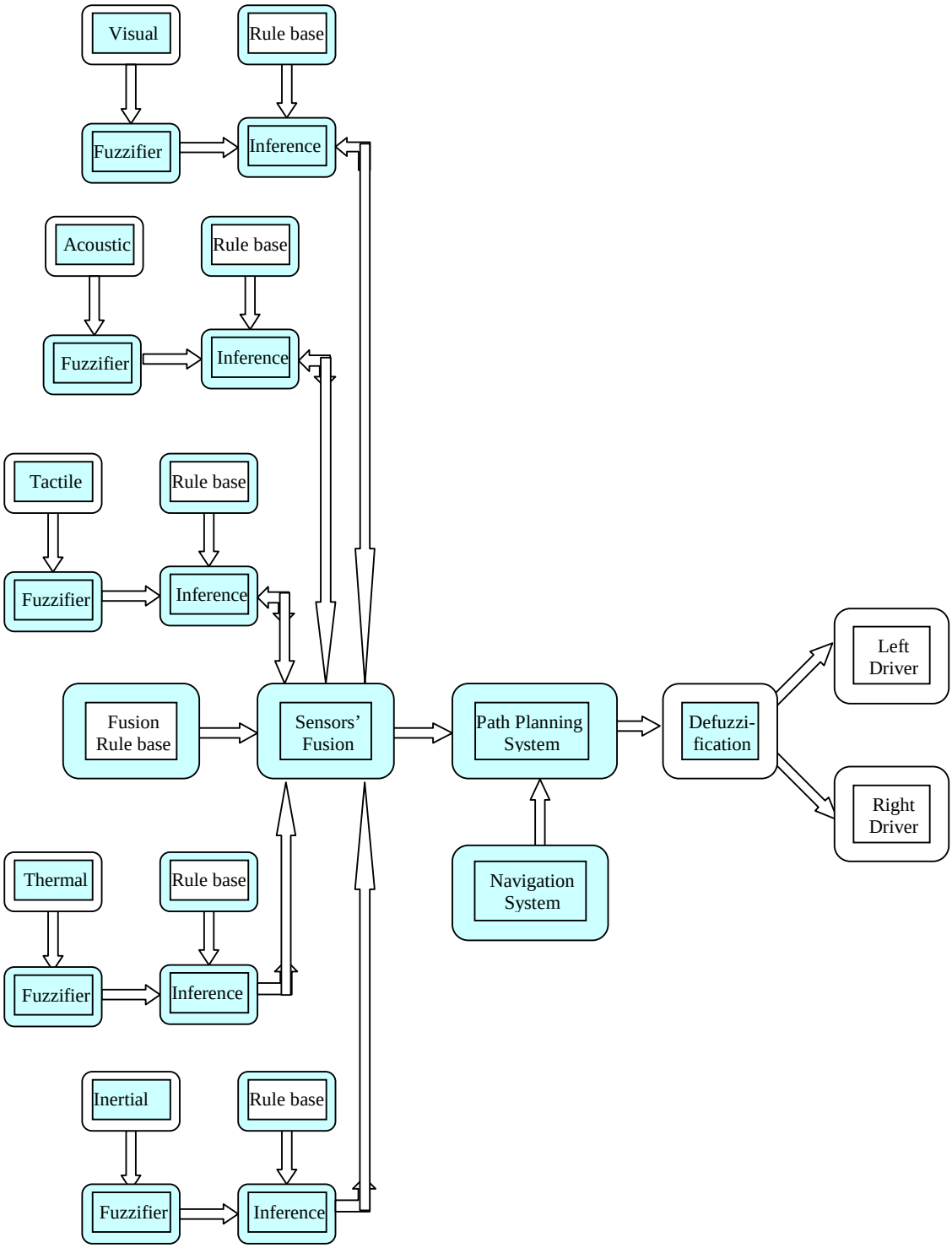


Fig. 7.Scheme of information fusion

Fuzzy Data Representation has five analogical sub-systems. They are combined to produce current envi-

ronment image of the robot position. Each of them consists of four blocks:

i. Input data that normalized measurements,
ii. Fuzzifier. It granulates crisp measurements in an unified for all sensors universe of discourse;
iii. Rule base that defines fuzzy inference assessment mechanism;
iv. Inference block that extracts fuzzy quality representation for every sensor measurement.

Path Planning System creates program motion based on current image and Navigation System. It interprets Navigation System inferences into short term goals. Then these goals are defuzzified and supply both left and right caterpillar drivers shown in Figure 7. Planning System also creates program motion in accordance with current image and Navigation System. Navigation System forms short (tactic) and long (strategic) goals of path planning, which supply Path Planning System. Navigation System utilises generalized environment image created by Fuzzy Representation System after data fusion. The fusion is realized in abstract five dimension space of sensed stimuli on the basis of vision, acoustic, tactile, thermal, and inertia sensors. At this stage we investigate a free trajectory without obstacles avoidance although the power of the system would be shown it its full strength in more complex tasks. Every stimulus obtains one of three fuzzy quality assessments after fuzzification of measured sensor values as: 'low', 'middle', and 'big'. We accept triangular atomic terms representation of fuzzy sets evenly distributed in their universes of discourse. Their fusion defines generalized quality assessment as point in five dimension space. The points present in its turn the quality of robot position in respect to our requirements, the next robot position and measurements. The next measurement produced short term prediction of motion on the basis of the best ambient measurements. The robot makes step in this direction. Every new measurement offers motion to the nearest goal from all the measurements. So far the robot works in a fee environment, i.e. without obstacle avoidance. It performs full scanning of working space of all distances and gives the full map of investigating field. On the tactic level Navigation System realizes zigzag motion as a result of short term planning. It offers motion in selected way optimal in respect to final goal implementation. The motion is a string of straight lines lying in the comfort zones of the investigated polygon. Strategic level is dedicated to perform compromise decisions of all the tactic point decisions for reaching of priority motion goal. The goal may be sought in curve of motion, shortest path, maximal speed, etc. when we involve obstacles and their optimal avoidance. To fulfill such goal the Navigation System performs compromises deteriorating quality in almost every step of planning. We used techniques similar to dynamic programming. The idea behind it is that optimal path is composed by optimal sub-trajectories connecting final and starting points in reverse direction. For this purpose already found strategic and tactic trajectories are estimated step by step in reverse direction from the goal to start point to define the true compromise trajectory. It is made by the following way: the last step of optimal strategic trajectory is compared with existing nearest around it. Here a compromise assessment of both, taking into account desired optimal one, is made. Since they have different representation a fuzzy systems is due to be involved again for such assessment. It consists of fuzzy rule base that calculates the final trajectory line based on both quality assessments. The quality of strategic steps are defined in terms of tactic one and as an output we obtain required compromise of optimal step. The new step is accepted as representative of calculated trajectory and the procedure is repeated until exhausting of the whole string of trajectory steps to initial point. As you can imagine the robot motion starts with short orientation in ambient environment, which corresponds to the studying and planning of human being in uncertain environment and its following motion.

4.....Case Study

In Figure 8 is shown a prototype of the robot.

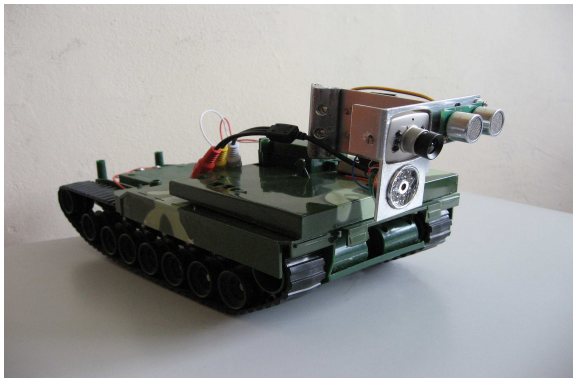


Fig. 8 A robot prototype

An experiment of robot motion in obstacle free polygon is presented in Figure 9.

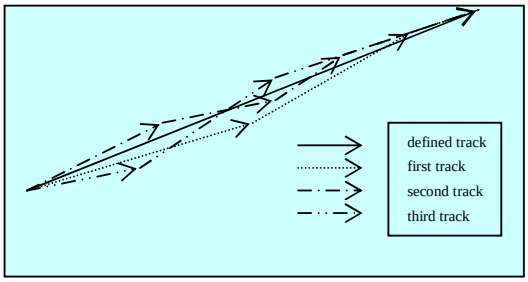


Fig. 9 Experimental robot trajectories

As you can see the differences between defined (ideal track) and combined strategic/tactic motion: the first, second, and third tracks are not significant for this specific application. The reason is that sensor information is gathered in obstacle free polygon, where sensor measurements are taken from the ambient environment outside polygon. Hence a compromise technique of back propagation corrections is realized among less contradictory variants. The situation is drastically different in case of obstacles existence into polygon, where compromises are heavy and there exists also forbidden trajectories. They principally have to be involved in priority cycle for their avoidance and elimination.

4 Discussion

The investigation demonstrates distributed application of relatively simple fuzzy rule base of systems for assessment, fusion, tactic and strategic decision making and control action performance. It is shown a solution of the task for generation of optimal trajectory as a compromise between real admissible by sensor system and ideal postulated as optimal by smoothness, speediness and other considerations. It turns out that there no complication of the task in case of obstacles existence, even in some cases they can simplify decision making as a sequence pruning of some required trajectories and their replacement with only possible ones.

Upcoming tasks:

i. Obviously, some more profound investigations are need in complicated spatial environment. It is very interesting the influence of: kind, type, and number and displacement of obstacles onto speed and quality of performance;

ii. The next task will be incorporation of searching with constraints in our multisensor system, especially suitable for different types of sensors: visual, acoustic, tactile, thermal and inertial;

iii. Improving of synchronization and organization. Up to now they are realized with distributed processing of information of autonomous sensor systems. So far, an improvement of synchronization and organization of the systems is sought in distributed information processing of autonomous units. They are incorporated into control PC for defining and processing of fuzzy variables. So far as the system uses a specialized programmable controllers all these functions can be imbedded in them, which will gain additional autonomy of the robot;

iv. All the while up to now these solutions mimic only Soft Computing Agents for using of single type protocol. As far as information processing is remote and inference systems are created beforehand, there no need they to be capsulated in a kernel of Soft Computing Agent, but future unification of these systems using and their protocols is excellent possibility for realization of simple Soft Computing Agents

v. Involvement of Soft Computing Agents organization would mean full compatibility and applicability of divers systems and different tasks. This is the path of standardization of these new solutions in independent platforms;

vi. Another important direction of investigation is involvement emergency system mimicking emotional tract of human activity, and why not in the conventional brain imitating artificial system of control loops. Even Latin emotional stands for movement, speed, which supposes improvement of the systems in respect to speedy and quality performance of robot reactions. It is referred already to creation of artificial feeling robot systems.

References.

[1] D. Lakov, Soft Computing Agents, Joint 9^{the} IFSA World Congress and 20^{the} NAFIPS International Conference ISA/NAFIPS 2001, July 25-28, 2001, Vancouver, Canada, 2585-2590.

[2] Lee .and P. Wang, Collision Avoidance by Fuzzy Logic Control for Automated Guided Vehicle Navigation, Journal of Robotic Systems, Vol.11 Aug. pp. 745-760, 1994.

[3] R. S. Arkin, Motor Schema-base Mobile Robot Navigation, International journal of robotic research, Vol. 8, 1989.

[4] S. G. Goodridge, R. C. Ruo, and M. G. Kay, Multilayer Fuzzy Behaviour Fusion for Real-time Control of Systems with many Sensors, in Proc. IEEE Intelligent Control of Multisensor Fusion and Integration for Intelligent Systems (MFI), Las Vegas, October 1994, pp. 272-279.

[5] P. Buschka, A. Saffiotti, .and Z. Wasik, Fuzzy Land-mark Localisation for a Ledged Robot, IEEE/RSJ International Conference of Intelligent Robots and Systems (IROS), Takamatsu, Japan, 2001, pp. 1205-1210.

[6] W. Matellan,.S. Fernández, and J. M. Molina, Genetic Learning of Fuzzy Reactive Controllers, ELSEVIER in series Robotics and Autonomous Systems, 1998, Vol. 25, pp. 33-41.

[7] G. Hailu, J. Bruske, and G. Sommer, Fuzzy Logic Control of Situated Agent, IFSA'97, Prague, Czech, June 25-29, 1997, pp. 496-500.

[8] D. J. Cook, Using Analytic and Genetic Methods to Learn Plans for Mobile Robots, in Proc. of the SPIE Conference on the Applications of Artificial Intelligence, 1993 pp. 327-336.

[9] F. G. Pin, Y. Watanabe, Driving a Car Using Reflexive Fuzzy Behaviours, in Proc. of the Second International Conference on Fuzzy Systems, San Francisco, USA 1993.

Session 3

Fuzzy Rules

How to use controller with conditionally firing rules

Mirko Navara
Center for Machine Perception,
Department of Cybernetics
Faculty of Electrical Engineering,
Czech Technical University
Technická 2,
CZ-166 27 Praha, Czech Republic
navara@cmp.felk.cvut.cz

Milan Petrík
Center for Machine Perception,
Department of Cybernetics
Faculty of Electrical Engineering,
Czech Technical University
Technická 2,
CZ-166 27 Praha, Czech Republic
petrikm@cmp.felk.cvut.cz

Abstract

In previous work, a fuzzy controller with conditionally firing rules was suggested as an advantageous alternative to Mamdani–Assilian controllers. Its features have been used without much understanding. Therefore we explicitly formulate the rules of its design in this tutorial which has not been available before.

Keywords: fuzzy controller, compositional rule of inference, firing rule, fuzzy relational equation, fuzzy hardware.

1 Introduction

A fuzzy controller with conditionally firing rules (CFR controller) was introduced in [10]. It is a generalization of the Mamdani–Assilian controller, using an inference rule that is not compositional. It was motivated by the effort to fit the behavior of the controller optimally to the meaning of the fuzzy rule base. Both theoretical and practical arguments support the expectation that the CFR controller allows a better performance without a change of the rule base. Thus it may improve the quality of control while keeping the order of computational complexity. It has been applied in several tasks [6, 8, 11, 13, 14], but the users implemented it sometimes in a simple manner, without much understanding its options and capabilities. Therefore we found it necessary to write a "user's manual" for this type of controller which we present here.

2 Theoretical analysis of fuzzy controllers

For basics on fuzzy controllers, we refer to [4, 5]. Let $\mathcal{X}$ and $\mathcal{Y}$ denote the input and the output space of a controller, respectively. They are supposed to be convex subsets of finite-dimensional real vector spaces. For a set $\mathcal{Z}$, we denote by $\mathcal{F}(\mathcal{Z})$ the set of all its fuzzy subsets. Throughout this paper, we identify fuzzy sets with their membership functions. The expert's knowledge may be expressed by a base Θ of rules of the form

$$\textbf{if } x \in A_i \textbf{ then } y \in C_i\,,$$

where $A_i \in \mathcal{F}(\mathcal{X})$ are *antecedents* and $C_i \in \mathcal{F}(\mathcal{Y})$ are *consequents*, $i \in \{1, \dots, n\}$ (see [4]). Following [15], the knowledge from the rule base Θ can be represented by a fuzzy relation $R_\Theta \in \mathcal{F}(\mathcal{X} \times \mathcal{Y})$. The Mamdani–Assilian controller [7] uses R_Θ in the form

$$R_\Theta(x, y) = \max_{i \le n} T(A_i(x), C_i(y))\,, \tag{1}$$

where T is a fixed t-norm modelling a fuzzy conjunction [5]. Then we apply a *compositional rule of inference* which assigns to a fuzzy input $X \in \mathcal{F}(\mathcal{X})$ a fuzzy output $Y \in \mathcal{F}(\mathcal{Y})$ by $Y = X \underset{T}{\circ} R_\Theta$, i.e.,

$$Y(y) = \sup_{x \in \mathcal{X}} T(X(x), R_\Theta(x, y))\,. \tag{2}$$

We denote the induced mapping by $\Phi_\Theta^{\mathrm{MA}}\colon \mathcal{F}(\mathcal{X}) \to \mathcal{F}(\mathcal{Y})$. Combining formulas (1), (2), it can be expressed as

$$\begin{aligned} Y(y) &= \Phi_\Theta^{\mathrm{MA}}(X)(y) \\ &= \sup_{x \in \mathcal{X}} T\Big(X(x), \max_{i \le n} T(A_i(x), C_i(y))\Big) \\ &= \max_{i \le n} T\Big(\sup_{x \in \mathcal{X}} T(X(x), A_i(x)), C_i(y)\Big)\,. \end{aligned}$$

The latter form allows for an effective calculation using the values

$$\mathcal{D}_T(X, A_i) = \sup_{x \in \mathcal{X}} T(X(x), A_i(x))\,, \tag{3}$$

called *degrees of overlapping* of X, A_i. The value $\mathcal{D}_T(X, A_i)$ is the height of the fuzzy intersection $X \underset{T}{\cap} A_i$. The output is

$$Y(y) = \max_{i \le n} T\Big(\mathcal{D}_T(X, A_i), C_i(y)\Big)\,. \tag{4}$$

The use of degrees of overlapping allows to avoid nested cycles over the spaces $\mathcal{X}$, $\mathcal{Y}$. Notice that the simplification is possible due to the associativity of T; it requires to use the same t-norm T in (1) and (2).

3 Requirements on fuzzy controllers

To formulate mathematically the desirable properties of a fuzzy controller, we have to introduce several basic notions and notations from fuzzy set theory. A fuzzy set is called *normal* if its membership function attains the value 1 at (at least) some point. The *support* of a fuzzy set $A \in \mathcal{F}(\mathcal{Z})$ is

$$\operatorname{Supp} A = \{z \in \mathcal{Z} : A(z) > 0\}\,.$$

For $x \in \mathcal{Z}$, let $\chi_x \in \mathcal{F}(\mathcal{Z})$ denote a singleton (crisp value), i.e.,

$$\chi_x(z) = \begin{cases} 1 & \text{if } z = x\,, \\ 0 & \text{if } z \neq x \end{cases}$$

for all $z \in \mathcal{Z}$. A fuzzy set is called *convex* if all its α-cuts are convex sets. For fuzzy sets $A_1, \dots, A_n \in \mathcal{F}(\mathcal{Z})$, their *convex hull* is the smallest (w.r.t. the pointwise ordering) convex fuzzy set $C \in \mathcal{F}(\mathcal{Z})$ satisfying $A_i(z) \le C(z)$ for all $z \in \mathcal{Z}$, $i = 1, \dots, n$.

Let $\Theta = (A_i, C_i)_{i=1}^n$ be a rule base. We want the control function, $\Phi_\Theta\colon \mathcal{F}(\mathcal{X}) \to \mathcal{F}(\mathcal{Y})$, of a fuzzy controller to satisfy the following properties:

[I1] If the input coincides with the ith antecedent, then the output coincides with the ith consequent, i.e., $\Phi_\Theta(A_i) = C_i$.

[I2] For each normal input $X \in \mathcal{F}(\mathcal{X})$, the output $\Phi_\Theta(X)$ is not contained in all consequents, i.e., $\Phi_\Theta(X) \not\le \min\limits_{j\le n} C_j$.

[I3] The output $\Phi_\Theta(X)$ belongs to the convex hull of consequents C_i of all rules such that $\operatorname{Supp} A_i \cap \operatorname{Supp} X \neq \emptyset$ (i.e., all firing rules).

[I4] Let $X = \chi_x$ for $x \in \mathcal{X}$ such that $A_i(x) = 1$ (for a fixed i). Then $\Phi_\Theta(X) = C_i$.

Condition [I2] avoids non-significant outputs bearing no information about the input. It compares the output with the fuzzy intersection of all consequents. Usually this intersection is empty, then [I2] simply means that the output is non-empty. Our formulation gives the respective meaning also in the (rare) case when the intersection of all consequents in non-empty.

Condition [I3] is a very weak requirement meaning that the output is obtained by *interpolation* of the consequents.

Condition [I4] concerns the (not so unusual) case when there is a totally firing rule for a crisp input. Then the output should be the corresponding consequent. This means that having a well-fitting rule, we do not need any other (which could only damage the correct output). As a consequence, we have to avoid a rule base in which this could occur simultaneously for more than one rule.

Condition [I1] deserves more attention. It is called *interaction* in [4]. It says that an antecedent as an input produces the corresponding consequent as an output. For a Mamdani–Assilian controller it corresponds to a system of fuzzy relational equations

$$C_i = A_i \underset{T}{\circ} R_\Theta\,, \qquad i \in \{1, \dots, n\}\,. \tag{5}$$

Necessary and sufficient conditions for its solvability by a Mamdani–Assilian controller are given in [3] (cf. also [12]). To formulate them, we use the *residuum* (*residuated implication*) induced by T, i.e., the operation $I_T\colon [0,1] \times [0,1] \to [0,1]$ such that

$$I_T(a,b) = \sup\{c \in [0,1] : T(a,c) \le b\}\,.$$

If T is continuous, we may take max instead of sup in the latter formula. The corresponding *biimplication* is

$$E_T(a,b) = \min\{I_T(a,b), I_T(b,a)\}\,.$$

For two consequents, C_i, C_j, we define their *degree of indistinguishability*, $\mathcal{E}_T(C_i, C_j)$, by

$$\mathcal{E}_T(C_i, C_j) = \inf_{z\in\mathcal{Z}} E_T(C_i(z), C_j(z))\,.$$

Recall that the degree of indistinguishability $\mathcal{E}_T\colon \mathcal{F}(\mathcal{Y}) \times \mathcal{F}(\mathcal{Y}) \to [0,1]$ is a fuzzy equality relation (similarity) on $\mathcal{F}(\mathcal{Y})$ with respect to T.

Theorem 1 [3] *Let $\Theta = (A_i, C_i)_{i=1}^n$ be a rule base with $A_i(x_i) = 1$ for pairwise different elements $x_1, \dots, x_n \in \mathcal{X}$. Then relation (1) satisfies the system (5) iff all $i, j \in \{1, \dots, n\}$ satisfy*

$$\mathcal{D}_T(A_i, A_j) \le \mathcal{E}_T(C_i, C_j)\,. \tag{6}$$

Condition (6) (and hence also [I1]) is violated in many fuzzy controllers. It depends on the choice of the t-norm T; choosing a smaller t-norm (e.g., the product instead of the minimum) could help to satisfy it. Nevertheless, this choice makes it more difficult to satisfy condition [I2]. Thus conditions [I1], [I2] are almost contradictory for a Mamdani–Assilian controller and typical shapes of membership functions. (See [2, 9, 10] for more detailed analysis of this phenomenon.)

4 CFR controller and its design

In [10], a controller is suggested which allows to satisfy the above requirements for rather general rule bases. It assumes the following properties of the rule base $\Theta = (A_i, C_i)_{i=1}^n$:

[C1] *"normality"*: each A_i is normal,

[C2] *"covering of antecedents"*: $\inf_{x\in\mathcal{X}} \max_{i\le n} A_i(x) > 0$,

[C3] *"significance of consequents"*: $\forall i \in \{1,\dots,n\}$: $C_i \not\le \min_{j\neq i} C_j$.

Conditions [C1] and [C2] ensure that we have normal antecedents covering the whole input space and [C3] is quite a weak requirement on significance of consequents. See [2, 9] for more detailed analysis of these conditions.

The new controller is obtained from the Mamdani–Assilian controller by the following modifications:

Step 1. Change the membership degrees in the input space by an increasing bijection $\varrho\colon [0,1] \to [0,1]$. This changes the degrees of overlapping $\mathcal{D}_T(X, A_i)$ to

$$\mathcal{D}_T(X\circ\varrho, A_i\circ\varrho) = \sup_{x\in\mathcal{X}} T(\varrho(X(x)), \varrho(A_i(x)))\,. \quad (7)$$

To keep the numerical complexity of the changes of scale as low as possible, we may choose $\varrho(t) = t^r$, $r \in N$, or a piecewise linear function. Choosing $\varrho(t) \le t$, we may reduce the degrees of overlapping to an arbitrary small value (provided that they were < 1).

Step 1 can be skipped (choosing the identity for ϱ) unless the degrees of overlapping $\mathcal{D}_T(A_i\circ\varrho, A_j\circ\varrho)$ are too large for some i, j, $i \neq j$.

Step 2. Choose $c < 1$ such that

$$c \ge \mathcal{D}_T(A_i\circ\varrho, A_j\circ\varrho) \quad (8)$$

for all i, j, $i \neq j$. Change the membership degrees in the output space by an increasing bijection $\sigma\colon [0,1] \to [c,1]$. To extend the inverse of σ to the whole interval $[0,1]$, we use its *pseudoinverse*

$$\sigma^{[-1]}(t) = \begin{cases} \sigma^{-1}(t) & \text{if } t \ge c\,, \\ 0 & \text{otherwise}\,. \end{cases} \quad (9)$$

We may choose for example $\sigma(t) = (1-c)\cdot t + c$.

Step 3. Replace the degree of overlapping in (4) with the *degree of conditional firing* of the ith rule defined as

$$\mathcal{C}_{T,i}(X) = \frac{\mathcal{D}_T(X\circ\varrho, A_i\circ\varrho)}{\max_{j\le n} \mathcal{D}_T(X\circ\varrho, A_j\circ\varrho)}\,. \quad (10)$$

(In fact, the value $\mathcal{C}_{T,i}(X)$ depends not only on T and i, but also on ϱ and on all antecedents A_j, $j = 1,\dots,n$. For each normal input X, the values $\mathcal{C}_{T,i}(X)$, $i = 1,\dots,n$, belong to $[0,1]$ and at least one of them is 1.)

Thus the rules are activated according to their *relative* degree of overlapping with the input.

After all these three changes, the output becomes

$$\begin{aligned} Y(y) &= \Phi_\Theta^{\mathrm{CFR}}(X)(y) \\ &= \sigma^{[-1]}\Big(\max_{i\le n} T\big(\mathcal{C}_{T,i}(X), \sigma(C_i(y))\big)\Big)\,, \end{aligned} \quad (11)$$

where $\Phi_\Theta^{\mathrm{CFR}}\colon \mathcal{F}(\mathcal{X}) \to \mathcal{F}(\mathcal{Y})$ denotes the induced input-output correspondence of this modified controller, called a *controller with conditionally firing rules*, or briefly *CFR controller*.

We do not deal here with defuzzification. Our basic observations refer to fuzzy subsets of the input and output space and they are independent of the method of defuzzification.

Theorem 2 [10] *Let a rule base Θ satisfy conditions* [Con1–3]. *Then the CFR controller given by* (11) *satisfies properties* [I1–4] *(for any increasing bijection $\varrho\colon [0,1] \to [0,1]$, any $c < 1$ satisfying (8), and any increasing bijection $\sigma\colon [0,1] \to [c,1]$).*

Condition (8) represents disjointness of antecedents needed to distinguish between firing rules (not more than one rule may fire totally as this would cause a contradiction in the rule base).

The weakening of conditions on rule base enabled by a CFR controller is mainly due to the following fact:

Theorem 3 [10] *Let $\Theta = (A_i, C_i)_{i=1}^n$ be a rule base with $A_i(x_i) = 1$ for pairwise different elements $x_1,\dots,x_n \in \mathcal{X}$. Then the input–output correspondence Φ_Θ^{CFR} of a CFR controller (with mappings ϱ, σ described above) satisfies* [I1] *iff all $i, j \in \{1,\dots,n\}$ satisfy*

$$\mathcal{D}_T(A_i\circ\varrho, A_j\circ\varrho) \le \mathcal{E}_T(C_i\circ\sigma, C_j\circ\sigma)\,. \quad (12)$$

Under the conditions of Theorem 2, we may choose the mapping $\varrho\colon [0,1] \to [0,1]$ so that the left-hand side of (12) becomes an arbitrarily small positive number; the lower bound c of the range of the mapping $\sigma\colon [0,1] \to [c,1]$ gives a lower estimate of the right-hand side of (12), so inequality (12) can be satisfied.

5 Simplification of CFR controller for crisp inputs

In general, it is not possible to substitute a CFR controller in its full generality by a Mamdani–Assilian

controller. Nevertheless, in the most important case of crisp inputs it admits such a modification [1].

A crisp input means that the values taken into account are represented by real numbers. Still they may be subject to some imprecision or uncertainty, but despite of this, the result is one value, not a fuzzy set of values. This is true for most measuring devices. In fact, it is very rare and difficult to measure a quantity in such a way that the result is given as a fuzzy set. Therefore a reduction to crisp inputs is assumed in most fuzzy controllers. Then fuzziness is reduced to the computations *inside* the controller, while it communicates with the world through crisp inputs and outputs. This approach is still general enough to be the basis of most applications of fuzzy control. In fact, this is what the current fuzzy hardware does; it handles only crisp inputs.

First we shall show how the controllers simplify if the input is crisp, i.e., $X = \chi_x$ for some $x \in \mathcal{X}$. Then the degrees of overlapping of X, A_i become

$$\mathcal{D}_T(X, A_i) = \sup_{x\in\mathcal{X}} T(X(x), A_i(x)) = A_i(x) \qquad (13)$$

(independently of the choice of the t-norm T). Let us denote by $\varphi_\Theta^{\mathrm{MA}}(x)$ the output $\Phi_\Theta^{\mathrm{MA}}(X)$ of a Mamdani–Assilian controller corresponding to a crisp input $X = \chi_x$. Thus $\varphi_\Theta^{\mathrm{MA}}\colon \mathcal{X} \to \mathcal{F}(\mathcal{Y})$ is the mapping fully describing the action of the Mamdani–Assilian controller on crisp inputs. We obtain

$$Y(y) = \varphi_\Theta^{\mathrm{MA}}(x)(y) = \max_{i\le n} T(A_i(x), C_i(y))\,. \qquad (14)$$

What remains is to defuzzify the fuzzy output Y.

In case of a CFR controller, we obtain

$$\begin{aligned}\mathcal{D}_T(X\circ\varrho, A_i\circ\varrho) &= \sup_{x\in\mathcal{X}} T(\varrho(X(x)), \varrho(A_i(x)))\\ &= \varrho(A_i(x))\,.\end{aligned}$$

(Notice that the membership degrees $X(x) \in \{0,1\}$ remain unchanged by ϱ.) The degrees of conditional firing are

$$\mathcal{C}_{T,i}(X) = \frac{\mathcal{D}_T(X\circ\varrho, A_i\circ\varrho)}{\max\limits_{j\le n}\mathcal{D}_T(X\circ\varrho, A_j\circ\varrho)} = \frac{\varrho(A_i(x))}{\max\limits_{j\le n}\varrho(A_j(x))}\,. \qquad (15)$$

These values depend only on $A_i\circ\varrho$, $i \in \{1,\dots,n\}$, which are known in advance (once the design of the controller is completed), and on x. Thus they can be represented internally by functions, similarly to A_i, $i \in \{1,\dots,n\}$. Explicitly, we define fuzzy sets (modified antecedents) $\overline{A}_i \in \mathcal{F}(\mathcal{X})$, $i \in \{1,\dots,n\}$, by

$$\overline{A}_i(x) = \frac{\varrho(A_i(x))}{\max\limits_{j\le n}\varrho(A_j(x))}\,. \qquad (16)$$

Under the assumption [C2], they are well defined (the denominator is nonzero). Then the desired degree of conditional firing may be expressed as

$$\mathcal{C}_{T,i}(X) = \overline{A}_i(x)\,. \qquad (17)$$

The transformation σ may be analogously represented by a modification of consequents. Explicitly, we define fuzzy sets (modified consequents) $\overline{C}_i \in \mathcal{F}(\mathcal{Y})$, $i \in \{1,\dots,n\}$, by

$$\overline{C}_i(y) = \sigma(C_i(y))\,. \qquad (18)$$

Let us denote by $\varphi_\Theta^{\mathrm{CFR}}(x)$ the output $\Phi_\Theta^{\mathrm{CFR}}(X)$ of a CFR controller corresponding to a crisp input $X = \chi_x$. Thus $\varphi_\Theta^{\mathrm{CFR}}\colon \mathcal{X} \to \mathcal{F}(\mathcal{Y})$ is the mapping fully describing the action of the CFR controller on crisp inputs. With the above modifications, we obtain

$$Y(y) = \varphi_\Theta^{\mathrm{CFR}}(x)(y) = \sigma^{[-1]}\Big(\max_{i\le n} T\big(\overline{A}_i(x), \overline{C}_i(y)\big)\Big)\,. \qquad (19)$$

Except for the final transformation $\sigma^{[-1]}$, this formula is of the same form as formula (14) for the Mamdani–Assilian controller. More exactly, if $\overline{\Theta} = (\overline{A}_i, \overline{C}_i)_{i=1}^n$ is the modified rule base and $\varphi_{\overline{\Theta}}^{\mathrm{MA}}$ is the respective input-output correspondence, then

$$\varphi_\Theta^{\mathrm{CFR}}(x) = \varphi_{\overline{\Theta}}^{\mathrm{MA}}(x)\circ\sigma^{[-1]}\,.$$

This proves the following result:

Theorem 4 *Suppose that we have a controller with conditionally firing rules with a rule base Θ which performs a mapping $\varphi_\Theta^{CFR}\colon \mathcal{X} \to \mathcal{F}(\mathcal{Y})$. Let $\sigma\colon [0,1] \to [c,1]$ be the mapping used in (11). Then there is a Mamdani–Assilian controller with a rule base $\overline{\Theta}$ which performs a mapping $\varphi_{\overline{\Theta}}^{MA}\colon \mathcal{X} \to \mathcal{F}(\mathcal{Y})$ such that $\varphi_\Theta^{CFR}(x) = \varphi_{\overline{\Theta}}^{MA}(x)\circ\sigma^{[-1]}$ for all $x \in \mathcal{X}$, where $\sigma^{[-1]}$ is the mapping given by (9).*

This means that the implementation of a CFR controller with crisp inputs can be done using the standard Mamdani–Assilian controller and a block performing the transformation $\sigma^{[-1]}$. As a consequence, hardware for Mamdani–Assilian controllers can be used to implement a CFR controller.

6 Example

In this section we present examples demonstrating the use of the CFR controller. The first example shows how the controller computes the output value for a given set of antecedents, consequents, and an input value. The second example gives the output of the Mamdani–Assilian controller with the same antecedents, consequents, and rule base. A comparison

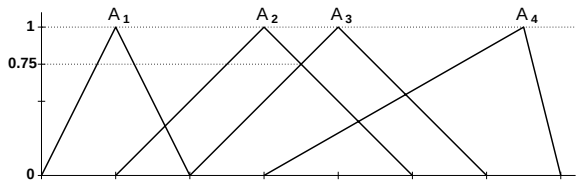


Figure 1: Example I: Fuzzy sets defining the antecedents.

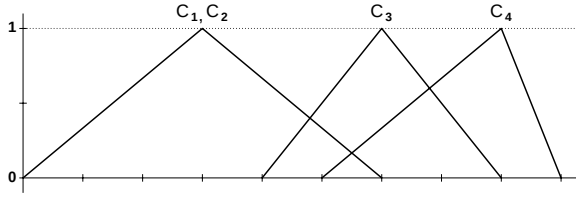


Figure 2: Example I: Fuzzy sets defining the consequents.

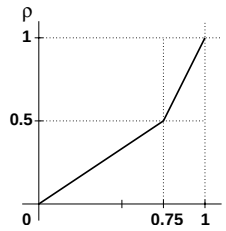


Figure 3: Example I: Graph of function ϱ.

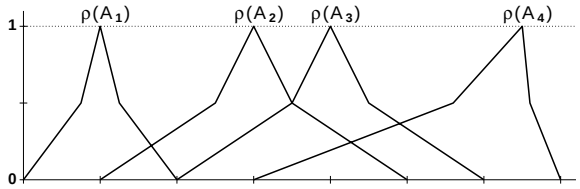


Figure 4: Example I: Antecedents modified by function ϱ.

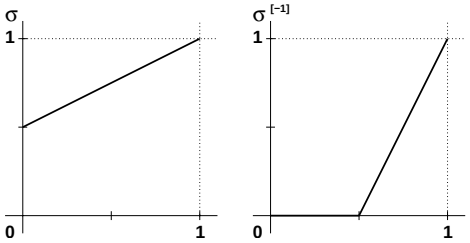


Figure 5: Example I: Graphs of functions σ and $\sigma^{[-1]}$.

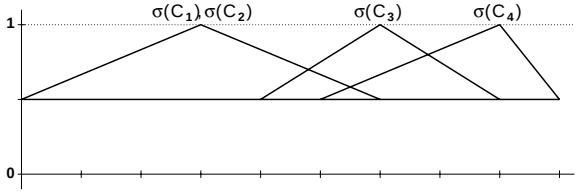


Figure 6: Example I and III: Consequents modified by function σ.

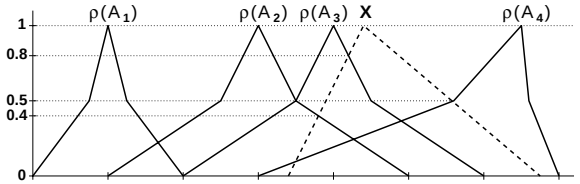


Figure 7: Example I: An input of the CFR controller represented by a fuzzy set X.

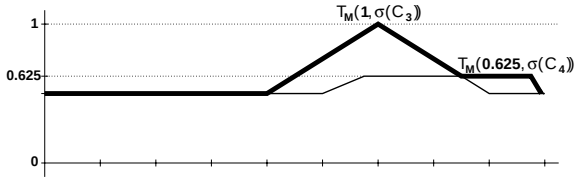


Figure 8: Example I: Consequents modified by function σ and combined with the degrees of conditional firing.

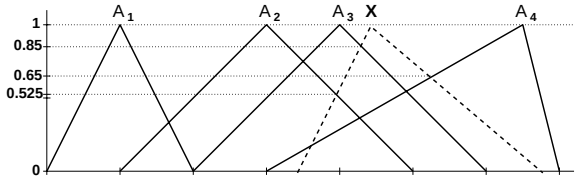


Figure 9: Example II: Antecedents, an input, and degrees of overlapping of the Mamdani–Assilian controller.

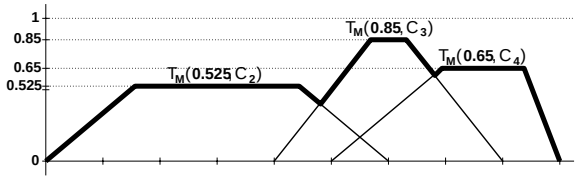


Figure 10: Example II: Consequents of the Mamdani–Assilian controller combined with the degrees of overlapping.

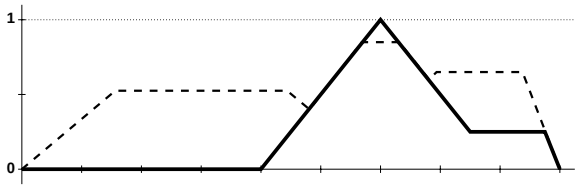


Figure 11: Example I and II: Resulting output fuzzy sets. The CFR controller is represented by the thick line; the Mamdani–Assilian controller is represented by the dashed line.

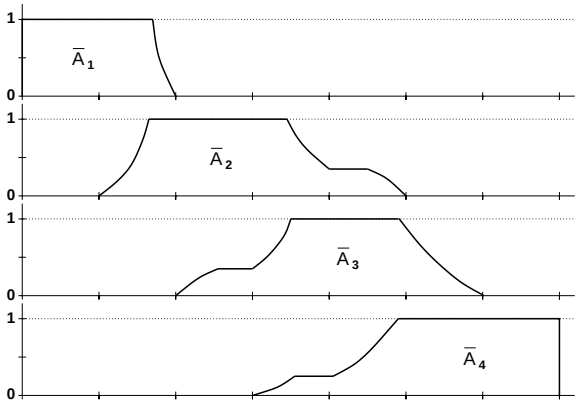


Figure 12: Example III: Antecedents of the derived Mamdani–Assilian controller.

of these two results is then made. The third example describes the transformation of a CFR controller into a Mamdani–Assilian controller for the case when the input can attain only crisp values. For simplicity, we demonstrate the principles in one dimension; the extension to more dimensions is easy.

6.1 Example I

We have a set of antecedents and consequents as given in Figure 1 and Figure 2. Further we have a set of fuzzy rules in the form:

1. **if** $x \in A_1$ **then** $y \in C_1$,
2. **if** $x \in A_2$ **then** $y \in C_2$,
3. **if** $x \in A_3$ **then** $y \in C_3$,
4. **if** $x \in A_4$ **then** $y \in C_4$,

To implement a CFR controller we have decided to use the standard (Gödel, minimum, ...) t-norm defined as $T_M(x, y) = \min\{x, y\}$. Now we need to find some suitable realizations for the functions ϱ and σ.

The function σ attains values from the interval $[c, 1]$ where $c \in (0, 1)$ is a number higher or equal to the highest degree of overlapping of the antecedents. Looking at Figure 1 we can see that the highest degree of overlapping is 0.75. Yet this value is too high; operating only in the interval $[0.75, 1]$ we could lose some precision. We may decide to have the constant c at most 0.5. For this, we choose the function ϱ as in Figure 3. Now the antecedents modified by ϱ look as shown in Figure 4, the highest degree of overlapping is 0.5, and we can choose $c = 0.5$. The functions σ and $\sigma^{[-1]}$ then may look as shown in Figure 5 and Figure 6 shows the consequents modified by σ.

Let us suppose that the input of the controller is a fuzzy set X from Figure 7. We are going to compute the output value.

First we determine the degrees of overlapping given by the fuzzy set X and the antecedents. The degrees are determined according to Figure 7 and they are listed in the fist row of Table 1. The highest degree of overlapping is 0.8 in our case. To compute the degrees of conditional firing, we divide each degree of overlapping by the highest degree, i.e. by 0.8.

	A_1	A_2	A_3	A_4
$\mathcal{D}(X)$	0	0.4	0.8	0.5
$\mathcal{C}(X)$	0	0.5	1	0.625

Table 1: Example I: Degrees of overlapping and degrees of conditional firing for the antecedents A_1, A_2, A_3, and A_4.

Having computed the degrees of conditional firing, we combine them with the consequents by the minimum t-norm, see Figure 8. (The first two consequents after the modification become less than 0.5 and they are omitted because they do not influence the output.) From these modified fuzzy sets the maximum is taken (shown in Figure 8 by a thick line) and this maximum is modified by the function $\sigma^{[-1]}$. This way, we obtain the resulting output fuzzy set shown in Figure 11 by a thick line. (It is subject to defuzzification which is not discussed here.)

Notice that although the degree of overlapping of the antecedent A_2, resp. A_4 was rather high, the influence of the corresponding consequent to the final result was none, resp. small.

6.2 Example II

For a comparison, let us now create a classical Mamdani–Assilian controller with the same antecedents, consequents, and rule base as in the first example. Let us suppose that the input of the controller is the same fuzzy set X (see Figure 9). We are going to compute the output value.

We determine the degrees of overlapping, given by the fuzzy set X and the antecedents, as can be seen in Figure 9; the degrees are listed in Table 2. We combine the degrees of overlapping with the consequents by the minimum t-norm, see Figure 10. (The first consequent after the modification becomes zero and it is omitted.) By taking maximum from these modified fuzzy sets (shown in Figure 10 by a thick line) we obtain the resulting output fuzzy set shown in Figure 11 by a dashed line.

	A_1	A_2	A_3	A_4
$\mathcal{D}(X)$	0	0.525	0.85	0.65

Table 2: Example II: Degrees of overlapping for the antecedents A_1, A_2, A_3, and A_4 in the case of the Mamdani–Assilian controller.

Now, in Figure 11, we can see the output fuzzy sets both for the CFR controller (thick line) and for the Mamdani–Assilian controller (dashed line) for the same antecedents, consequents, and rule base. As we can see the CFR controller gives a very specific result saying that the output is given mostly by the third consequent slightly influenced by the fourth consequent. On the contrary, the output of the corresponding Mamdani–Assilian controller is much more uncertain and blurred. The influence of all the three nonzero consequents is considerably high which makes the interpretation and the defuzzification more difficult.

6.3 Example III

In this example we are going to trasform a CFR controller to a Mamdani–Assilian controller according to Section 5. In general, it is not possible to substitute a CFR controller in its full generality by a Mamdani–Assilian controller. Nevertheless, in the most important case of crisp inputs it admits such a modification [1].

Note that this Mamdani–Assilian controller will work only with crisp input values. Let us take the same set of antecedents and consequents and the same rule base as in the first example. First we use (16) and modify the set of antecedents as it can be seen in Figure 12. Then, according to (18), we modify the set of consequents as it is shown in Figure 6. The rule base remains unchanged.

This way we have obtained new sets of antecedents and consequents. We can now apply the algorithm of the Mamdani–Assilian controller to the modified rule base. Then the mapping $\sigma^{[-1]}$ is applied to the output. Thus we obtain, for crisp inputs, the same behaviour as in the original CFR controller. It satisfies our requirements [I1–4] with respect to the *original* antecedents and consequents.

7 Discussion

Let us now discuss what Th. 2 says for the case of crisp input. (With the exception of [I1], the conclusions are valid for both the original formulation of a CFR controller and its reformulation described above—they are equivalent in this case.) Conditions [C1–4] remain unchanged. Property [I1] becomes meaningless, because we do not accept fuzzy inputs (the antecedents are usually fuzzy sets, not crisp singletons). Property [I2] simplifies, because crisp inputs are always normal (when considered as a special case of fuzzy sets). The new formulation can be:

[$\overline{\text{I2}}$] For each crisp input $x \in \mathcal{X}$, the output $\varphi_\Theta^{\mathrm{CFR}}(x)$ is not contained in all consequents, i.e., $\varphi_\Theta^{\mathrm{CFR}}(x) \not\leq \min_{j\leq n} C_j$.

Property [I3] can be also simplified:

[$\overline{\text{I3}}$] For each crisp input $x \in \mathcal{X}$, the output $\varphi_\Theta^{\mathrm{CFR}}(x)$ belongs to the convex hull of consequents C_i of all rules such that $x \in \mathrm{Supp}\,(A_i)$ (i.e., all firing rules).

Property [I4] remains unchanged.

Preprocessing means that instead of the original rule base Θ we have to use a *modified* rule base $\overline{\Theta} = (\overline{A}_i, \overline{C}_i)_{i=1}^n$ given by (16), (18). This rule base $\overline{\Theta}$ has to be downloaded into the Mamdani–Assilian controller. Nevertheless, the input-output correspondence satisfies the desired properties [I1–3] with respect to the *original* rule base $\Theta = (A_i, C_i)_{i=1}^n$. It is usually impossible to obtain this behavior by application of a Mamdani–Assilian controller directly to the rule base Θ.

8 Conclusion

Both theoretical motivation and practical experiments suggest that the controller with conditionally firing rules [10] allows us to enhance the performance in comparison to a Mamdani–Assilian controller without a change of the rule base. Its software implementation requires just adding three easy transformations. These require additional calculations, but do not extend the order of complexity.

Our treatment applies to an arbitrary finite dimension of the input and output spaces. This approach is now ready for further tests on practical control tasks.

Acknowledgement

The first author was supported by The Czech Government under the reseach program MSM6840770038. The second author was supported by the Czech Science Foundation under project 201/07/1136.

References

[1] P. Amato, A. Di Nola, M. Navara, Reformulation of fuzzy controller with conditionally firing rules, In: M. Mohammadian (Ed.) *Proc. Int. Conf. Computational Intelligence for Modelling, Control and Automation*, Vienna, Austria, 2003, 140–151.

[2] P. Amato, A. Di Nola, M. Navara, Mathematical aspects of fuzzy control, in: V. Gesù, F. Masulli, A. Petrosino (Eds.), *Fuzzy Logic and Applications: 5th Int. Workshop, WILF 2003*, Lecture Notes in Computer Science **2955**, Springer-Verlag, 2006, 80–88.

[3] B. De Baets, A note on Mamdani controllers, in: D. Ruan, P. D'hondt, P. Govaerts, E. Kerre (Eds.), *Intelligent Systems and Soft Computing for Nuclear Science and Industry*, World Scientific Publishing, Singapore, 1996, 22–28.

[4] D. Driankov, H. Hellendoorn, M. Reinfrank, *An Introduction to Fuzzy Control*, Springer, Heidelberg, 1993.

[5] S. Gottwald, *Fuzzy Sets and Fuzzy Logic*, Vieweg, Braunschweig, 1993.

[6] Z. Huang, Q. Shen, Fuzzy interpolative reasoning via scale and move transformations, *IEEE Trans. Fuzzy Syst.* **14** (2006) 340–359.

[7] E. H. Mamdani, S. Assilian, An experiment in linguistic synthesis of fuzzy controllers, *Int. J. Man-Mach. Stud.*, **7** (1975) 1–13.

[8] C. Manara, P. Amato, A. di Nola, M. Linawaty, I. Pedaci, Software implementation of fuzzy controller with conditionally firing rules, and experimental comparisons, in: *Fuzzy Logic and Applications*, (Lecture Notes in Artificial Intelligence, Vol. **3849**), 2006, 205–214.

[9] B. Moser, M. Navara, Which triangular norms are convenient for fuzzy controllers? *Proc. EUSFLAT-ESTYLF Joint Conf. 99*, Universitat de les Illes Balears, Palma (Mallorca), 1999, 75–78.

[10] B. Moser, M. Navara, Fuzzy controllers with conditionally firing rules, *IEEE Trans. Fuzzy Systems* **10** (2002) 340–348.

[11] M. Navara, D. Peri, Automatic generation of fuzzy rules and its applications in medical diagnosis, *Proc. 10th Int. Conf. Information Processing and Management of Uncertainty*, Perugia, Italy, Vol. 1, 2004, 657–663.

[12] I. Perfilieva, A. Tonis, Criterion of solvability of fuzzy relational equations system, *Proc. Congress IFSA 97*, Praha, 1997, 90–95.

[13] J. Singh, N. Singh, J.K. Sharma, Fuzzy modeling and control of HVAC systems - A review, *J. Sci. Ind. Res.* **65** (2006) 470–476.

[14] M. Takács, Residuum-based approximate reasoning with distance-based uninorms, in: E.P. Klement and E. Pap (Eds.), *Mathematics of Fuzzy Systems*, Univ. Linz, Austria, 2004, 209–212.

[15] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Syst. Man Cybern.* **3** (1973) 28–44.

On Additive and Multiplicative Fuzzy Models*

Martin Štěpnička
Institute for Research and Applications of Fuzzy Modeling
University of Ostrava
30. dubna 22, 701 03 Ostrava 1
Czech Republic
Martin.Stepnicka@osu.cz

Bernard De Baets
Department of Applied Mathematics, Biometrics, and Process Control
Ghent University
Coupure links 653, B-9000 Gent
Belgium
Bernard.DeBaets@UGent.be

Lenka Nosková
Institute for Research and Applications of Fuzzy Modeling
University of Ostrava
30. dubna 22, 701 03 Ostrava 1
Czech Republic
Lenka.Noskova@osu.cz

Abstract

Systems which use a fuzzy rule base and an inference mechanisms are quite frequently used in many applications. Fuzzy rules and inference mechanisms can be described by a system of fuzzy relation equations. A solution to a given system of fuzzy relation equations can serve us a proper model of fuzzy rules (fuzzy model for short). But only two particular solutions, let us call them disjunctive and conjunctive fuzzy models, have been mainly studied so far.

The paper aims at new solutions of systems of fuzzy relation equations. The first one - additive fuzzy model - which is derived from the disjunctive fuzzy model was motivated by several methods often used in practice. The second one — multiplicative fuzzy model — which is derived from the conjunctive fuzzy model was neglected up to now. The paper investigates both new solutions and describes existence of a direct connection between systems of fuzzy relation equations and different fuzzy models. The crucial role of the Ruspini partition is proved and also visually demonstrated.

Keywords: Systems of Fuzzy Relation Equations, Fuzzy Models, Fuzzy Rules, Inference, Interpolation.

1 Introduction

Fuzzy rule based (FRB) systems have been many times demonstrated to be powerful tools in modelling, decision making or automatic control. They basically consist of two main blocks: a fuzzy rule base determined by a set of n fuzzy rules and an inference mechanism. A correct choice of a fuzzy relation which models a fuzzy rule base and a corresponding inference mechanism is crucial for a proper performance of the whole FRB.

Consider arbitrary universes X and Y. The classes of fuzzy sets on X and Y are denoted by $\mathcal{F}(X)$ and $\mathcal{F}(Y)$, respectively. Then all information available in a given fuzzy rule base is contained in pairs of input-output fuzzy sets i.e. in $(\mathbf{A}_1, \mathbf{B}_1), \ldots, (\mathbf{A}_n, \mathbf{B}_n)$, telling us that, for $i = 1, ..., n$, the fuzzy set $\mathbf{A}_i \in \mathcal{F}(X)$ is assigned the fuzzy set $\mathbf{B}_i \in \mathcal{F}(Y)$ [31].

In general, there are two standard approaches to modelling a given fuzzy rule base by an appropriate fuzzy relation. Let $x \in X$, $y \in Y$, let operation $*$ be a left-continuous t-norm and finally, let $\to_*$ be its adjoint residuation operation [18]. Then the first approach consists in a construction of a fuzzy relation $\hat{\mathbf{R}}_* \in \mathcal{F}(X \times Y)$ given as follows

$$\hat{\mathbf{R}}_*(x, y) = \bigwedge_{i=1}^{n} (\mathbf{A}_i(x) \to_* \mathbf{B}_i(y)) \,. \tag{1}$$

As written in [11], "In the above view, each piece of information (fuzzy rule) is viewed as a constraint. This view naturally leads to a conjunctive way of merging the individual pieces of information since the more information, the more constraints and the less possible values to satisfy them."

This fact together with the fact that the minimum operation as well as other t-norms are appropriate interpretations of a conjunction (logical connective AND) and residuation operations are appropriate interpretations of an implication [1, 10, 15, 23, 30] leads to a conclusion that $\hat{\mathbf{R}}_*$ is a proper model of the following

*This research was sponsored by the Bilateral Scientific Cooperation Flanders–Czech Republic, sponsored by the Special Research Fund of Ghent University (project nr. 011S01106). Furthermore, we gratefully acknowledge partial support of the projects MSM6198898701 and DAR 1M0572 of the MŠMT ČR.

fuzzy rules

$$\begin{array}{lll} \textbf{IF } x \text{ is } \mathcal{A}_1 & \textbf{THEN} & y \text{ is } \mathcal{B}_1 \\ & \dots & \\ & \textbf{AND} & \\ & \dots & \\ \textbf{IF } x \text{ is } \mathcal{A}_n & \textbf{THEN} & y \text{ is } \mathcal{B}_n \end{array} \tag{2}$$

where $\mathcal{A}_i, \mathcal{B}_i$ are membership predicates represented by fuzzy sets $\mathbf{A}_i \in \mathcal{F}(X)$ and $\mathbf{B}_i \in \mathcal{F}(Y)$, respectively.

The second standard approach to modelling a fuzzy rule base, initiated by [20], consists in a construction of a fuzzy relation $\check{\mathbf{R}}_* \in \mathcal{F}(X \times Y)$ given as follows

$$\check{\mathbf{R}}_*(x,y) = \bigvee_{i=1}^{n} (\mathbf{A}_i(x) * \mathbf{B}_i(y)). \tag{3}$$

Obviously, fuzzy relation (3) can be hardly considered as a model of fuzzy rule base (2).

We again recall the work of Dubois, Prade and Ughetto [11]. "It seems that fuzzy rules modelled by (3) are not viewed as constraints but are considered as pieces of data. Then the maximum in (3) expresses accumulation of data".

This fact together with a commonly known fact that the maximum operation as well as other t-conorms are appropriate interpretations of a disjunction (logical connective OR) [15, 23] leads to a conclusion that $\check{\mathbf{R}}_*$ given by (3) is a proper model of the following set of fuzzy rules

$$\begin{array}{lll} x \text{ is } \mathcal{A}_1 & \textbf{AND} & y \text{ is } \mathcal{B}_1 \\ & \dots & \\ & \textbf{OR} & \\ & \dots & \\ x \text{ is } \mathcal{A}_n & \textbf{AND} & y \text{ is } \mathcal{B}_n. \end{array} \tag{4}$$

It is worth mentioning that distinguishing between the conditional (IF-THEN) form of fuzzy rules (2) and the Cartesian product (AND) form of fuzzy rules (4) on a syntactical level is not very common but it can be found e.g. in [3, 17, 22]. Usually only the form given by (2) is considered, because of several, e.g., historical reasons and the differences are taken into account only on the semantical level. But the differences can play a crucial role for further implementations and, therefore, they should be kept in mind. For a more detailed study concerning both rule forms we refer to [15, 23, 17] and to an exhaustive investigation in [9] and in [22].

2 Systems of Fuzzy Relation Equations

Besides a fuzzy rule base, (fuzzy) inference mechanism is an essential part of each FRB system. It is a deduction rule determining an output $\mathbf{B} \in \mathcal{F}(Y)$ based on an arbitrary input $\mathbf{A} \in \mathcal{F}(X)$. In particular, it is defined as an image of $\mathbf{A}$ under a fuzzy relation $\mathbf{R} \in \mathcal{F}(X \times Y)$, which models a given fuzzy rule base. In most cases, we deal with the *direct image* (sup-$*$ composition)

$$\mathbf{B} = \mathbf{A} \circ_* \mathbf{R} \tag{5}$$

which stems from the *compositional rule of inference* introduced by L.A. Zadeh [32]. It is given by

$$(\mathbf{A} \circ_* \mathbf{R})(y) = \bigvee_{x \in X} (\mathbf{A}(x) * \mathbf{R}(x,y)) \tag{6}$$

and it is worth mentioning that its logical background coincides with the *generalized modus ponens* [15].

A fuzzy rule base may be viewed as a partial mapping from $\mathcal{F}(X)$ to $\mathcal{F}(Y)$. Building a fuzzy inference module on the base of a rule base means extending this partial function to a total one in some "reasonable manner". It means we have to associate $\mathbf{B} \in \mathcal{F}(Y)$ with an arbitrary $\mathbf{A} \in \mathcal{F}(X)$ in such a way which having an input $\mathbf{A}_i$ would determine an output equal to $\mathbf{B}_i$ for $i = 1, \dots, n$.

It leads to the following *system of direct image fuzzy relation equations*

$$\mathbf{A}_i \circ_* \mathbf{R} = \mathbf{B}_i \quad i = 1, \dots, n. \tag{7}$$

A fuzzy relation $\mathbf{F} \in \mathcal{F}(X \times Y)$ which fulfills the equality (7) is a *solution* to the system of direct image equations.

Let us recall some basic results concerning the systems of direct image equations which can be found e.g. in [6, 16, 19].

Theorem 1 *System (7) is solvable if and only if $\hat{\mathbf{R}}_*$ is a solution to the system and moreover, by $\hat{\mathbf{R}}_*$ is the greatest solution to (7).*

Theorem 1 is a crucial theorem in the field of direct image equations since it is a necessary and sufficient condition of the solvability of system (7) and it determines a solution. Moreover, whenever we deal with fuzzy rules (2) interpreted by $\hat{\mathbf{R}}_*$, the direct image is the first choice for an inference mechanism since fuzzy relation $\hat{\mathbf{R}}_*$ has a unique position in the set of all possible solutions. There cannot appear any other solutions if $\hat{\mathbf{R}}_*$ is not a solution and it is the greatest solution if the corresponding system is solvable.

Let us recall a solvability criterion (sufficient condition) published in [7] and then independently re-found in [16] specifying conditions upon which even $\check{\mathbf{R}}_*$ can be a solution to system (7).

Theorem 2 *Let $\mathbf{A}_i$ for $i = 1, \dots, n$ be normal. Then $\check{\mathbf{R}}_*$ is a solution to (7) if and only if the following*

condition

$$\bigvee_{x\in X}(\mathbf{A}_i(x) * \mathbf{A}_j(x)) \le \bigwedge_{y\in Y}(\mathbf{B}_i(y) \leftrightarrow_* \mathbf{B}_j(y)) \quad (8)$$

holds for arbitrary $i, j \in \{1, \ldots, n\}$.

Theorem 2 specifies a condition under which $\check{\mathbf{R}}_*$, connected to the direct image inference mechanism, is an appropriate model of a fuzzy rule base. On the other hand, whenever $\check{\mathbf{R}}_*$ is an appropriate model, the fuzzy relation $\hat{\mathbf{R}}_*$ is an appropriate model too. So, $\check{\mathbf{R}}_*$ does not achieve such a unique position among other possible fuzzy models compared to $\hat{\mathbf{R}}_*$.

But it is not common approach to fix an inference mechanism and search for an appropriate model of a given fuzzy rule base which would be a solution to a corresponding system of fuzzy relation equations. Vice-versa, usually a fuzzy relation is chosen first based on certain, here more closely unspecified, reasons. Therefore, it is natural to adopt another inference mechanism which yields a system of fuzzy relation equations where $\check{\mathbf{R}}_*$ has an analogous position among other solutions to the system.

Let us recall by subdirect image (inf-$\rightarrow_*$ composition)

$$\mathbf{B} = \mathbf{A} \lhd_* \mathbf{R} \quad (9)$$

which rises from the *Bandler-Kohout subproduct* [2]. It is given by

$$(\mathbf{A} \lhd_* \mathbf{R})(y) = \bigwedge_{x\in X}(\mathbf{A}(x) \rightarrow_* \mathbf{R}(x, y)). \quad (10)$$

The subdirect image has no connection to the generalized modus ponens deductive rule and its motivation was different in [2]. On the other hand, as mentioned in [14], the inference mechanism needs not be necessarily logical but simply a mapping from $\mathcal{F}(X)$ to $\mathcal{F}(Y)$ fulfilling certain properties.

In the sequel, we will recall some basic facts [6, 19] about *systems of subdirect image fuzzy relation equations* given as follows

$$\mathbf{A}_i \lhd_* \mathbf{R} = \mathbf{B}_i \quad i = 1, \ldots, n \quad (11)$$

which should justify our further usage of the subdirect image as an inference mechanism.

Theorem 3 *System (11) is solvable if and only if* $\check{\mathbf{R}}_*$ *is a solution to the system and moreover,* $\check{\mathbf{R}}_*$ *is the least solution to system (11).*

Furthermore, both systems of fuzzy relation equations are dual [6].

Theorem 4 *[21] Let* $\mathbf{A}_i$ *for* $i = 1, \ldots, n$ *be normal. Then* $\hat{\mathbf{R}}_*$ *is a solution to (11) if and only if the condition*

$$\bigvee_{x\in X}(\mathbf{A}_i(x) * \mathbf{A}_j(x)) \le \bigwedge_{y\in Y}(\mathbf{B}_i(y) \leftrightarrow_* \mathbf{B}_j(y))$$

holds for arbitrary $i, j \in \{1, \ldots, n\}$.

Here we observe, that $\check{\mathbf{R}}_*$ has precisely the same position among other solutions to system (11) as the fuzzy relation $\hat{\mathbf{R}}_*$ had in the case of the direct image equations. So, if we adopt the idea of [14] that the inference mechanism can be understood only as a certain mapping between collections of fuzzy sets without deeper connection to a logical deduction rule, nothing prevents us from treating the subdirect image as a specific kind of finding a conclusion behaving as an inference mechanism.

Based on the facts and theorems above, we claim that fuzzy rules (2) predetermine the direct image inference while fuzzy rules (4) predetermine the subdirect image inference.

Remark 1 *It is worth mentioning that condition (8) appearing in Theorem 2 and Theorem 4 is from the practical point of view not very convenient. On the other hand, it is well known fact that if antecedent fuzzy sets form the so called* $*$ *semi-partition [8] it forces fulfilling the discussed condition in advance.*

3 Additive Models

A lot of work has been done in the field of fuzzy relation equations [6, 13, 16], mainly aiming at the solvability criteria, greatest and lower solutions, least and greater solutions, respectively and at determining complete solution sets.

Unfortunately, this work has not been followed by practitioners very often. This was perhaps caused, besides other unspecified reasons, by the fact that neuro-fuzzy systems [12] and Takagi-Sugeno rules [29] are the most frequently used fuzzy methods nowadays. But only two particular solutions $\hat{\mathbf{R}}_*$ and $\check{\mathbf{R}}_*$ have been mainly studied so far [24].

Let us only briefly recall that Takagi and Sugeno proposed fuzzy rules of the form

$$\textbf{IF } x \text{ is } \mathcal{A}_i \quad \textbf{THEN} \quad y \text{ is } f_i(x),\ i = 1, \ldots, n \quad (12)$$

which are modelled by the weighted arithmetic mean

$$y = \frac{\sum_{i=1}^{n} \mathbf{A}_i(x) f_i(x)}{\sum_{i=1}^{n} \mathbf{A}_i(x)}, \quad (13)$$

which has obviously nothing common with logical implication and so, they do not correspond to the fuzzy

rules in conditional form (2). But because of powerful approximation abilities, they became very popular in fuzzy community.

Let us stress that the consequent parts $f_i(x)$ of rules (12) are usually polynomial functions so, we can distinguish Takagi-Sugeno rules of the k-th order where $k \in \mathbb{N}$ denotes the order of the consequent polynomial functions.

It is worth mentioning that the so called *Ruspini* condition [28] given by

$$\sum_{i=1}^{n} \mathbf{A}_i(x) = 1, \quad \text{for } x \in X \tag{14}$$

is often required and therefore, the interpretation of the Takagi-Sugeno rules of the 0-th order is then given by the formula

$$\sum_{i=1}^{n} \mathbf{A}_i(x) b_i \tag{15}$$

where $b_i \in \mathbb{R}$ are consequents of rules (12).

Remark 2 *Takagi-Sugeno rules are basically data-driven i.e. determined on the basis of a finite input-output data set. Besides this standard approach to an identification of a model, there exists an integral (continuous) version of the identification of the Takagi-Sugeno rules of the* 0*-th order called fuzzy transform [25].*

Indeed, considering crisp values b_i to be singletons, i.e., special fuzzy numbers on Y, having in mind that the product is nothing else but a particular t-norm [18], and the fact that we assume the Ruspini condition leads to the following natural generalization of Takagi-Sugeno rules with fuzzy consequents

$$\mathbf{R}_*^{\oplus}(x,y) = \bigoplus_{i=1}^{n} (\mathbf{A}_i(x) * \mathbf{B}_i(y)), \tag{16}$$

where $\oplus$ is the Łukasiewicz t-conorm. Moreover, it coincides with standard fuzzy relations appearing in the neuro-fuzzy systems [12].

Fuzzy relation $\mathbf{R}_*^{\oplus} \in \mathcal{F}(X \times Y)$ is nothing else but a special case of a model of fuzzy rule base (4) derived from the disjunctive model $\check{\mathbf{R}}_*$ where the disjunction aggregating particular pieces of data is generally given by a t-conorm, in this particular case by the Łukasiewicz addition. Therefore, it will be called *additive* fuzzy model.

Remark 3 *Formalization of fuzzy rules (2) modelled by $\hat{\mathbf{R}}_*$ can be viewed in the "conjunctive normal forms" (CNF) and formalization of fuzzy rules (4) modelled by $\check{\mathbf{R}}_*$ can be viewed in the "disjunctive normal forms" (DNF), for both we refer to [26]. The normal forms were proposed to investigate fuzzy models from an approximation point of view. The additive models $\mathbf{R}_*^{\oplus}$ correspond to the "additive normal forms" (ANF) which were motivated by one particular additive normal form in [26] and then further studied in [4, 5]. Finally, relationship between additive normal forms and the fuzzy transform (and Takagi-Sugeno rules consequently) has been discussed in [4].*

Let us recall the definition of the generalized orthogonality [26] which was crucial for the additive normal forms [26, 5].

Definition 1 *We say that $\mathbf{A}_i \in \mathcal{F}(X)$, $i = 1, \ldots, n$ fulfill the orthogonality condition if*

$$\bigoplus_{\substack{i=1 \\ i \neq j}}^{n} \mathbf{A}_i(x) = 1 - \mathbf{A}_j(x). \tag{17}$$

The following lemma characterizes the orthogonality condition as equivalent to the Ruspini one.

Lemma 1 *Fuzzy sets $\mathbf{A}_i \in \mathcal{F}(X)$, $i = 1, \ldots, n$ fulfill the orthogonality condition if and only if*

$$\sum_{i=1}^{n} \mathbf{A}_i(x) = 1. \tag{18}$$

3.1 Subdirect Image Equations

Since the additive model corresponds to fuzzy rules (4) they are expected to be related to the subdirect image inference mechanism. This subsection investigates this relationship via subdirect image fuzzy relation equations.

Theorem 5 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Then system (11) is solvable and $\mathbf{R}_*^{\oplus}$ is a solution.*

Due to Theorem 3 we can state the following corollary of Theorem 5.

Corollary 1 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Then $\check{\mathbf{R}}_*$ is a solution to system (11).*

Moreover, ite can be demonstrated on the basis of the proof of Theorem 5, that $\mathbf{R}_*^{\oplus}$ is not the only additive model which is a solution to system (11).

Proposition 1 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Furthermore, let $\blacktriangle$ be a t-norm such that $* \leq \blacktriangle$. Then the fuzzy relation $\mathbf{R}_{\blacktriangle}^{\oplus}$ is a solution to (11).*

Let us briefly summarize results of this subsection. Theorem 5 provides us with an easy-to-check condition certifying a proper performance of an additive model connected to a subdirect image inference mechanism. Moreover, the assumptions refer only to the antecedent fuzzy sets so and their fulfillment can be ensured in advance before an identification process.

Since no solvability is assumed, the theorem impact is even in specifying sufficient solvability condition. This consequently means that $\check{\mathbf{R}}_*$ is a solution as well. Finally, a wide variety of t-norms can be used in the additive models.

3.2 Direct Image Equations

This subsection focuses on systems of direct image equations. Disjunctive model $\check{\mathbf{R}}_*$ has been proved to be a solution to system (7) assuming certain conditions. Similarly, we study the additive model on the same system.

Theorem 6 *Let $\mathbf{A}_i$ for $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Furthermore, let $* \leq \otimes$ where $\otimes$ is the Łukasiewicz t-norm. Then system (7) is solvable and $\mathbf{R}_*^{\oplus}$ is a solution.*

Due to Theorem 1 we can state the following corollary of Theorem 6.

Corollary 2 *Let $\mathbf{A}_i$ for $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Furthermore, let $* \leq \otimes$. Then $\hat{\mathbf{R}}_*$ is a solution to system (7).*

Theorem 6 requires to use a t-norm which is even weaker than the Łukasiewicz one which is already a very weak t-norm, see [18]. So, for practical applications, perhaps only the case when $* = \otimes$ is worth mentioning. In this case, the Łukasiewicz t-norm is used both for, the sup-$\otimes$ composition as an inference method as well as for connecting antecedent and consequent fuzzy sets in the corresponding fuzzy model $\mathbf{R}_{\otimes}^{\oplus}$.

The result is strengthened by the following theorem.

Proposition 2 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and $\mathbf{A}_i$ fulfill the Ruspini condition. Furthermore, let $\blacktriangle$ be an arbitrary t-norm and let $* \leq \otimes$. Then $\mathbf{R}_{\blacktriangle}^{\oplus}$ is a solution to (7).*

Proposition 2 allows us to deal with a t-norm weaker or equal to the Łukasiewicz one only in the inference mechanism but the fuzzy model can use an arbitrary t-norm $\blacktriangle$. Similarly to the subdirect image equations, only normality of antecedents and the Ruspini condition were assumed.

4 Multiplicative Models

As we have shown in Section 3, investigation of additive fuzzy models does not bring only results concerning proper usage of such models. Furthermore, it led to new results in the field of systems of fuzzy relation equations. Especially the fact, that fulfilling two, in practice very often required, conditions leads to the solvability of the corresponding fuzzy relation equations. These results are of a high practical importance since they put assumptions only on the antecedents. This enables to identify a fuzzy rule base in such a way to ensure the solvability of an adjoint system of fuzzy relation equations although the consequent fuzzy sets could be arbitrary e.g. identified from data by some algorithm.

It is a well known fact that both systems of fuzzy relation equations are dual [6]. This leads to a natural idea of introducing the following *multiplicative models*

$$\mathbf{R}_*^{\otimes}(x, y) = \bigotimes_{i=1}^{n} (\mathbf{A}_i(x) \rightarrow_* \mathbf{B}_i(y)). \tag{19}$$

Multiplicative model $\mathbf{R}_*^{\otimes}$ is a special case of a model of fuzzy rule base (2) derived from the conjunctive model $\hat{\mathbf{R}}_*$ where the conjunction aggregating particular rules is generally given by a t-norm, in this particular case by the Łukasiewicz t-norm. It is another arithmetic model, dual to the corresponding additive one.

Now we focus on searching for solutions to systems of direct image and subdirect image fuzzy relation equations given by (19).

4.1 Direct Image Equations

Theorem 7 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Then system (7) is solvable and $\mathbf{R}_*^{\otimes}$ is a solution.*

Due to Theorem 1 we can state the following corollary of Theorem 7.

Corollary 3 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Then $\hat{\mathbf{R}}_*$ is a solution to (7).*

Analogously to the case of subdirect image equations and additive models, we can use a wide variety of residuation operations in the multiplicative models.

Proposition 3 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Furthermore, let $\blacktriangle$ be a t-norm such that $* \leq \blacktriangle$. Then $\mathbf{R}_{\blacktriangle}^{\otimes}$ is a solution to (7).*

4.2 Subdirect Image Equations

Theorem 8 *Let $\mathbf{A}_i$ for $i = 1, \ldots, n$ be normal and fulfill the Ruspini condition. Furthermore, let $* \leq \otimes$. Then $\mathbf{R}_*^{\otimes}$ is a solution to system (11).*

Analogously to the dual case, to the case of an additive fuzzy model and the direct image equation, the Łukasiewicz t-norm is to weak and so $\mathbf{R}_{\otimes}^{\otimes}$ is the only case of a practical importance. And analogously to the dual case, Theorem 8 can be strengthen.

Proposition 4 *Let $\mathbf{A}_i$, $i = 1, \ldots, n$ be normal and fulfill the orthogonality condition. Furthermore, let $\blacktriangle$ be a t-norm such that $\blacktriangle \leq \otimes$. Then $\mathbf{R}_{\blacktriangle}^{\otimes}$ is a solution to (11).*

Proposition 4, allows us to deal with a residuation adjoint to a weak t-norm only in the fuzzy relation interpreting a fuzzy rule base but the respective inference machine can be based on an arbitrary residuation operation.

5 Examples

To demonstrate the results introduced in Section 3, let us consider the following example.

Example 1 *Let $\mathbf{A}_i \in \mathcal{F}(X)$, $i = 1, \ldots, n$ fulfill the Ruspini condition*

$$\sum_{i=1}^{n} \mathbf{A}_i(x) = 1, \quad \forall x \in X. \tag{20}$$

Moreover, let us assume normality of each fuzzy set $\mathbf{A}_i$.

Let there be a fuzzy rule base (4) where the antecedents are represented by the given fuzzy sets $\mathbf{A}_i$ and the consequents by arbitrary fuzzy sets $\mathbf{B}_i \in \mathcal{F}(Y)$.

Then due to Theorem 5, fuzzy relation

$$\mathbf{R}_{\otimes}^{\oplus}(x, y) = \bigoplus_{i=1}^{n} (\mathbf{A}_i(x) \otimes \mathbf{B}_i(y)) \tag{21}$$

is a solution to the system of fuzzy relation equations

$$\mathbf{A_i} \triangleleft_{\otimes} \mathbf{R} = \mathbf{B}_i, \quad i = 1, \ldots, n.$$

Furthermore, this system is also solved by fuzzy relations

$$\check{\mathbf{R}}_{\otimes}(x, y) = \bigvee_{i=1}^{n} (\mathbf{A}_i(x) \otimes \mathbf{B}_i(y)) \quad \textit{and}$$

$$\mathbf{R}_{\odot}^{\oplus}(x, y) = \bigoplus_{i=1}^{n} (\mathbf{A}_i(x) \odot \mathbf{B}_i(y)) = \sum_{i=1}^{n} \mathbf{A}_i(x)\mathbf{B}_i(y)$$

due to Corollary 1 and Proposition 1, respectively.

Moreover, due to Theorem 6, fuzzy relation (21) is also a solution to the system of fuzzy relation equations

$$\mathbf{A_i} \circ_{\otimes} \mathbf{R} = \mathbf{B}_i \quad i = 1, \ldots, n, \tag{22}$$

and due to Proposition 2, a solution to system (22) can be also found in the form of $\mathbf{R}_{\odot}^{\oplus}$.

This means, that fuzzy models $\check{\mathbf{R}}_{\otimes}$, $\mathbf{R}_{\otimes}^{\oplus}$ and $\mathbf{R}_{\odot}^{\oplus}$ are safe models [27] of fuzzy rules (4) from the fuzzy interpolation point of view when assuming the subdirect image with the Łukasiewicz t-norm as the corresponding inference mechanism. The two latter arithmetic fuzzy models $\mathbf{R}_{\otimes}^{\oplus}$ and $\mathbf{R}_{\odot}^{\oplus}$ are safe even in the case of the direct image with the Łukasiewicz t-norm.

To demonstrate the results introduced in Section 4, let us consider the following example.

Example 2 *Let $\mathbf{A}_i \in \mathcal{F}(X)$, $i = 1, \ldots, n$ fulfill the Ruspini condition. Moreover, let us assume normality of each fuzzy set $\mathbf{A}_i$.*

Let there be a fuzzy rule base (2) where the antecedents are represented by the given fuzzy sets $\mathbf{A}_i$ and the consequents by arbitrary fuzzy sets $\mathbf{B}_i \in \mathcal{F}(Y)$.

Then due to Theorem 7, fuzzy relation

$$\mathbf{R}_{\otimes}^{\otimes}(x, y) = \bigotimes_{i=1}^{n} (\mathbf{A}_i(x) \rightarrow_{\otimes} \mathbf{B}_i(y)) \tag{23}$$

is a solution to the system of fuzzy relation equations

$$\mathbf{A_i} \circ_{\otimes} \mathbf{R} = \mathbf{B}_i, \quad i = 1, \ldots, n.$$

Furthermore, this is system is also solved by fuzzy relations

$$\hat{\mathbf{R}}_{\otimes}(x, y) = \bigwedge_{i=1}^{n} (\mathbf{A}_i(x) \rightarrow_{\otimes} \mathbf{B}_i(y)) \quad \textit{and}$$

$$\mathbf{R}_{\odot}^{\otimes}(x, y) = \bigotimes_{i=1}^{n} (\mathbf{A}_i(x) \rightarrow_{\odot} \mathbf{B}_i(y))$$

due to Corollary 3 and Proposition 3, respectively.

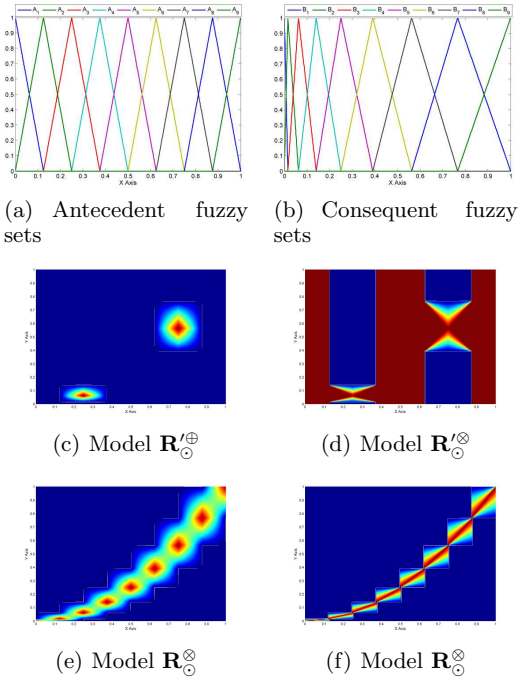
Moreover, due to Theorem 8, fuzzy relation (23) is also a solution to the system of fuzzy relation equations

$$\mathbf{A_i} \triangleleft_{\otimes} \mathbf{R} = \mathbf{B}_i \quad i = 1, \ldots, n, \tag{24}$$

and due to Proposition 4, even to the system

$$\mathbf{A_i} \triangleleft_{\odot} \mathbf{R} = \mathbf{B}_i \quad i = 1, \ldots, n. \tag{25}$$

This means, that the fuzzy models $\hat{\mathbf{R}}_{\otimes}$, $\mathbf{R}_{\odot}^{\otimes}$ and $\mathbf{R}_{\otimes}^{\otimes}$ are safe models [27] of fuzzy rules (2) from the fuzzy



(a) Antecedent fuzzy sets (b) Consequent fuzzy sets

(c) Model $\mathbf{R}'^{\oplus}_{\odot}$ (d) Model $\mathbf{R}'^{\otimes}_{\odot}$

(e) Model $\mathbf{R}^{\otimes}_{\odot}$ (f) Model $\mathbf{R}^{\otimes}_{\odot}$

Figure 1: Example of arithmetic fuzzy models.

interpolation point of view when assuming the direct image with the Łukasiewicz t-norm as the corresponding inference mechanism. The last arithmetic fuzzy models $\mathbf{R}^{\otimes}_{\otimes}$ is safe even in the case of the subdirect image with the Łukasiewicz t-norm or the subdirect image with the product t-norm.

In the sequel, we present several figures of the proposed fuzzy models.

Example 3 *Let us be given fuzzy pairs* $(\mathbf{A}_i, \mathbf{B}_i)$, $i = 1, \ldots, 9$ *on* $\mathcal{F}([0,1]) \times \mathcal{F}([0,1])$ *approximating the function* $y = x^2$. *The fuzzy sets* $\mathbf{A}_i$ *are triangular and form a uniform fuzzy partition of* $X = [0,1]$. *The fuzzy sets* $\mathbf{B}_i$ *are triangular with the kernel points equal to the precise solution of the equation* $y = x^2$ *for* x *being the kernel point of the corresponding* $\mathbf{A}_i$ *and they fulfil the Ruspini condition, see Figures 1(a)—1(b).*

Let us observe how the additive and the multiplicative fuzzy models, both with respect to the product t-norm, look if we take into account either only two or all fuzzy pairs.

The fuzzy models $\mathbf{R}'^{\oplus}_{\odot}$ *and* $\mathbf{R}'^{\otimes}_{\odot}$ *take into account only pairs of the fuzzy sets* $(\mathbf{A}_3, \mathbf{B}_3)$ *and* $(\mathbf{A}_7, \mathbf{B}_7)$, *the fuzzy models* $\mathbf{R}^{\oplus}_{\odot}$ *and* $\mathbf{R}^{\otimes}_{\odot}$ *take into account all the given fuzzy pairs.*

How the constructed fuzzy models look like (view from above) is displayed on Figure 1.

We can see that the, originally absolutely non-compatible, models converge to each other when the corresponding antecedent fuzzy sets converge to the Ruspini partition.

This observation completes known facts about relationship between partitions, solvability of systems of fuzzy relation equations and possible solutions (cf. Remark 1).

6 Conclusions

Two types of fuzzy rules have been discussed together with corresponding inference mechanisms. New fuzzy models of the rules have been proposed. Additive and multiplicative fuzzy models have been discussed from the fuzzy interpolation point of view, i.e., as possible solutions of corresponding systems of fuzzy relation equations. The crucial role of the Ruspini partition has been proved and visually demonstrated.

Acknowledgement

This paper is a tribute to the third author - Lenka Nosková - who tragically passed away in March 4, 2007. The paper was one of her last scientific works.

References

[1] Bandler, W. and Kohout, L.(1980). Fuzzy power sets and fuzzy implication operators. *Fuzzy Sets and Systems*, 4, pp. 183-190.

[2] Bandler, W. and Kohout, L.J.(1978). Fuzzy relational products and fuzzy implication operators. In: *Proc. of International Workshop on Fuzzy Reasoning Theory and Applications*, Queen Mary College, University of London.

[3] Daňková, M.(2007). On approximate reasoning with graded rules. *Fuzzy Sets and Systems*, to appear.

[4] Daňková, M. and Štěpnička, M.(2006). Fuzzy transform as an additive normal form. *Fuzzy Sets and Systems*, 157, pp. 1024-1035.

[5] Daňková, M. and Štěpnička, M.(2005). Genetic algorithms in fuzzy approximation. In: *Proc. of the Joint 4th EUSFLAT2005 and 11th LFA2005*, Barcelona, pp. 651-656.

[6] De Baets, B.(2000). Analytical solution methods for fuzzy relation equations. In: *The Handbook of Fuzzy Set Series*, (D. Dubois and H. Prade, Ed.) Vol. 1, Academic Kluwer Publ., Boston, pp. 291-340.

[7] De Baets, B.(1996). A note on Mamdani controllers, In: *Proc. of the 2nd International FLINS Workshop*, (R. Da, P. D'hondt, P. Govaerts and E. Kerre, Ed.), Mol, Belgium, World Scientific.

[8] De Baets, B. and Mesiar, R.(1998). T-partitions. *Fuzzy Sets and Systems*, 97, pp. 211-223.

[9] Dubois, D. and Prade, H.(1996). What are fuzzy rules and how to deal with them. *Fuzzy Sets and Systems*, 84, pp. 169185.

[10] Dubois, D. and Prade, H.(1984). A theorem on implication functions defined from triangular norms, *Stochastica*, 8, pp. 267-279.

[11] Dubois, D., Prade, H. and Ughetto, L.(1997). Checking the coherence and redundancy of fuzzy knowledge bases, *IEEE Transactions on Fuzzy Systems*, 5, pp. 398-417.

[12] Fuller, R.(2000). *Introduction to Neuro-fuzzy Systems*, (J. Kacprzyk (Eds.), Advances in Soft Computing). Physica-Verlag, Heidelberg.

[13] Gottwald, S.(2000). Generalized solvability behaviour for systems of fuzzy equations. In: *Discovering the World with Fuzzy Logic*, (V. Novák and I. Perfilieva, Ed., Studies in Fuzzines and Soft Computing), Springer Physica-Verlag, Heidelberg, New York, pp. 401-430.

[14] Gottwald, S.(2006). On a generalization of fuzzy relation equations. In: *Proc. of 11th IPMU 2006*, Paris, pp. 2572-2577.

[15] Hájek, P.(1998). *Metamathematics of Fuzzy Logic.* Kluwer, Dordrecht.

[16] Klawonn, F.(2000). Fuzzy points, fuzzy relations and fuzzy functions. In: *Discovering the World with Fuzzy Logic*, (V. Novák and I. Perfilieva, Ed., Studies in Fuzzines and Soft Computing), Springer Physica-Verlag, Heidelberg, New York, pp. 431-453.

[17] Klawonn, F. and Novák, V.(1996) The relation between inference and interpolation in the framework of fuzzy systems. *Fuzzy Sets and Systems*, 81, pp. 331-354.

[18] Klement, E.P., Mesiar, R. and E. Pap(2000). *Triangular Norms.* Kluwer, Boston, Dordrecht, London.

[19] Klir, G.J. and B. Yaun(1995). *Fuzzy Sets and Fuzzy Logic.* Prentice Hall, New Jersey.

[20] Mamdani, E.H. and Assilian, S.(1975). An experiment in lingustic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7, pp. 1-13.

[21] Nosková, L.(2005). Systems of fuzzy relation equation with inf→ composition: solvability and solutions. *Journal of Electrical Engineering*, 12(s), pp. 69-72.

[22] Novák, V. and Lehmke, S.(2006), Logical structure of fuzzy IF-THEN rules. *Fuzzy Sets and Systems*, 157, pp. 20032029.

[23] Novák, V., I. Perfilieva and J. Močkoř(1999). *Mathematical Prinicples of Fuzzy Logic.* Kluwer, Boston, Dordrecht, London.

[24] Perfilieva, I.(2004). Fuzzy function as an approximate solution to a system of fuzzy relation equations. *Fuzzy Sets and Systems*, 147, pp. 363383.

[25] Perfilieva, I.(2006). Fuzzy transforms: Theory and applications. *Fuzzy Sets and Systems*, 157, pp. 993-1023.

[26] Perfilieva, I.(2004). Normal forms in BL and ŁΠ algebras of functions. *Soft Computing*, 8, pp. 291-298.

[27] Perfilieva, I. and Lehmke, S.(2005). Safe modeling of fuzzy if-then rules. In: *Computational Intelligence, Theory and Applications*, (B. Reusch, Ed.), Springer Physica-Verlag, Heidelberg, pp. 235-240.

[28] Ruspini, E.H.(1969). A new approach to clustering. *Inform. and Control*, 15, pp. 22-32.

[29] Takagi, T. and Sugeno, M.(1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 15, pp. 116-132.

[30] Trillas, E. and Valverde, L.(1985). On implication and indistinguishability in the setting of fuzzy logic. In: *Management Decision Support Systems Using Fuzzy Sets and Possibility Theory*, (J. Kacprzyk and R. R. Yager, Ed.), Springer Physica-Verlag, Rheinland, Koln, pp. 198-212.

[31] Vetterlein, T. and Štěpnička, M.(2006). Completing fuzzy if-then rule bases by means of smoothing splines. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14, pp. 235-244.

[32] Zadeh, L.A.(1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man and Cybernet*, 3, pp. 28-44.

Automation of Human Reasoning in Economical Analysis

Antonín Dvořák
Institute for Research and Applications
of Fuzzy Modeling
University of Ostrava, 30. dubna 22,
701 03 Ostrava, The Czech Republic
antonin.dvorak@osu.cz

Vilém Novák
Institute for Research and Applications
of Fuzzy Modeling
University of Ostrava, 30. dubna 22,
701 03 Ostrava, The Czech Republic
vilem.novak@osu.cz

Abstract

An application of so-called perception-based logical deduction in the modeling of an economic analysis given in natural language is presented. Fuzzy IF-THEN rules and theory of evaluating linguistic expressions are used, and possibilities of automated reasoning are outlined.

Keywords: Fuzzy type theory, Fuzzy IF-THEN rules, Perceptions.

1 Introduction

In this paper, we will describe an application of perception-based logical deduction (PbLD) in the modeling of economic analysis given in natural language.

In paper [11], we used the fuzzy type theory (FTT) [7], theory of evaluating linguistic expressions [10] and perception-based logical deduction (PbLD) [8] in the analysis of a detective story. In this paper, we will apply this method to a model of economical analysis of macroeconomic situation. For this purpose, we use as an example a part of free economical analysis of the Czech Savings Bank. It describes the influence of economic growth and other factors to the change of unemployment rate in the Czech Republic. It turns out that influence of these auxiliary factors (non-motivating social system, rigid labor market, high tax load) is important and decreases the expected positive effect of high economic growth.

Of course, the model of this particular economic analysis should be taken as an illustration of our methodology based on FTT and PbLD. Its main features are:

- It works inside well-developed and sound formal theory.
- It extensively uses formal theory of the, so-called, *evaluating linguistic expressions*, which are natural-language expressions describing positions on an ordered scale. Examples of evaluating linguistic expressions are *small*, *more or less big*, etc. These expressions are used incessantly by people. Therefore, sound formal model of them behaving accordingly to human intuition is very important.
- Evaluating linguistic expressions are used in fuzzy IF-THEN rules and sets of them which we call *linguistic descriptions*. The model of meaning of linguistic descriptions is constructed using formal tools of FTT. It allows us to model the role of context (possible world) in accordance with intuition using standard notions of intension and extension (see Section 2).
- Due to its formal nature, our methodology is open to various kinds of improvements in the direction of higher proportion of automated extraction of linguistic descriptions, automated deduction, etc.

Our approach relates to the, so-called, precisiated natural language (PNL) introduced by L. A. Zadeh in [15]. It aims at formalizing natural language sentences or texts using tools from the field of soft computing. E. g., the authors in [3] showed some application of PNL in the analysis of simple economic sentence.

There are also connections to *common-sense reasoning* [13], originated in large part by John McCarthy [6]. It includes formalization of reasonings performed by humans, taking into account their nonmonotonicity and other features.

2 Preliminaries

2.1 Fuzzy type theory

As stated, the main tool for the construction of a model*) of our economic analysis is fuzzy type theory. In this section, we will very briefly overview some of its main points.

*)We are using the term *model* in two meanings: as a general description of some situation, system etc., and as a formal model in the sense of mathematical logic.

A detailed explanation of FTT can be found in [7]. The classical type theory is in details described in [1].

The *Types* is a set of types constructed iteratively from the atomic types ϵ (elements) and o (truth values). *Form*$_\alpha$ denotes a set of formulas of type $\alpha \in$ *Types*. If $A \in$ *Form*$_\alpha$ is a formula of type $\alpha \in$ *Types* then we write A_α.

Formulas of type o (truth value) can be joined by the following connectives (derived formulas): $\vee$ (disjunction), $\wedge$ (conjunction), $\&$ (strong conjunction), ∇ (strong disjunction), $\Rightarrow$ (implication). General ($\forall$) and existential ($\exists$) quantifiers are defined as special formulas. For the details on their definition and semantics — see [7].

If $A \in$ *Form*$_{o\alpha}$ then A represents a property of elements of the type α. By abuse of language, we will often say "A is a property" (of elements of type α) and similarly, $A_{(o\alpha)\alpha}$ is a relation (between elements of type α). We will freely write or omit the type when no misunderstanding may occur.

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.

The operator

$$\imath z_\alpha A_o := \iota_{\alpha(o\alpha)}(\lambda z_\alpha A_o)$$

picks up an element of type α such that the formula A_o is true in the degree 1 for it.

Semantics. The structure of truth values in this paper is the Łukasiewicz$_\Delta$ algebra and so, the corresponding FTT is Łukasiewicz (Ł-FTT). Δ is the Baaz delta [5]. Let J be a language of Ł-FTT. A *frame* for J is a tuple $\mathcal{M} = \langle (M_\alpha, =_\alpha)_{\alpha \in Types}, \mathcal{L}_\Delta \rangle$ where $\mathcal{L}_\Delta$ is Łukasiewicz$_\Delta$ algebra of truth values, $=_\alpha$ is a fuzzy equality on M_α.

Recall that if $\beta\alpha$ is a type then the corresponding set $M_{\beta\alpha}$ contains (not necessarily all) functions $f : M_\alpha \longrightarrow M_\beta$.

Let p be an assignment of elements from $\mathcal{M}$ to variables. An interpretation $\mathcal{I}^{\mathcal{M}}$ is a function that assigns every formula A_α, $\alpha \in$ *Types* and every assignment p a corresponding element, that is, a function of the type α. A general model is a frame $\mathcal{M}$ such that $\mathcal{I}^{\mathcal{M}}_p(A_\alpha) \in M_\alpha$ holds true.

The following is a special formula representing a non-zero truth value:

$$\Upsilon_{oo} := \lambda z_o \cdot \neg\Delta(\neg z_o).$$

Lemma 1
If $T \vdash \Upsilon z_o \,\&\, (z_o \Rightarrow y_o)$ *then* $T \vdash \Upsilon y_o$.

2.2 Evaluating linguistic expressions

Evaluating linguistic expressions (or, simply, evaluating expressions) are expressions of natural language, for example, *small, medium, big, about twenty five, roughly one hundred, very short, more or less deep, not very tall, roughly warm or medium hot, quite roughly strong, roughly medium size*, and many others. They form a small but very important part of natural language and they are present in its everyday use any time. The reason is that people very often need to evaluate phenomena around them. Moreover, they often make important decisions based on them, learn how to control, and many other activities.

Because of lack of space and quite complicated formalism, we will only touch this theory and refer to the contribution [9]. All the details can be found in [10].

Important role in the theory of evaluating expressions is played by the concept of the *context*. It characterizes range of possible values for (numerical) variables and is represented by a special type ω. For simplicity, we will suppose that in each model, the context is given by a triple $\langle v_L, v_M, v_R \rangle$, where $v_L, v_M, v_R \in \mathbb{R}$ and $v_L < v_M < v_R$. The values v_L, v_M, v_R characterize minimal, middle and maximal value of the given context, respectively. On syntactical level, the context is denoted by a variable $w \in$ *Form*$_\omega$ and represented by three constants $\bot_w, \dagger_w$ and $\top_w$.

In applications, we often need also more general type of context, which allows both positive as well as negative values. We will call this type of context *two-sided*. It is given by $\langle v_{NL}, v_{NM}, v_Z, v_{PM}, v_{PR} \rangle$, where $v_{NL} < v_{NM} < v_Z < v_{PM} < v_{PR}$. In this context, the sign is used before evaluating linguistic expressions, e.g. *negative small*, *positive big* (abbreviations: $-Sm$, $+Bi$), etc. There is also special evaluating linguistic expression *zero*. Contexts from the previous paragraph will be called *simple*.

Evaluating expressions are denoted by letters $\mathcal{A}, \mathcal{B}, \ldots$. *Intension* of $\mathcal{A}$ is a formula $\mathrm{Int}(\mathcal{A})$. Recall that intension means *a property* that is denoted by $\mathcal{A}$. It is important to note that intension does not depend on the context. For example, *very small* is a name of a property of being "very small" which may mean about 150 cm (and less) when speaking about people, about 3 mm when speaking about beetles, etc. The type of $\mathrm{Int}(\mathcal{A})$ is $(o\alpha)\omega$. The latter will often be denoted by φ.The formal theory of evaluating expressions is denoted by T^{Ev}. This theory provides means how the above concepts of context, intension and others can be effectively formalized.

2.3 Fuzzy IF-THEN rules and perception-based logical deduction

The perception-based logical deduction in the frame of FTT has been described in [8]. Though the method is more general, we will suppose that all considered linguistic expressions are evaluating ones.

A fuzzy IF-THEN rule is a linguistic expression of the form

$$\mathcal{R} := \ \mathsf{IF}\ X\ \mathsf{is}\ \mathcal{A}\ \mathsf{THEN}\ Y\ \mathsf{is}\ \mathcal{B}.$$

where $\mathcal{A}, \mathcal{B}$ are evaluating expressions. The linguistic predication 'X is $\mathcal{A}$' is called *antecedent* and 'Y is $\mathcal{B}$' is called

consequent.

Intension of a fuzzy IF-THEN rule $\mathcal{R}$ is (represented by) a formula

$$\mathrm{Int}(\mathcal{R}) := \lambda w\, \lambda w' \cdot \lambda x\, \lambda y \cdot A_{(o\alpha)\omega} wx \Rightarrow B_{(o\beta)\omega} w'y \tag{1}$$

A *linguistic description LD* is a set of fuzzy IF-THEN rules. Its *topic* is a set of linguistic expressions $\{\mathrm{Int}(\mathcal{A}_j) \mid j = 1, \ldots, m\}$ and its focus is $\{\mathrm{Int}(\mathcal{B}_j) \mid j = 1, \ldots, m\}$, where m is a number of IF-THEN rules in the linguistic description *LD*.

In practice, we usually need more antecedent variables. In this case, the antecedent variables are usually connected by linguistic connective AND that is interpreted by a logical connective $\wedge$.

In perception-based logical deduction, we must introduce several special formulas. We will give only their informal description and refer to [8] for their precise definitions. The formula $\prec$ denotes the relation of sharpness between (intensions of) evaluating expressions. For example, if x is, at least partly, "very big" in all contexts then it is also "big" in all of them, i.e. $\mathrm{Int}(very\ big) \prec \mathrm{Int}(big)$. We will also introduce formula $Eval_{o(\varphi\alpha\omega)}$ ($Eval\ wx\,\mathrm{Int}(\mathcal{A})$ expresses that an element x in context w is evaluated by $\mathcal{A}$).

One of principal paradigms of the concept of precisiated natural language is that the world knowledge, i.e. the knowledge accumulated by people during their life, is *perception based.* We will formalize the concept of perception using a formula *Perc* $\in$ *Form*$_{(o\varphi)\alpha}$ which expresses that an intension z_φ is a *perception* of $x_\alpha \in$ *Form*$_\alpha$.

If we consider a linguistic description *LD* then *Perc*^{LD} is a perception w.r.t. *Topic*^{LD}. For example, for a given linguistic description *LD* the formula

$$Perc^{LD} x Ev$$

means that the evaluating expression *Ev* from the topic of *LD* is a *perception* of x. More precisely, there is a context w in which x is evaluated in the best way by the sharpest *Ev* (among evaluating expressions belonging to the topic of *LD*). We will also introduce a *local perception LPerc*$_{o(\varphi\alpha\omega)}$ relative to a specific context w.

Lemma 2
(a) Let $Ev_1 \prec Ev_2$. *Then*

$$T^{Ev} \vdash Perc\ x\, Ev_1 \Rightarrow Perc\ x\, Ev_2.$$

(b) Let $T^{Ev} \vdash z_\varphi wx \Rightarrow z'_\varphi w'y$. *Then* $T^{Ev} \vdash Eval\ wxz_\varphi \Rightarrow Eval\ w'yz'_\varphi$.

The linguistic description *LD* characterizes a certain kind of dependence (relation) between features of objects (and, consequently, the objects themselves) using natural language. People use it when they want to describe a certain situation or process but they do not know it precisely. Therefore, the most important role of the linguistic description is to provide a conclusion about consequent objects Y when an information about antecedent objects X is given. Such an information has a character of *perception* of properties of the latter objects and so, the corresponding procedure is called *perception-based* logical deduction.

On the basis of the theory presented in [8], the following special inference rule of perception-based logical deduction can be introduced. Let *LD* be a linguistic description consisting of rules of the form (1) and $x \in$ *Form*$_\alpha$, $y \in$ *Form*$_\beta$, $w \in$ *Form*$_{\alpha o}$, $w' \in$ *Form*$_{\beta o}$. Then the following scheme is a valid special inference rule:

$$r_{PbLD} : \frac{LPerc^{LD} wx Ev_i^A, \quad LD}{Eval\, w'\, \hat{y}_i Ev_i^C} \tag{2}$$

where $\hat{y}_i \equiv \imath y \cdot Ev_i^A wx \Rightarrow Ev_i^C w'\, y$, $i \in \{1, \ldots, m\}$, $T \vdash Topic^{LD} Ev^A$ and $T \vdash Focus^{LD} Ev^C$.

This rule has the following interpretation: Let *LD* be a linguistic description consisting of fuzzy IF-THEN rules of the form (1). If we find a formula $\mathrm{Int}(\mathcal{A}_i) \equiv Ev_i^A$ of some expression from the topic *Topic*^{LD} and an element $\mathbf{u}^0$ in the context $\mathbf{w}^0$ such that $Ev_i^A \mathbf{w}^0 \mathbf{u}^0$ has a non-zero truth degree then (denoting $\mathbf{b}_i^0 \equiv Ev_i^A \mathbf{w}^0 \mathbf{u}^0$) we conclude that the element $\imath y \cdot \mathbf{b}_i^0 \Rightarrow Ev_i^C w'y$ which is *typical for the formula* $\mathbf{b}_i^0 \Rightarrow Ev_i^C w'y$, is evaluated by the linguistic expression $Ev_i^C \in Focus^{LD}$ in every context w'.

Note that the main result of r_{PbLD} is the element $\hat{y}_i$. Then in every model $\mathcal{M}$ we can find a specific element $\mathcal{I}_p^{\mathcal{M}}(\hat{y}_i) = v \in M_\beta$ using the operation $\mathcal{I}_p^{\mathcal{M}}(\iota_{\beta(o\beta)})$ which in fuzzy set theory is just the defuzzification function. In our case, the DEE method (Defuzzification of linguistic expressions) should be used. The detailed, less formal explanation of the perception-based logical deduction including examples is presented in [12].

Human reasoning that is based on a complex of experience, observation, logical reasoning and world knowledge is necessarily non-monotonic. Hence, our model must include also the theory of *non-monotonic reasoning* (cf. [2, 11]). We deal with a class of theories that themselves are consistent but when using them simultaneously, we may arrive at a contradiction or, at least, to a non-desirable result. Therefore, we consider a special preference relation that tells us which theory should be used in the given state (called *belief state* in [2]). At each state, we work in a special theory which in our case is determined by a linguistic description (one or more) and possibly also by some perception (recall that this is a formula representing intension of some evaluating linguistic expression).

In this paper, we are not going into details of this approach. We will suppose that our formal theories contain only those

parts of the theory of evaluating expressions T^{Ev}, that are necessary for deductions based on used linguistic descriptions and perceptions.

3 Example of a model of economic analysis

We present a model of a macroeconomic situation using PbLD. We use a part from free economical analysis of Czech Savings Bank for the fourth quarter of 2006:[†)]

> Our computations show, that acceleration of economy by one per cent causes unemployment rate decrease only by 0.3 per cent. Non-motivating social system, rigid labor market and high tax rate on labor expenses are by our opinion main culprits of structurally high rate of unemployment, although economy grows at high rate.

We will present an analysis of this quotation by means of the above outlined theory.

3.1 Method

As mentioned, our theory can be classified as a part of the methodology introduced by L. A. Zadeh in his papers [14, 15] and called *Precisiated Natural Language* (PNL). The latter is an attempt at developing a unified formalism for various tasks involving natural language propositions.

Two main premises of PNL are the following:

(a) Much of the world knowledge is perception based.

(b) Perception based information is intrinsically fuzzy.

The PNL methodology requires presence of the, so-called, *World Knowledge Database* (WKDB) which contains all the necessary information including perception based propositions describing the knowledge acquired by direct human experience and which can be used in the deduction process. A *multiagent, modular deduction database* (MDE) contains various rules of deduction.

Our version of PNL incorporates logical machinery. The translation from natural language to fuzzy IF-THEN rules is so far done manually. The formal frame is Łukasiewicz fuzzy type theory (Ł-FTT) and theories of evaluating linguistic expressions and fuzzy IF-THEN rules as described in the previous sections. Moreover, we apply some principles of non-monotonic reasoning. However, we are not going into details of the role of it in this paper.

We model the above natural-language macroeconomic analysis using two linguistic descriptions with hierarchic structure. We will use an intermediate variable *strength of auxiliary factors* for the overall influence of the non-motivating social system, rigid labor market and high tax rate. Hence, we will introduce linguistic description LD_{Aux} with three antecedent variables and one consequent variable, the above mentioned *strength of auxiliary factors*. This variable is then used as antecedent variable to the second linguistic description LD_{Un}, with the second antecedent variable *rate of economy acceleration*. The consequent variable in this linguistic description is *rate of unemployment change* – the result.

3.2 World knowledge

(i) $\vartheta \in$ *Types* represents a general feature of objects that can be characterized using grades. In the model, a set of this type can be, e.g., a subset of the real numbers.

(ii) $\beta \in$ *Types* represent objects of type (national) state.[†)]

(iii) $\gamma \in Type$ represents general abstract objects.

(iv) $Ss \in$ *Form*$_{\vartheta\beta}, Lm \in$ *Form*$_{\vartheta\beta}, Tl \in$ *Form*$_{\vartheta\beta}, Uc \in$ *Form*$_{\vartheta\beta}, Ea \in$ *Form*$_{\vartheta\beta}, Af \in$ *Form*$_{\vartheta\gamma}$, are special formulas characterizing state of social system, rigidity of labor market, level of tax load, unemployment change, economy acceleration (of objects of type β) and auxiliary factors (of object of type γ), respectively.

(v) In correspondence with the previous item, we also need to consider contexts $w_{Ss}, w_{Lm}, w_{Tl}, w_{Uc}, w_{Ea}, w_{Af} \in$ *Form*$_{\vartheta o}$.

(vi) Special constant concerning the story, namely: $\mathbf{c}_{CR} \in$ *Form*$_{\beta}$ representing in our case *the Czech Republic*. However, linguistic description below should be valid also for other countries.

3.3 Formalization and reasoning

(i) *Contexts.* Context can be either *simple* (items (a), (b), (c), (e)) or *two-sided* (items (d), (f)), see Section 2.2. Two-sided contexts are used for variables which can take both positive and negative values.

(a) Motivation of social system: $w_{Ss} = \langle 0, 0.5, 1\rangle$ (abstract degrees).

(b) Rigidity of labor market: $w_{Lm} = \langle 0, 0.5, 1\rangle$ (abstract degrees).

(c) Strength of tax load: $w_{Tl} = \langle 0, 15, 40\rangle$ (%).

(d) Rate of unemployment change: $w_{Uc} = \langle -1, -0.5, 0, 0.5, 1\rangle$ (%).

(e) Strength of auxiliary factors: $w_{Af} = \langle 0, 0.5, 1\rangle$ (abstract degrees).

(f) Rate of economy acceleration: $w_{Ea} = \langle -6, -3, 0, 5, 10\rangle$ (%).

[†)] http://www.csas.cz/banka/content/inet/internet/cs/treasury_ie.xml, in Czech

[†)] This type could also represent other territorial objects, e.g., regions, European Union etc.

(ii) *Linguistic descriptions*

(a) Linguistic description LD_{Aux} for the influence of auxiliary factors:[‡)]

$$\mathsf{IF}\ X_{Ss}\ \text{is}\ Sm\ \mathsf{AND}\ X_{Lm}\ \text{is}\ Bi\ \mathsf{AND}\ X_{Tl}\ \text{is}\ Bi\ \mathsf{THEN}\ X_{Af}\ \text{is}\ Bi \quad (3)$$

. . . . . .

The corresponding intensions of such rules are

$$\lambda w_{Ss}\,\lambda w_{Lm}\,\lambda w_{Tl}\,\lambda w_{Af}\cdot\lambda x_{1,\vartheta}\,\lambda x_{2,\vartheta}\,\lambda x_{3,\vartheta}\,\lambda y_{\vartheta} \cdot(Ev_1^A w_{Ss}\,x_{1,\vartheta})\wedge(Ev_2^A w_{Lm}\,x_{2,\vartheta})\wedge(Ev_3^A w_{Tl}\,x_{3,\vartheta}) \Rightarrow Ev^C w_{Af}\,y_{\vartheta}. \quad (4)$$

(b) Linguistic description LD_{Un} for the computation of unemployment change.

$$\mathsf{IF}\ X_{Ea}\ \text{is}\ +Bi\ \mathsf{AND}\ X_{Af}\ \text{is}\ Bi\ \mathsf{THEN}\ X_{Uc}\ \text{is}\ -Sm \quad (5)$$

. . . . . .

The corresponding intensions of such rules are

$$\lambda w_{Ea}\,\lambda w_{Af}\,\lambda w_{Uc}\cdot\lambda x_{1,\vartheta}\,\lambda x_{2,\vartheta}\,\lambda y_{\vartheta} \cdot(Ev_1^A w_{Ea}\,x_{1,\vartheta})\wedge(Ev_2^A w_{Af}\,x_{2,\vartheta}) \Rightarrow Ev^C w_{Uc}\,y_{\vartheta}. \quad (6)$$

(iii) Furthermore, we must construct a specific *model*. This will be based on a frame

$$\mathcal{M} = \langle (M_\alpha, =_\alpha)_{\alpha\in Types}, \mathcal{L}_\Delta\rangle.$$

where $M_o = [0,1]$ and $M_\alpha = \mathbb{R}$ for $\alpha = \vartheta$.

(iv) *Perceptions*

(P1) Social system is non-motivating. We can say that motivation of social system is *very small*. Formally,

$$\mathcal{I}_p^{\mathcal{M}}(LPerc\ w_{Ss}(Ss\mathbf{c}_{CR})\ Ve\ Sm) = 1. \quad (7)$$

(P2) Labor market is rigid. It means that rigidity of labor market is *big*. Formally,

$$\mathcal{I}_p^{\mathcal{M}}(LPerc\ w_{Lm}(Lm\mathbf{c}_{CR})\ Bi) = 1. \quad (8)$$

(P3) Tax load is high. Formally,

$$\mathcal{I}_p^{\mathcal{M}}(LPerc\ w_{Tl}(Ss\mathbf{c}_{CR})\ Bi) = 1. \quad (9)$$

(P4) Economy acceleration is positive big. Formally,

$$\mathcal{I}_p^{\mathcal{M}}(LPerc\ w_{Ea}(Ss\mathbf{c}_{CR}) + Bi) = 1. \quad (10)$$

(v) *Deductions* We use a formal theory T^{Uc} for the reasoning about our example. It includes appropriate parts of theory of evaluating expressions T^{Ev}, intensions of linguistic descriptions LD_{Aux} and LD_{Un} and perceptions (P1)–(P4).

$$T^{Uc} = \{LPerc\ w_{Ss}(Ss\mathbf{c}_{CR})\ Ve\ Sm,\ LPerc\ w_{Lm}(Lm\mathbf{c}_{CR})\ Bi,\ LPerc\ w_{Tl}(Ss\mathbf{c}_{CR})\ Bi,\ LPerc\ w_{Ea}(Ss\mathbf{c}_{CR}) + Bi, (4), (6)\} \cup T^{Ev} \quad (11)$$

Below, we give the main ideas of proofs.

(a)

$$T^{Uc} \vdash LPerc^{LD_{Aux}} w_{Ss}(Ss\mathbf{c}_{CR})\ Sm \quad (12)$$

It means that *small* is the perception of variable *motivation of social system* with respect to linguistic description LD_{Aux}.

PROOF: Note that there is evaluating expression $Ve\,Sm$ in perception (P1) and evaluating expression Sm in IF-THEN rule (3). Suppose that in linguistic description LD_{Aux} is no other rule which has in the variable X_{ss} evaluating expression$\langle$linguistic hedge$\rangle\langle$atomic expression$\rangle$, where $\langle$atomic expression$\rangle$ is *small*. Then (12) follows from the definition of *LPerc*. □

(b)

$$T^{Uc} \vdash LPerc^{LD_{Aux}} w_{Lm}(Lm\mathbf{c}_{CR})\ Bi\,. \quad (13)$$

It means that *big* is the perception of variable *rigidity of labor market* with respect to linguistic description LD_{Aux}.

PROOF: In (b) and (c), evaluating expression in perception (P2) and (P3) is equal to evaluating expression in corresponding variables of (3), respectively. Proofs are analogous to (a). □

(c)

$$T^{Uc} \vdash LPerc^{LD_{Aux}} w_{Tl}(Tl\mathbf{c}_{CR})\ Bi\,. \quad (14)$$

It means that *big* is the perception of variable *tax load* with respect to linguistic description LD_{Aux}.

(d)

$$T^{Uc} \vdash Eval\ w_{Af}(Af\mathbf{c}_{AF})\ Bi\,. \quad (15)$$

This deduction means that $\mathbf{c}_{AF}$ is evaluated by evaluating expression *big*. It represents the result

[‡)] Only one rule based on natural-language macroeconomic analysis is given here. However, more similar rules can be present.

of deduction over linguistic description LD_{Aux} – *strength of auxiliary factors* is *big*.

PROOF: Follows from r_{PbLD} and (4), where $\mathbf{c}_{AF}$ is a new constant representing the result of this deduction. □

(e)

$$T^{Uc} \vdash \mathit{LPerc}^{LD_{Un}} w_{Af}(Af\,\mathbf{c}_{AF})\; Bi \qquad (16)$$

It means that *big* is the perception of variable *strength of auxiliary factors* with respect to linguistic description LD_{Un}.

PROOF: Suppose that in linguistic description LD_{Un} is no other rule which has in the variable X_{af} evaluating expression *big*. Then (16) follows from the definition of *LPerc*. □

(f)

$$T^{Uc} \vdash \mathit{Eval}\; w_{Uc}(Uc\,\mathbf{c}_{UC})(-\,Sm) \qquad (17)$$

This deduction means that $\mathbf{c}_{UC}$ is evaluated by evaluating expression *negative small*. It represents the result of deduction over linguistic description LD_{Un} – *rate of unemployment change* is *negative small*.

PROOF: Follows from r_{PbLD} and (6), where $\mathbf{c}_{UC}$ is a new constant. □

Hence, we can conclude that change of unemployment is *negative small*.

3.4 Remarks

- In our example, there was one difference between evaluating expression from perception and evaluating expression in the corresponding variable in the linguistic description (namely *very small* in (P1) and *small* in (3)). Overall result (change of unemployment is *negative small*) corresponds to the evaluating expression in the consequent of IF-THEN rule (5). Hence, the result could seemingly be obtained without complex formal apparatus. However, more complicated situations (much more variables, linguistic descriptions etc.) are usually not transparent so much. Then, our formal apparatus is not only useful but necessary.
- An implementation of the theory of evaluating expressions and perception-based logical deduction is available as a part of the software system LFLC (Linguistic Fuzzy Logic Controller, demo version available at `http://irafm.osu.cz`) [4]. This system is designed in such a way that the user is not supposed to be aware of the complicated formal systems presented in previous sections. If he/she provides contexts, linguistic descriptions in the form of IF-THEN rules and perceptions, then deductions from subsection 3.3 are performed automatically. Hence, the method described in this paper is accessible also for users without advanced knowledge of fuzzy logic.

4 Conclusions

In this paper, a model of concrete macroeconomic analysis using perception-based logical deduction was presented. In the further research, we will concentrate on:

- Modeling of more complicated economic analyses.
- Methods for automated or semi-automated extraction of linguistic descriptions and deductions.
- Incorporation of further parts of theory of non-monotonic reasoning.

Acknowledgement

This work was supported by project MSM 6198898701 of the MŠMT ČR.

References

[1] P. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof*. Kluwer, Dordrecht, 2002.

[2] A. Bochman. *A logical theory of nonmonotonic inference and belief change*. Springer, Berlin, 2001.

[3] B. Diaz and T. Takagi. PNL applied to economics. In *Proc. Int. Conf. FUZZ-IEEE'2005*, Reno, USA, May 2005.

[4] A. Dvořák, H. Habiballa, V. Novák, and V. Pavliska. The software package LFLC 2000 - its specificity, recent and perspective applications. *Computers in Industry*, 51:269–280, 2003.

[5] P. Hájek. *Metamathematics of Fuzzy Logic*. Kluwer, Dordrecht, 1998.

[6] V. Lifschitz. *Formalizing Common Sense: Papers by John McCarthy*. Ablex Publishing, Norwood, 1990.

[7] V. Novák. On fuzzy type theory. *Fuzzy Sets and Systems*, 149:235–273, 2005.

[8] V. Novák. Perception-based logical deduction. In B. Reusch, editor, *Computational Intelligence, Theory and Applications*, pages 237–250. Springer, Berlin, 2005.

[9] V. Novák. Fuzzy logic theory of evaluating expressions and comparative quantifiers. In *Proc. 11th Int. Conf. IPMU, Paris, July 2006*, volume 2, pages 1572–1579, Les Cordeliers, Paris, 2006. Éditions EDK.

[10] V. Novák. A comprehensive theory of evaluating linguistic expressions. In S. Gottwald, Hájek. P., U. Höhle, and E. P. Klement, editors, *Proc. of 26th Linz Seminar on Fuzzy Set Theory*. (to appear). Research Report No. 71, http://irafm.osu.cz/.

[11] V. Novák and A. Dvořák. A model of complex human reasoning in fuzzy logic in broader sense. *Fuzzy Sets and Systems*, submitted.

[12] V. Novák and I. Perfilieva. On the semantics of perception-based fuzzy logic deduction. *International Journal of Intelligent Systems*, 19:1007–1031, 2004.

[13] Richmond Thomason. Logic and artificial intelligence. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2005.

[14] L. A. Zadeh. A note on web intelligence, world knowledge and fuzzy logic. *Data & Knowledge Engineering*, 50:291–304, 2004.

[15] L. A. Zadeh. Precisiated natural language. *AI Magazine*, 25:74–91, 2004.

Advanced Inference in Fuzzy Systems by Rule Base Compression

Alexander Gegov
University of Portsmouth
School of Computing, Buckingham Building
Portsmouth PO1 3HE, United Kingdom
alexander.gegov@port.ac.uk

Neelamugilan Gobalakrishnan
University of Portsmouth
School of Computing, Buckingham Building
Portsmouth PO1 3HE, United Kingdom
muges_3106@hotmail.com

Abstract

This paper describes a method for formal compression of fuzzy systems. This method compresses a fuzzy system with an arbitrarily large number of rules into a smaller fuzzy system by removing the redundancy in the fuzzy rule base. As a result of this compression, the number of on-line operations during the fuzzy inference process is significantly reduced without compromising the solution. This rule base compression method and its software implementation outperform significantly than all other known methods for fuzzy rule base reduction.

Keywords: Fuzzy systems, rule base reduction, rule base compression.

1 Introduction

Information based decision making is usually accompanied by uncertainty which is inherent to the environment in which the information is being gathered. Such uncertainty may seriously compromise the reliability of the information gathering process as well as the quality of any subsequent decisions made.

Fortunately, fuzzy systems are well suited for decision making tasks characterised by uncertainty [6]. The latter can be taken into account by means of the approximate reasoning and logical inference capabilities of fuzzy systems. However, there is often a problem in this case caused by the large number of rules which depends on the number of information based inputs. This usually leads to a significant increase in the qualitative complexity in terms of poor transparency and unclear interpretation of the fuzzy rules as well as the quantitative complexity in terms of increased number of operations during the fuzzy inference process. This point is illustrated further below by Equation (1).

A fuzzy system is usually represented by if-then rules of the form:

If i_1 is $v_{i1,1}$ and … and i_m is $v_{im,1}$
then o_1 is $v_{o1,1}$ and … and o_n is $v_{on,1}$
… …………………………….. .. (1)
If i_1 is $v_{i1,r}$ and … and i_m is $v_{im,r}$
then o_1 is $v_{o1,r}$ and … and o_n is $v_{on,r}$

Where m is the number of information based inputs, n is the number of decision making outputs and r is the number of fuzzy rules in the system. In this case, i_p, $p=1,m$ represents the p-th input, $v_{ip,s}$ $p=1,m$, $s=1,r$ is the linguistic value of the p-th input in the s-th rule, o_q, $q=1,n$ represents the q-th output and $v_{oq,s}$ $q=1,n$, $s=1,r$ is the linguistic value of the q-th output in the s-th rule.

The number of rules in a fuzzy system r is an exponential function of the number of the inputs m and the number of linguistic values k that these inputs can take [3]. In most cases, this exponential function is in the form

$$r = k^m \quad (2)$$

and it is shown explicitly in Figure 1.

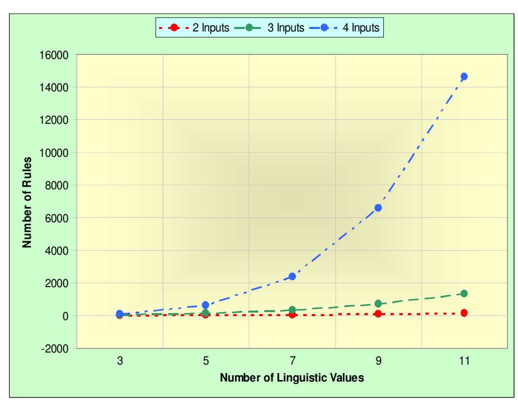


Figure 1: Number of rules for two, three and four inputs

So, the question that arises here is how to use fuzzy systems for tackling uncertain information without

making them fall into the trap of complexity and thus compromising their suitability for tackling the uncertainty in the first place. The assumption made in this case is that when it is impossible to improve the quality of information due to time or operation related constraints, it should still be possible to use this information in a reliable way by means of an enhanced decision making process which utilises the capabilities of fuzzy systems for dealing with uncertainty and simplifies their complexity at the same time [1].

2 Operation stages in fuzzy systems

Fuzzy inference is a process of mapping a given input to an output using the theory of fuzzy sets, as shown in Figure 2. The most commonly used fuzzy inference technique is the so-called Mamdani method and this is why it is used in this paper. This method consists of 3 major stages - fuzzification, inference and defuzzification. In the fuzzification stage, the authors have decided to use the two most widely used types of membership functions - triangular and trapezoidal. The inference stage is divided into 3 sub stages - application, implication and aggregation. The authors have decided to use the conjunctive method (MIN) in the application stage, the truncation method in the implication stage, and disjunctive method (MAX) in the aggregation stage. In the defuzzification stage, the authors have decided to apply the most widely used centroid method. In order to develop this artefact, the authors have decided to use the MATLAB Fuzzy Logic Toolbox due to its wide applicability in both academia and industry.

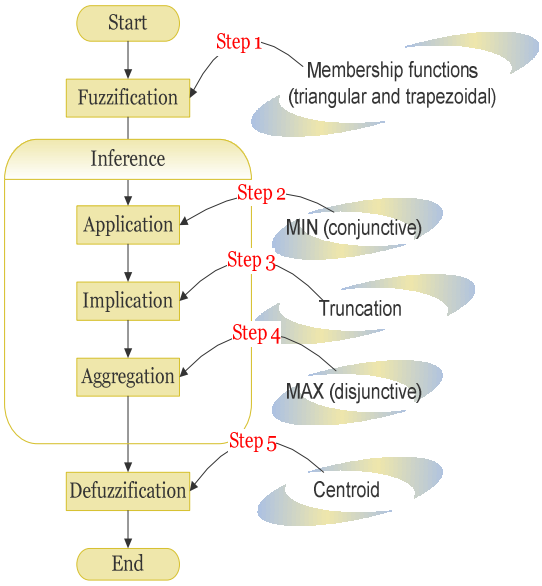


Figure 2: Operation stages in fuzzy systems

3 Algorithmic design of the rule base compression method

A detailed algorithm for rule base compression is shown in Figure 3.

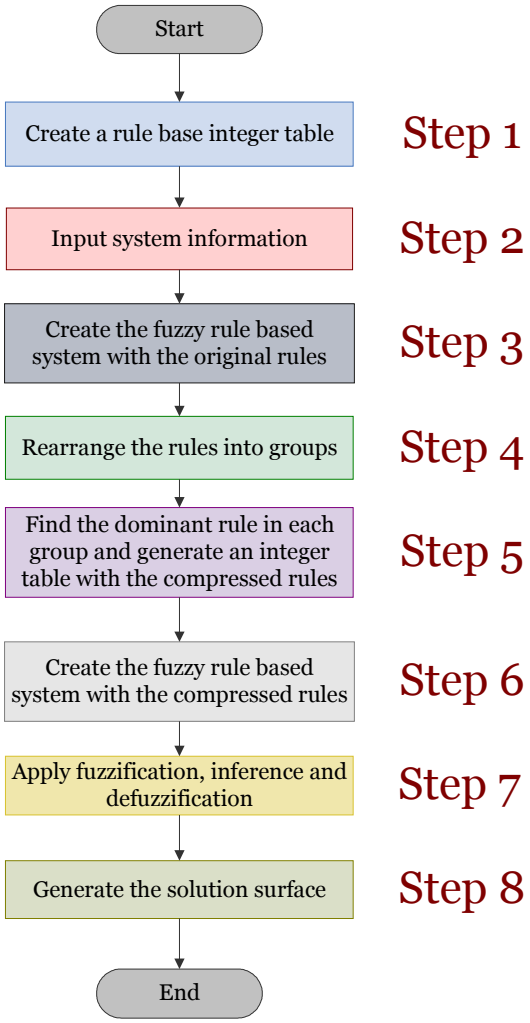


Figure 3: Algorithm for rule base compression

The above algorithm includes eight major steps. The first step is based on a dialogue with the user who is prompted to enter all the information about the rule base. The user is prompted to enter information such as the number of inputs and outputs, as well as the number of linguistic values for each input and output. The user also has to enter the output value for each possible combination of input linguistic values, which is displayed by the system. After the acquisition of all the necessary information, the algorithm creates an integer table with positive integer numbers.

The second step is again based on dialogue with the user who is prompted to enter all the fuzzy inference system information such as variable names and membership function definitions. In the third step, the software creates and saves the fuzzy system with the original rules.

The fourth step is to re-arrange the rules into groups. These groups are sorted in an increasing order with respect to the chosen output linguistic values. This aggregation process can be carried out entirely off-line. The next step is to find the dominant rule for each group. This step can only be applied on-line, which is due to the fact that the dominant rules can be found only after the completion of the fuzzification and application stages.

In the sixth step, the software creates a fuzzy rule base system with the compressed rules and saves it. Step seventh is the system evaluation process which uses the file saved in step six. This process evaluates the output of a fuzzy rule based system for given inputs. The final step is to generate the solution surface. This solution surface is created through a given number of points for the crisp input values and the defuzzified output values.

4 Software implementation of the rule base compression method

The MATLAB Fuzzy Logic Toolbox has been used for the software implementation of the rule base compression method. The software has been implemented with full functionality [2]. The implemented software is discussed in more detail below, where the rule base compression method is illustrated for a case study taken from [4].

A fuzzy system for the operation of a service centre for spare parts is described with 3 inputs i_1, i_2, i_3 and one output o_1. Whereby i_1 is the repair utilisation factor, i_2 is the number of servers, i_3 is the mean delay of service and o_1 is the number of spare parts. In this case, i_1 can take the three linguistic values low (L) = 1, medium (M) = 2 and high (H) = 3, i_2 can take the three linguistic values small (S) = 1, medium (M) = 2 and large (L) = 3, i_3 can take the three linguistic values very short (VS) = 1, short (S) = 2 and medium (M) = 3, whereas o_1 can take the seven linguistic values very small (VS) = 1, small (S) = 2, rather small (RS) = 3, medium (M) = 4 rather large (RL) = 5, large (L) = 6 and very large (VL) = 7.

The integer tables for the original and the compressed fuzzy system are shown below in Tables 1-2.

Table 1: Integer table for the original fuzzy system

Rule number	Input i_1	Input i_2	Input i_3	Output o_1
1	1	1	1	1
2	1	1	2	1
3	1	1	3	1
4	1	2	1	1
5	1	2	2	1
6	1	2	3	1
7	1	3	1	2
8	1	3	2	2
9	1	3	3	1
10	2	1	1	2
11	2	1	2	1
12	2	1	3	1
13	2	2	1	3
14	2	2	2	2
15	2	2	3	1
16	2	3	1	4
17	2	3	2	3
18	2	3	3	2
19	3	1	1	7
20	3	1	2	6
21	3	1	3	4
22	3	2	1	4
23	3	2	2	4
24	3	2	3	2
25	3	3	1	5
26	3	3	2	4
27	3	3	3	3

Table 2: Integer table for the compressed fuzzy system

Rule number	Input i_1	Input i_2	Input i_3	Output o_1
1 or 2 or 3 or 4 or 5 or 6 or 9 or 11 or 12 or 15	1 or 2	1 or 2 or 3	1 or 2 or 3	1
7 or 8 or 10 or 14 or 18 or 24	1 or 2 or 3	1 or 2 or 3	1 or 2 or 3	2
13 or 17 or 27	2 or 3	2 or 3	1 or 2 or 3	3
16 or 21 or 22 or 23 or 26	2 or 3	1 or 2 or 3	1 or 2 or 3	4
25	3	3	1	5
20	3	1	2	6
19	3	1	1	7

The integer Table 2 above is shown in the general form, whereby the rule numbers and the inputs are given is compound expressions of disjunctive terms.

Several output surfaces for these two systems are shown below whereby input 1 is fixed to 0, 0.5 and 1. These output surfaces are shown in Figures 4-9 and their numeric values are shown in Figures 10–15.

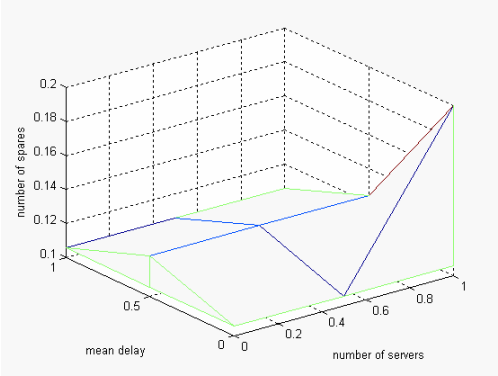


Figure 4: Output surface for the original system with input 1 fixed to value 0

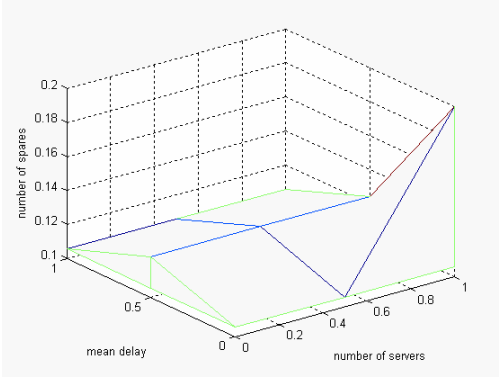


Figure 5: Output surface for the compressed system with input 1 fixed to value 0

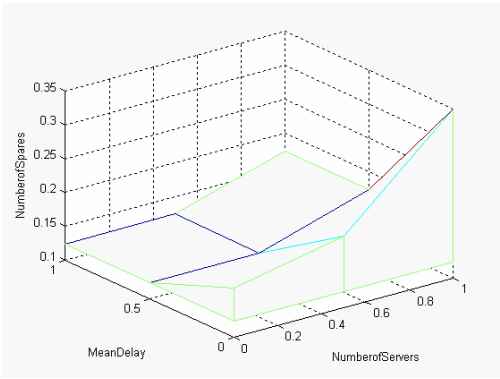


Figure 6: Output surface for the original system with input 1 fixed to value 0.5

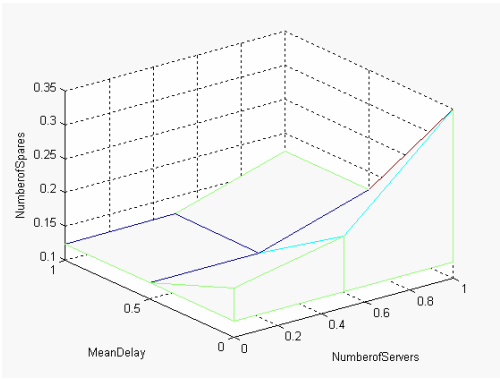


Figure 7: Output surface for the compressed system with input 1 fixed to value 0.5

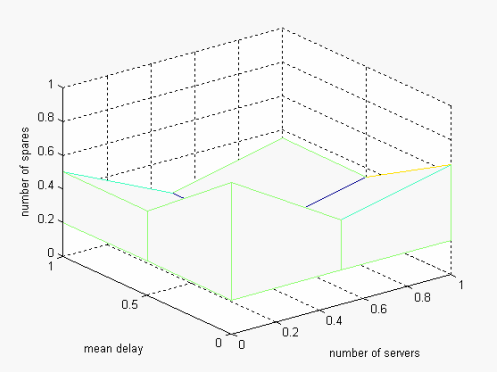


Figure 8: Output surface for the original system with input 1 fixed to value 1

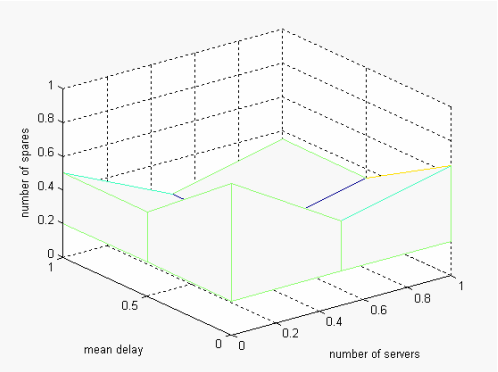


Figure 9: Output surface for the compressed system with input 1 fixed to value 1

The output surface values when :
Input1 fix to value: 0.00

Input2	Input3	Output1
0	0	0.1056
0.5000	0	0.1056
1.0000	0	0.2000
0	0.5000	0.1241
0.5000	0.5000	0.1241
1.0000	0.5000	0.1241
0	1.0000	0.1056
0.5000	1.0000	0.1056
1.0000	1.0000	0.1056

Figure 10: Numeric values for the original system with input 1 fixed to value 0

The output surface values when :
Input1 fix to value: 0.00

Input2	Input3	Output1
0	0	0.1056
0.5000	0	0.1056
1.0000	0	0.2000
0	0.5000	0.1241
0.5000	0.5000	0.1241
1.0000	0.5000	0.1241
0	1.0000	0.1056
0.5000	1.0000	0.1056
1.0000	1.0000	0.1056

Figure 11: Numeric values for the compressed system with input 1 fixed to value 0

The output surface values when :
Input1 fix to value: 0.50

Input2	Input3	Output1
0	0	0.1737
0.5000	0	0.2065
1.0000	0	0.3500
0	0.5000	0.1241
0.5000	0.5000	0.1241
1.0000	0.5000	0.1737
0	1.0000	0.1241
0.5000	1.0000	0.1241
1.0000	1.0000	0.1737

Figure 12: Numeric values for the original system with input 1 fixed to value 0.5

The output surface values when :
Input1 fix to value: 0.50

Input2	Input3	Output1
0	0	0.1737
0.5000	0	0.2065
1.0000	0	0.3500
0	0.5000	0.1241
0.5000	0.5000	0.1241
1.0000	0.5000	0.1737
0	1.0000	0.1241
0.5000	1.0000	0.1241
1.0000	1.0000	0.1737

Figure 13: Numeric values for the compressed system with input 1 fixed to value 0.5

The output surface values when :
Input1 fix to value: 1.00

Input2	Input3	Output1
0	0	0.8944
0.5000	0	0.5000
1.0000	0	0.6500
0	0.5000	0.5000
0.5000	0.5000	0.2000
1.0000	0.5000	0.3500
0	1.0000	0.5000
0.5000	1.0000	0.2000
1.0000	1.0000	0.3500

Figure 14: Numeric values for the original system with input 1 fixed to value 1

The output surface values when :
Input1 fix to value: 1.00

Input2	Input3	Output1
0	0	0.8944
0.5000	0	0.5000
1.0000	0	0.6500
0	0.5000	0.5000
0.5000	0.5000	0.2000
1.0000	0.5000	0.3500
0	1.0000	0.5000
0.5000	1.0000	0.2000
1.0000	1.0000	0.3500

Figure 15: Numeric values for the compressed system with input 1 fixed to value 1

More output surfaces for both systems are shown in Figures 16-21 whereby input 1 is fixed to 0, 0.5 and 1, whereas the variation of input 2 and input 3 is represented with 30 points each. The number of points for each input has been chosen to be 10 times greater than the corresponding number from the previous output surfaces (input 2 = 3x10 and input 3 = 3x10).

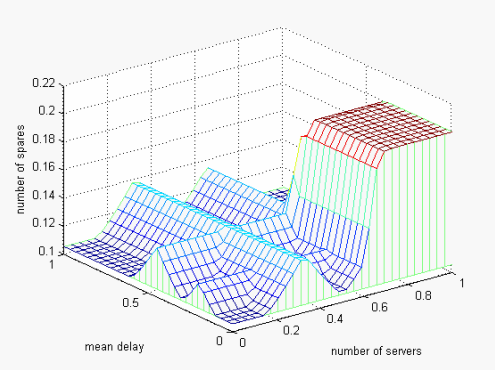


Figure 16: Output surface for the original system with input 1 fixed to value 0

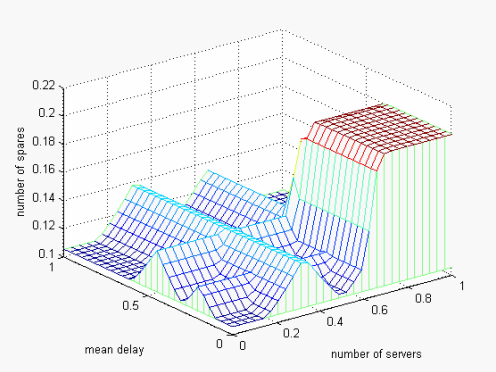


Figure 17: Output surface for the compressed system with input 1 fixed to value 0

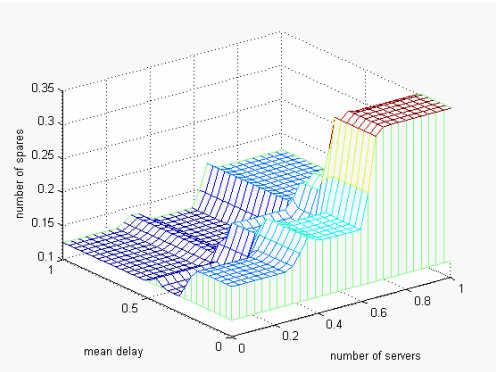


Figure 18: Output surface for the original system with input 1 fixed to value 0.5

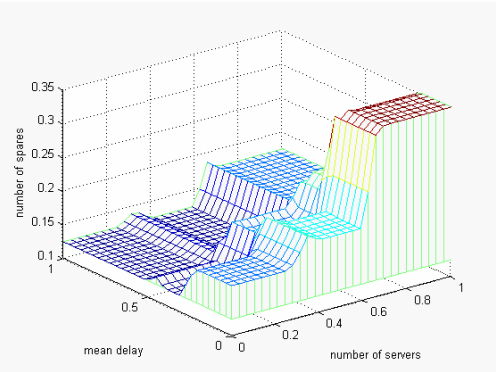


Figure 19: Output surface for the compressed system with input 1 fixed to value 0.5

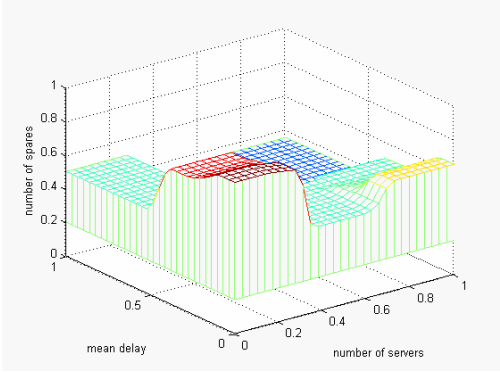


Figure 20: Output surface for the original system with input 1 fixed to value 1

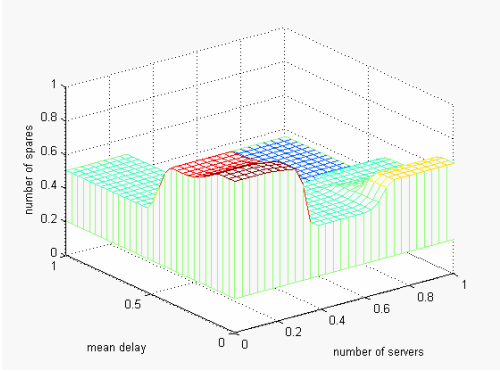


Figure 21: Output surface for the compressed system with input 1 fixed to value 1

5 Comparative evaluation

The evaluation approach used here is based on precise calculations and it is superior to the well known approximate BIG (O) approach. The proposed rule base compression method has been compared with the hierarchical method, which is the most advanced method available for rule base reduction due to its wide applicability [6].

The fuzzy systems implementing these two methods are compared in terms of exact amount of on-line operations, which determine the overall number of elementary operations (EO) such as addition, subtraction, multiplication, division and comparison.

The quantitative complexity for both systems is calculated for each stage and sub-stage in fuzzy rule based system such as fuzzification (EO_{FU}), inference that includes application (EO_{AP}), implication (EO_{IM}) and aggregation (EO_{AG}), and defuzzification (EO_{DE}). The compressed system has an additional process to determine the dominant rules (EO_{CO}).

The quantitative complexity is calculated for the hierarchical and the compressed system, as shown by the Equations (3)-(4)

$$EO^{HS} = EO_{FU}^{HS} + EO_{AP}^{HS} + EO_{IM}^{HS} + EO_{AG}^{HS} + EO_{DE}^{HS} = (m - 1) . [(8 . t + 1) . w^2 + 12 . w + 2 . t - 1] . n . h \quad (3)$$

$$EO^{FS} = EO_{FU}^{FS} + EO_{AP}^{FS} + EO_{IM}^{FS} + EO_{AG}^{FS} + EO_{DE}^{FS} + EO_{CO}^{FS} = (m - 1) . [(8 . t + 1) . w^2 + 12 . w + 2 . t - 1] . n . h \quad (4)$$

where m is number of inputs, w is number of linguistic values per input, n is number of outputs, t is number of elements in the discrete universe discourse for the output and h is number of simulation cycles.

The results from the comparative evaluation of the quantitative complexity for a hierarchical and a compressed system are presented in Table 3 and Figure 22 whereby a hierarchical system implements the hierarchical method and a compressed system implements the proposed method of rule base compression.

Table 3: Complexity of the hierarchical and the compressed system

Number of rules /fuzzy system	Hierarchical system	Compressed system
3^2 = 9	562	232
3^3 = 27	1,124	295
3^4 = 81	1,686	466
5^2 = 25	2,710	650
5^3 = 125	5,420	905
5^4 = 625	8,130	2,160
7^2 = 49	7,618	1,276
7^3 = 343	15,236	1,955
7^4 = 2,401	22,854	6,750
9^2 = 81	16,438	2,110
9^3 = 729	32,876	3,541
9^4 = 6,561	49,314	16,636
11^2 = 121	30,322	3,152
11^3 = 1,331	60,644	5,759
11^4 = 14,641	90,966	34,986

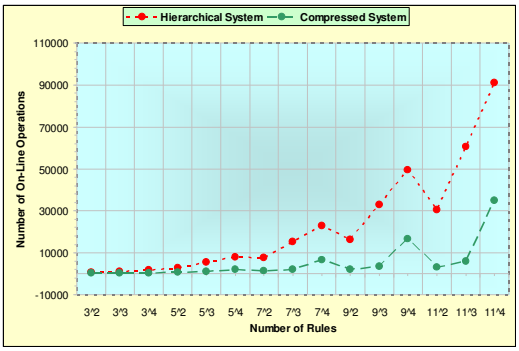


Figure 22: Complexity of the hierarchical and the compressed system

Table 3 and Figure 22 clearly show that the compressed system is superior to the hierarchical system for all considered permutations of inputs and their linguistic values. These permutations have been chosen on the basis of the most commonly used applications of fuzzy systems.

6 Conclusion

The proposed method and its programming implementation compress a fuzzy system with an arbitrarily large number of rules into a smaller fuzzy system by removing the redundancy in the fuzzy rule base. As a result of this compression, the number of on-line operations during the fuzzy inference process is substantially reduced without compromising the solution. This method outperforms significantly all other known methods for fuzzy rule base reduction.

References

[1] A.Gegov, *Complexity Management in Fuzzy Systems* (Springer, Berlin, 2007).

[2] N.Gobalakrishnan, *Software Tool for Advanced Inference in Fuzzy Systems* (BSc Thesis, University of Portsmouth, 2006)

[3] M.Jamshidi, *Large Scale Systems: Modelling, Control and Fuzzy Logic* (Prentice-Hall, Englewood Cliffs, 1997)

[4] M.Negnevitsky, *Artificial Intelligence: A Guide to Intellegent Systems* (Pearson, Harlow, 2002)

[5] G.Raju, J.Zhou and R.Kisner, Hierarchical fuzzy control, *International Journal of Control*, 54/5 (1991), 1201-1216

[6] T.Ross, *Fuzzy Logic with Engineering Applications* (John Wiley & Sons, Chichester, 2004).

Using Generalized Constraints and Protoforms to deal with adverbs

Andrés Soto
Computer Science Department,
Universidad Autónoma del Carmen, Ave. 31,
Ciudad del Carmen, CP 24160
Campeche, México
soto_andres@yahoo.com

José A. Olivas
Information Technologies and Systems Department
University of Castilla - La Mancha,
Paseo de la Universidad 4,
13071 Ciudad Real, Spain
JoseAngel.Olivas@uclm.es

Manuel E. Prieto
Information Technologies and Systems Department
University of Castilla - La Mancha,
Paseo de la Universidad 4,
13071 Ciudad Real, Spain
Manuel.Prieto@uclm.es

Abstract

Computation with information described in natural language (NL) has intrinsic importance because much of human knowledge is described using these languages. Soft Computing approach to NL-Computation concerns with semantic imprecision of natural languages through the use of NL precisiation, generalized constraints and prototypical forms. NLs are basically systems for describing perceptions. NL precisiation consist on expressing the meaning of propositions in NL as generalized constraints. In this paper, adverbs and adjectives in NL propositions are considered as constraints. Thus, generalized constraints are proposed to specify the attribute values associated to nouns by means of adverbs and adjectives. A prototypical form (or *protoform*) is suggested to represent an abstracted summary, a symbolic expression of a noun phrase, involving adverbs, adjectives and nouns. This type of protoform could be used to represent the use of adverbs and adjectives to describe or specify nouns. Then, these cases are considered protoform equivalent.

Keywords: generalized constraints, natural language, protoform, conceptual graph.

1 Introduction

World Wide Web has become the largest available repository in the history of Humankind with billions of Megabytes of information. Most of the knowledge stored in the Web is written in natural language. Existing search engines are able to retrieve millions of pages in just few seconds, but most of them are not relevant for the user. In order to stop retrieving irrelevant pages, it is necessary to deal with the semantic aspects of the information contained in those pages.

Furthermore, existing search engines are not able to synthesize a query answer from the pieces of information disperse across the Web, that is, they do not have deduction capabilities. There are several obstacles for search engines to be able to do that, but underlying them is the problem of natural language understanding, the basic problem of precisiation of meaning. To deal effectively with those obstacles, Zadeh [24] suggests new tools as: Precisiated Natural Language (PNL), Generalized Constraint Language (GCL) and Protoform Theory (PFT). Computation with information described in natural language, or NL-Computation, for short is of intrinsic importance because much of human knowledge is described in natural language. NL-Computation involves three modules: (a) Precisiation module; (b) Protoform module; and (c) Computation module. [16].

A prerequisite to mechanization of natural language understanding is precisiation of the meaning of concepts and propositions drawn from a natural language. A proposition in natural language is precisiable if it could be expressed as a generalized constraint [20].

The meaning of a natural language element is precisiated through translation into GCL and is expressed as a generalized constraint. Usually, the natural language element is a proposition, a system of propositions or a concept. And the generalized

constraint may be viewed as a model of meaning [16].

The Protoform module serves as an interface between Precisiation and Computation modules. Basically, its function is that of abstraction and summarization. Informally, a *protoform*, abbreviation of prototypical form, is an abstracted summary, a concept which is centred on the confluence of abstraction and summarization. More specifically, a protoform is a symbolic expression which defines the deep semantic structure of a construct such as a proposition, command, question, scenario, or a system of such constructs [19]. The Computation module serves to deduce an answer to a query [16].

A Precisiated Natural Language (PNL) is a sub language of precisiable propositions in NL, which primary function is serving as a part of NL which admits precisiation. A proposition, *p*, in NL is precisiable if it is translatable into a precisiated language, expressed in a mathematically well-defined language. [23].

In PNL a perception is equated to its description in a natural language. The point of departure in PNL is the assumption that the meaning of a proposition, p, in a natural language, NL, may be represented as a generalized constraint. Some of the main components of PNL are: a dictionary from NL to GCL and a dictionary from GCL to PFL [17]. This work could be seen as a first attempt in this direction: to define those dictionaries.

Conceptual graphs [12, 13], based on semantic networks and Peirce's existential graphs, combine the visual advantage of graphical languages and the expressive power of logic. The main motivation of conceptual graphs has been a smooth mapping between logic and natural language. The formal order-sorted logic foundation of conceptual graphs provides a rigorous basis for reasoning processes performed directly on them.

Morton [8] early recognized the advantages of both conceptual graphs and fuzzy logic and combined them into fuzzy conceptual graphs (FCGs). Cao [4] first introduced and studied FCG programs (FCGPs). Continuing that work, Cao [2] establish a sound and complete foundation for FCG programming.

Usually in dictionaries, the meaning of a word is explained by a descriptive definition, a statement which captures the use, the function and the essence of a term or a concept. Terms, which are included in the meaning definition, constitute a set of keywords associated with the meaning essence. For example, the term "megavitamin_therapy" which meaning in WordNet is "therapy based on a theory that taking very large doses of vitamins will prevent or cure physical or psychological disorders" could be described by the terms contained in the set *B*(*megavitamin_therapy*) = {therapy, theory, very large doses, vitamins, prevent, cure, physical, psychological, disorder}. Therefore, those terms should be used to describe what this kind of therapy is, because they constraint the essence of its meaning. It results evident from this example, that a noun phrase as *very large doses* describes better the therapy meaning than just the words *large* or *dose* used independently [10]. It is easy to realize that they constraint the meaning while describing it. In this paper, the role of adverbs and adjectives as noun constraints is analyzed and some ideas about how to manipulate them as generalized constraints are proposed. The role of adjectives as noun constraints by themselves was analyzed in a previous paper [11]. In this paper, we will concentrate in the use of adverbs as adjective modifiers.

2 Introducing adverbs as generalized constraints

A natural language is basically a system for describing perceptions, which are intrinsically imprecise, reflecting the limited human capacity to manage with detailed information [17]. For example, the proposition *p: Monika is very young* express that the age of Monika has a value that could characterize her as being *very young*. Therefore, in spite of her age is not precisely known, we know that it belongs to a set of age values which characterize very young persons (assuming that Monika is a person, of course). That way, the meaning of *very young* is clear, although the precise age is not known.

A prerequisite to the mechanization of natural language understanding is precisiation of the meaning of concepts and propositions drawn from a natural language. A proposition in natural language is precisiable if it could be expressed as a generalized constraint [19].

The concept of a generalized constraint serves as a bridge between linguistics and mathematics by providing a mean of precisiation of propositions and concepts drawn from a natural language [24]. A typical constraint is an expression of the form $X \in C$, where X is the constrained variable and C is the set of values which X is allowed to take. A generalized

constraint, GC, is defined as an expression of the form GC: X isr R where X is the constrained variable; R is a constraining relation which, in general, is nonbivalent; and r is an indexing variable which identifies the modality of the constraint, that is, its semantics. R will be referred to as a granular value of X. In the above example, a proposition *p*: *Monika is very young* in NL is transformed into a generalized constraint GC(p): *Age(Monika) isr very young* in GCL, where *very young* is a fuzzy relation representing the degree to which a numerical value of *age* fits the description of age as *very young* [17].

The main modalities of generalized constraints are [24]:

- Possibilistic (r = blank), where R defines the possible values of X.
- Probabilistic (r = p), with R playing the role of the probability distribution of X.
- Veristic (r = v), where R plays the role of a verity (truth) distribution of X.
- Usuality (r = u), the usuality constraint presupposes that X is a random variable, and that probability of the event $\{X$ isu $R\}$ is usually, where usually plays the role of a fuzzy probability which is a fuzzy number
- Random-set (r = rs), where X is a fuzzy-set-valued random variable and R is a fuzzy random set.
- Fuzzy-graph (r = fg), where X is a function, f and R is a fuzzy graph which constrains f.
- Bimodal (r = bm), where R is a bimodal distribution.
- Group (r = g), where X is a group variable, $G[A]$, and R is a group constraint on $G[A]$.

GCL serves as a precisiation language in order to express the meaning of propositions, commands as questions expressed in natural languages (NL). But, what is not clear is how to transform a proposition *p* from NL to a *GC(p)* in GCL. Therefore, it is necessary to identify methods and algorithms to do this transformation.

Each of the things in the world can be thought of as having a value on each of a set of dimensions such as size, colour, shape, taste, consistency, and other attributes. Each attribute is really a value on some dimension. To distinguish one thing from another, it is necessary to deal with these dimensions.

Lexicons define nouns to designate categories of things, which share a whole set of attributes that distinguish and characterize them. Common nouns represent a strong tendency for particular sets of values on different dimensions to co-occur. The noun *apple* itself cannot be said to refer to a particular apple; rather it designates a whole category, the category *apple*, which includes many possible individual apples. An apple isn't just an object of a particular shape; it has a characteristic range of sizes, tastes, consistencies, and locations. In other words, the category *apple* is a whole cluster of co-occurring features

In grammatical theory, a *noun phrase* (abbreviated NP) is a phrase whose head is a noun or a pronoun, optionally accompanied by a set of modifiers as adjectives [15]. Noun phrases could be used as the subject or object of a verb. Noun phrases are syntactical units that appear in most sentences expressing a NL constraint applied to nouns and bring information about them.

In a noun phrase as *extremely big apple*, the words *extremely big* define an attribute that characterizes some members of the *apple* category, one possible value on a conceptual dimension, size. The phrase attributes a big size to the apple that is being referred. The word designating the attribute *big* in the example is an *adjective*, while the word *extremely* is an adverb modifying the adjective.

Gasser [5] proposes that an English attributive phrase consisting of an adjective *Adj* designating an attribute *Att* followed by a noun *N* designating a thing category *C* designates the subcategory of *C* whose members have attribute *Att*. Then, if we have the phrase *young Monika*:

Young Monika
↑ ↑
<Adjective> <Noun>
↑ ↑
<Attribute> <Category>

This idea could be expressed using Conceptual Graphs (CG) in linear form [13, 7]:

[*Person*: *Monika*] → (Attr) → [*Young*]. (1)

In linear notation, concepts have [*square brackets*] around them, while relations have (*parentheses*) [9]. *Person* represents a concept type and *Monika* is a referent of this concept type, while Attr (attribute) represents a relation. Therefore, *Monika* should appear in the catalogue of individuals declared as a *Person*.

Fuzzy Conceptual Graphs (FCG) extends simple CGs with linguistic labels defined by fuzzy sets as individual markers [8, 3]. For example, the following FCG expresses that *Monika* has an attribute *Age* which has a measure *Young*. Young will be the linguistic label of a fuzzy set.

[Person: Monika] → (Attr) →[Age: *] → (Meas) → [Measure: Young]. (2)

Meas is a relation used to describe an attribute by a quantity. The quantity could be a measure or a degree [13]. There are some underlying problems as: How do we know that Monika is a person? Therefore, we need some contextual information about Monika. Another solution could be to assume it by default.

3 General aspects about adverbs.

Adverbs are a part of speech. They are used to modify any other language part: verbs, adjectives (including numbers), clauses, sentences and other adverbs, except for nouns [15]. Examples:

Mary sings *beautifully*

David is *extremely* clever

Adverbs typically answer such questions as *how?*, *when?*, *where?*, *in what way?*, or *how often?* This function is called the adverbial function, and is realised not just by single words (i.e., adverbs) but by adverbial phrases and adverbial clauses.

Adverbs and adjectives have important characteristics in common [14]:

- Their gradability,
- They have comparative and superlative forms.

The following words, together with their comparative and superlative forms, can be both adverbs and adjectives (see Table 1):

early, far, fast, hard, late

Table 1: Use of late as adjective and adverb

Adjective	**Adverb**
Robert catches the late train	Robert returns home late.

The comparative *better* and the superlative *best*, as well as some words denoting time intervals (*daily*, *weekly*, *monthly*), can also be adverbs or adjectives, depending on how they are used.

Like adjectives, many adverbs are ***gradable*** [14], which can be modified by *very* or *extremely* (see Table 2):

Table 2: Examples of gradable adverbs

softly	very *softly*
suddenly	very *suddenly*
slowly	extremely *slowly*

The modifying words *very* and *extremely* are themselves adverbs. They are called ***degree adverbs*** [14] because they specify the degree to which an adjective or another adverb applies. Degree adverbs include *almost*, *barely*, *entirely*, *highly*, *quite*, *slightly*, *totally*, and *utterly*.

Like adjectives, too, adverbs are inflected in terms of comparison. The comparative and superlative forms of adverbs are sometimes generated by adding *-er* and *-est*.

John works *hard* -- Mary works *harder* -- I work *hardest*

However, the majority of adverbs form the comparative using *more* and the superlative using *most* (see Table 3).

Table 3: Examples of adverbs in comparative form

Adverb	**Comparative**	**Superlative**
recently	*more recently*	*most recently*
effectively	*more effectively*	*most effectively*
frequently	*more frequently*	*most frequently*

Adverbs also take comparisons with *as ... as*, *less*, and *least*. In the formation of comparatives and superlatives, some adverbs are irregular (see Table 4)

Table 4: Examples of irregular adverbs

Adverb	**Comparative**	**Superlative**
well	*better*	*best*
badly	*worse*	*worst*
little	*less*	*least*
much	*more*	*most*

However, an important distinguishing feature is that adverbs do not modify nouns, either attributively or predicatively, as adjectives do.

Many adverbs are formed by adding *-ly* to an adjective (see Table 5). These adverbs are known as ***-ly adverbs*** [14], but also many adjectives also end in *-ly*, including *costly, deadly, friendly, kindly, likely, lively, manly*, and *timely*. In some cases, the suffix *–wise*, like *clockwise*, may be used to derive adverbs from typical nouns.

Table 5: Examples of –ly adverbs

Adjective	*slow*	*quick*	*soft*	*sudden*
Adverb	*slowly*	*quickly*	*softly*	*suddenly*

Taken as a whole, the adverb class is the most diverse of all the word classes, and its members exhibit a very wide range of forms and functions. Adverbs are considered a part of speech in traditional English grammar, which is derived from Latin grammar.

A logical approach for dividing words into classes relies on recognizing which words can be used in a certain context. For example, some adverbs can be used to modify an entire sentence, whereas others can not.

Following different adverbial functions are described briefly. In the following examples, the adverb is highlighted in bold, while the word it modifies is shown in italics.

Adverbs as an adjective-modifier typically express something about the degree of the adjective, such as `very'. Such adverbs are usually called **degree adverbs** [14] for obvious reasons.

- Monika is **very** *young*
- His poetry is **very** *beautiful*.
- The meaning of this passage is **abundantly** *clear*.
- That sign is **hardly** *visible*.

The value expressed by the adjective (i.e. *young*) is transformed by the adverb (i.e. *very*) in a new level, a new value. Using Fuzzy Conceptual Graphs, this sentence can be formulated:

[Person: Monika] → (Attr) →[Young] → (Meas) → [Degree: #very]. (3)

Then, it could be transformed in:

[Person: Monika] → (Attr) →[Age: *] → (Meas) → [Measure: VeryYoung]. (4)

where VeryYoung is the linguistic label of a fuzzy set

In general, it holds that adverbs that end in *-ly* can be modified by the same adverbs that can modify the adjective that results from removing the *-ly* ending.

- He writes **very** *clearly*.
- The sun came out **quite** *suddenly*.
- This species is the **slightly** *slower* growing one.

[Person: #he] ← (Thme) ← [write] → (Manr) → [Clearly] → (Meas) → [Degree: #very]. (5)

In the following examples the adverb modifies a preposition.

- She is standing **very** near the door.

- She's going **directly** to the store.

Adverbs as a verb-modifier convey information about the manner, time, or place of an event or action. They are known collectively as circumstantial adverbs and constitute the most distinctive adverbial classes [14]. They express one of the circumstances relating to an event or action - *how* it happened (manner), *when* it happened (time), or *where* it happened (place). They are often derived from adjectives by using the suffix -ly.

a) ***Manner*** adverbs tell us *how* an action is or should be performed:

- She sang *loudly* in the bath
- The sky *quickly* grew dark
- They whispered *softly*
- I had to run *fast* to catch the bus

b) ***Time*** adverbs denote not only specific times but also frequency:

- Robert returns *late*.
- I'll be checking out *tomorrow*
- Give it back, *now!*
- John *rarely* rings any more
- I watch television *sometimes*

c) ***Place*** adverbs indicate *where*:

- Put the box *there,* on the table
- I've left my gloves *somewhere*

English is a language which prefers a sequence of *subject–verb–object* in its simplest, unmarked declarative statements. A simple *complete sentence* consists of a *subject* and a *predicate*. The subject is typically a *noun phrase*. The predicate is a finite verb phrase.

A noun phrase is a phrase whose head is a noun or a pronoun, optionally accompanied by a set of modifiers [15]. The modifiers may be:

- **Determiners**: articles (*the*, *a*), demonstratives (*this*, *that*), numerals (*two*, *five*, etc.), possessives (*my*, *their*, etc.), and quantifiers (*some*, *many*, etc.); determiners are usually placed before the noun. Most English dictionaries still identify the determiners as adjectives.
- **Adjectives** (*the red ball*); or
- **Complements**, in the form of an ad positional phrase (such as: *the man with a black hat*), or a relative clause (*the books that I bought yesterday*).

Noun phrases could be expanded by adjectival phrases, which are phrases that have an adjective as their head. In turn, adjectives could be expanded by adverbial phrases, which are phrases that have an adverb as their head.

In English, noun phrases can be treated as single grammatical units. A noun phrase can play the role of a verb argument (such as the subject, the object) or the role of the predicate.

Noun phrases can be easily recognized by using a simple parser as the following, written in Prolog using Definite Clause Grammars (DCGs). DCGs constitute an extension of context free grammars that have proven useful for describing natural and formal languages, and that may be conveniently expressed and executed in Prolog. A more complete version could be found in [1]

```
sentence → nounphrase, verbphrase.
verbphrase → verb, nounphrase.
nounphrase → det(CV), adjnounph(CV).
adjnounph(CV) → noun(CV).
adjnounph(CV) → adjective(CV, Adj),
                noun(CV).
adjnounph(CV) → adjective(CV, Adj),
                adjective(CV, Adj2),
                {Adj \== Adj2}, noun(CV).
```

This program does not accept repeated adjectives such as in "*big big girl*". It is able also to check whether it is needed an '*an*' or an '*a*'. Linking this program with a dictionary as WordNet, it is easy to recognize a huge quantity of simple sentences using WordNet terms (nouns, verbs, adjectives, etc.).

When a noun phrase with an attributive adjective is detected as "*the very young girl*", it could be transformed to a predicative adjective form as "*the girl is very young*" and processed as it. In case of a

noun phrase with several adjectives as "*the extremely fascinating big red book*", it could be processed as if there were several sentences with predicative adjectives: "*the book is extremely fascinating*", "*the book is big*", and "*the book is red*".

Once a noun phrase is detected, the adjective and the noun could be used to determine the characteristic associated to that noun which is being described by the adjective, if there is enough contextual information. For example, the adjective *young* in the sentence *Monika is young* specifies the *age* of *Monika*. But if the sentence was *Monika is tall*, it will specify the *height* of *Monika*. Therefore an auxiliary relation will be needed to establish an association between nouns, adjectives and characteristics; this relation could be expressed by Conceptual Graphs.

[Person: *] → (Attr) → [Age: *] → (Meas) → [Measure: VeryYoung]. (6)

In order to determine the generalized constraint modality *r*, we considered that the usuality modality should be used by default taking into account that the concept of usual value is closer to our intuitive perception of "expected value" than the concept of expected value as it is defined in Probability Theory. [21]. The difference between the concepts of expected and usual values goes to the heart of the difference between precise and imprecise probability theories. The expected value is precisely defined and unique. The usual value is context-dependent and hence is not unique. However, its definition is precise if the natural language predicates which occur in its definition are defined precisely by their membership functions. In this sense, the concept of the usual value has a flexibility that the expected value does not have. Then, a new Prolog fact could be generated as 'Monika'.age is_u very_young. to represent the generalized constraint.

3 Defining protoforms to manage with adverbs.

Informally, a protoform, A, of an object, B, written as A=PF(B), is an abstracted summary of B. Usually, B is a lexical entity such as a proposition, question, command, scenario, decision problem, etc. More generally, B may be a relation, system, geometrical form or an object of arbitrary complexity. Usually, A is a symbolic expression, but, like B, it may be a complex object. The primary function of PF(B) is to place in evidence the deep semantic structure of B. [Zadeh, Aug-2005].

The prototypical form of the object '*Monika*'. *age is very_young* will be *A.B is C*, where *A* is an abstraction of *Monika*, *B* is an abstraction of *age*, and *C* is an abstraction of very_*young*. The same prototypical form could be applied to the objects *Monika is very_intelligent* and *the book is extremely_big*.

Abstraction has levels, just as summarization does. For this reason, an object may have a multiplicity of protoforms. Conversely, many objects may have the same protoform. Such objects are said to be protoform-equivalent. Based on the previous example, we consider that the protoform *A.B is C* could be applied to any noun phrase where the intrinsic semantic structure of that expression corresponds to specifying or describing the value of a subject attribute, including an adverb modifying the adjective. This kind of expression occurs when an adjective is used *attributively* or *predicatively* as in "*the very young Monika"* and "*Monika is very young*". Therefore, "the extremely big book" and "this woman is incredibly beautiful" are protoform equivalent with protoform type *A.B is C*.

4 Conclusions

The generalized-constraint-based computational approach to NL-Computation opens the door to a wide ranging enlargement of the role of natural languages in scientific theories [16]. In this paper, adverbs and adjectives in NL propositions are considered generalized constraints as soon as they describe or specify characteristics associated to the noun they describe.

Fuzzy Conceptual Graphs (FCGs) are used to express propositions in natural language, which use adjectives and adverbs. The visual advantage of these graphs and their logic expressiveness shown extremely helpful for defining dictionaries between NL and GCL.

In English, adjectives can be used either *attributively* or *predicatively*. In both cases, a generalized constraint could be used to express how the value of the attribute associated to the noun is conditioned in order to be acceptable. By using simple sentences as the ones presented here a lot of information could be retrieved from documents on the Web.

Furthermore a prototypical form is proposed to represent the deep semantic structure of a noun

phrase, involving adverbs, adjectives and nouns. That way, all of these cases share the same protoform and, therefore, are protoform equivalent.

Acknowledgement

This work was partially supported by SCAIWEB PAC06-0059 project, JCCM, Spain.

References

[1] [Aaby] A. Aaby, A Natural Language Processor, The Pennsylvania State University, 1999. http://moonbase.wwc.edu/~aabyan/LogicPgmg/CODE/DCG/doc/doc.html

[2] [TCao PhD] T. H. Cao, Foundations of Order-Sorted Fuzzy Set Logic Programming in Predicate Logic and Conceptual Graphs, PhD Thesis, University of Queensland, 1999

[3] [TCao 2001] T. H. Cao, Fuzzy Conceptual Graphs for the Semantic Web, BISC International Workshop on Fuzzy Logic and the Internet FLINT 2001.

[4] [TCao Master] T.H. Cao, Fuzzy Conceptual Graph Programs. Master's Thesis, Asian Institute of Technology, 1995.

[5] [Gasser] M. Gasser, How Language Works, Edition 3.0, Indiana University, 2006, http://www.indiana.edu/~hlw/book.html

[6] [Houghton Mifflin] E.D. Hirsch, Jr., J.F. Kett, and J. Trefil, Editors, The New Dictionary of Cultural Literacy, Third Edition, Houghton Mifflin Company, 2002

[7] [ISO] ISO, Conceptual Graphs, ISO/JTC1/SC 32/WG2 N 000, 2001-04-02

[8] [FCG] S. K. Morton, Conceptual graphs and fuzziness in artificial intelligence, PhD Thesis, University of Bristol, 1978

[9] [CG course] U. Petersen, H. Schärfe, and P. Øhrstrøm, Online Course in Knowledge Representation using Conceptual Graphs, http://www.huminf.aau.dk/cg, 2005

[10] [EUROCAST-07] A. Soto, J. A. Olivas, and M. E. Prieto, Fuzzy Measure of Meaning Presence in Documents, Eleventh International Conference on Computer Aided Systems Theory EUROCAST-07, Canary Islands, Spain, February, 2007 (submitted paper)

[11] [FUZZ-IEEE-07] A. Soto, J. A. Olivas, and M. E. Prieto, Using Generalized Constraints and Protoforms to deal with adjectives, FUZZ-IEEE 2007 IEEE International Conference on Fuzzy Systems, Imperial College, London, UK, 23-26 July, 2007, (submitted paper)

[12] [Sowa 76] J.F. Sowa, Conceptual graphs for a data base interface, IBM Journal of Research and Development 20:4, pp. 336-357, 1976.

[13] [Sowa 84] J.F. Sowa, Conceptual Structures - Information Processing in Mind and Machine. Addison-Wesley Publishing Company, Massachusetts. 1984.

[14] [UCL] UCL, Survey of English Usage, Department of English Language and Literature, University College London, London http://www.ucl.ac.uk/internet-grammar/adjectiv/inherent.htm

[15] [wikipedia] Wikipedia, the free enciclopedia, http://en.wikipedia.org/wiki/Main_Page

[16] [New Frontier] L. A. Zadeh, "A New Frontier in Computation – Computation with Information Described in Natural Language", CIS (Center for Intelligent Systems) Seminar, UC Berkeley, November 9, 2006.

[17] [Zadeh, 2004] L. A. Zadeh, "A note on web intelligence, world knowledge and fuzzy logic", Elsevier, Data & Knowledge Engineering 50, 2004, pp. 291–304

[18] [Zadeh Aug-05] L. A. Zadeh, "From Search Engines to Question- Answering Systems— The Problems of World Knowledge, Relevance, Deduction and Precisiation". Keynote Speech, IRI 2005 IEEE International Conference on Information Reuse and Integration, Las Vegas, Nevada, USA, August 15, 2005.

[19] [Zadeh Jul-05]. L. A. Zadeh, "Toward A Computational Theory Of Precisiation Of Meaning Based On Fuzzy Logic—The Concept Of Cointensive Precisiation". 11th World Congress of International Fuzzy Systems Association IFSA 2005, Beijing, China, July-2005.

[20] [Zadeh, Jan-2005]. L. A. Zadeh, "Toward a Generalized Theory of Uncertainty (GTU)—An Outline", Information Sciences 172, 2005.

[21] [Zadeh, 2002] L. A. Zadeh, "Toward a perception-based theory of probabilistic reasoning with imprecise probabilities", Elsevier Journal of Statistical Planning and Inference 105, 2002, pp. 233–264

[22] [Zadeh, Aug-2000] L. A. Zadeh, "Toward the Concept of Generalized Definability", Abstract of a lecture presented at the Rolf Nevanlinna Colloquium, University of Helsinki, Helsinki, Finland, Aug. 8, 2000

[23] [Zadeh & Nikravesh]. L. A. Zadeh, and M. Nikravesh, "Perception-Based Intelligent Decision Systems", ONR Summer 2002 Program Review, University of California Los Angeles UCLA, USA. July 30-August 1 2002.

[24] [Zadeh 2006 con Elie] L. A. Zadeh, From Search Engines to Question Answering Systems: The Problems of World Knowledge, Relevance, Deduction and Precisiation. In Fuzzy Logic and the Semantic Web, edited by Elie Sanchez, Elsevier 2006.

Session 4

Applications 2

A Less Cumulative Algorithm of Mining Linguistic Browsing Patterns in the World Wide Web

Krzysztof Dyczkowski
Adam Mickiewicz University
Faculty of Mathematics and Computer Science
Umultowska 87
Poznan, Poland
e-mail chris@amu.edu.pl

Abstract

Finding sequential patterns is one of important issues in data mining. This paper deals with linguistic (fuzzy) sequential patterns. The existing algorithms for discovering such patterns do involve usual sigma counts of fuzzy sets as measure of support. Unfortunately, a well-known side effect is then an undesirable cumulation of small membership values. We like to propose an improved approach based on generalized sigma counts, which in particular leads to a consequently lower cumulation, and considerably to better results. We test the modified algorithm using different generalized sigma counts based on different cardinality pattern functions.

Keywords: Data mining, Linguistic sequential patterns, Scalar Cardinality, Generalized sigma counts, Pattern function

1 Introduction

Internet servers especially WWW servers, generate a huge number of data. In particular, they log large amounts of information about connections, security events and users transfers. Finding regularities in such (usually) large data files leads to security improvement, brings knowledge about users behavior and, therefore, is of interest to administrators, marketing departments, etc.

There are two main approaches to discovering those singularities: association rules and sequential patterns (see [1, 2, 3, 4, 8, 9]). We like to focus on sequential patterns which are expression of the form $X_1 \rightarrow X_2 \rightarrow \ldots \rightarrow X_n$, where X_i is a set of some items. In the case of linguistic sequential patterns, the X_i's are equipped with some linguistic labels. In particular, a linguistic browsing pattern in the WWW is a sequence of web pages to which linguistic terms expressing the duration of browsing time are assigned. For instance, the linguistic browsing pattern $file1.html/short \rightarrow file2.html/long \rightarrow file3.html/medium$ says that a frequent behavior of users is that they watch the page file1.html for a short time, then they watch file2.html for a long time and, further, they watch file3.html for a medium time. It was Hong et al. [7] (cf.[8]) who proposed an algorithm for finding such patterns which is an adaptation of the ideas by Agrawal and Shirkant [1, 2, 3]. The following are three main steps of that algorithm:

1. Create fuzzy sets from transactions data using given membership functions.
2. Generate large itemsets by calculating the fuzzy cardinality of each candidate itemset.
3. Generate iterative sequential patterns from the large itemsets found in step 2.

By the way Hong et al. proposed in [9] a GA algorithm which dynamically constructs membership functions to model linguistic terms of browsing time.

The Hong's algorithm from [7] uses the usual sigma count as a measure of support. Unfortunately, a well-known side effect is then an undesirable cumulation of small membership values which can lead to wrong conclusions. For example, a fuzzy set A with 100 elements with membership values equal to 0.01 has the sigma count equal to 1, which is not coherent with intuition (see [5]).

In this paper we like to propose an improved approach based on generalized sigma counts, which in particular leads to considerably lower cumulation and, consequently, to better results.

2 Generalized Sigma Counts

Our further discussion requires some notions related to cardinalities of fuzzy sets. FFS will denote the family

of all finite sets in a universe $\mathcal{M}$. Let us recall the axiomatic definition of scalar cardinality according to Wygralak [13, 14]

A function $\sigma : FFS \to [0,\infty)$ will be called a scalar cardinality if the following postulates are satisfied for each $a, b \in [0,1]$, $A, B \in FFS$ and $x, y \in \mathcal{M}$:

(SC1) $\sigma(1/x) = 1$,
(SC2) $a \leq b \Rightarrow \sigma(a/x) \leq \sigma(b/y)$,
(SC3) $A \cap B = 1_\emptyset \Rightarrow \sigma(A \cup B) = \sigma(A) + \sigma(B)$.

If σ does fulfill the above axioms, the number $\sigma(A)$ is said to be a scalar cardinality of A.

The following characterization correspond to scalar cardinalities definied via (SC1)-(SC3).

Theorem 1 *A mapping $\sigma : FFS \to [0,\infty)$ is a scalar cardinality iff*

$$\forall A \in FFS:\ \sigma(A) = \sum_{x \in supp(A)} f(A(x)), \qquad (1)$$

where $f : [0,1] \to [0,1]$ is nondecreasing and $f(0) = 0$, $f(1) = 1$.

Each function f from Theorem 1 is called a cardinality pattern or pattern function. Since $f = id$ gives as $\sigma(A)$ the usual sigma count of A from [16], each $\sigma(A)$ from (1) will be called a generalized sigma count of A. To emphasize which pattern function f is used to generate a generalized sigma count, we will write $SC_f(A)$ instead of $\sigma(A)$.

In the following, we present some basic examples of pattern function together with the resulting generalized sigma counts. They are instances of pattern functions leading to a reduction of accumulation of small membership degrees.

Example 2.1 *Let $t \in (0,1]$*

$$f_1(a) = \begin{cases} 1, & \text{if } a \geq t, \\ 0, & \text{otherwise.} \end{cases} \qquad (2)$$

for $a \in [0,1]$. One can also see that choice of even door-sill 1 is justified.

It is easy to see that $SC_{f_1}(A) = |A_t|$, where A_t denotes the t-cut set of A. f_1 reduces the cumulation effect by reducing to zero each membership value which is less than a fixed threshold t, and by increasing to 1 each membership value $\geq t$ (see Fig. 1).

Example 2.2 *Let*

$$f_2(a) = a^p \qquad (3)$$

for $a \in [0,1]$, $p > 0$. (see Fig. 2)

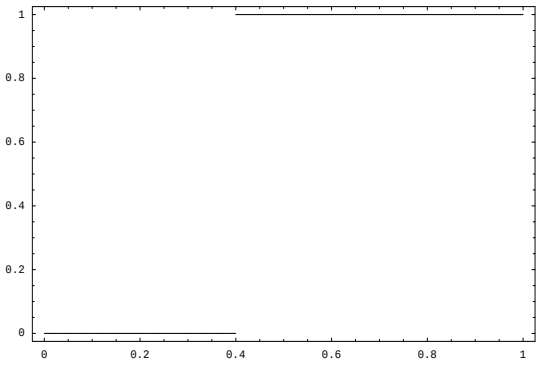


Figure 1: Pattern function f_1 with $t = 0.4$

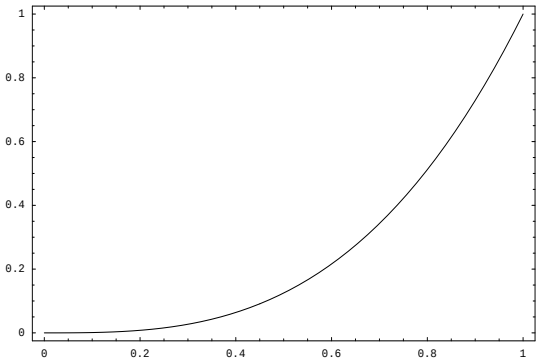


Figure 2: Pattern function f_2 with $p = 4$

Obviously, $p = 1$ gives the usual sigma count.

Example 2.3 *Let $b, c \in (0,1)$*

$$f_3(a) = \begin{cases} 0, & \text{if } a \leq b, \\ a & \text{if } a \in (b,c), \\ 1, & \text{otherwise.} \end{cases} \qquad (4)$$

for $a \in [0,1]$.

In this case, we are using two threshold levels b and c. The threshold b determines which low degrees of the membership are supposed to be 0 and c determines which big values of the membership are supposed to be replaced by 1. The values of the membership between b and c stays untouched. (see Fig. 3).

Example 2.4 *Let $b, c, d \in (0,1]$ and $b < c < d$*

$$f_4(a) = \begin{cases} 0, & \text{if } a \leq b, \\ 2(\frac{a-b}{c})^2, & \text{if } a \in (b,c], \\ 1 - 2(\frac{a-d}{c})^2, & \text{if } a \in (c,d), \\ 1, & \text{otherwise.} \end{cases} \qquad (5)$$

Function f_3 (4) is discontinuous contrast enhancement function. f_4 (5) is more sophisticated continuous variant of such function (see Fig 4). This kind of pattern

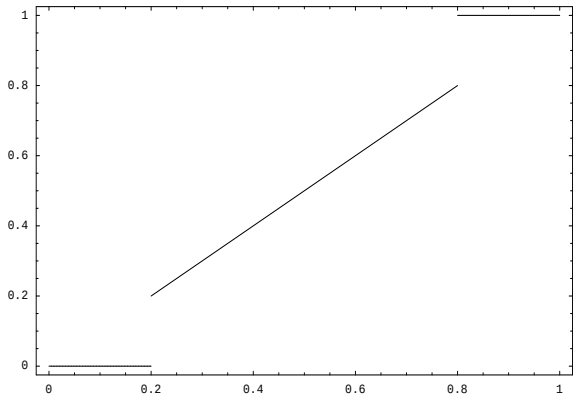


Figure 3: Pattern function f_3 for $b = 0.2$, $c = 0.8$

function promote higher values of membership function. We will take an advantage of this pattern function in examples.

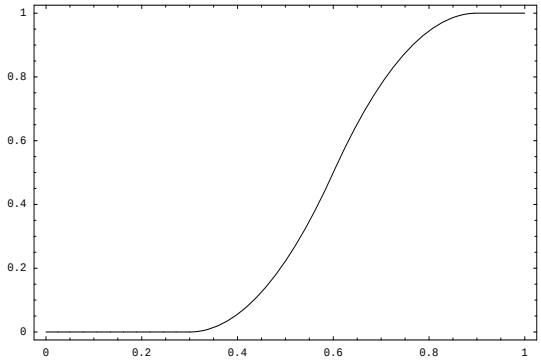


Figure 4: Pattern function f_4 with $b = 0.3$, $c = 0.6$, $d = 0.9$

Example 2.5 *Let* $b \in (0, 1]$

$$f_5(a) = \begin{cases} 0, & \text{if } a < b, \\ a, & \text{otherwise.} \end{cases} \tag{6}$$

for $a \in [0, 1]$.

In this case pattern function (6) works as identity function (as a usual sigma count) except values less than threshold point b which are treated as 0 (see Fig. 5). So b decides how small values will not be cumulated.

Example 2.6 *Let* $p \in (0, 1]$

$$f_6(a) = 1 - (1 - a)^p \tag{7}$$

for $a \in [0, 1]$.

Worth mentioning is that each normed generator of a nonstrict Archimedean t-conorm is a pattern function

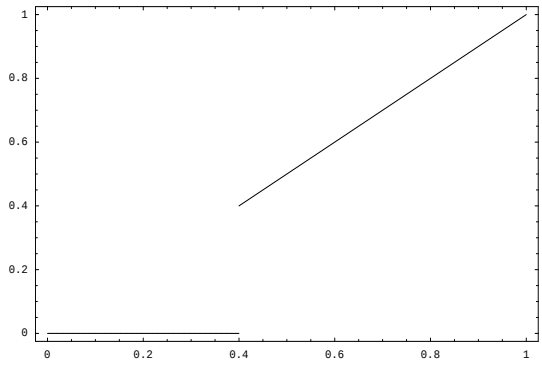


Figure 5: Pattern function f_5 with $b = 0.4$

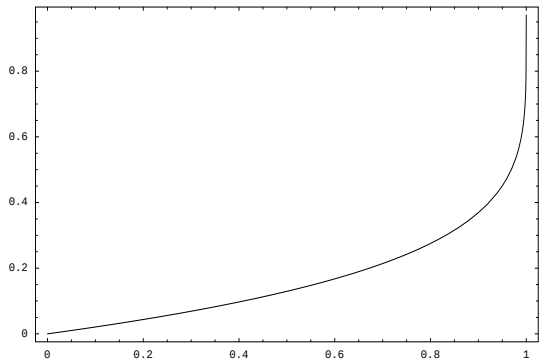


Figure 6: Pattern function f_6 with $a = 0.2$

(see [13]). As example, take the normed generator (7) of the Schweizer t-conorm $s_{S,\lambda}(a, b) := 1 - (0 \vee ((1 - a)^\lambda + (1 - b)^\lambda - 1))^{\frac{1}{\lambda}}$ whit $\lambda > 0$ (see Fig. 6).

Notice that function f_2 (3) is just normed generator of Yager t-conorm.

To show how values of cardinality depends on choice of cardinality function we show next example.

Example 2.7 *Let us consider a fuzzy set A to which 50 elements belong to degree 0.1, 30 elements to degree 0.2, 10 elements to degree 0.3, 6 elements to degree 0.4, 2 elements to degree 0.8 and 2 elements to degree 0. The sigma count of this fuzzy set equals* $SC_{id}(A)$*=19.8, which seems to be a inflated result in comparison with our intuition. Computing generalized sigma counts, we get more adequate results, for instance*

$SC_{f_1}(A) = 10$,
$SC_{f_2}(A) = 3.426$,
$SC_{f_3}(A) = 9.4$,
$SC_{f_4}(A) = 4.222$,
$SC_{f_5}(A) = 5.8$,
$SC_{f_6}(A) = 4.912$,

where the cumulation effect is reduced. As we see, the choice of an appropriate pattern function is a key issue when using generalized sigma counts.

3 An Improved algorithm of Mining Sequential Patterns

For the sake of simplicity, the algorithm will be presented in non-formalized form. As input data we take log file from WWW server. It is a set U of n transactions, where n can be large. Each transaction contains browsed information about opened files from WWW server. For the mining process the following files of transactions will be of interest: datetime, source IP and name of opened file. At the beginning we need to

Table 1: Example of transactions in the log

Client IP	Time	Web Page
194.114.147.2	22:03:57	index.html
194.114.147.2	22:04:11	offer.html
194.114.147.2	22:04:25	rocker.html
194.114.147.2	22:04:33	b8250.html
194.149.88.10	12:51:50	index.html
194.149.88.10	12:52:06	offer.html
194.149.88.10	12:52:28	rocker.html
194.149.88.10	12:52:28	bench.html
195.46.43.144	11:20:11	index.html
195.46.43.144	11:20:15	offer.html
195.46.43.144	11:20:20	rocker.html
212.122.214.145	20:07:18	index.html
212.122.214.145	20:07:23	offer.html
212.122.214.145	20:07:29	bench.html
212.122.214.145	20:07:40	rocker.html
212.122.214.145	20:07:48	chair.html
212.122.214.145	20:07:54	a0910.html

preprocess log data in three steps:

1. Select from log only transactions containing files that are web pages for example such as ".html", ".php", ".asp" etc. Example of part of such transformed log is shown in Table 1.
2. Replace each IP by unique integer (called encoded client ID). For convenience we will represent all files as capital letters eg. file "index.htm" as "A", "offer.html" as "B" and so on.
3. Sort by client ID and datetime fields.
4. For each client calculate the time durations of browsing each web page as time to open the next file. Last files for particular client ID will be omitted, because of the lack of information about the moment of disconnecting.

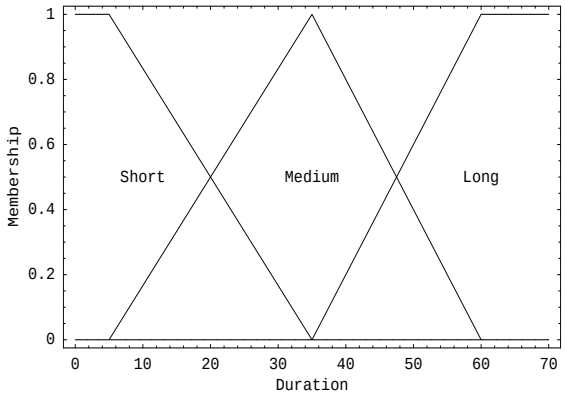


Figure 7: Example of linguistic terms of browsing time: Short, Medium, Long

Table 2: Example of transformed log into linguistic sequences

ID	Page	Short	Medium	Long
1	A	0,080	0,920	0,000
1	B	0,160	0,840	0,000
1	C	0,000	0,833	0,167
2	A	0,800	0,200	0,000
2	B	0,760	0,240	0,000
2	D	0,000	0,700	0,300
3	A	0,960	0,040	0,000
3	B	1,000	0,000	0,000
3	E	0,000	0,900	0,100
3	D	0,800	0,200	0,000
4	A	1,000	0,000	0,000
4	B	0,960	0,040	0,000
4	C	0,640	0,360	0,000
4	E	1,000	0,000	0,000
4	D	0,000	0,300	0,700
4	E	0,840	0,160	0,000
4	C	1,000	0,000	0,000
5	A	0,760	0,240	0,000
5	B	0,480	0,520	0,000
5	F	0,600	0,400	0,000
6	A	0,760	0,240	0,000
6	B	0,080	0,920	0,000
6	F	0,920	0,080	0,000
6	D	1,000	0,000	0,000
7	A	0,480	0,520	0,000
7	B	0,320	0,680	0,000
7	F	0,000	0,833	0,167
7	C	0,000	0,867	0,133
7	E	0,520	0,480	0,000

5. Construct browsing sequence by transforming each transaction into a fuzzy set $(\frac{x_1}{short} + \frac{x_2}{medium} + \frac{x_3}{long})$ using the given membership functions. It can be functions proposed in Figure 7.

Such linguistic browsing patterns will be a input for the mining algorithm described below:

1. Calculate membership value of each region for each browsing sequence using a fuzzy sum operator. A region is a web page with a linguistic value, for example: A.Short, B.Long, etc. For simplicity, we apply the max operation which, however, could be replaced by any t-conorm.

2. Calculate scalar cardinality of each regions in all browsing sequences. To avoid problems with cumulating small values we use a generalized sigma count with a given pattern function.

   For each file, select region with the value of cardinality greater than a fixed threshold α. We will use $\alpha = 1$. In [7], this step relies on choosing regions with the sigma count greater than or equal to α, where $\alpha = 2$.

   We collect all selected regions in the set L_1 which is called the large itemset.

3. If L_1 is empty, we stop the algorithm. Otherwise, we put $k := 2$, which denotes the length of current sequence, and we do following.

4. Generate the set of candidates C_k consisting of pairs of all possible combinations of elements from L_{k-1}.

5. For each candidate from set C_k we calculate membership value for each user sequence using a t-norm as intersection operator. Let us use the minimum operation for simplicity.

6. Calculate a generalized scalar cardinality of each k-sequence in C_k in the way shown in Step 2 with given cardinality pattern. We choose only these sequences which have counts grater than given threshold point α. The resulting set is denoted by L_k.

7. If L_k is not empty we construct set C_{k+1} from L_k by combining all possible k-sequences with sequences from L_2, and we go to Step 5 with k:=k+1.

8. Finally, take L_k as the set of longest sequences found by the algorithm.

4 Examples of Mining with Various Generalized Sigma Counts

To show the influence of the choice of a pattern function to the mining process, let us present some examples. We will use symbols "A.S", "A.M", "A.L" to denote linguistic items, where "A" represents file name and "S" stands for "Short", "M" for "Medium", "L" for "Long". All computations refer to the small data set from Table 2. We put threshold point $\alpha = 1$ and use the t-norm minimum together with the t-conorm maximum.

Example 4.1 *The values given in brackets are values of generalized cardinality.*

1. *Sets L_1 and L_2 generated by means of the standard sigma count SC_{id}*

Set L_1
A.S (4.84), B.S (3.76), C.M (2.06), D.S (1.80), E.M (1.54), F.S (1.52)
Set L_2
A.S→B.S (3.640)

2. *Sets L_1, L_2 and L_3 generated using a SC_{f_1} with $b = 0.4$*

Set L_1
A.S (6.00), B.M (5.00), C.M (2.00), D.S (2.00) E.S (2.00), F.S (2.00)
Set L_2
A.S→B.M (3.00), B.M→F.S (2.00)
Set L_3
A.S→B.M→F.S (2.00)

3. *Sets L_1 and L_2 generated by means of SC_{f_2} with $p = 4$*

Set L_1
A.S (2.98), B.S (2.25), C.M (1.06), D.S (1.41) E.S (1.07)
Set L_2
A.S→B.S (2.09)

4. *Sets L_1, L_2 and L_3 generated by means of SC_{f_3} with parameters as in Example 2.3*

Set L_1
A.S (5.00), B.M (3.87), C.M (2.36), D.S (2.00) E.S (1.52), F.S (1.60)
Set L_2
A.S→B.M (2.00), B.M→F.S (1.52)
Set L_3
A.S→B.M→F.S (1.28)

5. *Sets L_1, L_2 and L_3 generated by means of SC_{f_4} with parameters as in Example 2.4*

Set L_1
A.S (4.91), B.M (3.08), C.M (1.99), D.S (1.94) E.S (1.27), F.S (1.50)
Set L_2
A.S→B.M (1.34), B.M→F.S (1.27)
Set L_3
A.S→B.M→F.S (1.16)

6. *Set L_1, L_2 and L_3 generated by means of SC_{f_5} with $b = 0.4$*

Set L_1
A.S (4.76), B.M (3.39), C.M (1.70), D.S (1.80) E.S (1.52), F.S (1.52)
Set L_2
A.S→B.M (1.76), B.M→F.S (1.44)
Set L_3
A.S→B.M→F.S (1.28)

7. *Set L_1 and L_2 generated by means of SC_{f_6} with $b = 0.2$*

Set L_1
A.S (2.38), B.S (1.97), D.S (1.27) E.S (1.14)
Set L_2
A.S→B.S (1.43)

As we can see, the choice of the cardinality function plays a significant role. In the first example with the usual sigma count, cardinality values are relatively high but the length of maximal sequence is 2. In the next examples, values of cardinalities are smaller but sequences found are longer.

5 A Real Data Mining Example

Our mining algorithm, involving generalized sigma counts, and its implementation, have also been tested on real log data. Let us show some results for a log set of about 5.000 transactions concerning 900 clients who browsed about 200 web pages. We present only final sets of longest browsing patterns found for each cardinality pattern. As in previous example, we use the minimum and maximum operators, and $\alpha = 1$ as the threshold level.

Example 5.1 *As previously, values in brackets are cardinalities.*

1. *The use of SC_{id} (the usual sigma count)*

Set L_{10}
S.S→S2.S→T.S→T2.S→T3.S→T4.S→ B.S→B2.S→B3.S→B4.S (1.12)

2. *SC_{f_1} with $t = 0.4$*

Set L_5
A.S→S.S→T.S→B.S→W.S (4.00)

3. *SC_{f_2} with $p = 4$*

Set L_5
S.S→E1.S→E2.S→T.S→B.S (1.06)
W.S→S.S→A.S→C.S→B.S (1.04)

4. *SC_{f_3} with $b = 0.2$, $c = 0.8$*

Set L_4
F3.S→C3.S→C2.S→C.S (1.12)
G.L→H.L→O.S→CS (1.67)
C2.S→C3.S→L.S→C4.S→ (1.64)
B4.S→B5.S→F.S→V.S (1.20)
S.S→A.S→C.S→B.S (2.00)
S.S→T.S→B.S→W.S (8.00)

5. *SC_{f_4} with $b = 0.3$, $c = 0.6$, $d = 0.9$*

Set L_7
A.S→S.S→S2.S→T.S→T2.S→T3.S→T4.S (1.44)
S2.S→T.S→T2.S→T3.S→T4.S→B.S→B2.S (1.03)

6. *SC_{f_5} with $b = 0.4$*

Set L_5
O.S→B.S→T.S→S.S→A.S (1.44)

7. *SC_{f_6} with $b = 0.2$*

Set L_5
B.S→B2.S→B3.S→B4.S→B5.S (8.89)
T.S→T2.S→T3.S→T4.S→M.M (1.50)
S2.S→T.S→T2.S→T3.S→T4.S (1.05)

One sees that the usual sigma count leads to one long browsing pattern whose support is however weak, which is a consequence of the cumulation effect.

6 Conclusions

We presented a modified algorithm for mining linguistic browsing patterns in the WWW. Using generalized sigma counts we can reduce the effect of cumulating small values of membership function.

In our future work, we like to investigate methodological issues related to the choice of pattern functions generating generalized sigma counts. Also, we like to examine other t-norms and t-conorms for numerical interpretation of logical connectives in the mining procedure.

References

[1] Agrawal R, Imielinksi T, Swami A., Mining association rules between sets of items in large database, The 1993 ACM SIGMOD Conference, 207-216, 1993

[2] Agrawal R, Srikant R., Fast algorithm for mining association rules, The International Conference on Very Large Data Bases, 487-499, 1994

[3] Agrawal R, Srikant R., Mining sequential patterns, The Eleventh International Conference on Data Engineering, 314, 1995

[4] Cooley R, Mobasher B., Srivastava J., Web mining: information and pattern discovery on the world wide web, Ninth IEEE International Conference on Tools with Artificial Intelligence, 558-567, 1997

[5] Dubios D., Prade H., Sudkamp T., On the Representation, Measurment, and Discovery of Fuzzy Associations, IEEE Transactions on Fuzzy Systems, vol 13, 2:250–262, 2005

[6] Fiot C., Laurent A., Teisseire M., Web Access Log Mining With Soft Sequential Patterns, 7th International FLINS Conference on Applied Artificial Intelligence (FLINS'06), 2006.

[7] Hong T.P., Lin K.Y., Wang S.L., Mining linguistic browsing patterns in the world wide web, Soft Computing 6:329-336, Springer-Verlag, 2002

[8] Hong T.P., Lin K.Y., WangS.L., Mining fuzzy sequential patterns from quantitative transactions, Soft Computing 10:925-932, 2006

[9] Hong T.P., Chen C.H., Wu Y.L., Lee Y.C., A GA-based fuzzy mining approach to achieve a trade-off between number of rules and suitability of membership functions, Soft Computing, 10:1091-1101, 2006

[10] Larose D. T., Discovering Knowledge in Data: An Introduction to Data Mining, Wiley, 2005.

[11] Klement E.P., Meisar R., Pap E., Triangular Norms, Kluwer Academic Publishers, 2000.

[12] Wygralak M., Triangular operations, negations, and scalar cardinality of fuzzy set, in: L.A. Zadeh, J. Kacprzyk (eds.), Computing with Words in Information/Intelligent Systems Vol. 1. Foundations, Phisica-Verlag, Heidelberg New York, 326–341, 1999.

[13] Wygralak M., An axiomatic approach to scalar cardinalities of fuzzy sets, Fuzzy Sets and Systems, 110, 175–179, 2000.

[14] Wygralak M., Cardinalities of Fuzzy Sets Studies in Fuzziness and Soft Computing, vol 118, Springer Verlag, 2003.

[15] Zadeh. L. A., Fuzzy Sets, Inform. and Control, 8:338–353, 1965.

[16] Zadeh L. A., From computing with numbers to computing with words - form manipulation of measurements to manipulation of perceptions, IEEE Trans. on Circuits and Systems -I: Fundamental Theory and Appl., 45:105–119, 1999.

Skew-Tree Based Multistage Fuzzy Controller for Nonlinear Systems

Abbas Khosravi
Dept. of Elec., Informatics & Automatics
Universiy of Girona,
17071, Girona, Spain
khosravi@eia.udg.es

Jie Lu
Faculty of Information Technology
University of Technology Sydney
Broadway, NSW 2007
Australia
jielu@it.uts.edu.au

Xianyi Zheng
The ENSAIT Textile Institute,
9 rue de l'Ermitage
59100 Roubaix,
France
xianyi.zeng@ensait.fr

R. Barzamini
Dept. of Elec. Eng.,
Amirkabir University of Technology
15914, Tehran, Iran
barzamini@aut.ac.ir

Abstract

In this paper a multistage fuzzy controller is designed and implemented on the ball and beam system which is a well-known benchmark in the control area because of its nonlinear, unstable non-minimum phase behavior. The most important property of the proposed controller is its simplicity whiles it fully mimics human's actions for controlling this system. In comparison to classical controllers such as Linear Quadratic Regulator (LQR) or combinational controllers such as fuzzy proportional integrator (PI), the proposed controller makes the overall system more stable with a remarkable decrease in settling time. Furthermore it eliminates overshoots around the desired set point much better than mentioned controllers do. This leads to a very smooth movement of ball on the beam illustrated in some simulations.

Keywords: Ball and beam, Multistage fuzzy controller.

1 Introduction

In the last decade fuzzy algorithms have been used extremely in different areas such as image processing, control, communication, power plants and transport systems [1]. Fuzzy algorithms were first employed by Mamdani & King to control simple systems [2], [3]. After emerging these applications of fuzzy systems, many researchers devoted enormous effort and time to develop and extend fuzzy concepts to different areas such as manufacturing, robotics, and home utilities. A very good review of fuzzy applications has been presented in [1].

Classical controllers need a very precise linear or nonlinear model of the plant to do controlling task well. Almost in all practical cases, there is no accurate model of the plant or the existing model is full of uncertainties. In fuzzy algorithms, the same method that an expert operator uses to model and control a system is implemented without knowing any mathematical model of the plant. Besides use of human's experience in control process, fuzzy algorithms eliminate human weakness in making quick decisions and doing massive mathematical computations.

Rule base of commonplace fuzzy controllers are often developed directly using error (e) and error velocity ($\dot{e}$) [4]. Like other classical controllers, error acceleration ($\ddot{e}$) is not a suitable input for fuzzy controllers due to its lack of immunity to noise. Beside its calculation difficulties, it also increases the rule base dimension extremely which is source of some annoying problems such as cumbersome and time-consuming computations [5], [6]. In many cases for improving overall performance and elimination of steady state error, fuzzy controllers are combined with classical controllers such as PID [7]. Also some researchers have tried to link adaptive, supervisory, and optimal techniques with fuzzy systems to improve the results more [4]. In these methods, mass of mathematical computation and design complexity prevail use of human experiences which is in contrast with the core motive for use of fuzzy systems. Because of this complexity in their structure and having no transparency, subsequently, an ordinary operator easily can not understand and execute them [8]. Neural networks are an alternative that are not much popular in comparison to other intelligent methods. This is due to lack of transparency in their design and learning method used by them. Furthermore, they don't use human experiences to do controlling tasks.

This paper is an extension to the experimental project done in [5], [6]. In [5], authors have used a combination of fuzzy and classical PID controller to handle ball on the beam through turning the beam using an electrical DC motor. This motor provides required torque ordered by the controller for accomplishing that goal. To avoid dimension increasing, they have used a fuzzy PI controller with a resetting factor. The output of this controller is passed as input to a PID controller that eliminates steady state error. This design is almost complex and far from the concept of fuzzy controllers. In the other word, designers have not indeed employed human's experiences for controlling the ball on the beam.

In this paper a fuzzy multistage controller is designed for the famous ball and beam system considered in [5]. The underlying system possesses a highly nonlinear, unstable, non-minimum phase model which puts it as a challenging control problem. The proposed controller in this paper mimics human is method for controlling of this highly nonlinear system. We have avoided use of classical methods to improve the response of the closed-loop system. There is no need for having analytical model of the system which is a pleasing aspect. Furthermore, use of this design noticeably decreases number of rule bases leading to computational mass reduction.

The rest of this paper is organized as follows: Section 2 represents dynamic equations of the ball and beam system. Section 3 includes some information about two kinds of multistage controller and their structures. Proposed controller for the ball and beam system is expressed in section 4. The simulation results in section 5 highlight the capabilities of the proposed controller. In the end, section 6 concludes the paper.

2 Ball and Beam System

Nonlinearity and instability as well as being non-minimum phase make the control of ball and beam system a thought-provoking problem. In many cases, it has been considered as a challenging benchmark for verifying goodness of control algorithms. This section describes dynamics equation of this system used in simulation.

The ball and beam system includes three main parts: a ball with a fixed mass, a fixed length beam, and an electrical DC motor as depicted in Figure 1.

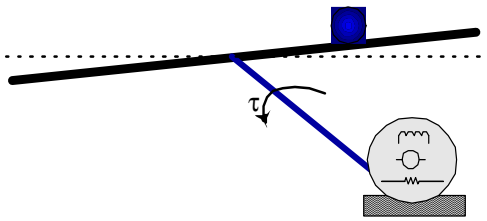


Figure 1: Ball and beam system schematic.

The electrical motor delivers an appropriate torque that turns beam around its center. By an adequate and quick turn of beam, the ball position is regulated on beam.

We used the well-known lagrangian equation in classical dynamics to obtain ball and beam dynamical equations. Lagrangian term of a system is defined as difference between kinematics energy and potential energy:

$$L = K - P \tag{1}$$

Lagrangian equations are as follow:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = Q_i \tag{2}$$

Q_i is an external input in q_i direction. For using (1) and (2), we need to calculate kinematics and potential energy for all parts of the system. While the overall kinematics energy of the system includes kinematics energies of both beam and ball, potential energy solely depends on the ball position on the beam. Using (2) and after some mathematical manipulations and simplification, dynamic equations for the ball and beam system are obtained as follow:

$$\begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \\ \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} \dot{\theta} \\ \dfrac{(\frac{k_i k_b}{R_a} - 2mx\dot{x} - B_m)\dot{\theta} - mgx\cos\theta + \frac{k_i}{R_a}u_a}{mx^2 + I + j_m} \\ \dot{x} \\ \dfrac{x\dot{\theta}^2 - g\sin\theta}{I + \frac{2}{5}C_r^2} \end{bmatrix} \tag{3}$$

In (3), m, M, L, R, x, θ and τ denote ball mass, radius of ball, beam mass, length of beam, ball position distance to the beam center, beam angle and torque input respectively. C_r is rotational coefficient of ball which is one for the pure rotational movement of ball on the beam. Ki, Kb, Jm and Ra are electrical DC motor parameters. All parameters belong to an experimental plant and their values have been presented in Table 1 [5].

Table 1: Parameter Values

Mechanical Parameters	Electrical Parameters
Ki = 2.1 Nm/A	m = 0.1017 Kg
Kb = 2.1 V/Rad/sec	M = 0.157 Kg
$R_a = 2\ \Omega$	l = 0.5 m
Jm = 0.6578 Kgm2	r = 0.011 m
Bm = 0.0606	---

The isolated equilibrium point of the system is $X^T = [0 \quad 0 \quad 0 \quad 0]$ and the system is quite unstable in other points. It means, ball goes out of the beam

length or the beam angle increases enormously. Besides this, the equilibrium point of the system has no attraction area and every faint disturbance will drive plant to undesired points.

The purpose of controlling task for the ball and beam system is to bring and keep ball in a specific position such as $X_0 = [0 \quad 0 \quad x_0 \quad 0]^T$. The nonlinear system in (3) is locally controllable and observable around this point. Furthermore, the linearized model has zeros and poles in the right side of plan. Therefore, the underlying nonlinear system is non-minimum phase and unstable.

One of the most important reasons for controlling this system is that the behavior of many systems is very similar to it. If we can overcome problems rising from high nonlinearity or instability in the ball and beam system, accordingly, we will be able to develop and apply the devised techniques to control other systems with similar troubles.

3 Multistage Control

In commonplace fuzzy controllers, usually the error and its derivative are frequently used as inputs to the rule base. There is a rule base (may be several) that produces control signals based on logical algorithms. In these methods, number of error inputs increases considerably if the system has more than one output. Oftentimes this leads to a very large and huge rule base which is undesirable for controlling tasks particularly in real time applications. Not only this, but also any modification or readjustment of such a rule base will be really a tiresome and time consuming task. To remedy this problem, an innovative method should be established that while it uses all error signals for control, dimension increment and related problems are subdued well. Two most famous techniques for overcoming this problem are decentralized control and multistage control. Conformity on human's experience and reasonable structure of multistage control technique has motivated some researchers to devise and implement controllers based this algorithm [7].

Multistage controllers have been addressed much for multi input-multi output system controlling tasks [9], [10], [11], [12]. In contrast with usual controllers that produce control signals in one stage, they produce the control signals in several stages. Multistage control has two renowned structures named skew-tree and ternary-tree. Figure 2 schematically represents these two structures. The number of rules and stages are equal in two structures [7], but logic of producing control signals is thoroughly different. There is a parallel mood in the ternary-tree structure; it means, some stages can be done and computed concurrently. In contrast, in skew-tree structure, each stage is dependent to outputs of former stages, so parallel calculation is not possible. While it seems that the

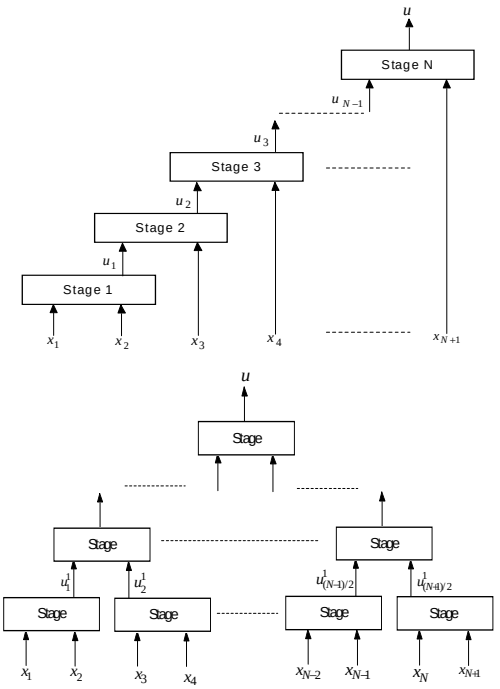


Figure 2: Skew-tree (top) and ternary-tree (bottom) structures.

ternary-tree structure is faster than the skew-tree structure, in most cases, skew-tree structure has more conformity on human's knowledge and experiences in sense of control. Which structure is used is determined based on the underlying system and controlling goals.

To control the ball and beam system, we have used the skew-tree structure to profit transparency in its structure. Next section presents the proposed fuzzy controller using this structure and human's experiences.

4 The Proposed Multistage Fuzzy Controller

The ball and beam system is a Single Input-Multi Outputs (SIMO) system. The motor voltage is the only input of the system while position and velocity of the ball on the beam, angle and angular velocity of the beam are four outputs of the system. Based on skew-tree structure and considering two inputs for every stage, number of the stage, should be three in the multistage controller. As a noticeable aspect of the proposed controller, we declined the number of stages by eliminating a middle stage. This elimination was made possible by a proper definition and selection of not only inputs, but also outputs in the structure of multi stage controller. The structure of the proposed controller has shown in Figure 3.

This controller is a mimicry of human strategy for keeping the ball on the beam. Every expert human

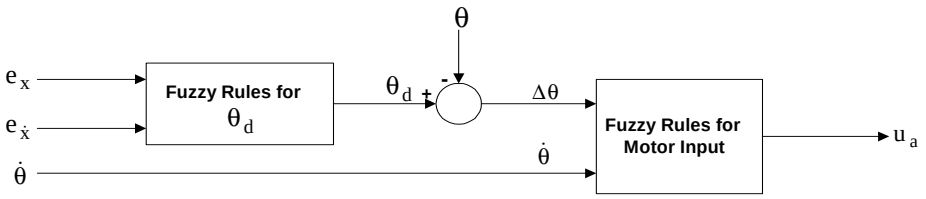


Figure 3: Schematic of proposed controller.

handles and manages beam angle in such a way that ball is compelled to go toward the beam center or any other desired point. If ball is going to exceed beam length, the proper task is to turn beam in a way to force the ball to change its movement direction. Amount of the beam turn directly depends on ball position, ball velocity, beam angle and angular velocity of beam. So, it is reasonable to use these signals as inputs to the multistage controller.

Use of human's experience in our design leads to a multistage controller which decides and performs like a human for keeping and manipulating ball on the beam. It means that the designed controller has the maximum analogy to an operator while its controlling capabilities are much more than human's ones.

In the first place, ball position on the beam and its velocity are passed as inputs to the first stage of multistage controller. This stage produces an appropriate angle which can drive ball to the favored positions on the beam. In the other word, this stage sends out the best value for the beam angle regarding to the ball position and its velocity. Actualization of this value for beam angle will steer ball to the desired positions properly. "Reset beam angle if $\dot{x}$ doesn't assist decreasing e_x" is the basic principle in construction of rule base of this stage given in Table 2. Figure 4 displays membership functions for input and output variables of this stage. For each variable we have considered 7 membership functions, namely NB, NM, NS, Z, PS, PM and PB. N, Z, P, S, M, and B are abbreviations for negative, zero, positive, small, medium, and big respectively. So, NB stands for a membership function with a highly negative discourse universe and so far.

Following human's procedure for controlling of ball and beam system and decreasing required stages can be fulfilled by turning beam in such a way that it quickly takes the output of the former stage. For realizing these goals as well as declining the number of stages, we consciously address the difference between the beam angle and desired angle calculated in the previous stage as one of the inputs to the stage ahead. This can be graphically understood from Figure 3. This difference and the actual angular velocity of beam are passed as inputs to the second stage which causes a turn in the beam assisting ball movement to the desired positions. Again, this stage is an intelligent implementation of human's experiences established based this simple but highly effective rule: "if present angular velocity leads to a bigger difference, reset motor voltage." Rules for this stage and membership functions of input/output variables are given in Table 3 and Figure 4 respectively.

The final point to note here is that there is an indirect control mood in the proposed design. Other control methods directly regulate motor voltage using position error and they can not be easily, if not possible, interpreted by human's experiences [5]. In our method, the first stage gives the best value for the beam angle which steers ball to the favored points. For realization of this angle, we skillfully tune motor voltage in the second stage with a complete sense about what's happening. In fact, we control ball position based on the beam angle which is similar to one a human does. This is possible due to the used multistage fuzzy controller.

Table 2: Fuzzy rules for the first stage

e_x \ $\dot{x}$	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	Z
NM	NB	NM	NM	NB	NS	Z	PS
NS	NB	NM	NS	NM	Z	PS	PS
Z	NM	NM	NS	Z	PS	PM	PM
PS	NS	NS	Z	PM	PS	PM	PB
PM	NS	Z	PS	PB	PM	PM	PB
PB	Z	PS	PM	PB	PB	PB	PB

Table 3: Fuzzy rules for the second stage

$\Delta\theta$ \ $\dot{\theta}$	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	Z
NM	NB	NB	NB	NM	NS	Z	PS
NS	NB	NB	NM	NS	Z	PS	PM
Z	NB	NM	NS	Z	PS	PM	PB
PS	NM	NS	Z	PS	PM	PB	PB
PM	NS	Z	PS	PM	PB	PB	PB
PB	Z	PS	PM	PB	PB	PB	PB

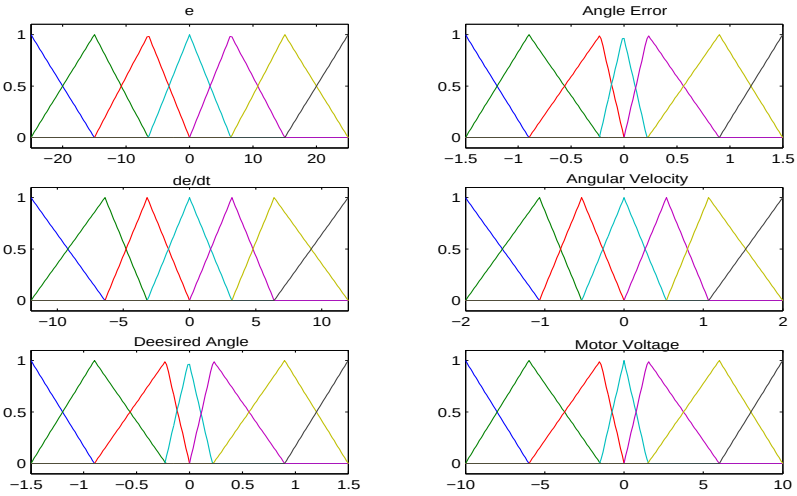


Figure 2: Membership functions for input/output variables in two stages.

5 Simulation

In simulations, we considered two scenarios. In the first one, the controller should drive ball to a desired point and keep it there. In the second one, it is desired that ball tracks a reference trajectory on the beam. To demonstrate capabilities of the proposed controller, we implement two other controllers (LQR and Fuzzy PI with a resetting factor) and compare our results to theirs.

In the first simulation, ball is initially in 20 cm of the beam center and controllers should bring the ball to the center of beam. Figure 5 shows ball move on the beam and beam angle for three controllers.

In the second simulation ball tracks a square periodic trajectory on the beam with a fixed period. In this case, the desired ball position changes from –10 cm to 10 cm and conversely with a 10s period. Ball position and beam angle for the three controllers are displayed in Figure 6.

The simulations results clearly reveal that performance of the multistage fuzzy controller is much better than the LQR and the fuzzy PI controllers. For a better sense about this matter, we compare the obtained results for these three controllers from different viewpoints such as settling time, steady state error, and overshoot which are the most important criteria for assessing of a controller performance.

The LQR and multistage fuzzy controller have a smaller settling time than the fuzzy PI controller. The settling time for the fuzzy PI controller is about 6s while it is about 4.5s for the rest. Furthermore, in the second simulation, the settling time for the LQR controller has increased, but it is approximately constant for the multistage controller.

From the other hand, the steady state error of the closed-loop system response is almost equal to zero for the three controllers in the first simulation. However with respect to the tracking error, the fuzzy controllers are much better than LQR controller and place ball to desired points more precisely.

Aside from transparency in structure and all mentioned benefits above, the most important advantage of the multistage controller over the two other controllers is that its implementation leads to a very smooth movement of ball on the beam. Ball goes from one side to another side of the beam without any overshoot or fluctuation around desired points in a short transient time than do the others. Other controllers produce some overshoots and variations during ball movement on the beam. Ball first exceeds the desired points and then controller efforts to bring it back to the desired points leading to some undesired fluctuation.

A look at beam angles in Figure 5 and Figure 6 also emphasizes strength of the proposed controller. It clearly shows that the multistage fuzzy controller requires the minimum beam turn for driving ball to the desired points. The turning range for this controller is [-0.5 0.5] degree while it is [-1.5, 0.5] and [-2, 1.5] degree for fuzzy PI and LQR controllers respectively. This means that ball moves more smoothly and uniformly in the case of using multistage fuzzy controller.

All these illustrations introduce the designed controller as a powerful one for handling the ball and beam system. The proposed controller not only hires fewer rules in comparison to the common techniques, but also outdoes over those traditional ones in term of reducing settling time and overshoot which lead to the smooth movement of ball on beam.

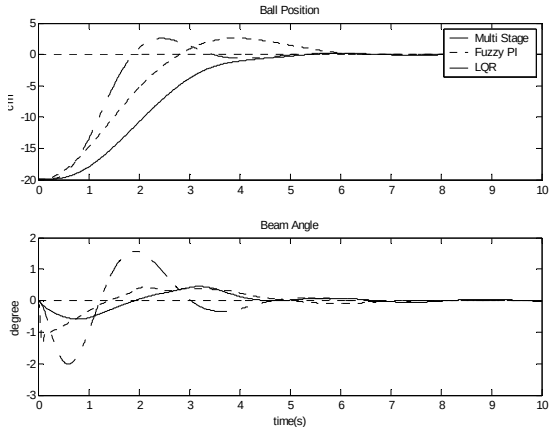


Figure 5: Ball position (top) and beam angle (bottom) in the first simulation.

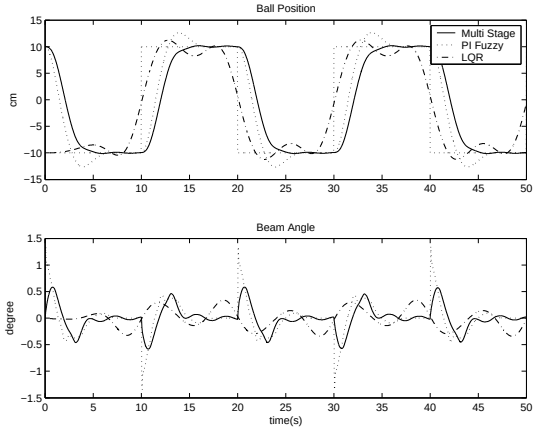


Figure 6: Ball position (top) and beam angle (bottom) in the second simulation.

6 Conclusion

In this paper a multistage fuzzy controller was designed for the classical ball and beam system. The most noticeable advantage of the proposed controller is its conformity on human's experiences and knowledge for controlling and managing difficult tasks such as moving a ball on the beam. In design of the proposed controller, we avoided combination of nonlinear controllers with fuzzy controllers to keep transparency in the controller structure. In comparison to classical controllers such as LQR and combinational controllers such as fuzzy PI, the proposed controller makes the overall system more stable and gives a very smooth movement of the ball on the beam. Furthermore it results to a remarkable decrease in settling time and eliminates overshoots around the desired points. Decreasing computational mass and having a few number of rule bases are other advantages of the proposed method achieved through proper input-output selection for stages of multistage controller. This method with some modification in structure can be applied to other dynamic systems to benefit from mentioned advantages.

Acknowledgement

The authors acknowledge the kind contribution of Ms R. Naghavizadeh during the preparation of this paper. The authors also acknowledge the support of Australian Research Council (ARC) under discovery grants DP0557154.

References

[1] Yasuhiko Dote, Seppo J. Ovaska, Industrial Applications of Soft Computing: A Review, Proceeding of IEEE, (2001) 1243–1265.

[2] P.J. King, E.H. Mamdani, The Application of Fuzzy Control Systems to Industrial Process, Automatica, Pergamon Press, Vol. 13 (1997) 235–242.

[3] E.H. Mamdani, Application of Fuzzy logic to Approximate Reasoning Using Linguistic Synthesis, IEEE Trans. on Computers, Vol. 26, No. 12, (1977) 1182–1191.

[4] L. Wang: A Course in Fuzzy System and Control, Prentice Hall, 1997.

[5] H. Hamidian, M.B. Menhaj: Design and Implementation of Fuzzy Controller for Ball & Beam System, MS Thesis (in Prtsian), Amir Kabir University, Iran, (1995).

[6] H. Hamidian, M.B. Menhaj, Design and Implementation of Ball and Beam System with Fuzzy Controller, 6th Iranian Conference on Electrical Engineering (in Persian), Iran, (1995) 189–194.

[7] Jihong Lee, On Methods for Improving Performance of PI-Type Fuzzy Logic Controllers, IEEE Tran. on Fuzzy Systems, Vol, 1, No. 4, (1993) 298–301.

[8] C. T. Lin, C.J. Lin, L.F. Chung, Neural Fuzzy Control of Unstable Nonlinear Systems, International Conference on Systems, Man and Cybernetics, Canada, (1995) 3666–3671.

[9] Moon G. Joo, A method of converting conventional fuzzy logic system to 2 layered, hierarchical fuzzy system, IEEE International Conference on Fuzzy Systems, (2003) 1357–1362.

[10] G. U. S. Raju, J. Zhou, R. A. Kinser, Hierarchical Fuzzy Control, International Journal of Control, Vol. 45, No. 5, (1991) 1201–1216.

[11] H. P. Chen, T. M. Parng. A New Approach of Multi-stage Fuzzy Logic Inference, Fuzzy Sets and Systems, Vol. 78, (1996) 51–72.

[12] M. Russo, L. Jain: Fuzzy Learning And Applications, CRC Press, USA, 2001.

Modifying the Classic Peak Picking Technique Using a Fuzzy Multi Agent to Have an Accurate P300-based BCI

Gholamreza Salimi Khorshidi
School of cognitive sciences, Institute for studies in theoretical physics and mathematics, Tehran, Iran
gsalimi@bme.aut.ac.ir

Ayyoub Jaafari
Faculty of biomedical engineering, Amirkabir University of technology (Tehran Polytechnic), Tehran, Iran
ajafari@bme.aut.ac.ir

Ali Motie Nasrabadi
Group of biomedical engineering, Faculty of Engineering, Shahed University, Tehran, Iran
nasrabadi@shahed.ac.ir

Mohammadreza Hashemi Golpayegani
Faculty of biomedical engineering, Amirkabir University of technology, Tehran, Iran
hashemi_g@morva.net

Abstract

EEG-based brain computer interface (BCI) provides a new communication channel between the human brain and a computer. The classification of EEG data is an important task in EEG-based BCI. In this paper we present a new modification on classic "Peak Picking" to make a better detection for a specific pattern in EEG. The new method shows a significant improvement in P300 detection which is a common approach in BCI systems. The proposed model uses more than one scalp electrode and combines the outputs with a fuzzy technique, to detect P300 cognitive component.

Keywords: EEG, ERP, BCI, Peak Picking Method (PPM), Fuzzy Membership Functions

1 Introduction

One of the most interesting fields of current researches is to develop a machine that can communicate with a brain, directly. Many methods for discrimination between different mental and cognitive activities have been developed so far, like; many techniques such as feature extraction, feature selection and classification. Each of these methods has their own benefits and drawbacks that make their use totally case-dependent. BCI systems use different methods to communicate with human brain. One of the most common communication techniques is to find a specific pattern in scalp-recorded EEG; like P300, N400, and N170 or generally ERP components. Different methods have been proposed to find a specific pattern in EEG, some like template matching and peak picking are classical ones while other methods like using a feature extraction block in combination with a classifier are modern ones [1, 2, 3].

The aim of this study is to show that a modification in a classical method (Peak Picking) can yield good results for P300 detection. This modification is using a simple fuzzy multi-agent to vote for a final decision based on decisions made by some Peak Picking blocks for some scalp electrodes (Fig.1).

So, this paper is organized as follows. After this introduction, in part 2, the EEG data will be introduced. Then in part 3, a brief review over Peak Picking will show its efficiency for P300 detection. In part 4, the proposed model for the classifier will be introduced and used to show the benefit with new model. Finally in part 5, the conclusion will be made to show the efficiency of this modification.

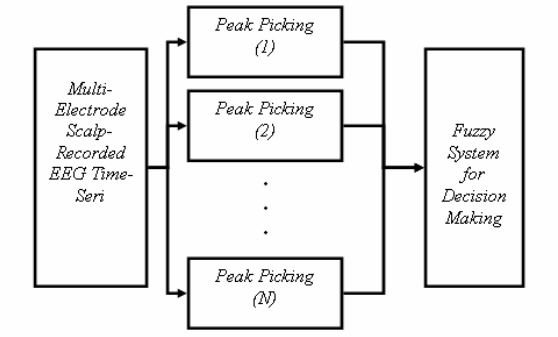


Figure1. Block diagram of the proposed classifier which comprises N template matching (TM) blocks and a fuzzy decision making system, (representing a typical multi-agent classifier)

2 EEG Data

In the present article, the EEG data is BCI Competetion2003 EEG dataset recorded from 64 scalp positions with P300 speller paradigm in which the subject was supposed to distinguish between two classes of stimuli; one class which contains a P300 cognitive component and the other one which lacks this component.

The aim of this competition is to detect the P300 component in scalp recorded EEG; which seems identical to our aim in P300 detection in this study. After data acquisition, there will be a 64 electrodes EEG data. Processing all these channels will be a time consuming task. So two important issues most be considered; 1) the efficiency in processing and 2) the classification accuracy. The first step to have an efficient classification is to know the scalp active regions during the P300 speller paradigm. In Fig.2, the average waveform for EEG data for a subject is shown for two classes of stimuli; with P300 (the upper panel, the dashed line) and without P300 (the upper panel, the solid line) for Cz electrode. The lower panel in Fig.2 shows the r value for these two classes of data. The r value (Eq.1) represents the difference between two classes of data with respect to time in which i is time sample, j is the trial number, and s is the EEG with P300, n is the EEG without P300.

$$\mathrm{r}_{\mathrm{i,j}} = s_{i,j} - n_{i,j} \quad (1)$$

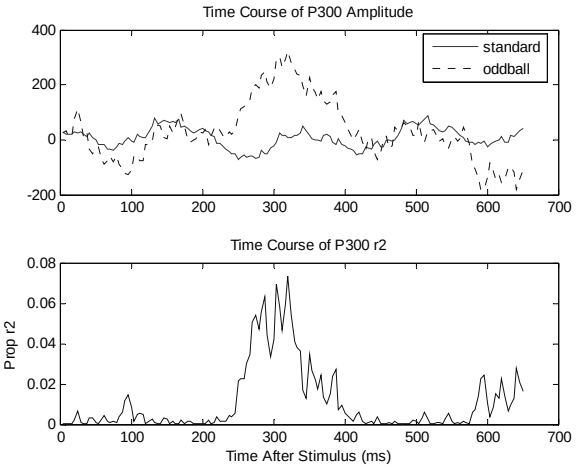


Figure2. (Upper panel) average EEG for two classes of stimuli over Cz electrode that shows a significant difference in about 300ms. (Lower panel) squared difference value (r) for each sample which confirms the importance of 300ms in class discrimination.

The r value shows that the major difference between two classes, happens around 300ms which is actually the time associated with P300 occurrence in the literature. Using plots in Fig.3, the major active regions during P300 occurrence can be estimated. This plot shows a head plot at 300ms after stimulus onset for one subject using all 64 scalp electrodes. Using this plot, it is easier to decide which electrodes to choose as the most efficient electrodes for P300 detection.

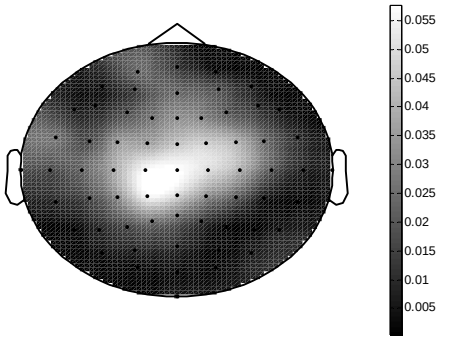


Figure3. The head plot for r value at 300ms after stimulus which can represent the activity at this time in head surface.

3 Using Classic Peak Picking for P300 Detection

As the name indicates, the aim of Peak Picking Method (PPM) is to detect a specific pattern based on its peak value in a region associated with a specific cognitive component. As far as the aim of this study is to use PPM for a P300 based BCI system, the aforementioned region is about 200 to 400ms which is the occurrence time of the P300 cognitive component. The mathematical background of this technique is quite simple and comprises a comparison between two values; a prespecified value (or the threshold) and a critical point's value (a point around 200-400).

If the point's value is bigger than this threshold, then the trial contains the P300 and if it's smaller than threshold, it doesn't contain the P300. In spite of its simplicity, this method yields good accuracy for a special P300 detection task; because of a dominant peak related to P300 occurrence around 300ms after stimuli onset [4].

Based on results from previous section, which adopted the time series values to show the active regions, it is obvious that using all 64 electrodes is not an efficient method. In addition to its long process time, this approach can cause a misclassification because of a bad detection in unrelated electrodes (like those near to Occipital or Parietal lobes).

As far as we need more than one scalp electrode for our modification, it seems necessary to find the best electrodes for a more accurate PPM based P300 detection. Before using PPM, data was preprocessed using band pass filtering (0.5-30Hz) which will cancel

the high frequency noises and keep the ERP related frequency content of the trial. Using this technique on Fz, Cz, Pz, Oz, F3, F4, C3, C4, P3 and P4 electrodes the results (Table1) show that the accuracy for this technique is not acceptable (comparing with the results of new researches using modern techniques). The P300 detection process in a pseudo code step by step form is as follows;

1- Select a trial from an EEG dataset.

2- Filter this trial with a band-pass filter (0.5-30Hz) to cancel the high frequency content of the trial.

3- For the selected electrode, use a calibration to find its corresponding threshold value. The reason for this task is that the ERP over each electrode can have a different peak value and occurrence time.

4- Select the time samples from 250 to 350 and find the average value for them (for the BCI2003 data, this time interval will contain about 24 points). For this part a test was done for some other intervals that finally, this interval yielded the best result.

5- If the resulted average value is bigger than the threshold calculated in section3, the trial contains P300 and if not, then the trial doesn't contain this component.

6- Calculate the "Target Accuracy" and "Non-target Accuracy" and "Total Accuracy" for each electrode.

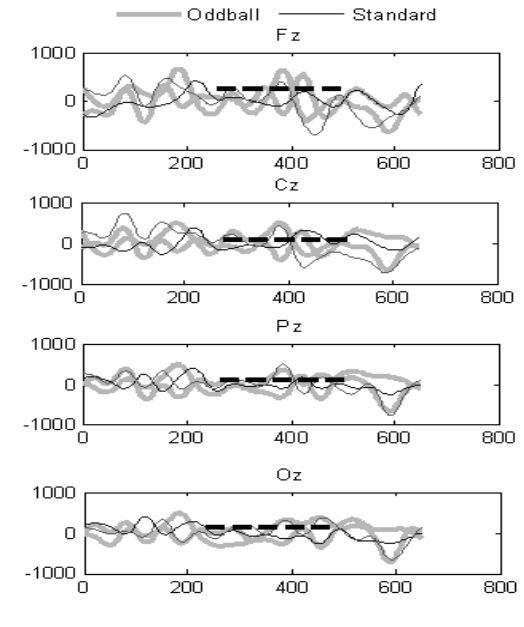


Figure4. Single trial plots for 4 scalp locations for target and non-target stimuli which indicate the difficulty with this method for an accurate P300 detection. The horizontal dashed bold lines show the threshold value.

According to percentage in columns 2 and 3 of Table1, it seems obvious that, for a specific trial, different channels can show different labels. This means that using a combination of outputs, based on the previous knowledge of the classifier, can make the results better. In the next section, using a fuzzy multi agent for this combination will be discussed.

Table 1: The accuracy for P300 detection using the classic PPM over some scalp electrodes (in percent) over 2000 trials (1000; target and 1000; non-target).

Electrode	Target Accuracy	Nontarget Accuracy	Total Accuracy
Fz	65.5	63.3	64.4
Cz	66.5	77	71.75
Oz	48.4	42.4	45.4
Pz	52.5	58.6	55.55
F3	61.3	60.5	60.9
F4	64.2	60	62.1
C3	67	70.6	68.8
C4	75.3	70.6	**72.95**
P3	52.6	55.6	54.6
P4	62	56.3	59.15

4 Modifying the Classic Peak Picking for P300 Detection

The spreading use of fuzzy logic shows its power in different task; specially inference and decision making. As shown in previous sections, a specific trial, can obtain different labels after going through each PPM block in Fig.1. This means that for each trial, some of PPM blocks can have accurate P300 detection while some others cannot.

In this section, we are going to modify PPM for a better and more accurate detection by simultaneously using it over some electrodes and voting over outputs to have a final label which will be better than a single PPM block. It seems to be a good idea to combine information acquired from different scalp positions to find the best combination for a more accurate classification [5].

For this reason, we used a fuzzy block as a voter or mixing block at the output to have all the PPM blocks' outputs as input to make the final output. As shown in Fig.5 and Table1, using a threshold based classification cannot yield good accuracy, at least for one scalp electrode. So it seems necessary to find a solution to combine the results from each electrode.

In this work, for good combination, the process was as follows. First of all, the classifier must get calibrated over a dataset (a train dataset). For training process, for each electrode we need a Max and Min value which represents the 0.95 and 0.05 of input range in Fig.6 respectively. Then using this Max and Min values, each average value (step 4 of previous section) must be normalized to a point between 0.05 and 0.95 (using Eq2).

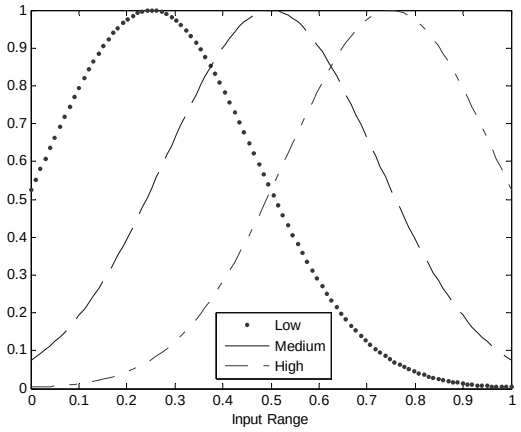


Figure6. Three Gaussian membership functions used to assign the outputs with a label; High, Low or Medium

$$NormVal = [\frac{(0.95)-(0.05)}{Max-Min} * (Val - Min)] + Min \quad (2)$$

After doing all these operations, average value (step4 of previous section) will be a member of three rules in Fig.6 with a specific membership.

Having membership values for each trial over each electrode and the rules in Table2, the final label for the input trial will be determined. Because of poor accuracy of Oz, Pz, P3 and P4 electrodes, they will not be taken into consideration (so we will use only Cz, C3, C4, Fz, F3 and F4).

Using this technique, for the same input trials as what used for section2, the accuracy result is shown in Table3. As can be seen, the results are really better than the previous ones.

Rules in Table2 are extracted based on some experimental trial and error procedure. This table shows a reduced set of rules that were the most effective and accurate ones.

Table2. Fuzzy rules used to determine the final label of an input trial based on its TM output over electrodes Cz, C3, C4, Fz, F3, and F4.

Rule#	If	Then
01	Sum (High)>Sum(Low)	Label=High
02	Sum(Low)>Sum(High)	Label=Low
03	Sum(Low)=Sum(High) C3=C4=Cz=Medium	Label=Low
04	Sum(Low)=Sum(High) C3=C4≠Medium	Label=Label(C3)
05	Sum(Low)=Sum(High) C3=Cz≠Medium	Label=Label(C3)
06	Sum(Low)=Sum(High) Cz=C4≠Medium	Label=Label(Cz)
07	Sum(Low)=Sum(High) C3=C4≠Medium	Label=Label(C3)
08	Sum(Low)=Sum(High) C3=C4=Medium Cz≠Medium	Label=Label(Cz)
09	Sum(Low)=Sum(High) Cz=C3=Medium C4≠Medium	Label=Label(C4)
10	Sum(Low)=Sum(High) C4=Cz=Medium C3≠Medium	Label=Label(C3)

Table3. Accuracy values for final system (Fig.1) over 2000 trials (1000 target and 1000 non-target).

System	Target Accuracy	Nontarget Accuracy	Total Accuracy
TMs + Fuzzy Voter	81.2	78.1	79.65

5 Review of the Method for a Typical Stimulus

Now that everything is described for this method, in this part we will evaluate a sample trial to detect whether it contains the P300 or not. The following step will be done till the final output (label) gets determined:

1- A trial must be selected for a particular stimulus for electrodes Cz, C3, C4, Fz, F3 and F4.

2- All six trials from previous step must get filtered for a noise cancellation (0.5-30Hz band-pass filter).

3- The average of samples ranging from 250 to 350 must be calculated. If it is bigger than the Max, then it contains the P300 and if not, it will get normalized using Eq2.

4- The output of step3 must go for membership functions in Fig.6 to determine its membership for three rules.

5- It will yield six memberships for six electrodes that will determine the final label based on rules in Table2.

Following these steps, we will have a simple but fairly accurate method for P300 detection. The final accuracy of 79.65% over 2000 samples is a good result for a simple classification procedure like our modified PPM.

6 Conclusion and Discussion

In this study, we used a new technique to modify a simple classification technique like Peak Picking. As shown in results, the poor PPM results got much better by a simple combination based on a fuzzy rule set. This idea can be used for better accuracy by combining some more accurate classifiers; such as SVM, LDA, ANN, etc.

The main advantage of this method that caused such good results was to combine information from different scalp places for a better classification.

References

[1] B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer, The BCI Competition 2003: Progress and perspectives in detection and discrimination of EEG single trials, in IEEE Trans. Biomed. Eng., vol. 51, pp. 1044–1051, June 2004.

[2] Handbook of Electroencephalographic Clinical Neurophysiology, vol. 3, Endogenous Event-Related Potentials, T. W. Picton and S. A. Hillyard, Eds., Elsevier, Amsterdam, The Netherlands, 1988, pp. 361–426.

[3] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, “A spelling device for the paralyzed,” Nature, vol. 398, pp. 297–298, 1999.

[4] Piotr Jaskowski, Rolf Verleger, An evaluation of methods for single trial estimation of P300 latency, Psychophysiology, vol.37, 2000, pp.153-162

[5] Fuzzy Expert Systems and Fuzzy Reasoning, William Siler, James J. Buckley, Wiley Interscience, 2006

Toward a Solution for "Genetic Algorithm"-"Support Vector Machine" Combination to Have a Reliable P300-based BCI

Gholamreza Salimi Khorshidi
School Of Cognitive Sciences, Institute For Studies In Theoretical Physics And Mathematics, Tehran, Iran
gsalimi@bme.aut.ac.ir

Ayyoub Jaafari
Faculty Of Biomedical Engineering, Amirkabir University Of Technology (Tehran Polytechnic), Tehran, Iran
ajafari@bme.aut.ac.ir

Ali Motie Nasrabadi
Group Of Biomedical Engineering, Faculty Of Engineering, Shahed University, Tehran, Iran
nasrabadi@shahed.ac.ir

Mohammadreza Hashemi Golpayegani
Faculty of biomedical engineering, Amirkabir University of technology, Tehran, Iran
hashemi_g@morva.net

Caro Lucas
School Of Cognitive Sciences, Institute For Studies In Theoretical Physics And Mathematics, Tehran, Iran
And
Control And Intelligent Processing Center Of Excellence, Electrical And Computer Engineering Department, University Of Tehran, Tehran, Iran
lucas@ipm.ir

Abstract

In this study, we present a new framework to combine the Genetic Algorithm (GA) and Support Vector Machine (SVM) for an accurate P300 detection. P300 is the most important cognitive component used in a Brain Computer Interface (BCI) occurring around 300ms after stimulus onset. To detect this component and discriminate it from ongoing EEG, a typical Continuous Wavelet Transform (CWT) feature extraction is adopted in combination with GA and SVM for feature selection and feature classification, respectively. SVM is a powerful classifier that yields good classification results in its optimum point. In this study we use a new technique in GA to select the best feature subset (optimum feature set) and tune the SVM (for the selected feature subset) simultaneously, which showed a significant improvement in classification results.

Keywords: BCI, P300, CWT, SVM, GA.

1 General formatting instructions

A reliable brain-computer interface (BCI) can be a big step forward for rehabilitation and different clinical and non-clinical applications. There are different approaches in designing a BCI system from its biological input from the subject; like MEG and EEG, among which EEG-based and especially ERP-based BCI systems are the most favorite ones from applicability, cost, time resolution and some other viewpoints [1]. The most common ERP component for a BCI system is P300 which is a positive peak occurring 300ms after stimulus onset. Compared to other well-known components, P300 has a robust and obvious waveform that makes it suitable for a typical BCI system. In a P300-based BCI system, the major task is to discriminate trials with P300 from those without it or to detect the P300 cognitive component. A typical P300 detector BCI, comprises following steps; EEG acquisition, EEG Preprocessing, Feature Extraction from EEG, Feature Selection and Classification. Using all or a subset of these blocks is a case dependent issue.

In this study, we used the BCI Competition 2003 EEG dataset for P300 speller paradigm (IIb dataset) and then used a CWT process for both preprocessing and feature extraction [2]. Because of their high resolution both in time and frequency domain, wavelet coefficients from the CWT process are good features for classification, especially for a P300 detection task. As a classifier, we adopted the SVM in three ways; Linear SVM (LSVM), SVM with Gaussian kernel (GSVM) and SVM with polynomial kernel (PSVM). For all these classifiers, GA is used to both select the optimum feature set and tune the SVM for its optimum point. To describe all these

issues, this paper is organized as follows. After a brief introduction, EEG dataset will be introduced and then the wavelet analysis procedure for feature extraction from EEG time series will be described. As the classifier, SVM, its parameters and their importance will be briefly introduced. Finally after describing the GA process, the results will be evaluated and discussed.

2 Materials and Methods

A typical P300 waveform can be extracted with a simple averaging technique. As shown in Fig.1, based on this technique, P300 waveform can have different waveforms (from its amplitude and shape viewpoints) in different scalp places which is because of different activities in different brain areas in response to each category of stimulation. The maximum peak and better discrimination between average waveforms can be seen in Cz and Fz channels while the worst discrimination is shown in Oz and Pz channels. Having this data, the next step is to extract the features from it to help the classifier discriminate between classes, easily.

In the present article, the EEG data is BCI Competition 2003 EEG dataset recorded from 64 scalp positions with P300 speller paradigm in which the subject was supposed to distinguish between two classes of stimulus; one class which induces P300 cognitive component and the other one which lacks this component. The aim of this competition is to detect the P300 component in scalp recorded EEG; which seems identical to our aim in P300 detection in this study. After data acquisition, there will be a 64-channel EEG data with number of trials equal to the number of stimuli. Each trial will contain 156 samples which are associated with a nearly 650ms time duration sampled with 256Hz.

3 Continuous Wavelet Transform

As a very useful and applicable method for EEG trials' analysis, wavelet transform is getting used in a wide range of data analyses in this field. In this study we used continuous wavelet transform (CWT) for extracting time-frequency characteristics of EEG trials as the features describing time series' dynamics. For this reason, we used four wavelets for a multi resolution analysis; two general purpose basis functions (Matlab's Bior3.9 and Db4) and two special purpose wavelets (Quiroga's manually tuned wavelet for ERP analyses and Glassman's SNAP wavelet based on EEG's origin). In this method, each EEG trial will be decomposed into its frequency sub-bands and in each sub-band there will be some coefficients representing the time interval associated with it. This process is shown in Fig.2 for a sample trial with P300 and a sample trial without P300. These coefficients can be used as our features. Using all the scalp channels and all the extracted coefficients, it's time to select the best wavelet and the best features associated with that particular wavelet. Fig.3 shows the D-values (Eq.1) for some features extracted by using all four aforementioned wavelets (first 4000 features sorted by their D-value). As Fig.3 shows, Quiroga's manually tuned wavelet can make more statistically discriminant features than three other wavelets.

$$D= (\mu_1-\mu_2)/(\sigma_1^2+\sigma_2^2)^{0.5} \quad (1)$$

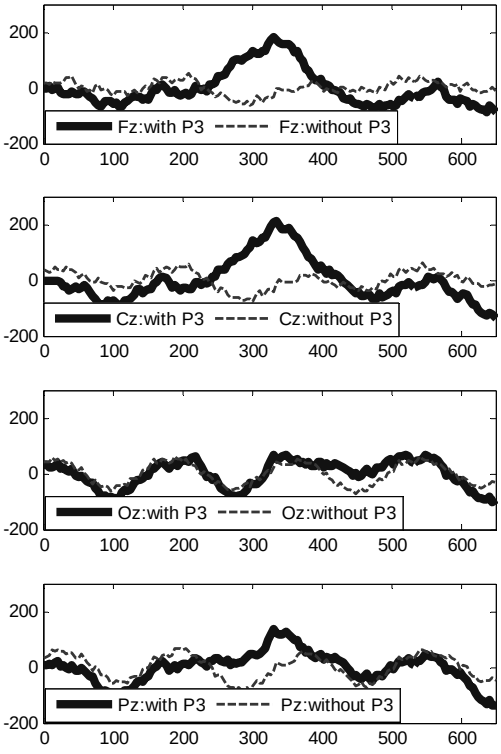


Fig. 1 Typical average waveform of trials with/without P300 over different scalp places which shows an obvious peak on central regions (Cz and Fz channels).

Having the result of Fig.3, we can now have a primary feature selection which will not yield the final optimum subset, but will result in a subset inside which we can have this optimum subset. So finally based on aforementioned D-value, we selected 150 most statistically discriminant features which will get fed into the "feature selection"-"classifier" block set.

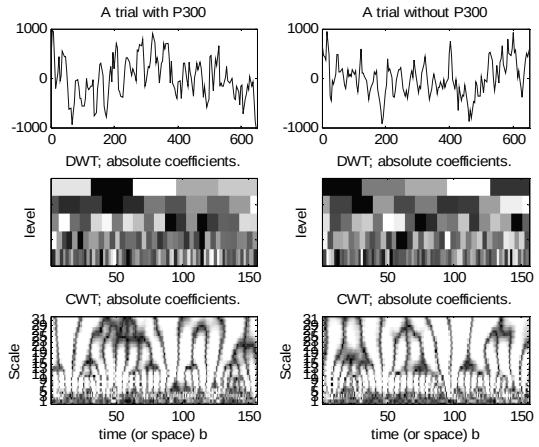


Fig. 2 A typical multi resolution analysis (MRA) on two sample EEG trials; one with P300 (a) and one without P300 (b) with Quiroga's wavelet. Obviously, we can say that the coefficients are different in low frequencies and in times around 300 ms. The DWT is for scales 2, 4, 8, 16, 32 and CWT is for scales 1 to 32.

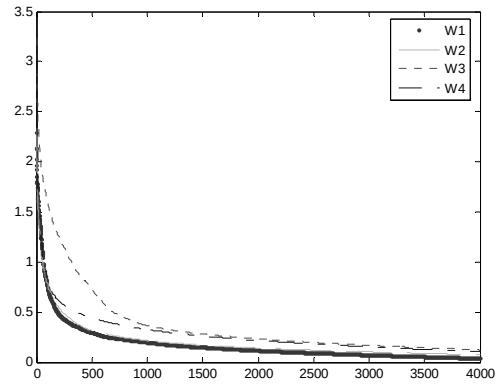


Fig. 3 Statistical discriminancy of 4000 best features (based on their D-value) for four wavelets; W1 (Db4), W2 (Bior3.9), W3 (Quiroga's) and W4 (SNAP).

4 "Genetic Algorithm"-"Support Vector Machine" Combination

The feature set resulting from primary statistical feature selection (using D-value) can not assure us to have the best classification accuracy. It means using all of these features together will not guarantee the best accuracy and also maybe destroy the result compared to an optimum subset of it. For this reason we used genetic algorithm for a final feature selection which will yield the best feature subset in order to make the most accurate classification.

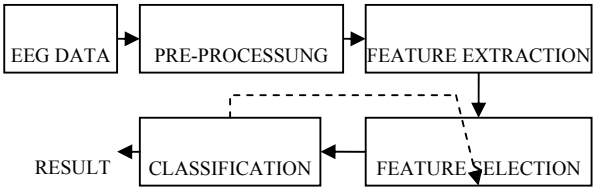


Fig.4 combination of GA and the SVM classifier in which the classification accuracy will affect the feature selection block's output. Finally this interaction between classifier and the feature selection block (GA) will make the system to evolve and select the most discriminant feature set.

SVM classifiers have two important parameters to be tuned; C and σ (standard deviation of the Gaussian kernel). In the easiest approach to tune these parameters, we can use a k-fold cross validation over a specific range of these parameters and finally find the best point from the classification accuracy viewpoint. This approach is not the best way of finding the optimum values for C (trade off coefficient) and σ (Gaussian kernel's variance) in an evolutionary process. Because, after some generations, the resulting feature set is not the same as the one in the first generations. It means, while using GA for feature selection, the input space is changing over generation from the dimension viewpoint which makes it require a new discriminant hyperplane for the optimum classification.

For this reason, we used another approach for determining the best point of using GA-SVM combination from feature set and classifier viewpoint. In a binary GA for feature selection (the classic viewpoint), the chromosome length is equal to the number of features (150 bits for this problem). Finally, bits with 1 value represent the selected features while those with 0 values represent the rejected features. To make the SVM tuning process an evolutionary process, we added 20 new bits to the end of chromosomes which were tuning-specific bits; 10 bits for C and 10 bits for σ. The decimal equivalent of these 10 bits divided by 10 will

represent the final value for each parameter. So the chromosomes will be of length 170 from which 150 bits are representing the selected features and 20 bits are representing the optimally tuned values for C and σ parameters (Fig.5).

150 bits (feature selection)	10 bits (C)	10 bits (σ)

Fig.5 a typical chromosome used for simultaneous feature selection and SVM tuning. As this figure shows, this chromosome comprises three parts; Feature Selection, C tuning and σ tuning bits.

There are some other important parameters in the GA that are supposed to get determined to have a better and more effective evolution; like mutation rate, crossover rate, initial population size and the number of generations to reach the end of evolution. For this reason, we used some tests over our dataset and selected the optimum parameters (as follows).

Crossover and Mutation Rate (c, m): to tune these two parameters, we selected four values for them; (0.3, 0.3), (0.5, 0.5), (0.7, 0.3) and (0.9, 0.1) for population size of 100, elite number of 10 and 2000 generations. The result is shown in Fig.6.

As the result in Fig.6 shows, increasing the crossover rate can make the mating more effective and faster for reaching its optimum point. The problem with having only the crossover is the lack of new evolutionary actions in the final generations as shown in Fig.6-A. In this figure, the fitness has a flat area after reaching a fairly good value and will remain unchanged. To evaluate the effect of mutation rate, we selected the values as in Fig.6-B in which the mutation rate is equal to 50% which shows a jump for better fitness at the middle of the evolution (after the flat area). This means that the mutation rate can show its effect in the final generation when the crossover can not make any new evolution. The other important point about Fig.6-B is that increasing the crossover rate can make the final fitness better. To confirm this idea, we increased the crossover rate to 0.9 with a small mutation rate which results the Fig.6-D in which there's no evolutionary change in the last generations (the small mutation rate) but reaches its optimum point very fast. So we can say having a big crossover can help us have a good evolution.

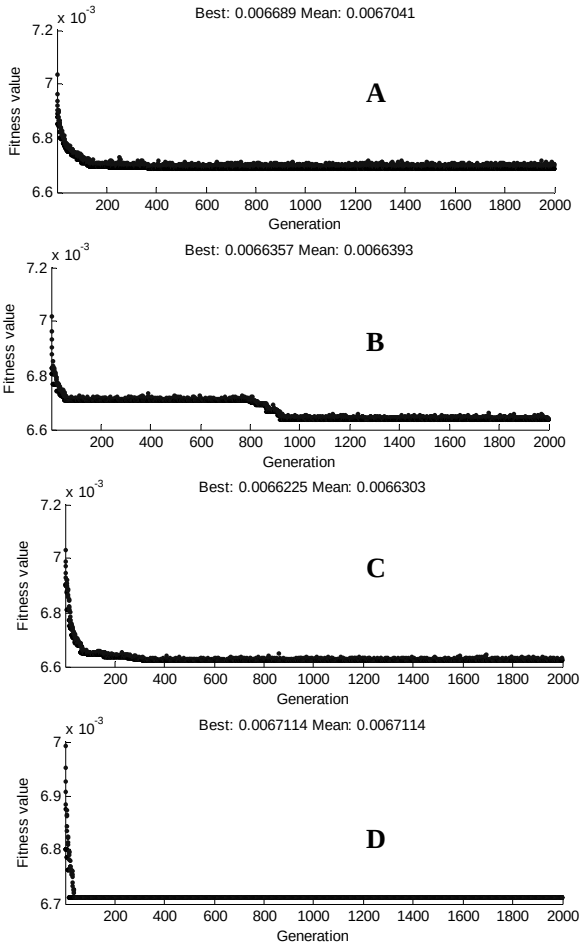


Fig.6 the best fitness value vs. the number of generations in the GA for different (c, m) values; A.(0.3, 0.3), B.(0.5, 0.5), C.(0.7, 0.3) and D.(0.9, 0.1).

So we finally selected the crossover rate and mutation rate the way in which we have both effects very well (a good evolution and effective mating with some new evolution in flat area of the fitness) which results the Fig.6-C. In this situation we have the best evolution in which we have all the things at the best point.

Population Size: the variety is another important issue to make better generations which is affected by the number of individuals in the initial population size. In this study, we tested this factor for two values of 50 and 100 which shows the better evolution for 100 individuals which confirms our initial idea (Fig.7).

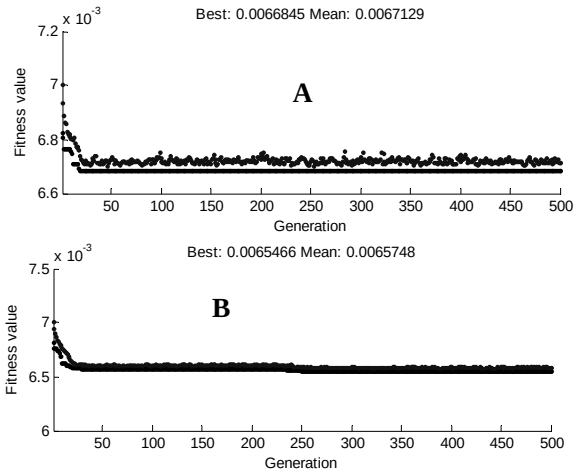


Fig.7 shows the effect of initial population size for two values of A. 50 individuals and B. 100 individuals. As the results show the bigger the initial population size, the better and the more effective the evolution is held here.

Having all these results together, we used the GA for both feature selection and SVM tuning with crossover rate of 0.7, mutation rate of 0.3, initial population size of 100 individuals, 1000 generations, binary chromosomes (individuals) with the length of 170 bits. The results of this process is described and discussed in the next section.

5 Results and Conclusion

To evaluate our method over real EEG data for a BCI application, we selected 2500 trials of aforementioned dataset. Having the EEG trials, we extracted the continuous time-frequency features from the data. Then selected 150 most statistically discriminant features for a simpler computation. Then GA was used with aforementioned parameters in combination with SVM classifier. Two different methods were used for SVM tuning; 1- evolutionary SVM tuning and 2- using a 10-fold cross validation for SVM tuning. For the first method this length was 170 and for the second method, the length of chromosome was 150.

Data was divided into two groups; training dataset (2000 trials: 1000 trials with P300 and 1000 trials without P300) and test dataset (500 trials: 250 trials with P300 and 250 trials without P300). Finally the classification task was done and evaluated. The result of this classification is shown in Table1. Since the same process can be done for different classes of SVM, we tested the method for linear SVM (LSVM), SVM with Gaussian kernel (GSVM) and SVM with polynomial kernel (PSVM). For PSVM, the chromosome length is different, because the range of σ is different from polynomial's degree. So we added 13 bits to the end of chromosomes (163 bits to tune the PSVM).

Table1. The result of classification for different methods of combining the SVM and GA

Method	Accuracy on train	Accuracy on test
LSVM+ First Method	76.8%	74.9%
PSVM+ First Method	81.8%	82.7%
GSVM+ First Method	84.6%	80.5%
PSVM+ Second Method	87.1%	81.9%
GSVM+ Second Method	91.2%	88.7%

As the result show, using the SVM tuning in a simultaneous manner with feature selection can yield a good result compared to the first method in which the SVM is tuned separated from the feature selection process. As mentioned before, this is as a result of tuning the SVM for the specific feature subset selected at the final level.

References

[1] B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer, "The BCI Competition 2003: Progress and perspectives in detection and discrimination of EEG single trials", in IEEE Trans. Biomed. Eng., vol. 51, pp. 1044–1051, June 2004.

[2] Handbook of Electroencephalographic Clinical Neurophysiology, vol. 3, "Endogenous Event-Related Potentials", T. W. Picton and S. A. Hillyard, Eds., Elsevier, Amsterdam, the Netherlands, 1988, pp. 361–426.

[3] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralyzed," Nature, vol. 398, pp. 297–298, 1999.

[4] Emanuel Donchin, Kevin M. Spencer, and Ranjith Wijesinghe, "The Mental Prosthesis: Assessing the Speed of a P300-Based Brain-Computer Interface," IEEE Transactions on

Rehabilitation Engineering, VOL. 8, NO. 2, Page(s):174 – 179, JUNE 2000.
[5] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003 - data set IIb: Support vector machines for the p300 speller paradigm," IEEE Transactions on Biomedical Engineering, VOL. 51, NO. 6, Page(s): 1073-1076, June, 2004.
[6] Vladimir Bostanov, "BCI Competition 2003—Data Sets Ib and IIb: Feature Extraction From Event-Related Brain Potentials with the Continuous Wavelet Transform and the t-Value Scalogram," IEEE Transactions on Biomedical engineering, VOL. 51, NO. 6, Page(s): 1057-1061, June 2004.
[7] Elena L. Glassman, "A Wavelet-Like Filter Based on Neuron Action Potentials for Analysis of Human Scalp Electroencephalographs", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 52, pp.1851-1862, 2005.
[8] R. Quian Quiroga , O.W. Sakowitz , E. Basar , M. Schurmann "Wavelet Transform in the analysis of the frequency composition of Evoked potentials" Brain Research Protocols vol.8, pp.16–24, 2001.
[9] Deon Garrett, David A. Peterson, Charles W. Anderson, Michael H. Thaut, "Comparison of Linear, Nonlinear and Feature Selection Methods for EEG signal Classification", IEEE transaction on neural systems and rehabilitation engineering, Vol.11, No.2, Page(s):141-152, June 2003.
[10] A. Schwaighofer. SVM Toolbox for Matlab. [Online].Available:http://www.cis.tugraz.at/igi/aschwaig/software.html
[11] N. Cristianini and B. Schölkopf, "Support vector machines and Kernel methods: The new generation of learning machines," *AI Mag.*, vol. 23, no. 3, pp. 31–41, 483, 484, 2002.
[12] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. (2003) A Practical Guide to Support Vector Classification. National Taiwan Univ., Dept. Comput. Sci Inf. Eng.. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
[13] Randy L. Haupt, Sue Ellen Haupt, *Practical genetic algorithm*, Wiley-Interscience; 2 edition May 31, 2004
[14] Schroder, M. Bogdan, M. Hinterberger, T. Birbaumer, N. "Automated EEG feature selection for brain computer interfaces", pp. 626-629, 20-22 March 2003.

Session 5

Measures of Vagueness

A characterization of the ambiguity and fuzziness by means of activity orders on the closed intervals in [0,1]

C. Alcalde
Escuela Universitaria Politécnica.
Dep. de Matemática Aplicada.
Universidad del País Vasco.
Plaza de Europa 1,
20018 San Sebastián, Spain.
c.alcalde@ehu.es

A. Burusco
Dep. Automática y Computación.
Universidad Pública de Navarra,
Campus de Arrosadía,
31006 Pamplona, Spain.
burusco@si.unavarra.es

R. Fuentes-González
Dep. Automática y Computación.
Universidad Pública de Navarra,
Campus de Arrosadía,
31006 Pamplona, Spain.
rfuentes@si.unavarra.es

Abstract

We take as a departure point the definition of activity orders, intending to measure the ambiguity and fuzziness of any closed interval in $[0, 1]$. To do it, we define a family of orders on the set of closed intervals of a distributive lattice L, that will allow us to set up some preorders associated to the ambiguity and fuzziness.
Keywords: activity orders, fuzziness and ambiguity.

1 Introduction. Activity orders

If $\mathcal{J}[L]$ represents the set of closed intervals in a distributive lattice L, then some order relations on $\mathcal{J}[L]$ are defined to set up preorders associated to the fuzziness and ambiguity.

To do it, we will start from the concept of *activity order* [10, 9, 11] that was introduced by Serra [10], in the context of the morphologic operators on images of $\mathbb{R}^n$ or $\mathbb{Z}^n$, as follows.

1.1 Definition

Let $\Omega = \{\Psi / \Psi : \wp(\mathbb{R}^2) \to \wp(\mathbb{R}^2)\}$ be a set of functions that transforms the images A of the plane $\mathbb{R}^2$ with the order $(\Psi_1 \leq \Psi_2) \Longleftrightarrow (\Psi_1(A) \subseteq \Psi_2(A), \ \ \forall A \in \mathbb{R}^2)$. Let $i \in \Omega$ be the identity operator: $i(A) = A, \ \ \forall A \in \mathbb{R}^2$. The activity order between operators Ψ_1, Ψ_2 is defined by:

$$\Psi_1 \preceq \Psi_2 \Longleftrightarrow \left\{ \begin{array}{c} (i \wedge \Psi_1) \geq (i \wedge \Psi_2) \\ \& \\ (i \vee \Psi_1) \leq (i \vee \Psi_2) \end{array} \right.$$

Later, Heijmans and Keshet [8] extended this definition to any lattice, generating a family of orders that are shown in the next:

1.2 Definition

Let $(L, \leq)$ be a lattice and $\omega_0 \in L$. The binary relation $\preceq_{\omega_0}$ is defined on L by the expression:

$$(x \preceq_{\omega_0} y) \Longleftrightarrow \left\{ \begin{array}{c} (\omega_0 \wedge y) \leq (\omega_0 \wedge x) \\ \& \\ (\omega_0 \vee y) \geq (\omega_0 \vee x) \end{array} \right. \quad (1)$$

where $\vee$ and $\wedge$ represent the joint and meet in L respectively.

1.3 Proposition

The relation $\preceq_{\omega_0}$ is reflexive and transitive for each $\omega_0 \in L$, that is:

$$(\alpha \preceq_{\omega_0} \alpha), \ \forall \alpha \in L$$
$$((\alpha \preceq_{\omega_0} \beta) \& (\beta \preceq_{\omega_0} \gamma)) \Longrightarrow (\alpha \preceq_{\omega_0} \gamma).$$

It is not necessarily antisymmetric, since the lattice in Figure 1 verifies: $\alpha \preceq_{\omega_0} \beta$, $\beta \preceq_{\omega_0} \alpha$ and $\alpha \neq \beta$. Consequently, the binary relation $\preceq_{\omega_0}$, in general, is not an order relation.

An immediate consequence of the studied theory in [8] is the following result that tells us when the previous relation is an order.

1.4 Theorem

If the ordered set $(L, \leq)$ is a distributive lattice $(L, \vee, \wedge)$, it is verified:

(i) For all $\omega_0 \in L$, the relation $\preceq_{\omega_0}$ defined in *(1)* is an order relation on L.

(ii) The partially ordered set $(L, \preceq_{\omega_0})$ associated to ω_0 is a meet-semilattice $(L, \curlywedge_{\omega_0}, \mathbf{0}_{\omega_0})$ with a minimum element $\mathbf{0}_{\omega_0} = \omega_{\mathbf{0}}$, and where the meet operation $\curlywedge_{\omega_0}$ is given, for all $(\alpha, \beta) \in L^2$, by

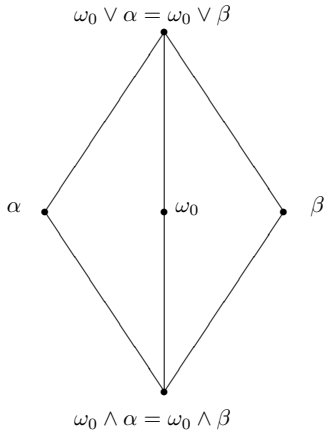


Figure 1: $\preceq_{\omega_0}$ is not an order

$$(\alpha \curlywedge_{\omega_0} \beta) = \xi, \text{ where}$$
$$\xi = [\omega_0 \wedge (\alpha \vee \beta)] \vee (\alpha \wedge \beta) = [\omega_0 \vee (\alpha \wedge \beta)] \wedge (\alpha \vee \beta).$$

Hence, we will work with the distributive lattices and with the following family of orders.

1.5 Definition

If $(L, \leq)$ is a distributive lattice and $\omega_0 \in L$, then the order relation $\preceq_{\omega_0}$ defined in *(1)*, is said to be the ω_0-activity order.

1.6 Example

In Figure 2, we represent some examples of the family of orders $(\preceq_\omega)_{\omega \in [0,1]}$ obtained in the chain $([0,1], \leq)$.

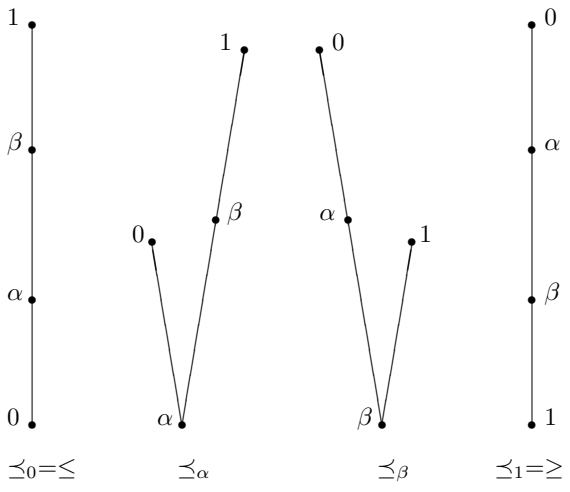


Figure 2: Orders of the family $(\preceq_\omega)_{\omega \in [0,1]}$

Note that $\preceq^{op}_{\omega_0}$, the dual order of $\preceq_{\omega_0}$, is given by:

$$(\alpha \preceq^{op}_{\omega_0} \beta) \Longleftrightarrow \begin{cases} (\omega_0 \wedge \alpha) \leq (\omega_0 \wedge \beta) \\ \& \\ (\omega_0 \vee \alpha) \geq (\omega_0 \vee \beta) \end{cases}$$

Therefore, $(L, \preceq^{op}_{\omega_0})$ is a joint-semilattice where

$$(\alpha \curlyvee^{op}_{\omega_0} \beta) = [\omega_0 \wedge (\alpha \vee \beta)] \vee (\alpha \wedge \beta) = [\omega_0 \vee (\alpha \wedge \beta)] \wedge (\alpha \vee \beta),$$

and ω_0 is the maximum element of $(L, \preceq^{op}_{\omega_0})$.

Note that, for every distributive lattice $(L, \leq)$ with a minimum element $\mathbf{0}$, it is verified: $(\preceq_{\mathbf{0}}) = (\leq)$. If there is a maximum element $\mathbf{1}$, then the relation $\preceq_{\mathbf{1}}$ is $\geq$, (opposite to $\leq$).

Let us see a particular case of this order relation that has already been defined.

1.7 Proposition

The "sharpened" order $\leq^s$ [5, 13] used in fuzziness degree measures [2] and in the theory of fuzzy entropies [12], is a particular case of the orders $\preceq_\omega$. Specifically, for $L = [0,1]$, $\omega = \frac{1}{2}$:

$$(\alpha \leq^s \beta) \Longleftrightarrow (\alpha \preceq^{op}_{1/2} \beta) \Longleftrightarrow$$
$$\begin{cases} \min(\frac{1}{2}, \alpha) \leq \min(\frac{1}{2}, \beta) \\ \& \\ \max(\frac{1}{2}, \alpha) \geq \max(\frac{1}{2}, \beta) \end{cases}$$

If we consider now an interval $[\omega_0]$ represented by $[\omega_0] = [\underline{\omega_0}, \overline{\omega_0}] \in \mathcal{J}[L]$, with $\underline{\omega_0} \leq \overline{\omega_0}$, we can characterize the family of orders that we have defined in the sets of intervals $(\mathcal{J}[L], \leq)$.

1.8 Proposition

If $(L, \leq)$ is a distributive lattice and $(\mathcal{J}[L], \leq)$ is the distributive lattice of the closed intervals in L with the order

$$([\underline{\alpha}, \overline{\alpha}] \leq [\underline{\beta}, \overline{\beta}]) \Longleftrightarrow (\underline{\alpha} \leq \underline{\beta}) \& (\overline{\alpha} \leq \overline{\beta}),$$

then

$$([\underline{\alpha}, \overline{\alpha}] \preceq_{[\omega_0]} [\underline{\beta}, \overline{\beta}]) \Longleftrightarrow (\underline{\alpha} \preceq_{\underline{\omega_0}} \underline{\beta}) \& (\overline{\alpha} \preceq_{\overline{\omega_0}} \overline{\beta})$$

Proof Given the interval $[\omega_0] = [\underline{\omega_0}, \overline{\omega_0}]$, the ω_0-activity order associated to the element ω_0 is

$$(\alpha \preceq_{\omega_0} \beta) \Longleftrightarrow \begin{cases} (\omega_0 \wedge \alpha) \geq (\omega_0 \wedge \beta) \\ \& \\ (\omega_0 \vee \alpha) \leq (\omega_0 \vee \beta) \end{cases}$$

On the other hand, if we take the order relation defined on the set of intervals

$$([\underline{\alpha},\overline{\alpha}] \preceq_{\omega_0} [\underline{\beta},\overline{\beta}]) \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} ([\underline{\omega_0},\overline{\omega_0}] \wedge [\underline{\alpha},\overline{\alpha}]) \geq ([\underline{\omega_0},\overline{\omega_0}] \wedge [\underline{\beta},\overline{\beta}]) \\ \& \\ ([\underline{\omega_0},\overline{\omega_0}] \vee [\underline{\alpha},\overline{\alpha}]) \leq ([\underline{\omega_0},\overline{\omega_0}] \vee [\underline{\beta},\overline{\beta}]) \end{cases}$$

Then, it follows that

$$([\underline{\alpha},\overline{\alpha}] \preceq_{\omega_0} [\underline{\beta},\overline{\beta}]) \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} [\underline{\omega_0} \wedge \underline{\alpha}, \overline{\omega_0} \wedge \overline{\alpha}] \geq [\underline{\omega_0} \wedge \underline{\beta}, \overline{\omega_0} \wedge \overline{\beta}] \\ \& \\ [\underline{\omega_0} \vee \underline{\alpha}, \overline{\omega_0} \vee \overline{\alpha}] \leq [\underline{\omega_0} \vee \underline{\beta}, \overline{\omega_0} \vee \overline{\beta}] \end{cases}$$

$$\Longleftrightarrow \begin{cases} \underline{\omega_0} \wedge \underline{\alpha} \geq \underline{\omega_0} \wedge \underline{\beta} \\ \& \\ \overline{\omega_0} \wedge \overline{\alpha} \geq \overline{\omega_0} \wedge \overline{\beta} \\ \& \\ \underline{\omega_0} \vee \underline{\alpha} \leq \underline{\omega_0} \vee \underline{\beta} \\ \& \\ \overline{\omega_0} \vee \overline{\alpha} \leq \overline{\omega_0} \vee \overline{\beta} \end{cases} \Longleftrightarrow \begin{cases} \underline{\alpha} \preceq_{\underline{\omega_0}} \underline{\beta} \\ \& \\ \overline{\alpha} \preceq_{\overline{\omega_0}} \overline{\beta} \end{cases}$$

□

1.9 Example

If $L = [0,1]$, then we obtain:

$$([\alpha] \preceq_{[0,1]} [\beta]) \Longleftrightarrow (\underline{\alpha} \preceq_0 \underline{\beta}) \& (\overline{\alpha} \preceq_1 \overline{\beta}),$$

that is

$$([\alpha] \preceq_{[0,1]} [\beta]) \Longleftrightarrow (\underline{\alpha} \leq \underline{\beta}) \& (\overline{\alpha} \geq \overline{\beta})$$

It proves that $\preceq_{[0,1]}$ is the inclusion order $\supseteq$ between intervals:

$$([\alpha] \preceq_{[0,1]} [\beta]) \Longleftrightarrow ([\alpha] \supseteq [\beta]).$$

Notation. Given a distributive lattice $(L, \leq)$, for every element $\omega_0 \in L$, the order relation dual of $\preceq_{\omega_0}$ will be denoted by $\leq_{\omega_0}$. That is, $\leq_{\omega_0} = \preceq_{\omega_0}^{op}$.
Therefore, we will have that:

$$\alpha \leq_{\omega_0} \beta \Longleftrightarrow \begin{cases} \omega_0 \wedge \alpha \leq \omega_0 \wedge \beta \\ \& \\ \omega_0 \vee \alpha \geq \omega_0 \vee \beta \end{cases}$$

As we said before, the set L with this order relation is a joint-semilattice where the joint of any two elements is given by

$$\alpha \vee_{\omega_0} \beta = \{\omega_0 \wedge (\alpha \vee \beta)\} \vee (\alpha \wedge \beta)$$

and where ω_0 is the maximum element.

1.10 Example

In particular, if $L = [0,1]$ and $(\mathcal{J}[L], \leq)$ is the distributive lattice of the closed intervals in L with the order

$$[a,b] \leq [c,d] \quad \Longleftrightarrow \quad a \leq c \quad \& \quad b \leq d$$

then

$$[a,b] \leq_{[1,1]} [c,d] \Longleftrightarrow \begin{cases} [a,b] \wedge [1,1] \leq [c,d] \wedge [1,1] \\ \& \\ [a,b] \vee [1,1] \geq [c,d] \vee [1,1] \end{cases}$$

$$\Longleftrightarrow [a,b] \leq [c,d]$$

that is, the order $\leq_{[1,1]}$ is the usual order between intervals $\leq$.

2 Preorder relation associated to the fuzziness of the intervals

In the Fuzzy sets theory, diverse magnitudes have been defined in order to measure the fuzziness [6, 7]. Similar measures have also been considered for interval-valued fuzzy sets [3].

From a different point of view, using the order relations $\leq_{\omega_0}$, we will introduce in this section some preorder relations on the set of closed intervals in $L = [0,1]$, that later we will allow us to define other measures of fuzziness and ambiguity on the interval-valued fuzzy sets.

In order to set up comparisons among the fuzziness of the elements of the set of closed intervals in $L = [0,1]$ we will remind the relation *'at least as fuzzy as'* defined by De Luca and Termini in [6].

2.1 Definition

Let $(L, \leq)$ be a distributive lattice endowed with a negation $'$. Given two elements $x, y \in L$, y is said to be *at least as fuzzy* as x, and we will represent by

$$x \prec_{fuz} y \Longleftrightarrow \begin{cases} x \wedge x' \leq y \wedge y' \\ \& \\ x \vee x' \geq y \vee y' \end{cases}$$

The intuitive idea of this definition is that y and y' are closer in L than x and x'.

2.2 Proposition

The relation $\prec_{fuz}$ is a preorder relation on L.

Proof The relation $\prec_{fuz}$ is reflexive:

$$\left.\begin{array}{c} x \wedge x' \leq x \wedge x' \\ \& \\ x \vee x' \geq x \vee x' \end{array}\right\} \Longleftrightarrow x \prec_{fuz} x,\ \forall x \in L$$

It is also transitive:

$$\left.\begin{array}{l} x \prec_{fuz} y \Rightarrow x \wedge x' \leq y \wedge y' \ \& \ x \vee x' \geq y \vee y' \\ \\ y \prec_{fuz} z \Rightarrow y \wedge y' \leq z \wedge z' \ \& \ y \vee y' \geq z \vee z' \end{array}\right\} \Longrightarrow$$

$$x \prec_{fuz} z$$

Therefore, the relation $\prec_{fuz}$ is a preorder relation. □

Remark. The previous relation is not an order relation because it is not an antisymetrical relation:

$$\left.\begin{array}{l} x \prec_{fuz} y \Rightarrow x \wedge x' \leq y \wedge y' \ \& \ x \vee x' \geq y \vee y' \\ \\ y \prec_{fuz} x \Rightarrow y \wedge y' \leq x \wedge x' \ \& \ y \vee y' \geq x \vee x' \end{array}\right\} \Longrightarrow$$

$$\left.\begin{array}{c} x \wedge x' = y \wedge y' \\ \& \\ x \vee x' = y \vee y' \end{array}\right\}$$

and this does not imply that, necessarily, $x = y$.

2.3 Proposition

If $L = [0, 1]$, it is proved in [6] that

$$x \prec_{fuz} y \Longleftrightarrow x \leq^s y \text{ or } x \leq^s y'$$

where $\leq^s$ is the sharpened order relation used in the Proposition 1.7.

Remark. It is immediate to prove that

1. The element 0.5 is the maximum element for this relation.
$$x \prec_{fuz} 0.5, \quad \forall x \in L$$
2. The elements 0 and 1 are the minimal elements.
$$\forall x \in L, \quad 0 \prec_{fuz} x, \ 1 \prec_{fuz} x$$
3. The following equivalences are verified
$$x \prec_{fuz} y \Longleftrightarrow x \prec_{fuz} y' \Longleftrightarrow x' \prec_{fuz} y' \Longleftrightarrow x' \prec_{fuz} y$$

In order to characterize the relation *'to be at least as fuzzy as'* in the set of closed intervals in $L = [0, 1]$, we define on $\mathcal{J}[L]$ the following order relation.

2.4 Definition

Given two intervals $[a, b], [c, d] \in \mathcal{J}[L]$,

$$[a, b] \leq_{[0.5,0.5]} [c, d] \Longleftrightarrow$$

$$\Longleftrightarrow \left\{\begin{array}{c} [a, b] \wedge [0.5, 0.5] \leq [c, d] \wedge [0.5, 0.5] \\ \& \\ [a, b] \vee [0.5, 0.5] \geq [c, d] \vee [0.5, 0.5] \end{array}\right.$$

$$\Longleftrightarrow \left\{\begin{array}{ll} [a, b] \leq [c, d] & \text{if} \quad [a, b], [c, d] \leq [0.5, 0.5] \\ a \leq c \ \& \ b \geq d & \text{if} \quad a, c \leq 0.5 \ \& \ b, d \geq 0.5 \\ [a, b] \geq [c, d] & \text{if} \quad [a, b], [c, d] \geq [0.5, 0.5] \end{array}\right.$$

2.5 Proposition

The order relation $\leq_{[0.5,0.5]}$ is an extension to the set of closed intervals in L of the 'sharpened order' $\leq^s$ defined on L.

Proof The order relation $\leq_{[0.5,0.5]}$ is the dual order of the ω_0-activity order $\preceq_{[0.5,0.5]}$ defined on set of intervals $\mathcal{J}[L]$ as

$$[a, b] \preceq_{[0.5,0.5]} [c, d] \Longleftrightarrow$$

$$\left\{\begin{array}{c} [a, b] \wedge [0.5, 0.5] \geq [c, d] \wedge [0.5, 0.5] \\ \& \\ [a, b] \vee [0.5, 0.5] \leq [c, d] \vee [0.5, 0.5] \end{array}\right.$$

As we proved in the Proposition 1.8, the ω_0-activity order between intervals $\preceq_{[0.5,0.5]}$ is an extension of the order $\preceq_{0.5}$

$$([a, b] \preceq_{[0.5,0.5]} [c, d]) \Longleftrightarrow (a \preceq_{0.5} c) \& (b \preceq_{0.5} d)$$

Considering the dual orders and taking into account that $\preceq_{0.5}$ is the dual order of the 'sharpened order' $\leq^s$ defined on $L = [0, 1]$, we obtain:

$$([a, b] \leq_{[0.5,0.5]} [c, d]) \Longleftrightarrow (a \leq^s c) \& (b \leq^s d)$$

That is, the order relation $\leq_{[0.5,0.5]}$ is an extension to the set of closed intervals in L of the 'sharpened order' $\leq^s$. □

2.6 Example

In Figure 3 we have represented the semilattice $(\mathcal{J}[\mathcal{C}], \leq_{[0.5,0,.5]})$ of the closed intervals in the chain $\mathcal{C} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.

The extension to the set of closed intervals in L of the order relation $\leq^s$ allows us to extend the relation *'to be at least as fuzzy as'* in the following way:

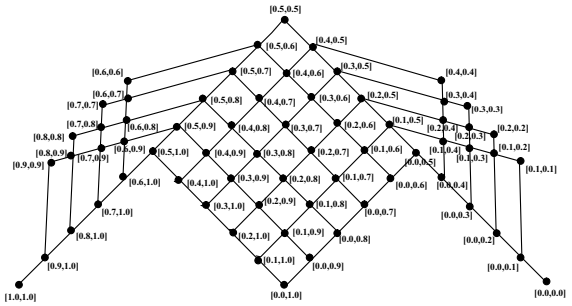


Figure 3: Semilattice $(\mathcal{J}[\mathcal{C}], \leq_{[0.5,0.5]})$

2.7 Definition

Given two intervals $[a,b],[c,d] \in \mathcal{J}[L]$, we will say that the interval $[c,d]$ is *at least as fuzzy as* the interval $[a,b]$, and we will denote it by:

$$[a,b] \prec_{fuz} [c,d] \Longleftrightarrow \begin{cases} [a,b] \leq_{[0.5,0.5]} [c,d] \\ \text{or} \\ [a,b] \leq_{[0.5,0.5]} [1-d,1-c] \end{cases}$$

that is:

$$[a,b] \prec_{fuz} [c,d] \Longleftrightarrow \begin{cases} [a,b] \leq [c,d] \leq [0.5,0.5] \\ \text{or} \\ a \leq c \leq 0.5 \ \& \ b \geq d \geq 0.5 \\ \text{or} \\ [a,b] \geq [c,d] \geq [0.5,0.5] \\ \text{or} \\ [a,b] \leq [1-d,1-c] \leq [0.5,0.5] \\ \text{or} \\ a \leq 1-d \leq 0.5 \ \& \ b \geq 1-c \geq 0.5 \\ \text{or} \\ [a,b] \geq [1-d,1-c] \geq [0.5,0.5] \end{cases}$$

Remarks.

1. The interval $[0.5,0.5]$ is at least as fuzzy as any other element of the set $\mathcal{J}[L]$.

   $$[a,b] \prec_{fuz} [0.5,0.5], \quad \forall [a,b] \in \mathcal{J}[L]$$

   Therefore, it will be the interval that represents the greatest fuzziness.

2. The intervals $[0,0]$, $[0,1]$ and $[1,1]$ are the minimal elements. However, it is not verified that any element is at least as fuzzy as the minimal elements, since they can be not comparable elements. For example, $[0,0] \not\prec_{fuz} [0,0.8]$ and $[0,0.8] \not\prec_{fuz} [0,0]$.

In next section we will try to do a similar study in order to be able to compare the ambiguity of the closed intervals in L.

3 Preorder relation associated to the ambiguity of the intervals

Intuitively, we associate the amplitude of the intervals to the ambiguity. The interval $[0,1]$ is the one that has the biggest amplitude and so, following this idea, it will be the one that would represent the greatest ambiguity. For this reason, we will do a development similar to the one done in the previous section to compare the fuzziness between too intervals using an order relation for which the interval $[0,1]$ is the maximum element.

3.1 Definition

Given two intervals $[a,b],[c,d] \in \mathcal{J}[L]$, the order relation $\leq_{[0,1]}$ is defined as:

$$[a,b] \leq_{[0,1]} [c,d] \Longleftrightarrow \begin{cases} [a,b] \wedge [0,1] \leq [c,d] \wedge [0,1] \\ \& \\ [a,b] \vee [0,1] \geq [c,d] \vee [0,1] \end{cases} \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} [0,b] \leq [0,d] \\ \& \\ [a,1] \geq [c,1] \end{cases} \Longleftrightarrow \begin{cases} b \leq d \\ \& \\ a \geq c \end{cases}$$

The order relation $\leq_{[0,1]}$ agrees with the inclusion $\subseteq$ between intervals:

$$[a,b] \leq_{[0,1]} [c,d] \Longleftrightarrow [a,b] \subseteq [c,d]$$

Remark. The order relation $\leq_{[0,1]}$ is the order relation dual of the ω_0-activity order $\preceq_{[0,1]}$, which, as we saw before, is the order of inclusion $\supseteq$ between intervals.

3.2 Example

Let $(\mathcal{J}[\mathcal{C}], \leq)$ be the lattice of closed intervals in the chain $\mathcal{C} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, then the semilattice $(\mathcal{J}[\mathcal{C}], \leq_{[0,1]})$ can be represented by the graphic of Figure 4.

Using this order relation defined on the set of intervals, we will define the relation *'to be at least as ambiguous as'*.

3.3 Definition

Given two intervals $[a,b],[c,d] \in \mathcal{J}[L]$, we will say that the interval $[c,d]$ is *at least as ambiguous as* the inter-

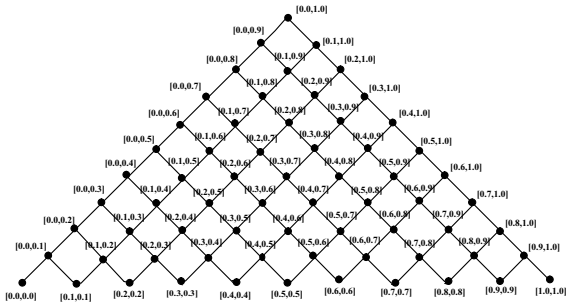


Figure 4: Semilattice $(\mathcal{J}[\mathcal{C}], \leq_{[0,1]})$

val $[a, b]$ and we will denote it by:

$$[a,b] \prec_{amb} [c,d] \Longleftrightarrow \begin{cases} [a,b] \leq_{[0,1]} [c,d] \\ \text{or} \\ [a,b] \leq_{[0,1]} [1-d, 1-c] \end{cases}$$

or what is the same:

$$[a,b] \prec_{amb} [c,d] \Longleftrightarrow \begin{cases} [a,b] \subseteq [c,d] \\ \text{or} \\ [a,b] \subseteq [1-d, 1-c] \end{cases}$$

Remarks.

1. The interval $[0, 1]$ is at least as ambiguous as any other element of the set $\mathcal{J}[L]$.

   $$[a,b] \prec_{amb} [0,1], \quad \forall [a,b] \in \mathcal{J}[L]$$

   This interval represents the greatest ambiguity.

2. The intervals of the type $[a, a]$ are the minimal elements and, therefore, there are the less ambiguous of the set of intervals. We will consider that the intervals of the type $[a, a]$ have the smallest ambiguity.

3. As it happened with the previous relation, also in this case we can find some elements that are not comparable. For example, $[0.2, 0.3] \not\prec_{amb} [0.5, 0.7]$ and $[0.5, 0.7] \not\prec_{amb} [0.2, 0.3]$.

4. It is immediate to prove that any pair of intervals verifies that:

   $[a, b] \prec_{amb} [c, d] \Leftrightarrow [a, b] \prec_{amb} [1-d, 1-c] \Leftrightarrow [1-b, 1-a] \prec_{amb} [1-d, 1-c] \Leftrightarrow [1-b, 1-a] \prec_{amb} [c, d]$

4 Future work

We are interested in being able to compare the ambiguity and the fuzziness of any pair of intervals and, hence, of an interval-valued L-Fuzzy set in order to work with the interval-valued L-Fuzzy contexts [4, 1]. However, that is not possible if we only use the preorder relations, as we remarked before. So, we will try to define an ambiguity and fuzziness degree of a closed interval in $[0, 1]$. Also, in future works, we will use these degrees in the interval-valued contexts area [4, 1] in order to mesure the ambiguity and fuzziness of an L-Fuzzy concept.

References

[1] C. Alcalde, A. Burusco and R. Fuentes-González: Treatment of the incomplete information in L-Fuzzy contexts. *EUSFLAT-LFA 2005.* Barcelona (2005), pp.518-523.

[2] C. Alsina and E. Trillas: Sur les measures du degré de flou. *Stochastica.* **3** (1979), 81-84.

[3] P. Burillo and H. Bustince: Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets and Systems.* **78** (1996), 305-316.

[4] A. Burusco and R. Fuentes-González: The study of the interval-valued contexts. *Fuzzy Sets and Systems.* **121** (2001), pp.69-82.

[5] A. De Luca and S. Termini: A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf. and Control.* **20** (1972), 301-312.

[6] A. De Luca and S. Termini: On some algebraic aspects of the Measures of Fuzzines. *Fuzzy Information and Decision Processes.* North-Holland Publishing Company (1982), 17-24.

[7] D. Dubois, W. Ostasiewicz and H. Prade: Fuzzy sets: History and basic notions. *In: D. Dubois, H. Prade (Eds),Fundamentals of Fuzzy Sets.* Kluwer, Boston, MA (2000), 21-124.

[8] H.J.A.M.Heijmans, R. Keshet: Inf-semilattice approach to self-dual morphology. *Journal of Mathematical Imaging and Vision.* **17** (2002), 55-80.

[9] F. Meyer and J. Serra: Contrast and activity lattice, *Signal process.* **16** (1989), 303-317.

[10] J. Serra: *Image Analysis and Mathematical Morphology.* Academic Press, vol. I (fourth printing 1990) and vol. II (second printing 1992).

[11] P. Soille: *Morphological Image Analysis. Principles and Applications.* Second Edition. Springer (2004).

[12] E. Trillas and T. Riera: Entropies in finite fuzzy sets. *Information Sciences.* **15** (1978), pp.159-168.

[13] E. Trillas: On the use of words and fuzzy sets. *Information Sciences.* **176** (2006), pp.1463-1487.

Extended distances between fuzzy points

J.Casasnovas
Dept. of Mathematics and Computer Science
University of the Balearic Islands
07122 Palma de Mallorca (Spain)
jaume.casasnovas@uib.es

J.Vicente Riera
Dept. of Mathematics and Computer Science
University of the Balearic Islands
07122 Palma de Mallorca (Spain)
jvicente.riera@uib.es

Abstract

In this paper, we tackle the extension of aggregations of distances in $\mathbb{R}^k$ to the fuzzy numbers and fuzzy points and we investigate the fullfilment of the conditions of a fuzzy metric.

Keywords: Distance, aggregation, fuzzy point, fuzzy number.

1 Introduction

Several distances are considered in [2, 3, 4, 6, 7, 9, 10, 12] and [5, 8] for the fuzzy sets. The fuzzy sets with finite support with N elements can be considered as points in Kosko's N-dimensional hypercube $[0,1]^N$ [7, 13], and a distance between them can be defined as a suitable aggregation of the values $|\mu(x)-\nu(x)|$, for each element x in the support. For instance, the well known euclidean distance: $\sqrt{\Sigma_{x\in S}|\mu(x)-\nu(x)|^2}$, the Hamming's distance: $\Sigma_{x\in S}|\mu(x)-\nu(x)|$, the Minkowski distances: $\sqrt[p]{\Sigma_{x\in S}|\mu(x)-\nu(x)|^p}$, even the maximum distance: $\bigvee_{x\in S}|\mu(x)-\nu(x)|$, where $\bigvee$ denotes the "supremum".

Fuzzy numbers are fuzzy subsets of $\mathbb{R}$ with continuous support. In this case, the additions can be replaced by integrations [5, 8] and, so, to define a distance between fuzzy numbers with values in $\mathbb{R}$. Nevertheless, in this paper we consider that, in a parallel way to the distance between numbers of $\mathbb{R}$ or points in $\mathbb{R}^2$ as a real number, the distance between fuzzy numbers or fuzzy points must be described by means of a fuzzy number. Namely, fuzzy numbers are fuzzy subsets of $\mathbb{R}$, where the euclidean distance: $|x-y|$ is known as well as the fuzzy points are fuzzy subsets of $\mathbb{R}^2$, where an aggregation of two euclidean pseudodistances is previously defined as euclidean, Hamming, maximum or Minkowski distances. Fuzzy points and fuzzy metrics have been introduced in [2] and [9] and the aggregation of distances in [11].

In order to reach the extension to the fuzzy numbers or the fuzzy points of a distance, an operation or even an aggregation of them, we work with the *extension principle* or with the equivalent α-cuts calculus [6] to define operations, distances and even aggregation of them for fuzzy numbers or fuzzy points.

In particular, if the fuzzy points have coordinates, i.e., they are defined from two fuzzy numbers, we can define fuzzy distances through two ways. On one hand, an aggregation of distances in $\mathbb{R}^2$ can be extended directly to the fuzzy points. On the other hand, the distance in $\mathbb{R}$ can be extended for those coordinates and the distance between fuzzy points can be constructed by aggregation of them. We investigate the relationship between the two methods of definition.

2 Preliminaries

In order to define distances between finite fuzzy sets, we consider a finite, n-element set

$$U=\{u_1,\ldots,u_n\},$$

and let $\mathcal{FP}(U)$ denote the set of its $[0,1]$-valued fuzzy subsets.

The mapping sending every $\mu\in\mathcal{FP}(U)$ to the vector $(\mu(u_1),\ldots,\mu(u_n))\in[0,1]^n$ is a bijection $\mathcal{FP}(U)\cong[0,1]^n$, and it allows to identify, in a one-to-one way, every fuzzy subset of U with a point of Kosko's n-dimensional hypercube [7]. Hypercubical calculus has been described in [13].

This identification allows the translation of operations and distances on $[0,1]^n$ into operations and distances on $\mathcal{FP}(U)$. So, for instance, if for any couple, μ,ν of fuzzy sets and each $i=1,\cdots,n$ we consider the *euclidean distance*: $d(\mu(u_i),\nu(u_i))=|\mu(u_i)-\nu(u_i)|$, then $d_i(\mu,\nu)=|\mu(u_i)-\nu(u_i)|$ is *pseudodistance* in $\mathcal{FP}(U)$ [3, 4].
Let $\mathbb{R}$ be the set of real numbers and let $\mathbb{R}^+$ be the subset of positive real numbers.

Definition 1 *[11] We will call* aggregation of distances *to a non-decreasing* $F : (\mathbb{R}^+)^k \to [0, +\infty]$*, such that:*

i) $F(0, \ldots, 0) = 0$.

ii) F is subadditive: if $c_i \leq a_i + b_i$ *for every* $i = 1, \ldots, k$*, then*

$$F(c_1, \ldots, c_k) \leq F(a_1, \ldots, a_k) + F(b_1, \ldots, b_k).$$

iii) $F(c_1, \ldots, c_k) = 0$ *implies* $c_1 = \cdots = c_k = 0$.

Proposition 1 *[11] If* $d_1, \ldots, d_k$ *are distances in* $\mathbb{R}$ *(hence they are pseudodistances in* $\mathbb{R}^k$*) and* $F : (\mathbb{R}^+)^k \to [0, +\infty]$ *is an aggregation of distances, then the mapping:*

$$F(d_1, \ldots, d_k) : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$$

such that: $F(d_1, \ldots, d_k)((x_1, \ldots, x_k), (y_1, \ldots, y_k)) = F(d_1(x_1, y_1), \ldots, d_k(x_k, y_k))$ *is a distance in* $\mathbb{R}^k$.

Example 1 *Let* $\omega = (\omega_1, \ldots, \omega_k) \in (\mathbb{R}^+)^k$ *be any vector of positive weights. It is easy to check that conditions (i) to (iii) of the definition 1 are satisfied, among others, by the following mappings:*

- $F_{2,\omega}(c_1, \ldots, c_k) = \sqrt{\sum_{i=1}^k \omega_i c_i^2}$,
- $F_{H,\omega}(c_1, \ldots, c_k) = \sum_{i=1}^k \omega_i c_i$
- $F_{\infty,\omega}(c_1, \ldots, c_k) = \bigvee_{i=1}^k \omega_i c_i$,
- $F_L(c_1, \ldots, c_k) = \min(\sum_{i=1}^k \omega_i c_i, 1)$

Definition 2 *[2, 12] A fuzzy set* $A : \mathbb{R} \to [0, 1]$ *is a fuzzy number if:*

- *A is upper semi-continous*
- $A(r) = 0$ *outside some interval* $[c, d]$
- *There are real numbers* $c \leq a \leq b \leq d$ *such that A(r) is increasing on interval* $[c, a]$*, A(r) decreasing on* $[b, d]$ *and* $A(r) = 1$ *on* $[a, b]$

Remark 1 *[6]*

- *Given a fuzzy subset* $B : U \to [0, 1]$ *of an ordinary set U and any number* $\alpha \in [0, 1]$*, the* alpha-*cut (*α*-cut) is the crisp set* $B^\alpha = \{x \in U;$ *such that* $B(x) \geq \alpha\}$.
- *In particular, given a fuzzy number A and any number* $\alpha \in [0, 1]$ *the* alpha-*cut,* A^α*, is the crisp set* $A^\alpha = \{x \in \mathbb{R}$ *such that* $A(x) \geq \alpha\}$.
- *The* α*-cuts of a fuzzy number are closed intevals.*
- *We will denote the* α*-cut of a fuzzy number A ,* A^α*, by* $[a_1^\alpha, a_2^\alpha]$*, with* $a_1^\alpha, a_2^\alpha \in \mathbb{R}$ *and* $a_1^\alpha \leq a_2^\alpha$.

Definition 3 *[2] Let A , B be two fuzzy numbers and let* $A^\alpha = [a_1^\alpha, a_2^\alpha]$*,* $B^\alpha = [b_1^\alpha, b_2^\alpha]$ *be their* α*-cut respectively for all* $\alpha \in [0, 1]$*. We can define:*

- Strong order*:*
  $A \leq_s B$ *if and only if* $a_1^\alpha \leq b_1^\alpha$ *and* $a_2^\alpha \leq b_2^\alpha$ *for all* $\alpha \in [0, 1]$
  This relationship is a partial order (reflexive, transitive, antisymmetric)
- Weak order*:*
  $A \leq_w B$ *if and only if* $a_2^\alpha \leq b_2^\alpha$ *for all* $\alpha \in [0, 1]$
  This relationship is reflexive and transitive.

Definition 4 *[2] A fuzzy set* $\tilde{P} : \mathbb{R}^2 \to [0, 1]$ *is a fuzzy point at (a,b) if:*

- $\tilde{P}$ *is upper semi-continous*
- $\tilde{P}(x, y) = 1$ *if and only if* $(x, y) = (a, b)$
- *The* α*-cut of the fuzzy set* $\tilde{P}$*,* $\tilde{P}^\alpha$*, is a compact, convex, subset of* $\mathbb{R}^2$*, for all* $\alpha \in [0, 1]$

Example 2 *[2] If A and B are fuzzy numbers so that* $A(x) = 1$ *if and only if* $x = a$ *and* $B(y) = 1$ *if and only if* $y = b$*, respectively, then* $\tilde{P}(x, y) = A(x) \wedge B(y)$ *is a fuzzy point at (a,b).*

Example 3 *We can consider a fuzzy point in* $(a, b) \in \mathbb{R}^2$*, such that its shape is a right circular cone with base* $(x - a)^2 + (y - b)^2 \leq 1/4$

Remark 2 *[6] If U and V are sets and* $O : U \times \cdots \times U \to V$ *is a mapping, the "extension principle" allows to define an* extended *mapping*

$$\tilde{O} : \mathcal{FP}(U) \times \cdots \times \mathcal{FP}(U) \to \mathcal{FP}(V)$$

as follows:

$$\tilde{O}(A_1, \ldots, A_s)(v) = \sup\{\min(A_1(u_1), \ldots, A_s(u_s)) \mid O(u_1, \ldots, u_s) = v\} \quad (1)$$

For instance, O can be an algebraic operation, a t-norm, a t-conorm, but also a distance or an aggregation.

Remark 3 *[6]*

The fuzzy set $\tilde{O}(A_1, \ldots, A_s)$ *can also be obtained through the* α*-cuts or* α*-levels. Namely, if* A_i *is a fuzzy*

set defined on U, let A_i^α be the α-cut, according the remark 1, and let's consider the crisp set:

$$O(A_1^\alpha, \ldots, A_s^\alpha) = \{O(x_1, \ldots, x_s) \mid x_i \in A_i^\alpha\} \qquad (2)$$

then $(\tilde{O}(A_1, \ldots, A_s))^\alpha = O(A_1^\alpha, \ldots, A_s^\alpha)$ and we have:

$$\tilde{O}(A_1, \ldots, A_s)(v) = \sup\{\alpha \mid v \in (\tilde{O}(A_1, \ldots, A_s))^\alpha = O(A_1^\alpha, \ldots, A_s^\alpha)\}$$

Remark 4 *The operations on $\mathbb{R}$: addition, multiplication, maximum and minimum can be extended to operations between fuzzy numbers.*

If A and B are fuzzy numbers with α-cuts $A^\alpha = [a_1^\alpha, a_2^\alpha]; B^\alpha = [b_1^\alpha, b_2^\alpha]$, then

$$(A \oplus B)(z) = \sup\{A(x) \wedge B(y) | x + y = z\}$$

$$(A \otimes B)(z) = \sup\{A(x) \wedge B(y) | xy = z\}$$

$$(A \bigvee B)(z) = \sup\{A(x) \wedge B(y) | x \vee y = z\}$$

$$(A \bigwedge B)(z) = \sup\{A(x) \wedge B(y) | x \wedge y = z\}$$

are fuzzy numbers and their α-cuts are:

$$(A \oplus B)^\alpha = [a_1^\alpha + b_1^\alpha, a_2^\alpha + b_2^\alpha]$$

$$(A \otimes B)^\alpha = [min\{a_1^\alpha b_1^\alpha, a_1^\alpha b_2^\alpha, a_2^\alpha b_1^\alpha, a_2^\alpha b_2^\alpha\}, max\{a_1^\alpha b_1^\alpha, a_1^\alpha b_2^\alpha, a_2^\alpha b_1^\alpha, a_2^\alpha b_2^\alpha\}]$$

$$(A \bigvee B)^\alpha = [a_1 \vee b_1, a_2 \vee b_2]$$

$$(A \bigwedge B)^\alpha = [a_1 \wedge b_1, a_2 \wedge b_2],$$

respectively.

3 Fuzzy Distances

Definition 5 *Let FN be the set of fuzzy numbers according to the definition 2 and $A, B \in FN$. A mapping $D : FN \times FN \to FN$ is a fuzzy distance between fuzzy numbers if the following properties hold:*

- $D(A, B) \geq \tilde{0}$
- $D(A, B) = D(B, A), \forall\ A, B \in FN$
- $D(A, C) \leq D(A, B) \oplus D(B, C), \forall\ A, B, C \in FN$
- $D(A, B) = \tilde{0}$ *if and only if* $A = B$

where $\oplus$ is the extended addition for fuzzy numbers and $\tilde{0}$ is any fuzzy number such that: $\tilde{0}(x) = 0$, for $x < 0$; $\tilde{0}(0) = 1$, $\tilde{0}(x)$ is decreasing for $0 < x < d$ for some $d > 0$ and $\tilde{0}(x) = 0$, for $x \geq d$

Moreover, $\leq$ is the binary relationship between fuzzy numbers, which we have called either weak order $\leq_w$ or the strong order $\leq_s$ defined in definition 3. If we use $\leq_s$ ($\leq_w$) we will say that D is strong (weak) fuzzy distance.

Is easy to check that if D is a fuzzy number such that $x > 0, \forall x \in SuppA$, then $\tilde{0} \leq_w D$ is equivalent to $\tilde{0} \leq_s D$

Remark 5 *If we only can assure: "as long as $A = B$, then $D(A, B) = 0$" , then D is called a fuzzy pseudodistance.*

Definition 6 *Let FP be the set of fuzzy points according to the definition 4. A fuzzy distance D between fuzzy points is a mapping $D : FP \times FP \to FN$ such that the following properties hold:*

- $D(\tilde{P}, \tilde{Q}) \geq \tilde{0}$
- $D(\tilde{P}, \tilde{Q}) = D(\tilde{Q}, \tilde{P}), \forall\ \tilde{P}, \tilde{Q} \in FP$
- $D(\tilde{P}, \tilde{Q}) = \tilde{0}$ *if and only if* $\tilde{P}$ *and* $\tilde{Q}$ *are both fuzzy points at* (a, b)
- $D(\tilde{P}, \tilde{Q}) \leq_w D(\tilde{P}, \tilde{R}) \oplus D(\tilde{R}, \tilde{Q}), \forall\ \tilde{P}, \tilde{Q}, \tilde{R} \in FP$

where $\oplus$ is the extended addition for fuzzy numbers and $\tilde{0}$ is any fuzzy number such that: $\tilde{0}(x) = 0$, for $x < 0$; $\tilde{0}(0) = 1$, $\tilde{0}(x)$ is decreasing for $0 < x < d$ for some $d > 0$ and $\tilde{0}(x) = 0$, for $x \geq d$

Moreover, $\leq_w$ is the binary relationship between fuzzy numbers: $\tilde{A} \leq_w \tilde{B}$, which is called weak order (see the definition 3).

Analogously to the case of fuzzy numbers, if D is a fuzzy number such that $x > 0, \forall x \in SuppA$, then $\tilde{0} \leq_w D$ is equivalent to $\tilde{0} \leq_s D$

4 Extended Distances

4.1 Extended Distances between fuzzy sets

If $d : U \times U \to \mathbb{R}$ is a distance defined in the universe U, it is possible to define an "extension of the distance" between fuzzy subsets of that universe by the "extension principle":

So, the extension of the distance, d, between two fuzzy sets, A, B, is a fuzzy subset of $\mathbb{R}$ defined as follows:

$$\tilde{d}(A, B)(r) = \sup\{A(x) \wedge B(y) \mid d(x, y) = r\}$$

if there exists $(x, y) \in U \times U$ such that $d(x, y) = r$ and $\tilde{d}(A, B)(r) = 0$, otherwise.

Notice that if $\tilde{d}(A, B)(r) > 0$, then there exist $(x, y) \in U \times U$ such that $d(x, y) = r$, therefore $r \geq 0$.

4.2 Extended Distances between fuzzy numbers

Definition 7 *Let FN the set of fuzzy numbers according to the definition 2 and $A, B \in FN$. Let $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the usual distance $d(x, y) = |x - y|$. We can extend:*

$$\tilde{d}(A, B)(z) = \sup\{A(x) \wedge B(y) || x - y| = z\}$$

Proposition 2 *For each couple of fuzzy numbers A, B the extension of the usual distance to the fuzzy numbers $\tilde{d}(A, B)$ yields a fuzzy number.*

Proof

According to the remark 3, the extended distance can be calculated through the α-cuts:

$$d(A^\alpha, B^\alpha) = \{r \in \mathbb{R} \mid \exists x \in A^\alpha; y \in B^\alpha, d(x, y) = r\}$$

and

$$\tilde{d}(A, B)(z) = \sup\{\alpha | z \in d(A^\alpha, B^\alpha)\}$$

then: $\tilde{d}(A, B)^\alpha = d(A^\alpha, B^\alpha)$.

We now argue that $\tilde{d}(A, B)$ is a fuzzy number (see definition 2).

a) Since the α-cuts of A and B are compact it is easily seen that $d(A^\alpha, B^\alpha)$ is a bounded closed interval for all α. Let $d(A^\alpha, B^\alpha) = [l(\alpha), r(\alpha)]$. It is known that [1], a mapping $f : \mathbb{R}^m \to \mathbb{R}$ is upper semi-continuous if and only if the set $\{x \mid f(x) \geq t\}$ is closed for all $t \in \mathbb{R}$.By reason of this result, as $\tilde{d}(A, B)^\alpha = d(A^\alpha, B^\alpha)$ is a closed interval for all α, hence $\tilde{d}(A, B)$ is upper semi-continuous.

b) Let $d(A^0, B^0) = [c, d]$. Then $[c, d] \subset [0, +\infty]$ and $\tilde{d}(A, B)(z) = 0$ outside of $[c, d]$

c) Let $d(A^1, B^1) = [a, b]$. Now, since $\tilde{d}(A, B)^\alpha = d(A^\alpha, B^\alpha) = [l(\alpha), r(\alpha)]$ for all α with $l(\alpha)$ increasing from c to a and $r(\alpha)$ decreasing from b to d, we obtain $\tilde{d}(A, B)$ is increasing on $[c, a]$ and decreasing on $[b, d]$.

Proposition 3 *The extension of the usual distance of the fuzzy numbers is not a fuzzy distance between fuzzy numbers according to the definition 5, because the fourth condition is not satisfied.*

Proof

- *$\tilde{d}(A, B) \geq \tilde{0}$, because $\tilde{d}(A, B)$ is fuzzy number and its support only contains positive or zero elements of $\mathbb{R}$. So, if $\tilde{d}(A, B) \neq \tilde{0}$ we can define a fuzzy number, $\tilde{0}$, such that $\tilde{d}(A, B) \geq \tilde{0}$*
- *Clearly, $\tilde{d}(A, B) = \tilde{d}(B, A)$*
- *Let $A^\alpha = [a_1^\alpha, a_2^\alpha]$; $B^\alpha = [b_1^\alpha, b_2^\alpha]$; $C^\alpha = [c_1^\alpha, c_2^\alpha]$. We only need to prove that: $a_2^\alpha \leq b_2^\alpha + c_2^\alpha$, for all $\alpha \in [0, 1)$*

$$a_2^\alpha = \sup\{d(x, y) | x \in A^\alpha, y \in B^\alpha\} \leq$$

$$\sup\{d(x, z) + d(z, y) | x \in A^\alpha, y \in B^\alpha, z \in C^\alpha\} \leq$$

$$\sup\{d(x, z) | x \in A^\alpha, z \in C^\alpha\} +$$

$$\sup\{d(z, y) | z \in C^\alpha, y \in B^\alpha\} \leq$$

$$b_2^\alpha + c_2^\alpha$$

- *If $\tilde{d}(A, B) = \tilde{0}$, then $0 \in \tilde{d}(A, B)^1 = d(A^1, B^1)$, but in general $A \neq B$. For instance, if $A^1 = B^1$ and $a_2^\alpha < b_2^\alpha, \forall \alpha \in [0, 1]$.*

Remark 6 *If A, B are triangular fuzzy numbers with $A^1 = a; B^1 = b$ and $\tilde{d}(A, B) = \tilde{0}$, then $a = b$*

Remark 7 *The following example shows that the extension of the euclidean distance is not a strong distance between fuzzy numbers, that is, a fuzzy distance between fuzzy numbers such that the order considered is the* the strong order *defined in 3:*

Let $A^0 = [1, 2]$; $B^0 = [3/2, 7/2]$; $C^0 = [3, 4]$; be the supports of A, B, C, respectively. Then:

$$\tilde{d}(A, B)^0 = [0, 5/2]; \ \tilde{d}(B, C)^0 = [0, 5/2];$$

$\tilde{d}(A, C)^0 = [1, 3]$ *hence:*

$$\tilde{d}(A, C)^0 \not\leq \tilde{d}(A, B)^0 + \tilde{d}(B, C)^0$$

because $3 \leq 5/2 + 5/2$, but $1 \leq 0 + 0$ is not true.

4.3 Extended Distances between fuzzy points

Definition 8 *Let d be a distance defined in $\mathbb{R}^2$ the "extension of d to the fuzzy points" is the mapping from couples of fuzzy points to the fuzzy subsets of $\mathbb{R}$:*

$$\tilde{d}(\tilde{P}, \tilde{Q})(r) =$$

$$\bigvee\{\tilde{P}(x_1, y_1) \wedge \tilde{Q}(x_2, y_2) \mid d((x_1, y_1), (x_2, y_2)) = r\}$$

Proposition 4 *For each couple of fuzzy points $\tilde{P}, \tilde{Q}$ the extension of the usual distance to the fuzzy points $\tilde{d}(\tilde{P}, \tilde{Q})$ yields a fuzzy number.*

Proof

According to the remark 3, the extended distance can be calculated through the α-cuts:

$$d(\tilde{P}^\alpha, \tilde{Q}^\alpha) = \{r \in \mathbb{R} \mid \exists (x_1, y_1) \in \tilde{P}^\alpha; (x_2, y_2) \in \tilde{Q}^\alpha |$$

$$d((x_1, y_1), (x_2, y_2)) = r\}$$

then

$$\tilde{d}(\tilde{P}, \tilde{Q})(r) = \bigvee \{\alpha | r \in d(\tilde{P}^\alpha, \tilde{Q}^\alpha)\}$$

and

$$\tilde{d}(\tilde{P}, \tilde{Q})^\alpha = d(\tilde{P}^\alpha, \tilde{Q}^\alpha)$$

We now argue that $\tilde{d}(\tilde{P}, \tilde{Q})$ is a fuzzy number (2).

a) Since the α-cuts of $\tilde{P}$ and $\tilde{Q}$ are compact it is easily seen that $d(\tilde{P}^\alpha, \tilde{Q}^\alpha)$ is a bounded closed interval for all α. Let $d(\tilde{P}^\alpha, \tilde{Q}^\alpha) = [l(\alpha), r(\alpha)]$. It is known that [1], a mapping $f : \mathbb{R}^m \to \mathbb{R}$ is upper semi-continuous if and only if the set $\{x \mid f(x) \geq t\}$ is closed for all $t \in \mathbb{R}$.By reason of this result, as $\tilde{d}(\tilde{P}, \tilde{Q})^\alpha = d(\tilde{P}^\alpha, \tilde{Q}^\alpha)$ is a closed interval for all α, hence $\tilde{d}(\tilde{P}, \tilde{Q})$ is upper semi-continuous.

b) Let $d(\tilde{P}^0, \tilde{Q}^0) = [c, d]$. Then $\tilde{d}(\tilde{P}, \tilde{Q})(z) = 0$ outside of $[c, d]$

c) Let $d(\tilde{P}^1, \tilde{Q}^1) = a$, where $a = d((a_1, b_1), (a_2, b_2))$ and $\tilde{P}(a_1, b_1) = 1$ and $\tilde{Q}(a_2, b_2) = 1$. Now, since $\tilde{d}(\tilde{P}, \tilde{Q})^\alpha = d(\tilde{P}^\alpha, \tilde{Q}^\alpha) = [l(\alpha), r(\alpha)]$ for all α with $l(\alpha)$ increasing from c to a and $r(\alpha)$ decreasing from a to d, we obtain $d(\tilde{P}^\alpha, \tilde{Q}^\alpha)$ is increasing on $[c, a]$ and decreasing on $[a, d]$.

Proposition 5 *The extension to the fuzzy points of a distance in $\mathbb{R}^2$ is a fuzzy distance between fuzzy points according to the definition 6.*

Proof The three first conditions of 6 can be proved analogously to the proposition 3.

For the fourth condition and since the remark 6 we have: $\tilde{d}(P, Q) = 0$ if and only if $(a_1, b_1) = (a_2, b_2)$, where $P^1 = (a_1, b_1)$ and $Q^1 = (a_2, b_2)$.

Remark 8 *An analogous example to 7 shows that the extension of a distance in $\mathbb{R}^2$ is not a strong distance [2].*

Proposition 6 *Let $\{d_i\}_{i=1,2}$ two metrics on $\mathbb{R}$ that define pseudometrics on $\mathbb{R}^2$ and let $F(d_1, d_2)$ an aggregation (see Proposition 1) on $\mathbb{R}^2$.*

We can extend each d_i to a fuzzy distance between fuzzy numbers:

$$\tilde{d}_i(A, B)(r) = \bigvee \{A(x) \wedge B(y) | d_i(x, y) = r\}; i = 1, 2$$

We can extend $F(d_1, d_2)$ to a fuzzy distance between fuzzy points:

$$\widetilde{F(d_1, d_2)}(P, Q)(r) =$$

$$\bigvee \{P(x_1, y_1) \wedge Q(x_2, y_2) | F(d_1, d_2)((x_1, y_1), (x_2, y_2)) = r\}$$

Proof *Straightforward.*

Proposition 7 *If $P(x, y) = A(x) \wedge B(y)$; $Q(x, y) = A'(x) \wedge B'(y)$, where A, B, A', B' are fuzzy numbers, we can extend the aggregation of the extended distances $\tilde{d}_i$ between fuzzy numbers to fuzzy points*

Proof

$$\tilde{F}(\tilde{d_1}(A, A'), \tilde{d_2}(B, B'))(r) =$$

$$\bigvee \{\tilde{d_1}(A, A')(r_1) \wedge \tilde{d_2}(B, B')(r_2) | F(r_1, r_2) = r\},$$

where:

$$\tilde{d_1}(A, A')(r_1) =$$

$$\bigvee \{A(x_1) \wedge A'(x_2) | d_1(x_1, x_2) = r_1\}$$

$$\tilde{d_2}(B, B')(r_2) =$$

$$\bigvee \{B(y_1) \wedge B'(y_2) | d_2(y_1, y_2) = r_2\},$$

Therefore:

$$\tilde{F}(\tilde{d_1}(A, A'), \tilde{d_2}(B, B'))(r) =$$

$$\bigvee \{A(x_1) \wedge A'(x_2) \wedge B(y_1) \wedge B'(y_2) |$$

$$F(d_1(x_1, x_2), d_2(y_1, y_2)) = r\}$$

i.e.:

$$\tilde{F}(\tilde{d_1}(A, A'), \tilde{d_2}(B, B'))(r) =$$

$$\bigvee \{A(x_1) \wedge B(y_1) \wedge A'(x_2) \wedge B'(y_2) |$$

$$F(d_1(x_1, x_2), d_2(y_1, y_2)) = r\}$$

hence:

$$\tilde{F}(\tilde{d_1}(A, A'), \tilde{d_2}(B, B'))(r) =$$

$$\bigvee \{P(x_1, y_1) \wedge Q(x_2, y_2) |$$

$$F(d_1, d_2)((x_1, y_1), (x_2, y_2)) = r\} =$$

$$\widetilde{F(d_1, d_2)}(P, Q)(r)$$

5 Conclusion and future work

The distances in $\mathbb{R}$ can be extended in several ways to the fuzzy numbers and, mainly, to the fuzzy points. We have investigated the relationship between two of these ways. As future work, we would investigate other possible ways. For instance:

Let d_1, d_2 be the two pseudodistances

$$d_1((x_1,y_1),(x_2,y_2)) = |x_1 - y_1|$$

$$d_2((x_1,y_1),(x_2,y_2)) = |x_2 - y_2|$$

defined in $\mathbb{R}^2$ and let $d = F(d_1, d_2)$ a distance defined by aggregation of pseudodistances, where F is a mapping included in example 1.

We have,

- On one hand the extended definition of distance for d_i and for d:
  $\tilde{d_1}(\tilde{P},\tilde{Q})(r) =$

  $$\bigvee\{\tilde{P}(x_1,y_1) \wedge \tilde{Q}((x_2,y_2) \mid |x_1 - y_1| = r\}$$

  $\tilde{d_2}(\tilde{P},\tilde{Q})(r) =$

  $$\bigvee\{\tilde{P}(x_1,y_1) \wedge \tilde{Q}((x_2,y_2) \mid |x_2 - y_2| = r\}$$

  $\widetilde{d_F}(\tilde{P},\tilde{Q})(r) = \bigvee\{\tilde{P}(x_1,y_1) \wedge \tilde{Q}((x_2,y_2) \mid d_F((x_1,y_1),(x_2,y_2)) = r\}$
  where

  $$d_F((x_1,y_1),(x_2,y_2)) = F(|x_1 - y_1|, |x_2 - y_2|)$$

- On the other hand the extended definition of the aggregation:

  $$\tilde{F}(\tilde{d_1},\tilde{d_2})(\tilde{P},\tilde{Q}) = \tilde{F}(\tilde{d_1}(\tilde{P},\tilde{Q}), \tilde{d_2}(\tilde{P},\tilde{Q}))$$

  as fuzzy numbers

i.e.

$$(\tilde{d_1} \oplus \tilde{d_2})(\tilde{P},\tilde{Q}) = \tilde{d_1}(\tilde{P},\tilde{Q}) \oplus \tilde{d_2}(\tilde{P},\tilde{Q})$$

where:

$$(\tilde{d_1}(\tilde{P},\tilde{Q}) \oplus \tilde{d_2}(\tilde{P},\tilde{Q}))(r) =$$
$$\bigvee\{\tilde{d_1}(\tilde{P},\tilde{Q})(x) \wedge \tilde{d_2}(\tilde{P},\tilde{Q}))(y) | x + y = r\}$$

or:

$$(\tilde{d_1} \bigvee \tilde{d_2})(\tilde{P},\tilde{Q})) = (\tilde{d_1}(\tilde{P},\tilde{Q}) \bigvee \tilde{d_2}(\tilde{P},\tilde{Q}))$$

Our future work will be the investigation of the relationship between

$$\tilde{d_F}(\tilde{P},\tilde{Q})$$

and

$$\tilde{F}(\tilde{d_1},\tilde{d_2})(P,Q)$$

References

[1] J.J. Buckley and Y. Qu, "Solving systems of linear fuzzy equations". *Fuzzy Sets and Systems*, volume 43,2(1991), pp. 33-43.

[2] J.J. Buckley and E.Eslami, "Fuzzy plane geometry I: Points and lines". *Fuzzy Sets and Systems*, volume 86,2(1997), pp. 179-188.

[3] J. Casasnovas, F. Rosselló, "Midpoints as average representations of pairs of descriptions by means of fuzzy subsets." *Proceedings IPMU04*(2004),pp.2157-2165.

[4] J. Casasnovas, F. Rosselló, "Averaging fuzzy genomes". *Fuzzy Sets and Systems*, volume 152(2005), pp.139-158.

[5] P. Grzegorzewski, "Metrics and orders in space of fuzzy numbers". *Fuzzy sets and Systems*, volume 97(1998), pp.83-94.

[6] Klir, George J., Yuan Bo "Fuzzy sets and fuzzy logic (Theory and applications)", Prentice Hall.

[7] B. Kosko, *Neural networks and fuzzy systems*. Prentice-Hall (1992).

[8] R.Lowen, W.Peeters "Distances between fuzzy sets representing grey level images". *Fuzzy Sets and Systems* 99(1998), pp. 135-149.

[9] J.N. Mordeson. P.S.Nair "Fuzzy Mathematics" Physica-Verlag(1998).

[10] L.Stefanini, L.Sorini, M.L. Guerra "Parametric representation of fuzzy numbers and application to the fuzzy calculus". *Fuzzy Sets and Systems*, volume 100(2006), pp.353-365.

[11] A. Pradera, E. Trillas, E. Castiñeira, On Distances Aggregation. *Proceedings of IPMU2000* (Madrid, July 3-7), pp. 693–700.

[12] W.Voxman, Some remarks on distances between fuzzy numbers. *Fuzzy Sets and Systems*, volume 100(1998), pp.353-365.

[13] M. Zaus, *Crisp and soft computing with hypercubical calculus*. Physica-Verlag (1999).

Conditional entropy for the union of fuzzy and crisp partitions

Doretta Vivona
Dipartimento di Metodi e Modelli Matematici
per le Scienze Applicate
v.A.Scarpa n.16 - 00161 ROMA(Italia)
vivona@dmmm.uniroma1.it

Maria Divari
Dipartimento di Metodi e Modelli Matematici
per le Scienze Applicate
v.A.Scarpa n.16 - 00161 ROMA(Italia)

Abstract

In this paper we introduce the conditional entropy without a fuzzy measure for the union of fuzzy partitions. We recall its properties and we solve the system of functional equations which derives from these conditions, taking into account the *locality principle* and the *independence axiom*.

Keywords: Fuzzy partitions, Entropy, Functional equations.

1 Introduction

In [4] B.Forte has introduced the so called locality property in the crisp setting; later, in [2] Divari and Pandolfi have given some compositive laws for crisp partitions. Later, the authors have used the locality property in the fuzzy setting in [8].

The aim of this paper is to find a class of conditional entropy for the union of fuzzy partitions, by using the locality property and the independence axiom.

Finally, we present the crisp case: the system is analogous.

2 Preliminaires

From algebraic point of view, similar concepts have been introduced in the setting of MV-algebras by Mundici in [5].

Let X be an abstract space and $\mathcal{F}$ a family of fuzzy sets F [9] with membership function $F(x)$, $0 \leq F(x) \leq 1$ for all $x \in X$. Now we recall some definitions which we will use later.

Let F_1 and F_2 be two fuzzy sets, with membership functions $F_1(x)$ and $F_2(x)$. We recall that $(F_1 \cap^+ F_2)(x) = F_1(x) \wedge F_2(x), \forall x \in X$ is the membership function of the intersection set $F_1 \cap^+ F_2$ and $(F_1 \cup^+ F_2)(x) = F_1(x) \vee F_2(x), \forall x \in X$ is the membership function of the union set $F_1 \cup^+ F_2$.

Moreover, we say that two fuzzy sets F_1 and F_2 are disjoint if $F_1 \cap^+ F_2 = \emptyset$. A family $\{F_1, ..., F_n\}$ is a finite partition of a set F, called support, if $F_i \neq \emptyset, F_i$ are disjoint sets and $\Sigma_{i=1}^n F_i(x) = F(x)\ \forall\ x \in X$. We shall indicate with $\mathcal{P}(\mathrm{F})$ the fuzzy partition with support F, with $\{F\}$ the fuzzy set thought as a partition and with $\mathcal{K}$ the family of all partitions of X.

We consider the whole space X as a fuzzy partition, which shall be indicated by $\{X\}$.

In [1] we have recognized that for every $\mathcal{P}(\mathrm{F})$ an algebra $\mathcal{C}_{\mathcal{P}(\mathrm{F})}$ of crisp sets is associated, whose elements are the inverse images of Borel sets of $[0,1]^n$ through the map $F \to [0,1]^n$.

Given two partitions $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}'(\mathrm{F})$ of the same set F :

$$\mathcal{P}(\mathrm{F}) = \{\mathrm{F}_1, ..., \mathrm{F}_\mathrm{i}, ..., \mathrm{F}_\mathrm{n} / \mathrm{F}_\mathrm{i} \cap^+ \mathrm{F}_\mathrm{h} = \emptyset, \mathrm{i} \neq \mathrm{h}, \quad (1)$$

$$\Sigma_{i=1}^n F_i(x) = F(x) \forall\ x \in X\},$$

$$\mathcal{P}'(\mathrm{F}) = \{\mathrm{F}'_1, ..., \mathrm{F}'_\mathrm{j}, ..., \mathrm{F}'_\mathrm{m} / \mathrm{F}'_\mathrm{j} \cap^+ \mathrm{F}'_\mathrm{k} = \emptyset, \mathrm{j} \neq \mathrm{k},$$

$$\Sigma_{j=1}^m F'_j(x) = F(x) \forall\ x \in X\}$$

we say that $\mathcal{P}'(\mathrm{F})$ is a less fine than $\mathcal{P}(\mathrm{F})$ ($\mathcal{P}'(\mathcal{F}) \preceq \mathcal{P}(\mathrm{F})$) if $\mathcal{C}_{\mathcal{P}'(\mathrm{F})} \subset \mathcal{C}_{\mathcal{P}(\mathrm{F})}$.

Later, as in [3], we shall use the operation algebraic joint between two partitions of X. Now, we consider two partitions $\mathcal{P}(\mathrm{F})$ as in (1) and $\mathcal{P}(\mathrm{G})$ in $\mathcal{K}$:

$$\mathcal{P}(\mathrm{G}) = \{\mathrm{G}_1, ..., \mathrm{G}_\mathrm{j}, ..., \mathrm{G}_\mathrm{m} / \mathrm{G}_\mathrm{j} \cap^+ \mathrm{G}_\mathrm{k} = \emptyset, \mathrm{j} \neq \mathrm{k},$$

$$\Sigma_{j=1}^m G_j(x) = G(x) \forall\ x \in X\}.$$

The algebraic joint $\mathcal{P}(\mathcal{F}) \nabla \mathcal{P}(\mathrm{G})$ is

$$\mathcal{P}(\mathrm{F}) \nabla \mathcal{P}(\mathrm{G}) = \{\mathrm{F}_\mathrm{i} \cdot \mathrm{G}_\mathrm{j} / 1 \leq \mathrm{i} \leq \mathrm{n}, \mathrm{i} \leq \mathrm{j} \leq \mathrm{m}\}$$

with lexicographic order, where $(F_i \cdot G_j)(x) = F_i(x) \cdot G_j(x), \forall\ x \in X$. We say that two fuzzy partitions $\mathcal{P}$ and $\mathcal{P}(\mathrm{F})$ are algebraically independent if

$$(F_i \cdot G_j)(x) \neq 0, \forall\ x \in X. \tag{2}$$

Now, we recall the definition of the union of two partitions, given in [6, 8].

Given two partitions $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}(\mathrm{G}) \in \mathcal{K}$, the union $\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})$ is defined by

$$\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G}) = \{\mathrm{F}_1, ..., \mathrm{F_i}, ..., \mathrm{F_n}, \mathrm{G}_1, ..., \mathrm{G_j}, ..., \mathrm{G_m}/$$

$$\forall x \in X,\ F_i(x) \wedge G_j(x) = 0\ (i \neq j),$$

$$\Sigma_{i=1}^{n} F_i(x) + \Sigma_{j=1}^{m} G_j(x) \leq 1\}. \tag{3}$$

In this case, we call $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}(\mathrm{G})$ compositive partitions. Let $\mathcal{H} \subset \mathcal{K}$ be the family of compositive partitions.

As in [7], we introduce the entropy without a fuzzy measure for a fuzzy partition $\mathcal{P}(\mathrm{F}) \in \mathcal{K}$ conditionated by $\mathcal{Q} \in \mathcal{K}$ with $H(\mathcal{Q}) \neq +\infty$ as a function

$$H_{\mathcal{Q}}^{*}(\cdot) : \mathcal{K} \to [0, +\infty]$$

with the following properties:

(I) $\mathcal{P}'(\mathrm{F}) \preceq \mathcal{P}(\mathrm{F}) \Rightarrow H_{\mathcal{Q}}^{*}(\mathcal{P}'(\mathrm{F})) \leq H_{\mathcal{Q}}^{*}(\mathcal{P}(\mathrm{F}))$,

$\mathcal{P}'(\mathrm{F}) \in \mathcal{K}$.

(II) $H_{\mathcal{Q}}^{*}(\mathcal{P}(\mathrm{F})) = H^{*}(\mathcal{P}(\mathrm{F}))$,

if $\mathcal{P}(\mathrm{F})$ is not conditioned by $\mathcal{Q}$.

It follows that : $H_{\mathcal{Q}}^{*}(\{X\}) = 0$ and $H_{\mathcal{Q}}^{*}(\{\emptyset\}) = +\infty$.

(III) $H_{\mathcal{Q}}^{*}(\mathcal{P}(\mathrm{F})\nabla\mathcal{P}(\mathrm{G})) = H_{\mathcal{Q}}^{*}(\mathcal{P}(\mathrm{F})) + H_{\mathcal{Q}}^{*}(\mathcal{P}(\mathrm{G}),$

if $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}(\mathrm{G})$ are algebraically independent and $\mathcal{P}(\mathrm{G}) \in \mathcal{K}$.

3 Conditional entropy for the union of the partitions

If $\mathcal{P}(\mathrm{F}), \mathcal{P}'(\mathrm{F}), \mathcal{P}(\mathrm{G}) \in \mathcal{H}, \mathcal{Q}, \mathcal{R} \in \mathcal{K}$ and $H^{*}(\mathcal{Q}) \neq +\infty$, the conditional entropy $H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right)$ for the union of the fuzzy partitions satisfies the following conditions:

$(\alpha)\ \ 0 \leq H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) \leq +\infty$

as $\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G}) = \{\mathrm{X}\} \to \mathrm{H}_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) = 0$ and $\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G}) = \{\emptyset\} \to \mathrm{H}_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) = +\infty$;

$(\beta)\ \ H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) = H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{G}) \cup^{+} \mathcal{P}(\mathrm{F})\right)$;

$(\gamma)\ \ \mathcal{P}'(\mathrm{F}) \preceq \mathcal{P}(\mathrm{F}) \Rightarrow$

$$H_{\mathcal{Q}}^{*}\left(\mathcal{P}'(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) \leq H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right),$$

$(\delta)\ \ H_{\mathcal{Q}}^{*}\left((\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G}))\nabla\mathcal{R}\right) =$

$$H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) + H^{*}(\mathcal{R}),$$

if $\mathcal{R}$ is algebraically independent from $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}(\mathrm{G})$ and it is not conditioned by $\mathcal{Q}$; the (δ) is a consequence of (II) and (III) seen above.

4 Statement of the problem

Taking into account the locality principle fuzzy partitions [7]:

$$H^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) - H^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \{\mathrm{G}\}\right) =$$

$$H^{*}\left(\{F\} \cup^{+} \mathcal{P}(\mathrm{G})\right) - H^{*}\left(\{F\} \cup^{+} \{G\}\right). \tag{4}$$

we put

$$H_{\mathcal{Q}}^{*}\left(\mathcal{P}(\mathrm{F}) \cup^{+} \mathcal{P}(\mathrm{G})\right) = \Psi\Big[\mathrm{H}^{*}\left((\mathcal{P}(\mathrm{F}) \cup^{+} \{\mathrm{G}\})\nabla\mathcal{Q}\right),$$

$$H^{*}\left((\{F\} \cup^{+} \mathcal{P}(\mathrm{G}))\nabla\mathcal{Q}\right), H^{*}\left((\{F\} \cup^{+} \{G\})\nabla\mathcal{Q}\right), H^{*}(\mathcal{Q})\Big] \tag{5}$$

where $\mathcal{P}(\mathrm{F}), \mathcal{P}'(\mathrm{F}), \mathcal{P}(\mathrm{G}) \in \mathcal{H}, \mathcal{Q}, \mathcal{R} \in \mathcal{K}$, with $H^{*}(\mathcal{Q}) \neq +\infty$ and $\Psi(\cdot) = [0, +\infty] \times [0, +\infty] \times [0, +\infty] \times [0, +\infty] \to [0, +\infty]$.

From $[(\alpha) - (\delta)]$, we have:

$(\mathbf{A})\ \Psi\Big[H^{*}\left(\{X\}\nabla\mathcal{Q}\right), H^{*}\left(\{X\}\nabla\mathcal{Q}\right),$

$$H^{*}\left(\{X\}\nabla\mathcal{Q}\right), H^{*}(\mathcal{Q})\Big] = 0;$$

$(\mathbf{B})\Psi\Big[H^{*}\left((\mathcal{P}(\mathrm{F}) \cup^{+} \{\mathrm{G}\})\nabla\mathcal{Q}\right), H^{*}\left((\{F\} \cup^{+} \mathcal{P}(\mathrm{G}))\nabla\mathcal{Q}\right),$

$$H^{*}\left((\{F\} \cup^{+} \{G\})\nabla\mathcal{Q}\right), H^{*}(\mathcal{Q})\Big] =$$

$$\Psi\Big[H^*\Big((\mathcal{P}(\mathrm{G})\cup^+\{\mathrm{F}\})\nabla\mathcal{Q}\Big), H^*\Big((\{G\}\cup^+\mathcal{P}(\mathrm{F}))\nabla\mathcal{Q}\Big), H^*\Big((\{G\}\cup^+\{F\})\nabla\mathcal{Q}\Big), H^*(\mathcal{Q})\Big];$$

$$(\mathbf{C})\ \Psi\Big[H^*\Big((\mathcal{P}'(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\mathcal{P}(\mathrm{G}))\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\{G\})\nabla\mathcal{Q}\Big), H^*(\mathcal{Q})\Big]$$
$$\leq \Psi\Big[H^*\Big((\mathcal{P}(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\mathcal{P}(\mathrm{G}))\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\{G\})\nabla\mathcal{Q}\Big), H^*(\mathcal{Q})\Big]$$

if $\mathcal{P}'(\mathrm{F}) \preceq \mathcal{P}(\mathrm{F})$;

$$(\mathbf{D})\ \Psi\Big[H^*\Big(((\mathcal{P}(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q})\nabla\mathcal{R}\Big), H^*\Big(((\{F\}\cup^+\mathcal{P}(\mathrm{G}))\nabla\mathcal{Q})\nabla\mathcal{R}\Big), H^*\Big(((\{F\}\cup^+\{G\})\nabla\mathcal{Q})\nabla\mathcal{R}\Big), H^*(\mathcal{Q})\Big] =$$
$$\Psi\Big[H^*\Big((\mathcal{P}(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\mathcal{P}(\mathrm{G}))\nabla\mathcal{Q}\Big), H^*\Big((\{F\}\cup^+\{G\})\nabla\mathcal{Q}\Big), H^*(\mathcal{Q})\Big] + H^*(\mathcal{R}),$$

if $\mathcal{R} \in \mathcal{K}$ is algebraically independent from $\mathcal{P}(\mathrm{F})$ and $\mathcal{P}(\mathrm{G})$ and it is not conditioned by $\mathcal{Q}$.

Setting:

$$H^*\Big((\{F\}\cup^+\mathcal{P}(\mathrm{F}))\nabla\mathcal{Q}\Big) = x,$$
$$H^*\Big((\mathcal{P}(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q}\Big) = y,$$
$$H^*\Big((\{F\}\cup^+\{G\})\nabla\mathcal{Q}\Big) = z,$$
$$H^*\Big((\mathcal{P}'(\mathrm{F})\cup^+\{\mathrm{G}\})\nabla\mathcal{Q}\Big) = y',$$
$$H^*(\mathcal{R}) = s, H^*(\mathcal{Q}) = t,$$

with $x, y, z, y' \in [0, +\infty], x \leq z, y \leq z$ and $s, t \in (0, +\infty)$. From $(A) - (D)$, we get the following system of functional equations:

$$\begin{cases} (a)\ \Psi(t,t,t,t) = 0\ , \\ (b)\ \Psi(x,y,z,t) = \Psi(y,x,z,t)\ , \\ (c)\ \Psi(x,y',z,t) \leq \Psi(x,y,z,t) \text{ if } y' \leq y\ , \\ (d)\ \Psi(x+s,y+s,z+s,t) = \Psi(x,y,z,t) + s\ . \end{cases}$$

5 Solution of the problem

In this paragraph we shall solve the system of functional equations seen above.

We look for a function Ψ continuous as a universal law, in the sense that the equations and the inequality about Ψ must be satisfied for all values of the variables in their proper spaces.

Proposition 5.1 *Let h be any bijettive and strictly increasing function, differentiable with its inverse, and with $h(0) = 0$; then every function Ψ of the type*

$$\Psi(x,y,z,t) = \frac{1}{3}\Big[h^{-1}\Big(h(x) - h(t)\Big) + h^{-1}\Big(h(y) - h(t)\Big) + h^{-1}\Big(h(z) - h(t)\Big)\Big]\ , \qquad (6)$$

enjoys the equations $(a), (b)$ and (c); it satisfies also (d) if and only if

$$\Psi(x,y,z,t) = \frac{x+y+z}{3} - t\ . \qquad (7)$$

Proof

The conditions $(a), (b), (c)$ are clearly satisfied by any function satisfying (6).

Moreover, it is obvious that the function (7) satisfies all the conditions $(a) - (d)$, and is of the form (6), for $h =$ identity map.

So, it remains to prove that every function Ψ of the form (6), satisfying (d), must be of the type (7).

By (6) the condition (δ) becomes

$$\frac{1}{3}\Big[h^{-1}\Big(h(x+s) - h(t)\Big) + h^{-1}\Big(h(y+s) - h(t)\Big) + h^{-1}\Big(h(z+s) - h(t)\Big)\Big] = \frac{1}{3}\Big[h^{-1}\Big(h(x) - h(t)\Big) + h^{-1}\Big(h(y) - h(t)\Big) + h^{-1}\Big(h(z) - h(t)\Big)\Big] + s\ . \qquad (8)$$

Putting

$$h^{-1}\Big(h(x) - h(t)\Big) = \varphi_t(x), \qquad (9)$$

then (8) becomes

$$\frac{1}{3}\Big[\varphi_t(x+s) + \varphi_t(y+s) + \varphi_t(z+s)\Big]$$

$$= \frac{1}{3}\Big[\varphi_t(x) + \varphi_t(y) + \varphi_t(z)\Big] + s, \quad \text{i.e.}$$

$$\varphi_t(x+s) + \varphi_t(y+s) + \varphi_t(z+s)$$

$$= \varphi_t(x) + \varphi_t(y) + \varphi_t(z) + 3s \ . \tag{10}$$

Choosing $y = x + s,\ z = y + s$, (10) gives rise to $\varphi_t(z+s) = \varphi_t(x) + 3s$, i.e.

$$\varphi_t(x+3s) = \varphi_t(x) + 3s. \tag{11}$$

Now, we calculate the derivative of (11) with respect to s :

$$3\frac{\partial\varphi_t}{\partial s}(x+3s) = 3, \quad \frac{\partial\varphi_t}{\partial s} = 1$$

i.e., by arbitrariness of x and s,

$$\varphi_t(s) = s + c(t) \ , \tag{12}$$

where $c(t)$ is an arbitrary function, which we can calculate by the derivative of the function (11) with respect to t. We get

$$\frac{\partial\varphi_t}{\partial t}(x+3s) = \frac{\partial\varphi_t}{\partial t}(x) \ , \forall x, s \text{ so } \frac{\partial\varphi_t}{\partial t} = (\text{cost}) = c_1$$

and from (12) $c'(t) = c_1$. The last equality implies $c(t) = c_1 t + c_2$, where c_2 is another constant. So, from (12), we have obtained the function

$$\varphi_t(x) = x + c_1 t + c_2. \tag{13}$$

By substituting (13) in (9), we get:

$$h(x) - h(t) = h(x + c_1 t + c_2), \forall x; \tag{14}$$

in particular for $x = t$ we obtain

$$0 = h((1 + c_1)t + c_2) \ \forall t. \tag{15}$$

Now (15) implies two possibilities: 1) $h(t) = 0, \forall t$ which we eliminate, 2) $c_1 = -1$ and $c_2 = 0$. From (14) we deduce $h(x) - h(t) = h(x - t)$, so the function h is linear: $h(x) = \alpha x$. In conclusion, from (6) it follows (7):

$$\Psi(x, y, z, t) = \frac{x + y + z}{3} - t \ .$$

6 Conditional entropy for crisp partitions

For crisp partitions [2], we can use the following locality property:

$$H\Big(\pi_A \cup \pi_B\Big) - H\Big(\pi_A \cup \{B\}\Big) =$$

$$H\Big(\{A\} \cup \pi_B\Big) - H\Big(\{A\} \cup \{B\}\Big) \quad , \tag{16}$$

and we put

$$H_{\pi'}(\pi_A \cup \pi_B) = \Phi\Bigg[H\Big((\pi_A \cup \{B\}) \cap \pi'\Big),$$

$$H\Big((\{A\} \cup \pi_B) \cap \pi'\Big),\ H\Big((\{A\} \cup \{B\}) \cap \pi'\Big),\ H(\pi')\Bigg] \tag{17}$$

$\pi_A, \pi_B, \pi' \in \mathcal{E}$, $\mathcal{E}$ is the family of all partitions of $A, B \subset X$, π' is the partition which conditiones $\pi_A \cup \pi_B$ and $H(\pi') \neq +\infty$, and $\Phi(\cdot) = [0, +\infty] \times [0, +\infty] \times [0, +\infty] \times [0, +\infty] \to [0, +\infty]$.

The system of functional equations are the same, so, also in the crisp setting, we get the same solution (7).

7 Conclusion

Taking into account the locality principle and the independence axiom, the measures of conditional entropy are:

- from (5) and (7), for fuzzy partitions:

$$H_{\mathcal{Q}}(\mathcal{P}(\mathrm{F}) \cup^+ \mathcal{P}(\mathrm{G})) = \frac{1}{3}\Bigg[H^*\Big((\{F\} \cup^+ \mathcal{P}(\mathrm{F}))\nabla\mathcal{Q}\Big) +$$

$$H^*\Big((\mathcal{P}(\mathrm{F}) \cup^+ \{\mathrm{G}\})\nabla\mathcal{Q}\Big) + H^*\Big((\{F\} \cup^+ \{G\})\nabla\mathcal{Q}\Big)\Bigg] - \mathrm{H}(\mathcal{Q})$$

- from (17) and (7), for crisp partitions:

$$H_{\pi'}(\pi_A \cup \pi_B)) = \frac{1}{3}\Bigg[H\Big((\{A\} \cup \pi_B) \cap \pi'\Big) + H\Big((\pi_A \cup \{B\}) \cap \pi'\Big) +$$

$$H\Big((\{A\} \cup \{B\}) \cap \pi'\Big)\Bigg] - H(\pi').$$

Acknowledgement

This research was supported by GNFM of MIUR (ITALY)

References

[1] P.BENVENUTI - D.VIVONA - M.DIVARI, Fuzzy partitions and entropy, Proc. III Linz Seminar on Fuzzy Set Theory: Uncertainty Measures, (Ed. P.Klement, S.Weber) (1991) , 14-18.

[2] M.DIVARI - M.PANDOLFI, Su una legge compositiva dell'Informazione, Rendiconti di Matematica, **3** (1971), 805-817.

[3] D.DUMITRESCU, Fuzzy measures and entropy of fuzzy partitions, J.Math. An. and Appl., **176** (1993), 359-373.

[4] B.FORTE, Measures of information: the general axiomatic theory, R.A.I.R.O (1969), 63-90.

[5] D.MUNDICI, Nonboolean partitions and their logic, Soft Computing, **2** (1998), p.18-21.

[6] D.VIVONA - M.DIVARI, Entropies of fuzzy compositive partitions, Proc. IPMU '06 Congress, Paris (2006), p.1811-1815.

[7] D.VIVONA - M.DIVARI, On a form of conditional information for crisp and fuzzy partitions, Proc. AGOP '05 Congress, Lugano (2005), p.147-149.

[8] D.VIVONA - M.DIVARI, Measures of local entropies for compositive fuzzy partitions (2006), submitted to I.J.U.F.

[9] L.A.ZADEH, Fuzzy sets, Inf. and Cont. **8** (1965), 338-353.

A construction of a fuzzy valued measure based on minimum T-norm

Vechislav Ruzha
University of Latvia
Department of Mathematics
Zellu street 8, Riga, LV - 1002, Latvia
vecislav.ruza@ge.com

Svetlana Asmuss
University of Latvia
Department of Mathematics
Zellu street 8, Riga, LV - 1002,Latvia
svetlana.asmuss@lu.lv

Abstract

In this paper we generalize the well known construction of "classic" measure theory to the fuzzy case. We consider such T-norm based classes of fuzzy sets as T-semirings and T-tribes (T - sigmaalgebras), such fuzzy valued functions as elementary T-measures, exterior T-measures and T-measures, and we suggest the construction of a $\mathrm{T_M}$-measure by extension of an elementary $\mathrm{T_M}$-measure on a $\mathrm{T_M}$-semiring to a $\mathrm{T_M}$-measure on a $\mathrm{T_M}$-tribe in the case of the minimum T-norm $\mathrm{T_M}$.
Keywords: fuzzy set, measure.

1 Preliminaries

Let X be a nonempty set, Φ a sigmaalgebra of its subsets and ν is a finite measure $\nu : \Phi \to [0, +\infty[$.
The class of all fuzzy subsets of X is denoted $F(X, [0, 1])$. The operations for fuzzy sets $A, B \in F(X, [0, 1])$ such as intersection, union and difference are defined by using a triangular norm T and its corresponding triangular conorm S :

$$(A \,\mathsf{T}\, B)(x) = T(A(x), B(x)),$$
$$(A \,\mathsf{S}\, B)(x) = S(A(x), B(x)),$$
$$(A \,\mathsf{D}\, B)(x) = T(A(x), B^c(x))$$

respectively (here ^{c} is the standart involution $\alpha^c = 1 - \alpha$).
The aim of the present paper is to construct a fuzzy valued measure of fuzzy sets by extension a "crisp" measure ν in the case of the minimum T-norm T_M: $T_M(x, y) = min(x, y)$ and $S_M(x, y) = max(x, y)$.

There are many works, where measures of fuzzy sets were studied. A detailed survey of the most important concepts and results obtained in this direction can be found in the paper by D. Butnariu and E.P. Klement [1]. As different from this approach our interest is in developing a theory where not only sets are fuzzy, but also the measure takes fuzzy real values. At this point we had to choose which one of the several (quite different) approaches to the concept of a fuzzy real number. We decided in favour of fuzzy real numbers, as they were first defined by B. Hutton [4], and then thoroughly studied in a series of papers (see e.g.[5],[6]). Although there are also some works were fuzzy valued measures are involved, they used an alternative (essentially different) definition of a fuzzy real number (see [2, 3]). Besides the authors of this works are mainly interested in finding appropriate definititions and stadying the properties of this measures. As different from this works our main interest is in presenting the construction of such measures.
We develope a construction of a fuzzy valued measure by extending a measure defined on a sigmaalgebra of crisp sets to the fuzzy-valued measure define on a T-tribe of fuzzy sets, where T is a given t-norm.
In order to realize the construction we need to introduce a concept of a semiring of fuzzy sets, as well as elementary and exterior measures are needed. In Section 3, we give the definition of a T-semiring of fuzzy sets. Then in Section 4, we consider the definition of a fuzzy-valued elementary T-measure defined on a T-semiring. Section 5 is devoted to the concept of a fuzzy-valued exterior T-measure defined on the class of all fuzzy sets $F(X, [0, 1])$.The main result (Section 8) is proved only in the case of a minimum t-norm.

2 Fuzzy real numbers

This section contains some definitions and properties of fuzzy number theory.

Definition 1. (Hutton [4]). A fuzzy real number is a function $z : \mathbb{R} \to [0, 1]$ such that

(i) z is non-increasing;

(ii) $\inf_x z(x) = 0$, $\sup_x z(x) = 1$;

(iii) z is left semi-continuous, i.e. $\inf_{t<x} z(t) = z(x)$.

The set of all fuzzy real numbers is called *the fuzzy real line* and it is denoted by $\mathbb{R}([0,1])$. A fuzzy number z is called *non-negative* if $z(0) = 1$. The set of all non-negative fuzzy real numbers we denote by $\mathbb{R}_+([0,1])$.

The ordinary real line $\mathbb{R}$ can be identified with the subspace $\{z_a \mid a \in \mathbb{R}\}$ of $\mathbb{R}([0,1])$ by assigning to a real number $a \in \mathbb{R}$ the fuzzy real number z_a defined by

$$z_a(x) = \begin{cases} 1, \text{ if } x \leq a\,, \\ 0, \text{ if } x > a\,. \end{cases}$$

The operation of fuzzy addition $\oplus$ defined by

$$(z_1 \oplus z_2)(x) = \sup_t \big(min\{z_1(t), z_2(x-t)\}\big)$$

whenever $z_1, z_2 \in \mathbb{R}([0,1])$, is a jointly continuous extension of addition + from the real line $\mathbb{R}$ to the fuzzy real line $\mathbb{R}([0,1])$.

3 Classes of fuzzy sets

In questions where T-tribes are involved, our main source of references is [1].

Definition 2. (Klement and Butnariu [1]) A subclass $\Sigma \subset F(X,[0,1])$ is called a *T-tribe* on X if the following properties are satisfied:

(i) $\emptyset \in \Sigma$;

(ii) for all $A \in \Sigma$ we have $A^c \in \Sigma$;

(iii) for all sequences $(A_n)_{n\in N} \subset \Sigma$ we have $\overset{\infty}{\underset{n=1}{\mathsf{T}}} A_n \in \Sigma$.

The operations for a sequence of fuzzy sets $(A_n)_{n\in\mathbb{N}}$ such as intersection and union are defined in the following way:

$$\overset{\infty}{\underset{n=1}{\mathsf{T}}} A_n = \inf_n \overset{n}{\underset{k=1}{\mathsf{T}}} A_n = \inf_n (A_1 \mathsf{T} A_2 \mathsf{T} \ldots \mathsf{T} A_n),$$

$$\overset{\infty}{\underset{n=1}{\mathsf{S}}} A_n = \sup_n \overset{n}{\underset{k=1}{\mathsf{S}}} A_n = \sup_n (A_1 \mathsf{S} A_2 \mathsf{S} \ldots \mathsf{S} A_n).$$

Next we introduce a class of fuzzy sets called a T-semiring.

Definition 3. A subclass $\wp \subset F(X,[0,1])$ is called a *T-semiring* on X if the following properties are satisfied:

(i) $\emptyset \in \wp$;

(ii) for all $A, B \in \wp$ we have $A \mathsf{T} B \in \wp$;

(iii) for all $A, B \in \wp$ there exist such T-disjoint fuzzy sets $A_1, A_2, \ldots, A_n \in \wp$, that $A \mathsf{D} B = \overset{n}{\underset{i=1}{\mathsf{S}}} A_i$.

A finite family of fuzzy sets $A_1, A_2, \ldots, A_n$ is said to be *T-disjoint* (see e.g. [1]) if for each $k \in \{1, \ldots, n\}$

$$(\overset{n}{\underset{j=1, j\neq k}{\mathsf{S}}} A_j) \mathsf{T} A_k = \emptyset.$$

Construction. In this survey the folowing example of a T- semiring plays a special role. For $M \in \Phi$, $\alpha \in [0,1]$ we define a fuzzy set A by

$$A(x) = \begin{cases} \alpha, x \in M, \\ 0, x \notin M. \end{cases}$$

In the sequel the set A will be denoted by $A(M,\alpha)$. We can prove that all such fuzzy sets form a T- semiring

$$\tilde{\wp} = \{A(M,\alpha) | M \in \Phi \text{ and } \alpha \in [0,1]\}.$$

4 T - measure

Definition 4. (Klement and Butnariu [1]) Let Σ be a T-tribe. A function $\mu : \Sigma \to \mathbb{R}_+([0,1])$ is called a fuzzy valued *T-measure* if it satisfies the following conditions:

(i) $\mu(\emptyset) = z_0$;

(ii) for all $A, B \in \Sigma$ it holds $\mu(A \mathsf{T} B) \oplus \mu(A \mathsf{S} B) = \mu(A) \oplus \mu(B)$;

(iii) $\sup_n \mu(A_n) = \mu(A)$, whenever $(A_n)_{n\in N} \subset \Sigma, (A_n)_{n\in N} \nearrow A$ and $A \in \Sigma$.

Construction. Our aim is to construct a fuzzy valued T - measure $\tilde{\mu}$ on a T -tribe by extension a "crisp" measure ν. To achieve this purpose we generalize the well known construction of "classic" measure theory (see e.g. [7]) to the fuzzy case.

5 Elementary T - measure

Definition 5. Let $\wp$ be a T-semiring. A function $m : \wp \to \mathbb{R}_+([0,1])$ is called an *elementary T-measure* if it satisfies the following conditions:

(i) $m(\emptyset) = z_0$;

(ii) for all $A, B \in \wp : A \mathsf{T} B = \emptyset$ we have $A \mathsf{S} B \in \wp \Rightarrow m(A \mathsf{S} B) = m(A) \oplus m(B)$.

Construction. We define an elementary T_M-measure $\tilde{m}$ on the T_M-semiring $\tilde{\wp}$ by the formula $\tilde{m}(A(M,\alpha)) = z_{\nu(M),\alpha}$, where

$$z_{\nu(M),\alpha}(t) = \begin{cases} 1, t \leq 0, \\ \alpha, 0 < t \leq \nu(M), \\ 0, t > \nu(M), \end{cases}$$

is the fuzzy real number. We proved that the function $\tilde{m}$ is countably T_M-semiaditive, i.e.

for all sequences $(A_n)_{n\in N} \subset \tilde{\wp}$ and for all $A \in \tilde{\wp} : A \subset \overset{\infty}{\underset{n=1}{\mathsf{S}_{\mathsf{M}}}} A_n$ we have $\tilde{m}(A) \leq \sum_{n=1}^{\infty} \tilde{m}(A_n)$.

6 Exterior T - measure

Definition 6. A function $m^* : F(X,[0,1]) \to \mathbb{R}_+([0,1])$ is called an *exterior T-measure* if it satisfies the following conditions:

(i) $m^*(\emptyset) = z_0$;

(ii) for all sequences $(A_n)_{n\in N} \subset F(X,[0,1])$ and for all $A \in F(X,[0,1]) : A \subset \overset{\infty}{\underset{n=1}{\mathsf{S}}} A_n$ it holds $m^*(A) \leq \sum_{n=1}^{\infty} m^*(A_n)$.

Construction. The function $\tilde{m}^* : F(X,[0,1]) \to \mathbb{R}_+([0,1])$ defined by

$$\tilde{m}^*(A) = inf\{\sum_{n=1}^{\infty} \tilde{m}(A_n) | (A_n)_{n\in\mathbb{N}} \subset \wp : A \subset \overset{\infty}{\underset{n=1}{\mathsf{S}_{\mathsf{M}}}} A_n\}$$

whenever $A \in F(X,[0,1])$, is the exterior T_M-measure.

7 T_M-tribe of m^*-measurable sets

Definition 7. Let m^* be an exterior T-measure. A fuzzy set $A \in F(X,[0,1])$ is called a m^*-measurable set, if it satisfies the following conditions for all sets $E \in F(X,[0,1])$:

(i) $m^*(A) \oplus m^*(E) = m^*(A \mathsf{T} E) \oplus m^*(A \mathsf{S} E)$;

(ii) $m^*(A^c) \oplus m^*(E) = m^*(A^c \mathsf{T} E) \oplus m^*(A^c \mathsf{S} E)$.

Construction. We denote $\Sigma_{\tilde{m}^*}$ the class of all $\tilde{m}^*$-measurable sets. It is proved that

(i) $\tilde{\wp} \subset \Sigma_{\tilde{m}^*}$;

(ii) $\Sigma_{\tilde{m}^*}$ is a T_M-tribe.

8 The restriction of an exterior T_M-measure to a T_M-tribe

Construction. We denote by $\tilde{\mu}$ the restriction of $\tilde{m}^*$ to $\Sigma_{\tilde{m}^*}$. It is proved that

(i) $\tilde{\mu}$ is a T_M-measure;

(ii) $\tilde{\mu}/\tilde{\wp} = \tilde{m}$;

(iii) $\tilde{\mu}/\Phi = \nu$.

The last equality means that for every $M \in \Phi$ it holds $\tilde{\mu}(A(M,1)) = z_{\nu(M)}$.

References

[1] Klement E. P., Butnariu D., *Triangular norm based measures.* In: Endre Pap (Hrsg.), *"Handbook of Measure Theory"*, Kapitel 23, Seite(n) 947-1010, Elsevier, Amsterdam, 2002

[2] Chen Degang, Zhang Liguo, *Signed Fuzzy-Valued Measure and Radon-Nikodym Theorem of Fuzzy-Valued Measurable Functions*, Southeaset Asian Bulletin of Mathematics (2002) 26: 375-385.

[3] Hsien-Chung Wu, *Fuzzy-valued integrals of fuzzy-valued measurable functions with respect to fuzzy-valued measures based on closed intervals*, Fuzzy Sets and Systems 87 (1997) 65-78.

[4] Hutton B., *Normality in fuzzy topological spaces*, J.Math.Anal.Appl., 50 (1975), 74-79.

[5] Gantner T.E., Steinlage R.C., Warren R.H. *Compactness in fuzzy topological spaces*, J.Math.Anal.Appl., 62 (1978), 547-562.

[6] Rodabaugh S.E., *Fuzzy addition and the L-fuzzy real line*, Fuzzy Sets and Systems, (1982), V.8 P.39-52.

[7] Halmosh P., *Measure theory*, Moscow, IL (1953) (in russian).

Compenzational Vagueness

Milan Mareš
Institute of information Theory and Automation
Academy of Sciences of the Czech Republic
P. O. Box 18, 182 08 Praha 8, Czech Republic
mares@utia.cas.cz

Abstract

Some manipulations with vague quantities consist in an aggregation of vague amounts where the resulting aggregated quantity (in our case a sum of vague summands) is expected to be (almost) crisp. The aim of this contribution is to analyze and briefly discuss various approaches to this problem which can be supported by the elementary theory of fuzzy quantities. The analysis is focused on the possibility of achieving the desired sum, on the fuzzy set theoretical methods applicable to this model, and also to the limits of regulation of some fuzzy summands during the aggregation process.

Keywords: Fuzzy quantity, extension principle, decomposition paradigm, shape, scale, compensation of vagueness, vagueness.

1 Introduction

The well handled processing of fuzzy quantities appears to be one of the most prospective trends of further development of fuzzy set theory and applications. It becomes important for advanced methods of management of uncertainty and vagueness in numerous economic, sociological and organizational procedures. The significancy of the multilevel investigation of fuzzy quantities is analyzed from the point of view of fuzzy methodology (see, e. g., [12, 13, 14]), its traditional methods are modified to be more adequate to the real problems (e. g., in [11]), and the fuzzy set theoretical model of quantitative phenomena is widely generalized (as, e. g., in [1]).

This development of the general theory implies an extension of the scale of its potential applications to new sorts of problems. In this paper, we discuss several approaches to the problem of aggregation of vague (it means fuzzy) components aiming to minimize the variability (fuzziness) of the result.

Such problem seems to be rather contradictoric with regard to the essence of fuzziness, nevertheless, it is not as unusual as it appears.

Let us consider, for example, a model of transport line passing one or more check-points where the time demanded for passing particular intervals is relatively free, meanwhile their total is given by a time-table as a narrow interval. Very similar situation appears when a container is part-wise filled by some material, where particular doses are in certain limits arbitrary but the total amount is given quite strictly. All such situations, formulated in the language of fuzzy set theory, mean the demand to get a crisp (or almost crisp) sum of fuzzy components. This is impossible in the classical model based on the extension principle. However, such situations exist and they are not very rare.

Some attempts to it were done in [7, 8], some others can be potentially derived by means of the decomposition model (see [11] and others). In this contribution, we discuss the effectivity and adequacy of some of them. Namely, we compare the models using extension and decomposition principles. The questions on which we focus our attention are:

— how to estimate the possibility that (after the realization of particular vague additions) the total sum will be equal or very near to the desired value,

— how to characterize the possibilities of compensation of declinations of consequently added elements,

— which of both principles of computation with fuzzy quantities appears more effective (and due

to which criteria) in modelling the real situations of the considered type.

2 Fuzzy Quantity

In the whole paper, we denote by R the set of all real numbers.

Due to [5, 6] and other works, we define *fuzzy quantity* $\boldsymbol{a}$ as a fuzzy subset of R with membership function μ_a such that:

(1) There exists $x_a \in R$ such that $\mu_a(x_a) = 1$.

(2) There exist x_1, $x_2 \in R$ such that $x_1 < x_a < x_2$ and $\mu_a(x) = 0$ for $x \notin [x_1, x_2]$.

Any member $x_a \in R$ fulfilling (1) is called a *modal value* of $\boldsymbol{a}$. The set of all fuzzy quantities will be denoted by $\mathbb{R}$. The properties of fuzzy quantities are summarized in many works. Let us mention at least [2, 3, 4, 5, 6, 9].

To simplify the notations, we denote by $\langle r \rangle$, where $r \in R$, the fuzzy quantity in $\mathbb{R}$ which is condensed in a single possible value, i. e.

$$\mu_{\langle r \rangle}(r) = 1, \quad \mu_{\langle r \rangle}(x) = 0 \quad \text{for } x \neq r. \tag{3}$$

If $\boldsymbol{a} \in \mathbb{R}$ then we denote by $(-\boldsymbol{a}) \in \mathbb{R}$ the *opposite* fuzzy quantity to $\boldsymbol{a}$, and

$$\mu_{(-a)}(x) = \mu_a(-x) \quad \text{for all } x \in R. \tag{4}$$

As the fuzzy quantities are, in fact, extensions of crisp real numbers, it is rational to extend even the algebraic operations used over R, on $\mathbb{R}$. It is usually done by so called *extension principle*. In this paper we need only the operation of addition which, extended over $\mathbb{R}$, is denoted by $\oplus$, and if $\boldsymbol{c} = \boldsymbol{a} \oplus \boldsymbol{b}$, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbb{R}$, then

$$\mu_c(x) = \sup_{y \in R} [\min(\mu_a(y),\, \mu_b(x - y))] \tag{5}$$

for any $x \in R$. The properties of this (and others) operation are summarized, e. g., in [2, 3, 5, 6]. It can be easily seen that they are rather similar but not identical with the group properties of the addition of crisp real numbers.

Some of the essential differences between both addition operations are related with the concept of *fuzzy zero*. Its seemingly natural definition as $\langle 0 \rangle$ (note, $\boldsymbol{a} \oplus \langle 0 \rangle = \boldsymbol{a}$ for every $\boldsymbol{a} \in \mathbb{R}$) does not fit with the concept of opposite element as $(-\boldsymbol{a})$. As shown in [5, 6] and further specified in [9], fuzzy zeros form a class of fuzzy quantities, closed for the operation $\oplus$, which we denote

$$\mathbb{S} = \{\boldsymbol{s} \in \mathbb{R} : \mu_s(x) = \mu_s(-x) \text{ for all } x \in R\}. \tag{6}$$

Fuzzy quantities from $\mathbb{S}$ are called *symmetric*. Due to [9], the class $\mathbb{S}$ can be narrowed to

$$\mathbb{Z} = \{\boldsymbol{z} \in \mathbb{R} : \boldsymbol{z} \in \mathbb{S},\, \mu_z(0) = 1\}, \tag{7}$$

whose elements are called *strongly symmetric*. The algebraic properties of symmetric and strongly symmetric fuzzy quantities are equivalent and they agree with the group axioms (including commutativity) under the condition that we substitute the identity between fuzzy quantities $a = b$, a, $b \in \mathbb{R}$,

$$\boldsymbol{a} = \boldsymbol{b} \quad \text{if} \quad \mu_a(x) = \mu_b(x) \quad \text{for all } x \in R, \tag{8}$$

by a weaker equivalence due to which $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{R}$ are equivalent iff they differ in fuzzy zeros, only (see, [5, 6, 9]).

The comparison and ordering relation between fuzzy quantities can be (and often is) defined in different ways, due to the character of the solved problem (a representative analysis of it can be found in [4], some possibilities are mentioned in [6], too). Here, we use the ordering relation most respecting the vague structure of fuzzy quantities. Namely, we define the *fuzzy ordering* relation $\succeq$ over $\mathbb{R}$ as a fuzzy relation with membership function $\nu_{\succeq} : \mathbb{R} \times \mathbb{R} \to [0, 1]$ such that for $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{R}$ the possibility of $\boldsymbol{a} \succeq \boldsymbol{b}$ is

$$\nu_{\succeq}(\boldsymbol{a}, \boldsymbol{b}) = \sup_{\substack{x, y \in R \\ x \geq y}} [\mu_a(x),\, \mu_b(y)], \tag{9}$$

and, analogously, we define *fuzzy equality* $\boldsymbol{a} \sim \boldsymbol{b}$ as fuzzy relation with $\nu_{\sim}(\boldsymbol{a}, \boldsymbol{b}) : \mathbb{R} \times \mathbb{R} \to [0, 1]$,

$$\nu_{\sim}(\boldsymbol{a}, \boldsymbol{b}) = \sup_{x \in R} [\mu_a(x),\, \mu_b(x)]. \tag{10}$$

3 Decomposition Paradigm

The algebraic operations based on the extension principle, including the addition (5), are natural. They reflect the expected properties of composition of vagueness of summands after the summation procedure in the standard situations. One of the evident results is, that the vagueness of the sum is higher than the one of the summands, where the vagueness is represented by the extent and values of the membership functions. This observation contradicts with the essential demand of the problem analyzed in this paper. Hence, it appears evident that the extension principle cannot serve as the method of its solution.

An alternative approach to the processing of fuzzy quantities increasing the flexibility of actually applied methods was suggested in several papers, including [7], and it was summarized in [11]. It is based on so called *decomposition paradigm* due to which any fuzzy quantity $\boldsymbol{a} \in \mathbb{R}$ consist of three components:

— the *crisp numerical value* x_a

— the *shape* function φ_a characterizing the structure of vagueness connected with $\boldsymbol{a}$,

— the increasing *scale* function f_a characterizing the structure of vagueness connected with the source of data.

The formal properties of all these components are presented, e. g., in [11]. The essential meaning of this decomposed model is that it admits the possibility to process each of the components separately, with regard to the demands of actual application. Such processing leads to the components of the resulting fuzzy quantity (let us denote it $\boldsymbol{c} \in \mathbb{R}$ with components (x_c, φ_c, f_c)), where the membership function μ_c is derived by means of

$$\mu_c(x) = \varphi_c\left(f_c(x) - f_c(x_c)\right), \quad x \in R. \tag{11}$$

The consequence of the decomposition paradigm significant for the model of addition which we are looking for is the admissibility of separate processing of shapes independent on the scales and crisp values. The shapes represent the main features of the membership functions – namely, they reflect the main structure of vagueness and its behaviour. In the following sections, we use especially the possibility to combine the shapes not only by means of the extension principle (or its very close analogy) but by other relevant fuzzy set theoretical operations. The choice of the one used in a specific model can be modified with respect to the described reality.

4 Composition of Vagueness

In the main section, we briefly discuss and compare some approaches to the problem formulated in the Introduction.

Let us consider fuzzy quantities $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbb{R}$ with membership functions μ_a, μ_b, μ_c, respectively. Let us denote, for any $\boldsymbol{d} \in \mathbb{R}$ with μ_d, the values

$$\mu_d^{\sup} = \sup\{x \in R : \mu_d(x) > 0\} \tag{12}$$

$$\mu_d^{\inf} = \inf\{x \in R : \mu_d(x) > 0\} \tag{13}$$

and

$$V_d = x_d^{\sup} - x_d^{\inf}. \tag{14}$$

In the following subsections, we analyze the demand that

$$\min(V_a, V_b) > V_c \tag{15}$$

including the possibility that $\boldsymbol{c} = \langle x_c \rangle$ for modal value x_c of c. By x_a, x_b we denote some of the (possibly many) modal values of $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively.

The main goal of this contribution is to discuss some of the possible compositions of $\boldsymbol{a}$ and $\boldsymbol{b}$ whose result is near to the fuzzy quantity $\boldsymbol{c}$. For the purpose of general problem discussion, we denote this composition by $\circ$, which means that we discuss the properties of the relations represented by formulas

$$\boldsymbol{a} \circ \boldsymbol{b} \sim \boldsymbol{c} \quad \text{or} \quad \boldsymbol{a} \circ \boldsymbol{b} \succeq \boldsymbol{c}. \tag{16}$$

Special attention is focused on the compositions of fuzzy quantities $\boldsymbol{a} \circ \boldsymbol{b}$ where (15) is fulfilled.

In the following subsections, we briefly discuss three specific cases of the composition operation $\circ$, especially from the point of view of the degree, in which they respect demands (15) and (16). Priority is given to the following questions:

— How large is the possibility that the composition $\boldsymbol{a} \circ \boldsymbol{b}$ is close to $\boldsymbol{c}$?

— What is the extent of possible values of the composed fuzzy quantity $\boldsymbol{a} \circ \boldsymbol{b}$ (i. e., the value of $V_{a \odot b}$)?

— Does the actual method include a possibility of compensation of the composition by means of proper modification of fuzzy quantity $\boldsymbol{b}$ in accordance with $\boldsymbol{a}$?

— Is it possible to compare the effectivity of particular composition methods.

4.1 Extension Principle Model

The first model of composition operation which we consider, is the classical extension principle (5), i. e., we consider the operation $\oplus$ in the position of $\circ$. It means that we compute, by means of the extension principle, the sum $\boldsymbol{a} \oplus \boldsymbol{b} \in \mathbb{R}$ with membership function $\mu_{a \oplus b}$.

Due to the above notation

$$\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{a} \oplus \boldsymbol{b}. \tag{17}$$

Unfortunately, this method evidently does not respect the most important general demand, namely the one demanding (see (14))

$$V_{a \oplus b} < \min(V_a, V_b).$$

Observation 1. The possibility of $\boldsymbol{a} \oplus \boldsymbol{b} \sim \boldsymbol{c}$ is equal to 1 iff there exist modal values x_a, x_b, x_c of $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, respectively, such that

$$x_a + x_b = x_c.$$

Observation 2. The possibility that $\boldsymbol{a} \oplus \boldsymbol{b} \sim \boldsymbol{c}$ is generally equal to $\mu_{\sim}(\boldsymbol{a} \oplus \boldsymbol{b}, \boldsymbol{c})$ due to (10).

Observation 3. Using (12), (13) and (14),

$$x^{\inf}_{a\oplus b} = x^{\inf}_a + x^{\inf}_b, \quad x^{\sup}_{a\oplus b} = x^{\sup}_a + x^{\sup}_b,$$

hence

$$V_{a\oplus b} \geq \max(V_a, V_b) > V_c.$$

Observation 4. The previous statements cannot be changed by any choice of $\boldsymbol{b}$. If we choose $\boldsymbol{b} = \langle x_c - x_a \rangle$ as a crisp quantity then $V_{a\oplus b} = V_a > V_c$, and $\nu_{\sim}(\boldsymbol{a} \oplus \boldsymbol{b}, \boldsymbol{c}) = 1$.

The previous observations show that the dogmatic application of the extension principle does not generally aim to the desirable model of the considered situation.

4.2 Complementation Model

The next model modifies the previous extension principle approach by accenting stronger stress on the potential possibility of compensation of vague declinations of the first component by proper modification of the second one.

Evidently, the space for the choice of the second summand (we are used to denote it $\boldsymbol{b} \in \mathbb{R}$) is not unlimited. Its boundaries are often given by physical or technological parameters, accessible stock of material or energy, qualification and accessibility of human resources, etc. Let us denote the *accessible set of the second summands* by $\mathbb{B} \subset \mathbb{R}$.

Using the above notation, the operation $\circ$ in this case means

$$\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{a} \oplus \boldsymbol{d} \quad \text{for some } \boldsymbol{d} \in \mathbb{B}, \tag{18}$$

or, more generally,

$$\boldsymbol{a} \circ \boldsymbol{b} = \{\boldsymbol{a} \oplus \boldsymbol{d} : \boldsymbol{d} \in \mathbb{B}\}.$$

The choice of $\boldsymbol{d} \in \mathbb{B}$ is to respect the following statements.

Summarizing the model, we can see that there exists a desirable terminal fuzzy quantity $\boldsymbol{c} \in \mathbb{R}$ which is to be as exact as possible, with a modal value x_c. Moreover, "the first" summand, a fuzzy quantity $\boldsymbol{a} \in \mathbb{R}$ is (vaguely) known. Our aim in this subsection is to choose "the second" summand, fuzzy quantity $\boldsymbol{b} \in \mathbb{B} \subset \mathbb{R}$ which is the best one from the point of view of the goal of this model. Let us suppose, still, that the considered composition operation $\circ$ is the extension principle summation $\oplus$. Then all observations introduced in Subsection 4.1 keep valid, and, moreover, the following ones can be seen whenever (18) is kept in mind.

Observation 5. If there exists $\boldsymbol{b} \in \mathbb{B}$ such that $x_b = x_c - x_a$then $\nu_{\sim}(\boldsymbol{a} \oplus \boldsymbol{b}, \boldsymbol{c}) = 1$ and, consequently, $\nu_{\sim}(\boldsymbol{a} \circ \boldsymbol{b}, \boldsymbol{c}) = 1$.

Observation 6. If the fuzzy quantity condensed in one possible value $\langle x_c - x_a \rangle \in \mathbb{B}$ and if we put $\boldsymbol{b} = \langle x_c - x_a \rangle$ then the statement of Observation 5 holds and, moreover $V_{a+b} = V_a$ which extent of variability is the nearest one to V_c.

Observation 7. If we put $\boldsymbol{b} = \boldsymbol{c} \oplus (-\boldsymbol{a})$ and if $\boldsymbol{b} \in \mathbb{B}$ then it fulfils assumptions of Observation 5.

4.3 Accessibility Testing Model

One of the questions to be answered whenever we try to solve the problem of (almost) deterministic total value of vague summands is the possibility that at least one value of the sum $\boldsymbol{a} \circ \boldsymbol{b}$ is equal to the desired total value $\boldsymbol{c}$.

Let $\boldsymbol{a}$, $\boldsymbol{b}$ be given, and let us define $\boldsymbol{d} = \boldsymbol{c} \oplus (-\boldsymbol{a})$. Then the eventual possibility of successful outcome of our problem, it means the possibility that

$$\boldsymbol{a} \oplus \boldsymbol{b} \sim \boldsymbol{c},$$

is a fuzzy phenomenon whose possibility is

$$\nu_{\sim}(\boldsymbol{d}, \boldsymbol{b}). \tag{19}$$

In a more lucid way, this possibility can be defined also by the formula

$$\max_{x \in R} [\min(\mu_d(x), \mu_b(x))]. \tag{20}$$

Observation 8. Analogously to Observation 1, possibilities (19) and (20) are equal to 1 iff there exist modal values x_a, x_b, x_c such that $x_a + x_b = x_c$, as in such case also $x_d = x_c - x_a = x_b$ and, consequently $x_c = x_a + x_d$, as well.

5 Extent of Vagueness

The problem of regulation and compensation of summands in order to maximize the possibility that the sum will be close to some not very extensive fuzzy quantity has also another aspect. In the previous section, the main attention is focused on the conditions, under which at least the modal values of the sum and goal quantity are identical. The second problem, to achieve the vagueness of the sum as narrow as possible, demands a qualitatively new approach based on the methods suggested (and analyzed) in [11, 7] and in several related papers. It means, to process the

crisp modal values and the shapes of the vagueness separately, and by means of other procedures than the extension principle.

Let us consider fuzzy quantities $\boldsymbol{a}$, $\boldsymbol{b}$ with modal crisp values x_a, x_b and shape functions φ_a, φ_b, respectively (cf. Section 2). To simplify the notations, we put the scales f_a, f_b (see Section 2 and (11))

$$f_a(x) = f_b(x) = x, \tag{21}$$

and, consequently, it is possible to construct the membership functions μ_a, μ_b, using (11)

$$\mu_a(x) = \varphi_a(x - x_a), \quad \mu_b(x) = \varphi_b(x - x_b), \tag{22}$$

where the shape functions fulfil:

- $\varphi_a(x) \in [0,1]$, $\varphi_b(x) \in [0,1]$, $x \in R$,
- $\varphi_a(0) = \varphi_b(0) = 1$,
- φ_a, φ_b are continuous, not decreasing for $x < 0$, and not increasing for $x > 0$,
- there exist x_1, $x_2 \in R$ such that $\varphi_a(x) = \varphi_b(x) = 0$ for $x \notin [x_1, x_2]$.

Evidently, the shape functions represent normalized forms of the membership functions.

Respecting the procedures, methods and conclusions discussed in [11], we may compute a composition $\boldsymbol{a} \circ \boldsymbol{b}$, due to Subsection 4.1, by means of separate composition of x_a, x_b and φ_a, φ_b. Let us keep the denotation

$$\boldsymbol{c} = \boldsymbol{a} \circ \boldsymbol{b}, \quad \boldsymbol{c} \text{ is defined by } (x_c, \varphi_c) \tag{23}$$

where φ_c fulfils the above properties of the shapes and the eventual scale f_c of $\boldsymbol{c}$ is equal to f_a and f_b.

In the case of our problem, it is natural to combine (or compose) fuzzy quantities in (23) by

$$x_c = x_a + x_b \quad \text{and} \quad \varphi_c = \varphi_a \circ \varphi_b, \tag{24}$$

where the composition of shapes $\varphi_a \circ \varphi_b$ can be relatively arbitrary, under the condition that it reflects the natural character of the modelled reality.

In our case, two following procedures appear natural and adequate to the problem:

$$\varphi_c(x) = \min\left(\varphi_a(x), \varphi_b(x)\right), \quad x \in R, \tag{25}$$

and

$$\varphi'_c(x) = \varphi_a(x) \cdot \varphi_b(x), \quad x \in R, \tag{26}$$

however, the possibility of other, more sophisticated, operations is admissible as well.

Observation 9. Evidently, $\mu_c(x) = \varphi_c(x - x_c)$ fulfills $\mu_c(x_c) = 1$ and

$$V_c = V_a \cap V_b,$$
$$\mu_c(x) = \min\left(\mu_a(x - x_b), \mu_b(x - x_a)\right), \quad x \in R,$$

which means that the demand of the minimization of the extent of the vagueness is fulfilled.

Observation 10. Similarly, if (26) is used and if $x_1 > -1$, $x_2 < 1$ then

$$V_c \subset V_a \cap V_b,$$
$$\mu_c(x) \leq \min\left(\mu_a(x - x_b), \mu_b(x - x_a)\right), \quad x \in R,$$

which means that the goals of the procedure are fulfilled, as well.

Observation 11. If $x_1 < -1$, or $x_2 > 1$, then the extents of vagueness V_c, V_a, V_b do not fulfill the previous Observation 10 but, the significantly large values of $\mu_c(x)$ are condensed near x_c.

6 Conclusive Remarks

In the last two sections, we have analyzed the problem of construction of an additive composition method resulting into a fuzzy quantity with minimal uncertainty, from two points of view. How to maximize the possibility of the most desired value, and how to minimize the uncertainty connected with the result. The second condition is very near to the demand to decrease the possibility of less desired values of the result.

In fact, the optimal approach to the problem formulated in Introduction is based on s simultaneous application of both views and by combination of their results. This represents the possible continuation of the research introduced here.

Acknowledgement

The research whose results are summarized above was partly supported by the project of Research Centre: Data, Algorithms, Decision–Making (DAR), No. 1M0572 donated by the Ministry of Education of the Czech Republic.

Moreover, part of the research was supported by the Institutional Research Plan No. AVOZ 10750506 donated by the Academy of Sciences of the Czech Republic.

References

[1] T. Calvo, G. Mayor, R. Mesiar (Eds.), Aggregation Operators. Physica–Verlag, Heidelberg, 2002.

[2] D. Dubois, E. E. Kerre, R. Mesiar, H. Prade, Fuzzy Interval Analysis, in: *Fundamentals of Fuzzy Sets*. Kluwer, Dordrecht, Vol. 1, 2000, pp. 483–581.

[3] D. Dubois, H. Prade, Fuzzy Numbers: An Overview, in: J. C. Bezdek (Ed.), *Analysis of Fuzzy Information*. CRC–Press, Boca Raton, Vol. 2, 1988, pp. 3–39.

[4] E. E. Kerre, X. Wang, Reasonable properties for the ordering of fuzzy quantities, Part I, Part II, Fuzzy Sets and Systems 118 (2001) 375–385, 387–405.

[5] M. Mareš, Computation Over Fuzzy Quantities. CRC–Press, Boca Raton 1994.

[6] M. Mareš, Weak arithmetics of fuzzy numbers, Fuzzy Sets and Systems 91 (1997), 2, 143–154.

[7] M. Mareš, Composition of Fuzzy Verbal Data, in: Proceedings of *EUSFLAT 2001*, Leicester 2001, pp. 463–465.

[8] M. Mareš, Quantitative aspects of fuzzy geographic information, Control and Information (submitted).

[9] M. Mareš, Two fuzzy zeros, two fuzzy units, International Journal of Inovative Computing, Information Control (submitted).

[10] M. Mareš, Vague Verbal Quantities in Transport Timetable, in: Proceedings of *IPMU 2002*, Annecy 2002, pp. 675–680.

[11] M. Mareš, R. Mesiar, Verbally Generated Fuzzy Quantities, in: T. Calvo, R. Mesiar, G. Mayor (Eds.), *Aggregation Operators*. Physica–Verlag, Heidelberg 2002, pp. 291–352.

[12] L. A. Zadeh, From computing with numbers to computing with words – from manipulation of measurements to manipulation of perception, IEEE Transactions on Circuits and Systems 45 (1999)105–119.

[13] L. A. Zadeh, Fuzzy logic = Computing with words, IEEE Transaction on Fuzzy Systems 2 (1996) 103–111.

[14] L. A. Zadeh, The Concept of Cointensive Precisation – A Key to Mechanization of Natural Language Understanding, in: Proceedings of *IMPU 2006*, Paris 2006, pp. 13–15.

On monotonic operations which are locally internal on some subset of their domain

Paweł Drygaś
Institute of Mathematics, University of Rzeszów,
35-310 Rzeszów, ul. Rejtana 16a, Poland.
e-mail: paweldr@univ.rzeszow.pl

Abstract

In the paper [4] a structure of operation locally internal on some subset of domain is given. Here we present additional properties of such operation and sufficient conditions to obtain such operation.

Keywords: Uninorms, monotonic operation, idempotent operation, ordinal sum of semigroup, locally internal operation.

1 Introduction

Binary operations in the unit interval have many applications in fuzzy set theory as multivalued logical connectives (cf.[9]). On the other hand, they are examples of aggregation operators in the unit interval (cf.[13]). For that reason it is important to examine and characterize such operations.

Some examples of binary operations $U : [0,1]^2 \to [0,1]$ are uninorms.
Uninorms were introduced by Yager and Rybalov [13] as a generalization of triangular norms and conorms letting the neutral element e to be an arbitrary point of the unit interval. In [7] Fodor, Yager and Rybalov examined a general structure of uninorms. For example, the frame structure of uninorms and characterization of representable uninorms were presented.
In the case $e = 1$ we obtain triangular norms and in the case $e = 0$ triangular conorms. In the case $e \in (0,1)$ we obtain operations considered in [2] - [7], [10]-[13].

In this paper we will consider more general class of operation as uninorm i.e. we omit the assumption of the commutativity.

The main goal of this paper is to present a characterization of operations locally internal on a some subset of domain.

In section 2 we present some notions and descriptions of the structure of uninorms and construction of ordinal sum of semigroup. In section 3 we put the characterization of idempotent uninorm given in [12]. Next, in section 4 we present the structure of operation which are locally internal on some subset of their domain, properties of the corresponding function and its influence on the structure of underlying operations (t-norm and t-conorm). At the end of paper we give full characterization of such operation.

2 Notion of uninorms

We discuss the structure of binary operations U in the class

$$\mathcal{U}(e) = \{U : [0,1]^2 \to [0,1] \ : U \text{ is increasing, associative, with the neutral element } e \in [0,1]\}$$

Definition 1 ([13])**.** An operation $U \in \mathcal{U}(e)$ (where $e \in [0,1]$) is called a uninorm if it is commutative.
A uninorm T (S) is called a triangular norm (triangular conorm) if it has the neutral element $e = 1$ ($e = 0$).

Definition 2 ([9])**.** Operation T (S) is called a triangular subnorm (triangular superconorm) if it is commutative, associative, increasing and fulfils the condition $T \leq \min$ ($S \geq \max$).

Example 1 (cf. [9])**.** Here are the basic triangular norms:
$T_M(x,y) = \min(x,y),\ x,y \in [0,1]$ (minimum),
$T_P(x,y) = x \cdot y,\ x,y \in [0,1]$ (product),
$T_L(x,y) = \max(x+y-1,0),\ x,y \in [0,1]$ (Łukasiewicz triangular norm).
Of course every triangular norms are triangular subnorms. Operations
$T_1(x,y) = 0, \quad x,y \in [0,1]$,
$T_2(x,y) = \frac{1}{2}x \cdot y, \quad x,y \in [0,1]$,
are triangular subnorms, but not triangular norms.

These operations we can use to construct new triangular norms.

Lemma 1 (cf.[8])**.** *Let $\{[a_k,b_k]\}_{k\in\mathcal{T}}$ be a countable family of nonoverlapping, closed, proper subintervals*

of $[0,1]$. *Let* T *be an operation in* $[0,1]$ *defined by*

$$T(x,y)=\begin{cases} a_k+(b_k-a_k)T_k\left(\frac{x-a_k}{b_k-a_k},\frac{y-a_k}{b_k-a_k}\right), \\ \qquad\qquad if \quad x,y\in(a_k,b_k], \\ \min(x,y) \qquad otherwise, \end{cases} \tag{1}$$

where T_i *are triangular subnorms. Moreover, we assume that operation* T_k *have neutral element* $e=1$, *if* $b_k=a_l$ *and* T_l *is with zero divisor, or* $b_k=1$. *Then* T *is a triangular norm.*

Operation given by (1) is called the ordinal sum of $\{([a_i,b_i],T_i)\}_{i\in\mathcal{T}}$.

Next we present the frame structure of operation from the class $\mathcal{U}(e)$

Theorem 1 (cf. [7]). *If an operation* U *is in the class* $\mathcal{U}(e)$ *for* $e\in(0,1)$, *then there exist an operation* $T\in\mathcal{U}(1)$ *and an operation* $S\in\mathcal{U}(0)$ *such that*

$$U(x,y)=\begin{cases} \varphi^{-1}\left(T\left(\varphi(x),\varphi(y)\right)\right) \ for \ \ x,y\in[0,e] \\ \psi^{-1}\left(S\left(\psi(x),\psi(y)\right)\right) \ for \ \ x,y\in[e,1] \end{cases}, \tag{2}$$

where

$$\begin{cases} \varphi(x)=x/e, & x,y\in[0,e] \\ \psi(x)=(x-e)/(1-e), & x,y\in[e,1] \end{cases}. \tag{3}$$

Let us denote

$$A(e)=[0,e)\times(e,1]\cup(e,1]\times[0,e). \tag{4}$$

Lemma 2. *If* $U:[0,1]^2\to[0,1]$ *is an increasing operation and has the neutral element* $e\in(0,1)$, *then*

$$\min\le U\le\max \ \ in\ A(e). \tag{5}$$

Furthermore, if U *is associative, then* $U(0,1),U(1,0)\in\{0,1\}$.

The frame structure of uninorms, by Theorem 1 and Lemma 2, is depicted on Figure 1.

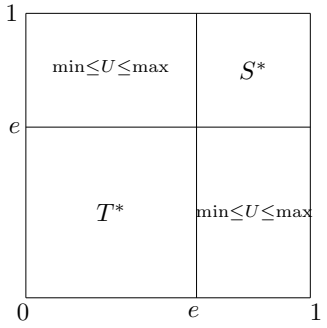


Figure 1: Structure of uninorm

Theorem 2 (cf.[10]). *Let* $e\in(0,1)$. *If* T *is an arbitrary triangular norm and* S *is an arbitrary triangular conorm, then formula* (2) *with* $U=\min$ *(or* $U=\max$*) in* $A(e)$ *gives a uninorm.*

Example 2 (cf.[7]). Formula

$$U(x,y)=\begin{cases} 0, & x=0 \text{ or } y=0, \\ \frac{xy}{(1-x)(1-y)+xy}, & x>0 \text{ and } y>0, \end{cases}$$

gives uninorm with $e=\frac{1}{2}$, $T(x,y)=\frac{xy}{2-(x+y-xy)}$, $S(x,y)=\frac{x+y}{1+xy}$, $x,y\in[0,1]$. T and S are arbitrary in Theorem 2, but here, T and S are dual (cf.[9], p.223).

3 Idempotent uninorms

In this section we present the characterization of an idempotent uninorm given in [12].

Theorem 3 ([12]). *Let* $e\in[0,1]$. *Operation* $U\in\mathcal{U}(e)$ *is idempotent iff there exists a decreasing function* $g:[0,1]\to[0,1]$ *with* $g(e)=e$, $g(x)=0$ *for all* $x>g(0)$, $g(x)=1$ *for all* $x<g(1)$, *satisfying for* $x\in[0,1]$

$$inf\{y:g(y)=g(x)\}\le g^2(x)\le\sup\{y:g(y)=g(x)\} \tag{6}$$

and such that for all $x,y\in[0,1]$

$$U(x,y)=\begin{cases} \min(x,y) & if \quad y<g(x) \ or \\ & \qquad y=g(x) \ and \ x<g^2(x) \\ \max(x,y) & if \quad y>g(x) \ or \\ & \qquad y=g(x) \ and \ x>g^2(x) \\ x \ or \ y & if \quad y=g(x) \ and \ x=g^2(x) \end{cases}. \tag{7}$$

Moreover, U *is commutative beyond the points* $(x,g(x))$ *such that* $x=g^2(x)$.

Example 3. Let $e\in(0,1)$. For the uninorms U_1 and U_2 given by (2) with $U_1=\min$ and $U_2=\max$ on $A(e)$ respectively, and $T=\min$, $S=\max$, the corresponding functions are of the form

$$g_{U_1}(x)=\begin{cases} 1, & \text{if } \ x\in[0,e) \\ e, & \text{if } \ x\in[e,1] \end{cases},$$

$$g_{U_2}(x)=\begin{cases} e, & \text{if } \ x\in[0,e] \\ 0, & \text{if } \ x\in(e,1] \end{cases}.$$

4 Operations locally internal in $A(e)$

Using additional assumption on $U\in\mathcal{U}(e)$ in $A(e)$ we get a representation of U by certain formula.

Lemma 3 ([3]). *If operation* U *is increasing with the neutral element* $e\in(0,1)$ *and* $U(x,y)\in\{x,y\}$ *in* $A(e)$, *then the formula*

$$g_U(x)=\begin{cases} \sup\{y:\ U(x,y)=x\}, & if \quad x\le e, \\ \inf\{y:\ U(x,y)=x\}, & if \quad x>e, \end{cases} \tag{8}$$

gives a decreasing function $g_U:[0,1]\to[0,1]$ *with fixed point* e, *satisfying for* $x\in[0,1]$ *equation* (6) *and such that in* $\overline{A(e)}$ *operation* U *is given by* (7). *Moreover,* U *is commutative in* $\overline{A(e)}$ *beyond the points* $(x,g(x))$ *such that* $x=g^2(x)$.

Unfortunately, values of U on $[0,e]^2 \cup [e,1]^2$, in above lemma, are undetermined.

Example 4. Let $e \in (0,1)$. For the uninorm U given in Theorem 2 which is given by minimum in $A(e)$, the corresponding function is the same as g_{U_1} in Example 3 In this case the corresponding t-norm and t-conorm may be arbitrary.

Moreover, for a given function g, not all t-norms and t-conorms leads to uninorms.

Example 5 ([3])**.** Let $g(x) = 1 - x$, $x \in [0,1]$. Operation $U : [0,1]^2 \to [0,1]$ given by

$$U(x,y) = \begin{cases} 2xy, & \text{if } x,y \in [0,\frac{1}{2}] \\ \max(x,y), & \text{if } y > g(x) \\ \min(x,y), & \text{otherwise} \end{cases}$$

is not associative. Indeed, for $x = y = \frac{1}{4}$, $z = \frac{13}{16}$ we have $U\left(U\left(\frac{1}{4},\frac{1}{4}\right),\frac{13}{16}\right) = U\left(\frac{1}{8},\frac{13}{16}\right) = \frac{1}{8}$, $U\left(\frac{1}{4},U\left(\frac{1}{4},\frac{13}{16}\right)\right) = U\left(\frac{1}{4},\frac{13}{16}\right) = \frac{13}{16}$. Therefore $U(U(x,y),z) \neq U(x,U(y,z))$. This shows, that t-norm and t-conorm in (2) cannot be arbitrary, if we want to construct a uninorm.

The partial answer, to the question, to what condition should fulfil underlying t-norm and t-conorm gives the next theorem.

Theorem 4 ([3])**.** *Let $g : [0,1] \to [0,1]$ be a decreasing involution ($g^2 = id$). The formula*

$$U(x,y) = \begin{cases} T(x,y), & x,y \in [0,e] \\ S(x,y), & x,y \in [e,1] \\ \min(x,y), & x < e < y \le g(x) \\ & \text{or } y \le g(x) < e < x \\ \max(x,y), & x < e < g(x) < y \\ & \text{or } g(x) < y < e < x \end{cases} \tag{9}$$

gives a uninorm, iff $T = \min$ and $S = \max$.

For that reason we give the properties of the function g_U, in case of operation U locally internal on $\overline{A(e)}$ and we describe its influence on the structure of the operation U (and underlying operations).

Lemma 4 ([4])**.** *Let $e \in (0,1)$, $U \in \mathcal{U}(e)$ be locally internal on $A(e)$ and $g = g_U$. If g is a strictly decreasing and continuous function on $(a,b) \subset [0,1]$ and $g((a,b)) = (c,d)$, then g is strictly decreasing and continuous function on (c,d), and $g^2(x) = x$ for $x \in (a,b) \cup (c,d)$ (cf. Figure 2 left).*

Lemma 5 (cf. [4])**.** *Let $e \in (0,1)$, $U \in \mathcal{U}(e)$ be locally internal on $A(e)$ and $g = g_U$. If $s \in [0,1]$ and p, q are define by*

$$p = \begin{cases} \lim\limits_{x \to s^+} g(x), & \text{if } s < 1 \\ 0, & \text{if } s = 1 \end{cases}, \tag{10}$$

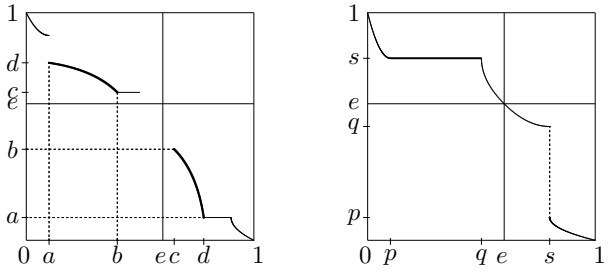


Figure 2: Subinterval (a,b) with strictly decreasing g (left) and the point of discontinuity of the function g (right)

$$q = \begin{cases} \lim\limits_{x \to s^-} g(x), & \text{if } s > 0 \\ 1, & \text{if } s = 0 \end{cases}, \tag{11}$$

then $g(x) = s$ for $x \in (p,q)$.

Remark 1. *The interval (p,q) in the above lemma may be nonempty only if $s = 0$ or $s = 1$ or s is a point of discontinuity of the function g.*

Lemma 6 (cf. [4])**.** *Let $e \in (0,1)$, $U \in \mathcal{U}(e)$ be locally internal on $A(e)$ and $g = g_U$. Let $s \in [0,1]$ and $B = \{x : g(x) = s\}$. If card $B \ge 2$ then $p := \inf B < \sup B =: q$, $p \le g(s) \le q$ and (10), (11) hold. Moreover, if $s \in (0,1)$, then s is a point of discontinuity of the function g.*

Lemma 7 ([4])**.** *Let $U \in \mathcal{U}(e)$ be locally internal on $A(e)$. If $g = g_U$ is strictly decreasing and continuous function on $(a,b) \subset [0,e]$ and $g((a,b)) = (c,d)$, then (cf. Figure 3)*

$$U = \min \quad on \quad [0,b] \times (a,e] \cup (a,e] \times [0,b], \tag{12}$$

$$U = \max \quad on \quad [e,d) \times [c,1] \cup [c,1] \times [e,d). \tag{13}$$

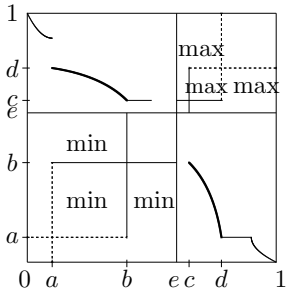


Figure 3: Subinterval (a,b) with strictly decreasing g

Lemma 8 ([4])**.** *Let $U \in \mathcal{U}(e)$ be locally internal on $A(e)$ and s be a point of discontinuity of the function $g = g_U$.*
If $s \in [0,e]$, then

$$U = \min \quad on \quad [0,s] \times (s,e] \cup (s,e] \times [0,s]. \tag{14}$$

If $s \in [e,1]$, then

$$U = \max \quad on \quad [e,s) \times [s,1] \cup [s,1] \times [e,s). \tag{15}$$

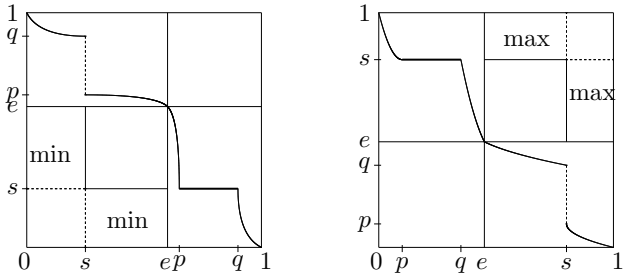


Figure 4: Dependence of the values of operation U under discontinuity points of the function g_U

Lemma 9 ([4]). *Let $U \in \mathcal{U}(e)$ be locally internal on $A(e)$ and s be a point of discontinuity of the function $g = g_U$, p, q be given by* (10), (11) *and $g(s) = r$, $p < r < q$.*
If $s \in [0, e)$, then $r > e$ and (cf. Figure 5, left)

$$U = \max \quad on \quad (p, r) \times [r, q) \cup [r, q) \times (p, r). \quad (16)$$

If $s \in (e, 1]$, then $r < e$ and (cf. Figure 5, right)

$$U = \min \quad on \quad (p, r] \times (r, q) \cup (r, q) \times (p, r]. \quad (17)$$

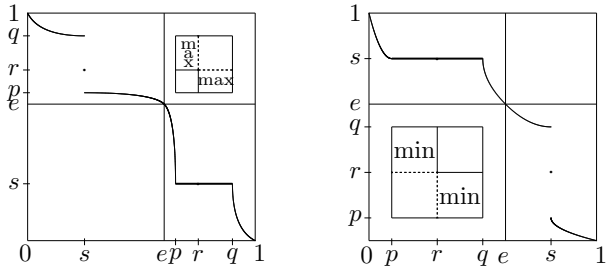


Figure 5: The values of the operation U in the points of discontinuity of the function g

If $s \in [0, 1]$ is a point such that $p < q$ (see (10), (11)) then we obtain one or two intervals, i.e. (p, q) or (p, r), (r, q), where $r = g(s)$. Moreover $g|_{(p,q)} = const$.

By this observation we obtain the fact that operation U on $[0, e]^2$ is an ordinal sum of $\{([a_i, b_i], T_i)\}_{i \in A_1}$, (see Lemma 1), such that function g is constant on $(a_i, b_i) \subset [0, e]$ for all $i \in A_1$. The operation U on $[e, 1]^2$ is an ordinal sum of $\{([c_j, d_j], S_j)\}_{j \in A_2}$, such that function g is constant on $(c_j, d_j) \subset [e, 1]$ for all $j \in A_2$.
Let us denote $B_1 = \{[a_i, b_i] : i \in A_1\}$, $B_2 = \{[c_j, d_j] : j \in A_2\}$.

So, by above, we have

Lemma 10. *Let $e \in (0, 1)$, $U \in \mathcal{U}(e)$ be locally internal in $A(e)$ and $g = g_U$. If $(a_i, b_i) \in B_1$, then U is internal operation in $(a_i, b_i]$ or in $[a_i, b_i]$.*

By this lemma we define operation T_i on the set $[a_i, b_i]$ as follows

$$T_i(x, y) = \begin{cases} a_i & \text{if } x = y = a_i, \\ U(x, y) & \text{otherwise,} \end{cases} \quad x, y \in [a_i, b_i]. \quad (18)$$

This operation is associative, increasing with zero element a_i and fulfills condition $T_i \leq \min$.

Similarly for $[e, 1]^2$ we have

Lemma 11. *Let $e \in (0, 1)$, $U \in \mathcal{U}(e)$ be locally internal in $A(e)$ and $g = g_U$. If $(c_j, d_j) \in B_2$, then U is internal operation in $[c_j, d_j)$ or in $[c_j, d_j]$.*

By this lemma we define operation S_j on the set $[c_j, d_j]$ as follows

$$S_j(x, y) = \begin{cases} d_j, & \text{if } x = y = d_j, \\ U(x, y) & \text{otherwise,} \end{cases} \quad x, y \in [c_j, d_j]. \quad (19)$$

This operation is associative, increasing with zero element d_j and fulfills condition $S_j \geq \max$.

The next theorem gives a description of operations belonging to the set $\mathcal{U}(e)$ and locally internal on $A(e)$.

Theorem 5 ([4]). *Let $e \in (0, 1)$. If operation $U \in \mathcal{U}(e)$ is locally internal on $A(e)$ then*

- *there exists a decreasing function $g : [0, 1] \to [0, 1]$ with fixed point e such that* (6) *holds, $g(x) = 0$ for all $x > g(0)$, $g(x) = 1$ for all $x < g(1)$ and U is given in $A(e)$ by formula* (7)*;*
- *$U|_{[0,e]^2}$ is an ordinal sum of $\{([a_i, b_i], T_i)\}_{i \in A_1}$, such that $(a_i, b_i) \subset [0, e] \setminus \{g(x) : x \in [e, 1]\}$ for all $i \in A_1$;*
- *$U|_{[e,1]^2}$ is an ordinal sum of $\{([c_j, d_j], S_j)\}_{j \in A_2}$, such that $(c_j, d_j) \subset [e, 1] \setminus \{g(x) : x \in [0, e]\}$ for all $j \in A_2$.*

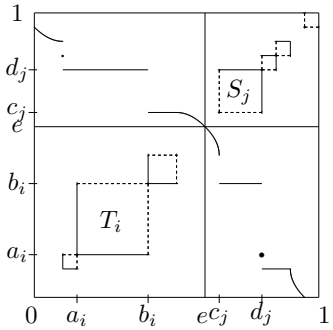


Figure 6: Structure of operation belonging to the class $\mathcal{U}(e)$ and locally internal on $A(e)$

But the converse theorem is not true

Example 6. Let $g : [0,1] \to [0,1]$ be given by
$$g(x) = \begin{cases} 1 & \text{if } x \in [0, 0.25) \\ 0.5 & \text{if } x \in [0.25, 0.5] \\ 0.25 & \text{if } x \in (0.5, 1] \end{cases}.$$
Operation given by
$$U(x,y) = \begin{cases} 0, & \text{if } x, y \in [0, 0.25] \\ \max(x,y), & \text{if } y > g(x) \\ \min(x,y), & \text{otherwise} \end{cases}$$
is increasing, has the neutral element $\frac{1}{2}$, but it is not associative, e.g. $U(\frac{1}{4}, U(\frac{3}{4}, \frac{1}{8})) = U(\frac{1}{4}, \frac{1}{8}) = 0$, $U(U(\frac{1}{4}, \frac{3}{4}), \frac{1}{8}) = U(\frac{3}{4}, \frac{1}{8}) = \frac{1}{8}$. Therefore, $U(\frac{1}{4}, U(\frac{3}{4}, \frac{1}{8})) \neq U(U(\frac{1}{4}, \frac{3}{4}), \frac{1}{8})$. This shows, that these conditions are not necessary.

We obtain only partial results

Lemma 12. *Let $e \in (0,1)$. If $g : [0,1] \to [0,1]$ is decreasing function such that $g(e) = e$ and (6) holds then U given by (7) on $\overline{A(e)}$ is increasing on $\overline{A(e)}$ and e is the neutral element of operation U.*

Proof. First we show, that U is increasing on $[0,e] \times [e,1]$. By (7) we see that $U(a,b) \in \{a,b\}$ for every point $(a,b) \in \overline{A(e)}$. Let $x < y \leq e \leq z$.
By monotonicity of the function g we get $g(x) \geq g(y)$.
If $z < g(x)$, then $U(x,z) = \min(x,z) = x$, $U(y,z) \in \{y,z\}$, thus $U(x,z) \leq U(y,z)$.
If $z > g(x)$, then $z > g(y)$ and $U(x,z) = z = U(y,z)$.
If $z = g(x)$ then we consider two cases, namely:
- if $z = g(y)$, then $g^2(x) = g^2(y)$ and one of the inequalities hold
$g^2(x) \leq x$, then $g^2(y) < y$, $U(y,z) = \max(y,z) = z$ and $U(x,z) \in \{x,z\}$, thus $U(x,z) \leq U(y,z)$,
$x < g^2(x)$, then $U(x,z) = \min(x,z) = x$ and $U(y,z) \in \{y,z\}$, thus $U(x,z) \leq U(y,z)$,
- if $z > g(y)$, then $U(y,z) = \max(y,z) = z$ and $U(x,z) \in \{x,z\}$, thus $U(x,z) \leq U(y,z)$.
It means that $U(x,z) \leq U(y,z)$ for all x, y, z such that $x < y \leq e \leq z$. Now we will prove the monotonicity of U on $[0,e] \times [e,1]$ with respect to the second variable. Let $z \leq e \leq x < y$.
If $x < g(z)$, then $U(z,x) = \min(z,x) = z$, $U(z,y) \in \{z,y\}$, thus $U(z,x) \leq U(z,y)$.
If $g(z) \leq x < y$, then $U(z,y) = \max(z,y) = y$, $U(z,x) \in \{z,x\}$, thus $U(z,x) \leq U(z,y)$.
The proof of monotonicity of U on $[e,1] \times [0,e]$ is similar. Therefore U is increasing on $\overline{A(e)}$.

By the fact that $g(e) = e$, and (7) we have $U(e,x) = \min(e,x) = x$ for $x < e$, $U(e,x) = \max(e,x) = x$ for $x > e$ and $U(e,e) = \min(e,e) = e$ or $U(e,e) = \max(e,e) = e$, thus $U(e,x) = x$ for $x \in [0,1]$. For $x < e$, we have $g(x) \geq e$. If $g(x) > e$, then by (7) we have $U(x,e) = \min(x,e) = x$. If $g(x) = e$, then by the fact that $g(e) = e$ we have $x < g^2(x)$ and by (7) we have $U(x,e) = U(x,g(x)) = \min(x,e) = x$. Similarly, we obtain $U(x,e) = x$ for $x > e$. Therefore, e is the neutral element of operation U. □

Lemma 13. *Let $e \in (0,1)$. If $g : [0,1] \to [0,1]$ is decreasing and fulfills the condition $g(e) = e$, $g(x) = 0$ for $x > g(0)$, $g(x) = 1$ for $x < g(1)$ and (6) holds, then U given by (7) on $\overline{A(e)}$ is commutative on $\overline{A(e)}$ beyond the points $(x, g(x))$ such that $x = g^2(x)$.*

To obtain complete results we need a few facts

Lemma 14. *Let $e \in (0,1)$, $U \in \mathcal{U}(e)$ be locally internal in $A(e)$ and $g = g_U$, $i \in A_1$, $(a_i, b_i) \in B_1$ and T_i be given by (18). Thus $g((a_i,b_i)) = \{s_i\}$, where $s_i \in [e,1]$ and:*

(i) $g(b_i) = s_i$ and ($U(x,s_i) = x$ on $(a_i,b_i]$ or $U(x,s_i) = s_i$ on $(a_i,b_i]$), or $U(b_i,s_i) = s_i$ and T_i has the neutral element b_i.

(ii) If T_i has the zero element divisors (where the zero element is equal to a_i), then $g(a_i) = s_i$ and ($U(x,s_i) = s_i$ on $[a_i,b_i)$ or $U(x,s_i) = x$ on $[a_i,b_i)$).

(iii) If b_i is the zero element divisor of operation T_i, then $g(a_i) = g(b_i) = s_i$ and ($U(x,s_i) = s_i$ on $[a_i,b_i]$ or $U(x,s_i) = x$ on $[a_i,b_i]$).

Proof. The condition $g((a_i,b_i)) = \{s_i\}$ follows directly from definition B_1.
If $U(b_i,s_i) = \min(b_i,s_i) = b_i$, then by (7) we have $g(b_i) \geq s_i$. By monotonicity of the function g and condition $g(x) = s_i$ for $x \in (a_i,b_i)$ we have $g(b_i) = s_i$. Moreover for $x \in (a_i,b_i)$ we have $U(x,s_i) \leq U(b_i,s_i) = b_i$ and $U(x,s_i) \in \{x,s_i\}$, thus $U(x,s_i) = x = \min(x,s_i)$ for $x \in (a_i,b_i]$.
If $U(b_i,s_i) = \max(b_i,s_i) = s_i$, then by (7) we have $g(b_i) \leq s_i$.
If $g(b_i) < s_i$, then for $z \in [e,1]$ such that $g(b_i) < z < s_i$ by associativity and commutativity U (without the graph of the function g) we have $U(x,b_i) \leq x < b_i$ and
$$U(x,b_i) = U(U(x,b_i),z) = U(x,U(b_i,z)) = U(x,z) = x,$$
$$U(b_i,x) = U(b_i,U(x,z)) = U(b_i,U(z,x)) = U(U(b_i,z),x) = U(z,x) = x$$
for $x \in (a_i,b_i)$. It means, that $U(x,b_i) = U(b_i,x) = x$.
If $g(b_i) = s_i$, then we have two possibilities:
$U(x,s_i) = \max(x,s_i) = s_i$ on $(a_i,b_i]$ or $U(x,s_i) = \min(x,s_i) = x$ on (a_i,b_i).
If $U(x,s_i) = \min(x,s_i) = x$ for $x \in (a_i,b_i)$, then
$U(x,b_i) = U(U(x,b_i),s_i) = U(x,U(b_i,s_i)) = U(x,s_i) = x$,
$U(b_i,x) = U(b_i,U(s_i,x)) = U(U(b_i,s_i),x) = U(s_i,x) = x$.
Moreover, $U(b_i,b_i) \leq b_i$ and $U(x,b_i) = x$ for $x \in (a_i,b_i)$. Therefore, $U(b_i,b_i) = b_i$. Similarly,

$U(a_i, b_i) \leq \min(a_i, b_i) = a_i$ and by Lemmas 7, 8 and 9, $U(x, a_i) = U(a_i, x) = a_i$ for $x \in (a_i, b_i)$, thus $U(a_i, b_i) = a_i$. It means that b_i is the neutral element of the operation T_i.

If T_i has zero element divisors, then there exist $x, y \in (a_i, b_i)$ such that $U(x, y) = a_i$. Suppose, that condition *(ii)* does not hold. By monotonicity of the function g we have two possibilities:

if $g(a_i) > s_i$, then taking $z \in (s_i, g(a_i))$, we have $U(U(x, y), z) = U(a_i, z) = a_i$, $U(x, U(y, z)) = U(x, z) = z$, which contradicts to the associativity of U.

if $g(a_i) = s_i$, $U(a_i, s_i) = a_i$ and $U(t, s_i) = \max(t, s_i)$ for $t \in (a_i, b_i)$, then $U(U(x, y), s_i) = U(a_i, s_i) = a_i$, $U(x, U(y, s_i)) = U(x, s_i) = s_i$, which contradicts to the associativity of U. So, we have *(ii)*.
In order to prove *(iii)*, first notice that, if b_i is the zero divisor of operation T_i, then b_i can not be the neutral element of operation T_i. By *(i)* we have that $g(b_i) = s_i$ and ($U(x, s_i) = \min(x, s_i)$ on $(a_i, b_i]$ or $U(x, s_i) = \max(x, s_i)$ on $(a_i, b_i]$).
The operation T_i has zero element divisors and by *(ii)* we have $g(a_i) = s_i$ and ($U(x, s_i) = s_i$ on $[a_i, b_i)$ or $U(x, s_i) = x$ on $[a_i, b_i)$).
By above results we have $g(a_i) = g(b_i) = s_i$ and ($U(x, s_i) = s_i$ on $[a_i, b_i]$ or $U(x, s_i) = x$ on $[a_i, b_i]$). □

Similarly, we obtain the following result

Lemma 15. *Let $e \in (0, 1)$, $U \in \mathcal{U}(e)$ be locally internal in $A(e)$. If $g = g_U$, $j \in A_2$, $(c_j, d_j) \in B_2$ and S_j be given by* (19)*, then $g((c_j, d_j)) = \{s_j\}$ for $s_j \in [0, e]$ and*

(i') $g(c_j) = s_j$ *and (*$U(x, s_j) = s_j$ *on* $[c_j, d_j)$ *or* $U(x, s_j) = x$ *on* $[c_j, d_j)$*)*
or $U(c_j, s_j) = s_j$ *and* S_j *has the neutral element* c_j*.*

(ii') *If* S_j *has zero divisors (where the zero element is equal* d_j*), then* $g(d_j) = s_j$ *and (*$U(x, s_j) = x$ *on* $(c_j, d_j]$ *or* $U(x, s_j) = s_j$ *on* $(c_j, d_j]$*).*

(iii') *If* c_j *is a zero divisor of operation* S_j*, then* $g(c_j) = g(d_j) = s_j$
and ($U(x, s_j) = x$ *on* $[c_j, d_j]$ *or* $U(x, s_j) = s_j$ *on* $[c_j, d_j]$*).*

Lemma 16. *Let $e \in (0, 1)$ and $g : [0, 1] \to [0, 1]$ be a decreasing function such that $g(e) = e$. Thus the conditions* (6) *and $g(x) = 0$ for all $x > g(0)$, $g(x) = 1$ for all $x < g(1)$ are equivalent to the following conditions*

- *If g is a strictly decreasing and continuous function on $(a, b) \subset [0, 1]$ and $g((a, b)) = (c, d)$, then g is a strictly decreasing and continuous function on (c, d), and $g^2(x) = x$ for $x \in (a, b) \cup (c, d)$.*
- *If $s \in [0, 1]$ is a point of discontinuity of the function g and p, q are given by* (10) *and* (11) *then $g(x) = s$ for $x \in (p, q)$.*
- *Let $s \in [0, 1]$ and $B = \{x : g(x) = s\}$. If card $B \geq 2$ then $p := \inf B < \sup B =: q$, $p \leq g(s) \leq q$ and* (10) *and* (11) *hold.*

Theorem 6. *Let $e \in (0, 1)$ and the following conditions hold:*

- *there exists a decreasing function $g : [0, 1] \to [0, 1]$ such that $g(e) = e$, $g(x) = 0$ for $x > g(0)$, $g(x) = 1$ for $x < g(1)$,* (6) *holds and U is given by* (7) *in $\overline{A(e)}$,*
- *$U|_{[0,e]^2}$ is an ordinal sum of $\{([a_i, b_i], T_i)\}_{i \in A_1}$, such that $(a_i, b_i) \subset [0, e] \setminus \{g(x) : x \in [e, 1]\}$ and T_i are increasing, associative operations fulfilling $T_i \leq \min$ for all $i \in A_1$,*
- *conditions (i), (ii), (iii) from Lemma 14 are fulfilled,*
- *$U|_{[e,1]^2}$ is an ordinal sum of $\{([c_j, d_j], S_j)\}_{j \in A_2}$, such that $(c_j, d_j) \subset [e, 1] \setminus \{g(x) : x \in [0, e]\}$ and S_j are increasing, associative operations fulfilling $S_j \geq \max$ for all $j \in A_2$,*
- *conditions (i'), (ii'), (iii') from Lemma 15 are fulfilled.*

Thus $U \in \mathcal{U}(e)$ and it is locally internal on $A(e)$.

Proof. By (7) we obtain that operation U is locally internal on $\overline{A(e)}$. Using Lemma 12 we can see that U is increasing and has the neutral element e.
Let us define the operation $\widetilde{U} : [0, 1]^2 \to [0, 1]$ as follows

$$\widetilde{U} = \begin{cases} U & \text{on } A(e) \\ \min & \text{on } [0, e]^2 \\ \max & \text{on } [e, 1]^2 \end{cases}$$

This operation is increasing, associative, idempotent, has the neutral element e. Moreover, $\widetilde{U}$ is commutative beyond the points $(x, g(x))$ such that $x = g^2(x)$ (see Theorem 3). Furthermore, $U = \widetilde{U}$ on the set $[0, 1]^2 \setminus D$ where $D = \left\{\bigcup_{i \in A_1} (a_i, b_i]^2 \cup \bigcup_{j \in A_2} [c_j, d_j)^2\right\}$. Directly from this fact we obtain that U is commutative on the set $[0, 1]^2 \setminus D$ beyond the points $(x, g(x))$ such that $x = g^2(x)$.
To prove the associativity of U we check the equality

$$U(U(x, y), z) = U(x, U(y, z)) \tag{20}$$

only on sets $\{x, y, z\}$ such that at least one of the pairs belonging to the set $\{x, y, z\} \times \{x, y, z\}$ belongs also to

the set D.
On the rest of the domain the associativity of U follows directly from the associativity of $\widetilde{U}$.
As a result, we need to consider only the case $x, y \in [a_i, b_i]$ for $i \in A_1$. Thus, there exists $s_i \in [e, 1]$ such that $g((a_i, b_i)) = \{s_i\}$. Moreover, $g(s_i) = a_i$ or $g(s_i) = b_i$.
Let us consider the following cases:

1. Let $x, y \in (a_i, b_i)$ and $U(x, y) \neq a_i$. Thus
   a) if $z < s_i$, then
   $U(U(x, y), z) = U(x, y) = U(x, U(y, z))$,
   $U(U(x, z), y) = U(x, y) = U(x, U(z, y))$,
   $U(U(z, x), y) = U(x, y) = U(z, U(x, y))$,
   $U(U(z, y), x) = U(y, x) = U(z, U(y, x))$,
   $U(U(y, x), z) = U(y, x) = U(y, U(x, z))$,
   $U(U(y, z), x) = U(y, x) = U(y, U(z, x))$,
   b) if $z > s_i$, then
   $U(U(x, y), z) = z = U(x, z) = U(x, U(y, z))$,
   $U(U(x, z), y) = U(z, y) = U(y, z) = z = U(x, z) = U(x, U(y, z)) = U(x, U(z, y))$,
   $U(U(z, x), y) = U(U(x, z), y) = U(z, y) = U(y, z) = z = U(U(x, y), z) = U(z, U(x, y))$,
   $U(U(z, y), x) = U(U(y, z), x) = U(z, x) = U(x, z) = z = U(U(y, x), z) = U(z, U(y, x))$,
   $U(U(y, x), z) = z = U(y, z) = U(y, U(x, z))$,
   $U(U(y, z), x) = U(z, x) = U(x, z) = z = U(y, z) = U(y, U(x, z)) = U(y, U(z, x))$,
   c) if $z = s_i$, then $U = \min$ on $(a, b) \times \{s_i\}$ or $U = \max$ on $(a, b) \times \{s_i\}$ and we get the case a) or b).
2. Let $x = a_i$, $y \in (a_i, b_i]$. Thus $U = \widetilde{U}$ on the set $\{x, y, z\}$ and condition (20) holds.
3. Let $x = b_i$, $y \in [a_i, b_i]$ and $U(x, y) \neq a_i$. Thus by *(i)* this case leads to the cases 1) or 2).
4. Let $x, y \in (a_i, b_i)$ and $U(x, y) = a_i$. Thus by *(ii)* this case is similar to the case 1).
5. Let $x = b_i$, $y \in (a_i, b_i]$ and $U(x, y) = a_i$. Thus and by *(iii)* this case is similar to the case 1).
6. Let $x = y = a_i$. Thus for $U(x, y) = a_i$ we have case 2) and for $U(x, y) < a_i$ we have $a_i = b_j$ for some $j \in A_1$ and we have cases 4) or 5) for the interval $[a_j, b_j]$.

Similarly, we may obtain the associativity condition (20) for $x < e < y, z$. Therefore, U is associative.
It means that U is the operation from the class $\mathcal{U}(e)$. □

5 Conclusion

In this paper we characterize uninorms with the neutral element $e \in (0, 1)$ such that $U(x, y) \in \{x, y\}$ on $A(e)$. We see that triangular norm and conorm from Theorem 1 have the form of the ordinal sum. An open problem is to characterize all uninorms, without the assumption, that $U(x, y) \in \{x, y\}$ in $A(e)$. In the case of the continuous triangular norm and the continuous triangular conorm given by (2) the characterization can be found in [6].

References

[1] E. Czogała, J. Drewniak, Associative monotonic operations in fuzzy set theory, Fuzzy Sets and Systems 12 (1984), 249–269.

[2] B. De Baets, Idempotent uninorms, European J. Oper. Res. 118 (1999), 631–642.

[3] J. Drewniak, P. Drygaś, On a class of uninorms. Internat, J. Uncertain. Fuzziness Knowledge-Based Syst. 10, Suppl. (2002), 5–10.

[4] P. Drygaś, Discusion of the structure of uninorms, Kybernetica 41 (2005), 213- -226.

[5] P. Drygaś, Remarks about idempotent uninorms. J. Electrical Engin. 57, (2006), 92–94.

[6] J. Fodor, B. De Baets, T. Calvo, Characterization of uninorms with given underlying t-norms and t-conorms (200x) (submitted).

[7] J. Fodor, R. Yager, A. Rybalov, Structure of uninorms, Internat. J. Uncertain. Fuzziness Knowledge-Based Syst. 5 (1997), 411–427.

[8] S. Jenei, A note on the ordinal sum theorem and its consequence for the construction of triangular norm, Fuzzy Sets and Systems 126 (2002), 199–205.

[9] E. P. Klement, R. Mesiar, E. Pap, Triangular norms, Kluwer Acad. Publ., Dordrecht, 2000.

[10] Y.-M. Li, Z.-K. Shi, Remarks on uninorm aggregation operators, Fuzzy Sets and Systems 114 (2000), 377–380.

[11] J. Martin, On a theorem of Czogała and Drewniak (1984), Proc. EUROFUSE PM'01, Granada 2001, 49–54.

[12] J. Martin, G. Mayor, J. Torrens, On locally internal monotonic operations, Fuzzy Sets and Systems 137 (2003), 27–42.

[13] R. Yager, A. Rybalov, Uninorm aggregation operators, Fuzzy Sets and Systems 80 (1996), 111–120.

Session 6

Fuzzy Logic 2

The logic of perfect MV-algebras

L. P. Belluce
Department of Mathematics,
British Columbia University,
Vancouver, B.C. Canada.
belluce@math.ubc.ca

A. Di Nola
DMI
University of Salerno,
Salerno, Italy
adinola@unisa.it

B. Gerla
DICOM
University of Insubria,
Varese, Italy
brunella.gerla@uninsubria.it

Abstract

In this paper we summarize results on perfect MV-algebras and we show that the class of first order Łukasiewicz formulas valid in all interpretations over perfect MV-chains is axiomatizable.

Keywords: MV-algebras, perfect algebra, first order Łukasiewicz logic.

1 Introduction

The class of MV-algebras arises as algebraic counterpart of the infinite valued Łukasiewciz sentential calculus, as Boolean algebras did with respect to the classical propositional logic. Due to the non-idempotency of the MV-algebraic conjunction, unlike Boolean algebras, MV-algebras can be non-archimedean and can contains elements x such that $x \odot \ldots \odot x$ (n times) is always greater than zero, for any $n > 0$ (here $\odot$ denotes the Łukasiewicz conjunction). In general, there are MV-algebras which are not semisimple, i.e. the intersection of their maximal ideals (the radical of A) is different from $\{0\}$. Non-zero elements from the radical of A are called *infinitesimals*. Perfect MV-algebras are those MV-algebras generated by their infinitesimal elements or, equivalently, generated by their radical. Hence perfect MV-algebras can be seen as extreme examples of non-archimedean MV-algebras.

An important example of a perfect MV-algebra is the subalgebra S of the Lindenbaum algebra L of first order Łukasiewicz logic generated by the classes of formulas which are valid when interpreted in $[0, 1]$ but non-provable. Hence perfect MV-algebras are directly connected with the very important phenomenon of incompleteness in Łukasiewicz first order logic (see [10], [2]).

As it is well known, MV-algebras form a category which is equivalent to the category of abelian lattice ordered groups (ℓ-groups, for short) with strong unit [9]. This makes the interest in MV-algebras relevant outside the realm of logic. Let us denote by Γ the functor implementing this equivalence. In particular each perfect MV-algebra is associated with an abelian ℓ-group with a strong unit. But, more has been proved. Namely the category of perfect MV-algebras is equivalent to the category of abelian ℓ-groups, see ([7], Theorem 3.5, p.420). Let us denote by $\mathcal{D}$ the functor implementing this equivalence. Hence $\mathcal{D}$ maps functorially each perfect MV-algebra to an abelian ℓ-group and vice versa, without the help of a strong unit. Here a curious remark has to be made. Indeed the Γ functor maps a non-equational class, the category of abelian ℓ-groups with strong unit, to an equational class, the variety of MV-algebras. On the other hand, the functor $\mathcal{D}$ maps an equational class, the category of abelian ℓ-groups, to a non-equational class, the category of perfect MV-algebras. However, as a consequence of using the functor $\mathcal{D}$, a surprising result was proved showing the equivalence between the category of perfect MV-algebras with a distinguished generator of the radical, whose morphisms preserve the distinguished element, and the category of all MV-algebras, cite4.

Perfect MV-algebras do not form a variety and contains non-simple subdirectly irreducible MV-algebras. It is worth stressing that the variety generated by all perfect MV-algebras is also generated by a single MV-chain, actually the MV-algebra C, defined by Chang in [5]. The MV-algebra C is therefore a prototypical perfect MV-algebra, and we shall see it plays a crucial role in the theory of perfect MV-algebras, as well as in the variety it generates.

In this paper we recount basic properties of perfect MV-algebras and we adapt results from [2] for establishing that the class of first order Łukasiewicz formulas valid on all interpretations over perfect MV-chains is axiomatizable.

2 Perfect MV-algebras

A structure $A = (A, 0, 1, \neg, \odot, \oplus)$ is an MV-algebra iff A satisfies the following equations:

1. $(x \oplus y) \oplus z = x \oplus (y \oplus z)$;
2. $x \oplus y = y \oplus x$;
3. $x \oplus 0 = x$;
4. $x \oplus 1 = 1$;
5. $\neg 0 = 1$;
6. $\neg 1 = 0$;
7. $x \odot y = \neg(\neg x \oplus \neg y)$;
8. $\neg(\neg x \oplus y) \oplus y = \neg(\neg y \oplus x) \oplus x$.

On A two new operations $\vee$ and $\wedge$ are defined as follows: $x \vee y = \neg(\neg x \oplus y) \oplus y$ and $x \wedge y = \neg(\neg x \odot y) \odot y$. The structure $(A, \vee, \wedge, 0, 1)$ is a bounded distributive lattice. We shall write $x \leq y$ iff $x \wedge y = x$. We say that an MV-algebra A is an MV-chain when, as a lattice, A is linearly ordered. Boolean algebras are just the MV-algebras obeying the additional equation $x \odot x = x$. For any MV-algebra A we denote by $B(A) = \{x \in A \mid x \odot x = x\}$ the biggest Boolean algebra contained in A. We write nx instead of $x \oplus ... \oplus x$ (n-times) and x^n instead of $x \odot ... \odot x$ (n-times). The least integer for which $nx = 1$ is called the *order* of x. When such an integer exists, we denote it by $ord(x)$ and say that x has *finite order*, otherwise we say that x has *infinite order* and write $ord(x) = \infty$.

Example. The unit interval of real numbers $[0, 1]$ with operations defined by $x \oplus y = \min\{1, x + y\}$, $x \odot y = \max\{0, x + y - 1\}$, and $\neg x = 1 - x$ is an MV-algebra. We shall refer to this MV-algebra as $[0, 1]$.

An *ideal* of an MV-algebra A is a non-empty subset I of A which is closed under $\oplus$ and such that if $x \leq y$ and $y \in I$ then $x \in I$. A *prime* ideal P of A is an ideal of A such that $x \wedge y$ implies $x \in P$ or $y \in P$. An ideal M of A is called *maximal* if $M \subseteq I$ implies $I = A$ or $I = M$, where I an ideal of A. Let M be a maximal ideal of A, then we say that M is *supermaximal* if $A/M \cong \{0, 1\}$. The set of all prime ideals of A shall be denoted by $Spec(A)$. For each element x of an MV-algebra A the set

$$id(x) = \{y \in A \mid y \leq nx, \text{ for some } n > 0\}$$

is the ideal of A generated by x. Each proper ideal is contained in a maximal ideal.

As MV-algebras form an equational class, the notions of MV-isomorphism, quotient, subalgebra, product, etc., are just the particular cases of the corresponding universal algebraic notions.

The intersection of all maximal ideals, the *radical* of A, will be denoted by $Rad(A)$.

An MV-algebra A such that $Rad(A) = 0$ is called *semisimple*. An MV-algebra A is called *simple* if and only if A is non trivial and $\{0\}$ is its only proper ideal. Every simple MV-algebra is isomorphic to a subalgebra of $[0, 1]$, (see, e.g., [6], Theorem 3.5.1, p. 70). Every non-zero element of a non trivial MV-algebra A has finite order if and only if A is simple.

If for every element x of the MV-algebra A there is an integer n such that nx is idempotent then A will be called *hyperarchimedean*. For all unexplained MV-algebraic notions we refer the reader to [6].

Definition 1. An MV-algebra A is *local* if A has a unique maximal ideal. The class of all local MV-algebras will be denoted by **Local**.

It is also well known that for each $x \in A$, x is a member of a proper ideal, hence a maximal ideal, if and only if the order of x is ∞.

It turns out that an MV-algebra A is local if and only if for every $x \in A$, $ord(x) < \infty$ or $ord(\neg x) < \infty$.

Definition 2. An MV-algebra A is called *perfect* if for every nonzero element $x \in A$ $ord(x) = \infty$ if and only if $ord(\neg x) < \infty$. The class of all perfect MV-algebras will be denoted by **Perfect**.

It is clear that the class of all MV-algebras is a variety, here denoted by MV. For any subclass K of elements from MV, $V(K)$ shall denote the subvariety of MV generated by K. If K has just one element A then we also write $V(A)$ for $V(K)$.

Definition 3. Chang's MV-algebra is defined on the set

$$C = \{0, c, \ldots, nc, \ldots, 1 - nc, \ldots, 1 - c, 1\}$$

by the following operations (consider $0 = 0c$):
$x \oplus y =$

- $(m + n)c$ if $x = nc$ and $y = mc$
- $1 - (m - n)c$ if $x = 1 - nc$ and $y = mc$ and $0 < n < m$
- $1 - (n - m)c$ if $x = nc$ and $y = 1 - mc$ and $0 < m < n$
- 1 otherwise;

$$\neg x = \begin{cases} 1 - nc & \text{if } x = nc \\ nc & \text{if } x = 1 - nc \end{cases}$$

C is a linearly ordered MV-algebra, $ord(nc) = \infty$ and $ord(1 - nc) < \infty$ for every n. So C is a perfect MV-algebra.

Definition 4. A proper ideal P of an MV-algebra A is called *perfect* if and only if for every $a \in A$, $a^n \in P$ for some $n \in \omega$ if and only if $(\neg a)^m \notin P$ for all $m \in \omega$.

P is a perfect ideal if and only if A/P is a perfect MV-algebra.

Proposition 5. *The following hold:*

(1) The only finite perfect MV-algebra is $\{0, 1\}$*;*

(2) Every nonzero element in a perfect MV-algebra $A \neq B(A)$ *generates a subalgebra isomorphic to the Chang MV-algebra* C*.*

(3) Linearly ordered MV-algebras in $V(C)$ *are all perfect MV-chains;*

(4) $V(\mathbf{Perfect}) = V(C)$*,*

(5) $\mathbf{Perfect} = V(C) \cap \mathbf{Local}$*;*

(6) $A \in V(C)$ *iff for every* $x \in A$*,* $2x^2 = (2x)^2$*;*

(7) A *is perfect iff* $A = alg(Rad(A))$*;*

(8) A *is perfect iff* $A = Rad(A) \cup \neg(Rad(A))$*; further* $x \in Rad(A)$ *iff* $ord(x) = \infty$*;*

(9) **Perfect** *is closed under homomorphic images and subalgebras;*

(10) A *is perfect iff any proper ideal of* A *is perfect iff* $\{0\}$ *is a perfect ideal.*

Proof. Properties (1) and (2) are immediate consequences of order of elements in a perfect MV-algebra.

The proof of properties $(3), (4), (5), (6)$ and of $(7), (8), (9), (10)$ can be found in [7] and in [4] respectively. □

Now we are going to show that the class of perfect MV-algebras is first order definable.

Consider the following well formed formulas in the first order language of MV-algebras containing the equality relation as predicate symbol, operations of MV-algebras as functional symbols and 0 and 1 as constant symbols. Further denote by $\&$, OR and $\Rightarrow$ the classical propositional connectives. Let σ be the wff $(\forall x)(x^2 \oplus x^2 = (x \oplus x)^2)$ and τ be the wff $(\forall x)(x^2 = x \Rightarrow (x = 0 \text{ OR } x = 1))$. Then,

Proposition 6. *Let* A *be an MV-algebra then the following are equivalent:*

(i) A *is perfect;*

(ii) A *satisfies* $\sigma \,\&\, \tau$*.*

Proof. $(i) \Rightarrow (ii)$. Let A be a perfect MV-algebra. Then by Proposition 5(4), $A \in V(\mathbf{C})$ hence A satisfies the equation σ satisfied by C. Further, by Proposition 5(2), A cannot contain idempotent elements different from 0 and 1.

$(ii) \Rightarrow (i)$. Assume A to be an MV-algebra satisfying the formula $\sigma \,\&\, \tau$, so $x^2 = x$ implies $x = 0$ or $x = 1$. Hence $B(A) = \{0, 1\}$ and $A \in V(C)$. Hence, by Chang Theorem (see [6], Theorem 1.3.3, p.20), Theorem 5.1 in ([7] p.424) and Corollary 5.2 in ([7] p.425) (cfr Proposition 5(3)):

$$A \hookrightarrow \prod_{P \in Spec(A)} A/P$$

where, for every $P \in Spec(A)$, A/P is a perfect MV-chain. Assume A is not perfect. Then there is $z \in A$ such that $z \notin Rad(A) \cup \neg(Rad(A))$. Therefore, there are $P, Q \in Spec(A)$ such that

$$z/P \in Rad(A/P) \qquad and \qquad z/Q \in \neg Rad(A/Q).$$

So we get $2(z^2/P) = 0/P$ and $2(z^2/Q) = 1/Q$, that is $2z^2 \in B(A) \setminus \{0, 1\}$, contradicting $B(A) = \{0, 1\}$. □

3 Łukasiewicz logic

In this section we consider first-order Łukasiewicz logic. We shall use definitions and axioms from [2] and [3], but apart from some technicalities, such definitions are analogous to those for example in [8].

Definition 7. Let $\mathcal{L}$ be a language containing symbols of variables $v_0, v_1, \ldots$, logical symbols $\rightarrow, \neg$; predicate symbols $R_0, R_1, \ldots$; a quantifier symbol $\exists$; improper symbols $(,\ )$; and a function $d : \mathbb{N} \rightarrow \mathbb{N}$, $\mathbb{N} = \{0, 1, 2, ...\}$.

The set of well-formed formulas of $\mathcal{L}$, WFF, is defined as usual, as follows: atomic formulas, $R_n(v_{i_1}, v_{i_2}, ..., v_{i_{d(n)}})$ are in WFF. If $\alpha, \beta \in WFF$ so are $(\alpha \rightarrow \beta)$ and $\neg\alpha$. If $\alpha \in WFF$ and x is a variable then $(\exists x)\alpha$ is in WFF.

Definition 8. Let A be an MV-algebra and X a nonempty set. An $\{A, X\} - model$ is a system $\langle A, X, (F_n)_{n\in\mathbb{N}}\rangle$ such that for each $n \in \mathbb{N}$ there is a function $F_n : X^{d(n)} \rightarrow A$.

An $\{A, X\} - model$ is *linear* if A is an MV-chain, is *canonical* if $A \subseteq [0, 1]$. Given an $\{A, X\} - model$ $\langle A, X, (F_n)_{n\in\mathbb{N}}\rangle$, an *assignment* is a function $f : Var \rightarrow X$, with $Var = \{v_0, v_1, \ldots\}$, i.e., Var is the set of variables of $\mathcal{L}$. If f is an assignment, $v \in Var$, $x \in X$, then f_{vx} is the assignment

$$f_{vx}(v_i) = \begin{cases} f(v_i), & \text{if } v_i \neq v \\ x, & \text{if } v_i = v. \end{cases}$$

If $S \subseteq A$ we define

$$\sum S = \begin{cases} \text{least upper bound of } S \text{ in } A \text{ (if it exists)} \\ A, \quad \text{otherwise.} \end{cases}$$

Given $\mathfrak{M} = \langle A, X, (F_n)_{n\in\mathbb{N}}\rangle$ we assign values to each $\alpha \in WFF$. We define, therefore a function

$$Val(\alpha, \mathfrak{M}, f) : WFF \to A \cup \{A\},$$

where f is an $\mathfrak{M}$-assignment inductively defined by the following conditions:

(1) $Val(R_n(v_{i_1}, v_{i_2}, ..., v_{i_{d(n)}}), \mathfrak{M}, f) = F_n(f(v_{i_1}), f(v_{i_2}), ..., f(v_{i_{d(n)}}))$;

(2) Assuming Val defined for $\alpha \in WFF$, then

$$Val(\neg\alpha, \mathfrak{M}, f) =$$

$$= \begin{cases} \neg Val(\alpha, \mathfrak{M}, f) & \text{if } Val(\alpha, \mathfrak{M}, f) \in A \\ A, \quad \text{otherwise;} \end{cases}$$

(3) Assuming Val defined for $\alpha, \beta \in WFF$, then

$$Val(\alpha \to \beta, \mathfrak{M}, f) =$$

$$= \begin{cases} \neg Val(\alpha, \mathfrak{M}, f) \oplus Val(\beta, \mathfrak{M}, f), \\ \quad \text{provided } Val(\alpha, \mathfrak{M}, f), Val(\beta, \mathfrak{M}, f) \in A \\ A, \quad \text{otherwise;} \end{cases}$$

(4) $Val((\exists v)\alpha, \mathfrak{M}, f) = \sum_{x\in X} Val(\alpha, \mathfrak{M}, f_{vx})$.

Call an $\mathfrak{M}$-assignment an *interpretation* if $Val(\alpha, \mathfrak{M}, f) \in A$ for all $\alpha \in WFF$.

Axioms of first order Łukasiewicz logic (see for example [2] or [8]), includes all axioms for Łukasiewicz propositional logic. We can hence define the Lindenbaum MV-algebra L of first order Łukasiewicz logic as the quotient of the MV-algebra of all formulas with respect to the congruence given by provable equivalence of formulas. An element from L shall be denoted by $[\alpha]$, with α a formula from $\mathcal{L}$.

Let $\beta_0, \beta_1, ...$ be an enumeration of all $\beta \in WFF$ of the form $\beta_n = (\exists v_i)\alpha_n$. For each $n, m \geq 0$, let $\beta(n, m)$ be defined from α_n by the following steps:

1) let v_j be the first variable not occurring in α_n;

2) replace all bounded occurrences of v_m in α_n by v_j;

3) replace all free occurrences of v_i in α_n by v_m.

An ideal I in L is said to *preserve sums* if for any n, if $[\beta(n, m)] \in I$ for all m, then $\sum_{m\in\mathbb{N}}[\beta(n, m)] \in I$.

The following results are from [2] and [3].

Theorem 9. *Let I be an ideal of L preserving sums. Then the map $\alpha \in Form \mapsto [\alpha]/I \in L/I$ is an interpretation over L/I.*

We know from [2] that the class of first order formulas valid in all interpretations over all linearly ordered MV-algebras is axiomatizable, while the class of formulas valid in interpretations over $[0, 1]$ is not (see [10]). We shall study now the class of first order formulas valid in all interpretation over perfect linearly ordered MV-algebras.

Add to **L** the axiom schema, $2\alpha^2 \leftrightarrow (2\alpha)^2$ and denote the resulting system by $\mathbf{L_p}$. Then we have,

Proposition 10. *A wff α is provable in $\mathbf{L_p}$, iff α is valid on all perfect MV-chains.*

Proof. The proof is based on the fact that for any formula β with $ord([\beta]) = \infty$ there exists prime ideal P of L containing $[\beta]$ and preserving sums, hence by Theorem 9 we can define an assignment of formulas into L/P mapping $[\beta]$ into 0. Since the axioms of $\mathbf{L_p}$ include the axioms of **L**, the same is true for the Lindenbaum algebra L_p of $\mathbf{L_p}$. But the latter algebra is now in the variety $V(C)$ generated by Chang's algebra C (Proposition 5(6)). Consequently each linearly ordered subalgebra of L_p is perfect (Proposition 5(3)). From this it follows that if α is valid on all perfect MV-chains and not provable in $\mathbf{L_p}$, then in L_p, $[\alpha] \neq 1$, hence $\neg[\alpha] \neq 0$. Thus for some prime ideal P we have $\neg[\alpha] \notin P$ and so in L_p/P we have $[\alpha]/P \neq 1$. But L_p/P is a perfect linearly ordered MV-algebra, hence α is not equal to one in a perfect MV-chain, in contradiction with our hypothesis. □

References

[1] L. P. Belluce, *Semisimple Algebras of Infinite-valued Logic* , Canadian J. Math., **38**(1986), 1356-1379.

[2] L. P. Belluce, C. C. Chang, *A weak completeness theorem for infinite valued predicate logic*, J. Symbolic Logic, **28**(1963), 43-50.

[3] L. P. Belluce, A. Di Nola, *The MV-algebra of First Order Łukasiewicz Logic*, Tatra Mt. Math. Publ., **27**(2003), 7-22.

[4] L. P. Belluce, A. Di Nola, A. Lettieri *Local MV-algebras* , Rend. Circ. Mat. Palermo. (2) **42**(1993), 347-361.

[5] C. C. Chang, *Algebraic Analysis of Many-Valued Logics*, Trans. Amer. Math. Soc., **88**(1958), 467-490.

[6] R. Cignoli, I. D'Ottaviano, D. Mundici *Algebraic Foundations of Many-valued Reasoning* , Trends in Logic, Vol. 7, Kluwer, Dordrecht, 2000.

[7] A. Di Nola, A. Lettieri *Perfect MV-algebras are Categorically Equivalent to Abelian ℓ-Groups*, Studia Logica, **88**(1958), 467-490.

[8] P. Hájek, *Metamathematics of fuzzy logic*. Kluwer 1998.

[9] D. Mundici, *Interpretation of AF C^*-Algebras in Lukasiewicz Sentential Calculus*, J. Funct. Analysis **65**, (1986), 15–63.

[10] B. Scarpellini, *Die Nichaxiomatisierbarkeit des unendlichwertigen Pradikatenkalkulus von Lukasiewicz*, Journ. Symb. Log.,**27**(1962),159-170.

An algebraic approach to states on MV-algebras

Tommaso Flaminio
Dip. di Matematica e Scienze Informatiche,
Pian dei Mantellini 44, 53100 Siena,
Italy
flaminio@unisi.it

Franco Montagna
Dip. di Matematica e Scienze Informatiche,
Pian dei Mantellini 44, 53100 Siena,
Italy
montagna@unisi.it

Abstract

In this paper we will introduce the variety of *state MV-algebras* by enlarging the language of MV-algebras with an unary operator σ, and by adding equations ensuring σ to satisfy some basic properties of a *state* on an MV-algebra in the sense of [4]. The main result of this paper says us that there is a standard way to define a state MV-algebra starting from a state on an MV-algebra and vice-versa, each state MV-algebra defines a state on an MV-algebra in its usual meaning.

Keywords: MV-algebra, state, tensor product, coherence.

1 Introduction

MV-algebras was introduced by Chang in [2] in order to show Łukasiewicz logic (see [3]) to be standard complete, i.e. complete with respect to evaluations of propositional variables in the real unit interval $[0, 1]$. Chang's Completeness Theorem says us that the real unit interval $[0, 1]$ is the standard domain of interpretation for the propositional variables of Łukasiewicz logic. Recall in fact that the prototypical example of an MV-algebra is the standard one

$$[0, 1]_{MV} = \langle [0, 1], \oplus, \neg, 0\rangle,$$

where for all $x, y \in [0, 1]$, $x \oplus y = \min\{1, x + y\}$, and $\neg x = 1 - x$.

MV-algebras provide a generalization of Boolean algebras in the sense that each Booelan algebra is an MV-algebra satisfying the equation $x \oplus x = x$. Therefore as to generalize probability on Boolean algebras, Mundici introduced in [4] the notion of state on MV-algebras. Roughly speaking a state on an MV-algebra $\mathcal{M}$ is a (finitely) additive (in the sense of $\mathcal{M}$) function $s : \mathcal{M} \to [0, 1]$ mapping the bottom element of $\mathcal{M}$ in 0. Hence each state s maps element of a given MV-algebra $\mathcal{M}$ into the domain of the standard MV-algebra $[0, 1]_{MV}$, for this reason it makes sense to consider a state as an additional operator on $\mathcal{M}$.

Following this idea, in this paper we will enlarge the language of MV-algebras by means of an unary operation σ and we will add, to those ones of MV-algebras, a class of equations ensuring σ to satisfy some basic properties of states. The resulting algebraic structure will be called *state MV-algebra*. These algebras will be introduced in Section 3, while in Section 5 we will show how each state MV-algebra induce (at least) a state (in its the usual meaning) on its MV-reduct, while each state on an MV-algebra $\mathcal{M}$ defines a state MV-algebra whose MV-reduct is the tensor product (see [5]) of $[0, 1]_{MV}$ by $\mathcal{M}$. In Section 6 we will apply state MV-algebras to study the probabilistic coherence problem on infinite-valued events. We will end with some remarks and open problems.

2 Preliminaries

Definition 2.1 *An MV-algebra is an algebra* $\mathcal{M} = \langle M, \oplus, \neg, 0, 1\rangle$ *such that:*

- $\langle M, \oplus, 0\rangle$ *is a commutative monoid, having* 0 *as neutral element,*
- $\neg\neg x = x$ *and* $x \oplus 1 = 1$ *hold for all* $x \in M$*,*
- $x \oplus \neg(x \oplus \neg y) = y \oplus \neg(y \oplus \neg x)$ *is satisfied for all* $x, y \in M$*.*

In any MV-algebra $\mathcal{M}$ one can define the following further operations:

$$x \to y = \neg x \oplus y,\ x \ominus y = \neg(x \to y),\ x \odot y = \neg(\neg x \oplus \neg y),\ x \leftrightarrow y = (x \to y) \odot (y \to x),$$
$$x \vee y = (x \to y) \to y,\ x \wedge y = \neg(\neg x \vee \neg x).$$

Finally any MV-algebra $\mathcal{M}$ is equipped with the order relation: $x \leq y$ iff $x \to y = 1$. An MV-algebra $\mathcal{M}$

is said *linearly ordered* iff the order relation defined as above is linear. Following the tradition, a linearly ordered MV-algebra will be also called an *MV-chain*.

Definition 2.2 ([4]) *A* state *on an MV-algebra $\mathcal{M}$ is a mapping $s : M \to [0,1]$ such that:*

- $s(1) = 1$,
- *if $x \odot y = 0$ for $x, y \in M$, then $s(x \oplus y) = s(x) + s(y)$.*

It is quite easy to see that each state s on an MV-algebra $\mathcal{M}$ is monotone and satisfies $s(0) = 0$. A state $s : M \to [0,1]$ is said *faithfull* if $s(x) > 0$, whenever $x \neq 0$.

3 State MV-algebras

Definition 3.1 *A* state MV-algebra *is a pair $(\mathcal{M}, \sigma)$ such that $\mathcal{M}$ is an MV-algebra and σ is a unary operation on M satisfying:*
(1) $\sigma(1) = 1$,
(2) $\sigma(\neg x) = \neg\sigma(x)$,
(3) $\sigma(x \oplus y) = \sigma(x) \oplus \sigma(y \ominus (x \odot y))$.
(4) $\sigma(\sigma(x) \oplus \sigma(y)) = \sigma(x) \oplus \sigma(y)$.

Obviously the class of state MV-algebras forms a variety. In the following Lemma we will prove some properties of state MV-algebras.

Lemma 3.2 *In a state MV-algebra $(\mathcal{M}, \sigma)$ the following conditions hold:*
(a) $\sigma(0) = 0$.
(b) *If $x \leq y$, then $\sigma(x) \leq \sigma(y)$.*
(c) $\sigma(x \oplus y) \leq \sigma(x) \oplus \sigma(y)$, *and if $x \odot y = 0$, then $\sigma(x \oplus y) = \sigma(x) \oplus \sigma(y)$.*
(d) $\sigma(x \ominus y) \geq \sigma(x) \ominus \sigma(y)$ *and if $y \leq x$, then $\sigma(x \ominus y) = \sigma(x) \ominus \sigma(y)$.*
(e) *Letting $d(x,y) = (x \ominus y) \oplus (y \ominus x)$, we have that $d(\sigma(x), \sigma(y)) \leq \sigma(d(x,y))$.*
(f) $\sigma(x) \odot \sigma(y) \leq \sigma(x \odot y)$. *Thus if $x \odot y = 0$, then $\sigma(x) \odot \sigma(y) = 0$.*
(g) $\sigma(\sigma(x)) = \sigma(x)$.
(h) *The image $\sigma(M)$ of M under σ is the domain of a MV-subalgebra of $\mathcal{M}$.*

Proof. (a) By (1) and (2).

(b) If $x \leq y$, then $y = x \oplus (y \ominus x)$, therefore since $x \odot (y \ominus x) = 0$, substituting in (2) $(y \ominus x)$ for y, we get $\sigma(y) = \sigma(x \oplus (y \ominus x)) = \sigma(x) \oplus \sigma(y \ominus x) \geq \sigma(x)$.

(c) By (b), $\sigma(y) \geq \sigma(y \ominus (x \odot y))$, therefore $\sigma(x \oplus y) = \sigma(x) \oplus \sigma(y \ominus (x \odot y)) \leq \sigma(x) \oplus \sigma(y)$. If $(x \odot y) = 0$, then $\sigma(x \oplus y) = \sigma(x) \oplus \sigma(y \ominus (x \odot y)) = \sigma(x) \oplus \sigma(y)$.

(d) Using (2) and (c), we obtain:

$\sigma(x \ominus y) = \sigma(\neg(\neg x \oplus y)) = \neg\sigma(\neg x \oplus y) \geq \neg(\sigma(\neg x) \oplus \sigma(y)) =$

$= \neg(\neg\sigma(x) \oplus \sigma(y)) = \sigma(x) \ominus \sigma(y)$. Moreover, if $y \leq x$, then $\neg x \odot y = 0$, and by (c), $\sigma(x \ominus y) = \neg\sigma(\neg x \oplus y) = \neg(\sigma(\neg x) \oplus \sigma(y)) = \sigma(x) \ominus \sigma(y)$.

(e) Since $(x \ominus y) \odot (y \ominus x) = 0$, by (d) and (c) we get

$\sigma(d(x,y)) = \sigma(x \ominus y) \oplus \sigma(y \ominus x) \geq (\sigma(x) \ominus \sigma(y)) \oplus (\sigma(y) \ominus \sigma(x)) = d(\sigma(x), \sigma(y))$.

(f) We have $x \odot y = x \ominus \neg y$, therefore by (e), $\sigma(x \odot y) \geq \sigma(x) \ominus \sigma(\neg y) = \sigma(x) \ominus (\neg\sigma(y)) = \sigma(x) \odot \sigma(y)$. If $x \odot y = 0$, then $0 = \sigma(x \odot y) \geq \sigma(x) \odot \sigma(y)$, therefore $\sigma(x \odot y) = 0$.

(g) By (a), $\sigma(0) = 0$, therefore using (4) we get $\sigma(\sigma(x)) = \sigma(\sigma(x) \oplus \sigma(0)) = \sigma(x) \oplus \sigma(0) = \sigma(x)$.

(h) By (g), the range of σ is constituted by all the fixed points of σ, therefore it is sufficient to prove that the set of such fixed points is closed under $\oplus$ and under $\neg$. Now closure under $\oplus$ follows from (4). As regards to closure under $\neg$, using (2) and (g), we get $\sigma(\neg\sigma(x)) = \neg(\sigma(\sigma(x)) = \neg(\sigma(x))$, and the claim follows. ■

Examples. (a) We start from a trivial example. Let $\mathcal{M}$ be any MV-algebra and σ be the identity on M. Then $(\mathcal{M}, \sigma)$ is a state MV-algebra.

(b) A slightly less silly example. Let σ be an idempotent endomorphism of an MV-algebra $\mathcal{M}$ (For example, we may take $\mathcal{M}$ to be a non-trivial ultrapower of the standard MV-algebra $[0,1]_{MV}$ and σ to be the standard part function). Then $(\mathcal{M}, \sigma)$ is a state MV-algebra.

(c) This is in our opinion a sufficiently general example. Let $\mathcal{M}$ be the algebra of all continuous and piecewise linear functions with real coefficients from $[0,1]^n$ into $[0,1]$. Then $\mathcal{M}$, with the point-wise application of Łukasiewicz $\oplus$ and $\neg$, forms an MV-algebra. Now let for $f \in M$, $\sigma(f) = \int_{[0,1]^n} f$. Then $(\mathcal{M}, \sigma)$ is a state MV-algebra. Note that $(\mathcal{M}, \sigma)$ is simple, therefore it is subdirectly irreducible, but it is not totally ordered. Although rather general, this algebra satisfies the quasi equation $\sigma(x) = 0$ implies $x = 0$, which is not valid in general.

4 Congruences, ideals, subdirectly irreducible elements

A σ-ideal of a state MV-algebra $(\mathcal{M}, \sigma)$ is an MV-ideal closed under σ. Since $\sigma(0) = 0$, the congruence classes of 0 are σ-ideals. Conversely, given a

σ ideal J of a state MV-algebra $(\mathcal{M},\sigma)$, the relation $\theta_J = \{(x,y) : d(x,y) \in J\}$ is a congruence of $(\mathcal{M},\sigma)$. That θ_J is a congruence of $\mathcal{M}$ is well-known. Moreover, if $(x,y) \in \theta_J$, then $d(x,y) \in J$, therefore $\sigma(d(x,y)) \in J$, as J is a σ-ideal. Moreover by Lemma 3.2, (e) $d(\sigma(x),\sigma(y)) \leq \sigma(d(x,y))$. Thus $d(\sigma(x),\sigma(y)) \in J$, and $(\sigma(x),\sigma(y)) \in \theta_J$. Thus θ_J is a congruence. Therefore:

Theorem 4.1 *The congruence lattice of a state MV-algebra is isomorphic to the lattice of its σ-ideals.*

Lemma 4.2 *The σ-ideal J_a generated by an element a of a state MV-algebra $(\mathcal{M},\sigma)$ is $\{x : \exists n \in \omega(x \leq n(a \oplus \sigma(a)))\}$.*

Proof. Let $I = \{x : \exists n \in \omega(x \leq n(a \oplus s(a)))\}$. By the definition of σ-ideal, we have that every element of I must be in J_a. Thus $I \subseteq J_a$. For the opposite inclusion, it suffices to show that I is an ideal which contains a. That $a \in I$ is clear, and the proof that I is an MV-ideal is routine. We show that I is closed under σ. By Lemma 3.2, (c) and (g), if $x \leq n(a \oplus \sigma(a))$, then $\sigma(x) \leq n(\sigma(a) \oplus \sigma(a)) \leq 2n(a \oplus \sigma(a))$, therefore $\sigma(x) \in J$. This ends the proof. ∎

Theorem 4.3 *(a) If $(\mathcal{M},\sigma)$ is a subdirectly irreducible, then $\sigma(M)$ is linearly ordered.*
(b) If $(\mathcal{M},\sigma)$ satisfies the quasi identity $\sigma(x) = 0$ implies $x = 0$, then $(\mathcal{M},\sigma)$ is subdirectly irreducible iff $\sigma(M)$ is subdirectly irreducible (as an MV-algebra).

Proof. (a) Let J be the minimum non-trivial σ-ideal of $\mathcal{M}$. By its minimality, J is mono-generated, therefore there is $a > 0$ such that $J = J_a = \{x : \exists n \in \omega(x \leq n(a \oplus \sigma(a)))\}$. Suppose that for some $\sigma(x),\sigma(y) \in \sigma(M)$ one has $\sigma(x) \not\leq \sigma(y)$ and $\sigma(y) \not\leq \sigma(x)$. Then the ideals $J_{\sigma(x)\ominus\sigma(y)}$ and $J_{\sigma(y)\ominus\sigma(x)}$ generated by $\sigma(x) \ominus \sigma(y)$ and by $\sigma(y) \ominus \sigma(x)$ respectively are non trivial, therefore they both contain J. Since $\sigma(M)$ is closed under σ, by Lemma 4.2, there is $n \in \omega$ such that $a \leq n((\sigma(x) \ominus \sigma(y)))$ and $a \leq n(\sigma(y) \ominus \sigma(x))$ (here we are using in a crucial way the fact that σ is the identity on $\sigma(M)$ and that $\sigma(M)$ is a subalgebra of A, cf Lemma 3.2 (g) and8 (h)). Hence $a \leq n((\sigma(x) \ominus \sigma(y)) \wedge (\sigma(y) \ominus \sigma(x))) = 0$, and a contradiction has been reached.

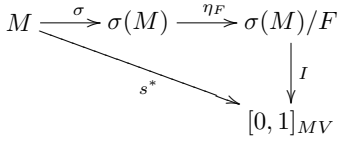
(b) If $\sigma(x) = 0$ implies $x = 0$, then the intersection of a non-trivial σ ideal J of $(\mathcal{M},\sigma)$ with $\sigma(M)$ is a non-trivial ideal of $\sigma(M)$. Moreover any ideal of $\sigma(M)$ is closed under σ, therefore in $\sigma(M)$ every ideal is a σ ideal. Hence if J is the minimum non-trivial σ ideal of $\sigma(M)$, then $J \cap \sigma(M)$ is the minimum non-trivial ideal of $\sigma(M)$. Indeed if I is another non-trivial ideal of $\sigma(M)$, then the σ ideal H of $(\mathcal{M},\sigma)$ generated by I contains J, and $I = H \cap \sigma(A) \supseteq J \cap \sigma(A)$. It follows that $J \cap \sigma(M)$ is minimal. Hence if $(\mathcal{M},\sigma)$ is subdirectly irreducible, then so is $\sigma(M)$.

Conversely, if I is the minimum non-trivial ideal of $\sigma(M)$, then the σ ideal K of $(\mathcal{M},\sigma)$ generated by I is the minimum non-trivial σ-ideal of $(\mathcal{M},\sigma)$. Indeed, if J is another non-trivial σ ideal of $(\mathcal{M},\sigma)$, then $J \cap \sigma(M) \supseteq K \cap \sigma(M) = I$. Thus K contains the σ ideal generated by I, i.e., $K \supseteq J$ and J is minimal. Hence if $\sigma(M)$ is subdirectly irreducible, then so is $(\mathcal{M},\sigma)$. ∎

Remark. As already noted, not all subdirectly irreducible state MV-algebras are linearly ordered. Moreover, linearly ordered state MV-algebras satisfy the identity $\sigma(x \vee y) = \sigma(x) \vee \sigma(y)$, which is not valid in general. Thus the variety of state MV-algebras is not generated by its totally ordered members. However, by Theorem 4.3, it is generated as a quasivariety by its member $(\mathcal{M},\sigma)$ such that $\sigma(M)$ is totally ordered.

5 State MV-algebras and states on MV-algebras

Let $(\mathcal{M},\sigma)$ be a state MV-algebra, and let F be a maximal MV-filter over $\sigma(M)$. Also, let for $x \in M$, $s^*(x) = \sigma(x)/F$. Note that $\sigma(M)/F$ (considered as an MV-algebra) is a simple MV-algebra, therefore it embeds into the standard MV-algebra $[0,1]_{MV}$ on $[0,1]$. Thus s^* can be regarded as a map from $\mathcal{M}$ into $[0,1]_{MV}$.



η_F being the quotient map of $\sigma(M)$ in the simple MV-algebra $\sigma(M)/F$, and I being the embedding of $\sigma(M)/F$ into the standard MV-algebra $[0,1]_{MV}$.

Theorem 5.1 *s^* is a state on $\mathcal{M}$.*

Proof. That $s^*(1) = 1$ is clear. Now let $x,y \in M$ such that $x \odot y = 0$. Then by Lemma 3.2 (c) and (f), we get that $s^*(x \oplus y) = s^*(x) \oplus s^*(y)$ and $\sigma(x) \odot \sigma(y) = 0$. Hence $s^*(x) \oplus s^*(y) = s^*(x) + s^*(y)$, and $s^*(x \oplus y) = s^*(x) + s^*(y)$. Thus s^* is a state. ∎

Now let s be a state on a MV-algebra $\mathcal{M}$. Consider the tensor product $\mathcal{T} = [0,1]_{MV} \otimes \mathcal{M}$, (see Mundici's paper [5]). Every element of $\mathcal{T}$ can be written as $\alpha \otimes a$ with $\alpha \in [0,1]$ and $a \in M$. (Warning: different expressions may denote the same element, for instance, for any $\alpha \in [0,1]$ and $a \in M$, both $\alpha \otimes 0$ and $0 \otimes a$ denote the bottom element of $\mathcal{T}$. We also warn the reader

that $\otimes$ is not an operation on $\mathcal{T}$, but a function from $[0,1] \times M$ into $\mathcal{T}$). Moreover, for $a, b, c, d \in M$ and $\alpha, \beta, \gamma, \delta \in [0,1]$, the following conditions are satisfied:

(i) $\mathcal{T}$ is a MV-algebra.

(ii) $(\alpha \oplus \beta) \otimes 1 = (\alpha \otimes 1) \otimes (\beta \otimes 1)$ and $(1 \otimes (a \oplus b)) = (1 \otimes a) \oplus (1 \otimes b)$.

(iii) $\neg\alpha \otimes 1 = \neg(\alpha \otimes 1)$ and $1 \otimes \neg a = \neg(1 \otimes a)$.

(iv) The maps $\alpha \mapsto \alpha \otimes 1$ and $a \mapsto 1 \otimes a$ are embeddings of $[0,1]_{MV}$ and of $\mathcal{M}$ respectively into $\mathcal{T}$. Thus we will identify α with $\alpha \otimes 1$ and a with $1 \otimes a$.

(v) If $\beta \odot \gamma = 0$ and $b \odot c = 0$, then $(\beta + \gamma) \otimes a = (\beta \otimes a) \oplus_T (\gamma \otimes a)$ and $\alpha \otimes (b \oplus c) = (\alpha \otimes c) \oplus_T (\alpha \otimes d)$.

(vi) $\alpha \otimes (a \ominus d) = (\alpha \otimes a) \ominus_T (\alpha \otimes d)$ and $(\alpha \ominus \delta) \otimes a = (\alpha \otimes a) \ominus_T (\delta \otimes a)$.

The top element of $\mathcal{T}$ is $1 = 1 \otimes 1$, and the bottom element is $0 \otimes 0 = 0 \otimes a = \alpha \otimes 0$.

We define an operation σ^+ on $\mathcal{T}$ as follows: $\sigma^+(\alpha \otimes a) = \alpha \cdot s(a)$ (note that $\alpha \cdot s(a) \in [0,1]_{MV}$ which is a subalgebra of $\mathcal{T}$).

Theorem 5.2 *σ^+ is well-defined and $(\mathcal{T}, \sigma^+)$ is a state MV-algebra.*

Proof. Let $\mathcal{M}^d$ be the divisible MV-algebra generated by $\mathcal{M}$. Hence every element of $\mathcal{M}^d$ has the form qa for some rational $q \in [0,1]$ and $a \in M$. There is only one way to extend $s : \mathcal{M} \to [0,1]$ to a state s^d on $\mathcal{M}^d$ into $[0,1]$. In fact it must be that $s^d(qa) = qs(a)$. Let now $\mathcal{T}_\mathbb{Q}$ be the MV-algebra $[0,1]_{MV} \cap \mathbb{Q} \otimes \mathcal{M}$. The following holds:

Claim 5.3 *$\mathcal{T}_\mathbb{Q}$ is isomorphic to $\mathcal{M}^d$ via the isomorphism $f : q \otimes a \mapsto qa$.*

Proof. (Claim) Let $\beta : [0,1] \cap \mathbb{Q} \times M \to M^d$ be defined as $\beta(q, a) = qa$. Then β is a bimorphism (see [5]). In fact the following holds:

- $\beta(1,1) = 1 \cdot 1 = 1$,
- $\beta(0, a) = \beta(q, 0) = 0$,
- $\beta(q, a_1 \vee a_2) = q(a_1 \vee a_2) = qa_1 \vee qa_2 = \beta(q, a_1) \vee \beta(q, a_2)$, and analogously the other cases.
- If $a_1 \odot a_2 = 0$, then
  - $\beta(q, a_1) \odot \beta(q, a_2) = qa_1 \odot qa_2 \leq a_1 \odot a_2 = 0$.
  - $\beta(q, a_1 \oplus a_2) = q(a_1 \oplus a_2)$. But $\cdot$ distributes on $\oplus$, and thus $\beta(q, a_1 \oplus a_2) = \beta(q, a_1) \oplus \beta(q, a_2)$

  Symmetrically the other cases when $q_1 \odot q_2 = 0$ hold.

Therefore β is a bimorphism. Now from [5] (Theorem 3.2 pp. 234–235), there exists a unique homomorphism $\lambda : [0,1]_{MV} \cap \mathbb{Q} \otimes \mathcal{M} \to \mathcal{M}^d$. Moreover it is easy to see that λ is a bijection. In fact if $\lambda(q \otimes a) \neq \lambda(r \otimes b)$, then $qa \neq rb$, and thus $q \otimes a \neq r \otimes b$, then λ is injective. To prove that λ is surjective is trivial. In fact notice that for each $qa \in A^d$, at least $\lambda(q \otimes a) = qa$. Thus the claim holds. ■

Therefore $q \otimes a$ can be identified with $qa = 1 \otimes qa$, and $\sigma^+(q \otimes a) = qs(a) = s^d(qa)$. It follows that if for some rationals q, r and for some $a, b \in M$ one has $r \otimes a = q \otimes b$, then $ra = qb$ and $\sigma^+(r \otimes a) = rs(a) = s^d(ra) = s^d(qb) = qs(b) = \sigma^+(q \otimes b)$. Thus σ^+ is well-defined on $\mathcal{M}^d$.

Now the restriction of σ^Q on $\mathcal{T}_\mathbb{Q}$ is a state on $\mathcal{T}_\mathbb{Q}$. In fact $\sigma^Q(1 \otimes 1) = 1s(1) = 1$, and if $(q \otimes a) \odot (r \otimes b) = 0$, then $qa \odot rb = 0$. Then $\sigma^Q((q \otimes a) \oplus (r \otimes b)) = s^d(qa \oplus rb) = s^d(qa) \oplus s^d(rb) = \sigma^Q(q \otimes a) \oplus \sigma^Q(r \otimes b)$. Finally, σ^+ extends σ^Q by continuity:

$$\sigma^+(\alpha \otimes a) = \sup\left\{\sigma^Q(r \otimes a) : r \in \mathbb{Q}; r \leq \alpha\right\},$$

and therefore the following hold:

- Clearly for each $\alpha \in [0,1]$ and $a \in M$, $\alpha \otimes a = \sup\{r \otimes a \mid r \in \mathbb{Q}, r \leq \alpha\}$. In fact for each fixed $a \in M$, the partial function $\otimes^a_\mathbb{Q} : q \mapsto q \otimes a = qa$ is a continuos map over $\mathbb{Q}$. Therefore for each $\alpha \in [0,1]$ and $a \in M$, $\alpha \otimes a = \sup\{r \in \mathbb{Q} \mid r \leq \alpha\} \otimes a = \sup\{r \otimes a \mid r \in \mathbb{Q}, r \leq \alpha\}$.
- σ^+ is well-defined. In fact if $\alpha \otimes a = \beta \otimes b$, then $\sup\{r \otimes a \mid r \leq \alpha\} = \sup\{q \otimes b \mid q \leq \beta\}$. Now, for any fixed $x \in A$, $\sigma^Q(\cdot \otimes x) : t \mapsto ts(x)$ is a continuous map over $\mathbb{Q}$. Therefore $\sigma^+(\alpha \otimes a) = \sup\{\sigma^Q(r \otimes a) \mid r \leq \alpha\} = \{\sigma^Q(q \otimes b) \mid q \leq \beta\} = \sigma^+(\beta \otimes b)$. Hence σ^+ is well-defined.
- σ^+ is a state on $\mathcal{T}$. In fact $\sigma^+(1,1) = \sigma^Q(1,1) = 1$. Moreover if $(\alpha \otimes a) \odot (\beta \otimes b) = 0$, then from the continuity of $\odot$, one has $\sup\{(r \otimes a) \odot (q \otimes b) \mid r \leq \alpha, b \leq \beta\} = 0$. Therefore $\sigma^Q((r \otimes a) \oplus (q \otimes b)) = \sigma^Q(r \otimes a) \oplus \sigma^Q(q \otimes b)$. Thus $\sigma^+((\alpha \otimes a) \oplus (\beta \otimes b)) = \sup\{\sigma^Q((r \otimes a) \oplus (q \otimes b)) \mid r \leq \alpha, q \leq \beta\} = \sup\{\sigma^Q(r \otimes a) \oplus \sigma^Q(q \otimes b) \mid r \leq \alpha, q \leq \beta\} = \sigma^+(\alpha \otimes a) \oplus \sigma^+(\beta \otimes b)$ given the continuity of $\oplus$.

Therefore (1), (2) and (3) of Definition 3.1 hold. Moreover if $\alpha \in [0,1]$, then $\sigma^+(\alpha) = \alpha s(1) = \alpha$. Since for $x, y \in T$, $\sigma^+(x) \oplus \sigma^+(y) \in [0,1]$, also (4) is verified. Thus $(\mathcal{T}, \sigma^+)$ is a state MV-algebra. ■

6 The coherence problem

Given a finite probabilistic assessment $\Pr(E_1) = \alpha_1$, $\ldots$, $\Pr(E_n) = \alpha_n$, where $E_1, \ldots, E_n$ are fuzzy events,

identified with sentences of Łukasiewicz logic, and $\alpha_1, \ldots, \alpha_n$ are real numbers in $[0,1]$, we consider the following problem: *Is there a state s on the Lindenbaum sentence algebra $L_{\mathbf{L}}$ of Łukasiewicz logic such that for $i = 1, \ldots, n$, $s([E_i]) = \alpha_i$,* (where $[E_i]$ denotes the equivalence class of E_i modulo provable equivalence)*?* If such a state exist, then we say that the assessment is *coherent*. We want to find an equational characterization of the coherence problem when $\alpha_1, \ldots, \alpha_n$ are rational numbers. To this purpose, note that since Łukasiewicz logic is algebraizable in the sense of Blok and Pigozzi (cf [1]), the formulas $E_1, \ldots, E_n$ can be regarded as terms of MV-algebras as well. Now suppose that the α_i are rational numbers, say $\alpha_i = \frac{n_i}{m_i}$. Let $x_1, \ldots, x_n$ be fresh variables, and consider for $i = 1, \ldots, n$ the formulas the equations $\varepsilon_i : (m_i - 1)x_i = \neg x_i$. Moreover, for $i = 1, \ldots, n$, we denote by δ_i the equation $\sigma(E_i) = n_i x_i$. Then we can prove:

Theorem 6.1 *Let χ be the assessment consisting of the conditions $\Pr(E_i) = \frac{n_i}{m_i}$, $i = 1, \ldots, n$. The following are equivalent:*
(a) χ is coherent.
(b) The equations ε_i and δ_i, $i = 1, \ldots, n$, are satisfiable in some non-trivial state MV-algebra.

Proof. (a) $\Rightarrow$ (b). Given a state s on $L_{\mathbf{L}}$ we can construct the state MV-algebra $(\mathcal{T}, \sigma^+)$ with $\mathcal{T} = [0,1]_{MV} \otimes L_{\mathbf{L}}$ as in Theorem 5.2, thus getting a state MV-algebra whose operator σ^+ extends s. Let v be an evaluation on $(\mathcal{T}, \sigma^+)$ such that $v(x_i) = \frac{1}{m_i}$ (note that $\frac{1}{m_i} = \frac{1}{m_i} \otimes 1 \in \mathcal{T}$). Then equations ε_i are satisfied by v. Moreover, $\sigma^+(E_i) = \sigma^+(1 \otimes E_i) = 1s(E_i) = s(E_i) = \frac{n_i}{m_i} = n_i v(x_i)$, and also equations δ_i are satisfied by v.

(b) $\Rightarrow$ (a). Let $(\mathcal{M}, \sigma)$ be a state MV-algebra and v be an evaluation on $(\mathcal{M}, \sigma)$ satisfying equations ε_i and δ_i, $i = 1, \ldots, n$. Without loss of generality, we may assume that $\mathcal{M}$ is countably (or even finitely) generated, therefore there is an epimorphism h of $L_{\mathbf{L}}$ onto $\mathcal{M}$ such that $h([x]) = v([x])$ for every propositional variable x. Thus $(m_i - 1)h([x_i]) = \neg h([x_i])$, and $\sigma(h([E_i])) = n_i h([x_i])$. Let I be a maximal MV-ideal of $\sigma(M)$, and define for every $[E] \in L_{\mathbf{L}}$, $s^*([E]) = \sigma(h(E))/I$. Since quotients preserve identities, we have that $s^*([E_i]) = (n_i h([x_i]))/I$ and $(m_i - 1)(h([x_i])/I) = \neg(h([x_i])/I)$. Thus the canonical embedding of $\sigma(M)/I$ into $[0,1]_{MV}$ maps $h([x_i])/I$ into $\frac{1}{m_i}$ and $s^*([E_i])$ into $\frac{n_i}{m_i}$.

Finally, we verify that (modulo the canonical embedding of $\sigma(M)/I$ into $[0,1]_{MV}$), s^* is a state. That $s^*(1) = 1$ is clear. Moreover if $[E] \odot [F] = 0$ in $L_{\mathbf{L}}$, then $h([E]) \odot h([F]) = 0$ holds in A, therefore $s^*([E]) \odot s^*([F]) = (\sigma(h([E])) \odot \sigma(h([F])))/I = 0/I =$ 0, and

$$s^*([E] \oplus [F]) = (\sigma(h([E]) \oplus h([F])))/I =$$
$$(\sigma(h([E])))/I \oplus (\sigma(h([F])))/I =$$
$$s^*([E]) \oplus s^*([F]).$$

Thus s^* is a state on $L_{\mathbf{L}}$ satisfying the assessment χ, and the proof is finished. ∎

7 Conclusions and open problems

We introduced a unary operator σ on MV-algebras as to deal with the notion of *states* on MV-algebras and so defining the class of state MV-algebras. The main result of this paper says us that there is a canonical way of associating to each state MV-algebra $(\mathcal{M}, \sigma)$ a state on $\mathcal{M}$, and vice-versa, given a state s on an MV-algebra $\mathcal{M}$, one can define a state MV-algebra whose MV-reduct is obtained by tensor product of the standard MV-algebra $[0,1]_{MV}$ with $\mathcal{M}$.

As we remarked at the end of Section 4, the variety of state MV-algebras is not generated by its linearly-ordered members. For this reason, unlike the most common case, it is not possible to prove that the variety of state MV-algebras is generated by those algebras whose MV-reduct is the standard MV-algebra. In our opinion the most appropriate candidates to be the standard state MV-algebras are those algebras defined as in Section 5, and hence those algebras in the form $(\mathcal{T}, \sigma^+)$, $\mathcal{T}$ being the tensor product $[0,1]_{MV} \otimes \mathcal{M}$ and $\mathcal{M}$ being any (not necessarily linearly ordered) MV-algebra. Unfortunately we was not able, since now, to prove that an equation $\varphi \approx 0$, written in the language of state MV-algebras, holds in every state MV-algebra iff $\varphi \approx 0$ holds in some $(\mathcal{T}, \sigma^+)$. Therefore it can be reasonably considered as an open problem.

Another open problem is that one of studying the computational complexity for state MV-algebras, and applying it to the problem of deciding the coherence of rational assessments by the equational characterization we proved in Section 6.

Acknowledgement

References

[1] Blok W., Pigozzi D., *Algebraizable Logics.* Mem. Amer. Math. Soc, **396** (77). Amer. Math Soc. Providence, 1989.

[2] Chang C. C., *Algebraic analysis of many-vaued logics.* Trans. Amer. Math. Soc., **88**, 467–490, (1958).

[3] Hájek P., *Metamathematics of Fuzzy Logic.* Kluwer Academy Press, 1998.

[4] MUNDICI D., *Averaging the truth-value in Łukasiewicz logic*. Studia Logica, **55**(1): 113–127, 1995.

[5] MUNDICI D., *Tensor Products and the Loomis-Sikorski Theorem for MV-Algebras*. Advances in Applied Mathematics, **22**, 227–248, 1999.

Free MV-algebras

Antonio Di Nola
Dipartimento di Matematica e Informatica
University of Salerno
e-mail: adinola@unisa.it

Revaz Grigolia
Institute of Computer Sciences
Tbilisi State University
e-mail: grigolia@yahoo.com

Abstract

We describe n-generated free MV-algebras as MV-algebras having the lattice reduct which is a direct limit in the category of distributive lattices.

Keywords: MV-algebras, Lattices.

1 Preliminaries

We recall that an algebra $A = (A; 0, 1, \oplus, \cdot, \sim)$, [1], is said to be an MV -algebra iff it satisfies the following equations:

1. $(x \oplus y) \oplus z = x \oplus (y \oplus z)$;
2. $x \oplus y = y \oplus x$;
3. $x \oplus 0 = x$;
4. $x \oplus 1 = 1$;
5. $\sim 0 = 1$;
6. $\sim 1 = 0$;
7. $x \cdot y = \sim (\sim x \oplus \sim y)$;
8. $\sim (\sim x \oplus y) \oplus y = \sim (\sim y \oplus x) \oplus x$.

Henceforth we shall write ab for $a \cdot b$ and a^n for $\underbrace{a \cdot \cdots \cdot a}_{n \text{ times}}$, for given $a, b \in A$. Every MV-algebra has an underlying ordered structure defined by

$$x \leq y \text{ iff } \sim x \oplus y = 1.$$

$(A; \leq, 0, 1)$ is a bounded distributive lattice. Moreover, the following property holds in any MV-algebra:

$$xy \leq x \wedge y \leq x \vee y \leq x \oplus y.$$

The unit interval of real numbers $[0, 1]$ endowed with the following operations: $x \oplus y = \min(1, x + y), x \cdot y = \max(0, x + y - 1), \sim x = 1 - x$, becomes an MV-algebra. Let Q denote the set of rational numbers, for $(0 \neq) m \in \omega$ we set $S_m = (S_m; \oplus, \cdot, \sim, 0, 1)$, where

$$S_m = \left\{0, \frac{1}{m}, \ldots, \frac{m-1}{m}, 1\right\}.$$

Recall that an algebra $A \in \mathbf{K}$ is said to be a *free algebra* in a variety $\mathbf{K}$, if there exists a set $A_0 \subset A$ such that A_0 generates A and every mapping f from A_0 to any algebra $B \in \mathbf{K}$ is extended to a homomorphism h from A to B. In this case A_0 is said to be *the set of free generators* of A. If the set of free generators is finite then A is said to be a *finitely generated free algebra* .

2 Representation of n-generated MV-algebras

Let us denote by DL the category, and the variety as well, of distributive lattices, named by DL-algebras in the sequel, and lattice homomorphisms.

Notice that any MV-algebra is bounded distributive lattice and the lattice operations $\vee$ and $\wedge$ are definable by means of MV-algebra operations $\oplus$, $\cdot$ and $\sim$. So, any MV-algebra is a DL-algebra with respect to the operations $\vee, \wedge$.

Let $(A, \oplus, \cdot, \sim, 1, 0)$ be an MV-algebra. Define a functor $\mathcal{D}$ transforming MV-algebra $(A, \oplus, \cdot, \sim, 1, 0)$ into DL-algebra in the following way:

$$\mathcal{D}((A, \oplus, \cdot, \sim, 1, 0)) = (A, \vee, \wedge).$$

Then we have functor from the category of MV-algebras MV to the category DL.

Let $F_{V_n}(m)$ be the m-generated free MV-algebra in the variety $V_n = \mathcal{V}(\{S_1, ..., S_n\})$. We can consider any MV-algebra as (unbounded) distributive lattice (with lattice operations $\vee$ and $\wedge$). It is clear that $F_{V_n}(m)$ is embedded into $F_{V_{n+1}}(m)$ as MV-algebra and, consequently, as a lattice.

Let $g_1^{(n)}, ..., g_m^{(n)} \in F_{V_n}(m)$ be free generators of $F_{V_n}(m)$.

Let us define the injective mapping

$$\varepsilon_{n(n+1)} : \mathcal{D}(F_{V_n}(m)) \to \mathcal{D}(F_{V_{n+1}}(m))$$

for any $n \in \omega/\{0\}$ in the following way:

$$g_1^{(n)} \mapsto g_1^{(n+1)}, ..., g_m^{(n)} \mapsto g_m^{(n+1)} \text{ for any } n \in \omega/\{0\}.$$

It is clear that in the case when $n = 1$, $g_1^{(1)}, ..., g_m^{(1)}$ generate $F_{V_1}(m)$ $(= D_1^{(m)} = \mathcal{D}(F_{V_1}(m)))$ using only operations $\vee, \wedge$ and $\sim$. Then $\varepsilon_{12}(g_i^{(1)}) = g_i^{(2)}$, for $i = 1, ..., m$, and generate the algebra $D_2^{(m)}$ inside $F_{V_2}(m)$ by means of $g_1^{(2)}, ..., g_m^{(2)}$ using only operations $\vee, \wedge$ and $\sim$. It easy to check that $D_2^{(m)}$ is isomorphic to $D_1^{(m)}(= \mathcal{D}(F_{V_1}(m)))$. $\varepsilon_{12}(x \vee y) = \varepsilon_{12}(x) \vee \varepsilon_{12}(y)$, $\varepsilon_{12}(x \wedge y) = \varepsilon_{12}(x) \wedge \varepsilon_{12}(y)$.

In particular, if $m = 1$, then $\varepsilon_{12}(1) = g_1^{(2)} \vee \sim g_1^{(2)}$ $(\neq 1)$ and $\varepsilon_{12}(0) = g_1^{(2)} \wedge \sim g_1^{(2)}$ $(\neq 0)$. Observe that $\varepsilon_{12}(\mathcal{D}(F_{V_1}(1)))$ generates $F_{V_1}(2)$ using MV-algebra polynomials containing operations $\vee, \wedge, \sim$ and MV-algebra operation $\oplus$ (or $\cdot$) only one time. It is enough, since in V_2 holds the axiom $2x = 3x$ (and $2x$ is a Boolean element). Indeed, by means of the generator $g_1^{(2)}$ using $\vee, \wedge, \sim$ and polynomials containing operation $\oplus$ (or $\cdot$) only one time, we obtain elements $(0,1,1), (1,0,1), (1,1,0), (1,0,0), (0,1,0), (0,0,1)$ and, hence, any element of $F_{V_1}(2)$ by means of $\vee, \wedge$ and $g_1^{(2)}, \sim g_1^{(2)}$. Notice, that the embedding ε_{12} is not MV-algebra embedding, it is DL-algebra embedding. So, we define a functor $\mathcal{L}_2$ transforming DL-algebra $\mathcal{D}(F_{V_2}(1))$ into MV-algebra $F_{V_2}(1)$. Further, $\varepsilon_{23}(g_1^{(2)}) = g_1^{(3)}$ and close $\{g_1^{(3)}\}$ by means of operations $\vee, \wedge, \sim$. Then we obtain DL-algebra $D_3^{(1)}$ which is isomorphic to $\mathcal{D}(F_{V_1}(1))$. Analogically, as in the previous case, close $D_3^{(1)}$ using $\vee, \wedge, \sim$ and polynomials containing operation $\oplus$ (or $\cdot$) only one time. Then we obtain DL-algebra $\varepsilon_{23}(\mathcal{D}(F_{V_1}(2)))$ which is isomorphic to $\mathcal{D}(F_{V_1}(2))$. Further, close $\varepsilon_{12}(\mathcal{D}(F_{V_1}(2)))$ using $\vee, \wedge, \sim$ and polynomials containing operation $\oplus$ (or $\cdot$) only not more than on two places. Then we obtain the algebra $\mathcal{D}(F_{V_1}(3))$. So, we can define a functor $\mathcal{L}_3$ transforming DL-algebra $\mathcal{D}(F_{V_3}(1))$ into MV-algebra $F_{V_3}(1)$. And so on.

Analogically it is shown in general case when $m > 1$.

Further, $\varepsilon_{12}(g_i^{(2)}) = g_i^{(3)}$, for $i = 1, ..., m$, and generate the algebra $D_3^{(m)}$ inside $F_{V_3}(m)$ by means of $g_1^{(3)}, ..., g_m^{(3)}$ using only operations $\vee, \wedge$ and $\sim$. It easy to check that $D_3^{(m)}$ is isomorphic to $F_{V_1}(m)$ (as a lattice with involution). Moreover, $D_3^{(m)}$ generates $F_{V_3}(m)$, if we use the MV-algebra operations $\sim$ and $\oplus$ only two times. In other words, for generating $F_{V_3}(m)$ we use polynomials containing the operation $\sim$ and $\oplus$ only in not more than two places. It is enough, since in V_3 holds the axiom $3x = 4x$.

Let us consider the direct system $(\mathcal{D}(F_{V_n}(m)), \varepsilon_{n(n+1)})_{n \in \omega/\{0\}}$. Let $\mathcal{D}(F(m))$ be the direct limit of the direct system $(\mathcal{D}(F_{V_n}(m)), \varepsilon_{n(n+1)})_{n \in \omega/\{0\}}$. The elements of $\mathcal{D}(F(m))$ are infinite sequences $(a_i, a_{i+1}, a_{i+2}, ...)$ for $i \in \omega$ such that $a_i \in \mathcal{D}(F_{V_i}(m))$ and a_i has no inverse image with respect to $\varepsilon_{(i-1)i}$, and, moreover, $\varepsilon_{i_1(i_1+1)}(a_{i_1}) = a_{i_1+1}$ for any $i_1 \geq i$. Now we define the operation $\oplus$ on $\mathcal{D}(F(m))$ in the following way.

Let $a = (a_i, a_{i+1}, a_{i+2}, ...), b = (b_j, b_{j+1}, b_{j+2}, ...)$ and $a, b \in F(m)$, and $i \leq j$. It is clear that a_i (respectively, b_j) represented by a valuation of an MV-algebra polynomial containing operation $\oplus$ only on $i-1$ (respectively, $j-1$) places. Then we consider elements $a_{max(i,j)}$ and $b_{max(i,j)}$. It is clear that they belong to $\mathcal{D}(F_{V_{max(i,j)}}(m))$. Choose the least number $k \geq max(i,j)$ such that $a_{max(i,j)} + b_{max(i,j)} \in \mathcal{D}(F_{V_k}(m))$. Then $a \oplus b = (a_k + b_k, a_{k+1} + b_{k+1}, a_{k+2} + b_{k+2}, ...)$. Then we convert $\mathcal{D}(F(m))$ into MV-algebra which we denote by $F_{MV}(m)$.

Theorem 2.1. *The MV-algebra $F_{MV}(m)$ is m-generated free MV-algebra.*

References

[1] C. C. Chang, Algebraic Analysis of Many-Valued Logics, Trans. Amer. Math. Soc. 88 (1958) 467-490.

Concept lattices under non-commutative conjunctors are generalized concept lattices

J. Medina*
Dept. Matemática Aplicada
Universidad de Málaga. Spain
jmedina@ctima.uma.es

M. Ojeda-Aciego†
Dept. Matemática Aplicada
Universidad de Málaga. Spain
aciego@ctima.uma.es

J. Ruiz Calviño‡
Dept. Matemática Aplicada
Universidad de Málaga. Spain
jorgerucal@ctima.uma.es

Abstract

Generalized concept lattices have been recently proposed to deal with uncertainty or incomplete information as a non-symmetric generalization of the theory of fuzzy formal concept analysis. On the other hand, concept lattices have been defined as well in the framework of fuzzy logics with non-commutative conjunctors.

The contribution of this paper is to prove that any concept lattice for non-commutative fuzzy logic can be interpreted inside the framewok of generalized concept lattices, specifically, it is isomorphic to a sublattice of the cartesian product of two generalized concepts lattices.

Keywords : formal concept analysis, concept lattices, Galois connections.

1 Introduction

The theory of Formal Concept Analysis has its beginnings in the works of Ganter and Wille [5] where an object-attribute view of data is developed. Specifically, a concept is defined as a pair of subsets which, respectively mean the extension (the subset of objects related to the concept) and the intension (the set of attributes which define the concept).

Ganter-Wille's approach was based on a classical setting, in that objects (resp. attributes) either crisply belong or not to the extension (resp. intension) of a concept. Since then, there have been several approaches aiming at introducing some kind of fuzziness, vagueness or uncertainty in the data. Fuzzy concept lattices were firstly introduced by Burusco and Fuentes-González in [4], later independently developed by Pollandt in [11] and Bělohlávek in [2] (which also considered fuzzy orderings). Later, Georgescu and Popescu [6] defined the notion of fuzzy concept lattice associated to fuzzy logic with a non-commutative conjunction. More recently, Krajči considered the so-called generalized concept lattices, which use different sets of truth-values to refer to the subset of objects, to the subset of attributes of a concept, as well as to the degree to which an object has an attribute.

The framework of generalized concept lattices is actually a generalization of Pollandt's concept lattice and Bělohlávek's concept lattice (when considering a classical ordering) and, furthermore, it has been shown to embed some other approaches, like the concept lattices with hedges [7].

In this paper, the frameworks of Krajči's generalized concept lattice and Georgescu-Popescu's construction for non-commutative conjunctions are related; specifically, we show the relationship between the residuated operators introduced by the latter with the notion of left-continuity introduced by the former. This relationship allows to interpret Georgescu-Popescu concept lattices as a particular case of the generalized concept lattices.

2 Generalized Concept Lattices

As stated in the introduction, generalized concept lattices are based on two complete lattices $(L_1, \preceq_1)$ $(L_2, \preceq_2)$, and a poset $(P, \leq)$, which are the different sets of truth-values to refer to the objects, to the attributes of a concept, as well as to the degree to which an object has an attribute.

In addition, a conjunction operator $\otimes\colon L_1 \times L_2 \to P$ is considered, which is assumed to be increasing and left-continuous in both arguments. The notion of left-continuity, see [8], is given below:

Definition 2.1 *Given $(L, \preceq)$ a complete lattice and*

*Partially supported by TIC2003-09001-C02-01
†Partially supported by TIN2006-15455-C03-01
‡Partially supported by P06-FQM-02049

$(P, \leq)$ a poset, an application $T\colon L \to P$ is left-continuous *when, given $p \in P$ and a non-empty subset $X \subseteq L$, the following condition holds:*

> *if $T(x) \leq p$ for every $x \in X$, then $T(\sup X) \leq p$*

The *context* under which concepts are defined is a tuple $(A, B, R, \otimes)$, where the sets A and B represent the attributes and objects, and $R\colon A \times B \to P$ is a P-fuzzy relation.

Given a context $(A, B, R, \otimes)$, two maps $^{\uparrow}\colon {L_2}^B \to {L_1}^A$ and $^{\downarrow}\colon {L_1}^A \to {L_2}^B$ can be defined as follows:

$$\begin{aligned} g^{\uparrow}(a) &= \sup\{x \in L_1 \mid (\forall b \in B)\ x \otimes g(b) \leq R(a,b)\} \\ f^{\downarrow}(b) &= \sup\{y \in L_2 \mid (\forall a \in A) f(a) \otimes y \leq R(a,b)\} \end{aligned}$$

Now, the concepts given by the context are defined as the elements (g, f) of ${L_2}^B \times {L_1}^A$ such that $g^{\uparrow} = f$ and $f^{\downarrow} = g$.

By relying on the fact that the pair $(^{\downarrow}, ^{\uparrow})$ is a Galois connection, Krajči proved that the set of concepts

$$\mathcal{G} = \{(g, f) \mid g^{\uparrow} = f \text{ and } f^{\downarrow} = g\}$$

is actually a complete lattice with the ordering $\preceq_{\mathcal{G}}$ defined as

$(g_1, f_1) \preceq_{\mathcal{G}} (g_2, f_2)$ if and only if $g_1 \preceq_2 g_2$

The lattice $\mathcal{G}$ is known as the *generalized concept lattice* associated to $(A, B, R, \otimes)$.

3 Concept lattice for non-commutative conjunctors

Georgescu and Popescu introduced in [6] a concept lattice which is based on the structure of complete biresiduated lattice[1] as underlying set for the truth-values of both the objects and attributes. The formal definition of this structure is given below:

Definition 3.1 *A* complete biresiduated lattice *is a tuple $(L, \preceq, \&, \swarrow, \nwarrow)$ satisfying the following conditions:*

1. *$(L, \preceq)$ is a complete lattice.*
2. *$(L, \&, \top)$ is a monoid.*
3. *The adjoint properties:*
   (a) $x \preceq z \swarrow y$ *if and only if* $x \,\&\, y \preceq z$
   (b) $y \preceq z \nwarrow x$ *if and only if* $x \,\&\, y \preceq z$

[1] The term used by Georgescu-Popescu is complete generalized residuated lattice, which we do not use here in order to avoid misunderstandings with Krajči's.

Note that the adjoint conditions are different precisely when the conjunctor & is non-commutative.

The study of implications and conjunctions related by adjointness has recently been the subject of extensive research, becoming an important branch of the theory of multiple-valued and fuzzy logic. Note that this structure was introduced in the framework of fuzzy logic programming in [10] and, simultaneously, under the name of implication triple, in [1].

In order to define the concept lattice, we have to introduce the notion of context. Given a complete biresiduated lattice $(L, \preceq, \&, \swarrow, \nwarrow)$, a (biresiduated) context is a tuple (A, B, R) where A, B are sets representing the attributes and the objects, respectively, and $R\colon A \times B \to L$ is a L-fuzzy relation.

Now, given a context (A, B, R) the mappings $^{\Uparrow}$, $^{\Uparrow}\colon L^B \to L^A$ and $^{\Downarrow}$, $^{\Downarrow}\colon L^A \to L^B$ are defined as follows:

$$\begin{aligned} g^{\Uparrow}(a) &= \inf\{R(a,b) \swarrow g(b) \mid b \in B\} \\ g^{\Uparrow}(a) &= \inf\{R(a,b) \nwarrow g(b) \mid b \in B\} \\ f^{\Downarrow}(b) &= \inf\{R(a,b) \nwarrow f(a) \mid a \in A\} \\ f^{\Downarrow}(b) &= \inf\{R(a,b) \swarrow f(a) \mid a \in A\} \end{aligned}$$

The concepts in this framework are triples $(g, f, f') \in L^{A \times B \times B}$ such that $g^{\Uparrow} = f$; $g^{\Uparrow} = f'$; $f^{\Downarrow} = g$; $f'^{\Downarrow} = g$; this is why we will call them *t-concepts*.

The fact that the pairs $(^{\Uparrow}, ^{\Downarrow})$ and $(^{\Uparrow}, ^{\Downarrow})$ form Galois connections is used in [6] to prove that the set of t-concepts, denoted as $\mathcal{L}$, is a complete lattice under the ordering $(g_1, f_1, f'_1) \preceq_{\mathcal{L}} (g_2, f_2, f'_2)$ if and only if $g_1 \preceq g_2$ (equivalently $f_2 \preceq f_1$ or $f'_2 \preceq f'_1$).

4 Relating both frameworks

In order to embed the concept lattice $\mathcal{L}$ into Krajči's framework, firstly we have to know when the operator &, defined in Section 3, is left continuous in both arguments.

By definition, we have that & has associated two "residuated" applications $\swarrow$ and $\nwarrow$ satisfying the adjoint properties. As a result we obtain that & is sup-preserving in both arguments, i. e., for all $x, y \in L$ and $X, Y \subseteq L$ we have that $\sup(X) \,\&\, y = \sup\{x' \,\&\, y \mid x' \in X\}$, and $x \,\&\, \sup(Y) = \sup\{x \,\&\, y' \mid y' \in Y\}$, see [1].

Once we know that & is sup-preserving in both arguments, the following step is to obtain left-continuity. This can be achieved as an application of the following result which characterises when an operator is sup-preserving in terms of the left-continuity.

Lemma 4.1 *Let $(L, \preceq)$ be a complete lattice and*

$\wedge\colon L \times L \to L$ an increasing operator then the following conditions are equivalent:

1. *$\wedge$ is sup-preserving in the first argument.*
2. *$\wedge$ is left-continuous in the first argument and $\bot \wedge y = \bot$ for every $y \in L$.*

Proof: (1 implies 2)

The proof of the boundary condition is trivial considering $X = \varnothing$ since $\bot \wedge y = \sup(X) \wedge y = \sup\{x \wedge y \mid x \in X\} = \bot$. Now, given y, $z \in L$ and a non-empty subset $X \subseteq L$, if $x \wedge y \preceq z$ for every $x \in X$ then $\sup\{x \wedge y \mid x \in X\} \preceq z$, so, by hypothesis:

$$\sup(X) \wedge y = \sup\{x \wedge y \mid x \in X\} \preceq z$$

therefore $\wedge$ is left-continuous in the first argument.

(2 implies 1)

Let $\varnothing \neq X \subseteq L$ and $y \in L$, the inequality $\sup\{x \wedge y \mid x \in X\} \preceq \sup(X) \wedge y$ follows directly from the increasing character of $\wedge$ and definition of supremum.

For the other inequality, since $x \wedge y \preceq \sup\{x \wedge y \mid x \in X\}$ for every $x \in X$, we can use the left-continuity in the first argument of $\wedge$, and obtain $\sup(X) \wedge y \preceq \sup\{x \wedge y \mid x \in X\}$.

If $X = \varnothing$, the equality is straightforward because of the boundary condition and $\sup(X) = \bot$. ∎

A similar lemma can be proved for the second argument, but in this case the boundary condition has to be modified as $x \wedge \bot = \bot$. As a consequence of Lemma 4.1, we get that $\&$ is left continuous in both arguments.

Remark 4.1 *Note that only the first implication is needed for our purposes. However, we have stated and proved the full equivalence in order to point out a gap in [3], in which there is no explicit mention to the need of the boundary conditions, which turns out to be essential.*

Finally, an alternative definition of the Galois connections of Section 3 can be given in terms of suprema, hence obtaining a definition more similar to that of Krajči's:

Lemma 4.2 *Given a complete biresiduated lattice $(L, \preceq, \&, \swarrow, \nwarrow)$ and a biresiduated context (A, B, R) we have that:*

$$\begin{aligned} g^{\uparrow\uparrow}(a) &= \sup\{x \in L_1 \mid (\forall b \in B) x \,\&\, g(b) \preceq R(a,b)\} \\ f^{\downarrow\downarrow}(b) &= \sup\{y \in L_2 \mid (\forall a \in A) f(a) \,\&\, y \preceq R(a,b)\} \end{aligned}$$

Proof: For the first equality we need to prove that

$$\sup\{x \in L_1 \mid (\forall b \in B) x \,\&\, g(b) \preceq R(a,b)\} = \inf\{R(a,b) \swarrow g(b) \mid b \in B\}$$

By the adjoint property of $\&$ with respect to $\swarrow$, and the characterisation of the infimum as the supremum of the lower bounds, we obtain

$$\begin{aligned} &\sup\{x \in L \mid (\forall b \in B) x \,\&\, g(b) \preceq R(a,b)\} = \\ &\quad = \sup\{x \in L \mid (\forall b \in B) x \preceq R(a,b) \swarrow g(b)\} \\ &\quad = \inf\{R(a,b) \swarrow g(b) \mid b \in B\} \end{aligned}$$

The other equality is proved in a similar way. ∎

Theorem 4.1 *Given a complete biresiduated lattice $(L, \preceq, \&, \swarrow, \nwarrow)$ and a biresiduated context (A, B, R), then there exist two generalized concept lattices, $\mathcal{G}_1$ and $\mathcal{G}_2$ such that the sublattice of $\mathcal{G}_1 \times \mathcal{G}_2$ defined by*

$$\mathcal{G}_{12} = \{((g_1, f_1), (g_2, f_2)) \in \mathcal{G}_1 \times \mathcal{G}_2 \mid g_1 = g_2\}$$

is isomorphic to the lattice of t-concepts $\mathcal{L}$.

Proof: By Lemma 4.2, we have that the definitions of the applications $(^{\uparrow\uparrow}, ^{\downarrow\downarrow})$ given in Section 3 coincide with the definitions of $(^{\uparrow}, ^{\downarrow})$ given in Section 2 considering the context $(A, B, R, \&)$.

Now, if we consider the operator $\&^{op}\colon L \times L \to L$, where $x \,\&^{op}\, y = y \,\&\, x$, we obtain similarly that the pair $(^{\Uparrow}, ^{\Downarrow})$ is equal to $(^{\uparrow^{op}}, ^{\downarrow_{op}})$ defined for the context $(A, B, R, \&^{op})$.

Finally, we simply have to take $\mathcal{G}_1$ and $\mathcal{G}_2$ as the generalized concept lattices associated to the contexts $(A, B, R, \&)$ and $(A, B, R, \&^{op})$ ∎

5 Conclusions

We have shown how the framework of concept lattices for non-commutative fuzzy logics introduced by Georgescu and Popescu naturally embeds in the framework of generalized concept lattices.

As future work we are planning to study the relationship between the generalized concept lattice and the recently introduced multi-adjoint concept lattice [9].

References

[1] A.A. Abdel-Hamid and N.N. Morsi. Associatively tied implicacions. *Fuzzy Sets and Systems*, 136(3):291–311, 2003.

[2] R. Bělohlávek. Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic*, 128:277–298, 2004.

[3] R. Bělohlávek and V. Vychodil. What is a fuzzy concept lattice? In *Intl Workshop on Concept Lattices and their Applications*, pages 34–45, 2005.

[4] A. Burusco and R. Fuentes-González. The study of L-fuzzy concept lattice. *Mathware & Soft Computing*, 3:209–218, 1994.

[5] B. Ganter and R. Wille. *Formal Concept Analysis: Mathematical Foundation*. Springer Verlag, 1999.

[6] G. Georgescu and A. Popescu. Concept lattices and similarity in non-commutative fuzzy logic. *Fundamenta Informaticae*, 55(1):23–54, 2002.

[7] S. Krajči. Every concept lattice with hedges is isomorphic to some generalized concept lattice. In *Intl Workshop on Concept Lattices and their Applications*, pages 1–9, 2005.

[8] S. Krajči. A generalized concept lattice. *Logic Journal of IGPL*, 13(5):543–550, 2005.

[9] J. Medina, M. Ojeda-Aciego, and J. Ruiz-Calviño. On multi-adjoint concept lattices: definition and representation theorem. In *Proc. Intl. Conf. on Formal Concept Analysis*, Lect. Notes in Artificial Intelligence 4390:197–209, 2007.

[10] J. Medina, M. Ojeda-Aciego, and P. Vojtás. Towards biresiduated multi-adjoint logic programming. In *Proc. CAEPIA'03*, Lect. Notes in Artificial Intelligence 3040:608–617, 2003.

[11] S. Pollandt. *Fuzzy Begriffe*. Springer, Berlin, 1997.

Mixed Integer Programming, General Concept Inclusions and Fuzzy Description Logics

Umberto Straccia
ISTI - CNR, Pisa, Italy
straccia@isti.cnr.it

Fernando Bobillo
Department of Computer Science and AI
University of Granada, Spain
fbobillo@decsai.ugr.es

Abstract

Fuzzy Description Logics (fuzzy DLs) have been proposed as a language to describe structured knowledge with vague concepts. In [23], a solution based on Mixed Integer Linear Programming has been proposed to deal with fuzzy DLs under Łukasiewicz semantics in which typical membership functions, such as triangular and trapezoidal functions, can be explicitly represented in the language.

A major theoretical and computational limitation so far is the inability to deal with *General Concept Inclusions* (GCIs), which is an important feature of classical DLs. In this paper, we address this issue and develop a calculus for fuzzy DLs with GCIs under various semantics: classical logic, "Zadeh semantics", and Łukasiewicz logic.

Keywords: Fuzzy Description Logics

1 Introduction

Description Logics (DLs) [1] are a logical reconstruction of the so-called frame-based knowledge representation languages, with the aim of providing a simple well established Tarski-style declarative semantics to capture the meaning of the most popular features of structured representation of knowledge. Nowadays, DLs have gained even more popularity due to their application in the context of the *Semantic Web*, as the theoretical counterpart of OWL DL (the W3C standard for specifying ontologies, see [12] for details).

Fuzzy DLs [18, 22, 28, 30] extend classical DLs by allowing to deal with *fuzzy/vague/imprecise concepts* for which a clear and precise definition is not possible. The problem to deal with imprecise concepts has been addressed several decades ago by Zadeh [31], which gave birth in the meanwhile to the so-called *fuzzy set and fuzzy logic theory* and a huge number of real life applications exists. Despite the popularity of fuzzy set theory, relative little work has been carried out in extending DLs towards the representation of imprecise concepts, notwithstanding DLs can be considered as a quite natural candidate for such an extension [2, 5, 7, 8, 13, 16, 19, 18, 20, 21, 22, 23, 25, 27, 28, 30, 32].

From a semantics point of view, most works rely on the Gödel conjunction and disjunction ($x \wedge y = \min(x, y)$, $x \vee y = \max(x, y)$), but with Łukasiewicz negation ($\neg x = 1 - x$), essentially the semantics of fuzzy sets operators proposed by Zadeh [31] (which we call here "Zadeh semantics"), while very few work address the case of a fully Łukasiewcz compliant semantics. Indeed, Hajek [6, 5] proposes a reasoning solution based on a reduction to Łukasiewicz propositional logic Ł, while Straccia [23] proposes a reasoning solution, which is based on a mixture of tableau rules and Mixed Integer Linear Programming (MILP). The latter solution has been proposed to deal with fuzzy DLs with so-called fuzzy concrete domains, i.e. the possibility to represent in fuzzy DLs concepts with explicit membership functions such as triangular, trapezoidal, left-shoulder and right-shoulder functions.

A major theoretical, computational and practical limitation so far of fuzzy DLs is the inability to deal with *General Concept Inclusions* (GCIs), which is an important feature of classical DLs; e.g., GCIs are necessary to represent domain and range constraints. We note that [29, 20] present a solution for fuzzy DLs under "Zadeh semantics", but which is hardly applicable in real world scenarios. The major problem relies on the fact that, informally, the algorithm generates a disjunction for each rational occurring in the knowledge base. This is clearly not feasible in practice.

In this paper, we address this issue and develop a calculus for fuzzy $\mathcal{ALCF}(\mathbf{D})$ with GCIs under various semantics, which have been considered in the running system *fuzzyDL*, available from Straccia's home page.

In the remainder, we proceed as follows. We next describe the basics of fuzzy DLs, then address the inference algorithm and finally we conclude.

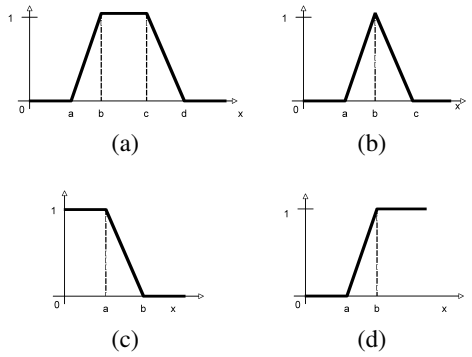


Figure 1: (a) Trapezoidal function; (b) Triangular function; (c) L-function; (d) R-function

2 Fuzzy DLs basics

In this section, we define fuzzy $\mathcal{ALCF}(D)$. We recall here the semantics given in [6, 5, 25, 26].

Syntax. In fuzzy $\mathcal{ALCF}(D)$, we allow to reason with concrete fuzzy data types, using so-called concrete domains. We recall that $\mathcal{ALCF}(\mathbf{D})$ is the basic DL $\mathcal{ALC}$ [17] extended with functional roles (also called attributes or features) and concrete domains [14] allowing to deal with data types such as strings and integers. In fuzzy $\mathcal{ALCF}(\mathbf{D})$, however, concrete domains are fuzzy sets.

A *fuzzy data type theory* $\mathbf{D} = \langle \Delta_{\mathbf{D}}, \cdot_{\mathbf{D}} \rangle$ is such that $\cdot_{\mathbf{D}}$ assigns to every n-ary data type predicate an n-ary fuzzy relation over $\Delta_{\mathbf{D}}$. For instance, as for $\mathcal{ALCF}(\mathbf{D})$, the predicate $\leqslant_{18}$ may be a unary crisp predicate over the natural numbers denoting the set of integers smaller or equal to 18. On the other hand, concerning non-crisp fuzzy domain predicates, we recall that in fuzzy set theory and practice, there are many functions for specifying fuzzy set membership degrees. However, the triangular, the trapezoidal, the L-function (left-shoulder function), and the R-function (right-shoulder function) are simple, but most frequently used to specify membership degrees. The functions are defined over the set of non-negative rationals $\mathbb{Q}^+ \cup \{0\}$ (see Fig. 1). Using these functions, we may then define, for instance, $Young\colon Natural \to [0,1]$ to be a fuzzy concrete predicate over the natural numbers denoting the degree of youngness of a person's age. The concrete fuzzy predicate $Young$ may be defined as $Young(x) = L(x; 10, 30)$.

We allow modifiers in fuzzy $\mathcal{ALCF}(\mathbf{D})$. Fuzzy modifiers, like $very$, $more_or_less$ and $slightly$, apply to fuzzy sets to change their membership function. Formally, a *modifier* is a function $f_m\colon [0,1] \to [0,1]$. For instance, we may define $very(x) = x^2$ and $slightly(x) = \sqrt{x}$. Modifiers have been considered, for instance, in [9, 28].

Now, let $\mathbf{A}$, $\mathbf{R}_A$, $\mathbf{R}_C$, $\mathbf{I}$, $\mathbf{I}_c$ and M be non-empty finite and pair-wise disjoint sets of *concepts names* (denoted A), *abstract roles names* (denoted R), i.e. binary predicates *concrete roles names* (denoted T), *abstract individual names* (denoted a), *concrete individual names* (denoted c) and *modifiers* (denoted m). Concepts may be seen as unary predicates, while roles may be seen as binary predicates. $\mathbf{R}_A$ also contains a non-empty subset $\mathbf{F}_a$ of *abstract feature names* (denoted r), while $\mathbf{R}_C$ contains a non-empty subset F_c of *concrete feature names* (denoted t). Features are functional roles.

The syntax of fuzzy $\mathcal{ALCF}(\mathbf{D})$ concepts is as follows:

$$\begin{aligned} C \quad := \quad & \top \mid \bot \mid A \mid C_1 \sqcap C_2 \mid C_1 \sqcup C_2 \mid \neg C \mid m(C) \mid \\ & \forall R.C \mid \exists R.C \mid \forall T_1, \ldots, T_n.D \mid \exists T_1, \ldots, T_n.D\,. \end{aligned}$$

Concerning axioms and assertions, similarly to [26], we define fuzzy axioms as follows.

Let $n \in (0,1]$. A *fuzzy TBox* $\mathcal{T}$ is a finite set of *fuzzy concept inclusion axioms* $\langle C \sqsubseteq D, n\rangle$, where C, D are concepts. Informally, $\langle C \sqsubseteq D, n\rangle$ states that all instances of concept C are instances of concept D to degree n. We write $C = D$ as a shorthand of the two axioms $\langle C \sqsubseteq D, 1\rangle$ and $\langle D \sqsubseteq C, 1\rangle$. For instance,

$$Minor = Person \sqcap \exists age.\leqslant_{18} \tag{1}$$

defines a person, whose age is less or equal to 18 (age is a concrete feature), i.e., it defines a minor,

$$YoungPerson = Person \sqcap \exists age.Young \tag{2}$$

defines a young person, while

$$SportsCar = Car \sqcap \exists speed.very(High)\,, \tag{3}$$

we may define the concept of sports car, where $very$ is a concept modifier, e.g. a linear hedge, and $High$ is a fuzzy concrete predicate over the domain of speed expressed in kilometres per hour and may be defined as $High(x) = R(x; 80, 250)$ ($speed$ is a concrete feature).

A *fuzzy ABox* $\mathcal{A}$ consists of a finite set of *fuzzy concept* and *fuzzy role assertion axioms* of the form $\langle a\colon C, n\rangle$, $\langle (a,b)\colon R, n\rangle$ and $\langle (a,c)\colon T, n\rangle$, where a, b are abstract individual constants, c is a concrete individual, and C, R and T are a concept, an abstract role and a concrete role, respectively.

Informally, from a semantical point of view, a fuzzy axiom $\langle \alpha, n\rangle$ constrains the membership degree of α to be at least n. Hence, $\langle jim\colon YoungPerson, 0.2\rangle$ says that jim is a $YoungPerson$ with degree at least 0.2, while $\langle (jim, 180)\colon hasHeight, 1\rangle$, where $hasHeight$ is a concrete feature, says that jim's height is 180. On the other hand, a fuzzy concept inclusion axiom of the form $\langle C \sqsubseteq D, n\rangle$ says that the subsumption degree between C and D is at least n.

An $\mathcal{ALCF}(\mathbf{D})$ *fuzzy knowledge base* $\mathcal{K} = \langle \mathcal{T}, \mathcal{A}\rangle$ consists of a fuzzy TBox $\mathcal{T}$, and a fuzzy ABox $\mathcal{A}$.

Semantics. The semantics extends [22]. The main idea is that concepts and roles are interpreted as fuzzy subsets of an interpretation's domain. Therefore, $\mathcal{ALCF}(\mathbf{D})$ axioms, rather than being satisfied (true) or unsatisfied (false) in an interpretation, become a degree of truth in $[0,1]$.

In the following, we use $\wedge, \vee, \neg$ and $\rightarrow$ in infix notation, in place of a t-norm, s-norm, negation function, and implication function. A *fuzzy interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ relative to a fuzzy data type theory $\mathbf{D} = \langle \Delta_{\mathbf{D}}, \cdot_{\mathbf{D}} \rangle$ consists of a nonempty set $\Delta^{\mathcal{I}}$ (called the *domain*), disjoint from $\Delta_{\mathbf{D}}$, and of a *fuzzy interpretation function* $\cdot^{\mathcal{I}}$ that coincides with $\cdot_{\mathbf{D}}$ on every data value, data type, and fuzzy data type predicate, and it assigns (i) to each abstract concept C a function $C^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \rightarrow [0,1]$; (ii) to each abstract role R a function $R^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0,1]$; (iii) to each abstract feature r a partial function $r^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0,1]$ such that for all $u \in \Delta^{\mathcal{I}}$ there is an unique $w \in \Delta^{\mathcal{I}}$ on which $r^{\mathcal{I}}(u,w)$ is defined; (iv) to each concrete role T a function $R^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta_{\mathbf{D}} \rightarrow [0,1]$; (v) to each concrete feature t a partial function $t^{\mathcal{I}}\colon \Delta^{\mathcal{I}} \times \Delta_{\mathbf{D}} \rightarrow [0,1]$ such that for all $u \in \Delta^{\mathcal{I}}$ there is an unique $o \in \Delta_{\mathbf{D}}$ on which $t^{\mathcal{I}}(u,o)$ is defined; (vi) to each modifier m the modifier function $f_m\colon [0,1] \rightarrow [0,1]$; (vi) to each abstract individual a an element in $\Delta^{\mathcal{I}}$; (vii) to each concrete individual c an element in $\Delta_{\mathbf{D}}$.

The mapping $\cdot^{\mathcal{I}}$ is extended to roles and complex concepts as specified in the Table 1 (where $x, y \in \Delta^{\mathcal{I}}$ and $v \in \Delta_{\mathbf{D}}$).

We comment briefly some points. The semantics of $\exists R.C$ $(\exists R.C)^{\mathcal{I}}(d) = \sup_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)$ is the result of viewing $\exists R.C$ as the open first order formula $\exists y.F_R(x,y) \wedge F_C(y)$ (where F is the obvious translation of roles and concepts into first-order logic (FOL)). Similarly, $(\forall R.C)^{\mathcal{I}}(x) = \inf_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \rightarrow C^{\mathcal{I}}(y)$ is related to the open first order formula $\forall y.F_R(x,y) \rightarrow F_C(y)$. However, unlike the classical case, in general, we do not have that $(\forall R.C)^{\mathcal{I}} = (\neg \exists R.\neg C)^{\mathcal{I}}$. For instance, it holds in Łukasiewicz logic, but not in Gödel logic. Also interesting is that (see [5]) the axiom $\top \sqsubseteq \neg(\forall R.A) \sqcap (\neg \exists R.\neg A)$ has no classical model. However, in [5], it is shown that in Gödel logic it has no finite model, but has an infinite model.

Finally, the mapping $\cdot^{\mathcal{I}}$ is extended to non-fuzzy axioms as specified in the following table (where a, b are individuals):

$$\begin{array}{rcl} (C \sqsubseteq D)^{\mathcal{I}} & = & \inf_{x \in \Delta^{\mathcal{I}}} C^{\mathcal{I}}(x) \rightarrow D^{\mathcal{I}}(x) \\ (a{:}C)^{\mathcal{I}} & = & C^{\mathcal{I}}(a^{\mathcal{I}}) \\ ((a,b){:}R)^{\mathcal{I}} & = & R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \ . \end{array}$$

Note here that e.g. the semantics of a concept inclusion axiom $C \sqsubseteq D$ is derived directly from its FOL translation, which is of the form $\forall x.F_C(x) \rightarrow F_D(x)$. This definition is clearly different from the approaches in which $C \sqsubseteq D$ is viewed as $\forall x.C(x) \leqslant D(x)$. This latter approach has the effect that the subsumption relationship is a boolean relationship, while the in former approach subsumption is determined up to a degree in $[0,1]$.

The notion of *satisfaction* of a fuzzy axiom E by a fuzzy interpretation $\mathcal{I}$, denoted $\mathcal{I} \models E$, is defined as follows: $\mathcal{I} \models \langle \alpha \geqslant n \rangle$, where α is a concept inclusion axiom, iff $\alpha^{\mathcal{I}} \geqslant n$. Similarly, $\mathcal{I} \models \langle \alpha \geqslant n \rangle$, where α is a concept or a role assertion axiom, iff $\alpha^{\mathcal{I}} \geqslant n$. We say that a concept C is *satisfiable* iff there is an interpretation $\mathcal{I}$ and an individual $x \in \Delta^{\mathcal{I}}$ such that $C^{\mathcal{I}}(x) > 0$.

For a set of fuzzy axioms $\mathcal{E}$, we say that $\mathcal{I}$ *satisfies* $\mathcal{E}$ iff I satisfies each element in $\mathcal{E}$. We say that $\mathcal{I}$ is a *model* of E (resp. $\mathcal{E}$) iff $\mathcal{I} \models E$ (resp. $\mathcal{I} \models \mathcal{E}$). $\mathcal{I}$ *satisfies* (is a *model* of) a fuzzy knowledge base $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$, denoted $\mathcal{I} \models \mathcal{K}$, iff $\mathcal{I}$ is a model of each component $\mathcal{T}, \mathcal{R}$ and $\mathcal{A}$, respectively.

A fuzzy axiom E is a *logical consequence* of a knowledge base $\mathcal{K}$, denoted $\mathcal{K} \models E$ iff every model of $\mathcal{K}$ satisfies E.

The interesting point is that according to our semantics, e.g., a minor is a young person to a degree and this relationship is obtained without explicitly mentioning it.

Example 2.1 Consider the knowledge base $\mathcal{K}$ with definitions (1) and (2). Then under Łukasiewicz logic we have that (see [24])

$$\begin{array}{rcl} \mathcal{K} & \models & \langle \mathit{Minor} \sqsubseteq \mathit{YoungPerson}, 0.6 \rangle \\ \mathcal{K} & \models & \langle \mathit{YoungPerson} \sqsubseteq \mathit{Minor}, 0.4 \rangle \ . \end{array}$$

□

Finally, given $\mathcal{K}$ and an axiom α of the form $C \sqsubseteq D$, $a{:}C$, $(a,b){:}R$ or $(a,c){:}T$, it is of interest to compute α's best lower degree value bound. The *greatest lower bound* of α w.r.t. $\mathcal{K}$ (denoted $glb(\mathcal{K}, \alpha)$) is

$$glb(\mathcal{K}, \alpha) = \sup\{n \mid \mathcal{K} \models \langle \alpha \geqslant n \rangle\} \ ,$$

where $\sup \emptyset = 0$. Determining the glb is called the *Best Degree Bound* (BDB) problem. For instance, the consequences in Example 2.1 are the best possible degree bounds. Another similar concept is the *best satisfiability bound* of a concept C and amounts to determine

$$glb(\mathcal{K}, C) = \sup_{\mathcal{I}} \sup_{x \in \Delta^{\mathcal{I}}} \{C^{\mathcal{I}}(x) \mid \mathcal{I} \models \mathcal{K}\} \ .$$

Essentially, among all models $\mathcal{I}$ of the knowledge base, we are determining the maximal degree of truth that the concept C may have over all individuals $x \in \Delta^{\mathcal{I}}$.

Example 2.2 Assume, that a car seller sells an Audi TT for \$31500, as from the catalog price. A buyer is looking for a sports car, but wants to pay no more than around \$30000. In classical DLs no agreement can be found. The problem relies on the crisp condition on the seller's and the buyer's price. A more fine grained approach would be (and usually happens in negotiation) to consider prices as concrete fuzzy sets instead. For instance, the seller may consider optimal to sell above \$31500, but can go down to \$30500. The buyer prefers to spend less than \$30000, but can go up to

$$
\begin{array}{rcl}
\bot^{\mathcal{I}}(x) & = & 0 \\
\top^{\mathcal{I}}(x) & = & 1 \\
(C_1 \sqcap C_2)^{\mathcal{I}}(x) & = & {C_1}^{\mathcal{I}}(x) \wedge {C_2}^{\mathcal{I}}(x) \\
(C_1 \sqcup C_2)^{\mathcal{I}}(x) & = & {C_1}^{\mathcal{I}}(x) \vee {C_2}^{\mathcal{I}}(x) \\
(\neg C)^{\mathcal{I}}(x) & = & \neg C^{\mathcal{I}}(x)) \\
(m(C))^{\mathcal{I}}(x) & = & f_m(C^{\mathcal{I}}(x)) \\
(\forall R.C)^{\mathcal{I}}(x) & = & \inf_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \rightarrow C^{\mathcal{I}}(y) \\
(\exists R.C)^{\mathcal{I}}(x) & = & \sup_{y \in \Delta^{\mathcal{I}}} R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y) \\
(\forall T_1, \ldots, T_n.D)^{\mathcal{I}}(x) & = & \inf_{y_1,\ldots,y_n \in \Delta_{\mathbf{D}}} (\bigwedge_{i=1}^{n} {T_i}^{\mathcal{I}}(x,y_i)) \rightarrow D^{\mathcal{I}}(y_1,\ldots,y_n) \\
(\exists T_1, \ldots, T_n.D)^{\mathcal{I}}(x) & = & \sup_{y_1,\ldots,y_n \in \Delta_{\mathbf{D}}} (\bigwedge_{i=1}^{n} {T_i}^{\mathcal{I}}(x,y_i)) \wedge D^{\mathcal{I}}(y_1,\ldots,y_n)\,.
\end{array}
$$

Table 1: Fuzzy DL semantics.

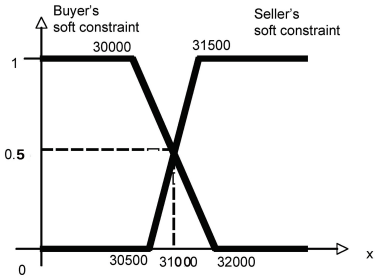


Figure 2: The soft price constraints.

$32000. We may represent these statements by means of the following axioms (see Figure 2):

$$
\begin{array}{rcl}
AudiTT & = & SportsCar \sqcap \exists hasPrice.R(x; 30500, 31500) \\
Query & = & SportsCar \sqcap \exists hasPrice.L(x; 30000, 32000)
\end{array}
$$

where $hasPrice$ is a concrete feature (a car has only one price, which is a number). Then we may find out that the highest degree to which the concept $C = AudiTT \sqcap Query$ is satisfiable is 0.5 (the *possibility* that the Audi TT and the query matches is 0.5). That is, $glb(\mathcal{K}, C) = 0.5$ and corresponds to the point where both requests intersects (i.e., the car may be sold at $31000).

3 Reasoning algorithm

Our procedure is inspired on [23]. We present here the method for reasoning in $\mathcal{ALCF}$ with GCI's under Łukasiewicz semantics. We leave out, for reasons of space, how to deal with fuzzy concrete domains and modifiers. The interested reader may consult [23]. We note however, that the procedure differs from the one presented in [23] in such a way that it will be better suited to deal with GCIs.

Note that $(\forall R.C)^{\mathcal{I}} = (\neg\exists R.\neg C)^{\mathcal{I}}$. This allows us to transform concept expressions into a semantically equivalent *Negation Normal Form* (NNF), *which is obtained by pushing in the usual manner negation on front of concept names, modifiers and concrete predicate names only*. With $nnf(C)$ we denote the NNF of concept C.

The basic idea behind our reasoning algorithm is as follows. Consider $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$. In order to solve the BDB problem, we combine appropriate DL tableaux rules with methods developed in the context of *Many-Valued Logics* (MVLs) [4]. In order to determine e.g. $glb(\mathcal{K}, a{:}C)$, we consider an expression of the form $\langle a{:}\neg C, 1-x \rangle$ (informally, $\langle a{:}C \leqslant x \rangle$), where x is a $[0,1]$-valued variable. Then we construct a tableaux for $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \cup \{\langle a{:}\neg C, \neg x \rangle\} \rangle$ in which the application of satisfiability preserving rules generates new fuzzy assertion axioms together with *inequations* over $[0,1]$-valued variables. These inequations have to hold in order to respect the semantics of the DL constructors. Finally, in order to determine the greatest lower bound, we *minimize* the original variable x such that all constraints are satisfied [1]. Similarly, for $C \sqsubseteq D$, we can compute $glb(\mathcal{K}, C \sqsubseteq D)$ as the minimal value of x such that $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \cup \{\langle a{:}C \sqcap \neg D, 1-x \rangle\} \rangle$ is satisfiable, where a is new abstract individual. Therefore, the BDB problem can be reduced to minimal satisfiability problem of a KB. Finally, concerning the satisfiability bound problem, $glb(\mathcal{K}, C)$ is determined by the maximal value of x such that $\langle \mathcal{T}, \mathcal{A} \cup \{\langle a{:}C, x \rangle\} \rangle$ is satisfiable.

Under Łukasiewicz logic we end up with a *bounded Mixed Integer Linear Program* (bMILP) optimization problem [15]. Interestingly, as for the MVL case, the tableaux we are generating contains *one* branch only and, thus, just *one* bMILP problem has to be solved.

Now, let $\mathbf{V}$ be a new alphabet of variables x ranging over $[0,1]$, $\mathbf{W}$ be a new alphabet of 0-1 variables y. We extend fuzzy assertions to the form $\langle \alpha, l \rangle$, where l is a linear expression over variables in $\mathbf{V}$, $\mathbf{W}$ and real values.

Similar to crisp DLs, our tableaux algorithm checks the satisfiability of a fuzzy KB by trying to build a fuzzy tableau,

[1] Informally, suppose the minimal value is $\bar{n}$. We will know then that for any interpretation $\mathcal{I}$ satisfying the knowledge base such that $(a{:}C)^{\mathcal{I}} < \bar{n}$, the starting set is unsatisfiable and, thus, $(a{:}C)^{\mathcal{I}} \geqslant \bar{n}$ has to hold. Which means that $glb(\mathcal{K}, (a{:}C)) = \bar{n}$

from which it is immediate either to build a model in case KB is satisfiable or to detect that the KB is unsatisfiable. The fuzzy tableau we present here can be seen as an extension of the tableau presented in [11], and is inspired by the one presented in [19, 20].

Given $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$, let $\mathbf{R}_\mathcal{K}$ be the set of roles occurring in $\mathcal{K}$ and let $sub(\mathcal{K})$ be the set of named concepts appearing in $\mathcal{K}$. A *fuzzy tableau* T for $\mathcal{K}$ is a quadruple $(\mathbf{S}, \mathcal{L}, \mathcal{E}, \mathcal{V})$ such that: $\mathbf{S}$ is a set of elements, $\mathcal{L} : \mathbf{S} \times sub(\mathcal{K}) \to [0, 1]$ maps each element and concept, to a membership degree (the degree of the element being an instance of the concept), and $\mathcal{E} : \mathbf{R}_\mathcal{K} \times (\mathbf{S} \times \mathbf{S}) \to [0, 1]$ maps each role of $\mathbf{R}_\mathcal{K}$ and pair of elements to the membership degree of the pair being an instance of the role, and $\mathcal{V} : \mathbf{I}_\mathcal{A} \to \mathbf{S}$ maps individuals occurring in $\mathcal{A}$ to elements in $\mathbf{S}$. For all $s, t \in \mathbf{S}$, $C, D \in sub(\mathcal{K})$, and $R \in \mathbf{R}_\mathcal{K}$, T has to satisfy:

1. $\mathcal{L}(s, \bot) = 0$ and $\mathcal{L}(s, \top) = 1$ for all $s \in \mathbf{S}$,
2. If $\mathcal{L}(s, \neg A) \geqslant n$, then $\mathcal{L}(s, A) \leqslant \neg n$.
3. If $\mathcal{L}(s, C \sqcap D) \geqslant n$, then $\mathcal{L}(s, C) \geqslant m_1$, $\mathcal{L}(s, D) \geqslant m_2$ and $n = m_1 \wedge m_2$, for some m_1 and m_2.
4. If $\mathcal{L}(s, C \sqcup D) \geqslant n$, then $\mathcal{L}(s, C) \geqslant m_1$, $\mathcal{L}(s, D) \geqslant m_2$ and $n = m_1 \vee m_2$, for some m_1 and m_2.
5. If $\mathcal{L}(s, \forall R.C) \geqslant n$, then $\mathcal{E}(R, \langle s, t \rangle) \leqslant \mathcal{L}(t, C) + 1 - n$ for all $t \in \mathbf{S}$.
6. If $\mathcal{L}(s, \exists R.C) \geqslant n$, then there exists $t \in \mathbf{S}$ such that $\mathcal{E}(R, \langle s, t \rangle) \geqslant m_1$, $\mathcal{L}(t, C) \geqslant m_2$ and $n = m_1 \wedge m_2$, for some m_1 and m_2.
7. If $\langle C \sqsubseteq D, n \rangle \in \mathcal{T}$, then $\mathcal{L}(s, C) \leqslant \mathcal{L}(s, D) + 1 - n$, for all $s \in \mathbf{S}$.
8. If $\langle a{:}C, n \rangle \in \mathcal{A}$, then $\mathcal{L}(\mathcal{V}(a), C) \geqslant n$.
9. If $\langle (a, b){:}R, n \rangle \in \mathcal{A}$, then $\mathcal{E}(R, \langle \mathcal{V}(a), \mathcal{V}(b) \rangle) \geqslant n$.
10. If $\langle (a, c){:}T, n \rangle \in \mathcal{A}$, then $\mathcal{E}(T, \langle \mathcal{V}(a), \mathcal{V}(c) \rangle) \geqslant n$.

Proposition 3.1 *$\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ is satisfiable iff there exists a fuzzy tableau for $\mathcal{K}$.*

Proof:[Sketch] For the *if* direction if $T = (\mathbf{S}, \mathcal{L}, \mathcal{E}, \mathcal{V})$ is a fuzzy tableau for $\mathcal{K}$, we can construct a fuzzy interpretation $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ that is a model of $\mathcal{A}$ and $\mathcal{T}$ as follows:

$\Delta^\mathcal{I} = \mathbf{S}$, $a^\mathcal{I} = \mathcal{V}(a)$, a occurs in $\mathcal{A}$
$\top^\mathcal{I}(s) = \mathcal{L}(s, \top)$, $\bot^\mathcal{I}(s) = \mathcal{L}(s, \bot)$, for all $s \in \mathbf{S}$
$A^\mathcal{I}(s) = \mathcal{L}(s, A)$, for all $s \in \mathbf{S}$
$R^\mathcal{I}(s, t) = \mathcal{E}(R, \langle s, t \rangle)$ for all $\langle s, t \rangle \in \mathbf{S} \times \mathbf{S}$

To prove that $\mathcal{I}$ is a model of $\mathcal{A}$ and $\mathcal{T}$, we can show by induction on the structure of concepts that $\mathcal{L}(s, C) \geqslant n$ implies $C^\mathcal{I}(s) \geqslant n$ for all $s \in \mathbf{S}$. Together with properties 7–10, this implies that $\mathcal{I}$ is a model of $\mathcal{T}$, and that it satisfies each fuzzy assertion in $\mathcal{A}$.

For the converse, we know from [5, 6] that if $\mathcal{K}$ has a model then it has a *witnessed model*. That is, $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ is a witnessed model of $\mathcal{K}$, if for all $x \in \Delta^\mathcal{I}$ there is $y \in \Delta^\mathcal{I}$ such that

$$(\exists R.C)^\mathcal{I}(x) = R^\mathcal{I}(x, y) \wedge C^\mathcal{I}(y)$$

and there is $x \in \Delta^\mathcal{I}$ such that

$$(C \sqsubseteq D)^\mathcal{I} = C^\mathcal{I}(x) \to D^\mathcal{I}(x) \, .$$

So, let $\mathcal{I}$ be a witnessed model of $\mathcal{K}$. Then a fuzzy tableau $T = (\mathbf{S}, \mathcal{L}, \mathcal{E}, \mathcal{V})$ for $\mathcal{K}$ can be defined as follows:

$\mathbf{S}{=}\ \Delta^\mathcal{I}$, $\mathcal{E}(R, \langle s, t \rangle) = R^\mathcal{I}(s, t)$, $\mathcal{L}(s, C) = C^\mathcal{I}(s)$, $\mathcal{V}(a) = a^\mathcal{I}$

It can be verified that T is a fuzzy tableau for $\mathcal{K}$. □

Now, in order to decide the satisfiability of $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ a procedure that constructs a fuzzy tableau T for $\mathcal{K}$ has to be determined. Like the tableaux algorithm presented in [20], our algorithm works on *completion-forests* since an ABox might contain several individuals with arbitrary roles connecting them. Due to the presence of general or cyclic terminology $\mathcal{T}$, the termination of the algorithm has to be ensured. This is done by providing a *blocking* condition for rule applications.

Let $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ be a fuzzy KB. A completion-forest $\mathcal{F}$ for $\mathcal{K}$ is a collection of trees whose distinguished roots are arbitrarily connected by edges.

Each node v is labelled with a sequence $\mathcal{L}(v)$ of expressions of the form $\langle C, l \rangle$, where $C \in sub(\mathcal{K})$, and l is either a rational, a variable x, or a negated variable, i.e. of the form $1 - x$, where x is a variable (the intuition here is that v is an instance of C to degree equal or greater than of the evaluation of l).

Each edge $\langle v, w \rangle$ is labelled with a sequence $\mathcal{L}(\langle v, w \rangle)$ of expressions of the form $\langle R, l \rangle$, where $R \in \mathbf{R}_\mathcal{K}$ are roles occurring in $\mathcal{K}$(the intuition here is that $\langle v, w \rangle$ is an instance of R to degree equal or greater than of the evaluation of l).

The forest has associated a set $\mathcal{C}_\mathcal{F}$ of constraints of the form $c \leqslant c', c = c', x_i \in [0, 1], y_i \in \{0, 1\}$, on the variables occurring the node labels and edge labels. c, c' are linear expressions.

If nodes v and w are connected by an edge $\langle v, w \rangle$ with $\langle R, l \rangle$ occurring in $\mathcal{L}(\langle v, w \rangle)$, then w is called an R_l-*successor* of v and w is called an R_l-*predecessor* of w.

A node v is an R-successor (resp. R-predecessor) of w if it is an R_l-successor (resp. R_l-predecessor) of w for some role R. As usual, *ancestor* is the transitive closure of *predecessor*.

We say that two non-root nodes v and w are *equivalent*, denoted $\mathcal{L}(v) \approx \mathcal{L}(w)$, if $\mathcal{L}(v) = [\langle C_1, l_1 \rangle, \ldots, \langle C_n, l_k \rangle]$, $\mathcal{L}(w) = [\langle C_1, l'_1 \rangle, \ldots, \langle C_n, l'_k \rangle]$, and for all $1 \leqslant i \leqslant k$, either both l_i and l'_i are variables, or both l_i and l'_i are negated variables or both l_i and l'_i are the same rational in $[0, 1]$ (the intuition here is that v and w share the same properties).

A node v is *directly blocked* iff none of its ancestors are blocked, it is not a root node, and it has an ancestor w such that $\mathcal{L}(v) \approx \mathcal{L}(w)$. In this case, we say w directly blocks v. A node v is *blocked* iff it is directly blocked or if one of its predecessor is blocked (the intuition here is that we need not further to apply rules to node v, as an equivalent predecessor node w of v exists).

The algorithm initializes a forest $\mathcal{F}$ to contain (i) a root node v_0^i, for each individual a_i occurring in $\mathcal{A}$, labelled with $\mathcal{L}(v_0^i)$ such that $\mathcal{L}(v_0^i)$ contains $\langle C_i, n\rangle$ for each fuzzy assertion $\langle a_i{:}C_i, n\rangle \in \mathcal{A}$, and (ii) an edge $\langle v_0^i, v_0^j\rangle$, for each fuzzy assertion $\langle (a_i, a_j){:}R_i, n\rangle \in \mathcal{A}$, labelled with $\mathcal{L}(\langle v_0^i, v_0^j\rangle)$ such that $\mathcal{L}(\langle v_0^i, v_0^j\rangle)$ contains $\langle R_i, n\rangle$. $\mathcal{F}$ is then expanded by repeatedly applying the completion rules described below. The completion-forest is complete when none of the completion rules are applicable. Then, the bMILP problem on the set of constraints $\mathcal{C}_\mathcal{F}$ is solved.

We also need a technical definition involving feature roles (see [14]). Let $\mathcal{F}$ be forest, r an abstract feature such that we have two edges $\langle v, w_1\rangle$ and $\langle v, w_1\rangle$ such that $\langle r, l_1\rangle$ and $\langle r, l_2\rangle$ occur in $\mathcal{L}(\langle v, w_1\rangle)$ and $\mathcal{L}(\langle v, w_2\rangle)$, respectively (informally, $\mathcal{F}$ contains $\langle (v, w_1){:}r, l_1\rangle$ and $\langle (v, w_2){:}r, l_2\rangle$). Then we call such a pair a *fork*. As r is a function, such a fork means that w_1 and w_2 have to be interpreted as the same individual. Such a fork can be deleted by adding both $\mathcal{L}(\langle v, w_2\rangle)$ to $\mathcal{L}(\langle v, w_1\rangle)$ and $\mathcal{L}(w_2)$ to $\mathcal{L}(w_1)$, and then deleting node w_2. A similar argument applies to concrete feature roles. At the beginning, we remove the forks from the initial forest. We assume that forks are eliminated as soon as they appear (as part of a rule application) with the proviso that newly generated nodes are replaced by older ones and not vice-versa.

We also assume a fixed rule application strategy as e.g. the order of rules below, such that the rule for $(\exists)$ is applied as last. Also, all expressions in node labels are processed according to the order they are introduced into the $\mathcal{F}$.

With x_α we denote the variable associated to the *atomic assertion* α of the form $a{:}A$ or $(a, b){:}R$. x_α will take the truth value associated to α, while with x_c we denote the variable associated to the concrete individual c. The rules are the following:

- (A). If $\langle A, l\rangle \in \mathcal{L}(v)$ then $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x_{v:A} \geqslant l\} \cup \{x_{v:A} \in [0, 1]\}$.
- $(\bar{A})$. If $\langle \neg A, l\rangle \in \mathcal{L}(v)$ then $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x_{v:A} \leqslant 1 - l\} \cup \{x_{v:A} \in [0, 1]\}$.
- (R). If $\langle R, l\rangle \in \mathcal{L}(\langle v, w\rangle)$ then $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x_{(v,w):R} \geqslant l\} \cup \{x_{(v,w):R} \in [0, 1]\}$.
- $(\sqcap)$. If $\langle C \sqcap D, l\rangle \in \mathcal{L}(v)$ then append $\langle C, x_1\rangle$ and $\langle D, x_2\rangle$ to $\mathcal{L}(v)$, and $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{y \leqslant 1 - l, x_i \leqslant 1 - y, x_1 + x_2 = l + 1 - y, x_i \in [0, 1], y \in \{0, 1\}\}$, where x_i, y are new variables.
- $(\sqcup)$. If $\langle C \sqcup D, l\rangle \in \mathcal{L}(v)$ then append $\langle C, x_1\rangle$ and $\langle D, x_2\rangle$ to $\mathcal{L}(v)$, and $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x_1 + x_2 = l, x_i \in [0, 1]\}$, where x_i are new variables.
- $(\forall)$. If $\langle \forall R.C, l_1\rangle \in \mathcal{L}(v)$, $\langle R, l_2\rangle \in \mathcal{L}(\langle v, w\rangle)$ and the rule has not been already applied to this pair then append $\langle C, x\rangle$ to $\mathcal{L}(w)$ and $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x \geqslant l_1 + l_2 - 1, x \leqslant y, l_1 + l_2 - 1 \leqslant y, l_1 + l_2 \geqslant y, x \in [0, 1], y \in \{0, 1\}\}$, where x, y are new variables. The case for concrete roles is similar.
- $(\sqsubseteq)$. If $\langle C \sqsubseteq D, n\rangle \in \mathcal{T}$ and v is a node to which this rule has not yet been applied then append $\langle nnf(\neg C), 1 - x_1\rangle$ and $\langle D, x_2\rangle$ to $\mathcal{L}(v)$, and $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{x_1 \leqslant x_2 + 1 - n\}$.
- $(\exists)$. If $\langle \exists R.C, l\rangle \in \mathcal{L}(v)$ and v is not blocked then create a new node w and append $\langle R, x_1\rangle$ to $\mathcal{L}(\langle v, w\rangle)$ and $\langle C, x_2\rangle$ to $\mathcal{L}(w)$, and $\mathcal{C}_\mathcal{F} = \mathcal{C}_\mathcal{F} \cup \{y \leqslant 1 - l, x_i \leqslant 1 - y, x_1 + x_2 = l + 1 - y, x_i \in [0, 1], y \in \{0, 1\}\}$, where x_i, y are new variables. The case for concrete roles is similar.

Let's see an example.

Example 3.1 *Consider $\mathcal{K} = \langle \mathcal{T}, \mathcal{A}\rangle$, where $\mathcal{T} = \{\langle \exists R.C \sqsubseteq D, 1\rangle\}$ and $\mathcal{A} = \{\langle (a, b){:}R, 0.7\rangle, \langle b{:}C, 0.8\rangle\}$. Let us show that $glb(\mathcal{K}, a{:}D) = 0.5$. To this end, we have to determine the minimal value for x such that $\langle \mathcal{T}, \mathcal{A} \cup \{\langle a{:}\neg D, 1 - x\rangle\}\rangle$ is satisfiable.*

To start with, we construct a forest $\mathcal{F}$ with two root nodes a and b (one for each individual in $\mathcal{A}$). We process first $\langle (a, b){:}R, 0.7\rangle$, then $\langle b{:}C, 0.8\rangle$ and finally $\langle a{:}\neg D, 1 - x\rangle$. Therefore, we set $\mathcal{L}(a) = [\langle C, 0.8\rangle, \langle \neg D, 1 - x\rangle]$, $\mathcal{L}(\langle a, b\rangle) = [\langle R, 0.7\rangle]$ and $\mathcal{C}_\mathcal{F} = \{x \in [0, 1]\}$.

We first process $\langle R, 0.7\rangle \in \mathcal{L}(\langle a, b\rangle)$, apply rule (R) and, thus, add $x_{(a,b):R} \geqslant 0.7$ and $x_{(a,b):R} \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Then we process $\langle C, 0.8\rangle \in \mathcal{L}(a)$, apply rule (A) and, thus, add $x_{a:C} \geqslant 0.8$ and $x_{a:C} \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

We next process $\langle \neg D, 1 - x\rangle \in \mathcal{L}(a)$, apply rule $(\bar{A})$ and, thus, add $x_{a:D} \leqslant x$ and $x_{a:D} \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Now, we apply $(\sqsubseteq)$ to a and $\langle \exists R.C \sqsubseteq D, 1\rangle$, and, thus, we append $\langle \forall R.\neg C, 1 - x_1\rangle$ and $\langle D, x_2\rangle$ to $\mathcal{L}(a)$, and we add $x_1 \leqslant x_2$ and $x_1, x_2 \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Similarly, we apply $(\sqsubseteq)$ to b: we append $\langle \forall R.\neg C, 1 - x_3\rangle$ and $\langle D, x_4\rangle$ to $\mathcal{L}(b)$, and we add $x_3 \leqslant x_4$ and $x_3, x_4 \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Next we process $\langle \forall R.\neg C, 1 - x_1\rangle \in \mathcal{L}(a)$, apply rule $(\forall)$ to it and $\langle R, 0.7\rangle \in \mathcal{L}(\langle a, b\rangle)$ and, thus, we append $\langle \neg C, x_5\rangle$ to $\mathcal{L}(b)$, and we add $\{x_5 \geqslant 1 - x_1 + 0.7 - 1, x_5 \leqslant y, 1 - x_1 - 0.3 \leqslant y, 1 - x_1 + 0.7 \geqslant y, x_5 \in [0, 1], y \in \{0, 1\}\}$ to $\mathcal{C}_\mathcal{F}$.

We next process $\langle D, x_2\rangle \in \mathcal{L}(a)$, apply rule (A) and, thus, add $x_{a:D} \geqslant x_2$ and $x_{a:D} \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Next we process $\langle \forall R.\neg C, 1 - x_1\rangle \in \mathcal{L}(b)$, but no rule is applicable to it. Finally, we process $\langle D, x_4\rangle \in \mathcal{L}(b)$, apply rule (A) and, thus, add $x_{b:D} \geqslant x_4$ and $x_{b:D} \in [0, 1]$ to $\mathcal{C}_\mathcal{F}$.

Now the forest $\mathcal{F}$ is complete as no more rule is applicable and we consider the set of inequations $\mathcal{C}_\mathcal{F}$. It remains to solve the bMILP problem on $\mathcal{C}_\mathcal{F}$. Indeed, it holds that $glb(\mathcal{K}, a{:}D) = \min x.\mathcal{C}_\mathcal{F}$. It can be verified that this value is 0.5.

Example 3.2 illustrates the behaviour on cyclic terminologies and shows how a potential infinite cyclic computation is blocked.

a $\quad \mathcal{L}(a) = [\langle A, 0.8\rangle, \langle \neg A, 1 - x_1\rangle, \langle \exists R.A, x_2\rangle]$

$\quad \mathcal{L}(\langle a, w_1\rangle) = [\langle R, x_3\rangle]$

w_1 $\quad \mathcal{L}(w_1) = [\langle A, x_4\rangle, \langle \neg A, 1 - x_5\rangle, \langle \exists R.A, x_6\rangle]$

$\quad \mathcal{L}(\langle w_1, w_2\rangle) = [\langle R, x_7\rangle]$

w_2 $\quad \mathcal{L}(w_2) = [\langle A, x_8\rangle, \langle \neg A, 1 - x_9\rangle, \langle \exists R.A, x_{10}\rangle]$

w_2 blocked by w_1

$$\begin{aligned}\mathcal{C}_{\mathcal{F}} = \{&x_i \in [0,1], y_i \in \{0,1\}, x_{a:A} \geqslant 0.8, x_1 \leqslant x_2,\\ &x_{a:A} \geqslant x_1, y_1 \leqslant 1 - x_2, x_3 \leqslant 1 - y_1, x_4 \leqslant 1 - y_1,\\ &x_3 + x_4 = x_2 + 1 - y_1, x_{(a,w_1):R} \geqslant x_3, x_{w_1:A} \geqslant x_4,\\ &x_5 \leqslant x_6, x_{w_1:A} \leqslant x_5, y_2 \leqslant 1 - x_6, x_7 \leqslant 1 - y_2,\\ &x_8 \leqslant 1 - y_2, x_7 + x_8 = x_6 + 1 - y_2,\\ &x_{(w_1,w_2):R} \geqslant x_7, x_{w_2:A} \geqslant x_8, x_9 \leqslant x_{10}, x_{w_2:A} \leqslant x_9\}\end{aligned}$$

Figure 3: Complete forest for Example 3.2.

Example 3.2 *Consider $\mathcal{K} = \langle \mathcal{T}, \mathcal{A}\rangle$, where $\mathcal{T} = \{\langle A \sqsubseteq \exists R.A, 1\rangle\}$ and $\mathcal{A} = \{\langle a{:}A, 0.8\rangle\}$. Let's show that $\mathcal{K}$ is satisfiable. The complete forest $\mathcal{F}$ in Figure 3 shows the computation.*

As we can notice, node w_2 is blocked by node w_1, meaning that w_1 and w_2 share the same properties. In order to build a model, we replace all occurrences of w_2 in $\mathcal{C}_{\mathcal{F}}$ with w_1 and then we find a solution to the inequalities. Our blocking condition is based on the fact that, in case of cyclic definitions, as the calculus is deterministic, the same sequence of expressions has to be generated *within a cycle.*

Proposition 3.2 (Termination) *For each KB $\mathcal{K}$, the tableau algorithm terminates.*

Proof: [Sketch] Termination is a result of the properties of the expansion rules, as in the classical case [11]. More precisely we have the following observations. (i) The expansion rules never remove nodes from the tree (except forks at the beginning) or concepts from node labels or change the edge labels. (ii) Successors are only generated by the rule $\exists$. For any node and for each concept these rules are applied at-most once. (iii) Since nodes are labelled with nonempty sequences of $sub(\mathcal{K})$, obviously there is a finite number of possible labelling for a pair of nodes. Thus, the blocking condition will be applied in any path of the tree and consequently any path will have a finite length. □

Proposition 3.3 (Soundness) *If the expansion rules can be applied to a KB $\mathcal{K}$ such that they yield a complete completion-forest $\mathcal{F}$ such that $\mathcal{C}_{\mathcal{F}}$ has a solution, then $\mathcal{K}$ has a fuzzy tableau for $\mathcal{K}$.*

Proof:[Sketch] Let $\mathcal{F}$ be a complete completion-forest constructed by the tableaux algorithm for $\mathcal{K}$. By hypothesis, $\mathcal{C}_{\mathcal{F}}$ has a solution. If x is a variable occurring in $\mathcal{C}_{\mathcal{F}}$, with $\bar{x}$ we denote the value of x in this solution. If the variable x does not occur in $\mathcal{C}_{\mathcal{F}}$ then $\bar{x} = 0$ is assumed. A fuzzy tableau $T = (\mathbf{S}, \mathcal{L}, \mathcal{E}, \mathcal{V})$ can be defined as follows:

$$\begin{aligned}\mathbf{S} &= \{v \mid v \text{ is a node in } \mathcal{F}, \text{ and } v \text{ is not blocked}\},\\ \mathcal{L}(v, \bot) &= 0, \text{ if } v \in \mathbf{S},\\ \mathcal{L}(v, \top) &= 1, \text{ if } v \in \mathbf{S},\\ \mathcal{L}(v, A) &= \bar{x}_{v:A}, \text{ if } v \text{ in } \mathcal{F} \text{ not blocked},\\ \mathcal{E}(R, \langle v, w\rangle) &= \bar{x}_{(v,w):R}, \text{ if } v, w \text{ in } \mathcal{F} \text{ not blocked}\\ \mathcal{E}(R, \langle v, w\rangle) &= \bar{x}_{(v,w'):R}, \text{ if } v \text{ in } \mathcal{F} \text{ not blocked, } w \text{ blocks } w'\\ \mathcal{V}(a_i) &= v_0^i, \text{ where } v_0^i \text{ is a root node}\end{aligned}$$

It can be shown that T is a fuzzy tableau for $\mathcal{K}$. □

Proposition 3.4 (Completeness) *Consider a KB $\mathcal{K}$. If $\mathcal{K}$ has a fuzzy tableau, then the expansion rules can be applied in such a way that the tableaux algorithm yields a complete completion-forest for $\mathcal{K}$ such that $\mathcal{C}_{\mathcal{F}}$ has a solution.*

Proof:[Sketch] Let $T = (\mathbf{S}, \mathcal{L}, \mathcal{E}, \mathcal{V})$ be a fuzzy tableau for $\mathcal{K}$. Using T, we can trigger the application of the expansion rules such that they yield a completion-forest $\mathcal{F}$ that is complete. Using $\mathcal{L}$ and $\mathcal{E}$ we can find a solution to $\mathcal{C}_{\mathcal{F}}$. □

4 Conclusions

In this work we presented a reasoning algorithm for $\mathcal{ALCF}(D)$ with general concept inclusions and explicit membership functions ($\mathcal{ALCF}$ is a guarded logic [3]), under Łukasiewicz semantics. Clearly, by using [23] the result applies also to the usually used "Zadeh semantics" of fuzzy DLs, and of course also for the classical boolean variant of $\mathcal{ALCF}$.

The result can be extended to more expressive fuzzy DLs, such as $\mathcal{SHIF}(D)$ and $\mathcal{SHOIN}(D)$, which are the DLs behind the web ontology description languages OWL-DL and OWL-Lite, by adapting our blocking condition similarly as done in [10]. The description of these blocking conditions will be the subject of an extend paper. So far, an implementation for fuzzy $\mathcal{SHIF}(D)$ can be found from Straccia's home page.

Acknowledgement

Fernando Bobillo holds a FPU scholarship from the Spanish Ministerio de Educación y Ciencia.

References

[1] F. Baader, D. Calvanese, D. McGuinness, D. Nardi and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.

[2] P. Bonatti and A. Tettamanzi Some complexity results on fuzzy description logics. In A. Petrosino V. Di Gesù, F. Masulli, editor, *WILF 2003 International Workshop on Fuzzy*

Logic and Applications, volume 2955 in Lecture Notes in Artificial Intelligence. 2003. Springer Verlag.

[3] E. Grädel, C. Hirsch and M. Otto. Back and forth between guarded and modal logics. *ACM Transactions on Computational Logic*, 3(3):418–463, 2002.

[4] R. Hähnle. Advanced many-valued logics. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 2. Kluwer, 2001.

[5] P. Hájek. Making fuzzy description logics more general. *Fuzzy Sets and Systems*, 154 (1), 1–15, 2005.

[6] P. Hájek. What does mathematical fuzzy logic offer to description logic? In Elie Sanchez, editor, *Capturing Intelligence: Fuzzy Logic and the Semantic Web*. Elsevier, 2006.

[7] S. Hölldobler, N. H. Nga and T. D. Khang. The fuzzy description logic ALC_{FH}. In *Proceeedings of the International Workshop on Description Logics (DL-05)*, 2005.

[8] S. Hölldobler, H.-P. Störr and T. D. Khang. The fuzzy description logic ALC_{FH} with hedge algebras as concept modifiers. *Journal of Advanced Computational Intelligence*, 7(3):294–305, 2003.

[9] S. Hölldobler, H.-P. Störr and T. D. Khang. The subsumption problem of the fuzzy description logic ALC_{FH}. In *Proceedings of the 10th International Conference on Information Processing and Managment of Uncertainty in Knowledge-Based Systems, (IPMU-04)*, 2004.

[10] I. Horrocks, U. Sattler and S. Tobies. Practical reasoning for very expressive description logics. *Logic Journal of the IGPL*, 8(3):239–263, 2000.

[11] I. Horrocks, U. Sattler and S. Tobies. Reasoning with individuals for the description logic SHIQ. In D. MacAllester, editor, *Proceedings of the 17th International Conference on Automated Deduction (CADE-17)*, volume 1831 in Lecture Notes in Artificial Intelligence, pages 482–496, Germany, 2000. Springer Verlag.

[12] I. Horrocks, P. F. Patel-Schneider and F. van Harmelen. From SHIQ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.

[13] Y. Li, B. Xu, J. Lu, D. Kang and P. Wang. Extended fuzzy description logic ALCN. In R. Khosla, R. J. Howlett, and L. C. Jain, editors, *Proceedings of the 9th International Conference, Knowledge-Based Intelligent Information and Engineering Systems (KES-05)*, volume 3684 of *Lecture Notes in Computer Science*, pages 896–902. Springer Verlag, 2005.

[14] C. Lutz. Description logics with concrete domains - A survey. In *Advances in Modal Logics Volume 4*. King's College Publications, 2003.

[15] H. Salkin and M. Kamlesh. *Foundations of Integer Programming*. North-Holland, 1988.

[16] D. Sánchez and A. G.B. Tettamanzi. Fuzzy quantification in fuzzy description logics. In E. Sanchez, editor, *Capturing Intelligence: Fuzzy Logic and the Semantic Web*. Elsevier, 2006.

[17] M. Schmidt-Schauß and G. Smolka. Attributive concept descriptions with complements. *Artificial Intelligence*, 48:1–26, 1991.

[18] G. Stoilos, G. Stamou, V. Tzouvaras, J. Z. Pan and I. Horrocks. The fuzzy description logic f-SHIN. *Proceedings of the ISWC2005 Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2005)*, 67–76, 2005.

[19] G.Stoilos, G. Stamou, V. Tzouvaras, J. Z. Pan and Ian Horrocks. A Fuzzy Description Logic for Multimedia Knowledge Representation. In *Proceedings of the International Workshop on Multimedia and the Semantic Web*, 2005.

[20] G. Stoilos, U. Straccia, G. Stamou and J. Z. Pan. General concept inclusions in fuzzy description logics. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI-06)*, pages 457–461. IOS Press, 2006.

[21] U. Straccia. A framework for the retrieval of multimedia objects based on four-valued fuzzy description logics. In F. Crestani and G. Pasi, editors, *Soft Computing in Information Retrieval: Techniques and Applications*, pages 332–357. Physica Verlag (Springer Verlag), 2000.

[22] U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.

[23] U. Straccia. Description logics with fuzzy concrete domains. In F. Bachus and T. Jaakkola, editors, *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI-05)*, pages 559–567, Edinburgh, Scotland, 2005. AUAI Press.

[24] U. Straccia. Fuzzy description logics with concrete domains. Technical Report 2005-TR-03, Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa, Italy, 2005.

[25] U. Straccia. Towards a fuzzy description logic for the semantic web (preliminary report). In *2nd European Semantic Web Conference (ESWC-05)*, volume 3532 in Lecture Notes in Computer Science, pages 167–181, Crete, Greece, 2005. Springer Verlag.

[26] U. Straccia. A fuzzy description logic for the semantic web. In E. Sanchez, editor, *Fuzzy Logic and the Semantic Web*, Capturing Intelligence, chapter 4, pages 73–90. Elsevier, 2006.

[27] U. Straccia. Uncertainty and description logic programs over lattices. In E. Sanchez, editor, *Fuzzy Logic and the Semantic Web*, Capturing Intelligence, chapter 7, pages 115–133. Elsevier, 2006.

[28] C. Tresp and R. Molitor. A description logic for vague knowledge. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI-98)*, 1998.

[29] J. L. Y. Li, B. Xu and D. Kang. Discrete tableau algorithms for SHI. In *Proceeedings of the International Workshop on Description Logics (DL-06)*. CEUR, 2006.

[30] J. Yen. Generalizing term subsumption languages to fuzzy logic. In *Proc. of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)*, pages 472–477, 1991.

[31] L. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

[32] L. Zhang, Y. Yu, J. Zhou, C. Lin and Y. Yang. An enhanced model for searching in semantic portals. In *WWW'05: Proc. of the 14th international conference on World Wide Web*, pages 453–462, New York, NY, USA, 2005. ACM Press.

From Fuzzy- to Bipolar- Datalog

Ágnes Achs
Department of Computer Science, Faculty of Engineering, University of Pécs,
Pécs, Boszorkány u. 2, Hungary
achs@witch.pmmf.hu

Abstract

Abstract. In this work we present several possible extensions of fuzzy Datalog. At first the concept of fuzzy Datalog will be summarized, then its extension for intuitionistic- and interval-valued fuzzy logic is given and the concept of bipolar fuzzy Datalog is introduced.

Keywords: intuitionistic fuzzy Datalog, interval-valued fuzzy Datalog, bipolar fuzzy Datalog

1 Introduction

The dominant portion of human knowledge can not be modeled by pure inference systems, because this knowledge is often ambiguous, incomplete and vague. The study of inference systems is tackled by several – and often very different – approaches. When knowledge is represented as a set of facts and rules, this uncertainty can be handled by means of fuzzy logic.

A few years ago in [8, 9] a possible combination of Datalog-like languages and fuzzy logic was presented. In these works the concept of fuzzy Datalog has been introduced by completing the Datalog-rules and facts by an uncertainty level and an implication operator. The level of a rule-head can be inferred from the level of the body, and the level of the rule by the implication operator of the rule. Based upon our previous works, later on a possible fuzzy knowledge-base was developed, the deduction mechanism of which is fuzzy Datalog [1]. Now a question has arisen: can the fuzzy knowledge-base be improved by any extension of deduction mechanism? In this paper we deal with some possible extensions of fuzzy Datalog without examining their influence on the knowledge-base; we extend it to intuitionistic- and interval-valued fuzzy logic and joining to the discussion in [2, 3] we introduce the concept of bipolar fuzzy Datalog.

On the next pages we summarize the concept of fuzzy Datalog, and then its new extensions will be presented.

2 Fuzzy Datalog

A Datalog program consists of facts and rules. In fuzzy Datalog (fDATALOG) we can complete the facts by an uncertainty level, the rules by an uncertainty level and an implication operator. This means that evaluating the fuzzy implication connecting to the rule, its truth-value according to the implication operator is at least the given uncertainty level. We can infer the level of a rule-head from the level of the body, and the level of the rule by the implication operator of the rule.

The notion of fuzzy rule is given in definition below.

Definition 1. An fDATALOG rule is a triplet r;β;I, where r is a formula of the form

$$A \leftarrow A_1,...,A_n \qquad (n \geq 0).$$

A is an atom (the head of the rule), $A_1,...,A_n$ are literals (the body of the rule); I is an implication operator and $\beta \in (0,1]$ (the level of the rule).

For getting finite result, all the rules in the program must be safe. An fDATALOG rule is safe if all variables occurring in the head also occur in the body, and all variables occurring in a negative literal also occur in a positive one. An fDATALOG program is a finite set of safe fDATALOG rules.

There is a special type of rule, called fact. A fact has the form $A \leftarrow ;\beta;I$. From now on, we refer to facts as (A, β), because according to implication I, the level of A easily can be computed.

The semantics of fDATALOG is defined as the fixed points of consequence transformations. Depending on these transformations, two semantics for fDATALOG can be defined [9]. The deterministic semantics is the least fixed point of the deterministic transformation DT_P, the nondeterministic semantics is the least fixed point of the nondeterministic transformation NT_P. According to the deterministic transformation, the rules of a program are evaluated in parallel, while in the nondeterministic case the rules are considered independently and sequentially. These transformations are the following:

Definition 2. Let B_P be the Herbrand base of the program P, and let $F(B_P)$ denote the set of all fuzzy sets over B_P. The consequence transformations DT_P: $F(B_P) \rightarrow F(B_P)$ and NT_P: $F(B_P) \rightarrow F(B_P)$ are defined as

$$DT_P(X) = \{\cup\{(A,\alpha_A)\}\} \cup X \text{ and} \quad (1)$$

$$NT_P(X) = \{(A, \alpha_A)\} \cup X, \quad (2)$$

where

$(A \leftarrow A_1,...,A_n ; \beta; I) \in \text{ground}(P)$, $(|A_i|, \alpha_{Ai}) \in X$, $1 \le i \le n$; $\alpha_A = \max(0, \min\{\gamma \mid I(\alpha_{body}, \gamma) \ge \beta\})$.

There ground(P) is the set of all possible rules of P the variables of which are replaced by ground terms of H_P, $|A_i|$ denotes the kernel of the literal A_i, (i.e., it is the ground atom A_i, if A_i is a positive literal, and $\neg A_i$, if A_i is negative) and $\alpha_{body} = \min(\alpha_{A_1},\ldots, \alpha_{A_n})$.

In [8] it is proved that starting from the set of facts, both DT_P and NT_P have fixed points which are the least fixed points in the case of positive P. These fixed points are denoted by $lfp(DT_P)$ and $lfp(NT_P)$. It was also proved, that $lfp(DT_P)$ and $lfp(NT_P)$ are models of P, so we could define $lfp(DT_P)$ as the deterministic semantics, and $lfp(NT_P)$ as the nondeterministic semantics of fDATALOG programs. For a function- and negation-free fDATALOG, the two semantics are the same, but they are different if the program has any negation. In this case the set $lfp(DT_P)$ is not always a minimal model, but the nondeterministic semantics – $lfp(NT_P)$ – is minimal under certain conditions. These conditions are referred to as stratification. Stratification gives an evaluating sequence in which the negative literals are evaluated first [10].

To compute the level of rule-heads, we need the next concept:

Definition 3. The uncertainty-level function is:

$$f(I, \alpha, \beta) = \min(\{\gamma \mid I(\alpha, \gamma) \ge \beta\}).$$

According to this function the level of a rule-head can be computed. In former papers several implications were detailed (the operators are detailed in [11]), for now three are chosen from these. The values of their uncertainty-level functions can be easily computed. They are the following:

Gödel:
$$I_G(\alpha, \gamma) = \begin{cases} 1 & \alpha \le \gamma \\ \gamma & \text{otherwise} \end{cases}$$

$$f(I_G, \alpha, \beta) = \min(\alpha, \beta)$$

Lukasiewicz:

$$I_L(\alpha, \gamma) = \begin{cases} 1 & \alpha \le \gamma \\ 1 - \alpha + \gamma & \text{otherwise} \end{cases}$$

$$f(I_L, \alpha, \beta) = \max(0, \alpha + \beta - 1)$$

Kleene-Dienes:

$$I_K(\alpha,\gamma) = \max(1-\alpha, \gamma)$$

$$f(I_K, \alpha, \beta) = \begin{cases} 0 & \alpha + \beta \le 1 \\ \beta & \alpha + \beta > 1 \end{cases}$$

Further on during the extensions of fDATALOG we deal only with these operators.

3 Extensions of fuzzy Datalog

In intuitionistic-(IFS) and interval-valued (IVS) fuzzy logic the uncertainty is represented by two values, $\mu = (\mu_1, \mu_2)$, while in "normal" fuzzy logic it is represented by a single value (μ). In the intuitionistic case the two elements must satisfy the condition $\mu_1+\mu_2 \le 1$, while in the interval-valued case the condition is $\mu_1 \le \mu_2$. In IFS μ_1 is the degree of membership, and μ_2 is the degree of non-membership, while in IVS the membership degree is between μ_1 and μ_2. It is obvious that the relation

$\mu'_1 = \mu_1$, $\mu'_2 = 1 - \mu_2$ create a mutual connection between the two systems.

In both cases an ordering relation can be defined, and according to this ordering a lattice is taking shape:

Definition 4.

$$L_F = \{(x_1,x_2)\in[0,1]^2 \mid x_1+x_2 \le 1\}, (x_1,x_2) \le_F (x_1,x_2) \Leftrightarrow x_1 \le y_1 \text{ and } x_2 \ge y_2$$

$$L_V = \{(x_1,x_2)\in[0,1]^2 \mid x_1 \le x_2\}, (x_1,x_2) \le_V (x1,x2) \Leftrightarrow x_1 \le y_1 \text{ and } x_2 \le y_2$$

are lattices of IFS and IVS respectively.

It can be proved that both L_F and L_V are complete lattices [4]. In both cases the extended fDATALOG is defined on these lattices, and the necessary concepts are generalizations of the ones presented in Definition 1 and Definition 2.

Definition 5. The extended fDATALOG program (efDATALOG) is a finite set of safe efDATALOG rules (r; $\underline{\beta}$; $\underline{I}_{FV}$);

– the extended consequence transformations eDT_P and eNT_P are formally the same as DT_P and NT_P in (1), (2) except:

$$\underline{\alpha}_A = \max(\underline{0}_{FV}, \min\{\underline{\gamma} \mid \underline{I}_{FV}(\underline{\alpha}_{body}, \underline{\gamma}) \ge_{FV} \underline{\beta}\}) \text{ and}$$

– the extended uncertainty-level function is

$$f(\underline{I}_{FV},\underline{\alpha},\underline{\beta}) = \min(\{\underline{\gamma} \mid \underline{I}_{FV}(\underline{\alpha}, \underline{\gamma}) \ge_{FV} \underline{\beta}\}),$$

where $\underline{\alpha}$, $\underline{\beta}$, $\underline{\gamma}$ are elements of L_F, L_V respectively,
$\underline{I}_{FV} = \underline{I}_F$ or $\underline{I}_V$ is an implication of L_F or L_V,
$\underline{0}_{FV}$ is $\underline{0}_F = (0,1)$ or $\underline{0}_V = (0,0)$ and $\ge_{FV}$ is $\ge_F$ or $\ge_V$.

As eDTP and eNTP are inflationary transformations over the complete lattices LF or LV, thus according to [11] they have an inflationary fixed point denoted by lfp(eDTP) and lfp(eNTP). If P is positive (without negation), then eDTP = eNTP is a monotone transformation, so lfp(eDTP) = lfp(eNTP) is the least fixed point. The next question is whether this fixed point is a model of P. The fixed point is an interpretation of P, which is a model, if for each $(A \leftarrow A_1,...,A_n; \underline{\beta}; \underline{I}) \in \text{ground}(P)$ $\underline{I}(\underline{\alpha}_{body}, \underline{\alpha}_A) \ge_{FV} \underline{\beta}$.

Proposition 1. $lfp(eDT_P)$ and $lfp(eNT_P)$ are models of P.

Proof: For T= eDTP or T= eNTP in ground(P) there are two kinds of rules:

a/ $(A \leftarrow A_1, ..., A_n; \underline{\beta}; \underline{I}); (A, \underline{\alpha}_A) \in lfp(T)$ and $(|A_i|, \underline{\alpha}_{Ai}) \in lfp(T)$, $1 \le i \le n$,

b/ $(A \leftarrow A_1, ..., A_n; \underline{\beta}; \underline{I}); \exists i : (|A_i|, \underline{\alpha}_{Ai}) \notin lfp(T)$

In the first case because of the construction of $\underline{\alpha}_A$ $\underline{I}_{FV}(\underline{\alpha}_{body}, \underline{\alpha}_A) \ge_{FV} \underline{\beta}$ holds, in the second case A_i is not among the facts, so $\underline{\alpha}_{Ai} = \underline{0}_{FV}$, therefore $\underline{\alpha}_{body} = \underline{0}_{FV}$ and $\underline{I}_{FV}(\underline{\alpha}_{body}, \underline{\alpha}_A) = \underline{1}_{FV} \ge_{FV} \underline{\beta}$ That is lfp(T) is a model.

According to the above statements the next theorem is true:

Theorem Both eDT_P and eNT_P have a fixed point, denoted by $lfp(eDT_P)$ and $lfp(eNT_P)$. If P is positive, then $lfp(eDT_P) = lfp(eNT_P)$ and this is the least fixed point. $lfp(eDT_P)$ and $lfp(eNT_P)$ are models of P; for negation-free fDATALOG this is the least model of the program.

A number of intuitionistic implications are established in [4] [6] and other papers, four of them are chosen for now, these are the extensions of the above operators. For these operators we have decided the suitable interval-values operators, and for both kinds of them we have deduced the uncertainty-level functions. Pressed for space the computations can not be shown, only the starting points and results are presented.

The connection between I_F and I_V and the extended versions of uncertainty-level functions are given below:

$$\underline{I}_V(\underline{\alpha},\underline{\gamma}) = (I_{V1}, I_{V2}); \qquad \underline{\alpha}'=(\alpha_1,1-\alpha_2); \underline{\gamma}'=(\gamma_1,1-\gamma_2)$$

$$I_{V1}=I_{F1}(\underline{\alpha}',\underline{\gamma}'); I_{V2}=1-I_{F2}(\underline{\alpha}',\underline{\gamma}')$$

$$f(\underline{I}_F, \underline{\alpha}, \underline{\beta}) =$$

$$(\min(\{\gamma_1 \mid I_{F1}(\underline{\alpha}, \underline{\gamma}) \ge \beta_1\}), \max(\{\gamma_2 \mid I_{F2}(\underline{\alpha}, \underline{\gamma}) \le \beta_2\}))$$

$$f(\underline{I}_V, \underline{\alpha}, \underline{\beta}) =$$

$$(\min(\{\gamma_1 \mid I_{V1}(\underline{\alpha}, \underline{\gamma}) \ge \beta_1\}), \max(\{\gamma_2 \mid I_{V2}(\underline{\alpha}, \underline{\gamma}) \ge \beta_2\}))$$

The studied operators and the related uncertainty-level functions are the following:

3.1 Extension of Kleene-Dienes implication

One possible extension of Kleene-Dienes implication for IFS is:

$$\underline{I}_{FK}(\underline{\alpha},\underline{\gamma}) = (\max(\alpha_2,\gamma_1), \min(\alpha_1,\gamma_2))$$

The appropriate computed elements are the following:

$$\underline{I}_{VK}(\underline{\alpha}, \underline{\gamma}) = (\max(1-\alpha_2, \gamma_1), \max(1-\alpha_1, \gamma_2))$$

$$f_1(\underline{I}_{FK}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 0 & \alpha_2 \geq \beta_1 \\ \beta_1 & \text{otherwise} \end{cases}$$

$$f_2(\underline{I}_{FK}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 1 & \alpha_1 \leq \beta_2 \\ \beta_2 & \text{otherwise} \end{cases}$$

$$f_1(\underline{I}_{VK}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 0 & 1-\alpha_2 \geq \beta_1 \\ \beta_1 & \text{otherwise} \end{cases}$$

$$f_2(\underline{I}_{VK}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 0 & 1-\alpha_1 \geq \beta_2 \\ \beta_2 & \text{otherwise} \end{cases}$$

3.2 Extension of Lukasiewicz implication

One possible extension of Lukasiewicz implication for IFS is:

$$\underline{I}_{FL}(\underline{\alpha},\underline{\gamma}) = (\min(1,\alpha_2+\gamma_1), \max(0,\alpha_1+\gamma_2-1))$$

The appropriate computed elements are the following:

$$\underline{I}_{VL}(\underline{\alpha}, \underline{\gamma}) = (\min(1, 1-\alpha_2+\gamma_1), \min(1, 1-\alpha_1+\gamma_2))$$

$$f_1(\underline{I}_{FK}, \underline{\alpha}, \underline{\beta}) = \min(1-\alpha_2, \max(0, \beta_1-\alpha_2))$$

$$f_2(\underline{I}_{FK}, \underline{\alpha}, \underline{\beta}) = \max(1-\alpha_1, \min(1, 1-\alpha_1+\beta_2))$$

$$f_1(\underline{I}_{VK}, \underline{\alpha}, \underline{\beta}) = \max(0, \alpha_2+\beta_1-1)$$

$$f_2(\underline{I}_{VK}, \underline{\alpha}, \underline{\beta}) = \max(0, \alpha_1+\beta_2-1)$$

3.3 Extensions of Gödel implication

There are several alternative extensions of Gödel implication, now we present two of them:

$$\underline{I}_{FG1}(\underline{\alpha},\underline{\gamma}) = \begin{cases} (1,0) & \alpha_1 \leq \gamma_1 \\ (\gamma_1,0) & \alpha_1 > \gamma_1, \alpha_2 \geq \gamma_2 \\ (\gamma_1,\gamma_2) & \alpha_1 > \gamma_1, \alpha_2 < \gamma_2 \end{cases}$$

$$\underline{I}_{FG2}(\underline{\alpha},\underline{\gamma}) = \begin{cases} (1,0) & \alpha_1 \leq \gamma_1, \alpha_2 \geq \gamma_2 \\ (\gamma_1,\gamma_2) & \text{otherwise} \end{cases}$$

The appropriate computed elements are:

$$\underline{I}_{VG1}(\underline{\alpha},\underline{\gamma}) = \begin{cases} (1,1) & \alpha_1 \leq \gamma_1 \\ (\gamma_1,1) & \alpha_1 > \gamma_1, \gamma_2 \geq \alpha_2 \\ (\gamma_1,\gamma_2) & \alpha_1 > \gamma_1, \gamma_2 < \alpha_2 \end{cases}$$

$$\underline{I}_{VG2}(\underline{\alpha},\underline{\gamma}) = \begin{cases} (1,1) & \alpha_1 \leq \gamma_1, \alpha_2 \leq \gamma_2 \\ (\gamma_1,\gamma_2) & \text{otherwise} \end{cases}$$

$$f_1(\underline{I}_{FG1}, \underline{\alpha}, \underline{\beta}) = \min(\alpha_1, \beta_1)$$

$$f_2(\underline{I}_{FG1}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 1 & \alpha_1 \leq \beta_2 \\ \max(\alpha_2,\beta_2) & \text{otherwise} \end{cases}$$

$$f_1(\underline{I}_{VG1}, \underline{\alpha}, \underline{\beta}) = \min(\alpha_1, \beta_1)$$

$$f_2(\underline{I}_{VG1}, \underline{\alpha}, \underline{\beta}) = \begin{cases} 0 & \alpha_1 \leq \beta_2 \\ \min(\alpha_2,\beta_2) & \text{otherwise} \end{cases}$$

$$f_1(\underline{I}_{FG2}, \underline{\alpha}, \underline{\beta}) = \min(\alpha_1, \beta_1)$$

$$f_2(\underline{I}_{FG2}, \underline{\alpha}, \underline{\beta}) = \max(\alpha_2, \beta_2)$$

$$f_1(\underline{I}_{VG2}, \underline{\alpha}, \underline{\beta}) = \min(\alpha_1, \beta_1)$$

$$f_2(\underline{I}_{VG2}, \underline{\alpha}, \underline{\beta}) = \min(\alpha_2, \beta_2)$$

An extremely important question is whether the resulting degrees satisfy the conditions referring to IFS and IVS respectively. Unfortunately for implications other than G2, the resulting degrees do not fulfill these conditions in all cases. (It is possible, that as a result of further research the current results can be used for eliminating the contradiction between the facts and rules of the program.) For now the next proposition can easily be proven:

Proposition 2. For $\underline{\alpha} = (\alpha_1, \alpha_2)$, $\underline{\beta} = (\beta_1, \beta_2)$

$$\text{if } \alpha_1+\alpha_2 \leq 1, \beta_1+\beta_2 \leq 1$$

$$\text{then } f_1(\underline{I}_{FG2}, \underline{\alpha}, \underline{\beta}) + f_2(\underline{I}_{FG2}, \underline{\alpha}, \underline{\beta}) \leq 1$$

$$\text{if } \alpha_1 \leq \alpha_2, \beta_1 \leq \beta_2$$

$$\text{then } f_1(\underline{I}_{VG2}, \underline{\alpha}, \underline{\beta}) \leq f_2(\underline{I}_{VG2}, \underline{\alpha}, \underline{\beta})$$

A further question is whether the fixed-point algorithm terminates or not, that is the consequence transformations reach the fixed point in finite steps or not. As P is finite, the fixed point contains only finite many elements. The only problem may occur with the uncertainty levels of recursive rules. But it can be seen that it does not occur if $f(\underline{I}_{FV}, \underline{\alpha}, \underline{\beta}) \leq_{FV} \underline{\alpha}$ for each $\underline{\alpha} \in L_{FV}$. G2 satisfies this condition so:

Proposition 3. In the case of G2 operator the fixed-point algorithm terminates.

4 Bipolar extension of fuzzy Datalog

The above mentioned problem of extended implications other than G2, and the results of certain psychological researches have led to the idea of bipolar fuzzy Datalog. The intuitive meaning of intuitionistic degrees is based on psychological observations, namely on the idea that concepts are more naturally approached by separately envisaging positive and negative instances [2] [6] [8]. Taking a further step, there are differences not only in the instances but also in the way of thinking as well. There is difference between positive and negative thinking, between deducing positive or negative uncertainty. The idea of bipolar Datalog is based on the previous observation: we use two kinds of ordinary fuzzy implications for positive and negative inference, namely we define a pair of consequence transformations instead of single one. Since in the original transformations lower bounds are used with degrees of uncertainty, therefore starting from IFS facts, the resulting degrees will be lower bounds of membership and non-membership respectively, instead of the upper bound for non-membership. However, if each non-membership value μ is transformed into membership value $\mu'=1-\mu$, then both members of head-level can be inferred similarly. Therefore, two kinds of bipolar evaluations have been defined.

Definition 6. The bipolar fDATALOG program (bfDATALOG) is a finite set of safe bfDATALOG rules $(r; (\beta_1, \beta_2); (I_1,I_2))$;

– In variant "a" the elements of bipolar consequence transformations $bDT_P = (DT_{P1}, DT_{P2})$ and $bNT_P = (NT_{P1}, NT_{P2})$ are the same as DT_P and NT_P in (1), (2), in variant "b" in DT_{P2} and NT_{P2} the level of rule's head is:

$$\alpha'_{A\,2}=\max(0, \min\{\gamma'_2 \mid I_2(\alpha'_{body\,2}, \gamma'_2) \geq \beta'_2\});$$

$$\alpha'_{body\,2}=\min(\alpha'_{A_1 2},\ldots,\alpha'_{A_n 2})$$

The uncertainty-level functions are:

$$f_a=(f_{a1},f_{a2});\ f_b=(f_{b1},f_{b2});$$

$$f_{a1} = f_{b1} = \min\{\gamma_1 \mid I_1(\alpha_1, \gamma_1) \geq \beta_1\})$$

$$f_{a2} = \min\{\gamma_2 \mid I_2(\alpha_2, \gamma_2) \geq \beta_2\});$$

$$f_{b2} =1-\min\{\gamma'_2 \mid I_2(\alpha'_2, \gamma'_2) \geq \beta'_2\})$$

It is evident, that applying the transformation $\mu'_1 = \mu_1$, $\mu'_2 = 1 - \mu_2$, for all IFS levels of the program, the B variant can be applied to IVS degrees as well.

Contrary the results of efDATALOG, the resulting degrees of most variant of bipolar fuzzy Datalog satisfy the conditions referring to IFS and IVS respectively.

Proposition 4. For $\underline{\alpha} = (\alpha_1, \alpha_2)$, $\underline{\beta} = (\beta_1, \beta_2)$ and for $(I_1,I_2) = (I_G,I_G)$; $(I_1,I_2) = (I_L,I_L)$; $(I_1,I_2) = (I_L,I_G)$; $(I_1,I_2) = (I_K,I_K)$; $(I_1,I_2) = (I_L,I_K)$

$$\text{if } \alpha_1+\alpha_2 \leq 1,\ \beta_1+\beta_2 \leq 1$$

$$\text{then} \quad f_{a1}(I_1, \alpha_1, \beta_1) + f_{a2}(I_2, \alpha_2, \beta_2) \leq 1$$

$$\text{and} \quad f_{b1}(I_1, \alpha_1, \beta_1) + f_{b2}(I_2, \alpha_2, \beta_2) \leq 1$$

further on

$$f_{a1}(I_G, \alpha_1, \beta_1) + f_{a2}(I_L, \alpha_2, \beta_2) \leq 1$$

$$f_{a1}(I_G, \alpha_1, \beta_1) + f_{a2}(I_K, \alpha_2, \beta_2) \leq 1$$

$$f_{a1}(I_K, \alpha_1, \beta_1) + f_{a2}(I_G, \alpha_2, \beta_2) \leq 1$$

$$f_{a1}(I_K, \alpha_1, \beta_1) + f_{a2}(I_L, \alpha_2, \beta_2) \leq 1$$

Although there are more results for variant "a", it seems the model realised by variant "b" is more useful. Because of the construction of bipolar consequence transformations the following proposition is evident.

Proposition 5. Both variations of bipolar consequence transformations have a least fixed point, which are models of P in the following sense:

a/ for each $(A \leftarrow A_1,\ldots,A_n; \underline{\beta}; \underline{I}) \in$ ground(P)

$$I(\alpha_{body1}, \alpha_{A1}) \geq \beta_1;\ I(\alpha_{body2}, \alpha_{A2}) \geq \beta_2.$$

b/ for each $(A \leftarrow A_1,\ldots,A_n; \underline{\beta}; \underline{I}) \in$ ground(P)

$$I(\alpha_{body1}, \alpha_{A1}) \geq \beta_1;\ I(\alpha'_{body2}, \alpha'_{A2}) \geq \beta'_2.$$

The termination of the consequence transformations based on these three implication operators was proven in the case of fDATALOG [5], and since this property does not change in bipolar case, the bipolar consequence transformations terminate as well.

Example 1 Consider the next program:

(p(a), (0.6,0.25)).

(r(a), (0.7, 0.3)).

$q(x) \leftarrow p(x),r(x)$; $\underline{I}$; (0.8, 0.15).

q(a); (0.7,0.1).

Let $\underline{I}=\underline{I}_{FK}(\underline{\alpha},\underline{\gamma}) = (\max(\alpha_2,\gamma_1), \min(\alpha_1,\gamma_2))$. Then $\underline{\alpha}_{body}= \min((0.6,0.25), (0.7, 0.3)) = (0.6, 0.3)$. As $0.3 < 0.8$ and $0.6 > 0.15$, so $f_1(I_{FK}, \underline{\alpha}, \underline{\beta}) = 0.8$, $f_2(I_{FK}, \underline{\alpha}, \underline{\beta}) = 0.15$, that is the level of rule's head is (0.8, 0.15). Allowing the other levels of q(a), its resulting levels are the union of them: max((0.8, 0.15), (0.7,0.1)) = (0.8, 0.1). So the fixed point of the program is: {(p(a), (0.6, 0.25)), (r(a), (0.7, 0.3)), (q(a), (0.8, 0.1))}

Now the program let be evaluated in bipolar manner and let $\underline{I}=(I_L,I_G)$, then in variant "b"

$\alpha_{body1} = \min(0.6, 0.7) = 0.6$,

$\alpha'_{body2} = \min(1-0.3, 1-0.25) = 0.7$

$f_{b1}(I_L, \alpha_1, \beta_1) = \max(0, \alpha_1+\beta_1-1)=0.8+0.6-1 = 0.4$, $f_{b2}(I_G, \alpha'_2, \beta'_2)=1-\min(\alpha'_2,1-\beta_2)=1-\min(0.7,0.85) = 0.3$, allowing for the other levels of q(a), its resulting levels are (max(0.4, 0.7), 1-max(1-0.3, 1-0.1)) = (0.7, 0.1), so the fixed point is: {(p(a), (0.6, 0.25)), (r(a), (0.7, 0.3)), (q(a), (0.7, 0.1))}.

Example 2. Consider the next – recursive – program:

(p(a, b), (0.6, 0.2)).

(p(a, c), (0.7, 0.3)).

(p(b, d), (0.5, 0.3)).

(p(d, e), (0.8, 0.1)).

$q(x, y) \leftarrow p(x,y)$; $\underline{I}_1$; (0.75, 0.2).

$q(x, y) \leftarrow p(x,z), q(z, y)$; $\underline{I}_2$; (0.7,0.2).

Let $\underline{I}_1=\underline{I}_2= \underline{I}_{FG2}(\underline{\alpha}, \underline{\gamma}) = \begin{cases}(1,0) & \alpha_1 \le \gamma_1, \alpha_2 \ge \gamma_2 \\ (\gamma_1, \gamma_2) & \text{otherwise}\end{cases}$

Then $f_1(\underline{I}_{FG2},\underline{\alpha},\underline{\beta})=\min(\alpha_1,\beta_1)$, $f_2(I_{FG2},\underline{\alpha},\underline{\beta})=\max(\alpha_2,\beta_2)$. So the fixed point is: {(p(a,b), (0.6, 0.2)), (p(a,c), (0.7, 0.3)), (p(b,d), (0.5, 0.3)), (p(d,e), (0.8, 0.1)), (q(a,b), (0.6, 0.2)), (q(a,c), (0.7, 0.3)), (q(b,d), (0.5, 0.3)), (q(d,e), (0.75, 0.2), (q(a,d), (0.5,0.3)), (q(b,e), (0.5, 0.3)), (q(a,e), (0.5, 0.3))}

Now let the program be evaluated in bipolar manner according to variant "b" and let $\underline{I}_1= (I_L,I_K)$, $\underline{I}_2=(I_G,I_G)$ that is

$f(I_G, \alpha, \beta) = \min(\alpha, \beta)$, $f(I_L, \alpha, \beta) = \max(0, \alpha + \beta -1)$,

$f(I_K, \alpha, \beta) = \begin{cases}0 & \alpha+\beta \le 1 \\ \beta & \alpha+\beta > 1\end{cases}$

Then the first coordinates of computed uncertainties are:

(q(a,b), (max(0, 0.6+0.75-1)=0.35, _)), (q(a,c), (0.45, _)), (q(b,d), (0.25, _)), (q(d,e), (0.55, _); q(a,d): as min(0.6,0.25)+0.7≤1 so (q(a,d), (0,_)), (q(b,e), (0.7, _)), (q(a,e), (0.7, _)).

The second coordinates are computed after the appropriate transformations $\alpha'_2=1- \alpha_2$. So

(q(a,b), (_,1-min(1-0.2.1-0.2)=0.2)), (q(a,c), (_,0.3)), (q(b,d), (_,0.3)), (q(d,e), (_,0.2), (q(a,d), (_,0.3)), (q(b,e), (_,0.3)), (q(a,e), (_,0.3)).

So the fixed point is: {(p(a,b), (0.6, 0.2)), (p(a,c), (0.7, 0.3)), (p(b,d), (0.5, 0.3)), (p(d,e), (0.8, 0.1)), (q(a,b), (0.35, 0.2)), (q(a,c), (0.45, 0.3)), (q(b,d), (0.25, 0.3)), (q(d,e), (0.55, 0.2); (q(a,d), (0, 0.3)), (q(b,e), (0.7, 0.3)), (q(a,e), (0.7, 0.3)).

5 Conclusions

In this paper we have presented several possible extensions of fuzzy Datalog to intuitionistic- and interval-valued fuzzy logic, and the concept of bipolar fuzzy Datalog has been introduced. Our propositions were proven for negation-free fDATALOG. The extension of stratified fDATALOG and the examination of other implication operators will be the subject of further research as well.

References

[1] Ágnes Achs (2006) Creating and evaluating of fuzzy knowledgebase, Journal of Universal Computer Science, 12:9, http://www.jucs.org/jucs_12_9

[2] D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk, H. Prade (2005) Terminological difficulties in fuzzy set theory – The case of "Intuitionistic Fuzzy Sets", Fuzzy Sets and Systems 156: 485 491

[3] Krassimir Atanassov (2005) Answer to D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk and H. Prade's paper "Terminological difficulties in fuzzy set theory—the case of "Intuitionistic Fuzzy Sets", Fuzzy Sets and Systems 156: 496 – 499

[4] C. Cornelis, G. Deschrijver, E.E. Kerre (2004) Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application, International Journal of Approximate Reasoning 35: 55-95

[5] Achs Ágnes-Kiss Attila: A Datalog fuzzy kiterjesztése, Alkalmazott Matematika Lapok, 1998., Budapest (111-138.)

[6] D. Dubois, P. Hajek, H. Prade (2000) Knowledge-Driven versus Data-Driven Logics, Journal of Logic, Language, and Information 9: 65-89

[7] K. Atanassov, G. Gargov (1998) Elements of intuitionistic fuzzy logic. Part I, Fuzzy Sets and Systems 95: 39-52

[8] J.T. Cacioppo, W.L. Gardner, G.G. Berntson (1997) Beyond bipolar conceptualization and measures: the case of attitudes and evaluative spaces, Personality and Social Psychol. Rev. 1: 3 25.

[9] Ágnes Achs, Attila Kiss (1995) Fixed point query in fuzzy Datalog, Annales Univ. Sci. Budapest, Sect. Comp. 15: 223-231

[10] Ágnes Achs, Attila Kiss (1995) Fuzzy extension of Datalog, Acta Cybernetica Szeged 12: 153 166

[11] Didier Dubois, Henri Prade (1991) Fuzzy sets in approximate reasoning, Part 1: Inference with possibility distributions, Fuzzy Sets and Systems 40: 143-202

[12] S. Ceri, G. Gottlob, L.Tanca (1990) Logic Programming and Databases, Springer-Verlag Berlin

Session 7

Neuro-fuzzy, Predictions and Evolutionary Approach

Training Neurofuzzy Networks with Participatory Learning

Michel Hell
UNICAMP
FEEC - DCA
Campinas, SP
Brazil, 13083-970
mbhell@dca.fee.unicamp.br

Rosângela Ballini
UNICAMP
IE - DTE
Campinas, SP
Brazil, 13083-970
ballini@eco.unicamp.br

Pyramo Costa Jr.
PUC-MG
PPGEE
Belo Horizonte, MG
Brazil, 30535-610
pyramo@pucminas.br

Fernando Gomide
UNICAMP
FEEC - DCA
Campinas, SP
Brazil, 13083-970
gomide@dca.fee.unicamp.br

Abstract

This paper introduces a new approach to adjust a class of neurofuzzy networks based on the idea of participatory learning. Participatory learning is a mean to learn and revise beliefs based on what is already known or believed. The performance of the approach is verified with the Box and Jenkins gas furnace modeling problem, and with a short-term load forecasting problem using actual data. Comparisons with alternative training procedures suggested in the literature are included to shown the effectiveness of participatory learning to train neurofuzzy networks.

Keywords: Participatory Learning, Fuzzy Systems, Neurofuzzy Networks.

1 Introduction

Neurofuzzy systems have been widely used in different areas of science and engineering, including medicine, economics, fuzzy mathematics, game theory, systems modeling to mention a few.

As part of computational intelligence, neurofuzzy systems combine two major paradigms, neural networks and fuzzy systems. The combination results in systems that integrate the representation richness of fuzzy systems with the learning ability of neural networks. The central idea behind neurofuzzy systems is the capability to tune fuzzy systems using learning procedures and data. Suitable methodologies and learning algorithms are among the most important issues when designing neurofuzzy systems [1].

In the past years several methods to train neurofuzzy systems, especially neurofuzzy networks, have been proposed. In general these methods are based on gradient descent and/or reinforcement approaches [1, 2, 3, 4].

These techniques produce satisfactory results in many cases but currently there is no consensus on which one is the best since they depend heavily on the data set characteristics [5].

In 1990 a new learning approach, namely Participatory Learning (PL), was introduced by Yager as a model of learning that captures many of the salient features of human learning [6]. This approach assumes that learning and beliefs about an environment depend on what the system already knows about the environment [7].

Recently, the participatory learning paradigm was used to develop an efficient unsupervised fuzzy cluster algorithm [5] and to find rule base structures in adaptive, evolving fuzzy modeling procedures [8].

Alternative methods, such as Bayesian analysis [9], have been developed to combine current belief with new knowledge about the environment coming from observations. In this work we adopt an approach based on the assumption that probability density functions are not easily available or computable, what precludes the use of Bayesian approaches.

More precisely, this paper suggests the use of PL idea to develop a new training procedure for hybrid neurofuzzy networks. The procedure does not dependent on the data set because the current knowledge about the environment is part of the learning process itself and influences the way in which new observations are used for learning [5].

The paper is organized as follows. Next section reviews the main ideas and concepts of participatory learning. Section three presents the neurofuzzy network structure. Session 4 details the training procedure suggested in the paper. Session 5 summarizes simulation results and comparison studies. Section 6 concludes the paper summarizing its contributions and issues for further investigation.

2 Participatory Learning

The main characteristic of participatory learning is that an exogenous observation impact in causing learning or belief revision depends on its compatibility with the current system belief. In particular, observations conflicting with the current belief are discounted [6].

Let $v \in [0,1]^n$ be a variable that represents the belief of a system. The aim of the participatory learning is to learn the value of this variable based on a sequence of observations $x^k \in [0,1]^n$ that encodes the knowledge about the value of the variable v. In this sense x^k is a manifestation of value of v in the $k-th$ observation. Thus we use the vector x^k as a means to learn valuations of v. The learning process is participatory if the usefulness of each observation x^k in contributing to the learning process depends upon its acceptance by the current estimate of the values of v as being valid observation [6]. Participatory learning means that, to be relevant for the learning process, x^k must be close to v^k. A mechanism for updating the current beliefs of v is a smoothing like algorithm:

$$v^{k+1} = v^k + \alpha\rho_k(x^k - v^k) \tag{1}$$

where $k = 1, \ldots, P$, and P is the number of observations; v^{k+1} is the new system belief; v^k is the current belief; x^k is the current observation; $\alpha \in [0,1]$ is the learning rate; and $\rho_k \in [0,1]$ is the compatibility degree between x^k and v^k. As an example, ρ_k can be computed as $\rho_k = 1 - \frac{1}{n}\sum_{i=1}^{n} d_i^k$ where $d_i^k = \left|x_i^k - v_i^k\right|$.

Notice that, if $\rho_k = 0$ (observation too far from current belief) we get from (1) that $v^{k+1} = v^k$ and the system is maximally closed to learn. On the oder hand, if $\rho_k = 1$ we see from (1) that $v^{k+1} = v^k + \alpha(x^k - v^k)$, therefore the system is maximally open for learning. However, in other for ρ_k to equal 1 it must be the case that $x_i^k - v_i^k = 0$ for all i. This condition implies that $v^k = x^k$ and hence we get $v^{k+1} = v^k$, which implies that no learning occurs [10].

One concern that can be raised about the described learning procedure is that it ignores the situation when a stream of low ρ_k's arises during a long period of time. In this case, the system should become more open to learning because it may be the case that the current belief is wrong, not the new observations. In [6] Yager proposes a mechanism that monitors the compatibility of the current beliefs with the observations. This information is translated into an arousal index used to influence the learning process, as shown in Figure 1. The higher the arousal rate, the less confident is the system with the current belief, and conflicting observations become important to update the beliefs. Le us denote $a_k \in [0,1]$ as the arousal index. The higher a_k the more aroused the system will be. The dynamics of

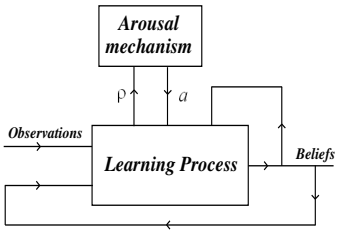


Figure 1: Participatory Learning with Arousal

the arousal index adaptation is specified as follows:

$$a_{k+1} = a_k + \lambda((1 - \rho_{k+1}) - a_k) \tag{2}$$

where $\lambda \in [0,1]$ is a constant that controls the rate of change arousal; the closer λ is to one, the faster the system is to sense compatibility variations. The arousal index can be seen as the complement of the confidence in the belief structure currently held.

Expression (1) can be rewritten to incorporate the arousal mechanism as follows:

$$v^{k+1} = v^k + \alpha\rho_k^{1-a_k}(x^k - v^k) \tag{3}$$

As it can be noted in (3), while ρ_k measure how much the system changes its credibility in its own beliefs, the arousal index a_k acts as a critic to remind when the current belief should be modified in front of new evidences. Moreover, apart from the term $\rho_k^{1-a_k}$ Equation (3) can be viewed as a standard learning rule used to training neural networks. However, in the presence of spurious samples in training data set, a learning rule without the term $\rho_k^{1-a_k}$ may not converge or converge to an incorrect value.

3 Neurofuzzy Network Model

This section introduces the complete structure of the hybrid neurofuzzy network (RNF), which is composed by two parts. The first part has an input and two hidden layers and assembles a fuzzy inference system. The second is a classic neural network whose purpose is to aggregate the outputs of the fuzzy inference system. Figure 2 shows a neurofuzzy network with n inputs and m outputs. The input layer supplies the inputs to the neurons in the next layer. The first hidden layer consists of neurons whose activation functions are membership functions of fuzzy sets that form the input space partition.

For each dimension x_i^k of a n-dimensional input vector x^k there are N_i fuzzy sets $A_i^{\lambda_i}, \lambda_i = 1, \ldots, N_i$ whose membership functions are activation functions of corresponding input layer neurons. The variable k is the

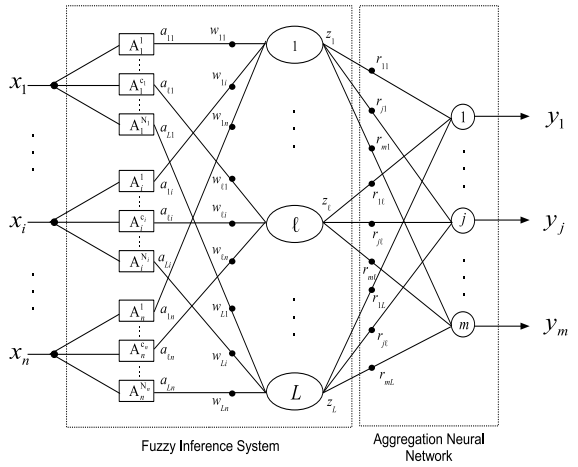


Figure 2: Hybrid Neurofuzzy Network Structure

discrete-time, that is, $k = 1, 2, \ldots$, but it will be omitted in the rest of the paper to simplify notation. Thus, the outputs of the first hidden layer are the membership degrees associated with the input values, that is, $a_{\ell i} = \mu_{A_i^{\lambda_i}}(x_i),\ i = 1,\ \ldots,\ n$ and $\ell = 1,\ \ldots,\ L$; where L is the number of neurons in the second hidden layer.

The neurons of the second hidden layer are fuzzy set-based neurons called *or* neurons (Figure 3) whose inputs $a_{\ell i}$ are weighted by $w_{\ell i}$ where $w_{\ell i}$ corresponding to the credibility of the input $a_{\ell i}$. The neurons use s-norms to process the inputs. For different instances of $s-$norms see [2] and references therein.

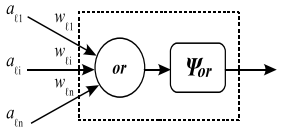


Figure 3: Fuzzy Set *or* Neurons

Fuzzy set-based neurons assume membership degrees $a_{\ell i}$ inputs and weights $w_{\ell i}$ in $[0, 1]$. Therefore they perform nonlinear mappings $[0, 1]^n \to [0, 1]$. The activation function ψ_{or} of a *or* neuron is, in general, a nonlinear function, but here we assume ψ_{or} to be the identity function $\psi_{or}(u) = u$.

It is easy to see that the network structure embeds a set of *if-then* rules $R = \{R_k, k = 1,\ \ldots,\ L\}$ of the form:

$$\begin{aligned} R_k \quad : \quad & \text{If } (x_1 \text{ is } A_1^{\lambda_1} \text{ with credibility } w_{\ell 1}) \ldots \\ & \textit{or } (x_i \text{ is } A_i^{\lambda_i} \text{ with credibility } w_{\ell i}) \ldots \\ & \textit{or } (x_n \text{ is } A_i^{\lambda_n} \text{ with credibility } w_{\ell n}) \\ & \text{Then, } z \text{ is } z_\ell. \end{aligned}$$

where

$$z_\ell = \mathop{\mathbf{S}}_{i=1}^{n} (w_{\ell i}\ \mathbf{t}\ a_{\ell i})\ .$$

Therefore, there is a close correspondence between the structure of the first part of the neurofuzzy network and a set of fuzzy rules or, alternatively, a fuzzy rule base. In addition, the processing scheme induced by the network structure agrees with the principles of fuzzy set theory and approximate reasoning [2].

As mentioned above, the second part is a classical neural network composed by conventional neurons whose output y_j is an aggregation of the inputs z_ℓ and the weights $r_{j\ell}, \ell = 1,\ \ldots,\ L$, $j = 1,\ \ldots,\ m$.

The neurofuzzy network processing can be summarized as follows:

1. N_i is the number of fuzzy sets that composes the partition of $i-$th input;
2. $a_{\ell i} = \mu_{A_i^{\lambda_i}}(x_i)$ is the membership degree of x_i in the fuzzy set $A_i^{\lambda_i}$, where $a_{\ell i}$ is the $i-$th input of neuron ℓ of the second hidden layer;
3. z_ℓ is the output of the $\ell-$th neuron of the second hidden layer:

$$z_\ell = \mathop{\mathbf{S}}_{i=1}^{n} (w_{\ell i}\ \mathbf{t}\ a_{\ell i}) \tag{4}$$

4. y_j is the output of $j-th$ nonlinear neuron of the output layer given by:

$$y_j = f(u_j) = f\left(\sum_{k=1}^{L} (v_{jk} z_k)\right) \tag{5}$$

   where, $f : \Re^L \to [0, 1]$ is a nonlinear, monotonically increasing function. In this paper, we use the logistic function $f(u_j) = 1/(1 + \exp(-u_j))$;
5. $w_{\ell i}$ is the weight between the $\ell-$th *or*-neuron and the $i-$th neuron of the first hidden layer;
6. $r_{j\ell}$ is the weight between the output y_j of the network and the $\ell-$th neuron of the second hidden layer;

The neurofuzzy network architecture is very flexible because it allows appropriate triangular norms to be chosen and offers the possibility to extract linguistic fuzzy rules directly from its topology.

4 Training Procedure

The training procedure suggested in this paper involves three main phases. The first phase uses the fuzzy c-means algorithm to granulate the input space. The next phase use the gradient descent to update the

weights related to the classic neural network ($r_{j\ell}$'s). The last phase use the participatory learning paradigm summarized in section 2 to update the weights connected to the *or* neurons ($w_{\ell i}$'s). The essential implementation steps are detailed next.

4.1 Generation of Membership Functions

To generate the membership functions of the first hidden layer of the network shown in Figure 2 the c-means clustering algorithm is used. Information about input and output spaces are included in the cluster process. The modal values of the membership functions are the projection of the clusters centers in its respective universes. We assume Gaussian-shaped function as membership functions in the first hidden layer neurons. The dispersions of the Gaussians are adjusted to form a cognition frame of the input space [2].

4.2 Weight Updating

The first step in the weight updating process is to estimate the network output $\hat{y} \in [0,1]^m$ for a given input $x \in [0,1]^n$. This corresponds to the fuzzification of the input pattern, and successive computation the outputs of the remaining neurons layers using Equations (4) and (5). Next, the learning process aims to minimizing an error measure between the actual network output and a desired output for each input pattern, that is, to minimize

$$e = \frac{1}{2}\sum_{j=0}^{m}(y_j - \hat{y}_j)^2 \tag{6}$$

where $\hat{y}_j$ is the value of the output unit j and y_j is the desired value of the output unit j.

We update the output layer weights using a gradient descent method, that is

$$\Delta r_{j\ell} = \eta(y_j - \hat{y}_j)f'(u_j)z_\ell \tag{7}$$

where $f'(u_j) = f(u_j)(1-f(u_j))$ is the derivative of the activation function evaluated at u, $u_j = \sum_{\ell=1}^{L}(r_{j\ell}z_\ell)$, and η is the learning rate.

The next step update the weights of the logical neurons using the participatory learning idea, as follows.

Participatory Learning Procedure

As discussed in the Section 3, the fuzzy inference system encodes in the fuzzy set-based neurons structures can be seen as a fuzzy relation $G : A \times Z \rightarrow [0,1]$ where A and Z are two finite universes and G is any subset of the Cartesian product of these two universes. In particular, for $a^k \in A$ and $G \subset A \times Z$ we have that it is possible to compute an element $z^k \in Z$ related with a^k of the form $z^k = a^k \circ G$.

If we choose as $s-$norm the maximum to the *or* neurons we have that the operator '$\circ$' becomes $max_\ t$ ($sup_\ t$) and for a given pair (a^k, z^k) the relation G can be found by solving the fuzzy-relational estimation problem [2] whose solution is:

$$G = a^{k\ T}\ \varphi\ z^k \tag{8}$$

where ^{T} denote the transpose and φ is such that

$$(a^k_{\ell i}\ \varphi\ z^k_\ell) = sup(c \in [0,1] \mid a^k_{\ell i}\ t\ c \leq z^k_\ell) \tag{9}$$

We can see that the relational matrix $G = [g_{\ell i}]$ have strong relation with the weight matrix $W = [w_{\ell i}]$. In fact, for a given pattern $[x_k, y_k]$ and a matrix $R = [r_{j\ell}]$ of output layer weight's the better approximation of the mapping $f :\ x_k \rightarrow y_k$ performed by the network shown of Figure 2 occurs when $W \equiv G$.

To find the relation G we need the output of logical layer neurons z_k values. This can be estimated from y_k and R solving the constrained linear least-squares problem:

$$\min_{z^k}\left\{\frac{1}{2}\left|Rz^k - f^{-1}(y^k)\right|^2\right\} \text{ subject to } 0 \leq z^k \leq 1$$
where $f^{-1}(\cdot)$ is the inverse of function f in (5). (10)

Next, we compute the fuzzy relation G using (9).

Once G is found we can update weights $w_{\ell i}$'s based on participatory learning paradigm as follows.

$$\Delta W^k = \alpha\left(\rho^k\right)^{1-a_k}(G^k - W^k) \tag{11}$$

where α is the learning rate; $k = 1,\ \ldots\ P$ and P is the number of training patterns; a_k is calculated according (2), and

$$\rho^k = 1 - \frac{1}{Ln}\sum_{\ell=1}^{L}\sum_{i=1}^{n}|g^k_{\ell i} - w^k_{\ell i}|. \tag{12}$$

The procedure to train the neurofuzzy networks using PL approach (NFPL) proposed in this paper can be summarized in the pseudo code presented in Figure 4.

5 Experimental Results

To evaluate the performance of the learning algorithm introduced in this paper, we compare our approach with alternative models suggested in the literature. The examples given here include nonlinear system identification and time series prediction problems. We adopt the maximum $s-$norm and the algebraic product as the $t-$norm in the *or* neurons.

Box and Jenkins Gas Furnace

The Box and Jenkins system identification is a well-known benchmark problem. Identification uses of 290 pairs of input/output data taken from a laboratory furnace [11]. Each data sample consists of the methane flow rate, the process input variable, x^k, and the CO_2 concentration in off gas, the process output, y^k. This is a dynamical process with one input x^k and one output y^k. The aim is to predict current output $y(k)$ using

```
Begin NFPL
  Read Training Data;
  Granularize the input space (section 4.1);
  p:= n° of training patterns;
  WHILE err ≥ error tolerance DO
      t:=1;
      WHILE t ≤ p DO
          Read (x^k, y^k);
          Estimate the network output y^k (section 3);
          Compute the approximation error e using (7);
          Compute the variation of the weights r_jl's using (8);
          Update the r_jl's weights;
          t:=t+1;
      END WHILE
      t:=1;
      WHILE t ≤ p DO
          Read (x^k, y^k);
          Compute the z^k from y^k and R using (11);
          Compute the fuzzy relation G using (9) ;
          Determine the compatibility degree ρ_k using (13);
          Determine the arousal index using (3);
          Compute the variation of the weights w_ri using (12);
          Update the w_ri's weights;
          t:=t+1;
      END WHILE
      Calculate the general approximation error err
  END WHILE

END NFPL
```

Figure 4: NPFL Approach

past input and output values, with the lowest error. Different studies indicate that the best model structure for this system is $y^k = f\left(y^{k-1}, x^{k-4}\right)$.

Figure 5 depicts the model output obtained by the NFPL approach. Comparison results with previous fuzzy/neurofuzzy approaches that works with the same inputs is given in Table 1. We notice that NFPL perform as well as the one of the best model.

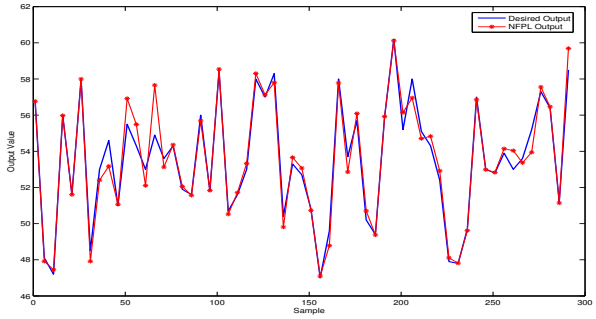


Figure 5: Box and Jenkins Gas Furnace Identification

Load Forecasting

An application example that shows the effectiveness of the participatory learning in neuroffuzy training is the short-term load forecasting. The task load forecasting is to accurately predict the 24 hourly loads of the next operation day.

We use a training data set composed by hourly loads from May of 2000 to February of 2001 of an electrical utility located at the Southeast region of Brazil. The input variables of the model are L_{h-1} and L_{h-2}, which represent the load at times $h-1$ and $h-2$, respectively.

Table 1: Gas Furnace Identification

Model	N° Rules	Mean Square Error
Pedrycz [12]	81	0.5656
Xu and Lu [13]	25	0.5727
Delgado [14]	4	0.4100
Yoshinari [15]	6	0.5460
Rutkowski [16]	9	0.4912
NFPL	**9**	**0.4528**

To show the efficiency and the data-independent nature of the PL approach during training, two cases are considered. The first case (Case 1) data was collected without any measurement error. In the second case (Case 2) we introduce uniform random noise in selected samples of the training data. Four different neural network models were trained using data of the two cases, namely, a classic multilayer perceptron with 2 hidden layers with 12 and 8 neurons, respectively, trained with backpropagation algorithm (MLP), an adaptive neurofuzzy system [17] using the same number of fuzzy sets to granulate each input variable (ANFIS), and the neurofuzzy network presented in Figure 2 with 64 neurons in second hidden layer. The RNF was trained in two different ways. The first way uses participatory learning as detailed in Section 4 (NFPL). The second one adopts a non-participatory version (NF-NPL) of the learning procedure proposed in section 4, that is, we exclude the participatory with arousal term $\rho_k^{1-a_k}$ from Equation (11). The Mean Absolute Error (MAE) is used to evaluate the results.

$$MAE = \frac{1}{P}\sum_{k=1}^{P}\left|\hat{y}^k - y^k\right|\frac{100}{y^k} \quad (\%) \qquad (13)$$

where $\hat{y}^k$ is the k-th predicted value, y^k is k-th real value and P is the number of values to predict. The results obtained are summarized in Table 2. Figure 6 shows the load forecast results provided by the neural models for September 12, 2001 in both, Case 1 (a) and Case 2 (b).

Table 2: Load Forecast

Model	MAE (%) - Case 1	MAE (%) - Case 2
MLP	7.20	12.95
ANFIS	5.74	11.47
NF-NPL	6.86	12.07
NFPL	**3.98**	**4.42**

As it can be seen in Figure 6 and Table 2, all neural models show comparable performance in Case 1. However, in Case 2 the participatory model improves the forecast substantially and outperforms the remaining neural models. The participatory nature of NFPL is

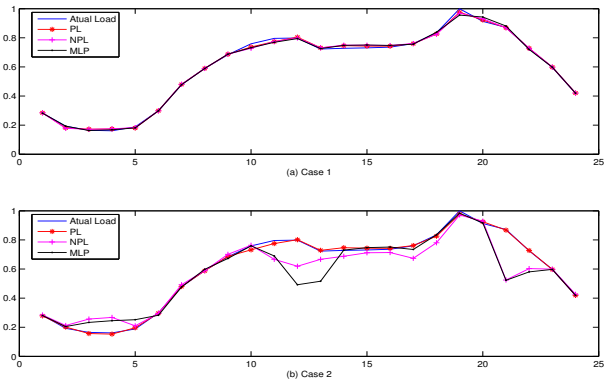


Figure 6: Short Term Load Forecasting

able to smooth the effect of noise using the compatibility measure.

6 Conclusions

In this paper we have introduced a new training procedure for a class of hybrid neurofuzzy network based on participatory learning paradigm. Experimental results show that participatory learning provide an effective alternative for supervised training of neurofuzzy networks that do not dependent of training data set and is as efficient as alternative fuzzy/neurofuzzy approaches addressed in the literature.

Acknowledgement

The authors acknowledge FAPESP, the Research Foundation of the State of São Paulo, for fellowship 03/05042-1 and grant 03/10019-9, and CNPq, the Brazilian National Research Council, for grant 304857/2006-8, and the anonymous reviewers for their constructive comments and suggestions that help to improve this paper.

References

[1] M. Figueiredo, R. Ballini, S. Soares, M. Andrade, F. Gomide, Learning Algorithms for a class of Neurofuzzy Network and Application, *IEEE Trans. on System, Man and Cybernetics - Part C*, vol. 34, no. 3, pp. 293–301, 2004.

[2] W. Pedrycz, F. Gomide, Fuzzy Systems Engineering: Toward Human-Centric Computing, *Wiley Interscience*, Hoboken, NJ, USA, 2007.

[3] L. X. Wang, Adaptive Fuzzy Systems and Control, *Prentice-Hall*, NJ, USA, 1994.

[4] C. T. Lin, C. S. Lee, Neural Fuzzy Systems, *Prentice Hall*, NJ, USA, 1996.

[5] L. Silva, F. Gomide, R. R. Yager, Participatory Learning in Fuzzy Clustering, *Proc. 14th IEEE International Conference on Fuzzy Systems*, Reno, USA, May 2005, pp. 857–861.

[6] R. R. Yager, A model of Participatory Learning, *IEEE Trans. on System, Man and Cybernetics*, vol. 20, no. 5, pp. 1229–1234, 1990.

[7] R. R. Yager, Participatory Learning: A Paradigm for More Human Like Learning, *Proc. 2004 IEEE International Conference on Fuzzy Systems*, vol. 1, July 2004, pp. 79–84.

[8] E. Lima, F. Gomide, R. Ballini, Participatory Evolving Fuzzy Modeling, *Proc. 2006 International Symposium on Evolving Fuzzy Systems*, Lake district, UK, Sept. 2006, pp. 36–41.

[9] W. A. Wright, Bayesian approach to neural-network modeling with input uncertainty, *IEEE Trans. on Neural Networks*, vol. 10, no. 6, pp. 1261–1270 , 1999.

[10] R. R. Yager, D. P. Filev, A fuzzy logic controller view of participatory learning , *Proc. Second IEEE International Conference on Fuzzy Systems*, vol. 2, April 1993, pp. 912–917.

[11] G. E. P. Box, G. M. Jenkins, Time Series Analysis- Forecasting and Control, 2.ed, *Holden Day*, CA, USA, 1976

[12] W. Pedrycz, A Identification Algorithm in Fuzzy Relational Systems, *Fuzzy Sets and Systems*, vol. 13, pp. 153–167, 1984.

[13] W. Xu, Y. Z. Lu, Fuzzy Model Identification and Self-learning for Dynamic Systems, *IEEE Trans. on System, Man and Cybernetics*, vol. SMC-17, pp. 683–689, July 1987.

[14] M. Delgado, A. F. Gmez-Skarmeta, F. Martin, Fuzzy Clustering based Rapid Prototyping for Fuzzy Rule-based Modeling, *IEEE Trans. on Fuzzy Systems*, vol. 5, pp. 223–233, May 1997.

[15] Y. Yoshinari, W. Pedrycz, K. Hirota, Construction of Fuzzy Models Through Clustering Techniques, *Fuzzy Sets and Systems*, vol. 54, pp. 157–165, 1993.

[16] L. Rutkowski, and K, Cpalka, Designing and Learning of Adjustable Quasi-Triangular Norms With Applications to Neuro-Fuzzy Systems, *IEEE Trans. on Fuzzy Systems*, vol. 13, no. 1, pp. 140–151, Feb. 2005.

[17] R. Jang, ANFIS: Adaptive network based fuzzy inference system, *IEEE Trans. on Systems Man and Cybernetics*,vol. 23, no. 3, pp. 665–685, 1993.

A New Preprocessing Approach to Preparation of Binary Patterns for FAM Neural Networks

M. Chitsaz
Azad. Islamic University, Aliabad Katul Branch,Iran
Saba.sanat@gmail.com

N. Sadati
AmirKabir University of Technology, Tehran, Iran
sadati@sina.sharif.edu

R. Barzamini
AmirKabir University of Technology, Tehran, Iran
Barzamini@aut.ac.ir

J. Jouzdani
AmirKabir University of Technology, Tehran, Iran
Javid@aut.ac.ir

A. Khosravi
Universiy of Girona, Girona, Spain
khosravi@eia.udg.es

Abstract

The patterns which are presented to a Fuzzy ARTmap network should be preprocessed in such a way that the data are of appropriate clearance. In order to decrease the degree of similarity between the normalized binary patterns, a new method is proposed in which the binary patterns are converted to fuzzy (analogue) patterns. This is done by weighting and normalizing the values, determining the binary pattern, according to the number of ones which surround that value. By using this method, the noise in the input patterns is rejected and therefore noisy patterns are not considered as new training patterns. This increases the efficiency of the proposed method.

Keywords: Paper, Instructions.

1 Introdction

Usually the binary patterns are presented to a neural network in the form of an $(m \times n)$ matrix and are converted to $(mn \times 1)$ normalized vectors. In normalizing a binary pattern, the number of *ones* plays a key role. When two patterns have the same number of ones, the normalized values have equal weights and therefore, during the training process they may be placed in one group. Furthermore, when the number of ones in an input pattern is equal to the number of ones in two (or more) training patterns, the prediction of the network is the pattern which is presented to the network before the other pattern(s). Hence, the prediction of the network depends on the order of training patterns when they are presented to the network. Such property is not suitable for prediction and may lead to an incorrect prediction. In this paper a method is presented in which the binary values are converted to fuzzy values to generate a three dimensional pattern and finally the patterns are normalized. In what follows, first the structure of the neural network is explained and then the proposed method is presented. In the end the simulation results and the conclusions are provided.

2 Fuzzy ARTmap Technique

Fuzzy ARTmap is an incremental supervised learning algorithm which combines fuzzy logic and Adaptive Resonance Theory (ART) neural network for recognition of pattern categories and multidimensional maps in response to input vectors presented in an arbitrary order. It realizes a new minmax learning rule which conjointly minimizes predictive error and maximizes code compression, and therefore gives generalization. This is achieved by a match tracking process that increases the ART vigilance parameter (fuzzy degree of membership of the input with respect to the category templates) by the minimum amount needed to correct a predictive error (PE). The Fuzzy ARTmap neural network is composed of two Fuzzy ARTmap modules [3], i.e., $fuzzy\ ART_a$ and $fuzzy\ ART_b$ [2]. Which are typified in Fig. 1 and are essentially the same as those described by Carpenter et al in [1]. The interactions mediated by the map field F^{ab} may be operationally characterized as follows.

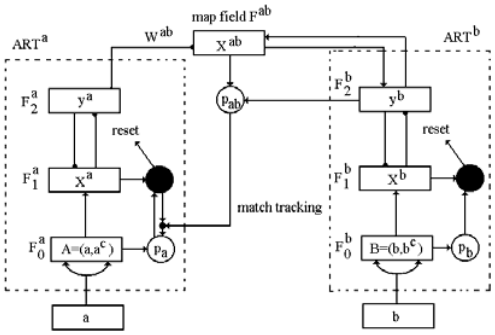


Figure 1: Typical Fuzzy ARTmap

Inputs to ART_a and ART_b are in the complement code form: for ART_a. $I = A = (a, a^c)$; for ART_b. $I = B = (b, b^c)$ (see Fig. 1). Variables in ART_a or ART_b are designated by subscripts "a" and "b" respectively. For ART_a, let $x^a = \{x_1^a ... x_{2Ma}^a\}$ denote the F_1^a output vector, let $y^a = \{y_1^a ... y_{Na}^a\}$ denote F_2^a, and let $w_j^a = \{w_{j1}^a, w_{j2}^a, ..., w_{j2Ma}^a\}$ denote the jth ART_a weight vector. For ART_b, let $x^b = \{x_1^b ... x_{2Ma}^b\}$ denote F_1^b output vector and let $y^b = \{y_1^b ... y_{Na}^b\}$ denote F_2^b. And let $w_k^b = \{w_{k1}^b, w_{k2}^b, ..., w_{k2Mb}^b\}$ denote the kth ART_b weight vector. For the map field, let $x^{ab} = \{x_1^{ab} ... x_{Na}^{ab}\}$ denote the F^{ab} output vector, and let $w_j^{ab} = \{w_{j1}^{ab}, ..., w_{jNb}^{ab}\}$ denote the weight vector from the jth F_2^a node to F^{ab}. Vectors x^a, y^a, x^b, y^b, and x^{ab} are set to 0 between input presentations.

3 Map Field Activation

The map field F^{ab} is activated whenever one of the ART_a or ART_b categories is active. If node J of F_2^a is chosen, then its weights w_j^{ab} activate F^{ab}. If node K in F_2^b is active, then node K in F^{ab} is activated by 1-to-1 pathways between F_2^b and F^{ab}. If both ART_a and ART_b are active, then F^{ab} becomes active only if ART_a predicts the same category as ART_b via the weights w_j^{ab}. The F^{ab} output vector x^{ab} obeys the following:

$$x^{ab} = \begin{cases} y^b \wedge w_j^{ab}, & IF \ the \ Jth \ F_2^a \ node \ is \ active \ and \ F_2^b \ is \ active. \\ w_j^{ab} & IF \ the \ Jth \ F_2^a \ node \ is \ active \ and \ F_2^b \ is \ inactive. \\ y^b & IF \ F_2^a \ is \ inactive \ and \ F_2^b \ is \ active. \\ 0 & IF \ F_2^a \ is \ inactive \ and \ F_2^b \ is \ inactive. \end{cases} \quad (1)$$

From (1), $x^{ab} = 0$ if the prediction w_J^{ab} is disconfirmed by y^b. Even such a mismatch triggers and ART_a search for a better category, as follows.

4 Match Tracking

At the start of each input presentation, the ART_a vigilance parameter ρ_a equals to baseline vigilance ρ_a. The map field vigilance parameter is ρ_{ab}

$$IF \quad |x^{ab}| < \rho_{ab} |y^b|. \quad (2)$$

Then ρ_a is increased until it is slightly larger than $|A \wedge w_J^a| |A|^{-1}$, where A is the input to F_1^a, in complement coding form, and

$$|x^a| = |A \wedge w_J^a| < \rho_a |A|. \quad (3)$$

Where *J* is the index of the active F_2^a node. When this occurs, ART_a search leads either to activation of another F_2^a node *J* with

$$|x^a| = |A \wedge w_J^a| \geq \rho_a |A|. \quad (4)$$

and

$$|x^a| = |y^b \wedge w_J^{ab}| \geq \rho_a |y^b|. \quad (5)$$

Or, if no such node exists, i.e., the input pattern to F_2^a layer does not match any pattern; the input pattern is classified as a new pattern.

5 Comparison of Two Input Patterns

Let's consider the numbers 2 and 3 as 5×5 binary input patterns (see Fig. 2).

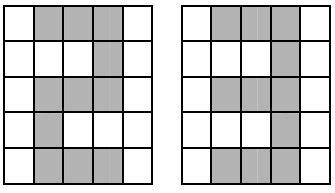


Figure 2: Two input patterns

The gray areas have a value of 1 and other areas have a value of 0. The sum of the values for the two patterns is equal to 11. In order to normalize each

pattern, each cell of the patterns is divided by 11; the result is 0.09 for all the cells.

After the pattern for the number 2 and then the pattern for number 3 are presented to the network, the output was observed by providing the network with an input pattern similar to 2. Assume a case in which the pattern for 2 is learned by the network before the pattern for 3. In this case if the input pattern depicted in Fig. 3 is presented to the network, the predicted output will be 2. If the pattern for 3 is learned first, 3 will be the output of the network if the input pattern is (presented to the network) is the pattern in Fig. 3.

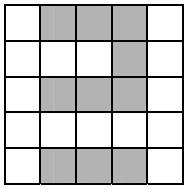


Figure 3: The input Pattern

Therefore, the behavior of the network in this case, depends completely on the order of training patterns and the winning pattern is not predictable. In the next section, the weighted function is presented and compared with the regular function.

5 Weighting Equation for the Input Binary Pattern

The binary pattern, including only zeros and ones, is converted to an analogue pattern through the following equation [4]:

$$a_{m\times n}^{new} = \left(\sum_{i=m-1}^{m+1} \sum_{j=n-1}^{n+1} a_{ij} - 1 \right) \times a_{m\times n}^{old} \quad . \qquad (6)$$

The above equation is explained by using an example: Consider the input pattern depicted in Fig. 2. At first, the pattern is like the pattern in Fig. 4-(a).

0	1	1	1	0
0	0	0	1	0
0	1	1	1	0
0	1	0	0	0
0	1	1	1	0

(a)

0	1	3	2	0
0	0	0	4	0
0	2	4	2	0
0	4	0	0	0
0	2	3	1	0

(b)

Figure 4: (a) Crisp pattern, (b) Weighting the input pattern (Fuzzified)

Using (6) the input pattern is converted to the pattern shown in Fig. 4-(b). Each of the obtained values is calculated by multiplying each value in the pattern (zero or one) by the number of ones in the 8 positions surrounding that value. For example for $a_{2\times4}$ is obtained by using the following equation:

$$\begin{aligned} a_{2\times4}^{new} = a_{2\times4}^{old} \times (& [a_{1\times3} + a_{1\times4} + a_{1\times5}] \\ & + [a_{2\times3} + a_{2\times4} + a_{2\times5}] \\ & + [a_{3\times3} + a_{3\times4} + a_{3\times5}] - 1) = 4 \quad . \end{aligned} \qquad (7)$$

Similarly the obtained pattern for the number 3 is as follows:

0	1	3	2	0
0	0	0	4	0
0	1	4	3	0
0	0	0	4	0
0	1	3	2	0

Figure 5: Weighting of the pattern for the number 3

Fig. 6 shows the weighted pattern for the input pattern depicted in Fig. 3.

0	1	3	2	0
0	0	0	4	0
0	1	3	2	0
0	0	0	0	0
0	1	2	1	0

Figure 6: The weighted input pattern

The normalized patterns for the patterns for the number 2 and 3 and the input pattern shown in Fig. 3 are depicted in Fig. 7-(a), 7-(b) and 7-(c).

The normalized value are obtained by dividing the summation calculated in (6) by the summation of all the values in the pattern. The equation is as follows:

$$a_{m\times n}^{norm} = \frac{a_{m\times n}^{new}}{\sum_{i=1}^{k} \sum_{j=1}^{l} a_{ij}^{new}} \quad . \qquad (8)$$

k is the number of rows and l is the number of columns of the pattern. It is noticeable that in the preprocessing layer and before (6) is used, the dimensions of the matrix by which the pattern is presented, should be extended to $k+2$ and $l+2$ and after the calculations the dimensions should be returned to their original values.

0	0.03	0.1	0.07	0
0	0	0	0.14	0
0	0.07	0.14	0.07	0
0	0.14	0	0	0
0	0.07	0.1	0.03	0

(a)

0	0.03	0.1	0.07	0
0	0	0	0.14	0
0	0.03	0.14	0.1	0
0	0	0	0.14	0
0	0.03	0.1	0.07	0

(b)

0	0.03	0.1	0.07	0
0	0	0	0.14	0
0	0.03	0.1	0.07	0
0	0	0	0	0
0	0.03	0.07	0.03	0

(c)

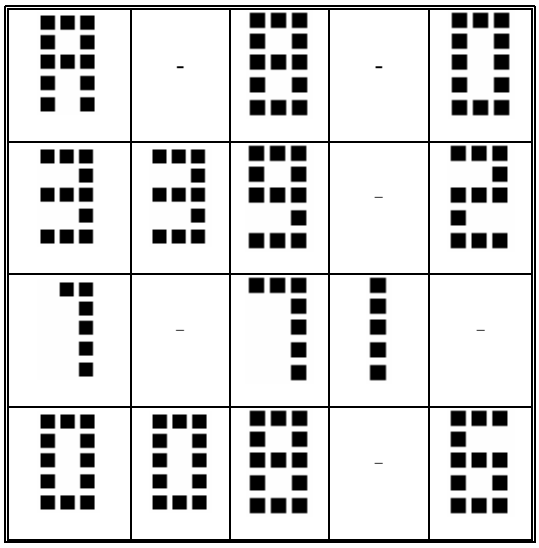


Figure 7: Weighted patterns for 2, 3 and the input pattern

5 Simulation Results

As mentioned in section 6, the input patterns are weighted by using (6). The training patterns are given in Table 1.

Table 1: Example of title of a table

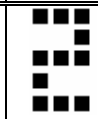
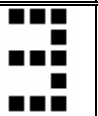
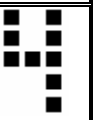
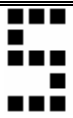
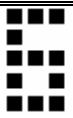
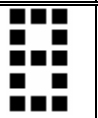
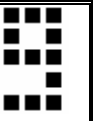
0	1	2	3	4
				
5	6	7	8	9
				

Table 2 depicts the outputs when specific input patterns are fed into the network. .

Table 2: Example of title of a table

Test Input Patterns	Priority of recognition			
	First	Second	Third	Fourth
	Equal	Decrease	Increase	Transmit

For the first pattern in Table 2, the network examines the patterns which are “Equal” to the input pattern. If the pattern is not a complete match, then the output is the pattern which when “Decreased” the result is the input pattern. If these trials fail, the network gives the output which when “Increased”, the result will be the input pattern. For the last attempt, the network suggests the pattern which when one of its values is “Transmitted” the input pattern is obtained. If by using these techniques no suggestions are made by the network, the input pattern is considered as new training pattern and learns that pattern in an on-line manner.

0	1	1	1	0
0	0	0	1	0
0	1	0	1	0
0	0	0	1	0
1	0	0	1	0

Figure 8: The noisy input pattern

0	1	3	2	0
0	0	0	3	0
0	0	0	2	0
0	0	0	2	0
0	0	0	1	0

Figure 9: The preprocessed pattern

Generally, the applied techniques are as Equal-Increase-Decrease-Transmit and therefore the proposed method is called “EDIT”. By using this technique, the random prediction behavior of the

network is enhanced to present a prioritized and organized prediction behavior.

Another advantage of this method is compression of data. This is done by deletion of some of noise values/data in the input patterns. For example if the noisy pattern shown in Fig. 8 is fed into the network, in the preprocessing layer, the proposed method converts it to the pattern depicted in Fig. 9.

As can be seen in Fig. 9, the noise in patterns is rejected. This property increases the speed of learning during the training process because the noisy patterns are not considered as new training patterns. This also leads to a faster selection of the winning neuron and reduces the error.

5 Conclusions

In training Fuzzy ARTmap neural networks, a well-organized and prioritized prediction method is proposed. The prioritization takes place when the EDIT technique is applied. If this method is not used, the outputs selection is done almost randomly or at most according to the order of presentation of training patterns to the network in the learning process.

Another advantage of using this method is converting the binary patterns to analogue (Fuzzified) patterns besides eliminating the noise that may exist in the input patterns and the noise in patterns is rejected. This property increases the speed of learning during the training process.

References

[1] G. A Carpenter, & S Grossberg, (1987b). " ART 2: Stable self organization of pattern recognition codes for analog input paterns." Applied Optics, 26, 4919-4930.

[2] G. A Carpenter, & S Grossberg, (1990b). " ART 3: search using chemical transmitters in self – organization pattern recognition architectures," Neural Networks, 3, 129-152.

[3] G. A Carpenter, S Grossberg, & J, H Reynolds (1990b). "ARTMAP: Supervised real time learning and classification of nonstationary data by a self-organizing neural network," Neural Networks, 4, 565-588.

[4] Shaun Marriot , Rpbert F .Harrisson "A Modified Fuzzy ARTMAP Architecture For The Approximation of Noisy Mapings"IEEE Neural Networks, Vol.8.No.4, pp.619_641,1995.

[5] G. A Carpenter. S Grossberg, N Markuzon. J. H. Reynolds J.H (1991a) "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network." Neural Networks, 4(5), 565-588.

[6] F, Y, Liang, C, C, & Lin, W, C. (1990). "ART 1 neural network for reducing search space in 3-D object recognition uslng multiple views," International Joint Conference on Neural Networks (San Diego), II (pp, 931-936), Piscataway, NJ: IEEE Service Center.

Time series prediction by Perturbed Fuzzy Model

Paulo Salgado
UTAD, Vila Real, Portugal
psal@utad.pt

Getúlio Igrejas
IPB, Bragança, Portugal
igrejas@ipb.pt

Fernando Gouveia
UTAD, Vila Real, Portugal
fgouveia@utad.pt

Abstract

This paper presents a fuzzy system approach to the prediction of nonlinear time series and dynamical systems based on a fuzzy model that includes its derivative information. The underlying mechanism governing the time series, expressed as a set of IF–THEN rules, is discovered by a modified structure of fuzzy system in order to capture the temporal series and its temporal derivative information. The task of predicting the future is carried out by a fuzzy predictor on the basis of the extracted rules and by the Taylor ODE solver method. We have applied the approach to the benchmark Mackey-Glass chaotic time series.

Keywords: Derivative approximation, Times Series, Fuzzy modelling.

1 Introduction

Time series is widely observed in many aspects of our lives. Daily temperature, stock market and so forth are examples of time series. A time series is a continuous $x(t)$ or a discrete sequence of measured quantities $x_1, x_2, \cdots, x_n$ taken from human activity data or some physical system. Basically, there are three main goals in time series analysis: prediction, modelling and characterization [6]. The goal of prediction is to accurately forecast the short-term evolution of the system, modelling aims to precisely capture the features of the system's long-term behaviour, and the purpose of system characterization is to determine some underlying fundamental properties of the system and the nature of observations. Forecasting refers to a process by which the future behaviour of a dynamical system or data series is estimated based on our understanding and characterization of the system and the laws of its nature.

Much effort has been devoted over the past several decades to develop and improve the time-series forecasting models. This task is most complex and hard due to multiple reasons, such as: high sensibility to initial conditions in the unstable dynamical system [4]; difficulty on the determination of trends or on the recognition of patterns in presence of stochastic noise on observable sequence data; or just the natural uncertainty, vagueness and incompleteness of data. However, in the absence of something better, there are some statistical or empirical solutions to make reasonable predictions [11]. Most of these linear approaches, such as the well-known Box–Jenkins method, have shortcomings [3][12]: all have a natural default in that they lack the ability to directly incorporate the natural linguistic information in their modelling or in their strategies, even to extract relevant linguistic information from the data series.

More recently, neural networks and fuzzy logic modelling have been applied to the problem of forecasting complex time series. The main advantage of these methodologies is that we do not need to specify a priori the structure of a model, which is clearly needed in the classical regression analysis [14][11]. Also, both models are nonlinear in nature and they can more easily approximate complex nonlinear systems than simple linear statistical models. Of course, there are also disadvantages to statistical regressions models, where we can use the information given by their parameters to understand the process. This problem can be greatly reduced if alternatively the relationships of the process are expressed by linguistic relationships, which are transparent and easily read by an expert, through the use of a fuzzy modelling.

Fuzzy Systems (FS) have been successfully applied to a number of scientific and engineering fields in recent years, but their performance is

highly dependent on their structure (hierarchical or flat structure, number of inputs, the partition of the input and output spaces by membership functions, and their shape), their inference mechanism, their aggregation operations, and their fuzzification and defuzzification methodology. Various schemes of fuzzy modelling, with specific algorithms, have been proposed to capture some specificity of the problem at hand and thus contribute to the efficiency of the overall identification schemes [18], if necessary by taking advantage of numeric experimental data and of neurofuzzy techniques [8][15]. Various fuzzy predictors (FP's) of time series have been introduced and have demonstrated their success in accurate predictions [2][7][16][19], whose main research effort was made in the creation of appropriate fuzzy rules and memberships through a learning process.

However, this effort generally does not guarantee a model with reasonable dynamical information, i.e. the derivative information of the time series. So, we propose a new fuzzy system structure that is capable of approximating regular functions as well as their derivatives on compact domains with linguistic information. Here, the linguistic information is associated to the translation process of fuzzy sets within the fuzzy relationships, that when modelled is able to describe local trends of the fuzzy models (temporal or positional derivatives). With derivative models of the temporal series, for the attractive regions of work, it is possible to make a Taylor series that can approximate a solution of ordinary differential equations or of a temporal series in distinct regions of the space. With this result it is possible to use the traditional ODE method to solve dynamical equations or to model temporal series.

This paper is organized as follows. Section 2 describes the perturbed fuzzy system. Section 3 presents theoretical aspects of the approximation of a function and its derivatives by the perturbed fuzzy system. Section 4 presents an ODE Fuzzy solver method based on the perturbed fuzzy system. The proposed algorithm will be used to benchmark prediction problems of the Mackey-Glass chaotic time series (Section 5). Finally, a conclusion is drawn.

2 The Perturbed Fuzzy System

Fuzzy systems modelling [9] provides a framework for modelling complex nonlinear relations, using a rule-based methodology. Consider a system $y = f(\boldsymbol{x})$; y is the output (or consequent) variable and $\boldsymbol{x} = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$ is the input vector (or antecedent) variable. Let $\boldsymbol{U} = U_1 \times \cdots \times U_n$ be the domain of the input vector $\boldsymbol{x} \in \mathbb{R}^n$ and V the output space.

A linguistic model relating variables x and y can be written as a collection of rules that link terms $A_{i,j_i} \in U_i$, $j_i = 1, \ldots, N_i$, $i = 1, \ldots, n$, and $B_j \in V$, $j = (j_1, \ldots, j_n)$, where $A_{i,j_i}(x_i)$ and $B_j(y)$ represent the descriptor sets associated, respectively, to variables x_i, $i = 1, \ldots, n$ and y. In fuzzy systems modelling, this relationship is represented by a collection $\Im$ of fuzzy IF–THEN rules:

$$\begin{aligned} R_j&: \text{IF } x_1 \text{ is } A_{1,j_1} \text{ and ... and } x_n \text{ is } A_{n,j_n} \\ &\text{THEN } y \text{ is } B_j \end{aligned} \tag{1}$$

where $j = (j_1, \ldots, j_n)$ is the index of rule, which belongs to the index set:

$$J = \{ j_1, \cdots, j_n \mid j_i = 1, \cdots, N_i; i = 1, \cdots, n \}.$$

The input space $\boldsymbol{U} = U_1 \times \cdots \times U_n$ and the output space V are being partitioned in $N = \prod_{i=1}^{n} N_i$ fuzzy regions, in which it is possible to define N fuzzy rules of the form of (1). The rule base can be represented by the fuzzy relation defined on the Cartesian product $A \times B$ [18]. If each input space U_j (for $j = 1, \ldots, n$) is completely partitioned by N_j fuzzy sets, then there is always at least one active rule. Given values for the input variables $\boldsymbol{x} = \boldsymbol{x}^*$, the value of y is calculated as a fuzzy subset G by a fuzzy inference process:

1. For each rule j, find the firing level of the rule:

$$A_j(\boldsymbol{x}) = A_{j_1}(x_1) \bigstar A_{j_2}(x_2) \bigstar \cdots \bigstar A_{j_n}(x_n) \tag{2}$$

With the linguistic connective "and" of the antecedent of rule (1) defined as T-norm operation, "$\bigstar$", A_j can be viewed as the fuzzy set $\times_{i=1}^{n} A_{i,j_i}$ with membership functions $A_j(\boldsymbol{x})$.

2. The fuzzy implication of each rule R_j: $A_j \mapsto B_j$ is a fuzzy set in $U \times V$ that is defined as $R_{j:A \mapsto B}(\boldsymbol{x}, y) = A_j(\boldsymbol{x}) \otimes B_j(y)$, where "$\otimes$" is an operator rule of fuzzy implication. For each rule j, calculate the effective output value G_j, based of sup-star composition:

$$G'_j(y) = \sup_{\boldsymbol{x} \in U} \left[A'(\boldsymbol{x}) \bigstar R_{j:A \mapsto B}(\boldsymbol{x}, y) \right]$$

3. Combine (unite) the individual outputs of the activated rules to find the overall system output $G = \bigcup_{j=1}^{N} G_j$.

Generally, $A'(\boldsymbol{x})$ is considered as a singleton set. For the arithmetic inference process, the output of each rule *j* is given by $G_j(y) = A_j(\boldsymbol{x}) \cdot B_j(y)$. In many situations, e.g. in series prediction and modelling applications, it is desirable to have a crisp value *y** for the output of a fuzzy system, instead of a fuzzy value *G*(*y*). The defuzzifier maps the fuzzy sets in *V* to crisp points in *V*. In this paper, a centre-average defuzzifier [17] is used, and the output expression of the Fuzzy System can be written as

$$g(\boldsymbol{x}) = \sum_{j=1}^{N} A_j(\boldsymbol{x}) \cdot \theta_j \Big/ \sum_{j=1}^{N} A_j(\boldsymbol{x}) \tag{3}$$

where θ_j is the centroid point in *V* for which the membership function $B_j(y)$ achieves its maximum value, being assumed that $B_j(y)$ is a normal fuzzy set, i.e., $B_j(\theta_j) = 1$.

Fuzzy identification systems are able to integrate information from different sources, namely from human experts and from experimental observation. However, the process of translating this knowledge to linguistic IF–THEN rules is made as a sequence of static or instantaneous pictures of the modelled process, and so the dynamical information is discarded. However, the state variables of a dynamical process or temporal series are not static: in each instant they possess an instantaneous value *and* a trend of evolution. This trend, which contains information of the derivatives, should also be modelled by the diffuse system. A simple way to do this is to ensure that each input and output fuzzy set captures the trend, obtaining what we call a perturbed fuzzy set. These perturbed fuzzy sets are traditional fuzzy sets, characterized by their static position and shape, added with *potential* velocity, acceleration, etc. In the context of this paper, we are concerned with a special type of perturbed fuzzy sets, due both to a translation process and an additive process.

Definition 1: The nonlinear translation of a fuzzy set *A* of *U* by $h \in U$, denoted A_h, is the fuzzy subset of *U* defined as $A_h(x) = A(x - \sigma(h))$, where $\sigma(h)$ is a nonlinear homogenous translation function of the perturbation variable *h*, i.e., $\lim_{h \to 0} \sigma(h) = 0$.

Perturbation *h* moves fuzzy set *A* from its natural position to another position in the neighbourhood. As a special and known case we have $\sigma(h) = h$.

Definition 2: Let $\rho(h, x)$ be an additive perturbed function resulting from the product of a translation function weight $\sigma(h, x)$ by the membership function $A(x)$:

$$\rho(h, x) = \sigma(h, x) A(x)$$

where $\sigma(h, x)$ is a homogenous, nonlinear function of variable *x* and perturbation *h*, i.e., $\lim_{h \to 0} \sigma(h, x) = 0$ and $\sigma(h, x) = 0$. The additive perturbed fuzzy set of *A* is $A_h(x) = \varphi(x, h) A(x)$, where $\varphi(x, h) = 1 + \sigma(h, x)$.

For convenience of representation, consider that $\sigma_x(h) = \sigma(h, x)$. Both previously defined perturbed fuzzy sets obey the following axiom.

Axiom 1: Let $1 \le p \le \infty$ and $f \in L^p(\mathbb{R}^n)$. For $h \in \mathbb{R}^n$, let $A_h(x)$, the perturbed function of Definitions 1 and 2, be continuous with respect to variable *h*. Then $\lim_{h \to 0} \| A_h(x) - A(x) \| = 0$.

Let $A_{h,j}(\boldsymbol{x}) = A_{h,j_1}(x_1) \star \cdots \star A_{h,j_n}(x_n)$ be the aggregation of perturbed membership functions. For the product *T*-norm operation and perturbations of the additive type, we have $A_{h,j}(\boldsymbol{x}) = A_{h,j}(\boldsymbol{x}) \varphi_j(\boldsymbol{x}, \boldsymbol{h})$, where $\varphi_j(\boldsymbol{x}, \boldsymbol{h}) = \varphi_{h_1, j_1}(x_1, h_1) \cdots \varphi_{h_n, j_n}(x_n, h_n)$ and $\boldsymbol{h} = (h_1, \cdots, h_n)^T$ is the perturbation vector.

The fuzzy relationships that involve fuzzy set *A* are consequently also perturbed, and that reflects into the fuzzy system. The result of the perturbed fuzzy sets is a perturbed fuzzy system that is equal to the static fuzzy system when the perturbation variables *h* are null.

Definition 3: A perturbed fuzzy system, PFS, results from the perturbation of input and output fuzzy sets of fuzzy system (3). Let the input fuzzy sets of rules be of the additive type, i.e., $A_{h,j}(x) = \varphi_j(x, h) A_j(x)$, and let the output fuzzy sets be of the nonlinear translation type, $B_{\boldsymbol{h},j}(x) = B_j(x - \sigma_j(\boldsymbol{h}))$. The perturbed version of fuzzy system (3) is:

$$g(\boldsymbol{x}, \boldsymbol{h}) = \frac{\sum_{j=1}^{N} A_{\boldsymbol{h},j}(\boldsymbol{x}) \cdot (\theta_j + \sigma_j(\boldsymbol{h}))}{\sum_{j=1}^{N} A_{\boldsymbol{h},j}(\boldsymbol{x})} \tag{4}$$

Remark: If the input fuzzy sets of the PFS (4) are not perturbed and the perturbation vector *h* is

obtained by difference between vector x and a constant vector $\boldsymbol{x}_0$ (for example, the null vector), $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{x}_0$, then the PFS given by (4) is the well-known TSK model. Moreover, if all input fuzzy sets of the PFS (4) are perturbed by the same additive function, i.e., $\varphi_j(x,h) = \varphi_i(x,h)$, $\forall i, j \in J$, then the PFS of equation (4) is also the well-known TSK model.

3 The Derivative Approximation

In this section we present the sufficient condition for the perturbed fuzzy systems as universal approximators, in order to answer the questions: "Given a real, continuous and differentiable function in $\mathbb{C}^v$, is there a perturbed fuzzy system that can approximate it up to its v^{th} derivative? How to perturb the membership functions and how many fuzzy sets (or fuzzy rules) are needed to ensure the desired approximation accuracy?"

The design of the static fuzzy system $g(\boldsymbol{x})$ is made by choosing the appropriate partition of the input space, the shape of the membership function and its position in the input space $\boldsymbol{U}$ and output space $\boldsymbol{V}$. These structural learning settings are of great importance to approximate the zero-order function $f(\boldsymbol{x})$. The derivative information could be included in the fuzzy modelling by associating it to the potential perturbation of its membership function. Without lost of generality, we will consider that the additive perturbed function $\varphi_j(x,h) = \varphi_j(h)$ is independent of variable $\boldsymbol{x}$. Furthermore, we assume that the perturbed functions are approximated by multivariate polynomials of the multivariate variable perturbation $\boldsymbol{h}$.

Definition 4: Let the perturbed functions $\sigma_j(\boldsymbol{h})$ and $\varphi_j(\boldsymbol{x},\boldsymbol{h})$, presented on Definition 1 and 2, be multivariate polynomials of degree r and s, respectively, defined on compact set $U \subset \mathbb{R}^n$, i.e.:

$$Q_{s,j}(\boldsymbol{h}) = \varphi_j(\boldsymbol{h}) = \sum_{d_1=0}^{s_1} \cdots \sum_{d_n=0}^{s_n} a^j_{d_1 \cdots d_n} h_1^{d_1} \cdots h_n^{d_n} \qquad (5)$$

$$P_{r,j}(\boldsymbol{h}) = \theta_j + \sigma(\boldsymbol{h}) = \sum_{d_1=0}^{r_1} \cdots \sum_{d_n=0}^{r_n} b_{d_1 \cdots d_n} h_1^{d_1} \cdots h_n^{d_n} \quad (6)$$

where $\sum_{i=1}^{n} r_i = r$, $\sum_{i=1}^{n} s_i = s$, $v=r+s$, $b^j_{0,\cdots,0} = \theta_j$ and $a^j_{0,\cdots,0} = 1$, for $j = 1, \cdots, N$.

The perturbed fuzzy system, PFS, is now a rational function of polynomials of variable h:

$$g(\boldsymbol{x},\boldsymbol{h}) = \frac{\sum_{j=1}^{N} A_j(\boldsymbol{x}) Q_{s,j}(\boldsymbol{h}) P_{r,j}(\boldsymbol{h})}{\sum_{j=1}^{N} A(\boldsymbol{x}) Q_{s,j}(\boldsymbol{h})} \qquad (7)$$

Two new theorems show us that fuzzy system (7) can: first, approximate a v^{th}-order polynomial to any degree of approximation; second, extrapolate the last result to approximate any nonlinear function.

Theorem 1: Consider the perturbed fuzzy system of equation (7). It can approximate any N distinct polynomials of order v, $T_{v,\boldsymbol{x}_j}(\boldsymbol{h}) = \sum_{|\boldsymbol{\alpha}|=1}^{v} c_{\boldsymbol{\alpha}}(\boldsymbol{x}_j)\boldsymbol{h}^{\boldsymbol{\alpha}}$, in N distinct nodes $\boldsymbol{x}_j \in S$. Also, the i^{th} derivative of $g(\boldsymbol{x}_j,\boldsymbol{h})$ in respect to $\boldsymbol{h}$ can approximate the i^{th} derivative of $T_{v,\boldsymbol{x}_j}(\boldsymbol{h})$, for $i = 1, \cdots, v$; i.e:

$$\left|E^{(i)}_{\boldsymbol{h}}(\boldsymbol{x})\right| = \left|\frac{\partial^{(i)} T_{r,\boldsymbol{x}}(\boldsymbol{h})}{\partial \boldsymbol{h}^{(i)}} - \frac{\partial^{(i)} g(\boldsymbol{x},\boldsymbol{h})}{\partial \boldsymbol{h}^{(i)}}\right| \le \left|\frac{\partial^{(i+1)} T_{r,\boldsymbol{x}}(\boldsymbol{h})}{\partial \boldsymbol{h}^{(i+1)}}\right|_{\infty} \boldsymbol{h}$$

and $E^{(i)}_{\boldsymbol{0}}(\bar{\boldsymbol{x}}_j) = \lim_{\boldsymbol{h}\to\boldsymbol{0}} E^{(i)}_{\boldsymbol{h}}(\bar{\boldsymbol{x}}_j) = 0$.

Theorem 2: Suppose that N_j overlapped and equally distributed fuzzy sets are assigned to each input variable of Fuzzy System (7). Then, for any given real, continuous and differentiable function $f(\boldsymbol{x})$ defined in $\mathbb{C}^v$ and approximation error bound $\varepsilon > 0$, there is a perturbed fuzzy system (7) with perturbed functions of Definition 4 that guarantees:

i) $g(x) = \lim_{h\to 0} g(x,h)$

ii) $\sup_{x \in \sup(f)} \left(\lim_{h\to 0} \left\| f(x) - g(x,h) \right\|\right) < \varepsilon$

iii) $\sup_{x \in \sup(f)} \lim_{h\to 0} \left\| \frac{\partial^{(i)} f(x)}{\partial x^{(i)}} - \frac{\partial^{(i)} g(x,h)}{\partial h^{(i)}} \right\| < \varepsilon$ for $i = 1,..,v$

4 The Fuzzy ODE Taylor Series Method

A continuous, autonomous, stationary, nonlinear dynamical system can be described by a set of ordinary differential equations, ODE, $dx(t)/dt = F(x(t))$, where $x(t)$ is the vector of system states and F is the system vector field.

By expanding the solution of ODE to the initial value problem, $x(t_0) = x_0$, in a Taylor series about t_0, one obtains a local solution that is valid within its radius of convergence, R_0. If the series is evaluated at t_1, where $t_1 < R_0$, we obtain an

approximation to $x(t_1)=x_1$, and the solution may then be expanded in a new series about t_1. The solution may of course then be extended to point t_2, and so forth, so that by a process of "analytical continuation" one obtains a piecewise polynomial solution to the ODE problem.

Perhaps the simplest one-step methods of order *P* are based on Taylor series expansion of the solution $x(t)$. If $x^{(p+1)}(t)$ is continuous on $[a,b]$, then Taylor's formula gives

$$x_{k+1}=x_k+f(t_k,x_k)h+\cdots+f^{(p-1)}(t_k,x_k)\frac{h^p}{p!}+O(h^{p+1}) \quad (8)$$

where $O(h^{p+1})=f^{(p+1)}(\xi_k,x(\xi_k))\frac{h^{(p+1)}}{p!}$, with $t_k\le\xi_k\le t_{k+1}$, and the total derivatives of f are defined recursively by:

$$f^{(i)}(t,x)=f_t^{(i-1)}(t,x)+f_y^{(i-1)}(t,x)\,f(t,x),\ i=1,2,\cdots.$$

To solve the ODE problem with (8), in each observable sample point x_k, it is necessary to (analytically or numerically) estimate the value of the dynamical system's derivatives. With a perturbed fuzzy system, a multi-derivative modelling can be created in order to be used in solving the prediction problem. The result is a fuzzy system (based on linguistic representation structure) that describes the temporal series as well as its derivatives.

With a set of $N_p=v\times N$ points of the temporal series, with *N* and *v* as in **Definition 4**, where we now use the values of the local Taylor series terms (up to the k^{th}-order continuous derivative), a multivariate Padé approximation is used to identify the coefficients of the polynomials.

Definition 5: Consider a function $f(\boldsymbol{x})$, through its series expansion at a certain point $\boldsymbol{x}_k$ in *U*,

$$f(\boldsymbol{x}_k+\boldsymbol{h})=\sum_{|\boldsymbol{i}|=0}^{\infty}c_{\boldsymbol{i}}(\boldsymbol{x}_k)\boldsymbol{h}^{\boldsymbol{i}} \quad (9)$$

The Fuzzy Padé approximant, $g(\boldsymbol{x},\boldsymbol{h})$, is that rational fuzzy function of degree *v=r+s* in the numerator and *s* in the denominator (polynomials *Q*'s, $Q_j^s(\boldsymbol{h})$), whose power series expansion agrees with a given power series to the highest needed degree of f. If the rational function is of the type of (7), then $g(\boldsymbol{x},\boldsymbol{h})$ is said to be a Fuzzy Padé approximation to series (9), which in those points satisfies the conditions:

$$g(\boldsymbol{x},0)=f(\boldsymbol{x}) \quad (10)$$

and

$$\left.\frac{\partial^{(i)}g(\boldsymbol{x}+\boldsymbol{h})}{\partial\boldsymbol{h}^{(i)}}\right|_{\boldsymbol{h}=\boldsymbol{0}}=\left.\frac{\partial^{(i)}f(\boldsymbol{x}+\boldsymbol{h})}{\partial\boldsymbol{x}^{(i)}}\right|_{\boldsymbol{h}=\boldsymbol{0}},\ 0\le|\boldsymbol{i}|\le v \quad (11)$$

These last two equations provide $v+1$ algebraic equations that involve $(v+1)\times N$ unknown parameters. This approach can be applied in a set of *N* points of space *U*, as much as the number of fuzzy rules of the PFS. The resulting equations allow us to solve the problem of finding the parameters. In this way, coefficients $a^{j}_{d_1,\cdots,d_n}$ of polynomials $Q_j^s(\boldsymbol{h})$ and coefficients $b_{d_1,\cdots,d_n}$ of $P_j^r(\boldsymbol{h})$ will be determined. The Padé approximant is the "best" approximation of a function by a rational function of the given order [1]. A Padé approximant often yields better approximation of the function than truncating its Taylor series and it may still work where the Taylor series does not converge.

5 Numerical Example

The PFS approach has been evaluated for a nonlinear system identification problem, the Mackey-Glass chaotic time-series prediction problem.

The Mackey-Glass time series has been widely used as a standard benchmark for prediction algorithms (Crowder [5], Lapedes and Farber [10], Moody and Darken [13], …). The time series is generated by integrating the delay differential equation,

$$x'(t)=f(\boldsymbol{y}(t)) \quad (12)$$

where $x'(t)=dx(t)/dt$, $\boldsymbol{y}(t)=(x(t),x(t-\tau))^T$ and $f(\boldsymbol{y}(t))=a\,x(t-\tau)/(1+x^c(t-\tau))-b\,x(t)$. With $a=0.2$, $b=0.1$, $c=10$ and $\tau=17$, the time series is chaotic, exhibiting a cyclic but non-periodic behaviour. The upper order of temporal derivatives of state variable $x(t)$ can be recursively defined by:

$$x''(t)=f'_{x(t)}\cdot x'(t)+f'_{x(t-\tau)}\cdot x'(t-\tau)$$

$$x'''(t)=f'_{x(t)}\cdot x''(t)+f''_{x(t-\tau)}\cdot x'(t-\tau)^2+f'_{x(t-\tau)}\cdot x''(t-\tau)$$

They will be used for validating the FPS derivative approximation of the series.

For the calculation of the 3rd derivative of $x(t)$ a number of $3\times\tau$ past samples are necessary to save.

The task's goal is to use known values of the time series up to the point $x = t$ to predict the values at some point in the future $x = t + h$. The standard method for this type of prediction is to create a mapping from *D* points of the time series spaced *h* apart, i.e., vector $\boldsymbol{y}(t) = \left(x\left(t-\left(n_d-1\right)h\right),\ldots,x(t)\right)$, to a predicted future value $x(t+1)$. To allow comparison with earlier work [5][10][13], we used n_d = 4 and h = 6. All other simulation settings in this example were purposely arranged to be as close as possible to those reported in [8]. The numerical solution of equation (12) is obtained by the fourth-order Runge-Kutta with time step 0.1, initial condition $x(0) = 1.2$, and we assume $x(t) = 0$ for $t < 0$. The time series thus generated consisted of 3000 data values, 2700 of which were used as training patterns and the other 300 as test data. The domain space is partitioned by a grid of triangular memberships. The total number of rules created, after excluding all unfired rules, was *N*=137.

Next, the PFS was used for simultaneously modelling the time series $x(t)$ and the derivative time series $x'(t)$, $x''(t)$ and $x'''(t)$. The fuzzy model has consequent polynomials of order 3 and antecedents of order zero. With these models the time series $x(t)$ can be predicted by Taylor ODE solver:

$$x(k+1) = g(\boldsymbol{y}) + g'(\boldsymbol{y})T + g''(\boldsymbol{y})\frac{T^2}{2} + g'''(\boldsymbol{y})\frac{T^3}{6} \quad (13)$$

Note that, in the predicted process the past system output terms $x(k)$, $x(k-6)$, $x(k-12)$, and $x(k-18)$ were replaced by the respective model predictions. This free-running fuzzy model tests the stability of the obtained model.

Figure 1 shows the approximation of Mackey-Glass time series done by the ODE method (13), where $g^{(i)}(x) = \lim_{h\to 0} G^{(i)}(x,h)$, for $i = 0,1,2,3$. In all cases, the resulting time series (from zero to third order) are practically coincident with the correspondent analytical time series. The mean square errors of this time series approximation and its derivatives are shown in Table 1.

6 Conclusions

The ability of the Perturbed Fuzzy Systems, PFS, to approximate any sufficiently smooth function, reproducing its derivatives up to any order, has been demonstrated. The PFS proved its ability to simultaneously estimate functions and their derivatives using information contained in finite numerical samples extracted from the data series, as well as its use in solving ODE problems.

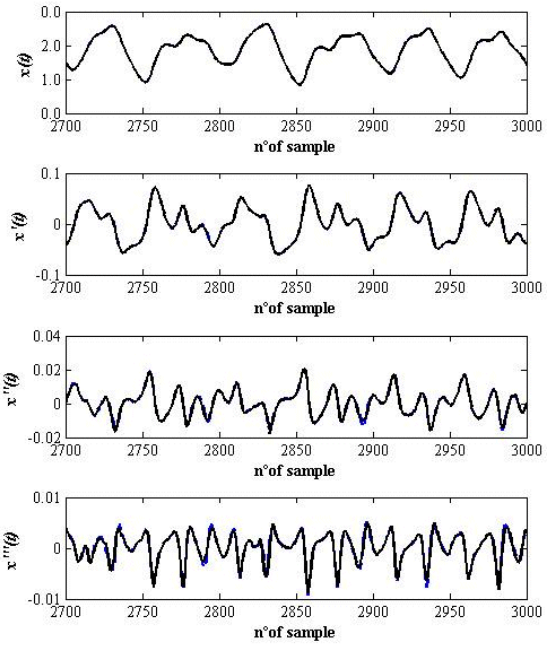


Figure 1: Modelling of the Mackey-Glass time series and its temporal derivative series. The exact time series (solid line) and the fuzzy approximation by PFS (dashed line) are practically coincident.

Table 1: Mean Square Error of the Approximation

ORDER OF DERIVATIVE	MACKEY-GLASS TIME SERIES
0	1.095×10^{-3}
1	9.808×10^{-7}
2	4.538×10^{-7}
3	2.044×10^{-7}

Acknowledgment

This work was supported by the Portuguese Fundação para a Ciência e a Tecnologia (FCT) under grant POSI/SRI/41975/2001.

References

[1] G. A. Baker, Jr. and P. Graves-Morris, *Padé Approximants*, Cambridge University Press, New York, NY, 1996.

[2] D. Benachenhou, Smart trading with (FRET), in: G. J. Deboeck (Ed.), *Trading on the Edge, Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, New York, NY, 1994, pp. 215–242.

[3] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting, and Control*, Holden-Day, San Francisco, CA, 1970.

[4] W. A. Brock, D. A. Hsieh, and B. LeBaron, *Nonlinear Dynamics, Chaos and Instability*, MIT Press, Cambridge, MA, 1991.

[5] R. S. Crowder III, Predicting the Mackey-Glass time series with cascade-correlation learning, in: D. Touretzky, G. Hinton, and T. Sejnowski (Eds.), *Proceedings of the 1990 Connectionist Models Summer School*, Carnegie-Mellon University, Pittsburgh, PA, 1990, pp. 117–123.

[6] James D. Hamilton, *Time series analysis*, Princeton University Press, Princeton, NJ, 1994.

[7] J.-S. R. Jang and C.-T. Sun, Predicting chaotic time series with fuzzy IF–THEN rules, in: *Proceedings of the 2nd IEEE International Conference on Fuzzy Systems* (FUZZ-IEEE'93), San Francisco, CA, 1993, Vol. 2, pp. 1079–1084.

[8] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice-Hall, Englewood Cliffs, NJ, 1997.

[9] George J. Klir and Bo Yuan, *Fuzzy Sets and Fuzzy Logic, Theory and Applications*, Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[10] A. S. Lapedes and R. Farber, Nonlinear signal processing using neural networks: prediction and system modelling, in: *Technical Report LA-UR-87-2662*, Los Alamos National Laboratory, Los Alamos, NM, 1987.

[11] G. S. Maddala, *Introduction to Econometrics*, Prentice-Hall, Englewood Cliffs, NJ, 1996.

[12] M. Magdon-Ismail, A. Nicholson, and Y. S. Abu-Mostafa, Financial markets: very noisy information processing, in: *Proceedings of the IEEE*, Vol. 86, No. 11, Nov. 1998, pp. 2184–2195.

[13] J. Moody and C. J. Darken, Fast learning in networks of locally-tune d processing units, in: *Neural Computation*, Vol. 1, 1989, pp. 281–294.

[14] A.-P. N. Refenes, A. N. Burgess, and Y. Bentz, Neural networks in financial engineering: a study in methodology, in: *IEEE Transactions on Neural Networks*, Vol. 8, No. 6, Nov. 1997, pp. 1222–1267.

[15] L. H. Tsoukalas and R. E. Uhrig, *Fuzzy and Neural Approaches in Engineering*, Wiley, New York, NY, 1997.

[16] L.-X. Wang and J. M. Mendel, Generating fuzzy rules by learning from examples, in: *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, Nov. 1992, pp. 1414–1427.

[17] Li-Xin Wang, *A course in fuzzy systems and control*, Prentice-Hall PTR, Upper Saddle River, NJ, 1997.

[18] R. R. Yager and D. Filev, *Essentials of Fuzzy Modeling and Control*, Wiley, New York, NY, 1994.

[19] Z. Ye and L. Gu, A fuzzy system for trading the Shanghai stock market, in: G. J. Deboeck (Ed.), *Trading on the Edge, Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, New York, NY, 1994, pp. 207–214.

Competitive Self-adaptation in Evolutionary Algorithms

Josef Tvrdík
University of Ostrava
josef.tvrdik@osu.cz

Ivan Křivý
University of Ostrava
ivan.krivy@osu.cz

Abstract

Heuristic search for the global minimum is studied. This paper is focused on the adaptation of control parameters in differential evolution (DE) and in controlled random search (CRS). The competition of different control parameter settings is used in order to ensure the self-adaptation of parameter values within the search process. In the generalized CRS the self-adaptation is ensured by several competing local-search heuristics for the generation of a new trial point. DE was experimentally compared with other adaptive algorithms on a benchmark, self-adaptive CRS was compared in estimation of regression parameters on NIST nonlinear regression datasets. The competitive algorithms outperformed other algorithms both in the reliability and in the convergence rate.

Keywords: Global optimization, Differential evolution, Controlled random search, Self-adaptation.

1 Introduction

We deal with the heuristic search for the global optimization problem defined as follows: for a given objective function

$$f : S \to \mathbb{R}, \quad S \subset \mathbb{R}^D\,,$$

the point $\boldsymbol{x}^*$ is to be found such that $\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in S} f(\boldsymbol{x})$. The point $\boldsymbol{x}^*$ is called the global minimum point and S is the search space. We focus on the problems, where the objective function is continuous and the search space is closed compact set, $S = \prod_{d=1}^{D}[a_d, b_d],\ a_d < b_d,\ d = 1, 2, \ldots, D$. This specification of S is called box constrains.

The problem of the global optimization is hard. That is why stochastic (heuristic) algorithms are used for its solution, see e.g. [2, 12]. The authors of many stochastic algorithms claim the efficiency and the reliability of the searching for the global minimum. The reliability means that the point with minimum function value found in the search process is sufficiently close to the global minimum point and the efficiency means that the algorithm finds a point sufficiently close to the global minimum point at reasonable time. However, when using such algorithms, we face the problem of their control-parameters setting. The efficiency and the reliability of many algorithms is strongly dependent on the values of control parameters. A user is supposed to be able to change the parameter values according to the results of trial-and-error preliminary experiments with the search process. Such attempt is time-consuming.

Adaptive robust algorithms reliable enough at reasonable time-consumption without the necessity of fine tuning their input parameters have been studied in recent years. Theoretical analysis done by Wolpert and Macready [21] implies, that any heuristic search algorithm cannot outperform the others for all objective functions. In spite of this fact, there is empirical evidence, that some algorithms can outperform others for relatively wide range of problems both in the convergence rate and in the reliability of finding the global minimum point. This paper presents an adaptive procedure of control-parameters setting in the algorithm of differential evolution and self-adaptive controlled random search algorithm, where its adaptation is based on the competition of local-search heuristics.

2 Differential Evolution and Adaptation of its Control Parameters

The differential evolution (DE) introduced by Storn and Price [13] is simple but powerful evolutionary algorithm for global optimization over the box-constrained search space. The algorithm of DE in pseudo-code is

shown as Algorithm 1.

Algorithm 1. Differential evolution

```
1   generate P = (x_1, x_2, ..., x_NP); (points in S)
2   repeat
3       for i := 1 to NP do
4           compute a mutant vector u;
5           create y by the crossover of u and x_i;
6           if f(y) < f(x_i) then insert y into Q
7                            else insert x_i into Q
8           endif;
9       endfor;
10      P := Q;
11  until stopping condition;
```

There are several strategies how to generate the mutant point $\boldsymbol{u}$. One of the most popular variant (called DE/rand/1/bin in the literature [3, 10, 13]) generates the point $\boldsymbol{u}$ by adding the weighted difference of two points

$$\boldsymbol{u} = \boldsymbol{r}_1 + F\,(\boldsymbol{r}_2 - \boldsymbol{r}_3)\,, \tag{1}$$

where $\boldsymbol{r}_1 \neq \boldsymbol{r}_2 \neq \boldsymbol{r}_3 \neq \boldsymbol{x}_i$, $\boldsymbol{r}_1, \boldsymbol{r}_2$ and $\boldsymbol{r}_3$ are taken at random from P and $F > 0$ is an input parameter. Another variant, called DE/best/2/bin generates the point $\boldsymbol{u}$ according to formula

$$\boldsymbol{u} = \boldsymbol{x}_{\min} + F\,(\boldsymbol{r}_1 + \boldsymbol{r}_2 - \boldsymbol{r}_3 - \boldsymbol{r}_4)\,, \tag{2}$$

where $\boldsymbol{r}_1 \neq \boldsymbol{r}_2 \neq \boldsymbol{r}_3 \neq \boldsymbol{r}_4 \neq \boldsymbol{x}_{\min}$, $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ and $\boldsymbol{r}_4$ are taken randomly from P (not coinciding with the current $\boldsymbol{x}_i$), $\boldsymbol{x}_{\min}$ is the point of P with minimum function value, and $F > 0$ is an input parameter.

In both strategies mentioned above, the elements y_d, $d = 1, 2. \ldots, D$ of trial point $\boldsymbol{y}$ are built up by the crossover of its parents $\boldsymbol{x}_i$ and $\boldsymbol{u}$ using the following rule

$$y_d = \begin{cases} u_d & \text{if} \quad U_d \leq CR \quad \text{or} \quad d = l \\ x_{id} & \text{if} \quad U_d > CR \quad \text{and} \quad d \neq l\,, \end{cases} \tag{3}$$

where $l \in \{1, 2, \ldots, D\}$ is a randomly chosen integer, $U_1, U_2, \ldots, U_D$ are independent random variables uniformly distributed in [0, 1), and $CR \in [0, 1]$ is an input parameter influencing the number of elements to be exchanged by crossover. Rule (3) ensures that at least one element x_{id} of $\boldsymbol{x}_i$ is replaced by u_d, even if $CR = 0$.

The differential evolution has become one of the most frequently used algorithms for solving the continuous global optimization problems in recent years [10]. But it is also known that the efficiency of searching for the global minimum is very sensitive to the setting of values F and CR. The recommended values are $F = 0.8$ and $CR = 0.5$, but even Storn and Price in their principal paper [13] use $0.5 \leq F \leq 1$ and $0 \leq CR \leq 1$ depending on the results of preliminary tuning. They also used the population size less than recommended $NP = 10\,D$ in many test tasks.

Many papers deal with the setting of control parameters for differential evolution. Ali and Törn [1] suggested to adapt the value of the scaling factor F within the search process according to the equation

$$F = \begin{cases} \max(F_{\min}, 1 - |\frac{f_{\max}}{f_{\min}}|) & \text{if } |\frac{f_{\max}}{f_{\min}}| < 1 \\ \max(F_{\min}, 1 - |\frac{f_{\min}}{f_{\max}}|) & \text{otherwise}\,, \end{cases} \tag{4}$$

where $f_{\min}$, $f_{\max}$ are the minimum and maximum function values in the population and $F_{\min}$ is an input parameter ensuring $F \in [F_{\min}, 1]$. According to [1] this calculation of F reflects the demand to make the search more diversified at early stage and more intensified at latter stages, i.e. to produce rather larger values of F for large difference $f_{\max} - f_{\min}$, and rather smaller values of F otherwise. The rule (4) works properly only for $f_{\min} > 0$. When $f_{\max} > 0$ and $f_{\min} < 0$, especially if $|f_{\max}| < |f_{\min}|$, the values of F fluctuate very rapidly in $[F_{\min}, 1]$, even if the changes in $f_{\max}$ or $f_{\min}$ are small. However, from practical point of view, it occurs only as a short episode of the search process in most optimization tasks.

Zaharie [22] derived the critical interval for the control parameters of DE. This interval ensures to keep the mean of population variance non-decreasing, which results in the following relationship

$$2\,p\,F^2 - \frac{2p}{NP} + \frac{p^2}{NP} > 0\,, \tag{5}$$

where $p = \max(1/D, CR)$ is the probability of "differential perturbation" according to (3). The relationship (5) implies that the mean of population variance is non-decreasing, if $F > \sqrt{1/NP}$, but practical reason of such result is very limited, because it brings no new information when we compare this result with the minimum value of $F = 0.5$ used in [13] and in other applications of differential evolution.

Some other attempts to the adaptation of DE control parameters have appeared, recent state of adaptive parameter control in differential evolution is summarized by Liu and Lampinen [5].

New idea of self-adaptation of control parameters F and CR in differential evolution was proposed by Brest et al. [3]. The values of F and CR can be changed in each generation with probability τ_1, τ_2, respectively. Thus, the successful values are used more frequently due to the fact, that individuals with lower function level survive longer. New values of F are distributed uniformly in $[\,F_l,\, F_u]$ and new CR are also uniform random values from $[\,0,\, 1]$.

The setting of the control parameters can be made self-adaptive through the implementation of a compe-

tition into the algorithm. This idea, similar to the competition of local-search heuristics in evolutionary algorithm [14] or in controlled random search [15], was proposed recently [17].

Let us have H settings (different values of F and CR used in the statements on line 4 and 5 of Algorithm 1) and choose among them at random with the probability $q_h,\ h = 1, 2, \ldots, H$. The probabilities can be changed according to the success rate of the settings in preceding steps of search process. The h-th setting is successful, if it generates such a trial point $\mathbf{y}$ that $f(\mathbf{y}) < f(\mathbf{x}_i)$. When n_h is the current number of the h-th setting successes, the probability q_h can be evaluated simply as the relative frequency

$$q_h = \frac{n_h + n_0}{\sum_{j=1}^{H}(n_j + n_0)}\ , \tag{6}$$

where $n_0 > 0$ is a constant. The setting of $n_0 \geq 1$ prevents a dramatic change in q_h by one random successful use of the h-th parameter setting. In order to avoid the degeneration of process the current values of q_h are reset to their starting values ($q_h = 1/H$), if any probability q_h decreases bellow a given limit $\delta > 0$.

It is supposed that such a competition of different settings will prefer successful settings. The competition provides an self-adaptive mechanism of setting control parameters to appropriate values for the problem actually solved.

Four variants of such competitive differential evolution were implemented and tested on benchmark in [19]. The benchmark consists of six functions commonly used, see e.g. [1, 10, 13] at four levels of dimension D of search spaces, namely $D = 2$, $D = 5$, $D = 10$ and $D = 30$:

- Ackley's function - multimodal, separable

$$f(\boldsymbol{x}) = -20\exp\left(-0.02\ \sqrt{\tfrac{1}{D}\sum_{d=1}^{D} x_d^2}\right) - \exp\left(\tfrac{1}{D}\sum_{d=1}^{D}\cos 2\pi x_d\right) + 20 + \exp(1)$$

  $x_d \in [-30, 30]$, $\boldsymbol{x}^* = (0, 0, \ldots, 0)$, $f(\boldsymbol{x}^*) = 0$

- First De Jong's function (sphere model) - unimodal, continuous, convex

$$f(\boldsymbol{x}) = \sum_{d=1}^{D} x_d^2$$

  $x_d \in [-5.12, 5.12]$, $\boldsymbol{x}^* = (0, 0, \ldots, 0)$, $f(\boldsymbol{x}^*) = 0$

- Griewank's function - multimodal, nonseparable

$$f(\boldsymbol{x}) = \sum_{d=1}^{D} \frac{{x_d}^2}{4000} - \prod_{d=1}^{D} \cos\left(\frac{x_d}{\sqrt{d}}\right) + 1$$

  $x_d \in [-400, 400]$, $\boldsymbol{x}^* = (0, 0, \ldots, 0)$, $f(\boldsymbol{x}^*) = 0$

- Rastrigin's function - multimodal, separable

$$f(\boldsymbol{x}) = 10\,D + \sum_{d=1}^{D}[{x_d}^2 - 10\cos(2\pi x_d)]$$

  $x_d \in [-5.12, 5.12]$, $\boldsymbol{x}^* = (0, 0, \ldots, 0)$, $f(\boldsymbol{x}^*) = 0$

- Rosenbrock's function (banana valley) - unimodal, nonseparable

$$f(\boldsymbol{x}) = \sum_{d=1}^{D-1}\left[100(x_d^2 - x_{d+1})^2 + (1 - x_d)^2\right]$$

  $x_d \in [-2.048, 2.048]$, $\boldsymbol{x}^* = (1, 1, \ldots, 1)$, $f(\boldsymbol{x}^*) = 0$

- Schwefel's function - multimodal, the global minimum distant from the next best local minima

$$f(\boldsymbol{x}) = \sum_{d=1}^{D} x_d \sin(\sqrt{|\ x_d\ |})$$

  $x_d \in [-500, 500]$, $\boldsymbol{x}^* = (s, s, \ldots, s)$, $s = 420.9687$, $f(\boldsymbol{x}^*) = -418.9829\,D$

The most reliable results (and at the same time the second best ones in time consumption) provided the algorithm DEBR18 with 18 competing settings of control parameters, all the combinations of ($CR = 0$, $CR = 0.5$, and $CR = 1$) and ($F = 0.5$, $F = 0.8$, and $F = 1$) in the both DE/rand/1/bin and DE/best/2/bin strategy. Parameters for the competition of settings were set to $n_0 = 2$, and $\delta = 1/(5\,H)$. Here we present the comparison of this algorithm with standard DE and three other adaptive DE algorithms published recently. The tests were preformed for all the functions at four levels of dimension D of search spaces like in [19]. One hundred of independent runs were carried out for each function and level of D. The search for the global minimum was stopped, if $f_{\max} - f_{\min} < 1e-07$ or the number of objective function evaluations exceeds the input upper limit $20000\,D$. Population size was set to $NP = \max(20, 2\,D)$ in all the tested algorithms except BREST2, where population size recommended in [3], i.e. $NP = 10\,D$ was used. The input parameters for self-adaptation in BREST1 and BREST2 algorithm were set as follows, $\tau_1 = \tau_2 = 0.1$, $F_l = 0.1$, and $F_u = 0.9$.

In the standard DE the recommended control parameter values $F = 0.8$ and $CR = 0.5$ were used. In adaptive DE-Ali the control parameters were set to $CR = 0.5$ and $F_{\min} = 0.45$.

Table 1: Comparison of DE algorithms

Algorithm		DEBR18		DER		DE-Ali		BREST1		BREST2	
Function	D	ne	RP	rne	RP	rne	RP	rne	RP	rne	RP
ackley	2	2409	100	-2	100	2	100	-19	99	-18	100
dejong1	2	1162	100	-1	100	12	100	-18	100	-19	100
griewank	2	2876	100	25	78	5	94	1	96	8	95
rastrig	2	1778	100	-2	99	-2	99	-18	100	-18	99
rosen	2	1956	100	105	100	150	100	72	100	71	100
schwefel	2	1640	100	-3	100	-21	100	-18	99	-18	98
ackley	5	6401	100	1	99	-9	100	-25	100	91	100
dejong1	5	3176	100	-3	100	14	100	-26	100	90	100
griewank	5	8686	100	14	70	9	80	-15	84	127	100
rastrig	5	4989	100	16	95	20	99	-16	98	117	100
rosen	5	6256	100	528	100	331	32	155	97	511	100
schwefel	5	4564	98	-3	98	-33	97	-25	96	91	100
ackley	10	13569	100	14	99	-38	96	-35	97	248	100
dejong1	10	6973	100	6	100	-2	100	-34	100	252	100
griewank	10	13153	99	18	78	-15	83	-37	81	260	100
rastrig	10	10711	100	104	82	78	84	-4	96	414	100
rosen	10	20524	100	429	100	-6	0	130	97	729	100
schwefel	10	9964	99	9	96	-41	90	-31	88	264	100
ackley	30	142208	100	164	100	-58	100	-49	100	179	100
dejong1	30	78664	100	141	100	-23	100	-48	100	182	100
griewank	30	103095	100	174	100	-25	100	-46	98	191	100
rastrig	30	110071	100	445	0	445	0	62	100	445	0
rosen	30	381972	100	57	0	50	0	6	98	57	0
schwefel	30	108050	100	206	100	-41	100	-35	99	246	100

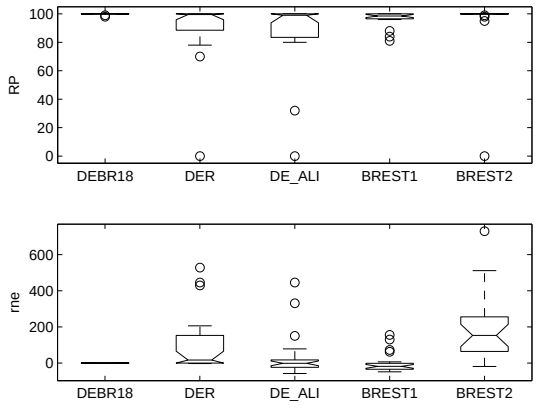


Figure 1: Comparison of DE algorithms – Boxplots of RP and rne

The results of comparison of the algorithms are summarized in Fig. 1. The values for all the test tasks are given in Table 1 in more detail. The reliability of the search in columns RP is the percentage of runs, where the minimum value of objective function obtained by the search duplicates at least four digits, when compared with the right result. The time consumption is expressed as the average number (ne) of the objective function evaluations needed to reach the stopping condition. For easier comparison of the algorithms the (ne) values are given only for DEBR18 algorithm, and the relative change of ne in percents when compared with DEBR18 is presented for the other algorithms in the columns denoted rne in the tables. Therefore, the negative values of rne mean smaller time consumption with respect to DEBR18, the positive values larger one. For example, the value $rne = -50$ means half ne, the value $rne = 100$ means that ne is twice larger, when compared with DEBR18. The number of the objective function evaluations can be easily recalculated as $ne = ne_0 \times (1 + rne/100)$, where ne_0 is the appropriate number of objective function evaluations for DEBR18.

As it is apparent from Fig. 1, DEBR18 is the most reliable algorithm among the algorithms in the test. The second best is BREST1, both in reliability and in convergence rate. BREST2 is highly reliable except two tasks with $D = 30$, but its convergence is much slower in comparison with DEBR18. Competitive setting of control parameters proved to be efficient adaptive scheme.

3 Control Random Search with Competing Heuristics

In an additive nonlinear regression model, the elements of random vector $\boldsymbol{Y}$ are expressed as follows

$$Y_i = g(\boldsymbol{x}_i, \boldsymbol{\beta}) + \varepsilon_i, \quad i = 1, 2, \ldots, n\,, \tag{7}$$

where $\boldsymbol{x}_i^T = (x_1, x_2, \ldots, x_k)$ is i-th row of regressor matrix $\boldsymbol{X}$, $\boldsymbol{\beta}$ is vector of parameters, g is a given function nonlinear in parameters, and ε_i's are *iid* random variables with zero means. The estimation of parameters by the least squares method means to find such estimates of $\boldsymbol{\beta}$ that minimize the residual sum of squares $Q(\boldsymbol{\beta})$ given by the following equation

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[Y_i - g(\boldsymbol{x}_i, \boldsymbol{\beta})\right]^2 . \tag{8}$$

Due to the fact that the function $Q(\boldsymbol{\beta})$ need not be unimodal, the estimation of $\boldsymbol{\beta}$ is the global optimization problem. Iterative deterministic algorithms (e.g. Levenberg-Marquardt) used in standard statistical packages often fail when searching for the true solution of the problem. Several statistical packages were tested [16] on NIST tasks of higher-level-difficulty [8]. For approximately one half of the tasks, the algorithms either completely failed or resulted in a significant disagreement with the true parameter values.

A stochastic algorithm based on controlled random search (CRS) was proposed for the estimation of nonlinear-regression parameters. The CRS algorithm was published originally by Price [9]. The reflection known from simplex method [7] was used for generating a new trial point. There are several modifications of the CRS algorithm, which were successfully used in solving the global optimization problems [1]. This algorithm can be written in pseudo-code as follows.

Algorithm 2. Control Random Search

```
1 generate P (population of N points in D);
2 find x_max (point with max function value);
3 repeat
4    generate a new trial point y using a heuristic;
5    if f(y) < f(x_max) then
6       x_max := y;
7       find new x_max;
8    endif
9 until stopping condition;
```

The role of a heuristic mentioned at line 4 can play any non-deterministic rule generating a new trial point $\boldsymbol{y} \in S$. There are many different heuristics that can be used and, moreover, the heuristics can alternate during the course of search.

Four competing heuristics were used in the implementation of the CRS algorithm for the nonlinear-regression parameter estimation. Three of them are based on a randomized reflection in the simplex Σ ($d+1$ points chosen from P) proposed by [4]. A new trial point $\boldsymbol{y}$ is generated from the simplex by the relation

$$\boldsymbol{y} = \boldsymbol{g} + U\,(\boldsymbol{g} - \boldsymbol{x}_H)\,, \tag{9}$$

where $\boldsymbol{x}_H = \arg\max_{\boldsymbol{x} \in \Sigma} f(\boldsymbol{x})$ and $\boldsymbol{g}$ is the centroid of remaining d points of the simplex. The multiplication factor U is a random variable distributed uniformly in $[s, \alpha - s)$, $\alpha > 0$ and s being input parameters, $0 < s < \alpha/2$. All the $d+1$ points of simplex are chosen at random from P in two heuristics: the first heuristic uses $\alpha = 2$, and $s = 0.5$, the second one uses $\alpha = 5$, and $s = 1.5$. Regarding the third heuristic, one point of the simplex is the point of P with the minimum objective function value and the remaining d points of the simplex S are chosen at random from remaining points of P. Input parameters of this heuristic are set to $\alpha = 2$ and $s = 0.5$. The fourth competing heuristic is based on differential evolution, see (3) and (4) with $F_{\min} = 0.4$ and $C = 0.9$. The algorithm based on the competition of all four heuristics is denoted CRS4.

The rules for the competition of heuristics are similar to those described in Section 2, the success is weighted by its relative change in objective function value

$$w_h = \frac{f_{\max} - \max(f(\boldsymbol{y}), f_{\min})}{f_{\max} - f_{\min}}\,. \tag{10}$$

Thus $w_h \in (0,\ 1]$ and the corresponding probability q_h is evaluated as

$$q_h = \frac{W_h + w_0}{\sum_{j=1}^{H} (W_j + w_0)}\,, \tag{11}$$

where W_h is the sum of w_h in previous searching steps and $w_0 > 0$ is an input parameter of the algorithm.

The collection of NIST datasets [8] was used as a benchmark. This collection contains 27 nonlinear regression datasets (tasks) ordered according to their level of difficulty (lower – 8 tasks, average – 11 tasks, and higher – 8 tasks). The corresponding nonlinear regression models are of exponential or rational types, the number of parameters ranging from 2 to 9. Most of the datasets (18 of 27 tasks) result from experimental studies, the remaining ones are generated artificially. The total number of observations varies in a wide range, from 6 to 250.

Several algorithms for the estimation of parameters were compared. One of them is a modification of Levenberg-Marquardt algorithm implemented in *nlinfit* procedure of Statistical Toolbox [6]. with default values of its control parameters. The other algorithms

Table 2: Comparison of algorithms in estimation of nonlinear-regression parameters

Algorithm		*nlinfit*	CRS4		CRS4e		DER		REFL1	
Task	Level	Result	*RP*	*ne*	*RP*	*rne*	*RP*	*rne*	*RP*	*rne*
chwirut1	lower	OK	100	3008	100	-35	100	574	84	47
chwirut2	lower	OK	100	2987	100	-35	100	571	71	74
danwood	lower	OK	100	1620	100	-28	100	300	81	3
gauss1	lower	OK	100	14137	100	-35	100	530	24	243
gauss2	lower	OK	98	14726	98	-36	98	1039	38	280
lanczos3	lower	OK	100	29810	100	2	0	705	0	444
misra1a	lower	OK	100	2157	100	-17	100	325	2	59
misra1b	lower	OK	100	1861	100	-19	100	386	9	57
enso	middle	OK	87	19220	86	-30	100	722	86	56
gauss3	middle	OK	100	15908	99	-35	99	1912	4	532
hahn1	middle	OK	93	16509	93	-26	0	1596	0	992
kirby2	middle	OK	100	8508	100	-23	100	2251	34	261
lanczos1	middle	L	0	28361	100	639	0	746	0	523
lanczos2	middle	OK	55	28251	100	8	0	750	0	475
mgh17	middle	X	100	11023	100	-18	25	1714	0	381
misra1c	middle	OK	100	2104	100	-11	100	435	0	62
misra1d	middle	OK	100	2043	100	-12	100	414	0	73
nelson	middle	OK	100	5904	100	-17	100	619	0	88
roszman1	middle	OK	100	5301	100	-36	100	1245	91	51
bennett5	higher	L	100	41335	100	-11	5	190	1	-69
boxbod	higher	X	100	1308	100	-37	100	206	90	-1
eckerle4	higher	L	100	2629	100	-35	100	140	94	3
mgh09	higher	L	100	10422	100	-15	100	1427	0	184
mgh10	higher	L	100	20761	100	1	0	478	0	-50
rat42	higher	OK	100	2942	100	-35	100	474	81	42
rat43	higher	OK	100	4807	100	-39	100	904	87	83
thurber	higher	OK	100	13915	100	-30	0	1912	1	824

in the tests are population-based, differential evolution (DER) and two variants of the CRS with competition(CRS4 and CRS4e) and one variant without competition consisting of one local-search heuristic, randomized reflection with $\alpha = 2$, and $s = 0.5$ (REFL1).

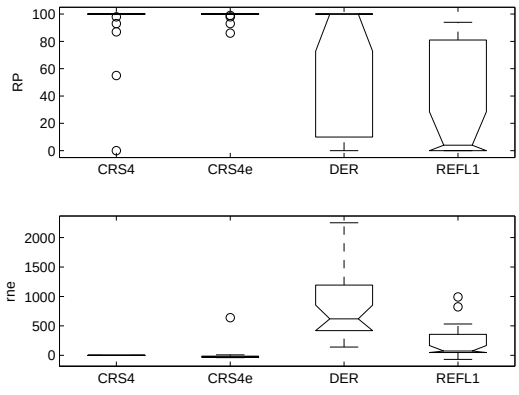


Figure 2: Comparison of stochastic algorithms used in nonlinear regression – Boxplots of *RP* and *rne*

These algorithms were tested in 100 repeated runs for each task. Search spaces S for the individual tasks are given in [18]. The common control parameters of all the population-based algorithms were set up as follows:

- population size $N = 10\,d$,
- stopping condition $R^2_{\max} - R^2_{\min} < \varepsilon$ or $ne \geq 40000\,d$,

where $\varepsilon = 1 \times 10^{-15}$ (except `CRS4e`, see later), $R^2_{\max}$ and $R^2_{\min}$ are maximum and minimum values of the determination index R^2 in population, $R^2 = 1 - Q(\hat{\boldsymbol{\beta}})/\sum_{i=1}^{n}(Y_i - \overline{Y})^2$.

The algorithm CRS4e (described in [18]) contains the same four heuristics as CRS4 and, moreover, provides a procedure for adapting the value ε in the stopping condition. The value ε starts from input value ε_0 and can be decreased, if $1 - R^2_{\max} < \gamma \times \varepsilon$, where $\gamma \gg 1$ is

another input parameter. Values of input parameters of CRS4e were set to $\varepsilon_0 = 1 \times 10^{-9}$ and $\gamma = 1 \times 10^7$.

The results of our experiments are summarized in Table 2. The quantities reported in this table have similar meaning as in the Section 2, *rne* denoting relative change in *ne*, when compared with CRS4. The overall comparison of all the stochastic algorithms used in testing is shown in Fig. 2.

The algorithm CRS4 is reliable enough for most tasks except lanczos1 and lanczos2 (caused by their extremely low values of $1 - R^2$). As regards the other tasks, RP is less than 100 only in four tasks, but it always exceeds 85 %. The best results in both RP and ne were achieved by using CRS4e, where the use of adaptive stopping condition decreases ne by one tenth to one third in most cases and increases RP in tasks with small value of $1 - R^2$.

4 Conclusions

Several self-adaptive evolutionary algorithms were described and the experimental results were briefly presented. The results showed that the algorithms, where the search strategy is adapted by competition (using the different control-parameters settings or using several local-search heuristics), outperform the other algorithms in most test tasks.

Some of these self-adaptive evolutionary algorithms are implemented in the Matlab program library [20] available on website[1]. This program library is free software. Any user of Matlab can download selected function of the library and use and/or modify it under the terms of the GNU General Public License. Although all the algorithms included into the Matlab program library were tested extensively and proved to be high reliable, often with significantly lower time consumption in the comparison with other stochastic algorithms, there is no guarantee of their right performance in other global optimization problems. However, the advantage of these algorithms consists in the fact, that they can run with default values of their control parameters for a wide range of optimization tasks.

The presentation of this non-fuzzy topic at fuzzy-experts forum follows two objectives. The first of them is a challenge for fuzzy approach to the self-adaptation in evolutionary algorithms. The second reason is to offer a new robust optimization procedure for application. The self-adaptive evolutionary algorithms described in this paper can be also applied to the optimization of the Takagi-Sugeno fuzzy models. When optimizing them, it is desirable (see [11]):

- to realize partitioning the premise space of the models, preferably from real data by using an efficient fuzzy clustering algorithm,
- to optimize the model successively by combining an evolutionary algorithm with the fuzzy rule base reduction and simplification,
- to check for meeting both partition and search space constraints during the optimization process.

[1] http://albert.osu.cz/oukip/optimization/

Acknowledgement

The research was supported by the grant 201/05/0284 of the Czech Grant Agency and by the research scheme MSM 6198898701 of the Institute for Research and Applications of Fuzzy Modeling.

References

[1] M. M. Ali , Törn, Population set based global optimization algorithms: Some modifications and numerical studies, Computers and Operations Research 31 (2004) 1703 – 1725.

[2] T. Bäck, Evolutionary Algorithms in Theory and Practice, Oxford University Press, New York, 1996.

[3] J. Brest, S. Greimer, B. Boškovič, M. Mernik, V. Žumer, Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems,. IEEE Transactions on Evolutionary Computation 10 (2006) 646 – 657.

[4] I. Křivý, J. Tvrdík, The controlled random search algorithm in optimizing regression models, Comput. Statist. and Data Anal. 20 (1995) 229 – 234.

[5] J. Liu, J. Lampinen, A Fuzzy Adaptive Differential Evolution Algortithm. Soft Computing 9 (2005) 448 – 462.

[6] MATLAB, version 2006b, The MathWorks, Inc., 2006.

[7] J. A. Nelder, R. Mead, A simplex method for function minimization, Computer J. 7 (1964) 308 – 313.

[8] NIST, Statistical Reference Datasets: nonlinear regression, NIST Information Technology Laboratory, 2001. http://www.itl.nist.gov/div898/strd/.

[9] W. L. Price, A controlled random search procedure for global optimization, Computer J. 20, (1977) 367 – 370.

[10] K. V. Price, R. Storn, J. Lampinen, Differential Evolution: A Practical Approach to Global Optimization, Springer-Verlag, 2005.

[11] H. Roubos, M. Setnes, Compact Fuzzy Models and Classifiers through Model Reduction and Evolutionary Optimization. In: Lance Chambers (Ed.) The Practical Handbook of Genetic Algorithms Applications , chapter 2. New York: Chapman & Hall/CRS 2000, pp. 31–59.

[12] J. C. Spall, Introduction to Stochastic Search and Optimization, Wiley-Intersience, 2003.

[13] R. Storn, K. V. Price, Differential evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, J. Global Optimization 11 (1997) 341 – 359.

[14] J. Tvrdík, L. Mišík, I. Křivý, Competing Heuristics in Evolutionary Algorithms, in: P. Sinčák et al. (Eds.), *2nd Euro-ISCI, Intelligent Technologies - Theory and Applications*. IOS Press, Amsterdam 2002, pp. 159 – 165.

[15] J. Tvrdík, Generalized controlled random search and competing heuristics, in: R. Matoušek and P. Ošmera (Eds.), *MENDEL 2004, 10th International Conference on Soft Computing*. Technical University Press, Brno, 2004, pp. 228 – 233.

[16] J. Tvrdík, I. Křivý, Comparison of algorithms for nonlinear regression estimates, in: J. Antoch (Ed.), *COMPSTAT 2004*. Physica-Verlag, Heidelberg, 2004, pp. 1917 – 1924.

[17] J. Tvrdík, Competitive Differential Evolution, in: R. Matoušek and P. Ošmera (Eds.), *MENDEL 2006, 12th International Conference on Soft Computing*. Technical University Press, Brno, 2006, pp. 7 – 12.

[18] J. Tvrdík, I. Křivý, L. Mišík, Adaptive Population-based Search: Application to Estimation of Nonlinear Regression Parameters, Computational Statistics and Data Analysis (2007), In Press, Corrected Proof available online since 9 November 2006, (http://www.sciencedirect.com/science/article/B6V8V-4M9HWXT-3/2/e4fa1077aa80c154b130396b3c486286)

[19] J. Tvrdík, Differential Evolution with Competitive Setting of its Control Parameters. TASK Quarterly 11 (2007) (In Press).

[20] J. Tvrdík, H. Habiballa, V. Pavliska, Matlab Program Library for Box-Costrained Global Optimization. in: *APLIMAT 2007, 7th International Conference on Applied Math*. Slovak Technical University, Bratislava, 2007, pp. 463 – 470.

[21] D. H. Wolpert, W. G. Macready, No Free Lunch Theorems for Optimization, IEEE Transactions on Evolutionary Computation 1 (1997) 67 – 82.

[22] D. Zaharie, Critical Values for the Control Parameter of the Differential Evolution Algorithms, in: R. Matoušek and P. Ošmera (Eds.), *MENDEL 2002, 8th International Conference on Soft Computing*, Technical University Press, Brno, 2002, pp. 62 – 67.

Differential Evolution methods for LU-fuzzy Arithmetic

Luciano Stefanini
University of Urbino "Carlo Bo", Urbino, ITALY.
lucste@uniurb.it

Abstract

The paper illustrates a differential evolution (DE) algorithm to calculate the level-cuts of the fuzzy extension of a multidimensional real valued function to fuzzy numbers. The method, in the frame of the LU parametrization, decomposes the fuzzy extension engine into a set of "nested" min and max box-constrained optimization problems that can be solved by appropriate forms of the DE algorithm, i.e. based on multi populations which cooperate during the search phase and specialize to find the global min (corresponding to Lower branch) and the global max (Upper branch), both gaining efficiency from the work done for a level-cut to the subsequent ones. A special version of the algorithm is designed to the case of differentiable functions, for which a representation of the fuzzy numbers is used to improve efficiency and quality of calculations. Some preliminary results indicate DE methods as promising tools: in a number of test problems, its computational complexity grows on average superlinearly (of degree less than 1.5) in the number of variables of the function to be extended.

Keywords: LU-Fuzzy numbers, Fuzzy extension principle, Differential evolution method.

1 Introduction

In general, the arithmetic operations on fuzzy numbers can be approached either by the direct use of the membership function (by the Zadeh extension principle) or by the equivalent use of the $\alpha - cuts$ representation. By the $\alpha - cuts$ approach, it is possible to define a parametric representation with the advantage of obtaining a wide family of fuzzy numbers (see [14]). It is well known that the fuzzy extension principle requires to solve a set of optimization problems and different heuristic methods have been proposed to obtain good solutions. Well known fundamental algorithms are the vertex method and its modifications (see [18] and [12]); the transformation method (see [8]) in its general or reduced versions (see [10] for an efficient implementation); a sparse grids method (see [11]). We suggest here two procedures based on the differential evolution (DE) method of Storn and Price (see [16], [17], [13]) and adapted to take into account both the nested property of $\alpha - cuts$ and the min and max problems over the same domains. In particular, we use simultaneous multiple populations that collaborate each other and specialize during the process to find all the required solutions for a specified finite number of $\alpha - cuts$. Computational results are reported that indicate the DE method as a promising tool, as it exhibits, on average for 22 test problems having dimension from 2 to 64, superlinear computational complexity (of degree less than 1.5) in the number of variables.

2 Fuzzy numbers and extension principle

We will consider fuzzy numbers and intervals, i.e. fuzzy sets defined over the field $\mathbb{R}$ of real numbers having a particular form. A general *fuzzy set* over $\mathbb{R}$ is usually defined by its membership function $\mu : \mathbb{R} \longrightarrow \mathbb{T} \subseteq [0,1]$ and a fuzzy (sub)set u of $\mathbb{R}$ is uniquely characterized by the pairs $(x, \mu_u(x))$ for each $x \in \mathbb{R}$; the value $\mu_u(x) \in [0,1]$ is the membership grade of x to the fuzzy set u. Denote by $\mathcal{F}(\mathbb{R})$ the collection of the fuzzy sets over $\mathbb{R}$. Elements of $\mathcal{F}(\mathbb{R})$ will be denoted by letters u, v, w and the corresponding membership functions by μ_u, μ_v, μ_w.

Fundamental concepts in fuzzy theory are the *support*, the *level-sets* (or *level-cuts*) and the *core* of a fuzzy set (see [3] and [4]). Let $\mu_u(x)$ be the membership

function of a fuzzy set u over $x \in \mathbb{R}$. The *support* of u is the (crisp) subset of points of $\mathbb{R}$ at which the membership grade $\mu_u(x)$ is positive: $supp(u) = \{x|x \in \mathbb{R}, \mu_u(x) > 0\}$. For $\alpha \in]0,1]$, the *α-level cut* of u (or simply the $\alpha - cut$) is defined by $[u]_\alpha = \{x|x \in \mathbb{R}, \mu_u(x) \geq \alpha\}$ and for $\alpha = 0$ by the closure of the support $[u]_\alpha = cl\{x|x \in \mathbb{R}, \mu_u(x) > 0\}$. The *core* of u is the set of elements of $\mathbb{R}$ having membership grade one, i.e. $core(u) = \{x|x \in \mathbb{R},\ \mu_u(x) = 1\}$ and we say that u is normal if $core(u) \neq \emptyset$. It is well-known that the $level - cuts$ are "nested", i.e. $[u]_\alpha \subseteq [u]_\beta$ for $\alpha > \beta$.

A particular class of fuzzy sets $u \in \mathcal{F}(\mathbb{R})$ is when the support is a convex and closed set and the membership function is quasi-concave, i.e. $supp(u)$ is convex and $\mu_u((1-t)x' + tx'') \geq \min\{\mu_u(x'), \mu_u(x'')\}$ for every $x', x'' \in supp(u)$ and $t \in [0,1]$. Equivalently, μ_u is quasi-concave if the level sets $[u]_\alpha$ are convex for all $\alpha \in [0,1]$. Finally, if the membership function is upper semi-continuous, then the level-cuts are closed. A fuzzy set u is a fuzzy quantity if the $\alpha - cuts$ are nonempty, compact intervals of the form $[u]_\alpha = [u_\alpha^-, u_\alpha^+] \subset \mathbb{R}$. If there exist $\widehat{u}^-, \widehat{u}^+ \in \mathbb{R}$ such that $core(u) = [\widehat{u}^-, \widehat{u}^+]$, u is called a fuzzy interval and if $\widehat{u}^- = \widehat{u}^+ = \widehat{u}$ it is a fuzzy number.

We denote by $\mathbb{F}$ the set of fuzzy quantities. The "nested" property is the basis for the LU representation (L for lower, U for upper).

Definition 1 *An LU-fuzzy quantity (number or interval) u is completely determined by any pair $u = (u^-, u^+)$ of functions $u^-, u^+ : [0,1] \longrightarrow \mathbb{R}$, defining the end-points of the $\alpha-cuts$, satisfying the three conditions:(i) $u^- : \alpha \longrightarrow u_\alpha^- \in \mathbb{R}$ is a bounded monotonic nondecreasing left-continuous function $\forall \alpha \in]0,1]$ and right-continuous for $\alpha = 0$;(ii) $u^+ : \alpha \longrightarrow u_\alpha^+ \in \mathbb{R}$ is a bounded monotonic nonincreasing left-continuous function $\forall \alpha \in]0,1]$ and right-continuous for $\alpha = 0$;(iii) $u_1^- \leq u_1^+$.*

We refer to the functions $u_{(.)}^-$ and $u_{(.)}^+$ as the lower and upper branches on u, respectively. If the two branches $u_{(.)}^-$ and $u_{(.)}^+$ are continuous invertible functions then $\mu_u(.)$ is formed by two continuous branches, the left being the increasing inverse of $u_{(.)}^-$ on $[u_0^-, u_1^-]$ and, the right, the decreasing inverse of $u_{(.)}^+$ on $[u_1^+, u_0^+]$. There are many choices for $u_{(.)}^-$ and $u_{(.)}^+$. Define a shape function as an increasing $s : [0,1] \to [0,1]$ such that $s(0) = 0$ and $s(1) = 1$. If we start with two shape functions $p(.)$ and $q(.)$ and with four numbers $u_0^- \leq u_1^- \leq u_1^+ \leq u_0^+$ defining the support and the core of u then we can model $u_{(.)}^-$ and $u_{(.)}^+$ by $u_\alpha^- = u_0^- + (u_1^- - u_0^-)p(\alpha)$ and $u_\alpha^+ = u_0^+ + (u_1^+ - u_0^+)q(\alpha)$ for all $\alpha \in [0,1]$.

The simplest fuzzy quantities have linear branches: a trapezoidal fuzzy interval, denoted by $u = \langle a, b, c, d \rangle$, where $a \leq b \leq c \leq d$, has $[u]_\alpha = [a + \alpha(b-a), d - \alpha(d-c)]$, $\alpha \in [0,1]$, obtaining a triangular fuzzy number if $b = c$.

Consider now the extension of function $f : \mathbb{R}^n \to \mathbb{R}$ to a vector of n fuzzy numbers $u = (u_1, u_2, ..., u_n) \in (\mathbb{F})^n$, with $k-th$ component $u_k \in \mathbb{F}$ given by $[u_k]_\alpha = [u_{k,\alpha}^-, u_{k,\alpha}^+]$ for $k = 1, 2, ..., n$ or $\mu_{u_k} : supp(u_k) \longrightarrow [0,1]$ for $k = 1, 2, ..., n$ and denote $v = f(u_1, u_2, ..., u_n)$.

The extension principle (Zadeh in [19]) is the basic tool for fuzzy calculus; it states that μ_v is given by

$$\mu_v(y) = \begin{cases} \sup\{\min\{\mu_{u_1}(x_1), ..., \mu_{u_n}(x_x)\}| \\ y = f(x_1, ..., x_n)\} \end{cases} \quad (1)$$

where $\mu_v(y) = 0$ if $y \notin Range(f) = \{y \in \mathbb{R} | \exists (x_1, ..., x_n) \in \mathbb{R}^n \text{ s.t. } y = f(x_1, ..., x_n)\}$.

For a continuous function $f : \mathbb{R}^n \to \mathbb{R}$, the $\alpha - cuts$ of the fuzzy extension v are obtained by solving the following box-constrained global optimization problems ($\alpha \in [0,1]$)

$$v_\alpha^- = \min\{f(x_1, x_2, ..., x_n) | x_k \in [u_k]_\alpha,\ k = 1, 2, ..., n\} \quad (2)$$

$$v_\alpha^+ = \max\{f(x_1, x_2, ..., x_n) | x_k \in [u_k]_\alpha,\ k = 1, 2, ..., n\}. \quad (3)$$

If the function $f(x_1, x_2, ..., x_n)$ is sufficiently simple, the analytical expressions for v_α^- and v_α^+ can be obtained, as it is the case for many unidimensional elementary functions.

For general functions, we need to solve numerically the global optimization problems (2) and (3) above; general methods have been proposed and a very extended scientific literature is available. It is clear that, except for simple cases, we have only the possibility of fixing a finite set of values $\alpha \in \{\alpha_0, ..., \alpha_M\}$ and obtain the corresponding v_α^- and v_α^+ pointwise; a sufficiently precise calculation requires M in the range from 10 to 100 or more (depending on the application and the required precision) and the computational time may become very large. To reduce these difficulties, various specific heuristic methods have been proposed; among others, the vertex method and its variants (see [5], [1] and [12]), the transformation method (see [8], [9], [10]), the interval arithmetic optimization with sparse grids (see [11]).

All the specific methods try to take computational advantage from the specific structure of "nested" optimizations (2)-(3) intrinsic in the properties of the $\alpha - cuts$. We will see that, at least in the differentiable case, the advantages of the LU representation appear to be quite interesting, based on the fact that a small number of α points is in general sufficient to

obtain good approximations (this is the essential gain in using the slopes to model fuzzy numbers), so reducing the number of constrained *min* (2) and *max* (3) problems to be solved directly. On the other hand, finding computationally efficient extension solvers is still an open research field in fuzzy calculations.

2.1 LU-fuzzy numbers

In this section we present the basic elements of the LU parametric representation of the shape functions proposed in [7] and [14] based on monotonic Hermite-type interpolation. A "standardized" model is given by any *differentiable monotonic* shape function $p : [0,1] \longrightarrow [0,1]$ such that $p(0) = 0$ and $p(1) = 1$ with $p(t)$ increasing on $[0,1]$.

Valid shape functions can be obtained by p satisfying the four Hermite interpolation conditions $p(0) = 0$, $p(1) = 1$ and $p'(0) = \beta_0$, $p'(1) = \beta_1$ for any value of the two nonnegative parameters $\beta_i \geq 0$, $i = 0, 1$.

To explicit the parameters, we denote the interpolating function by $t \longrightarrow p(t; \beta_0, \beta_1)$ for $t \in [0,1]$. We recall here two of the basic forms illustrated in [14]:

∘ (2,2)-rational spline:

$p(t; \beta_0, \beta_1) = \frac{t^2 + \beta_0 t(1-t)}{1 + (\beta_0 + \beta_1 - 2)t(1-t)}$;

∘ mixed exponential spline:

$p(t; \beta_0, \beta_1) = \frac{1}{a}[t^2(3-2t) + \beta_0 - \beta_0(1-t)^a + \beta_1 t^a]$

where $a = 1 + \beta_0 + \beta_1$.

Note that in both forms we obtain a linear $p(t) = t$, $\forall t \in [0,1]$ if $\beta_0 = \beta_1 = 1$ and a quadratic $p(t) = t^2 + \beta_0 t(1-t)$ if $\beta_0 + \beta_1 = 2$.

In order to produce different shapes we can either fix the slopes β_0 and β_1 (if we have information on the first derivatives at $t = 0$, $t = 1$) or we can estimate them by knowing values of $p(t)$ in additional points.

The standardized functions can be used to model "piecewise" the lower an upper branches of a fuzzy number on a decomposition of the interval $[0,1]$ into N subintervals $0 = \alpha_0 < \alpha_1 < ..... < \alpha_{i-1} < \alpha_i < ... < \alpha_N = 1$. It is convenient to use the same subdivision for both the lower u_α^- and upper u_α^+ branches (we can always reduce to this situation by the union of two different subdivisions). In each subinterval $I_i = [\alpha_{i-1}, \alpha_i]$, the values and the slopes of the two functions are $u^-_{(\alpha_{i-1})} = u^-_{0,i}$, $u^+_{(\alpha_{i-1})} = u^+_{0,i}$, $u^-_{(\alpha_i)} = u^-_{1,i}$, $u^+_{(\alpha_i)} = u^+_{1,i}$, $u'^-_{(\alpha_{i-1})} = d^-_{0,i}$, $u'^+_{(\alpha_{i-1})} = d^+_{0,i}$, $u'^-_{(\alpha_i)} = d^-_{1,i}$, $u'^+_{(\alpha_i)} = d^+_{1,i}$ and by the transformation $t_\alpha = \frac{\alpha - \alpha_{i-1}}{\alpha_i - \alpha_{i-1}}$, $\alpha \in I_i$, each subinterval I_i is mapped into the standard $[0,1]$ interval to determine each piece independently.

Globally continuous or more regular $C^{(1)}$ fuzzy numbers can be obtained directly from the data (for example, $u^-_{1,i} = u^-_{0,i+1}$, $u^+_{1,i} = u^+_{0,i+1}$ for continuity and $d^-_{1,i} = d^-_{0,i+1}$, $d^+_{1,i} = d^+_{0,i+1}$ for differentiability at $\alpha = \alpha_i$). So, we obtain a fuzzy number by

$$\begin{aligned} u_\alpha^- &= u^-_{0,i} + (u^-_{1,i} - u^-_{0,i}) p_i^-(t_\alpha) \\ u_\alpha^+ &= u^+_{0,i} + (u^+_{1,i} - u^+_{0,i}) p_i^+(t_\alpha). \end{aligned} \tag{4}$$

for $\alpha \in [\alpha_{i-1}, \alpha_i]$ and $i = 1, 2., , , N$.

A first parametrization is obtained on the trivial decomposition of interval $[0,1]$, with $N = 1$ (without internal points) and $\alpha_0 = 0, \alpha_1 = 1$. In this simple case, u can be represented by a vector of 8 components (the slopes corresponding to u_i^- are denoted by δu_i^-, etc.)

$$u = (u_0^-, \delta u_0^-, u_0^+, \delta u_0^+; u_1^-, \delta u_1^-, u_1^+, \delta u_1^+) \tag{5}$$

with u_0^-, δu_0^-, u_1^-, δu_1^- for the lower branch u_α^- and u_0^+, δu_0^+, u_1^+, δu_1^+ for the upper branch u_α^+. On a decomposition $0 = \alpha_0 < \alpha_1 < ... < \alpha_N = 1$ we proceed piecewise so that, for example, a differentiable shape function requires $4(N+1)$ parameters

$$\begin{aligned} u &= (\alpha_i; u_i^-, \delta u_i^-, u_i^+, \delta u_i^+)_{i=0,1,...,N} \text{ with} \\ u_0^- &\leq u_1^- \leq ... \leq u_N^- \leq u_N^+ \leq u_{N-1}^+ \leq ... \leq u_0^+ \\ \delta u_i^- &\geq 0, \delta u_i^+ \leq 0; \end{aligned} \tag{6}$$

the branches are computed according to (4) (see [7] and [14] for details).

3 Differential Evolution algorithms for fuzzy arithmetic

We adopt an algorithmic approach to describe the application of differential evolution methods to calculate the fuzzy extension of multivariable function, associated to the LU representation.

Let $v = f(u_1, u_2, ..., u_n)$ denote the fuzzy extension of a continuous function f in n variables; it is well known that the fuzzy extension of f to normal upper semicontinuous fuzzy intervals (with compact support) has the level-cutting commutative property (see [6]), i.e. the $\alpha - cuts$ $v_\alpha = [v_\alpha^-, v_\alpha^+]$ of v are the images of the $\alpha - cuts$ of $(u_1, u_2, ..., u_n)$ and are obtained by solving the box-constrained optimization problems

$$k = 1, ..., n \begin{cases} v_\alpha^- = \min \left\{ f(x_1, ..., x_n) \mid x_k \in [u^-_{k,\alpha}, u^+_{k,\alpha}], \right\} \\ v_\alpha^+ = \max \left\{ f(x_1, ..., x_n) \mid x_k \in [u^-_{k,\alpha}, u^+_{k,\alpha}], \right\}. \end{cases} \tag{7}$$

We will consider differentiable functions f and, for simplicity, we will illustrate the case of differentiable representations (6).

If $u_k = (u^-_{k,i}, \delta u^-_{k,i}, u^+_{k,i}, \delta u^+_{k,i})_{i=0,1,...,N}$ are the LU-fuzzy representations of the n input quantities and $v = (v_i^-, \delta v_i^-, v_i^+, \delta v_i^+)_{i=0,1,...,N}$, the $\alpha - cuts$ of v are obtained by solving (7).

For each $\alpha = \alpha_i$, $i = 0, 1, ..., N$ the min$\{\}$ and the max$\{\}$ can occur either at a point whose components $x_{k,i}$ are internal to the corresponding intervals $[u^-_{k,i}, u^+_{k,i}]$ or are coincident with one of the extremal values; denote by $\widehat{x}_i^- = (\widehat{x}^-_{1,i}, ..., \widehat{x}^-_{n,i})$ and $\widehat{x}_i^+ = (\widehat{x}^+_{1,i}, ..., \widehat{x}^+_{n,i})$ the points where the min and the max take place; then $v_i^- = f(\widehat{x}^-_{1,i}, ..., \widehat{x}^-_{n,i})$ and $v_i^+ = f(\widehat{x}^+_{1,i}, ..., \widehat{x}^+_{n,i})$ and the slopes δv_i^-, δv_i^+ are computed (as f is differentiable) by

$$\delta v_i^- = \sum_{\substack{k=1 \\ \widehat{x}^-_{k,i}=u^-_{k,i}}}^{n} \frac{\partial f(\widehat{x}^-_{1,i}, ..., \widehat{x}^-_{n,i})}{\partial x_k} \delta u^-_{k,i} + \sum_{\substack{k=1 \\ \widehat{x}^-_{k,i}=u^+_{k,i}}}^{n} \frac{\partial f(\widehat{x}^-_{1,i}, ..., \widehat{x}^-_{n,i})}{\partial x_k} \delta u^+_{k,i} \quad (8)$$

$$\delta v_i^+ = \sum_{\substack{k=1 \\ \widehat{x}^+_{k,i}=u^-_{k,i}}}^{n} \frac{\partial f(\widehat{x}^+_{1,i}, ..., \widehat{x}^+_{n,i})}{\partial x_k} \delta u^-_{k,i} + \sum_{\substack{k=1 \\ \widehat{x}^+_{k,i}=u^+_{k,i}}}^{n} \frac{\partial f(\widehat{x}^+_{1,i}, ..., \widehat{x}^+_{n,i})}{\partial x_k} \delta u^+_{k,i}. \quad (9)$$

If, for some reasons, the partial derivatives of f at the solution points are not available we can produce an estimation of the shapes δv_i^- and δv_i^+.

The idea of *DE* to find *min* or *max* of $\{f(x_1, ..., x_n) | (x_1, ..., x_n) \in \mathbb{A} \subset \mathbb{R}^n\}$ is simple: start with an initial "population" $x^{(1)} = (x_1, ..., x_n)^{(1)}, ..., x^{(p)} = (x_1, ..., x_n)^{(p)} \in \mathbb{A}$ of p feasible points for each generation (i.e. for each iteration) to obtain a new set of points by recombining randomly the individuals of the current population and by selecting the best generated elements to continue in the next generation. The initial population is chosen randomly and should try to cover uniformly the entire parameter space.

Denote by $x^{(k,g)}$ the $k-$th vector of the population at iteration (generation) g and by $x_j^{(k,g)}$ its $j-$th component ($j = 1, ..., n$).

At each iteration, the method generates a set of candidate points $y^{(k,g)}$ to substitute the elements $x^{(k,g)}$ of the current population, if $y^{(k,g)}$ is better.

To generate $y^{(k,g)}$ two operations are applied: recombination and crossover.

A typical recombination operates on a single component $j \in \{1, ..., n\}$ and generates a new perturbed vector of the form $v_j^{(k,g)} = x_j^{(r,g)} + \gamma[x_j^{(s,g)} - x_j^{(t,g)}]$, where $r, s, t \in \{1, 2, ..., p\}$ are chosen randomly and $\gamma \in]0, 2]$ is a constant (eventually chosen randomly for the current iteration) that controls the amplification of the variation.

The potential diversity of the population is controlled by a crossover operator, that construct the candidate $y^{(k,g)}$ by crossing randomly the components of the perturbed vector $v_j^{(k,g)}$ and the old vector $x_j^{(k,g)}$:

$$y_j^{(k,g)} = \begin{cases} v_j^{(k,g)} & \text{if } j \in \{j_1, j_2, ..., j_h\} \\ x_j^{(k,g)} & \text{if } j \notin \{j_1, j_2, ..., j_h\} \end{cases}$$

with h random between 1 and n and 0 with a crossover probability q and $j_1, j_2, ..., j_h$ being random components if h is not 0.

So, the components of each individual of the current population are modified to $y_j^{(k,g)}$ by a given probability q.

Typical values are $\gamma \in [0.2, 0.95]$, $q \in [0.7, 1.0]$ and $p \geq 5n$ (the higher p, the lower γ).

The candidate $y^{(k,g)}$ is then compared to the existing $x^{(k,g)}$ by evaluating the objective function at $y^{(k,g)}$: if $f(y^{(k,g)})$ is better than $f(x^{(k,g)})$ then $y^{(k,g)}$ substitutes $x^{(k,g)}$ in the new generation $g + 1$, otherwise $x^{(k,g)}$ is retained.

Many variants of the recombination schemes have been proposed and some seem to be more effective than others (see [16] and [17], [2] for constraints handling).

Examples are:

DE/rand/1: $v_j^{(k,g)} = x_j^{(r,g)} + \gamma[x_j^{(s,g)} - x_j^{(t,g)}]$;

DE/best/1: $v_j^{(k,g)} = x_j^{(best,g)} + \gamma[x_j^{(s,g)} - x_j^{(t,g)}]$ where $x^{(best,g)}$ is the current best solution;

DE/rand-best/2: $v_j^{(k,g)} = x_j^{(k,g)} + \gamma[x_j^{(best,g)} - x_j^{(k,g)} + x_j^{(s,g)} - x_j^{(t,g)}]$;

DE/best/2: $v_j^{(k,g)} = x_j^{(best,g)} + \gamma[x_j^{(s_1,g)} + x_j^{(s_2,g)} - x_j^{(t_1,g)} - x_j^{(t_2,g)}]$, $s_1, t_1, s_2, t_2 \in \{1, 2, ..., p\}$ are random;

DE/rand/2: $v_j^{(k,g)} = x_j^{(r,g)} + \gamma[x_j^{(s_1,g)} + x_j^{(s_2,g)} - x_j^{(t_1,g)} - x_j^{(t_2,g)}]$.

To take into account the particular nature of our problem, we modify the basic procedure and examine two different strategies.

Let $[u^-_{k,i}, u^+_{k,i}]$, $k = 1, 2, ..., n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be given; we have to find v_i^- and v_i^+ according to (7)

for $i = 0, 1, ..., N$. The slope parameters δv_i^-, δv_i^+ are computed by (8)and (9).

The first strategy is implemented in algorithm 1. Function $ran(0,1)$ generates a random uniform number in [0,1].

SPDE (Single Population DE procedure): start with the $(\alpha = 1) - cut$ back to the $(\alpha = 0) - cut$ so that the optimal solutions at a given level can be inserted into the "starting" populations of lower levels; use two distinct populations and perform the recombinations such that, during generations, one of the populations specializes to find the minimum and the other to find the maximum.

Algorithm 1: (Frame of SPDE).

Choose $p \approx 10n$, $g_{\max} \approx 500$, q and γ.

Select $(x_1^{(l)}, ..., x_n^{(l)})$, $x_k^{(l)} \in [u_{k,N}^-, u_{k,N}^+]$

$\forall k,\ l = 1, ..., 2p$ evaluate $y^{(l)} = f(x_1^{(l)}, ..., x_n^{(l)})$
for $i = N, N-1, ..., 0$
  for $g = 1, 2, ..., g_{\max}$
  (up to $g_{\max}$ generations or other stopping rule)
    for $l = 1, 2, ..., 2p$
      select (randomly) $r, s, t \in \{1, 2, ..., 2p\}$
      and $j^* \in \{1, 2, ..., n\}$
      for $j = 1, 2, ..., n$
        if ($j = j^*$ or $ran(0,1) < q$)
          then $x'_j = x_j^{(r)} + \gamma[x_j^{(s)} - x_j^{(t)}]$
          else $x'_j = x_j^{(l)}$
        ensure that $u_{j,i}^- \le x'_j \le u_{j,i}^+$
      end
      evaluate $y = f(x'_1, ..., x'_n)$
      if $l \le p$ and $y < y^{(l)}$ then
        substitute $(x_1, ..., x_n)^{(l)}$ with $(x'_1, ..., x'_n)$
      if $l > p$ and $y > y^{(l)}$ then
        substitute $(x_1, ..., x_n)^{(l)}$ with $(x'_1, ..., x'_n)$
    end
  end
  $v_i^- = y^{(l^*)} = \min\{y^{(l)} | l = 1, 2, ..., p\}$
  $(\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-) = (x_1, ..., x_n)^{(l^*)}$
  $v_i^+ = y^{(l^{**})} = \max\{y^{(p+l)} | l = 1, 2, ..., p\}$
  $(\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+) = (x_1, ..., x_n)^{(l^{**})}$
  if $i < N$
    select $(x_1^{(l)}, ..., x_n^{(l)})$, $x_k^{(l)} \in [u_{k,i-1}^-, u_{k,i-1}^+]$
    $\forall k,\ l = 1, ..., 2p$
    including $(\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-)$ and $(\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+)$
  endif
end

The second strategy is implemented in algorithm 2.

MPDE (Multi Populations DE procedure): use $2(N+1)$ populations to solve simultaneously all the box-constrained problems (7); $N+1$ populations specialize for the min and the others for the max and the current best solution for level α_i is valid also for levels $\alpha_0, ..., \alpha_{i-1}$.

Algorithm 2: (Frame of MPDE).

Choose $p \approx 10n$, $g_{\max} \approx 500$, q and γ.

Select $(x_1^{(l,i)}, ..., x_n^{(l,i)})$, $x_k^{(l,i)} \in [u_{k,i}^-, u_{k,i}^+]$

$\forall k,\ l = 1, ..., 2p,\ i = 0, 1, ..., N$

let $y^{(l,i)} = f(x_1^{(l,i)}, ..., x_n^{(l,i)})$
let $v_i^- = \min\left\{y^{(l,j)} | j = 0, ..., i, \forall l\right\}$
let $v_i^+ = \max\left\{y^{(l,j)} | j = 0, ..., i, \forall l\right\}$
let $\widehat{x}_i^-, \widehat{x}_i^+ \in \mathbb{R}^n$ the points where v_i^-, v_i^+ are taken
for $g = 1, 2, ..., g_{\max}$
  (up to $g_{\max}$ generations or other stopping rule)
  for $i = N, N-1, ..., 0$
    for $l = 1, 2, ..., p$
      select (randomly) $r, s, t \in \{1, 2, ..., p\}$
      and $k^* \in \{1, 2, ..., n\}$
      for $k = 1, 2, ..., n$
        if ($k = k^*$ or $ran(0,1) < q$) then
          $x'_k = x_k^{(r,i)} + \gamma[x_k^{(s,i)} - x_k^{(t,i)}]$
          $x''_k = x_k^{(p+r,i)} + \gamma[x_k^{(p+s,i)} - x_k^{(p+t,i)}]$
          ensure $u_{k,i}^- \le x'_k,\ x''_k \le u_{k,i}^+$
        else
          $x'_k = x_k^{(l,i)},\ \ x''_k = x_k^{(p+l,i)}$
        endif
      end
      let $y' = f(x'_1, ..., x'_n)$ and $y'' = f(x''_1, ..., x''_n)$;
      if $y' < y^{(l,i)}$ (population for min)
        substitute $(x_1, ..., x_n)^{(l,i)}$ with $(x'_1, ..., x'_n)$
      if $y'' > y^{(p+l,i)}$ (population for max)
        substitute $(x_1, ..., x_n)^{(p+l,i)}$ with $(x''_1, ..., x''_n)$
      if y' or y'' are better
        update values $\{v_j^-, v_j^+, \widehat{x}_j^-, \widehat{x}_j^+ | j = 0, ..., i\}$
      endif
    end
  end
end

An efficient constraint handling method has been examined in [2], where a tournament selection is proposed, based on three criteria:

(1) any feasible solution is preferred to any infeasible solution;

(2) among two feasible solutions, the one with better objective value is preferred;

(3) among two infeasible solutions, the one having smaller constraints violation is preferred.

In our case, as we have simple box-constraints, it is easy to produce feasible starting populations, as we have to generate random numbers $x_j^{(k,0)}$ between the lower $u_{j,i}^-$ and the upper $u_{j,i}^+$ values.

During the iterations, we use a variant of the method

above, where the $y^{(k,g)}$ are progressively forced to be feasible or with small infeasibilities and a penalty is assigned to infeasible values:

(i) modify $y_j^{(k,g)}$ to fit $[u_{j,i}^- - \frac{\varepsilon}{g^2}, u_{j,i}^+ + \frac{\varepsilon}{g^2}]$, $j = 1, 2, ..., n$ with small $\varepsilon \sim 10^{-2}(u_{j,i}^+ - u_{j,i}^-)$, so that the eventual infeasibilities decrease rapidly during the generation process;

(ii) if the candidate point $y^{(k,g)}$ is infeasible and has a value $f(y^{(k,g)})$ better than the current best feasible value $f(x^{(best,g)})$ then a penalty is added and the value of $y^{(k,g)}$ is elevated to $f(x^{(best,g)}) + \varepsilon'$ (for the min problems) or reduced to $f(x^{(best,g)}) - \varepsilon'$ (for the max problem), being $\varepsilon' \sim 10^{-3}$ a small positive number.

To decide that a solution is found, we use the following simple rule: choose a fixed tolerance $tol \sim 10^{-3}, 10^{-4}$ and a number $\widehat{g} \sim 20, 30$ of generations; if for $\widehat{g}$ subsequent iterations all the values v_i^- and v_i^+ are changed less than tol, then the procedure stops and the found solution is assumed to be optimal. In any case, no more than 500 iterations are performed (but this limit was never reached during the computations).

4 Computational results

The two algorithms SPDE and MPDE have been implemented using C++ and Matlab and executed on a set of test functions with different dimension $n = 2, 4, 8, 16, 32$.The first table contains a set of ten 2-dimensional problems. Tests of higher dimensions (table 2) use the following functions:

1. Rastrigin:

$$f(x_1, ..., x_n) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10\right],$$

$x_i \in [0, 3]^n$;

2. Ackley:

$$f = 20 + e - 20\exp\left[-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right]$$

$$-\exp\left[\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right]$$

$x_i \in [-1, 3]^n$;

3. Modified Rosenbrock:

$$f(x_1, ..., x_n) = \sum_{i=1}^{n-1} \left[10(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$$

$x_i \in [-0.2, 0.2]^n$.

Table 1: 2-dim problems

#	$f(x_1, x_2)$ $supp(u_1) \times supp(u_2)$
1	$x_2 \cos(\pi x_1)$ $[0, 5] \times [1, 5]$
2	$\sqrt{(x_1 - 0.1)^4 + (x_2 - 0.1)^4}$ $[-1, 1] \times [0, 2]$
3	$\frac{1}{0.2+(x_1-2)^4+(x_2-2)^2}$ $[0, 5] \times [1, 5]$
4	$1 + \frac{1}{2}x_1 + \sin(2x_1 - \pi/2) + 2\cos(x_2)$ $[-2, 2]^2$
5	$(x_1^2 - x_2)^2 + 0.01(1 - x_1)^2$ $[-2, 2]^2$
6	$(1 - \sqrt{x_1^2 + x_2^2})\sin(\pi(x_1 + \frac{1}{2}))$ $[-1, 1] \times [-2, 2]$
7	$20\cos(x_1 + x_2) - x_1^2 - x_2^2$ $[-4, 4]^2$
8	$\cos(2x_1 + \sin(x_2))$ $+ \cos(x_2) - 0.1(x_1^2 + x_2^2)$ $[-4, 4]^2$
9	$20 + e - 20e^{-\frac{1}{20}(x_1^2+x_2^2)}$ $-e^{\frac{1}{4}\cos(2\pi x_1)+\cos(2\pi x_2)}$ $[-1, 3]^2$
10	$100(x_2 - x_1^2)^2 + (x_1 - 1)^2$ $[-0.2, 0.2]^2$

Table 2: Other $n \geq 4$ test problems

#	n	$f(x_1, ..., x_n)$
11	4	*Rastrigin* $n = 4$
12	4	*Ackley* $n = 4$
13	4	*Mod. Rosenbrock* $n = 4$
14	8	*Rastrigin* $n = 8$
15	8	*Ackley* $n = 8$
16	8	*Mod. Rosenbrock* $n = 8$
17	16	*Rastrigin* $n = 16$
18	16	*Ackley* $n = 16$
19	16	*Mod. Rosenbrock* $n = 16$
20	32	*Rastrigin* $n = 32$
21	32	*Ackley* $n = 32$
22	32	*Mod. Rosenbrock* $n = 32$

In the computations, the input fuzzy numbers are triangular and symmetric of the form $u_i = \langle a_i, \frac{a_i+b_i}{2}, b_i\rangle$ with support given by the interval $[a_i, b_i]$; the applied supports are illustrated in the tables.

If the extension algorithm is used in combinations with the LU-fuzzy representation for differentiable membership functions (and differentiable extended functions), then the number $N+1$ of $\alpha-cuts$ (and correspondingly of min/max optimizations) can be sufficiently small. Experiments in [7] and [14] motivated that $N = 10$ is in general quite sufficient to obtain good approxima-

tions.

Tables 3. and 4. report the number of function evaluations FE_{SPDE} and FE_{MPDE} needed to the two algorithms $SPDE$ and $MPDE$ to reach the solution of the nested min/max optimization problems corresponding to the eleven $\alpha-cuts$ of the uniform $\alpha-$decomposition $\alpha_i = \frac{i}{10}$, $i = 0, 1, ..., 10$ ($N = 10$ subintervals).

Table 3: Results for 2-dim problems

#	n	FE_{SPDE}	FE_{MPDE}
1	2	15400	8800
2	2	12760	6380
3	2	16280	6600
4	2	12760	6380
5	2	12320	5720
6	2	11000	5500
7	2	16720	7040
8	2	13200	8140
9	2	16280	6380
10	2	11440	6600

Table 4: Results for other test problems

#	n	FE_{SPDE}	FE_{MPDE}
11	4	40040	32560
12	4	27720	16280
13	4	19360	18920
14	8	72160	47520
15	8	51392	25344
16	8	20416	19712
17	16	292160	186560
18	16	255552	98560
19	16	59136	63360
20	32	1122176	560384
21	32	283008	252032
22	32	250624	243584

A difficult case is the Rastrigin function; for $n = 2$ it has the representation in Fig. 1.

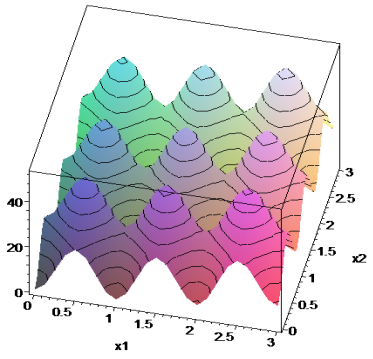


Figure 1: 2-dim Rastrigin function

Its extension to non symmetric triangular fuzzy numbers $u_i = \langle 0, 4, 5 \rangle$, for $i = 1, 2, ..., n = 8$ by MPDE required $FE = 103830$ function evaluations; the extended fuzzy number is represented if figure 2.

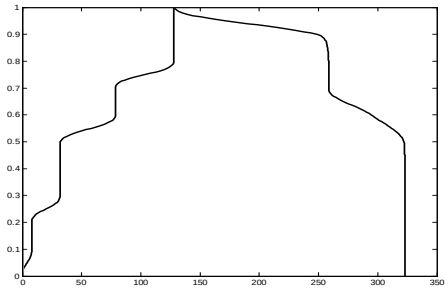


Figure 2. Extension of 8-dim Rastrigin function

Figure 3. represents the logarithm of the number of function evaluations vs the logarithm of the number n of arguments. It appears an almost linear relationship $\ln(FE_{SPDE}) = a + b\ln(n)$ and $\ln(FE_{MPDE}) = c + d\ln(n)$.

The estimated coefficients are $a = 8.615, b = 1.20$ and $c = 7.869, d = 1.34$. The computational complexity of the proposed algorithms (on average for the 22 test problems) grows less then quadratically with the dimension n (SPDE is less efficient but grows slowly than MPDE). This is an interesting result, as all the existing methods for the fuzzy extension of functions are essentially exponential in n.

Extended results and complete descriptions of the algorithms are discussed in [15]. The C++ source codes are available on request to the author; also a MatLab implementation is available. Possible further research is related to experimentation of different constraints handling strategies, different recombination operators, different selections for the parameters of the DE procedure (e.g. varying and situation-sensitive parameters) or different ways of cooperation between populations.

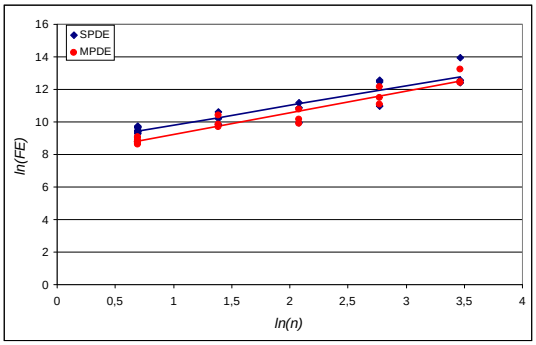


Figure 3. $(\ln(n), \ln(FE))$ for 22 test problems

References

[1] H. K. Chen, W. K. Hsu, W. L. Chiang, A comparison of vertex method with JHE method, *Fuzzy Sets and Systems*, 95, 1998, 201-214.

[2] K. Deb, An efficient constraint handling method for genetic algorithms, *Computer methods in applied mechanics and engineering*, 186, 2000, 311-338.

[3] P. Diamond, P. Klöden, *Metric Spaces of Fuzzy Sets*, World Scientific, Singapore, 1994.

[4] P. Diamond, P. Klöden, Metric topology of fuzzy numbers and fuzzy analysis, in [6], 2000, 583-641.

[5] W. M. Dong, H. C. Shah, Vertex method for computing functions of fuzzy variables, *Fuzzy Sets and Systems*, 24, 1987, 65-78.

[6] D. Dubois, H. Prade (Ed), *Fundamentals of Fuzzy Sets*, Kluwer, Boston, The Handbooks of Fuzzy Sets Series, 2000.

[7] M. L. Guerra, L. Stefanini, Approximate Fuzzy Arithmetic Operations Using Monotonic Interpolations, *Fuzzy Sets and Systems*, 150, 2005, 5-33.

[8] M. Hanss, The transformation method for the simulation and analysis of systems with uncertain parameters, *Fuzzy Sets and Systems*, 130, 2002, 277-289.

[9] M. Hanss, A. Klimke, On the reliability of the influence measure in the transformation method of fuzzy arithmetic, *Fuzzy Sets and Systems*, 143, 2004, 371-390.

[10] A. Klimke, An efficient implementation of the transformation method of fuzzy arithmetic, Extended Preprint Report, 2003/009, Institute of Applied Analysis and Numerical Simulation, University of Stuttgard, Germany, 2003.

[11] A. Klimke, B. Wohlmuth, Computing expensive multivariate functions of fuzzy numbers using sparse grids, *Fuzzy Sets and Systems*, 153, 2005, 432-453.

[12] E. N. Otto, A. D. Lewis, E. K. Antonsson, Approximating $\alpha - cuts$ with the vertex method, *Fuzzy Sets and Systems*, 55, 1993, 43-50.

[13] K. Price, An introduction to differential evolution, in D. Corne, M. Dorigo, F. Glover (Ed.), *New Ideas in Optimization*, McGraw Hill, 1999, 79-108.

[14] L. Stefanini, L. Sorini, M. L. Guerra, Parametric representation of fuzzy numbers and application to fuzzy calculus, *Fuzzy Sets and Systems*, 157, 2006, 2423 – 2455.

[15] L. Stefanini, Differential Evolution Methods for the Fuzzy Extension of Functions, Working Paper Series EMS n. 103, University of Urbino, 2006; available online at the *RePEc* web page, *http://ideas.repec.org/f/pst233.html*.

[16] R. Storn, K. Price, Differential Evolution: a simple and efficient heuristic for global optimization over continuous spaces, ICSI technical report TR-95-012, Berkeley University, 1995. Also, *Journal of Global Optimization*, 11, 1997, 341-359.

[17] R. Storn, System design by constraint adaptation and differential evolution, *IEEE Transactions on Evolutionary Computation*, 3, 1999, 22-34.

[18] K. L. Wood, K. N. Otto, E. K. Antonsson, Engineering design calculations with fuzzy parameters, *Fuzzy Sets and Systems*, 52, 1992, 1-20.

[19] L. A. Zadeh, Fuzzy Sets, *Information and Control*, 8, 1965, 338-353.

Selection of Optimal Set of Diagnostic Tests with Use of Evolutionary Approach in Intelligent Systems

A.E. Yankovskaya
Tomsk State University of Architecture and Building, Tomsk, Russia
yank@tsuab.ru

Y. R. Tsoy
Tomsk Polytechnic University, Tomsk, Russia
qai@mail.ru

Abstract

This paper concerns problem of selection of optimal subset of irredundant unconditional diagnostic tests by means of evolutionary approach. The method of correction of features' weight, cost and damage coefficients for the test patterns recognition is proposed. The suggestion is made that evolutionary programming approach would be more appropriate than genetic algorithm because of disadvantage of crossover use for multi-objective problems solution. Future research tasks to compare different evolutionary algorithms for solving problem at hand are outlined.

Keywords: optimal subset selection, evolutionary multi-objective optimization, diagnostic test, intelligent systems

1 Introduction

Selection of "good" irredundant unconditional diagnostic tests (IUDT) is of great importance for decision making in intelligent systems, since quality of obtained solutions depends significantly on properties of the used tests. However such a selection doesn't lead necessarily to an optimal solution because total number of features in selected tests set can be too large as well as time consumption and cost. Also one should take into consideration damage (risk), caused in result of features measuring for the object under investigation, for example, in geoecological or biomedicine problems.

This research continues our previous work on optimal subset of IUDTs selection. For the first time the optimization criteria and undoubtedly actual problem of optimal tests subset selection has been formulated in the paper [1]. In the paper [2] logical-combinatorial algorithm for optimal IUDTs subset selection was presented. In the paper [3] optimization criteria were further elaborated and three algorithms providing satisfaction of those criteria were proposed: logical-combinatorial with sequent satisfaction of the prescribed criteria, algorithm of optimal tests set selection on the base of hierarchies analysis method, and genetic algorithm (GA).

In this paper we propose correction of weights, costs and damages for characteristic (descriptive) features from the informative viewpoint. Analog of informative weight of features entering all IUDTs was used in Y.I. Zhuravlev's researches [4]. This correction is necessary for more comprehensive estimation of each test to improve recognition accuracy with use of as separate IUDT so and their set (subset).

For solution of the optimal IUDT subset selection problem we will use evolutionary algorithm (EA) which presents heuristic search concept similar to trial-and-error method.

During last decade a number of models of GAs were developed, such as NSGA-II [5], PAES [6], SPEA2 [7], PPREA [8] to solve multi-objective optimization (MOO) problems. Also alternative approaches on a basis of particles swarm optimization [9] and differential evolution [10] were proposed. Some researches are aimed at reduction of the optimization criteria number (see for

example [11]) and this certainly appears to be promising for the optimization results, though search of competent universal method of reduction of criteria number is rather challenging (if possible at all) due to great variety of existing MOO problems.

One of the critical conditions for the success of EA in MOO problem is preserving as many predominated (incomparable) solutions in one population as possible. Such solutions correspond to different points on the Pareto front. An idea of grouping of individuals according to some similarity/difference measure emerges in various forms to reach this, for example, as niching, or as specific non-dominated selection [5]. Considering this condition the idea of GAs use to solve MOO problems looks rather contradictory, from the authors point of view, since the main searching operator in GA is crossover and use of this operator involves risk of recombination of incompliant values of the optimization parameters due to crossing of different parent individuals, though the last can be situated rather close to each other in parameters space (according to the metrics selected).

We are planning to examine this by investigation of MOO problem solution using evolutionary programming (EP) algorithm, which doesn't adopt crossing of individuals. The results of EP optimization will be compared with those ones of GA.

2 Basic Notions and Definitions

Let's introduce a number of definitions [1, 2, 12] and notations used later.

Test is a set of features distinguishing any pairs of objects belonging to different patterns.

Test is called *irredundant* if after the removal of any feature the test is not a test.

Let $\mathbf{T} = \{t_{ij} : i = 1,...,n, j = 1,...,m\}$ be the matrix of IUDTs and $\mathbf{T}_i$ corresponds to the i^{th} IUDT (the i^{th} row of matrix **T**). We denote set of characteristic features as $\mathbf{z} = \{z_j : j = 1,...,m\}$ and for each feature z_j we define its weight w_j [12], cost w'_j [2] and damage w''_j [13].

The case of binary matrix **T** is considered therefore weight of the i^{th} IUDT is $W_i = \sum_j w_j t_{ij}$. Then average test weight along all tests in IUDT matrix equals to: $\overline{W} = \frac{\sum_i W_i}{n}$. Number h_i of features in each test is given by $h_i = \sum_j t_{ij}$ and average number of features along all tests in **T** is $\overline{h} = \frac{\sum_i h_i}{n}$.

3 Modification of Weight, Cost and Damage Values

Here we will use terms weight, cost and damage of features instead of the terms weight and cost coefficients and coefficient of damage.

Consider modification of weight for the j^{th} feature z_j. Let $\mathbf{T}(z_j)$ be the subset of IUDT possessing j^{th} feature and $n_j = |\mathbf{T}(z_j)|$ is its power. The average weight $\overline{W}_j$ of IUDTs in $\mathbf{T}(z_j)$ is given by

$$\overline{W}_j = \frac{\sum_{\mathbf{T}_i \in \mathbf{T}(z_j)} W_i}{n_j},$$

where $\sum_{\mathbf{T}_i \in \mathbf{T}(z_j)}$ denotes sum along all tests in $\mathbf{T}(z_j)$. Average number $\overline{h}_j$ of features in IUDTs in $\mathbf{T}(z_j)$ equals to:

$$\overline{h}_j = \frac{\sum_{\mathbf{T}_i \in \mathbf{T}(z_j)} h_i}{n_j}.$$

We introduce correction factor $k_j = \frac{\overline{W}_j}{\overline{W}} \frac{\overline{h}}{\overline{h}_j}$ to define modified weight of the j^{th} feature:

$$\hat{w}_j = k_j w_j.$$

Then the modified value of weight is used together with original weight value for recalculation of the tests weight. Choice between w_j and $\hat{w}_j$ values is managed according to selection probabilities as follows:

$$p(\hat{w}_j) = m(\hat{w}_j), \quad p(w_j) = m(w_j),$$

where $m(\cdot)$ is membership function:

$$m(\hat{w}_j) = \frac{n_j}{n}, \qquad m(w_j) = 1 - \frac{n_j}{n}.$$

The same way is used to calculate modified values of cost $\hat{w}'_j$ and damage $\hat{w}''_j$ of features and also to make a probabilistic choice between original and modified values of weight, cost and of features.

4 Setting of a Problem

For the given tests matrix **T** with defined values of features attributes it is necessary to find such submatrix $\mathbf{T}_0$ of height n_0, which corresponds to the set $\mathbf{N}^0$ of tests that would provide satisfaction of the following criteria (in order of significance descend):

1. $\mathbf{N}^0$ should contain as many pseudo-obligatory features as possible.

2. $\mathbf{N}^0$ should contain in total as little number features as possible.

3. $\mathbf{N}^0$ should have maximum possible total weight.

4. $\mathbf{N}^0$ should have minimum possible total cost.

5. $\mathbf{N}^0$ should have minimum possible total damage.

Statement of this problem accounting 5 optimization criteria was firstly introduced in the paper [13]. Since solving of the problem at hand is considered with use of evolutionary algorithm, which is known to be a heuristic search method, then as a consequence there is no guarantee that the optimal submatrix $\mathbf{T}_0$ (subset of IUDTs) will be found. In other words obtained solution is most likely suboptimal.

5 Genetic encoding and objective function

Candidate-solutions are encoded in binary chromosomes (strings) of length l, where each i^{th} symbol denotes inclusion ("1") or exclusion ("0") of the i^{th} IUDT in/from the resulting set of tests. Initial population of solutions is generated at random.

We will calculate fitness of the individual with chromosome h by evaluation of quality of corresponding submatrix $\mathbf{T}(h)$ as follows [13]:

$$f_h = \sum_{k=1}^{5} v_k e_h^{(k)} + 100\left(U(h) - n_0\right)^2,$$

where v_k is a weight coefficient for the k^{th} optimization criterion which corresponds to its significance; $U(\boldsymbol{\psi})$ gives number of units in binary string $\boldsymbol{\psi}$; $e_h^{(k)}$ is a penalty function for violation of the k^{th} criterion:

$$e_h^{(1)} = \frac{m - U_c\left(\mathbf{T}_0(h)\right)}{m}, \quad e_h^{(2)} = \frac{U_d\left(\mathbf{T}_0(h)\right)}{m},$$

$$e_h^{(3)} = \frac{S_W(\mathbf{T}) - S_W(\mathbf{T}_0(h))}{S_W(\mathbf{T})}, \quad e_h^{(4)} = \frac{S_{W'}\left(\mathbf{T}_0(h)\right)}{S_{W'}(\mathbf{T})},$$

$$e_h^{(5)} = \frac{S_{W''}\left(\mathbf{T}_0(h)\right)}{S_{W''}(\mathbf{T})},$$

where $S_W(\boldsymbol{\Psi}), S_{W'}(\boldsymbol{\Psi})$ and $S_{W''}(\boldsymbol{\Psi})$ – total weight, cost and damage correspondingly along all tests of the set of IUDTs corresponding to matrix $\boldsymbol{\Psi}$; $U_c(\boldsymbol{\Psi}) = U\left(\bigwedge_i \boldsymbol{\psi}_i\right)$ and $U_d(\boldsymbol{\Psi}) = U\left(\bigvee_i \boldsymbol{\psi}_i\right)$ – correspondingly number of units in conjunction and disjunction along all rows of binary matrix $\boldsymbol{\Psi}$. Evolutionary search is aimed at minimization of f.

In order to respect priorities of criteria formulated in Section 4 we will reduce weights of penalties with growth of penalty number k. Then the following penalties weights will be used: $v_1 = 40$, $v_2 = 30$, $v_3 = 15$, $v_4 = 10$, $v_5 = 5$. Note that penalties weights depend on the specific application.

6 Research Description

We consider following variants of objective function depending on whether original or modified values of weights, costs and damages of features are used:

1. $f = f(w, w', w'')$ – using only original values for weights, costs and damages.

2. $f_{fuz} = f(w, w', w'', \hat{w}, \hat{w}', \hat{w}'')$ – using probabilistic choice between original and modified values of weights, costs and damages of features.

3. $\hat{f} = f(\hat{w}, \hat{w}', \hat{w}'')$ – using only modified values for weights, costs and damages of features.

We are planning to make series of experiments to compare results of GA and EP corresponding to different variants of the objective function. Note that the second variant f_{fuz} of the objective function presents noisy function but as it was shown in the paper [14] use of population of candidate-solutions allows to compensate appearing uncertainty of solution quality evaluation.

Conclusion

The new method of modification of weights, costs and damages of features is presented in this paper. The method of probabilistic selection between original and modified weights, costs and damages of features attributes is proposed.

In result of critical analysis of application of GA for solution of MOO problems and suggested deficiencies involved by crossover operator, use of EP algorithm instead of GA is proposed.

Future work is connected with solution of the formulated problem of optimal IUDTs subset selection using both GA and EP for different variants of the objective function.

Implemented algorithms will be used in instrumental intelligent tool IMSLOG [15] for regularities revealing and decision making on the basis of test pattern recognition.

Acknowledgment

This research is funded by Russian Foundation for Basic Researches grant (projects no. 06-01-00144 and no. 07-01-00452) and by Russian Humanitarian Scientific Foundation (project no. 06-06-12603V).

References

[1] Yankovskaya A.E. Construction of logical tests of prescribed properties and logic-combinatorial pattern recognition on them // Abstracts of the Intellectualization of Information Processing (IIP2002). Simferopl, 2002. P. 100-102. (in Russian)

[2] Yankovskaya A.E., Mozheiko V.I. Optimization of a set of tests selection satisfying the criteria prescribed // 7th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-7-2004). Conference Proceedings. Vol. I. – St. Petersburg: SPbETU 2004. – Pp.145-148

[3] Kolesnikova S.I., Mozheiko V.I., Tsoy Y.R., Yankovskaya A.E. Algorithms of selection of optimal set of irredundant diagnostic tests in intelligent decision making systems // Proceedings of the First International conference "System analysis and information technologies" (SAIT-2005). Vol. 1. Moscow: KomKniga, 2005. P. 256-262.

[4] Zhuravlev Yu.I., Gurevitch I.B. Pattern Recognition and Image Analysis // Artificial Intelligence in 3 books. Book 2. Models and Methods: Reference Book / By ed. D.A. Pospelov. – M: Radio and Comm. – 1990. – P. 149-191. (in Russian)

[5] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2):182–197, 2002.

[6] Joshua D. Knowles and David W. Corne. Approximating the non-dominated front using the Pareto archived evolution strategy. Evolutionary Computation Journal, 8(2):149–172, 2000.

[7] Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the strength pareto evolutionary algorithm for multi-objective optimization. InK. C. Giannakoglou, D. T. Tsahalis, J. P'eriaux, K. D. Papailiou, and T. Fogarty, editors, Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems, pages 95–100, Athens, Greece, 2001. International Center for Numerical Methods in Engineering (Cmine).

[8] Hallam N., Graham K., Blanchfield P. Solving multi-objective optimization problems using the potential pareto regions evolutionary algorithm // T.P. Runarsson et al. (eds.): PPSN IX, LNCS 4193, pp. 503-512. Springer-Verlag, 2006.

[9] Alvarez-Benitez J.E., Everson R.M., Fieldsen J.E. A MOPSO algorithm based exclusively on pareto dominance concepts // In C. Coello-Coello et al. (eds.): Evolutionary Multi-Criterion Optimization, vol. 3410, pp. 459-473. Springer, 2005.

[10] Becerra R.L., Coello Coello C.A. Solving hard multi-objective optimization problems using *e*-constraint with cultured differential evolution // T.P. Runarsson et al. (eds.): PPSN IX, LNCS 4193, pp. 543-552. Springer-Verlag, 2006.

[11] Brockhoff D., Zitzler E. Are all objectives necessary? On dimensionality reduction in evolutionary multi-objective optimization // T.P. Runarsson et al. (eds.): PPSN IX, LNCS 4193, pp. 533-542. Springer-Verlag, 2006.

[12] Yankovskaya A.E. Design of Optimal Mixed Diagnostic Test With Reference to the Problems of Evolutionary Computation // Proceedings of the First International Conference on Evolutionary Computation and Its Applications (EVCA'96). – Moscow, 1996. – P. 292-297.

[13] Yankovskaya A. E, Tsoy Y.R. Optimization of a set of tests selection satisfying the criteria prescribed using compensatory genetic algorithm // Proceedings of IEEE East-West Design & Test Workshop (EWDTW'05) Odessa, Ukraine, September 15-19, 2005 – Kharkov: SPD FL Stepanov V.V. – P. 123-126.

[14] Arnold D.V., Beyer H.-G. Performance Analysis of Evolution Strategies with Multi-Recombination in High-Dimensional $\mathbf{R}^N$-Search Spaces Disturbed by Noise. Reihe CI 94/00, SFB 531, University of Dortmund, 2000. See also: http://sfbci.cs.uni-dortmund.de

[15] Yankovskaya A.E., Gedike A.I., Ametov R.V., Bleikher A.M. IMSLOG-2002 Software Tool for Supporting Information Technologies of Test Pattern Recognition// Pattern Recognition and Image Analysis. – 2003. – Vol. 13. – No. 4. – pp. 650-657.

Session 8

Fuzzy Numbers

Object Oriented Approach in Fuzzy Arithmetic Using Parametric Representation of Fuzzy Numbers

Frantisek Hunka
University of Ostrava, Institute of Fuzzy Applications
Computer Science Dept., Czech Republic
email: frantisek.hunka@osu.cz

Abstract

Parametric representation of fuzzy numbers (PFN) with derived methods for performing fuzzy arithmetic proves to have better accuracy and similar computational speed to using triangular fuzzy numbers (TFN). These fuzzy numbers (PFN) can be considered a special case of triangular fuzzy numbers (TFN) as they have three additional parameters and their methods for arithmetic operations are further specialized in regard to TFNs. In the paper object oriented approach and achieved results applied to the topic are presented. Generally object oriented approach provides higher abstraction, which makes the problem better apprehensible and further extensible in more specialized cases. The other benefit of this approach is reusability of created modules and rather simpler programming code formalization. This paper draws from rich conceptual means that provides object oriented software tools.

In the paper virtual mechanism, polymorphism and information hiding are the main features used for object oriented approach in this topic.

Keywords: object oriented approach, fuzzy arithmetic, parameterized fuzzy number

1 Introduction

One of the possibilities to model engineering design and manufacturing problems is the use of fuzzy numbers and fuzzy arithmetic. Direct implementation of extended arithmetic operators on fuzzy numbers is computationally complex and is similar to solving a nonlinear programming problem. For this reason either triangular fuzzy numbers (TFN) or trapezoidal fuzzy numbers (TrFN) are used. Unfortunately the TFN shape is not closed under multiplication and division because the result of these operators is polynomial membership function and triangular shape only approximates the actual result. To solve the problem a parametric representation of fuzzy numbers and their arithmetic operators was proposed and described [1,2]. Together with the proposal Giachetti and Young [1,2] also introduced a new approximation, which is consistent, computationally fast and accurate estimation for arithmetic operations.

The PFN contains further three parameters and its arithmetic operations can be regarded as further extended operators of the same applied on TFNs. The result of fuzzy multiplication and fuzzy division applied on the TFNs is nonlinear operations with a polynomial membership function.

Our aim is to show facilitating object oriented approach in design and implementation of the fuzzy arithmetic applied on fuzzy numbers expressed either in parametric or triangular form. Viewing the whole problem we may consider PFNs as an incremental extension of TFN. This view covers of course both attribute extension (specification) and arithmetic operation extension. As was showed in other publications e.g. [6, 7] object oriented approach can be used with benefit for fuzzy applications.

2 Fuzzy numbers and arithmetic

Basic fuzzy number arithmetic operations are defined in Table 1. The operators of fuzzy addition and subtraction are exact defined in the Table 1. However fuzzy multiplication and division operators defined in the Table 1 are only approximations to the actual results. In operations addition and subtraction standard approximation can be replaced by actual result, which is found by rewriting the membership function to define a set of closed intervals.

Commonly used approximations for the standard operators for TFNs showed that incorrect results could be obtained.

The discretization approaches were not a suitable alternative since they did not have a concise representation, required internal storage of many discrete points to reconstruct membership functions, and suffer from computational complexity.

Arithmetic operations	Definition
$\tilde{A} \oplus \tilde{B} =$	$\langle a_1 + a_2, b_1 + b_2, c_1 + c_2 \rangle$
$\tilde{A} \ominus \tilde{B} =$	$\langle a_1 - a_2, b_1 - b_2, c_1 - c_2 \rangle$
$\tilde{A} \otimes \tilde{B} =$	$\langle a_1 \cdot a_2, b_1 \cdot b_2, c_1 \cdot c_2 \rangle$
$\tilde{A} \oslash \tilde{B}$	$\langle \frac{a_1}{c_2}, \frac{b_1}{b_2}, \frac{c_1}{a_2} \rangle$

Table 1 Arithmetic operations on TFNs and their definition

The new approximation designed by Giachetti and Young is built upon six parameters which describe a parameterized fuzzy number. The presentation of a PFN is

$$\tilde{A} \rightarrow \langle a, b, c, \lambda, \rho, n \rangle \quad (1)$$

where *a*, *b* and *c* has the same meaning as with TFNs. λ and ρ parameters are the *spread ratios* and are defined

$$\lambda_i = \frac{b_i}{a_i}, \qquad \rho_i = \frac{b_i}{c_i} \quad (2)$$

Spread ratios are included because they characterize the approximation error. The term *n* is the order of the polynomial expression for the membership function. The definition for using these six parameters for performing arithmetic operations by [1] is shown in the Table 2.

Arithmetic operations	Definition
$\tilde{A} \oplus \tilde{B} =$	$\langle a_1 + a_2, b_1 + b_2, c_1 + c_2, \sqrt[\max(n_1,n_2)]{\lambda_1^{n_1} \cdot \lambda_2^{n_2}}, \sqrt[\max(n_1,n_2)]{\rho_1^{n_1} \cdot \rho_2^{n_2}}, \max(n_1, n_2) \rangle$
$\tilde{A} \ominus \tilde{B} =$	$\langle a_1 - a_2, b_1 - b_2, c_1 - c_2, \sqrt[\max(n_1,n_2)]{\lambda_1^{n_1} \frac{1}{\rho_2^{n_2}}}, \sqrt[\max(n_1,n_2)]{\rho_1^{n_1} \frac{1}{\lambda_2^{n_2}}}, \max(n_1, n_2) \rangle$
$\tilde{A} \otimes \tilde{B} =$	$\langle a_1 \cdot a_2, b_1 \cdot b_2, c_1 \cdot c_2, \sqrt[n_1+n_2]{\lambda_1^{n_1} \lambda_2^{n_2}}, n_1 + n_2 \rangle$
$\tilde{A} \oslash \tilde{B} =$	$\langle \frac{a_1}{c_2}, \frac{b_1}{b_2}, \frac{c_1}{c_2}, \sqrt[n_1+n_2+1]{\lambda_1^{n_1} (1/\rho_2)^{n_2+1}}, \sqrt[n_1+n_2+1]{\rho_1^{n_1} (1/\lambda_2)^{n_2+1}}, n_1 + n_2 + 1 \rangle$

Table 2 Fuzzy arithmetic with parameterized fuzzy numbers

It is recognized that the main source of error between the actual and approximated results of fuzzy multiplication and division is the difference between the polynomial shape and the straight line approximation. The new approximation [1,2] is based on the fact that a better approximation than a straight line approximation is a polynomial approximation. A generalized polynomial which closely matches the actual results was proposed [1]. The generalized polynomial is then scaled and added to the original linear result.

The scaling factor times the generalized polynomial is added to standard approximation to obtain a new approximation for each α-cut. For multiplication the alpha-cut expression are

$$P_{N(L)} = P_L + G(\alpha, n)\tau_L(n, \bar{\lambda})(b - a) \quad (3a)$$

$$P_{N(R)} = P_R + G(\alpha, n)\tau_R(n, \bar{\rho})(c - b) \quad (3b)$$

where symbol $\bar{\lambda}$ is geometric mean of λ.

And for the division

$$Q_{N(L)} = D_L + G(\alpha, n)\tau_L(n, \bar{\lambda})(b - a) \quad (4a)$$

$$Q_{N(R)} = D_R + G(\alpha, n)\tau_R(n, \bar{\rho})(c - b) \quad (4b)$$

The scaling expressions are for the left segment,

$$\tau_L = (n, \bar{\lambda}) = 0.568\bar{\lambda} + 0.11n - 0.859 \quad (5a)$$

And the right segment

$$\tau_R(n, \bar{\rho}) = -1.85\,\bar{\rho} + 0.144\,n + 1.19 \quad (5b)$$

The generalized polynomial expression, which closely tracks the shape of the polynomial for the actual multiplication and division has the following expression

$$G = \left[(-\alpha)^{\frac{(-1)^n - 1}{2}} \sum_{i=2}^{n} (-\alpha)^i \right] - \left[\frac{(-1)^n + 1}{2} \right] \alpha \quad (6)$$

In the derived expression there is an important aspect useful for object oriented approach that is that the new approximation introduced by Giachetti and Young [1,2] is composed from the standard approximation, expressed by e.g. $P_L, P_R, or D_L, D_R$, which is added to a new part. The new part as can be seen from (3, 4) is

created by multiplication of scaling expression and generalized polynomial expression. The generalized polynomial expression is a function of two arguments α and n.

All these aspects may be used in object oriented design of fuzzy arithmetic for fuzzy numbers.

3 Object-oriented approach and its tools

Generally, there are three main benefits of object-oriented approach: real world apprehension, stability of design and reusability of both design and implementation.

Real world apprehension - object oriented approach is widely accepted because it is close to our own natural perception of the real world. Object oriented analysis is based upon concepts such as objects and attributes, classes and members, wholes and parts. Generally it means communicating about phenomena and concepts of the application domain.

Stability of design - this aspect stresses instead focusing on the functionality of the system, the first step in the system's development should be making physical model of the real world with which the system is concerned. This model then forms the basis for the different functions that the system may have. Functions may later be changed, and new functions may be added without changing the underlying model. Object oriented tools provide a natural framework for modeling the application domain.

Reusability represents not only pure software component reuse but, in general, it also represents incremental program modification (development). It is possible to define a new component as an incremental extension of an existing one and thus preserving the relationship between the two components. In this way most object oriented systems are based on the class/subclass mechanism (inheritance), which enables to create programs by means of incremental extension. This feature is frequently considered to be the main advantage of object oriented approach. In case the meaning of reusability is overrated the final application may suffer from the fact that the relations between components are mainly dictated by maximal code sharing, often conflicting with the modeling requirements.

During our work we drew from books or articles, which describe design and implementation of object oriented approach in fuzzy sets. E.g. Schmucker [6] provides detailed Pascal-like code for some of the simpler fuzzy set operations. Lalonde and Pugh [7] devote a chapter to model a FuzzySet class in Smalltalk environment. We also applied our practical experience obtained from object oriented model in cluster analysis application [5]. In the designed class a general fuzzy set with discretized membership function is presented. For this reason the FuzzySet class inherits from a container like class Dictionary. Presented approach represents simple and intuitive solution.

Three main steps in object oriented approach represent identification of objects, classification and composition. Identification of objects is naturally given as we want to describe fuzzy number, triangularFN or parameterizedFN and so on. By classification we mean creation of proper hierarchy of classes (class / subclass system) that fulfills our demands. Observing arithmetic operations for triangular FN and Parameterized FN one can see that the part of the operations remain the same and the other part is extended (Table 1, 2). This feature can be used for object oriented description in the sense that a method describing a given operation in super class can be further extended (specialized) in its subclass. In our case triangular FN is the super class and parameterized FN is the subclass.

The same situation can be seen with the operations for new approximation that considerably improves standard approximation of alpha cuts. In addition the new approximation is created by addition of the standard approximation and a given polynomial extension. All can be seen in (3, 4). Standard approximation might be described in the super class and its extension in the subclass. There are tools in object oriented perspective that enable to create these requirements. Described approach has other advantages, which is implementation of the next classes that express other forms of fuzzy numbers such as trapezoidal FN class and so on.

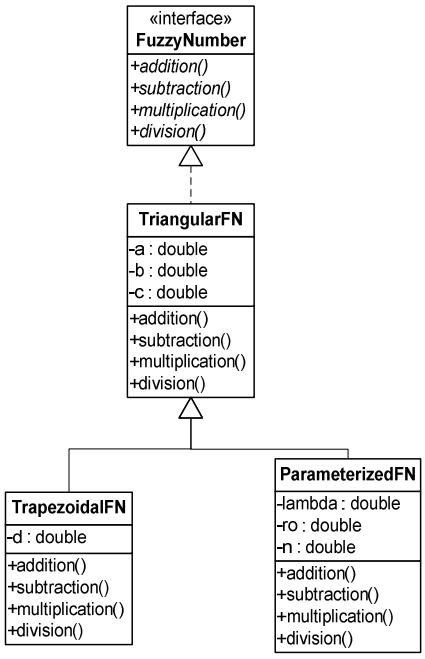
Composition can be applied e.g. when we would want to work within fuzzy number with a membership function, which could be represented by a discretized membership function. In this case a class describing such situation would have an attribute that would be declared as a container. The fuzzy number class and the container create composition.

4 Design and implementation of fuzzy arithmetic

Object oriented design of fuzzy arithmetic should cover arithmetic operation, which might be implemented in a different way depending on a concrete class of fuzzy number. For this reason a declaration of an interface including arithmetic operations will be the first step. The interface actually acts as an abstract super class for all other declared classes. FuzzyNumber interface describe arithmetic operations and type of necessary arguments as well as type of return objects from the operations.

Created model should cover both TFNs, PFNs potentially other fuzzy number expression such as trapezoidal fuzzy number but of course it could cover fuzzy numbers expressed by its discretized membership function. As in our paper we are primaryly interested in fuzzy by arithmetic all these operation are placed in the common interface called FuzzyNumber. This mechanism enables to implement different operations in a different ways. On the other hand it also facilitates to use further specialization principle for declared operations. It means that operations declared in the super class can be further extended in its subclasse.

In object oriented perspective of PFN can be considered as an extension of the TFN in that way that PFN has next specific attributes and the basic arithmetic operation are in this sense also extended. For this reason ParameterizedFN class is the subclass of TriangularFN class. TriangularFN class is also the superclass for all derived classes including fuzzy number expressed by an explicit membership function. The whole structure of designed classes and interface can be seen in the picture 1 in the form of UML class diagram.



Picture 1.: Fuzzy number arithmetic – class diagram

TriangilarFN class represents object oriented description of triangular fuzzy number. In the class three attributes *a, b, c,* representing triangular fuzzy number are declared. Apart from declared attributes the class contains implementation of the arithmetic methods described in the table 1.

ParameterizedFN class further extends TriangularFN superclass by declaring next data attributes λ, ρ, and *n*. Their meaning is explained shortly in the second part of the acticle or in [1] in a more detailed way. As attributes λ and ρ express so called spread ratios and can be derived from the attributes *a, b,* and *c* they are not explicitly entered by the user but they are calculated at object creation (instantiation). The term *n* is the attribute representing the order of the polynomial expression for the membership function. The arithmetic operations for the parameterizedFN are declared in the Table 2. Their implementation is following. The first part of the calculation is the same as for the triangularFN so it is used from the superclass. The other part differs so it is declared as extension in the superclass method.

The other arithmetic operations declared in this class could also include scalar addition and scalar multiplication. Scalar addition represents addition of PFN and a scalar (crisp number) and in the same way scalar multiplication means multiplication of PFN and a scalar (crisp number).

Some applications require using of alpha-cuts for various calculations. Just using alpha-cuts declared in ParameterizedFN enable to get far precise results than the use of standard approximation for alpha cuts. Calculating error is mainly seen when the multiplication or division arithmetic operation are repeatedly executed. In that case the use of parameterized alpha-cuts is considerable. It has its reason in the fact that multiplication and division have both polynomial membership function and the alpha cuts far better follow actual shape of the membership function.

One of the basic aspects of object oriented approach is information hiding, which means that not all part of the class (attributes and methods) can be “visible” from outside of the class. This is a powerful means that helps check the access inside the class. The possibility of hiding information is used for auxiliary operations such as calculation of the generalized polynomial or polynomial part of the new approximation that are assessable only within the class, not from the outside. On the other hand basic arithmetic methods and left and right segment of the new approximation are accessible both from inside and outside of the given class.

TrapezoidalFN class represents other possible classes derived from the description of the triangular fuzzy number class. We made this decision on the basis of the next additional parameter but we have not done extensive experiments with the class yet as we focused mainly on TriangularFN and ParameterizedFN classes.

4.1 Example of using object oriented approach in parameterized fuzzy numbers

At the beginning there is a need to create an object or instance of fuzzy number class, in our case it will be

ParameterizedFN class. During this phase all necessary parameters has to be added or entered. This means that parameters *a*, *b*, *c* and *n* has to be given. Java code for simple declaration of PFN1 <1, 4, 9, 1> can be:

```
ParameterizedFN pfn1 =
        new ParameterizedFN(1, 4, 9, 1);
```

Object *pfn1* is a parameterized fuzzy number with above declared parameters. For a real arithmetic operation it is necessary to have at least two fuzzy numbers on which the operation is applied. All basic arithmetic operations (in object oriented jargon 'methods') are created on the similar principle. Inside the arithmetic method a new (empty) object of the same class is created. First of the arithmetic operand is the 'receiver' of the message and the second operand is the parameter of the message, which is sent to the receiver. (The message then finds proper method). The required arithmetic operation is applied on the 'receiver' of the method and argument of the method. The result of the arithmetic operation is stored into a new object. Finally a reference of the new object is returned from the method and possibly assigned to other object.

```
ParameterizedFN pfn2 = new ParameterizedFN(1, 2, 3, 1);
```

```
ParametrerizedFN pfn3 = pfn1.division(pfn2);
```

where

pfn1 – 'receiver' of the message,

pfn2 – argument of the message,
second operand

pfn3 – result of the arithmetic operation.

Object *pfn3* is the result of arithmetic division. By this formalization even a complex fuzzy arithmetic operation can be expressed in a rather simple way. Let us consider an expression:

$E = (A + B) * (C / D)$

All upper cased letters represent e.g. parameterized fuzzy numbers. Formalization in Java language is:

```
ParameterizedFN E =
        (A.addition(B)).multiplication(C.division(D));
```

As it can be seen from the transcript, formalized code description is simple and quite comfortable. On the other hand it shifts the level of abstraction a bit higher.

	a	b	c	λ	ρ	n
operand 1	25.0	40.0	55.00	1.60	0.73	1
operand 2	4.0	10.0	16.00	2.50	0.63	1
addition	29.0	50.0	71.0	4.0	0.45	1
subtraction	21.0	30.0	39.0	2.56	0.29	1
multiplication	100.0	400.0	880.0	2.0	0.67	2
division	1.56	4.0	13.7	1.6	0.002	3

Table 3: Results from the basic fuzzy arithmetic operations

5 Results

Implementation of the FuzzyNumber interface and TriangularFN and ParameterizedFN classes was made in Java object oriented language. To prove the implementation is correct we use the similar examples as were used in [1, 2]. Achieved results for basic arithmetic operations are presented in the Table 3. New approximation results in alpha-cuts form is presented in the Table 4.

For the results in the table 4 there are given two PFNs:

$\tilde{x} \rightarrow \langle 70, 100, 130, 1 \rangle$

$\tilde{y} \rightarrow \langle 4, 10, 16, 1 \rangle$

	Actual quotient		New approximation	
Alpha	Left segm.	Right segm.	Left segm.	Right segm.
1	10.0	10.0	10.0	10.0
0.9	9.2	11.0	9.3	10.9
0.8	8.4	12.0	8.6	12.0
0.7	7.7	13.3	7.9	13.5
0.6	7.1	14.7	7.3	15.3
0.5	6.5	16.4	6.7	17.4
0.4	6.0	18.4	6.2	19.8
0.3	5.6	20.9	5.7	22.5
0.2	5.1	23.8	5.2	25.5
0.1	4.7	27.6	4.8	28.9
0	4.4	32.5	4.4	32.5

Table 4.: Comparing actual quotient with new approximation

6 Conclusion

Object-oriented approach offers next abstract level that enable user (designer) better means for modeling specific domain. This is achieved by encapsulating details inside objects and by strict distinction between interface and implementation. Thus user (designer) may focus on interface rather on the implementation. There are many formal descriptions of arithmetic operations on fuzzy numbers. Some parts are common the other are more and more specific. This is just what can be modeled with benefits using object-oriented perspective.

PFNs connected with polynomial approximation were introduced and invented by Giachetti and Young [1,2] as a better alternative to TFNs mainly in engineer's models. The advantage of using PFNs is their relative easy implementation and considerably better results in applying operations of multiplication and division. In our paper we tried to show object oriented approach in modeling fuzzy arithmetic by parameterized fuzzy numbers. UML class diagram of the designed solution and results achieved by Java object oriented language implementation is presented. In our approach we draw from the benefits of object oriented paradigm mainly from using classification, further specialization in class/subclass hierarchy rather than overriding and a mechanism of interface. Basic object oriented methods cover arithmetic operations and alpha cuts calculations.

Described approach could be also used for trapezoidal form of fuzzy numbers and finally also for fuzzy matrix and their operations. Object oriented approach applied in the implementation ensures clear arrangement and flexible solution that enables further possible modifications.

Acknowledgement

The paper was supported by the research scheme of the Institute for Research and Application of Fuzzy Modeling No: MSM6198898701

References

[1] Giachetti, R.E., Young, R.E.: A parametric representation of fuzzy numbers and their arithmetic operators. Fuzzy sets and systems. vol 91, No. 2, 1997

[2] Giachetti, R.E., Young, R.E.: Analysis of the error in the standard approximation used for multiplication of triangular and trapezoidal fuzzy numbers and the development of a new approximation. Fuzzy Sets and Systems. vol 91, (1997) 1-13

[3] Dubois, D., Prade,H.: Operations on fuzzy numbers. Internat. J. Systems Science 9 (6) (1978) 613-626

[4] Novak, V.: Zaklady fuzzy modelovani. BEN – technicka literatura Praha 2000

[5] Hunka, F.: Object Oriented Approach in Cluster Analysis. Acta Electronica et Informatica. No. 2, Vol. 3, 2003. pp55-59

[6] Schmucker, K. J.: Fuzzy Sets, Natural Language Computations, and Risk Analysis. Computer Science Press. Rockwille. Maryland, 1984

[7] Lalonde, W., Pugh, J.: Smalltalk V: Practice and Experience. Prentice Hall. Englewood Cliffs, New Jersey 1994

Inverse arithmetic operators for fuzzy intervals

Reda Boukezzoula
LISTIC, Université de Savoie
Domaine Universitaire, BP80439
74944 Annecy le Vieux Cedex
Reda.Boukezzoula@univ-savoie.fr

Sylvie Galichet
LISTIC, Université de Savoie
Domaine Universitaire, BP80439
74944 Annecy le Vieux Cedex
Sylvie.Galichet@univ-savoie.fr

Laurent Foulloy
LISTIC, Université de Savoie
Domaine Universitaire, BP80439
74944 Annecy le Vieux Cedex
Laurent.Foulloy@univ-savoie.fr

Abstract

Fuzzy arithmetic is a powerful tool in many engineering problems such as decision making, control theory, fuzzy systems and approximate reasoning. However, it is well known that the practical use of standard fuzzy arithmetic operators gives results more imprecise than necessary or in some cases, even incorrect. This problem is due to the overestimation effect induced by computing fuzzy arithmetic operations. In this paper a modified implementation for fuzzy unimodal interval arithemtics is defined where new subtraction and division operators are proposed. These new operators are exactly the inverse of the addition and multiplication operators. The effectiveness of the proposed methodology is illustrated by simulation examples.

Keywords: Fuzzy arithmetic, Inverse operators

1 Introduction

Nowadays, fuzzy numbers and fuzzy values are widely used in engineering applications because of their suitability for representing uncertain information. In this context, fuzzy arithmetics using fuzzy numbers and fuzzy intervals [5][12] is a frequently encountered problem, especially in decision making, control theory, fuzzy systems and approximate reasoning problems.

It is known that real numbers arithmetic operations are computed using unique rules that are independent of what is represented by the numbers involved. That is, the result of each particular arithmetic operation on real numbers depends only on the numbers involved and not on the entities represented by the numbers [19][21]. As is well known, the validity of this simple principle is also tacitly assumed in fuzzy arithmetic operations [4][16][17][24]. Indeed, the usual arithmetic operations on real numbers can be extended to the ones defined on fuzzy numbers by means of Zadeh's extension principle [26][27]. In this context, direct implementation of this principle in fuzzy arithmetics is computationally expensive due to the requirement of solving a nonlinear programming problem [2][18]. To overcome this deficiency, many researchers have investigated this problem by observing the fuzzy numbers as a collection of α-levels [1][9][10][20]. In this case, fuzzy arithmetics are performed using conventional interval arithmetics according to the α-cut representation.

In this framework, it has been proved [6][7] that the conventional interval arithmetic operations $\{\oplus, \ominus, \otimes, \oslash\}$ defined for intervals [13][14] can be directly extended to fuzzy intervals using the left and right profiles instead of the lower and upper limits of the intervals. That is the approach used in this paper for computing fuzzy arithmetic operations where unimodal fuzzy intervals are considered.

It is well known that standard interval arithmetic operations give a guaranteed enclosure of the solution, which is usually too large or more imprecise (the so-called "dependence effect"). The origin of the problem is the memoryless nature of interval arithmetic operations [13][14]. So, the independence property assumed in real number arithmetics holds not true for computing interval arithmetics as well as fuzzy interval arithmetics [21][22]. As mentioned by Zhou in [25] and stated by Piegat in [19]: «*the standard fuzzy arithmetic does not take into account all the information available, and the obtained results are more imprecise than necessary or in some cases, even incorrect*». Indeed, as discussed in [21], the fuzzy arithmetic operations tackle into account only the information contained in the operands and completely ignore the additional information that may emanate from the meaning of these operands. It follows that fuzzy arithmetic operations may produce some counterintuitive results. For example, when considering a fuzzy interval A, according to standard fuzzy arithmetics we have $A \ominus A \neq 0$ and $A \oslash A \neq 1$. However, in engineering applications, it can be desirable to have crisp values for $A \ominus A$, and $A \oslash A$,

i.e. the crisp values 0 and 1 respectively.

Moreover, it can be stated that the X solution of the fuzzy linear equation $B \oplus X = A$ is not, as we would expect, the value $X = A \ominus B$, and the solution of the fuzzy equation $B \otimes X = A$ is not the fuzzy interval $X = A \oslash B$. Consequently, the operations of addition and subtraction (multiplication and division) of fuzzy intervals are not inverse operations.

In order to overcome these problems, the authors [21][19] propose to tackle into account all constraints which represent additional information that are imposed on variable values in the considered problem. If these constraints (equality constraints, inequality constraints, ...) are respected, fuzzy arithmetics gives exact operations. However, if they are ignored, the obtained results are less precise than necessary. An elegant study of this problem has been developed in [19][21].

In this paper another way is explored to determine exact fuzzy arithmetics operations. The proposed methodology is not based on additional information constraints on the variables but on the characterization of the volume and/or the quantity of uncertainties contained in the handled variables, for which exact fuzzy arithmetics can be obtained.

When considering unimodal fuzzy intervals the originality of this work lies in the answers given to the following questions:

- *Is it possible to develop arithmetic operators for fuzzy intervals so that addition and multiplication are exactly inverse operations of subtraction and division?*
- *Is it possible to determine conditions for guaranteeing the existence of these exact operators?*

This paper is structured in the following way. In section 2, the concepts of fuzzy intervals and fuzzy profiles are introduced. Section 3 is devoted to the presentation of conventional fuzzy arithmetic operations according to the fuzzy profiles principle. In section 4, new fuzzy arithmetic operators are proposed with their existence conditions. Concluding remarks are finally given in Section 5.

2 Fuzzy Intervals and Fuzzy Profiles

An interval a can be considered as a set of elements to which a rectangular membership function $\mu_a(x)$ is associated (see figure 1). It can thus be viewed as a special fuzzy number whose membership function takes the value 1 over the interval and 0 anywhere else. Therefore, interval analysis can be considered as a subset of fuzzy set theory [15][23].

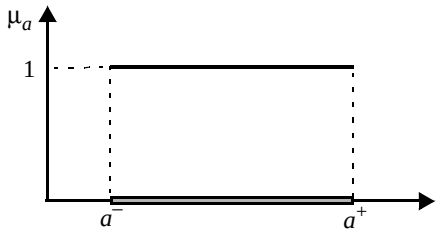


Figure 1: Interval Representation

The interval representation supposes that all possible values of the interval a belong to it with the same membership degree. When using a unimodal fuzzy interval representation, a possibility distribution, represented by a membership function, is associated with the fuzzy interval A whose support is the interval of all possible values and whose kernel value is the one and only best value (see figure 2).

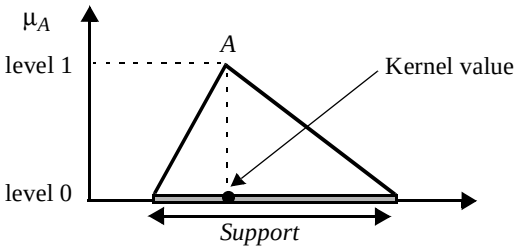


Figure 2: Fuzzy Interval Representation

In order to extend interval arithmetics to fuzzy intervals, one has to consider two dimensions. The first one (horizontal dimension) is similar to that used in interval representation, that is the real line $\Re$. The second one (vertical dimension) is related to the handling of the membership degrees and thus restricted to the interval [0, 1]. In this context, two kinds of information are required for completely defining fuzzy intervals and thus extending the conventional interval operations to the latter. Both pieces of information, called support and kernel value, are defined on the horizontal dimension, but are associated to two different levels (level 0 and level 1) on the vertical dimension (see figure 2).

Generally, an interval valued number is defined by the set of elements lying between its lower and upper limits as $a = \{x \mid a^- \le x \le a^+, x \in \Re\}$. Given an interval a, its midpoint $M(a)$ and its radius $R(a)$ are defined by:

$$M(a) = (a^- + a^+)/2 \quad \text{and} \quad R(a) = (a^+ - a^-)/2 \quad (1)$$

Let us now consider a unimodal fuzzy interval A [5] whose membership function is denoted μ_A and whose above mentioned features, i.e. support and kernel value, are respectively denoted $S_A = [S_{A^-}, S_{A^+}]$ and K_A. Obviously, since S_A is an interval, we have:

$$S_{A^-} = M(S_A) - R(S_A) \quad \text{and} \quad S_{A^+} = M(S_A) + R(S_A) \quad (2)$$

In order to specify the fuzzy interval shape, two additional functions are used to link the support with the kernel value according to the vertical dimension. These functions, called left and right profiles and denoted respectively A^- (the increasing part) and A^+ (the decreasing part) for the fuzzy interval A, are defined as:

$$\begin{aligned} A^-(\lambda) &= Inf\ \{x |\ \mu_A(x) \geq \lambda\ ;\ x \geq S_{A^-}\} \\ A^+(\lambda) &= Sup\ \{x |\ \mu_A(x) \geq \lambda\ ;\ x \leq S_{A^+}\} \end{aligned} \quad (3)$$

where $\lambda \in [0, 1]$ represents the vertical dimension. It can be easily stated that:

$$K_A = A^-(1) = A^+(1), \quad S_A = [A^-(0), A^+(0)]. \quad (4)$$

Finally, the fuzzy interval A can be univoquely defined by its left and right profiles. Thus, in the same way that the conventional interval a is denoted $[a^-, a^+]$, the fuzzy interval A will be denoted $[A^-, A^+]$.

Reciprocally, the profile representation $[A^-(\lambda), A^+(\lambda)]$ defines a unimodal fuzzy interval A if and only if the following conditions are satisfied:

(***a***) $A^-(\lambda)$ is increasing with respect to $\lambda \in [0, 1]$,
(***b***) $A^+(\lambda)$ is decreasing with respect to $\lambda \in [0, 1]$,
(***c***) $A^-(1) = A^+(1) = K_A$,
(***d***) $A^-(0) \leq K_A \leq A^+(0)$. (5)

It can be noted that condition (***d***) is induced by properties (***a***),(***b***) and (***c***).

3 Conventional Fuzzy Arithmetic Operations with Fuzzy Profiles

The conventional interval arithmetic operations $\{\oplus, \ominus, \otimes, \oslash\}$ defined for intervals [13][14] can be directly extended to fuzzy intervals using the left and right profiles instead of the lower and upper limits of the intervals [6][7].

3.1 Profile-based definitions

Let $A = [A^-(\lambda), A^+(\lambda)]$ and $B = [B^-(\lambda), B^+(\lambda)]$ be two fuzzy intervals, the classical four arithmetic operations are given by:

$$(A \oplus B)(\lambda) = [A^-(\lambda) + B^-(\lambda), A^+(\lambda) + B^+(\lambda)] \quad (6)$$

$$(A \ominus B)(\lambda) = [A^-(\lambda) - B^+(\lambda), A^+(\lambda) - B^-(\lambda)] \quad (7)$$

$$(A \otimes B)(\lambda) = [min\ Z(\lambda), max\ Z(\lambda)] \quad (8)$$

where $Z(\lambda) = \{A^-(\lambda).B^-(\lambda), A^-(\lambda).B^+(\lambda), A^+(\lambda).B^-(\lambda), A^+(\lambda).B^+(\lambda)\}$

For B such that $0 \notin S_B$

$$(A \oslash B)(\lambda) = [A^-(\lambda), A^+(\lambda)] \otimes [1/B^+(\lambda), 1/B^-(\lambda)] \quad (9)$$

3.2 Illustrative Example

For the sake of simplicity, let us consider two triangular fuzzy intervals A and B (see figure 3) given by:

$$A(\lambda) = [A^-(\lambda), A^+(\lambda)] = [1 + 2\lambda, 7 - 4\lambda]$$

$$B(\lambda) = [B^-(\lambda), B^+(\lambda)] = [-3 + \lambda, -1 - \lambda].$$

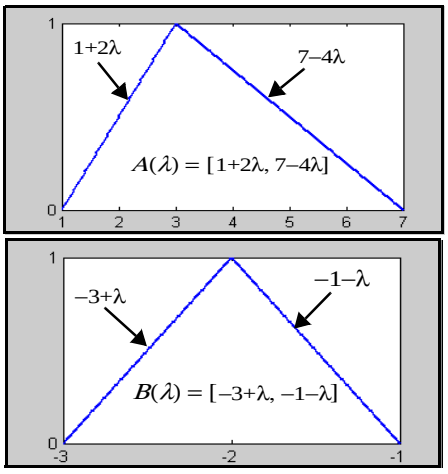


Figure 3: Fuzzy Intervals A and B.

Using the standard operations on the fuzzy profiles of A and B, the following results are obtained:

$$(A \oplus B)(\lambda) = [-2 + 3\lambda, 6 - 5\lambda] \quad (10)$$

$$(A \ominus B)(\lambda) = [2 + 3\lambda, 10 - 5\lambda] \quad (11)$$

$$(A \otimes B)(\lambda) = [-4\lambda^2 + 19\lambda - 21, -2\lambda^2 - 3\lambda - 1] \quad (12)$$

$$(A \oslash B)(\lambda) = [(7-4\lambda) / (-1-\lambda), (1+2\lambda) / (-3+\lambda)] \quad (13)$$

The obtained results are illustrated in the following figures (see figure 4 and figure 5).

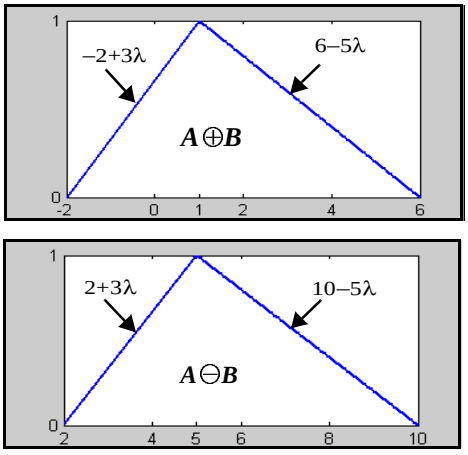


Figure 4: Addition and subtraction of A and B.

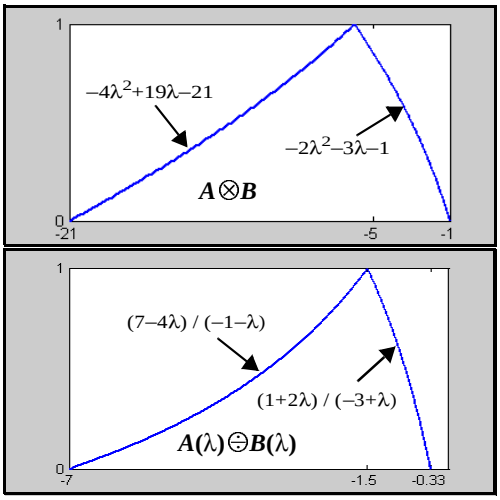


Figure 5: Multiplication and division of A and B.

3.3 Remarks

While being based on standard fuzzy arithmetics, these results are not controversial. However, as illustrated by the following computations:

$$(A \ominus A)(\lambda) = [-6 + 6\lambda, 6 - 6\lambda]$$
$$(A \oslash A)(\lambda) = [(1+2\lambda)/(7-4\lambda), (7-4\lambda)/(1+2\lambda)],$$

it is obvious that the defined operators produce counterintuitive results. Indeed, as mentioned in [19][21], $A \ominus A \neq 0$ and $A \oslash A \neq 1$, where 0 and 1 are singletons whose fuzzy representation is based on constant profiles, that is $\forall \lambda \in [0, 1]$, $0^-(\lambda) = 0^+(\lambda) = 0$ and $1^-(\lambda) = 1^+(\lambda) = 1$. It follows that the X solution of the fuzzy linear equation $A \oplus X = D$ is not, as we would expect, $X = D \ominus A$. For example, by choosing $D = [-2 + 3\lambda, 6 - 5\lambda]$, we would expect according to equation (10) that $X = D \ominus A$ be equal to B, i.e. $[-3+\lambda, -1-\lambda]$, which is unfortunately not the case:

$$X = D \ominus A = [-2+3\lambda, 6-5\lambda] \ominus [1+2\lambda, 7-4\lambda]$$
$$= [-9+7\lambda, 5-7\lambda] \neq B.$$

The same annoyance appears when solving the fuzzy equation $A \otimes X = E$ whose solution is not given by $X = E \oslash A$ as expected. For example, by choosing $E = [-4\lambda^2 + 19\lambda - 21, -2\lambda^2 - 3\lambda - 1]$ according to equation (12), the computation of $X = E \oslash A$ leads to a result different from B:

$$X = E \oslash A$$
$$= [-4\lambda^2 + 19\lambda - 21, -2\lambda^2 - 3\lambda - 1] \oslash [1+2\lambda, 7-4\lambda]$$
$$= [-4\lambda^2+19\lambda-21, -2\lambda^2-3\lambda-1] \otimes [1/7-4\lambda, 1/1+2\lambda]$$
$$= [-4\lambda^2 + 19\lambda - 21 / 1+2\lambda, -2\lambda^2 - 3\lambda - 1/ 7-4\lambda]$$
$$\neq B.$$

To sum up, the addition and subtraction (resp. multiplication and division) of fuzzy intervals are not reciprocal operations. According to this statement, it is not possible to solve inverse problems exactly using standard fuzzy arithmetic operators. In this context, the sequel of the paper is devoted to the definition of modified fuzzy arithmetic operators that allows exact inversion.

4 Exact Fuzzy Inverse Operators

Our objective is now to develop new subtraction and division operators denoted respectively $\ominus_{\approx}$ and $\oslash_{\approx}$ which are inverse operations of the addition $\oplus$ and multiplication $\otimes$ operators.

4.1 The modified $\ominus_{\approx}$ Operator

Definition 1: Let A and B be two unimodal fuzzy intervals with profile representation $A(\lambda) = [A^-(\lambda), A^+(\lambda)]$ and $B(\lambda) = [B^-(\lambda), B^+(\lambda)]$.

The modified difference operator $\ominus_{\approx}$ is defined by:

$$(A \ominus_{\approx} B)(\lambda) = [A^-(\lambda) - B^-(\lambda), A^+(\lambda) - B^+(\lambda)] \quad (14)$$
$$= [\Phi^-(\lambda), \Phi^+(\lambda)] = \Phi(\lambda)$$

when the result Φ actually represents a unimodal fuzzy interval.

Otherwise, $\ominus_{\approx}$ is undefined.

4.2 Properties of the operator $\ominus_{\approx}$

i. *Inverse operator of* $\oplus$

When defined, $\ominus_{\approx}$ is the inverse operator of $\oplus$, that is $B \oplus (A \ominus_{\approx} B) = (A \ominus_{\approx} B) \oplus B = A$.

Proof

$$[B^-(\lambda), B^+(\lambda)] \oplus [\Phi^-(\lambda), \Phi^+(\lambda)]$$
$$= [B^-(\lambda), B^+(\lambda)] \oplus [A^-(\lambda) - B^-(\lambda), A^+(\lambda) - B^+(\lambda)]$$
$$= [A^-(\lambda), A^+(\lambda)] = A(\lambda).$$

ii. *Commutativity*

The operator $\ominus_{\approx}$ is not commutative, that is $A \ominus_{\approx} B \neq B \ominus_{\approx} A$.

iii. *Multiplication by a scalar*

$\omega \,.\, (A \ominus_{\approx} B) = (\omega \,.\, A \ominus_{\approx} \omega \,.\, B)$ for any scalar ω.

iv. *Neutral element*

The singleton 0 defined by a constant profile equal to 0 is a right neutral element of $\ominus_{\approx}$, that is: $A \ominus_{\approx} 0 = A$.

v. *Associativity*

The operator $\ominus_{\approx}$ is associative, that is:

$A \ominus_{\approx} B \ominus_{\approx} C = A \ominus_{\approx} (B \ominus_{\approx} C) = (A \ominus_{\approx} B) \ominus_{\approx} C$.

vi. *Inverse element*

Any fuzzy interval is its own inverse under the modified subtraction $\ominus_{\approx}$, i.e. $A \ominus_{\approx} A = 0$.

vii. *Regularity*

$A \ominus_{\approx} B = A \ominus_{\approx} C \implies B = C.$

viii. *Pseudo-distributivity with regard to* $\oplus$

$(A \oplus B) \ominus_{\approx} (C \oplus D) = (A \ominus_{\approx} C) \oplus (B \ominus_{\approx} D),$

$(A \ominus_{\approx} B) \oplus (C \ominus_{\approx} D) = (A \oplus C) \ominus_{\approx} (B \oplus D).$

4.3 Necessary existence condition for $\ominus_{\approx}$

According to equation (14), it is clear that $A \ominus_{\approx} B$ can always be computed. However, the difference operation is only valid when the obtained result $\Phi(\lambda)$ is a unimodal fuzzy interval, which means that the operator $\ominus_{\approx}$ exists only if $\Phi(\lambda)$ satisfies the four conditions expressed in (5). The aim of this subsection is to formulate a simple necessary existence condition of $\ominus_{\approx}$ that only depends on some features of the operands A and B.

Proposition 1: The operator $\ominus_{\approx}$ exists only if the following condition is satisfied:

$$| \Delta K - (M(S_A) - M(S_B)) | \leq R(S_A) - R(S_B) \qquad (15)$$

with : $\Delta K = K_A - K_B$.

Proof

$\Phi(\lambda)$ is an interval fuzzy number only if conditions (5) are verified.

According to (14), $\Phi^-(1) = A^-(1) - B^-(1)$ and $\Phi^+(1) = A^+(1) - B^+(1)$. As A and B are unimodal fuzzy intervals, it follows that $\Phi^-(1) = \Phi^+(1) = K_A - K_B = \Delta K$. So, condition (5)(***c***) always holds.

Satisfying condition (5)(***d***) requires that

$$\Phi^-(0) \leq \Delta K \leq \Phi^+(0) \quad \Leftrightarrow$$

$$A^-(0) - B^-(0) \leq \Delta K \leq A^+(0) - B^+(0) \quad \Leftrightarrow$$

$$(M(S_A) - R(S_A)) - (M(S_B) - R(S_B)) \leq \Delta K \leq (M(S_A) + R(S_A)) - (M(S_B) + R(S_B)) \quad \Leftrightarrow$$

$$- R(S_A) + R(S_B) \leq \Delta K - (M(S_A) - M(S_B)) \leq R(S_A) - R(S_B)) \quad \Leftrightarrow$$

$$| \Delta K - (M(S_A) - M(S_B)) | \leq R(S_A) - R(S_B),$$

which concludes the proof that condition (15) is a necessary condition for the definition of operator $\ominus_{\approx}$. ❑

Remarks

- It can be stated that the proposed difference operator is an adapted version of the Hukuhara difference definition for fuzzy intervals. Indeed, the Hukuhara difference of two sets $X \in C$ and $Y \in C$, if it exists, is a set $Z \in C$ such that $X = Y + Z$, where C is the family of all nonempty convex compact subsets [3][11].
- When triangular fuzzy numbers are considered, it can be easily proved that condition (15) is a sufficient condition for the existence of the operator $\ominus_{\approx}$. In other words, the satisfaction of (15) guarantees that Φ^- (resp. Φ^+) is increasing (resp. decreasing) with respect to $\lambda \in [0, 1]$.
- If the two fuzzy intervals A and B are symmetric, i.e. $K_A = M(S_A)$ and $K_B = M(S_B)$, then the existence condition (15) is reduced to the support condition $R(S_A) - R(S_B) \geq 0$.

4.4 The modified $\ominus_{\approx}$ Operator

Definition 2: Let A and B be two unimodal fuzzy intervals with $0 \notin S_B$, the modified division operator $\ominus_{\approx}$ is defined by:

$$A(\lambda) \ominus_{\approx} B(\lambda) = \Phi(\lambda) = [\Phi^-(\lambda), \Phi^+(\lambda)] \qquad (16)$$

where:

$$\Phi^-(\lambda) = \frac{Num^-(\lambda)}{Den^-(\lambda)} = \frac{M(A(\lambda)) - R(A(\lambda)) \cdot sign(M(S_B))}{M(B(\lambda)) - R(B(\lambda)) \cdot sign(Num^-(\lambda))} \qquad (17)$$

and

$$\Phi^+(\lambda) = \frac{Num^+(\lambda)}{Den^+(\lambda)} = \frac{M(A(\lambda)) + R(A(\lambda)) \cdot sign(M(S_B))}{M(B(\lambda)) + R(B(\lambda)) \cdot sign(Num^+(\lambda))} \qquad (18)$$

when the result Φ actually represents a unimodal fuzzy interval.

Otherwise, $\ominus_{\approx}$ is undefined.

4.5 Properties of the operator $\ominus_{\approx}$

i. *Inverse operator of* $\otimes$

When defined, $\ominus_{\approx}$ is the inverse operator of $\otimes$, that is $B \otimes (A \ominus_{\approx} B) = (A \ominus_{\approx} B) \otimes B = A$.

Proof

$$B(\lambda) \otimes (A(\lambda) \ominus_{\approx} B(\lambda)) = B(\lambda) \otimes \Phi(\lambda) = [(B(\lambda) \otimes \Phi(\lambda))^-, (B(\lambda) \otimes \Phi(\lambda))^+] \qquad (19)$$

By substituting the expression of Φ given by (17) and (18) in equation (19) and according to the sign of S_B, two cases can be distinguished:

1) If $S_B > 0$, the expression is given by:

$(B \otimes \Phi)^- = M(B).\Phi^- - R(B).\Phi^-.\ sign(\Phi^-)$

$(B \otimes \Phi)^+ = M(B).\Phi^+ + R(B).\Phi^+.\ sign(\Phi^+)$

otherwise expressed as:

$$(B \otimes \Phi)^{-} = \frac{M(B) \cdot A^{-}}{M(B) - R(B) \cdot sign(Num^{-})} - \frac{R(B) \cdot A^{-}}{M(B) - R(B) \cdot sign(Num^{-})} \cdot sign(\Phi^{-}) \quad (20)$$

and

$$(B \otimes \Phi)^{+} = \frac{M(B) \cdot A^{+}}{M(B) + R(B) \cdot sign(Num^{+})} + \frac{R(B) \cdot A^{+}}{M(B) + R(B) \cdot sign(Num^{+})} \cdot sign(\Phi^{+}) \quad . \quad (21)$$

As $sign(Num^{-}) = sign(A^{-}) = sign(\Phi^{-})$
$sign(Num^{+}) = \text{sign}(A^{+}) = sign(\Phi^{+})$
equations (20) and (21) can be rewritten as:

$$(B \otimes \Phi)^{-} = \frac{(M(B) - R(B) \cdot sign(\Phi^{-})) \cdot A^{-}}{(M[B] - R[B] \cdot sign(\Phi^{-}))} = A^{-}$$
$$(B \otimes \Phi)^{+} = \frac{(M(B) + R(B) \cdot sign(\Phi^{+})) \cdot A^{+}}{(M(B) + R(B) \cdot sign(\Phi^{+}))} = A^{+} \quad (22)$$

2) If $S_B < 0$, it can be stated that:

$(B \otimes \Phi)^{-} = M(B).\Phi^{+} - R(B).\Phi^{+}.\ sign(\Phi^{+})$,
$(B \otimes \Phi)^{+} = M(B).\Phi^{-} + R(B).\Phi^{-}.\ sign(\Phi^{-})$.

By adopting the same principle as in the previous case, the following expression are obtained:

$$(B \otimes \Phi)^{-} = A^{-} \quad . \quad (23)$$
$$(B \otimes \Phi)^{+} = A^{+}$$

❑

ii. *Commutativity*

The operator $\ominus_{\approx}$ is not commutative, that is $A \ominus_{\approx} B \neq B \ominus_{\approx} A$.

iii. *Multiplication by a scalar*

$\omega.(A \ominus_{\approx} B) = (\omega . A \ominus_{\approx} \omega . B)$ for any scalar ω.

iv. *Neutral element*

The singleton 1 defined by a constant profile equal to 1 is a right neutral element of $\ominus_{\approx}$, that is $A \ominus_{\approx} 1 = A$.

v. *Associativity*

The operator $\ominus_{\approx}$ is associative, i.e.

$A \ominus_{\approx} B \ominus_{\approx} C = A \ominus_{\approx} (B \ominus_{\approx} C) = (A \ominus_{\approx} B) \ominus_{\approx} C$.

vi. *Inverse element*

Any fuzzy interval is its own inverse under the modified division $\ominus_{\approx}$, i.e. $A \ominus_{\approx} A = 1$.

vii. *Regularity*

$A \ominus_{\approx} B = A \ominus_{\approx} C \quad \Rightarrow \quad B = C$.

viii. *Distributivity with regard to* $\oplus$

$(A \oplus B) \ominus_{\approx} C = (A \ominus_{\approx} C) \oplus (B \ominus_{\approx} C)$.

4.6 Necessary existence condition for $\ominus_{\approx}$

Proposition 2: The operator $\ominus_{\approx}$ exists only if the following condition is satisfied:

$$\frac{1 - \left| Rex(S_A) \right|}{1 - S \cdot \left| Rex(S_B) \right|} \leq \frac{\delta K}{\delta M} \leq \frac{1 + \left| Rex(S_A) \right|}{1 + \left| Rex(S_B) \right|} \quad (24)$$

where $\delta K = K_A / K_B$, $\delta M = M(S_A) / M(S_B)$,
$Rex(.) = R(.) / M(.)$ and $S = sign(1 - |Rex(S_A)|)$.

Proof

$\Phi(\lambda)$ is an interval fuzzy number only if conditions (5) are verified.

According to (17), as $R(A(1)) = R(B(1)) = 0$,
$\Phi^{-}(1) = M(A(1)) / M(B(1)) = K_A / K_B = \delta K$

In the same way, according to (18),
$\Phi^{+}(1) = M(A(1)) / M(B(1)) = K_A / K_B = \delta K$

So, $\Phi^{-}(1) = \Phi^{+}(1) = \delta K$, which means that continuity between left and right profiles is guaranteed. In other words, condition (5)(***c***) always holds.

Satisfying condition (5)(***d***) requires that

$$\Phi^{-}(0) \leq \delta K \leq \Phi^{+}(0) \quad (25)$$

Two situations have to be considered according to the sign of B that is:

1) If B is positive, according to equations (17) and (18), the inequality (25) becomes:

$$\delta M \frac{1 - Rex(S_A)}{1 - s_1 Rex(S_B)} \leq \delta K \leq \delta M \frac{1 + Rex(S_A)}{1 + s_2 Rex(S_B)} \quad (26)$$

where $s_1 = sign\,(M(S_A) - R(S_A))$ and
$s_2 = sign\,(M(S_A) + R(S_A))$.

2) If B is negative, the inequality (25) becomes:

$$\delta M \frac{1 + Rex(S_A)}{1 - s_2 Rex(S_B)} \leq \delta K \leq \delta M \frac{1 - Rex(S_A)}{1 + s_1 Rex(S_B)} \quad . \quad (27)$$

By regrouping inequalities (26) and (27), it follows:

$$\frac{1 - \left|Rex(S_A)\right|}{1 - s_1 s_2 \left|Rex(S_B)\right|} \leq \frac{\delta K}{\delta M} \leq \frac{1 + \left|Rex(S_A)\right|}{1 + \left|Rex(S_B)\right|} . \quad (28)$$

As $s_1 s_2 = sign\,(M(S_A)^2 - R(S_A)^2) = sign\,(1 - Rex(S_A)^2) = sign\,(1 - |\,Rex(S_A)\,|\,) = S$, it follows that condition (5)(***d***) is satisfied if and only if constraint (24) holds.

❑

Remark

- The inequality (24) is a necessary condition for the existence of $A \ominus_{\approx} B$ but not a sufficient one. Whenever condition (24) holds, the existence of $A \ominus_{\approx} B$ further required that Φ^{-} (resp. Φ^{+}) be

increasing (resp. decreasing) with respect to $\lambda \in [0, 1]$. Actually, it is possible that condition (24) is verified for all the level-cuts of A and B without Φ being a fuzzy subset. In other words, the monotonicity of Φ^- and Φ^+ not only depends on the support and kernel of A and B but also on their shape. For example, it can be verified that condition (24) is sufficient for triangular fuzzy intervals A and B. On the contrary, when A is the product of two triangular fuzzy intervals, condition (24) is no more sufficient.

4.7 Illustrative Example

Case 1: Let us illustrate the behavior of the proposed modified operators for computing with the two fuzzy intervals A and B previously defined in section 3.2

It can be verified that the condition (15) is respected, i.e.

$$\Delta K = K_A - K_B = 3 - (-2) = 5$$

$$M(S_A) - M(S_B) = M([1,7]) - M([-3,-1]) = 6$$

$$R(S_A) - R(S_B) = 3 - 1 = 2$$

$$|\Delta K - (M(S_A) - M(S_B))| = |5 - 6| \leq R(S_A) - R(S_B) = 2$$

In this case, the modified difference between A and B is defined and given by:

$$A(\lambda) \ominus_{\approx} B(\lambda) = [1+2\lambda, 7-4\lambda] \ominus_{\approx} [-3+\lambda, -1-\lambda] = [4 + \lambda, 8-3\lambda]$$

As expected, it can be verified that:

$$[4+\lambda, 8-3\lambda] \oplus [-3+\lambda, -1-\lambda] = A(\lambda)$$

In the same way, let us evaluate $A \oslash_{\approx} B$. Firstly let us verify that the existence condition (24) is respected.

$$\delta K = K_A / K_B = 3/(-2) = -6$$

$$\delta M = M(S_A) / M(S_B) = 4/(-2) = -8$$

By substituting these values into (24), it follows:

$$0.5 \leq \frac{6}{8} \leq \frac{7}{6},$$

which means that (24) is verified.

Then, computing $A \oslash_{\approx} B$, we get:

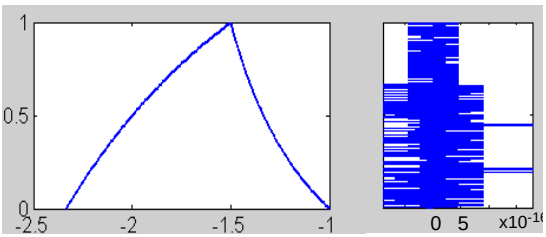
$$[1+2\lambda, 7-4\lambda] \oslash_{\approx} [-3+\lambda, -1-\lambda] = [\Phi^-(\lambda), \Phi^+(\lambda)]$$

with:

$$\Phi^-(\lambda) = \frac{(4-\lambda) - (3-3\lambda)\cdot(-1)}{(-2) - (1-\lambda)\cdot(1)} = \frac{7-4\lambda}{-3+\lambda}$$

$$\Phi^+(\lambda) = \frac{(4-\lambda) + (3-3\lambda)\cdot(-1)}{(-2) + (1-\lambda)\cdot(1)} = \frac{1+2\lambda}{-1-\lambda}$$

It can be verified that $\oslash_{\approx}$ is the exact inverse of $\otimes$ as shown in Figure 6 where the right plot illustrates that $A \ominus_{\approx} [(A \oslash_{\approx} B) \otimes B] = 0$ (with some computation noise).



$A \oslash_{\approx} B$ $\quad$ $A \ominus_{\approx} [(A \oslash_{\approx} B) \otimes B]$

Figure 6: Exact fuzzy inverse operators

Case 2: Let us now consider two other fuzzy intervals A and B:

$$A(\lambda) = [A^-(\lambda), A^+(\lambda)] = [1 + 0.5\lambda, 5 - 3.5\lambda]$$

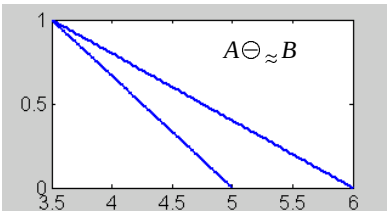
$$B(\lambda) = [B^-(\lambda), B^+(\lambda)] = [-4 + 2\lambda, -1 - \lambda].$$

It can be stated that condition (15) is violated. Indeed, $|3.5 - 5.5| \geq 0.5$. In this case, the difference operator is undefined. Actually, computing $A \ominus_{\approx} B$ would produce $[5 - 1.5\lambda, 6 - 2.5\lambda]$, that is a non valid result since it does not represent a fuzzy interval (see Figure 7(a)).

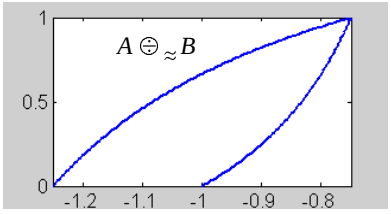
In the same way, it can be verified that the existence condition for the division operator is not respected:

$$\frac{5}{6} \leq \frac{10}{9} \geq \frac{50}{48}.$$

By computing the result that would be obtained, we get $[(5 - 3.5\lambda) / (-4 + 2\lambda), (1 + 0.5\lambda) / (-1 - \lambda)]$, again a non valid result as illustrated by Figure 7(b).



(a) Subtraction



(b) Division

Figure 7: Non valid operations

5 Conclusion

This paper presents a new methodology for the implementation of the substraction and the division operators between fuzzy intervals with their existence conditions. The proposed operators are exact inverses of the addition and multiplication operators. In some cases, these new modified operators can solve the overestimation problem well known in fuzzy arithmetics.

References

[1] Bodjanova S., «Alpha-bounds of fuzzy numbers», Information Sciences, Vol. 152, pp. 237-266, 2003.

[2] Dong W. M. and Wong F. S., «Fuzzy weighted averages and implementation of the extension principle», Fuzzy Sets and Systems, N°. 21, pp. 183-199, 1987.

[3] Diamond P. and R. Körner R., «Extended fuzzy linear models and least squares estimates», Computers&Mathematics with Applications, Vol. 33, N°. 9, pp. 15-32, 1997.

[4] Dubois D. and Prade H., «Operations on fuzzy numbers», Journal of Systems Science, No. 9, pp. 613-626, 1978.

[5] Dubois D., Kerre E., Mesiar R. and Prade H., «Fuzzy interval analysis», in: D. Dubois, H. Prade (Eds.), Fundamentals of Fuzzy Sets, The Handbooks of Fuzzy Sets Series, Kluwer, Boston, pp. 483-581, 2000.

[6] Dubois D., Fargier H. and Fortin J., «The Empirical Variance of a Set of Fuzzy Intervals», Proc. of the Fuzzy IEEE Conference, pp. 885-890, 2005.

[7] Dubois D., Fargier H. and Fortin J., «A Generalized Vertes Method for Computing with Fuzzy Intervals», Proc. of the Fuzzy IEEE Conference, pp. 541-546, 2004.

[8] Filev D.P. and Yager R.R., «Operations on fuzzy numbers via fuzzy reasoning», Fuzzy Sets and Systems, Vol. 91, No. 2, pp. 137-142, 1997.

[9] Giachetti R.E. and Young R.E., «A parametric representation of fuzzy numbers and their arithmetic operators», Fuzzy Sets and Systems, Vol. 91, No. 2, pp. 185-202, 1997.

[10] Giachetti R.E, Young R.E., «Analysis of the error in the standard approximation used for multiplication of triangular and trapezoidal fuzzy numbers and the development of a new approximation», Fuzzy Sets and Systems, N°. 91, pp. 1–13, 1997.

[11] Hukuhara M. «Integration des applications measurables dont la valeur est compact convexe», Funkcialaj Ekvacioj 10, pp. 205-223, 1967.

[12] Li S-G., «Fuzzy intervals», Applied Mathematics Letters, Vol. 14, No. 6, pp. 737-740, 2001.

[13] Moore R.E., «Interval Analysis», Prentice-Hall, Englewood Cliffs, NJ. 1966.

[14] Moore R.E., «Methodes and applications of interval analysis», SIAM, philadelphia, 1979.

[15] Moore R. and Lodwick W., «Interval analysis and fuzzy set theory», Fuzzy Sets and Systems, Vol. 135, No. 1, pp. 5-9, 2003.

[16] Mizumoto M. and Tanaka K., «The four operations of arithmetic on fuzzy numbers», Systems Comput. Controls, Vol. 7 (5), pp. 73-81, 1976.

[17] Mizumoto M. and Tanaka K., «Some properties of fuzzy numbers», in Gupta M.M., Ragade R.K. and Yager R.R. (Edts), Adavances in Fuzzy Sets theory and Applications, North-Holland, pp. 156-164, 1979, Amesterdam.

[18] Oussalah M. and De Schutter J., «Approximated fuzzy LR computation», Information Sciences, Volume 153, pp. 155-175, 2003.

[19] Piegat A., «Cardinality Approach to Fuzzy Number Arithmetic», IEEE Transactions on Fuzzy Systems, Vol. 13, N°. 2, pp. 204-215, 2005.

[20] Kaufmann A. and Gupta M.M., «Introduction to fuzzy arithmetic: Theory and Applications», Van Nostrand Reinhold Company Inc. Nex York, 1991.

[21] Klir G.J., «Fuzzy arithmetic with requisite constraints», Fuzzy Sets and Systems, Vol. 91, No. 2, pp. 165-175, 1997.

[22] Kulpa Z., «Diagrammatic representation for interval arithmetic», Linear Algebra and its Applications, Vol. 324, No. 1-3, pp. 55-80, 200.

[23] Lodwick W. A. and Jamison K.D., «Special issue: interfaces between fuzzy set theory and interval analysis», Fuzzy Sets and Systems, Vol. 135, No.1, pp. 1-3, 2003.

[24] Yager R.R., «On the lack of inverses in fuzzy arithmetic», Fuzzy Sets Syst. (1980) 73–82.

[25] Zhou C., «Fuzzy-arithmetic-based Lyapunov synthesis to the design of stable controllers: A computing with words approach», Appl. Math. Comput. Sci., Vol. 12, N°. 3, pp. 411-421, 2002.

[26] Zadeh L.A., «Fuzzy Sets», Information and Control, No. 8, pp. 338-353, 1965.

[27] Zadeh L.A., «Fuzzy sets as a basis for a theory of possibility», Fuzzy Sets and systems, N° 1, pp. 3-28, 1978.

Representing Incomplete Knowledge about Fuzzy Sets

Maciej Wygralak
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Umultowska 87, 61-614 Poznań, Poland
e-mail: wygralak@amu.edu.pl

Abstract

This paper deals with interval-valued fuzzy sets and I-fuzzy sets, Atanassov's intuitionistic fuzzy sets. They are treated in a systematic way as two, formally equivalent, natural extensions of methods of representing incomplete knowledge about sets. We define and investigate triangular norm-based areas of uncertainty of interval-valued fuzzy sets and I-fuzzy sets, and study some properties of related uncertainty degrees of their elements.

Keywords: Interval-valued fuzzy set, I-fuzzy set, Incomplete knowledge, Area of uncertainty.

1 Introduction

In recent years, there is a growing interest in and an intensification of research on two generalizations of the idea of a fuzzy set: interval-valued fuzzy sets and I-fuzzy sets. Both of them are tools for representing incompletely known fuzzy sets, i.e. tools for modeling a mixture of two important phenomena: imprecision and incompleteness of information.

Interval-valued fuzzy sets were introduced by Zadeh in 1975 ([17]). Surprisingly, they were also proposed in the same year by Sambuc [13] and by some other researchers; see [4] for further references. I-fuzzy sets were defined by Atanassov in 1983 under the name "intuitionistic fuzzy sets" ([1, 2]). Since there is a terminological conflict between his objects and "true" intuitionistic fuzzy sets, intuitionistic logic-based fuzzy sets by Takeuti and Titani [14], some alternative names have been proposed, e.g. "bipolar fuzzy sets", "IF-sets", and just "I-fuzzy sets" used in this paper; see [4] for details about that dispute. Speaking formally, interval-valued fuzzy sets and I-fuzzy sets are equivalent concepts. The idea of I-fuzzy sets, however, puts emphasis on positive and negative information. It seems that just this feature of Atanassov's objects is especially attractive and stimulates their applications in decision making and other areas (see e.g. [10]).

We like to look at interval-valued fuzzy sets and I-fuzzy sets in a systematic way as at two natural extensions of some methods referring to a simpler issue, namely representation of incompletely known sets. For both of these extensions, the notion of a triangular norm-based area of uncertainty and degree of uncertainty will be introduced and investigated.

2 Incomplete Knowledge about Sets and its Representation

Let us take a look at the issue of incomplete knowledge about a set $\boldsymbol{A} \subset \boldsymbol{M}$. The status of some elements $x \in \boldsymbol{M}$ (membership/nonmembership or fulfilment/nonfulfilment of a crisp property, in other words) is then practically unknown or uncertain. This implies a division of the universe $\boldsymbol{M}$ into the following three pairwise disjoint classes:

$\boldsymbol{A}^+$ - composed of those x's whose belonging to $\boldsymbol{A}$ is known (sure),

$\boldsymbol{A}^-$ - contains the x's that surely do not belong to $\boldsymbol{A}$,

$\boldsymbol{A}^?$ - the *area of uncertainty* of $\boldsymbol{A}$, composed of all x's with unknown (uncertain) status.

Since each of these classes is uniquely determined by the remaining two, incompleteness of knowledge about $\boldsymbol{A}$ can be modeled by specifying any of the pairs

$$(\boldsymbol{A}^+, \boldsymbol{A}^?),\ (\boldsymbol{A}^-, \boldsymbol{A}^?),\ (\boldsymbol{A}^+, \boldsymbol{A}^-)$$

of sets.

If the representation $(\boldsymbol{A}^+, \boldsymbol{A}^?)$ is chosen, it is convenient to rewrite it in the form $(\boldsymbol{A}^+, \boldsymbol{A}^+ \cup \boldsymbol{A}^?)$. So, speaking

more formally, incompleteness of knowledge about $\boldsymbol{A}$ can be expressed by means of a pair

$$(\boldsymbol{B}, \boldsymbol{C}) \quad \text{with} \quad \boldsymbol{B} \subset \boldsymbol{C}, \tag{2.1}$$

where the set $\boldsymbol{B}$ contains sure elements of $\boldsymbol{A}$, and $\boldsymbol{C}$ represents sure elements together with those with unknown or uncertain status. Then

$$\boldsymbol{A}^{?} = \boldsymbol{B}' \cap \boldsymbol{C}. \tag{2.2}$$

Recollect in the context of (2.1) that each nondecreasing, uncountable family of (cut) sets

$$(A_{1-t})_{t \in [0,1)}$$

is identical with a fuzzy set $A: \boldsymbol{M} \to [0,1]$ representing imprecise information, where $A_w = \{x: A(x) \geq w\}$. The two-element families from (2.1) are isomorphic to fuzzy sets $\boldsymbol{M} \to \{0, 1/2, 1\}$.

Pairs (2.1) were introduced by Gentilhomme in 1968 as *flou sets* ([6]; see also [9]). Again, independently and using different motivations, they were defined at the same time by Klaua as *partial sets* ([7]). Pawlak's *rough sets* from [11, 12] are also constructions collapsing in essence to (2.1). However, contrary to flou sets, their sums and intersections cannot be constructed by performing suitable set-theoretic operations on the components. Recall that a rough set

$$\Re_{\boldsymbol{A}} = (\underline{\boldsymbol{A}}, \overline{\boldsymbol{A}}) \tag{2.3}$$

is a pair of approximating sets in which

$$\underline{\boldsymbol{A}} = \{x \in \boldsymbol{M}: [x]_{\rho} \subset \boldsymbol{A}\}$$

is the *lower approximation* of $\boldsymbol{A}$, whereas

$$\overline{\boldsymbol{A}} = \{x \in \boldsymbol{M}: [x]_{\rho} \cap \boldsymbol{A} \neq \emptyset\}$$

forms its *upper approximation*. The equivalence relation ρ is treated as an indiscernibility relation of objects x described in terms of given attributes and their values. The area of uncertainty (2.2),

$$\underline{\boldsymbol{A}}' \cap \overline{\boldsymbol{A}},$$

is then called the *boundary* of $\boldsymbol{A}$ with respect to ρ.

An interesting generalization of the representation from (2.1) is the concept of a *twofold fuzzy set*

$$\mathbb{T}_{\boldsymbol{A}} = (S_{\boldsymbol{A}}, P_{\boldsymbol{A}}) \quad \text{with} \quad S_{\boldsymbol{A}} \subset (P_{\boldsymbol{A}})_1 \tag{2.4}$$

introduced in [5]. $S_{\boldsymbol{A}}$ is a fuzzy set of more or less sure elements of $\boldsymbol{A}$, whereas $P_{\boldsymbol{A}}$ forms a fuzzy set of its more or less possible elements. The inclusion in (2.4) implies $S_{\boldsymbol{A}} \subset P_{\boldsymbol{A}}$, and means that elements being sure to a positive degree must be fully possible. The corresponding area of uncertainty

$$S_{\boldsymbol{A}}' \cap P_{\boldsymbol{A}}$$

is interpreted as a fuzzy set of *dubious elements* of $\boldsymbol{A}$, i.e. elements whose belonging to $\boldsymbol{A}$ is possible but not sure.

Let us move on to the other two variants of representations of an incompletely known set $\boldsymbol{A}$. It seems that the form $(\boldsymbol{A}^{-}, \boldsymbol{A}^{?})$, dual in a way to $(\boldsymbol{A}^{+}, \boldsymbol{A}^{?})$ and focusing on negative information, is not used in the subject literature. Formally, it can be rewritten as (2.1), too, but the semantics of the components is then different: $\boldsymbol{B}$ is a set of impossible elements of $\boldsymbol{A}$, whereas $\boldsymbol{C}$ contains elements which are impossible or uncertain.

Much more interesting seems to be the third form, namely $(\boldsymbol{A}^{+}, \boldsymbol{A}^{-})$. It is bipolar by putting emphasis on positive $(\boldsymbol{A}^{+})$ and negative $(\boldsymbol{A}^{-})$ information, on positive and negative examples of elements of $\boldsymbol{A}$, in other words. This way of representing incomplete knowledge is exemplified by the idea of *sub-definite sets* by Narin'yani ([8]). In a more formalized notation, incomplete knowledge about $\boldsymbol{A}$ is now represented by a pair

$$(\boldsymbol{B}, \boldsymbol{D}) \quad \text{with} \quad \boldsymbol{B} \subset \boldsymbol{D}'. \tag{2.5}$$

$\boldsymbol{B}$ is a set of sure elements of $\boldsymbol{A}$, and $\boldsymbol{D}$ contains sure elements of $\boldsymbol{A}'$. We thus have

$$\boldsymbol{B} \subset \boldsymbol{A} \subset \boldsymbol{D}', \tag{2.6}$$

i.e. $\boldsymbol{B}$ and $\boldsymbol{D}'$, respectively, are a lower evaluation and an upper evaluation of $\boldsymbol{A}$. The set

$$\boldsymbol{A}^{?} = \boldsymbol{B}' \cap \boldsymbol{D}' \tag{2.7}$$

forms the area of uncertainty of $\boldsymbol{A}$.

3 Representing Incompletely Known Fuzzy Sets

Trying to model fuzzy sets with incompletely known membership grades, a reasonable and natural way of doing is to pattern oneself upon the representations of incompletely known sets from Section 2 with special reference to (2.1) and (2.5). And this idea finds its materialization in the concepts of interval-valued fuzzy sets and I-fuzzy sets (cf. also [15]). Let us look at them a bit closer drawing special attention to related areas of uncertainty.

3.1 Using Interval-Valued Fuzzy Sets

As one knows, an interval-valued fuzzy set is a pair $E = (A_l, A_u)$ of fuzzy sets $A_l, A_u: \boldsymbol{M} \to [0, 1]$ with

$$A_l \subset A_u.$$

A_l and A_u, respectively, form a lower bound and an upper bound of an incompletely known fuzzy set A, i.e. $A(x) \in [A_l(x), A_u(x)]$. What we deal with is thus a direct extension of (2.1) to fuzzy sets.

For a triangular norm $\boldsymbol{t}$ and a strong negation ν, let us define (cf. (2.2))

$$U_E = A_l^{\nu} \cap_{\boldsymbol{t}} A_u. \tag{3.1}$$

The fuzzy set U_E is the area of uncertainty of E, whereas

$$U_E(x) = \nu(A_l(x)) \,\boldsymbol{t}\, A_u(x) \tag{3.2}$$

is a *degree of uncertainty* concerned with x. Worth considering are properties of (3.2) for two important classes of t-norms. To this end, however, we need some additional notation. Denote by $\nu_{\boldsymbol{t}}$ the negation induced by $\boldsymbol{t}$, $\nu_{\boldsymbol{t}}(a) = \bigvee\{c: a \,\boldsymbol{t}\, c = 0\}$. On the other hand, $\boldsymbol{t}^{\nu}$ will denote the t-conorm which is ν-dual to $\boldsymbol{t}$, $a \,\boldsymbol{t}^{\nu}\, b = \nu(\nu(a) \,\boldsymbol{t}\, \nu(b))$.

If $\boldsymbol{t}$ is nonstrict Archimedean and $\nu = \nu_{\boldsymbol{t}}$, then $U_E(x)$ expresses our uncertainty as to membership degree of x in an incompletely known fuzzy set A represented by E. Notice that

$$\begin{aligned} &U_E(x) = 0 \;\Leftrightarrow\; A_l(x) = A_u(x), \\ \text{i.e.}\quad &\\ &U_E = 1_{\varnothing} \;\Leftrightarrow\; E \text{ is a fuzzy set.} \end{aligned} \tag{3.3}$$

Interpreting (3.2) becomes a more complex task whenever we use a strong negation ν and a t-norm $\boldsymbol{t}$ having no zero divisors, say, a strict $\boldsymbol{t}$ or $\boldsymbol{t} = \wedge$. This time $U_E(x)$ seems to be uncertainty as to membership degree of x in A combined with fuzziness of that degree. Indeed, we have

$$\begin{aligned} &U_E(x) = 0 \;\Leftrightarrow\; (A_l(x), A_u(x)) \in \{(0, 0), (1, 1)\}, \\ \text{i.e.}\quad &\\ &U_E = 1_{\varnothing} \;\Leftrightarrow\; E \text{ is a set.} \end{aligned} \tag{3.4}$$

If there is no uncertainty as to membership degrees in A, then E collapses to a fuzzy set, $E = (A, A)$, and

$$U_E(x) = \nu(A(x)) \,\boldsymbol{t}\, A(x). \tag{3.5}$$

So, $U_E(x)$ becomes an ordinary $\boldsymbol{t}$-based index of fuzziness of x.

Theorem 3.1. *Let $E = (A_l, A_u)$. If $\boldsymbol{t}$ is a nonstrict Archimedean t-norm, $\nu = \nu_{\boldsymbol{t}}$ and h is normed generator of $\boldsymbol{t}^{\nu}$, then*

$$\forall x \in \boldsymbol{M}:\; U_E(x) = h^{-1}(h(A_u(x)) - h(A_l(x))).$$

In the case of nonstrict Archimedean t-norms, uncertainty degrees $U_E(x)$ are thus isomorphic to the lengths of the intervals $[A_l(x), A_u(x)]$. Denote by $\nu_{Ł}$ and $\boldsymbol{t}_{Ł}$, respectively, the Łukasiewicz negation and the Łukasiewicz triangular norm, respectively, where $\nu_{Ł}(a) = 1 - a$ and $a \,\boldsymbol{t}_{Ł}\, b = 0 \vee (a + b - 1)$. Applying $\boldsymbol{t} = \boldsymbol{t}_{Ł}$, we get $\nu_{\boldsymbol{t}} = \nu_{Ł}$ and $h = id$, which gives

$$U_E(x) = A_u(x) - A_l(x).$$

More generally, one has

$$U_E(x) = ((A_u(x))^p - (A_l(x))^p)^{1/p}$$

when a Schweizer-Sklar t-norm $\boldsymbol{t}_{S,p}$ is used, where

$$a \,\boldsymbol{t}_{S,p}\, b = [0 \vee (a^p + b^p - 1)]^{1/p}, \quad p > 0.$$

Theorem 3.2. *Let $E = (A_l, A_u)$ be an interval-valued fuzzy set. One has*

$$\forall x \in \boldsymbol{M}:\; A_l(x) \,\boldsymbol{t}^{\nu}\, U_E(x) = A_u(x)$$

if $\boldsymbol{t}$ is nonstrict Archimedean and $\nu = \nu_{\boldsymbol{t}}$. For $\boldsymbol{t} = \wedge$ with any strong negation ν, we get

$$\forall x \in \boldsymbol{M}:\; A_l(x) \,\boldsymbol{t}^{\nu}\, U_E(x) \le A_u(x).$$

For $\boldsymbol{t} = \boldsymbol{t}_{Ł}$, the equality in Theorem 3.2 collapses to the standard

$$A_l(x) + U_E(x) = A_u(x).$$

There is no general relationship between $A_l(x) \,\boldsymbol{t}^{\nu}\, U_E(x)$ and $A_u(x)$ and if a strict $\boldsymbol{t}$ is used.

3.2 Using I-Fuzzy Sets

By an *I-fuzzy set* one means a pair $\mathscr{E} = (A^+, A^-)$ of fuzzy sets $A^+, A^-: \boldsymbol{M} \to [0, 1]$ such that

$$A^+ \subset (A^-)^{\nu}, \tag{3.6}$$

where ν denotes a strong negation. This notion is thus a direct extension of (2.5). By virtue of (3.6), each I-fuzzy set is formally equivalent to an interval-valued fuzzy set. However, some care as to that equivalence is necessary if triangular norm-based operations are taken into account. Basic relations for and operations on I-fuzzy

sets $\mathscr{E}=(A^+, A^-)$ and $\mathscr{F}=(B^+, B^-)$ with (3.6) are then defined as follows:

$$\begin{aligned}
&\mathscr{E}=\mathscr{F} \;\Leftrightarrow\; A^+=B^+ \;\&\; A^-=B^-,\\
&\mathscr{E}\subset\mathscr{F} \;\Leftrightarrow\; A^+\subset B^+ \;\&\; B^-\subset A^-,\\
&\mathscr{E}\cup_{\boldsymbol{t},\boldsymbol{s}}\mathscr{F}=(A^+\cup_{\boldsymbol{s}} B^+, A^-\cap_{\boldsymbol{t}} B^-),\\
&\mathscr{E}\cap_{\boldsymbol{t},\boldsymbol{s}}\mathscr{F}=(A^+\cap_{\boldsymbol{t}} B^+, A^-\cup_{\boldsymbol{s}} B^-),\\
&\mathscr{E}'=(A^-, A^+),
\end{aligned} \tag{3.7}$$

where $\boldsymbol{t}$ and $\boldsymbol{s}$, resp., denote a t-norm and a t-conorm, $(F\cap_{\boldsymbol{t}} G)(x)=F(x)\,\boldsymbol{t}\,G(x), (F\cup_{\boldsymbol{s}} G)(x)=F(x)\,\boldsymbol{s}\,G(x)$. By the way, $\boldsymbol{t}$ and $\boldsymbol{s}$ can be, respectively, replaced by an interval-valued (non-representable) t-norm and t-conorm from [3]. There are good reasons for taking $\boldsymbol{s}=\boldsymbol{t}^{\nu}$ in (3.7) (see [10, 16] for details). Indeed, just the use of $\boldsymbol{s}=\boldsymbol{t}^{\nu}$ guarantees that sums and intersections of I-fuzzy sets coincide with $\boldsymbol{s}$-based sums and $\boldsymbol{t}$-based intersections of interval-valued fuzzy sets corresponding to them, where

$$(A_l, A_u)\cup_{\boldsymbol{s}}(B_l, B_u)=(A_l\cup_{\boldsymbol{s}} B_l, A_u\cup_{\boldsymbol{s}} B_u)$$

and

$$(A_l, A_u)\cap_{\boldsymbol{t}}(B_l, B_u)=(A_l\cap_{\boldsymbol{t}} B_l, A_u\cap_{\boldsymbol{t}} B_u).$$

What is more, sums and intersections of I-fuzzy sets collapsing to ordinary fuzzy sets are then (and only then) still fuzzy sets expressed as I-fuzzy sets. Notice that, on the other hand, the very idea of interval-valued fuzzy sets with $\boldsymbol{s}$-based sums and $\boldsymbol{t}$-based intersections does not require the restriction $\boldsymbol{s}=\boldsymbol{t}^{\nu}$ in this context.

Unlike interval-valued fuzzy sets, the idea of I-fuzzy sets puts emphasis on positive (A^+) and negative (A^-) information, on positive and negative examples of elements having an imprecise property. $\mathscr{E}=(A^+, A^-)$ represents incomplete knowledge about a fuzzy set A and its complement A^{ν} which are such that (cf. (2.6))

$$A^+\subset A\subset(A^-)^{\nu} \;\text{ and, hence, }\; A^-\subset A^{\nu}\subset(A^+)^{\nu}.$$

So, A^+ is a lower bound on A, A^- forms a lower bound on A^{ν},

$$A(x)\in[A^+(x), \nu(A^-(x))]$$

and

$$A^{\nu}(x)\in[A^-(x), \nu(A^+(x))].$$

In the original commonly used Atanassov's terminology from [1, 2], A^+ is understood as a *membership function*, whereas A^- is interpreted as a *nonmembership function*. So, $A^+(x)$ forms a *membership degree* of x, and $A^-(x)$ is a *nonmembership degree*. Condition (3.6) guarantees that

$$A^+(x)\le a^* \;\text{ or }\; A^-(x)\le a^*$$

for each x with a^* denoting a unique fixed point of the negation ν. In other words, the membership and the nonmembership degrees of x must not be "too large" at the same time.

Drawing inspiration from Pawlak's rough set theory, the following two types of approximate equalities can be defined on I-fuzzy sets:

$$\mathscr{E}=_{pos}\mathscr{F}\Leftrightarrow A^+=B^+,$$

($\mathscr{E}$ *positively equal to* $\mathscr{F}$)

$$\mathscr{E}=_{neg}\mathscr{F}\Leftrightarrow A^-=B^-.$$

($\mathscr{E}$ *negatively equal to* $\mathscr{F}$)

Both $=_{pos}$ and $=_{neg}$ are obviously equivalence relations, and we have

$$\mathscr{E}=\mathscr{F} \;\Leftrightarrow\; \mathscr{E}=_{pos}\mathscr{F} \;\&\; \mathscr{E}=_{neg}\mathscr{F}.$$

Similarly, approximate inclusions

$$\mathscr{E}\subset_{pos}\mathscr{F}\Leftrightarrow A^+\subset B^+$$

($\mathscr{E}$ *positively contained in* $\mathscr{F}$)

and

$$\mathscr{E}\subset_{neg}\mathscr{F}\Leftrightarrow B^-\subset A^-$$

($\mathscr{E}$ *negatively contained in* $\mathscr{F}$)

can be introduced as two partial order relations.

3.3 Uncertainty Areas of I-Fuzzy Sets

The uncertainty area of an I-fuzzy set $\mathscr{E}=(A^+, A^-)$ is usually called the *hesitation area* of $\mathscr{E}$, and is defined as ([10, 16])

$$\chi_{\mathscr{E}}=(A^+)^{\nu}\cap_{\boldsymbol{t}}(A^-)^{\nu}=(A^+\cup_{\boldsymbol{t}^{\nu}} A^-)^{\nu} \tag{3.8}$$

with a t-norm $\boldsymbol{t}$ and a strong negation ν (see (2.7)). Its membership degree

$$\chi_{\mathscr{E}}(x)=\nu(A^+(x)\;\boldsymbol{t}^{\nu}\,A^-(x)) \tag{3.9}$$

is interpreted as a *degree of hesitation* concerned with x. Notice that

$$\forall x\in\boldsymbol{M}:\; \phi(A^+(x)\;\boldsymbol{t}^{\nu}\,A^-(x))+\phi(\chi_{\mathscr{E}}(x))=1 \tag{3.10}$$

with ϕ being an automorphism of $[0,1]$ generating ν, $\nu(a)=\phi^{-1}(1-\phi(a))$.

The original Atanassov's formulation of I-fuzzy sets from [1, 2] is a particular case of the presented general approach when $\cup_{\wedge,\vee}$ and $\cap_{\wedge,\vee}$ are used in (3.7) as basic operations, whereas $\boldsymbol{t}=\boldsymbol{t}_{Ł}$ and $\nu=\nu_{Ł}$ are employed in (3.6) and (3.8). Indeed, (3.6), (3.9) and (3.10), respectively, then collapse to (see also Theorem 3.3)

$$A^+(x) + A^-(x) \le 1,$$

$$\chi_{\mathscr{E}}(x) = 1 - A^+(x) - A^-(x),$$

$$A^+(x) + A^-(x) + \chi_{\mathscr{E}}(x) = 1$$

for each $x \in \boldsymbol{M}$.

Let us restrict our further discussion to Archimedean t-norms, including $\boldsymbol{t} = \wedge$. First, we present a few additional examples of hesitation degrees. For $\boldsymbol{t} = \boldsymbol{t}_{S,p}$ with $\nu = \nu_{\boldsymbol{t}}$, the basic condition (3.6) collapses to

$$\forall x \in \boldsymbol{M}\colon (A^+(x))^p + (A^-(x))^p \le 1,$$

and we obtain

$$\chi_{\mathscr{E}}(x) = [1 - (A^+(x))^p - (A^-(x))^p]^{1/p}.$$

Using the product t-norm with $\nu = \nu_{Ł}$, one gets

$$\chi_{\mathscr{E}}(x) = 1 - A^+(x) - A^-(x) + A^+(x) \cdot A^-(x),$$

whereas $\boldsymbol{t} = \wedge$ gives

$$\chi_{\mathscr{E}}(x) = 1 - A^+(x) \vee A^-(x).$$

Similarly to (3.3), if $\boldsymbol{t}$ is nonstrict Archimedean and $\nu = \nu_{\boldsymbol{t}}$, then $\chi_{\mathscr{E}}(x)$ expresses our hesitation as to membership degree of x in an incompletely known fuzzy set A modeled by $\mathscr{E}$. One has

$$\chi_{\mathscr{E}}(x) = 0 \;\Leftrightarrow\; A^+(x) = \nu(A^-(x))$$

for each x, which leads to

$$\chi_{\mathscr{E}} = 1_{\emptyset} \;\Leftrightarrow\; \mathscr{E} \text{ is a fuzzy set.}$$

As a corollary from Theorem 3.1, we have

Theorem 3.3. *Let $\mathscr{E} = (A^+, A^-)$. If $\boldsymbol{t}$ is nonstrict Archimedean, $\nu = \nu_{\boldsymbol{t}}$, and h is normed generator of $\boldsymbol{t}^{\nu}$, then*

$$\forall x \in \boldsymbol{M}\colon \chi_{\mathscr{E}}(x) = h^{-1}(1 - h(A^+(x)) - h(A^-(x))).$$

Again, applying a strong negation ν with a strict t-norm $\boldsymbol{t}$ or $\boldsymbol{t} = \wedge$, $\chi_{\mathscr{E}}(x)$ becomes hesitation as to membership degree of x combined with fuzziness of that degree, namely (cf. (3.4), (3.5))

$$\chi_{\mathscr{E}}(x) = 0 \Leftrightarrow (A^+(x), A^-(x)) \in \{(1, 0), (0, 1)\},$$

i.e.

$$\chi_{\mathscr{E}} = 1_{\emptyset} \Leftrightarrow \mathscr{E} \text{ is a set.}$$

Further, if $\mathscr{E}$ collapses to a fuzzy set, $\mathscr{E} = (A, A^{\nu})$, and thus there is no hesitation as to membership degrees in A, we have

$$\chi_{\mathscr{E}}(x) = A(x)\ \boldsymbol{t}\ \nu(A(x)),$$

i.e. $\chi_{\mathscr{E}}(x)$ forms an ordinary $\boldsymbol{t}$-based index of fuzziness of x.

Theorem 3.4. *Let $\mathscr{E} = (A^+, A^-)$. If $\nu = \nu_{\boldsymbol{t}}$ with a nonstrict Archimedean t-norm $\boldsymbol{t}$, then*

$$A^+ \cup_{\boldsymbol{t}^{\nu}} \chi_{\mathscr{E}} = (A^-)^{\nu}.$$

For $\boldsymbol{t} = \wedge$ with any strong negation ν,

$$A^+ \cup_{\boldsymbol{t}^{\nu}} \chi_{\mathscr{E}} \subset (A^-)^{\nu}.$$

This analogue of Theorem 3.2 refers to the identity $\boldsymbol{A}^+ \cup \boldsymbol{A}^? = (\boldsymbol{A}^-)'$ for incompletely known sets. As previously, if a strict t-norm is used, there is no general relationship between $A^+ \cup_{\boldsymbol{t}^{\nu}} \chi_{\mathscr{E}}$ and $(A^-)^{\nu}$.

Finally, we like to present an illustrative example of applications of triangular norm-based I-fuzzy sets and their hesitation areas in GDM, group decision making. Let $\boldsymbol{P} = \{p_1, \ldots, p_m\}$ be a set of $m \ge 1$ individuals, and let $\boldsymbol{S} = \{s_1, \ldots, s_n\}$ denote a set of $n \ge 2$ options (decision alternatives). Each individual p_k, $k = 1, \ldots, m$, formulates his/her individual preferences over $\boldsymbol{S}$ by means of a binary fuzzy relation R_k in $\boldsymbol{S}$. R_k can be presented in the form of an $n \times n$ matrix

$$R_k = [\, r_{ij}^k \,]$$

in which $r_{ij}^k \in [0, 1]$ is a *degree to which p_k prefers s_i to s_j*; one puts $r_{ii}^k = 0$. Intuitively speaking, r_{ij}^k and r_{ji}^k cannot be "too large" at the same time and, thus, we assume that

$$r_{ij}^k \le \nu(\, r_{ji}^k)$$

with a strong negation ν (cf. (3.6)). The number

$$h_{ij}^k = \nu(r_{ij}^k)\ \boldsymbol{t}\ \nu(r_{ji}^k)$$

with $i \ne j$ and a t-norm $\boldsymbol{t}$ expresses a *hesitation of p_k* as to his/her preference concerning s_i and s_j (cf. (3.9)); additionally, one defines $h_{ii}^k = 0$. This formal model is then a starting point for constructing a family of general algorithms of GDM, including the direct and indirect approach. In the direct case, they generate a solution being a fuzzy set S_Q of options such that a soft majority, Q individuals, is not against them, where Q denotes a "most"-type relative linguistic quantifier. Using the indirect approach, first the individual preference matrices $R_1, \ldots, R_m$ are aggregated into a social preference $n \times n$ matrix R representing common preferences of the whole group $\boldsymbol{P}$ over $\boldsymbol{S}$. Second, on the basis of R, one derives a solution W_Q understood in this case as a fuzzy set of

options being preferred by the group over Q other options from $\boldsymbol{S}$. The reader is referred to [10] for a detailed presentation and discussion of this issue.

4 Concluding Remarks

In this paper, first, we have dealt with the issue of sets with incomplete knowledge for which the status of some elements is practically unknown or uncertain. Three mathematically equivalent methods of representing that incompleteness and related uncertainty are then possible. Further, we have considered interval-valued fuzzy sets and I-fuzzy sets, Atanassov's intuitionistic fuzzy sets, with triangular norms. Those objects were treated as two formally equivalent tools for representing incomplete knowledge about fuzzy sets which are natural extensions of the methods of representing incompletely known sets. We have defined and investigated triangular norm-based uncertainty areas of interval-valued fuzzy sets and I-fuzzy sets. In general, their properties and semantics depend on the t-norm being used. In the case of a t-norm without zero divisors, an interpretation involving fuzziness degrees seems to be appropriate.

References

[1] K. T. Atanassov, Intuitionistic Fuzzy Sets - Theory and Applications, Physica-Verlag, Heidelberg, 1999.

[2] K. T. Atanassov, S. Stoeva, Intuitionistic fuzzy sets, in: *Proc. Polish Symp. on Interval and Fuzzy Mathematics*, Poznań, Poland, 1983, pp. 23-26.

[3] G. Deschrijver, C. Cornelis, E. E. Kerre, On the representation of intuitionistic fuzzy t-norms and t-conorms, IEEE Trans. on Fuzzy Systems 12 (2004) 45-61.

[4] D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk, H. Prade, Terminological difficulties in fuzzy set theory – The case of "Intuitionistic Fuzzy Sets", Fuzzy Sets and Systems 156 (2005) 485-491.

[5] D. Dubois, H. Prade, Twofold fuzzy sets and rough sets - Some issues in knowledge representation, Fuzzy Sets and Systems 23 (1987) 3-18.

[6] Y. Gentilhomme, Les ensembles flous en linguistique, Cahiers dé Linguistique Theorique et Appliquée 5 (1968) 47.

[7] D. Klaua, Partiell definierte Mengen, Deut. Akad. Wiss. Berlin 10 (1968) 571-578.

[8] A. S. Narin'yani, Sub-definite sets - New data type for knowledge representation, Memo no. 4-232, Computing Center, Novosibirsk, 1980 (in Russian).

[9] C. V. Negoita, D. A. Ralescu, L-fuzzy sets and L-flou sets, Elektron. Informationsverarb. und Kybernet. 12 (1976) 599-605.

[10] A. Pankowska, M. Wygralak, General IF-sets with triangular norms and their applications to group decision making, Information Sciences 176 (2006) 2713-2754.

[11] Z. Pawlak, Rough sets. Basic notions, ICS PAS Reports, vol. 431, Warsaw, 1981.

[12] Z. Pawlak, Rough Sets - Theoretical Aspects of Reasoning About Data, Kluwer, Dordrecht, 1991.

[13] R. Sambuc, Fonctions Φ-Floues – Application á l'Aide au Diagnostic en Pathologie Thyroïdienne, Ph. D Thesis, Univ. Marseille, France, 1975.

[14] G. Takeuti, S. Titani, Intuitionistic fuzzy logic and intuitionistic fuzzy set theory, J. Symb. Logic 49 (1984) 851-866.

[15] M. Wygralak, Vaguely Defined Objects. Representations, Fuzzy Sets and Nonclassical Cardinality Theory, Kluwer, Dordrecht, 1996.

[16] M. Wygralak, I-fuzzy Sets with Triangular Norms, their Hesitation Areas and Cardinalities, in: K. T. Atanassov, J. Kacprzyk, M. Krawczak, E. Szmidt (Eds.), *Issues in the Representation and Processing of Uncertain and Imprecise Information. Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets, and Related Topics*. EXIT Publ., Warsaw, 2005, pp. 397-408.

[17] L. A. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning I, Information Sciences 8 (1975) 199-249.

Session 9

Decision Making 1

Two consensus protocols based on an acceptance threshold in Group Decision Making

Christophe Labreuche
THALES Research and Technology
RD 128, 91767 Palaiseau Cedex, France
christophe.labreuche@thalesgroup.com

Abstract

We define two negotiation protocols for Group Decision Making, which main feature is the existence of acceptation thresholds. In order to predict which consensus are expected to arise, we study the equilibria of these protocols in the sense of Game Theory. We also investigate their sensitivity to manipulation by the actors.

Keywords: Group Decision Making, Preferences, Game Theory, Equilibria, Manipulation.

1 Introduction

By Group Decision Making (GDM in short) we mean a set of decision makers (DM) that have jointly to single out one option among several from the preferences of each DM. These preferences are assumed to be represented by a cardinal utility function.

A first class of approaches to previous problem consists in defining an automatic procedure, also called mechanism, in which the DMs do not participate directly. The most well-known ones are derived from voting rules, using Fuzzy Set Theory [10]. A GDM problem is indeed similar to voting problems, except that the preference representation is not cardinal in these latters [1]. One can also find extensions of some concepts of Cooperative Game Theory, such as the Core or the bargaining set, to GDM problems [17]. Whatever the solution that is proposed, there is always the possibility that a DM complains about the outcome of the mechanism. This leads to the second class of approaches where the idea is to let the DMs discuss together and reach by themselves a consensus. The GDM community has mainly focuses on defining measures of soft consensus [11], on understanding the reason why individual preferences differ and on providing recommendations to the DMs who are the most distant to the consensual position, of how to change their preferences to reduce the overall disagreement [16]. Another reason why the preferences of the individuals can vary over time is the fact that each agent is influenced by the other ones. From a theoretical standpoint, one can assume that such influence is known and modeled with, for instance De Groot's linear model [3] or the soft consensus non linear model [6]. Consensual dynamics consists then in predicting what will be the result of such influences.

In the Multi-Agent and Artificial Intelligence communities, the problem of finding a consensus is seen as a bargaining problem where the consensus results from a negotiation *protocol* [14, 9, 5]. This approach has not received much attention in the GDM community [4]. The goal of this paper is to study a type of protocol that tackle the GDM problem.

A consensus protocol is an iterative process that formalizes the exchanges that are allowed among the DMs. At each run of a protocol, one proposal is made by one DM or a group of DMs, the other DMs vote on the acceptation or rejection of that offer, and the process goes on if the offer is rejected.

We consider two protocols in this paper. In these two protocols, the main feature is the existence of an acceptance threshold such that any offer more attractive for a DM than that threshold must be accepted by him. This threshold decreases over time, compelling the DMs to accept more and more offers over time. The only difference between these two protocols is that individual DMs are actors of the protocol in the first one whereas coalitions of DMs are actors in the second protocol.

It is not easy to see whether a protocol is a priori expected to lead to fair consensual options. Non-Cooperative Game Theory is a powerful tool to perform such prediction under rationality assumptions. One indeed looks for the *equilibria* of the protocols. Interestingly, the equilibria for the first protocol lead to

the usual maximin solution, whereas the second protocol leads to a weak version of the counterpart in GDM problems of the Nucleolus known in Cooperative Game Theory [13].

Another important issue often studied for voting rules is the possibility for the DMs to manipulate the outcome of the consensus by announcing preferences that are not their true ones [7, 15, 12]. Our two protocols are easily manipulable. The origin comes from the cardinal nature of the preference representation. Two possible proposals to reduce this sensitivity are made.

The rest of the paper is organized as follows. Section 2 gives some preliminaries. The first protocol is described and analyzed in Section 3. The sensitivity of this protocol to manipulations is discussed in Section 4. Section 5 depicts the second protocol.

2 Description of the problem

Let N be the set of DMs and A the set of options. A is supposed to be finite. We assume that the preferences of DM i can be expressed as a utility function u_i such that the larger the better for DM i. Without loss of generality, utilities can be assumed to belong to the $[0,1]$ interval. The set of functions from A onto $[0,1]$ is denoted by $\mathcal{U}(A)$. A consensus mechanism is a function $M : \mathcal{U}(A)^n \to A$.

It is not easy to predict the behavior of the DMs face to a negotiation protocol. *Subgame Perfect Equilibria* (SPE) [13] give realistic predictions under the assumption that DMs behave rationally and that they have perfect information on all preferences. The difficulty comes from the fact that, during the protocol, a DM cannot readily assess the consequences of his choices since he does not know the choice of the next DMs in the protocol. The only case where this is simple is for the last DM to give a decision. The second last DM can then assess the consequences of his choices, assuming that the last DM considers his best choices. And so on. This backward induction algorithm defines SPEs [13].

3 Description and analysis of the first protocol

In a negotiation protocol, a DM may stick to his initial preferences without taking into account the arguments of the other players. As a result, the protocol never converges to an agreement. In the literature, it is often assumed that options become less and less attractive over time so that early consensus are encouraged [14]. However, assuming that the utility of DMs for an option decays with the time is not always valid or may be complex to represent in practice. Another possibility is to force DMs to accept proposals for which their utility is larger than a given threshold. This threshold is large at the beginning of the protocol so that players have the complete freedom to refuse, and the threshold decreases with time, enforcing players to accept more and more potential proposals, and thus to make more and more concession. This reduction of the freedom of the DMs is motivated by fairness (it is not fair that a player accepts a not so good option to him and that another player rejects an option with a similar utility) and cooperation (make sure to converge to an agreement) reasons.

We assume that the order in which DMs are proposers is defined beforehand. More precisely, DM τ_k is the proposer at iteration k. We introduce the following protocol.

Protocol ***P1.*** At iteration $k \in \mathbb{N}_*$:

- Player τ_k makes an offer $a \in A$.
- All other players say *Yes* or *No*. However, a player $i \in N \setminus \{\tau_k\}$ must necessarily accept if $u_i(a) \geq \rho_i(k)$.
- If at least one player rejects the offer, we go to the next iteration. Otherwise, a is chosen and the protocol ends.

The option that is chosen is called *outcome* of the protocol. We introduce the following condition.

Axiom **Thresholds (T)**. For all $i \in N$, $\rho_i(k)$ is non-increasing w.r.t. k and reaches value 0 in a finite time $T^\star$ (i.e. $\rho_i(T^\star) = 0$).

Thanks to that axiom, after a given time (that is large enough), any proposal is necessarily accepted by all players, so that ***P1*** ends after a finite time.

Lemma 1 *Under* **(T)**, *protocol* ***P1*** *has at least one SPE.*

Let $A^\star(k) = \{a \in A\,,\ \forall i \in N\ \ u_i(a) \geq \rho(k)\}$. At an iteration k, any offer in $A^\star(k)$ is necessarily accepted as the outcome. By **(T)**

$$A^\star(1) \subseteq A^\star(2) \subseteq \cdots \subseteq A^\star(T^\star) = A\ .$$

Lemma 2 *Assume that* **(T)** *holds. Let K be the smallest integer such that $A^\star(K) \neq \emptyset$. Assume furthermore that all players are proposer at least once at or after iteration K. Let K' be the smallest integer larger than K such that*

$$\{\tau_K, \tau_{K+1}, \ldots, \tau_{K'}\} = N\ .$$

Then the resulting outcome of any SPE of protocol ***P1*** *belongs to $A^\star(K')$.*

Player τ_K that is proposer at iteration K is the first player that can be sure to get an acceptable offer. However the options of $A^\star(K)$ may not be very attractive to him so that he can make an unacceptable offer and hope that a better option will be proposed at a further iteration.

The following result shows that we can restrict ourself to SPEs that ends at iteration K'. The other SPEs use unnecessary extra iterations.

Corollary 3 *Under the assumptions of Lemma 2, for any SPE of protocol **P1**, there exists a SPE of protocol **P1** that ends at an iteration in $\{1,\dots,K'\}$ and that leads to the same outcome.*

By Corollary 3, it is enough for restrict to iteration in $\{K,\dots,K'\}$, in the computation of the SPE. We define the following algorithm.

Algorithm ***Algo4***. Let us define $R_k \subseteq A$ on induction of $k \in \{1,\dots,K'\}$ as follows.

- Set $R_{K'} = \{a \in A^\star(K')\,,\ u_{\tau_{K'}}(a) = \max_{b\in A^\star(K')} u_{\tau_{K'}}(b)\}$
- For k iterating from $K'-1$ downwards to 1,
  - For any $a \in A$ and $a^\star \in R_{k+1}$, define
$$F_k(a,a^\star) = \begin{cases} a & \text{if } \forall i \in A\setminus\{\tau_k\} \text{ either} \\ & u_i(a) \geq \rho(k) \text{ or } u_i(a) \geq u_i(a^\star) \\ a^\star & \text{otherwise} \end{cases}$$
  - Set
$$R_k = \bigcup_{a^\star\in R_{k+1}} \{F_k(a,a^\star)\ ,\ a \in A \text{ and } u_{\tau_k}(F_k(a,a^\star)) = \max_{b\in A} u_{\tau_k}(F_k(b,a^\star))\}$$

The set of options resulting from that algorithm is R_K.

Algorithm ***Algo4*** has a polynomial complexity. More precisely, the maximal number of operations that Algorithm ***Algo4*** can perform is of order $K' p^2 (2n+p)$.

Proposition 4 *Under the assumptions of Lemma 2, R_K resulting from Algorithm **Algo4** is the set of possible outcomes of the SPEs of protocol **P1**.*

The following axiom assumes that any player is periodically a proposer.

Axiom **Periodical Proposers (PP)**.
There exists $T \in \mathbb{N}$ such that for any $k \in \mathbb{N}_*$ and any $i \in N$, there exists $j \in \{k, k+1,\dots,k+T\}$ such that $\tau_j = i$.

Let $A^\star = \left\{a \in A\,,\ \min_{i\in N} u_i(a) = \max_{b\in A}\min_{i\in N} u_i(b)\right\}$ be the maximin solution.

We consider now families $\{\rho_i^\delta\}_\delta$ of functions ρ_i parametrized by a value $\delta \in (0,1)$ such that

$$\forall k \in \mathbb{N}\ ,\quad \left|\rho_i^\delta(k) - \rho_i^\delta(k+1)\right| \leq c\,\delta \tag{1}$$

For these functions, one can attain any arbitrarily small differences between two successive thresholds. Note that the upper bound for the maximal time of an end to the protocol depends on δ.

We show that under previous condition, equilibria correspond to the maximin solutions.

Proposition 5 *Let $\{\rho_i^\delta\}_\delta$ be families of functions satisfying (1), **(PP)** and **(T)** such that $\rho_1^\delta = \cdots = \rho_m^\delta$ and $\rho_1^\delta(0) = \cdots = \rho_m^\delta(0) = 1$. Then there exists $\Lambda \in (0,1)$ such that for any $\delta \in (0,\Lambda)$, the SPEs of protocol **P1** are elements of $A^\star$.*

The idea of the proof is that if δ is small enough, then all players will be proposers at least once while ρ^δ lies in-between $\underline{u}(A\setminus A^\star)$ and $\underline{u}(A^\star)$. Before those iterations, there is always at least one player that is not forced to accept a proposal. For those iterations, the only proposal for which all players are forced to accept are $A^\star$. After those iterations, more options can be necessarily accepted but such options are disadvantageous for at least one player. This player could have avoided getting this bad option by just proposing any element of $A^\star$ during his turn in previous iterations.

4 Strategy proofness of Protocol *P1*

The Gibbard-Satterthwaite theorem [7, 15] shows that there is no democratic voting rule that is not manipulable. A voting rule is manipulable when the outcome of the vote is better to a DM if he announces preferences that are not his true ones, compared to when he reveals his true preferences. This definition of manipulation can be extended to consensus mechanisms.

Definition 6 *A consensus mechanism M is* subject to manipulation *if*

$$\exists i \in N\ ,\ \exists u_i' \in \mathcal{U}(A) \quad \text{such that} \quad u_i\left(M(u_i', u_{-i}\right) > u_i\left(M(u_1,\dots,u_n)\right)$$

In this definition u_i represents the true utility of i, and u_i' is what DM i announces.

If $a^\star$ is the best option of player i (according to his true utility u_i), and if player i announces the following utility

$$\tilde{u}_i(a) = \begin{cases} \varepsilon & \text{if } a \neq a^\star \\ 1 & \text{if } a = a^\star \end{cases} \tag{2}$$

then for ε small enough, $a^\star$ is the outcomes of all SPE of protocol ***P2***. As a result, protocol ***P2*** is subject very much to manipulation. An easy manipulation for a DM consists thus to over-estimate his best options and under-estimate the other ones. This strategy does not require to know the private information of the other DMs.

This drawback may be see as acceptable in the context of cooperation. However, even if the DMs accept to cooperate, they may wish to adopt a strategy that will favor their personal interest. So they may exaggerate their preferences or change them a little bit.

4.1 Detecting exaggerations

Protocol ***P1*** assumes that utility functions u_i represent the true preferences of the DMs and that these utility functions are *commensurate* among the DMs. This implies that a score say 0.6 given by one DM has the same meaning that the same score 0.6 for another DM. This assumption is very hard to obtain. Multi-Criteria Decision Analysis and Measurement Theory provide a solution to this problem [2, 8]. The idea is to construct utility function u_i as an interval scale. This is a numerical representation of the preferences of the DM in which the notion of difference makes sense. This means that the difference of satisfaction between options a and b is $\frac{u_i(b)-u_i(a)}{u_i(d)-u_i(c)}$ times more important than the difference of satisfaction between options c and d. An interval scale is given up to an affine transformation. The idea of the Macbeth approach is then to identify in A two references elements and to fix entirely the interval scale by fixing the utility for these two particular elements. These two elements are chosen so that they have the same meaning throughout the DM. The first element corresponds to an option $\mathbf{0}_i \in A$ that is considered as completely unacceptable to DM i. The second element is an option $\mathbf{1}_i \in A$ that is considered as completely satisfactory to DM i. Elements $\mathbf{0}_1, \ldots, \mathbf{0}_n$ have the same meaning, and so do elements $\mathbf{1}_1, \ldots, \mathbf{1}_n$. We furthermore set $u_i(\mathbf{0}_i) = 0$ and $u_i(\mathbf{1}_i) = 1$ for all $i \in N$.

Previous methodology is designed to construct utility functions that really represent the preferences of the DM and to ensure commensurateness between the DMs. Note that it is possible that each utility function u_i is constructed from multiple criteria [2, 8].

It is always possible that a DM has sheeted during the elicitation of his utility function. This means that the utility function he obtains is somehow in-between his true one and the one described in (2). We want to stress the fact that it is possible to some extend to guess whether a DM has exaggerated his preferences.

An easy way to exaggerate his preferences is to reduce the support of the utility function, making it sharper. Let us show this in a simple case. Assume that the options are described by one attribute and are ranked according to this attribute in the following order $a_1, a_2, \ldots, a_p$. We assume furthermore that the options are uniformly spread over the values of the attribute. Then $\Delta_l u_i := u_i(a_{l+1}) - u_i(a_l)$ is proportional to the derivative of u_i at the value of the attribute corresponding to a_l. According to previous remark, the more DM i exaggerates his preferences, the sharper u_i in some parts, and thus the less uniform $\Delta_l u_i$ over l. It appears thus that a measure of uniformity of $\Delta_l u_i$ can assess the degree of exaggeration of DM i. Since the Shanon entropy is classically used to measure such uniformity, we make the following definition

$$\mathcal{W}_i := \sum_{l \in \{1,\ldots,p-1\}} -\Delta_l u_i \, \log(\Delta_l u_i) \; .$$

Comparing the value of $\mathcal{W}_i$ for all DMs, one can detect the more exaggerating ones, i.e. the ones with the smallest values of $\mathcal{W}_i$. They can be asked to revise their utility function, in the spirit of [18, 4].

4.2 On the manipulation of protocol *P1*

We have seen that when commensurateness among DMs is assumed, the protocol is very sensitive to manipulation. It is not so easy to renormalize cardinal utility functions in order to obtain this property. In (2), a DM can exaggerate the difference of preferences and thus of utility between two options. One way to get rid of this problem is to construct a utility function that is not based on difference of intensities in the preferences but just on an ordinal information which is the ranking of the options according to a DM.

More precisely, we consider an isomorphism τ_i from A onto the integer set $\{1, \ldots, p\}$. The set of all isomorphisms from A onto the integer set $\{1, \ldots, p\}$ is denoted by $\mathcal{I}(A)$. τ_i is assumed to represent the ordinal preferences of DM i over A. This representation is valid iff the preference relation of DM i is a total order (i.e. having no ex equo). Then $a \in A$ is preferred to $b \in A$ according to DM $i \in N$ if $\tau_i(a) > \tau_i(b)$. Then we define $u_i(a) := \frac{1}{p}\, \tau_i(a)$. Set $\tau = (\tau_1, \ldots, \tau_n) \in \mathcal{I}(A)^n$.

Remark that τ cannot easily be determined when A contains many elements, for instance when A is a combinatorial space.

Each DM has thus to define the ordering τ_i he wants to be applied in the game. The overall problem can thus be seen as a two-stage game. Firstly each DM has to provide his ranking on the set A of options. This corresponds to a distortion game [12]. Secondly, the DMs enters in the protocol with fixed utility functions defined from τ. One can assume that the solutions to

the second game are the Maximin solutions $A^\star(\tau)$.

A vector $\tau^\star \in \mathcal{I}(A)^n$ is a Nash equilibrium of the first game iff for all $i \in N$ and all $\tau_i \in \mathcal{I}(A)$,

$$u_i(A^\star(\tau^\star)) \geq u_i(A^\star(\tau_i, \tau^\star_{-i})) \ .$$

Set for $S \subseteq N$, $\tau^\wedge_S(a) := \min_{k\in S} \tau_k(a)$, and, for $i \in N$, $\tau^\wedge_{-i}(a) := \tau^\wedge_{N\setminus\{i\}}(a)$. Let

$$A^\star(\tau) := \left\{ a \in A \, , \tau^\wedge_N(a) = \max_{b\in A} \tau^\wedge_N(b) \right\}$$

Lemma 7 *Let $\tau_{N\setminus\{i\}} \in \mathcal{I}(A)^{n-1}$ and $a \in A$. Then there exists $\tau_i \in \mathcal{I}(A)$ such that $A^\star(\tau) = \{a\}$ iff*

$$\begin{aligned} &\left|\left\{b \in A \setminus \{a\} \, , \ \tau^\wedge_{-i}(b) \geq \tau^\wedge_{-i}(a)\right\}\right| < \tau^\wedge_{-i}(a) \\ &\text{and } \tau^\wedge_{-i}(a) > 1 \end{aligned} \tag{3}$$

Let

$$A_i(\tau_{-i}) := \{a \in A \, , \ \ (3) \text{ is satisfied}\} \ .$$

Corollary 8 *Let $\tau_{N\setminus\{i\}} \in \mathcal{I}(A)^{n-1}$. Set $\tilde{D}_k = \{a \in A \, , \ \tau^\wedge_{-i}(a) \geq k\}$.*

There exists $\tau_i \in \mathcal{I}(A)$ such that $A^\star(\tau)$ is reduced to a singleton iff there exists $k \in \{2, \ldots, p\}$ such that

$$|\tilde{D}_k| \leq k \ . \tag{4}$$

Let K be the smallest integer k satisfying (4). Then $A_i(\tau_{-i}) = \tilde{D}_K$. Moreover, $|A_i(\tau_{-i})| \leq \left\lceil \frac{p}{2} \right\rceil$, where $\lceil r \rceil$ denotes the smallest integer larger or equal to r.

Lemma 9 *$A^\star(\tau)$ is reduced to a singleton iff there exists $a \in \tilde{D}_K$ such that*

$$\begin{aligned} &\forall b \in D_a \, , \ \tau_i(b) < \min(\tau_i(a), \tau^\wedge_{-i}(a)) \\ &\text{and either } \tau_i(a) \geq \tau^\wedge_{-i}(a) \text{ or } \forall b \in A \setminus (D_a \cup \{a\}) \, , \\ &\qquad \min(\tau_i(b), \tau^\wedge_{-i}(b)) < \tau_i(a) \end{aligned} \tag{5}$$

Assume that $A_i(\tau_{-i}) \neq \emptyset$. Then player i cannot act in such a way to obtain an option not in $A_i(\tau_{-i})$. Consider the option $a^\star(i)$ of $A_i(\tau_{-i})$ that has the largest value of τ_i. This is the best option that DM i can get. If the conditions of Lemma 7 are not satisfied for option $a^\star_i$, then DM i can announce another $\tilde{\tau}_i$. He can set $\tilde{\tau}_i(a^\star(i)) = \max(\tau_i(a^\star(i)), \tau^\wedge_{-i}(a^\star(i)))$. Then (5) becomes

$$\forall b \in D_{a^\star(i)} \, , \ u_i(b) < u_i(a^\star(i)) \ .$$

Since $|D_{a^\star(i)}| < \tau_i(a^\star(i))$, it is always possible to define $\tilde{\tau}_i(b)$ for $b \in D_{a^\star(i)}$ such that previous relation holds.

We study here in the situation of all DMs except one are giving their true preferences, and we investigate the interest for the last player to lie on his true preferences. The major difference between the cardinal case and the ordinal case is that a DM cannot obtain his preferred option by choosing strategy (2). In the ordinal case, he can only force to obtain one of the most preferred options for the remaining DMs. This set is some kind of consensus of the remaining DMs.

5 A protocol based on coalitions of players

5.1 Protocol

In Protocol ***P1***, each individual DM is actor of the protocol. It is well-known that individual DMs become stronger when they group themselves. This lead to the case where the actors correspond to a coalition structure, that is a partition of N. One can also think of a generalization of that, considering all coalitions that can form. Let us denote by $\mathcal{W} \subseteq 2^N$ the set of admissible coalitions. One can think of the set of *winning coalitions* in the sense of a voting rule. We are thus interested in a protocol where the actors that give their opinion are in fact all coalitions in $\mathcal{W}$.

The second protocol is similar to ***P1***, except that the actors are coalitions of $\mathcal{W}$ rather than individual DMs. The proposers and responders are elements of $\mathcal{W}$. The sequence of proposers is supposed to be known in advance and is denoted by $\Pi_1, \Pi_2, \ldots$

Protocol ***P2***. At iteration $k \in \mathbb{N}_*$:

- Coalition Π_k makes an offer $a \in A$ (a can already be proposed earlier).
- All other coalitions in $\mathcal{W} \setminus \{\Pi_k\}$ say *Yes* or *No*. However, a coalition S must accept a if $e(S, a) \geq \rho_S(k)$.
- If at least one coalition rejects the offer, we go to the next iteration. Otherwise, the option is chosen and the protocol ends.

Lemma 10 *Under* **(T)**, *protocol **P2** has at least one Subgame Perfect Nash Equilibrium (SPE).*

5.2 A generalization of Minimum Regret Solution

The idea of the maximin solution is to select the option for minimize the dissatisfaction of the player that enjoys the less the option. The maximin solution is the option (or the set of options) a such that

$$\min_{i\in N} \ u_i(a) = \max_{a'\in A} \ \min_{i\in N} \ u_i(a') \ .$$

We will see in a second that the Nucleolus proposes another definition of dissatisfaction. Instead of looking at the worse score of an option over the players, the idea of the Nucleolus is to look at the difference between what a set of players get with an option and what they could get at best.

The bargaining sets are based on the notion of *justified objection*. An objection is justified whenever there is no counterobjection. One sees that this definition is purely qualitative. There is indeed no notion of intensity of an objection or a counterobjection. This may explain why one cannot obtain nonemptyness in the general case [17]. In order to avoid the consequences of Arrow's theorem, one may think of quantitative concepts, such as the *Nucleolus*. This notion is classically based on the concept of *excess*. For a classical TU game, the excess for a given coalition and an imputation is the difference between what the game will provide for this coalition and what the imputation promises to give. In our case, the excess is defined by

$$e(S,a) = \max_{b\in A} \sum_{i\in S} u_i(b) - \sum_{i\in S} u_i(a) \ .$$

$e(S,a)$ is the difference between the maximal possible total satisfaction degree to players S and what they get with a. $e(S,a)$ cannot be negative. This notion of excess generalizes the notion of *regret* known in Social Choice [1]. In Social Choice, N represents the society, and $\sum_{i\in N} u_i(a)$ represents the total worth represented by a on the whole society. The option that is *best for society* is thus the option a that maximizes $\sum_{i\in N} u_i(a)$. The regret associated to an option a is then defined as $e(N,a)$. In Social Choice, the *minimum regret* option procedure corresponds to selecting the option that minimizes regret $e(N,a)$.

A strictly positive excess $e(S,a)$ may be interpreted as a dissatisfaction of the coalition S when faced with the proposal a. Hence, vector $\{e(S,a)\}_{S\subseteq N}$ measures the dissatisfaction of the subcoalitions of N about a. Instead of expressing dissatisfaction as an objection that one player can make to another one, as in the bargaining set, another option is to consider the alternatives a whose highest complaint $\max_{S\subseteq N} e(S,a)$ is the smallest one. Define thus $\succeq_{\max}$ by $(e(S,b))_{S\subseteq N} \succeq_{\max} (e(S,a))_{S\subseteq N}$ iff $\max_{S\subseteq N} e(S,b) \geq \max_{S\subseteq N} e(S,a)$.

A first idea is to select the option that minimizes the The leads to the maximin solution on the excess :

$$\mathcal{K} = \left\{ a\in A \,,\ \max_{S\subseteq N} e(S,a) = \min_{b\in A} \max_{S\subseteq N} e(S,b) \right\} \ ,$$

and

$$\mathcal{K}_{\mathcal{W}} = \left\{ a\in A \,,\ \max_{S\in \mathcal{W}} e(S,a) = \min_{b\in A} \max_{S\in \mathcal{W}} e(S,b) \right\} \ ,$$

where $\mathcal{W}$ is a subset of 2^N.

We will then show that under some assumptions, all equilibria of this protocol are exactly the elements of the maximin set $\mathcal{K}_{\mathcal{W}}$ on the excess. This gives a theoretical characterization of $\mathcal{K}_{\mathcal{W}}$. Since the Nucleolus $\mathcal{N}_{\mathcal{W}}$ is a refinement of $\mathcal{K}_{\mathcal{W}}$, it also justifies this result. We want to stress that Protocol ***P2*** does not pretend to be realistic in practice. It serves as a bargaining characterization of $\mathcal{K}_{\mathcal{W}}$.

Lemma 11 *Let $\{\rho_S^\delta\}_\delta$ be families of functions satisfying (1),* **(PP)** *and* **(T)** *such that ρ_S^δ does not depend on coalition S, and $\rho_S^\delta(0) = 1$ for all coalitions S. Then there exists $\Delta \in (0,1)$ such that for any $\delta \in (0,\Delta)$, the SPEs of protocol **P2** are the elements of $\mathcal{K}_{\mathcal{W}}$.*

5.3 Nucleolus

It is well-known that the max ordering is not quite discriminative since the two payoff vector are judged similar as soon as their maximal value is the same. It is possible to refine this $\succeq_{\max}$ ordering, in such a way to be more discriminative. We say that an ordering $\succeq'$ is a refinement of $\succeq_{\max}$ if $x \succeq' y$ whenever $x \succeq_{\max} y$, $x \succ' y$ whenever $x \succ_{\max} y$, and there exists z,t such that $z \succ' t$ and $z \sim_{\max} t$. The *leximax* ordering defined below is a refinement of the max ordering.

Define the *lexicographic* ordering $\succeq_{lex}$ on vectors in $\mathbb{R}^d$ (for some $m \in \mathbb{N}$) by : $x \succeq_{lex} y$ iff

$$\exists p \in \{1,\dots,m\} \,,\ x_j = y_j \text{ for all } j \in \{1,\dots,p\} \\ \text{and } x_{p+1} > y_{p+1} \ .$$

Let $x,y \in \mathbb{R}^d$ with τ,π permutations on $\{1,\dots,d\}$ such that $x_{\tau(1)} \geq \dots \geq x_{\tau(d)}$ and $y_{\pi(1)} \geq \dots \geq y_{\pi(d)}$. The *leximax* ordering is defined by

$$x \succeq_{leximax} y \ \text{ iff } \ (x_{\tau(1)},\dots,x_{\tau(d)}) \succeq_{lex} (y_{\pi(1)},\dots,y_{\pi(d)}) \,.$$

Define $\succ_{\mathcal{N}}$ by

$$a \succeq_{\mathcal{N}} b \quad \Longleftrightarrow \quad (e(S,b))_{S\subseteq N} \succeq_{leximax} (e(S,a))_{S\subseteq N}$$

The *prenucleolus* set is the option of A such that the associated vector $(e(S,a))_{S\subseteq N}$ in $\mathbb{R}^{2^N}$ is the smallest element in the leximax sense [13] :

$$\mathcal{N} = \{ a\in A \,,\ \ \forall b \in A \ \ a \succeq_{\mathcal{N}} b \}$$

Lemma 12 $\mathcal{N} \neq \emptyset$.

Example 13 *Consider 4 options and 3 players.*

	Option a	*Option b*	*Option c*
Player 1	1	0.4	0.7
Player 2	0.5	0.3	0.2
Player 3	0.2	0.3	0.7

The excess values are given in the following array.

$e(\cdot,\cdot)$	*Option a*	*Option b*	*Option c*
$\{1\}$	0	0.6	0.3
$\{2\}$	0	0.2	0.3
$\{3\}$	0.5	0.4	0
$\{1,2\}$	0	0.8	0.6
$\{1,3\}$	0.3	0.8	0
$\{2,3\}$	0.2	0.3	0
$\{1,2,3\}$	0	0.7	0.1

One has $\mathcal{N} = \{a\}$ whereas b is the maximin solution. The choice of option a has much less complaint from players and groups of players than the selection of b.

Lemma 14 *If $a, b \in \mathcal{N}$ then*

$$\sum_{i\in N} u_i(a) = \sum_{i\in N} u_i(b) \ .$$

In the general case, the Nucleolus is not necessarily reduced to a singleton, as shown in the following result.

Lemma 15 *Let $a, b \in A$ such that $\sum_{i\in N} u_i(a) = \sum_{i\in N} u_i(b)$, and for all $c \in A \setminus \{a,b\}$ and all $i \in N$,*

$$u_i(c) < u_i(a) \quad \text{and} \quad u_i(c) < u_i(b) \ .$$

Then

$$\mathcal{N} = \{a, b\} \ .$$

However, as the following example shows, the elements of the Nucleolus do not necessarily have the largest value of $\sum_{i\in N} u_i(a)$, or equivalently the smallest value of the overall excess $\sum_{S\subseteq N} e(S,a)$. This comes from the fact that the Nucleolus is derived from the max ordering.

Example 16 *Consider 4 options and 3 players.*

	a	b	c	d
Player 1	1	0.5	0.5	0.6
Player 2	0	0.5	0	0.1
Player 3	0.2	0.2	0.7	0.3

The excess values are given in the following array.

$e(\cdot,\cdot)$	a	b	c	d
$\{1\}$	0	0.5	0.5	0.4
$\{2\}$	0.5	0	0.5	0.4
$\{3\}$	0.5	0.5	0	0.4
$\{1,2\}$	0	0	0.5	0.3
$\{1,3\}$	0	0.5	0	0.3
$\{2,3\}$	0.5	0	0	0.3
$\{1,2,3\}$	0	0	0	0.2

One has $\mathcal{N} = \{d\}$. However,

$$\sum_{S\subseteq N} e(S,d) = 2.3 > \sum_{S\subseteq N} e(S,a) = \sum_{S\subseteq N} e(S,b) = \sum_{S\subseteq N} e(S,c) = 1.5$$

So, one sees that the overall excess on all possible coalitions is much worse for d than for a, b and c.

It is possible to restrict the definition of the Nucleolus to winning coalitions. This means that one does not care about the excess for losing coalitions.

$$\mathcal{N}_{\mathcal{W}} = \{a \in A \ , \ \ \forall b \in A \quad (e(S,b))_{S\in\mathcal{W}} \succeq_{leximax} (e(S,a))_{S\in\mathcal{W}}\} \ .$$

Following Example 16, the overall excess $\sum_{S\in\mathcal{W}} e(S,a)$ is an interesting measure together with the leximax ordering. To this end, the leximax ordering has to be transformed into a utility model. It is well-known that the *leximax* ordering cannot in the general case be described by a utility model, i.e.

$$x \succeq_{leximax} y \quad \Longleftrightarrow \quad F_{leximax}(x) \geq F_{leximax}(y) \ .$$

This is possible when the scale in discrete, which is not our case. However, we replace the leximax ordering by that induced by such function $F_{leximax}$. This gives a coarsening of the leximax ordering but it is still a refinement of the max ordering.

Define thus for $\tau \in [0,1]$ for $x : \mathcal{W} \to \mathbb{R}$

$$G_\tau(x) = \tau \sum_{S\in\mathcal{W}} x(S) + (1-\tau)\, F_{leximax}(x) \ .$$

Function G_τ describes a tradeoff between the largest excess and the mean excess. Then define

$$\mathcal{N}^\star_{\mathcal{W}} = \{a \in A \ , \ \ \forall b \in A \quad G_\tau\left(\{e(S,b)\}_{S\in\mathcal{W}}\right) \geq G_\tau\left(\{e(S,a)\}_{S\in\mathcal{W}}\right)\} \ .$$

Example 17 *Consider 3 options and 3 players.*

	Option a	*Option b*	*Option c*
Player 1	0.5	0	0
Player 2	0.1	0.5	0.7
Player 3	0.5	0.5	0

It is easy to seen that $a \succ_{\mathcal{N}} b$. Assume now that option c is withdrawn. Removing c from A leads to the reverse relation $b \succ_{\mathcal{N}} a$. We conclude that the presence of c influences the preference between a and b. Hence the Nucleolus does not satisfy the Independence to Irrelevant Alternatives *axiom [1].*

6 Conclusion

We have introduced in Section 3 a negotiation protocol that is mainly characterized by an acceptance threshold that obliges the DMs to accept an offer having a better utility than that threshold. In order to predict the outcome of this protocol, we have studied its SPEs. In the case where the thresholds decreases slowly enough, we show that the SPEs correspond to the maximin solution. In the general case, we have designed an algorithm that compute the outcomes of all possible SPEs.

We notice in Section 4 that the protocol is subject to manipulation by the DMs. Two possible proposals to reduce this sensitivity are made. The first one consists in the detection of a possible distortion of the initial preferences. The second one proposes an ordinal version of the protocol, which reduces the influence of manipulation.

Previous protocol has been generalized in Section 5 allowing coalitions to interact. This leads to a solution concept that generalizes the notion of minimum regret and that seems to be new.

Acknowledgement

This work is partly funded by the ICIS research project under the Dutch BSIK Program (BSIK 03024).

References

[1] K.J. Arrow, A.K. Sen, K. Suzumura (Eds.). Handbook of Social Choice and Welfare. Elsevier Science, Amsterdam, 2002.

[2] C. Bana e Costa, J.M. De Corte, J.C. Vansnick. On the mathematical foundations of MACBETH. In: "Multiple Criteria Decision Analysis: state of the art surveys" J. Figueira, S. Greco, M. Ehrgott eds. Springer, New York, 2005.

[3] M.H. DeGroot. Reaching a consensus. J. of American Statistical Association, 69 (1974), 118-121.

[4] P. Eklund, A. Rusinowska, H. de Swart. Consensus reaching in committees. European J. of Operational Research, to appear in 2006.

[5] P. Faratin, C. Sierra, N. R. Jennings. Negotiation Decision Functions for Autonomous Agents. Int. Journal of Robotics and Autonomous Systems 24 (1998) 159-182.

[6] M. Fedrizzi, M. Fedrizzi, R. Marques Pereira. Soft consensus and network dynamics in group decision making. Int. J. of Intelligent Systems, 14 (1999), 63-77.

[7] A. Gibbard, "Manipulation of voting schemes: a general result", Econometrica, Vol. 41, No. 4 (1973), pp. 587601.

[8] M. Grabisch, Ch. Labreuche. Fuzzy Measures and Integrals in MCDA. In: J. Figueira, S. Greco, M. Ehrgott. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, New York, 2005. pp. 563-608.

[9] N.R. Jennings, P. Faratin, A.R. Lomuscio, S. Parson, M. Wooldridge, C. Sierra. Automated negotiation: Prospects, Methods and Challenges. Group Decision and Negotiation 10 (2001), 199-215.

[10] J. Kacprzyk and M. Fedrizzi (Eds.). Multiperson Decision Making Models using Fuzzy Sets and Possibility Theory, Kluwer, Dordrecht, 1990.

[11] J. Kacprzyk, H. Nurmi, M. Fedrizzi (Eds.). Consensus under Fuzziness. Kluwer, Boston, 1996.

[12] M. Kurz. Income distribution and distortion of preference: the l-commodity case. Journal of Economic Theory 22 (1980) 99-10.

[13] G. Owen. Game Theory. Academic Press, New York. 1995.

[14] A. Rubinstein. Perfect equilibrium in a bargaining model. Econometrica 50 (1982) 97-109.

[15] M. Satterthwaite. Strategy-Proofness and Arrows Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. Journal of Economic Theory 10 (1975) 187-217.

[16] Sengupta, K. and D. Teeni. Incorporating Multiple Levels of Information Processing in CSCW: An Integrated Design Approach. In: "Cooperative Work and GroupWare Systems", D. Shapiro, M. Tauber and R. Traunmuller (Eds.). Elsevier Science, North-Holland (1996), pp 119-134.

[17] P. Shenoy. On committee decision making: a game theoretical approach. Management Science 26 (1980) 387-400.

[18] R. Vetschera. Integrating databases and preference evaluations in group decision support: a feedback-oriented approach. Decision Support Systems 46 (2000) 1585-1601.

Modeling the context in Multi-Criteria Decision Analysis

Christophe Labreuche
THALES Research and Technology
RD 128, 91767 Palaiseau Cedex, France
christophe.labreuche@thalesgroup.com

Abstract

In Multiple Criteria Decision Analysis, the preferences of the decision maker regarding each criterion are classically modelled by a utility function depending on one single variable representing the point of view attached to the criteria. There are situations in which the preferences regarding an attribute depend on more variables. In this case, we propose an extended model taking into account all contributing variables. The main asset of our approach is that it does not requires much more data from the decision maker than the classical model. We also propose an explanation framework.

Keywords: decision support systems, multiple criteria analysis, utility functions, context.

1 Introduction

Multi-Criteria Decision Analysis (MCDA) aims at helping a Decision Maker (DM) in selecting one option among several or in assessing several options, on the basis of a family $N = \{1, \ldots, n\}$ of viewpoints characterizing the various consequences of this selection or these assessments. An essential part of this analysis is the construction of a multi-criteria model deemed to represent the preferences $\succeq$ of the DM. For two options a and b, $a \succeq b$ means that a is at least as good as b to the DM. We consider here the transitive decomposable model [7]

$$a \succeq b \iff H(u_1(a), \ldots, u_n(a)) \geq H(u_1(b), \ldots, u_n(b))$$

where $u_k(a)$ measures the satisfaction of the DM regarding option a considering only viewpoint $k \in N$, and H is an aggregation function.

MCDA is very often applied to problems where the options of interest $A = \{a^1, \ldots, a^m\}$ are either known in advance or constructed during the early stage of the analysis. In this case, the multi-criteria model is usually specified only on the options of A. In particular, utility function u_k is constructed only on the elements of A, and thus $u_k : A \to \mathbb{R}$.

There are many applications where an explicit representation of the utility functions on the set of options is not possible. Let us describe three examples. In the first example, the number of options is either infinite or very large so that it is not possible to elicit the satisfaction degree u_k of the DM explicitly on all options. In the two other examples, the multi-criteria model must be constructed without a knowledge of the options on which it will be applied later. The second example corresponds to bid evaluation in public call for tenders [1, 2]. The model must thus be specified before the competition begins, and must be made public to the potential candidates. It defines the *rules* under which the candidates will be assessed. In the last example, the multi-criteria model is implemented in decision support system as an embedded function that has to provide an assessment each time a new situation arises. One can mention automatic assessment of trainees in a simulator or automatic situation assessment in a surveillance system.

These applications are very different from previous case since the model shall produce meaningful comparisons or assessments on all potential options that can arise. All these options may cover very different situations that have to be anticipated beforehand. The model that is constructed must thus be generic enough to take all these situations into consideration.

In previous examples, the consequence of potential options on each viewpoint $k \in N$ must necessarily be described by a set X_k called *descriptor* or *attribute*. This set corresponds to a mono-dimensional variable that is the indicator or metric that is used to repre-

sent the viewpoint. An option a is thus described by a value $a_k \in X_k$ on each attribute, and utility function u_k is now a function from X_k onto $\mathbb{R}$.

The representation of a viewpoint $k \in N$ by one mono-dimensional attribute X_k is not always satisfactory. It happens indeed that a same value of an attribute X_k corresponds to different situations and is interpreted differently by the DM. Let us give an example. Consider a training simulator in which trainees practice shooting on targets with a riffle. The main viewpoint to be considering when evaluating the quality of a shot is the sighting shift which is the distance between the point that the trainee has aimed with the riffle and the center of the target. This variable corresponds to the descriptor associated with the main viewpoint. The smaller this value the better. However, a value of, say $10\,cm$ is not always judged in the same manner by the instructor. If the target is at short range and is not moving, $10\,cm$ is judged bad, whereas if the target is at long range and is moving, this same figure is judged very good. This means that in order to assess attribute *sighting shift*, one needs to know the distance of the target as well as its speed. These two variables characterize the context under which the attribute must be analyzed.

More generally, an attribute X_k must sometimes be analyzed with other variables called hereafter *contextual variables* described the sets Y_k^1, ..., $Y_k^{p_k}$. Let $P_k = \{1, \ldots, p_k\}$ be the contextual variables for viewpoint k. We want to stress that the contextual variables do not measure the consequence on another viewpoint. The utility function u_k are thus defined as $u_k : X_k \times Y_k^1 \times \cdots \times Y_k^{p_k} \to \mathbb{R}$.

From now on, we consider the case where set X_k is continuous. We propose in this paper a method to construct a utility function $u_k : X_k \times Y_k^1 \times \cdots \times Y_k^{p_k} \to \mathbb{R}$, requiring a number of data provided by the DM which is not much more than what is done for usual mono-dimensional utility functions.

The rest of the paper is organized as follows. Section 2 describes the Macbeth approach that aims in specifying mono-dimensional utility functions. Section 3 shows some naive ways to tackle the multidimensional aspect of utility functions and gives their drawbacks. The model of representation of the multi-dimensional utility functions is explained in Section 4. We explain the data that must be provided by the DM to specify the model. The interpolation method that enables the DM to compute the utility value once the model has been specified is given in Section 5. The model that is proposed in this paper is a little bit more complex than the mono-dimensional one for a user. It becomes thus necessary to help the DM in understanding the result of the evaluation by proposing an automatic generation of an explanation in a Decision Support System. This is precisely the goal of Section 6.

2 Construction of mono-dimensional utility function

Let us describe in this section the usual case where the preferences of the DM over viewpoint k depends only on attribute X_k. We denote by $\succeq_k$ the preference relation of the DM over X_k. Under the Birkhoff-Milgram theorem, if $\succeq_k$ is a weak order and $X_k/\sim_k$ contains a countable order-dense subset, then $\succeq_k$ can be represented by a function $u_k : X_k \to \mathbb{R}$, that is [7]

$$x \succeq_k x' \iff u_k(x) \geq u_k(x') \,. \tag{1}$$

Function u_k satisfying (1) corresponds to an ordinal scale. In order to obtain a cardinal scale, one needs to possess a quaternary relation $\succeq_k^\star$ over $X_k \times X_k$. For $x, x', t, t' \in X_k$, $x\,x' \succeq_k^\star t\,t'$ means that the intensity of preference between x and x' is at least as good as that between t and t'. Under some reasonable conditions on $\succeq_k^\star$, we have the following equivalence [7]

$$xx' \succeq_k^\star tt' \iff u_k(x) - u_k(x') \geq u_k(t) - u_k(t') \,. \tag{2}$$

AHP [9] and Macbeth [3] can be used to construct utility functions. We only describe Macbeth since it is well-founded from a measurement standpoint, and ensures both (1) and (2). Assume that X_k is an interval of $\mathbb{R}$. Utility function u_k is assumed to be piecewise linear. It is thus enough to determine u_k on a finite set of elements $\{x^1, \ldots, x^q\} \subset X_k$, with $x^1 \leq \cdots \leq x^q$. u_k is indeed affine between two points x^i and x^{i+1}:

$$u_k(x) = I(x; x^1, u_k(x^1), \ldots, x^q, u_k(x^q))$$

where

$$I(x; x^1, u^1, \ldots, x^q, u^q) = \begin{cases} u^1 & \text{if } x \leq x^1 \\ u^i + (u^{i+1} - u^i)\,\frac{x - x^i}{x^{i+1} - x^i} & \text{if } x^i \leq x < x^{i+1} \\ u^q & \text{if } x \geq x^q \end{cases}$$

There remains to explain how $u_k(x^1), \ldots, u_k(x^q)$ are determined in the Macbeth approach. The construction of $\succeq_k^\star$ on the q selected elements $x^1, \ldots, x^q$ requires the comparison of $q(q-1)/2$ quantities (namely all differences $u_k(x^j) - u_k(x^l)$ where $x^j \succ_k x^l$), which demands roughly $q^4/8$ operations. In order to reduce the complexity, the idea is ask the DM to assess directly the $q(q-1)/2$ differences of satisfaction $u_k(x^j) - u_k(x^l)$ between two values x^j and x^l, for any $j \neq l$ such that $x^j \succ_k x^l$. Since the DM cannot assess

such quantities with precise numbers, the Macbeth approach proposes an ordinal assessment scales composed of only 6 elements: {*very small, small, mean, large, very large, extreme*}=: $\mathcal{E}$. The information asked in practice is thus quite similar to $\succeq_k^\star$. The advantage of asking $u_k(x^j)-u_k(x^l)$ is that it leads to a redundant information that will enable to compute accurate but non-unique values of $u_k(x^j)$ for $j \in \{1,\ldots,q\}$. The second advantage is that it is easier for a human being to give some relative information regarding a difference (for instance $u_k(x^j) - u_k(x^l)$) than to give some absolute information (for instance $u_k(a^j)$).

Previous method defines an interval scale. Such scale is unique up to an affine transformation. u_k can be uniquely determined by fixing its value at two reference points of X_k. We assume that there exists on X_k an element denoted by $\mathbf{0}_k$ that is considered as completely unacceptable by the DM, and an element denoted by $\mathbf{1}_k$ that is considered perfectly satisfactory by the DM. In order to specify entirely utility function u_k, the utilities of reference elements $\mathbf{0}_k$ and $\mathbf{1}_k$ are fixed to values 0 and 1 respectively.

$$u_k(\mathbf{0}_k) = 0 \quad , \quad u_k(\mathbf{1}_k) = 0 \ . \tag{3}$$

3 Naive ways to construct multi-dimensional utility functions

The first naive method to construct multi-dimensional utility function $u_k : X_k \times Y_k^1 \times \cdots \times Y_k^{p_k} \to \mathbb{R}$ consists in applying exactly the approach described in Section 2. Utility function u_k can indeed be seen as a function on the Cartesian product $X_k \times Y_k^1 \times \cdots \times Y_k^{p_k}$. Assume that each set $X_k, Y_k^1, \ldots, Y_k^{p_k}$ can be characterized by q elements. Hence one needs to determine u_k on q^{p_k+1} elements to determine u_k everywhere. Applying Macbeth, this means that the DM must provide $\frac{q^{p_k+1}(q^{p_k+1}-1)}{2} \approx \frac{1}{2}q^{2(p_k+1)}$. The generalization of mono-dimensional approaches to the multi-dimensional case would lead to a number of questions that is far too big for the DM. So this method is not possible in practice.

An alternative solution would be to synthesize all variables into one unique. In previous example, let us denote by x the sighting shift, by y_1 the distance to target and by y_2 the speed of the target. This amounts to defining a new variable such as

$$z = x - \frac{1}{10}\, y_1 - \frac{1}{3}\, y_2$$

and construct a one-dimensional utility function on z. This is what B. Roy have called *arithmo-morphism* [8]. The construction of z is very empiric and arbitrary [8].

We propose in this paper a method to construct a utility function $u_k : X_k \times Y_k^1 \times \cdots \times Y_k^{p_k} \to \mathbb{R}$ that is not doing arithmo-morphism, with a number of data provided by the DM which is not much more than what is done for usual mono-dimensional utility functions.

4 Representation of the multi-dimensional utility function

Point of view $k \in N$ is described by a main variable x called metric which values belong to the set X_k. The interpretation of each value of that metric depends on a context depicted by a set $y_1, \ldots, y_{p_k}$ of contextual variables taking values in the sets $Y_k^1, \ldots, Y_k^{p_k}$ respectively.

We denote by $\succeq_k$ the preference relation of the DM over $X_k \times Y_k^1 \times \cdots \times Y_k^{p_k}$. Set $y = (y_1, \ldots, y_{p_k}) \in Y_k := Y_k^1 \times \cdots \times Y_k^{p_k}$. In the following, for $S \subseteq P_k$, we will denote (y_S, y'_{-S}) the compound vector z with $z_j = y_j$ for $j \in S$ and $z_j = y'_j$ for $j \in P_k \setminus S$. For $y \in Y_k$ fixed, define the preference relation $\succeq_{k,y}$ over X_k, which is just the restriction of $\succeq_k$ on the set $X_k \times \{y\}$.

Preference relation $\succeq_{k,y}$ is a mono-dimensional order relation that can be constructed as described in Section 2. By (1), (2) and (3), we obtain

$$x \ \succeq_{k,y} \ x' \iff U_{k,y}(x) \geq U_{k,y}(x') \ . \tag{4}$$

$$x\, x' \ \succeq_{k,y}^\star \ t\, t' \iff U_{k,y}(x) - U_{k,y}(x') \geq U_{k,y}(t) - U_{k,y}(t') \ . \tag{5}$$

$$\exists \mathbf{0}_{k,y}, \mathbf{1}_{k,y} \in X_k \ , \quad U_{k,y}(\mathbf{0}_{k,y}) = 0 \quad , \quad U_{k,y}(\mathbf{1}_{k,y}) = 0 \ . \tag{6}$$

Recall that $\mathbf{0}_{k,y}$ and $\mathbf{1}_{k,y}$ correspond to two reference points.

The preferences of the DM on viewpoint k is mainly influenced by attribute X_k. The contextual variables just strengthen or weaken this link. They cause the judgment of the DM being more or less tolerant. One feels that the preferences as described by $\succeq_{k,y}$ can be derived from $\succeq_{k,y'}$. More precisely, we look for an homomorphism from $(X_k, \succeq_{k,y})$ and $(X_k, \succeq_{k,y'})$. Under some order-dense conditions between $(X_k, \succeq_{k,y})$ and $(X_k, \succeq_{k,y'})$, which states that $(X_k, \succeq_{k,y})$ and $(X_k, \succeq_{k,y'})$ are of the same dimension, there exists an homorphism $\phi_{y,y'} : X_k \to X_k$ such that

$$\forall x, x' \in X_k \ , \quad x \succeq_{k,y} x' \iff \phi_{y,y'}(x) \succeq_{k,y'} \phi_{y,y'}(x') \ . \tag{7}$$

$\phi_{y,y'}(x)$ is the value of X_k at context y' that is equivalent to value x at context y.

$$\forall y, y' \in Y_k \quad , \quad (x, y) \sim_k (\phi_{y,y'}(x), y') \ , \tag{8}$$

Function $\phi_{y,y'}$ shall be surjective so that $Range(\phi_{y,y'}) = X_k$. Function $\phi_{y,y'}$ shall be strictly increasing so that $\succeq_{k,y}$ and $\succeq_{k,y'}$ represent the same ordinal preferences up to shrinking and dilations of the X_k scale.

Since $\phi_{y,y'}$ is an homomorphism, it shall preserve the reference elements :

$$\forall x \in X_k \quad \phi_{y,y'}(\mathbf{0}_{k,y'}) = \mathbf{0}_{k,y} \quad , \quad \phi_{y,y'}(\mathbf{1}_{k,y'}) = \mathbf{1}_{k,y} \ . \tag{9}$$

Lemma 1 *Under (4), (5), (6), (8) and (9), one has for all $x \in X_k$ and $y, y' \in Y_k$*

$$U_{k,y}(x) = U_{k,y'}(\phi_{y,y'}(x)) \ .$$

Previous lemma shows that the utility function for one context y' can be deduced from the utility function for another context y.

We now assume that $\phi_{y,y'}$ is continuous and strictly increasing, and that X_k is an interval (X_k is convex in $\mathbb{R}$).

Lemma 2 *For all $x \in X_k$ where $L(x) = \{x' \in X_k \ , \ U_{k,y}(x') = U_{k,y}(x)\}$ is finite, then*

$$\phi_{y,y'}(x) = {\phi_{y',y}}^{-1}(x) \ .$$

For the other values of x, $U_{k,y}$ is constant and if ${\phi_{y,y'}}(x) = {\phi_{y',y}}^{-1}(x)$, one would have exactly the same utility value.

One can assume that for all $x \in X_k$

$$\phi_{y,y'}(x) = {\phi_{y',y}}^{-1}(x) \ . \tag{10}$$

Previous lemma shows that it is not necessary to construct the interval scale $U_{k,y}$ for all $y \in Y_k$. If transformation ϕ is known, one needs only to construct the interval scale for only one $y \in Y_k$. This particular element $y^\star$ of Y_k is called the *reference context*. This reference context corresponds to a usual situation where the DM can easily construct the interval scale $U_{k,y^\star}$. We construct the utility function $U_{k,y^\star}$ using what is described in Section 2 since it corresponds to a mono-dimensional utility function. One needs to identify in X_k several elements $\{x^1, \ldots, x^q\} \subset X_k$, with $x^1 \leq \cdots \leq x^q$. From, what is described in Section 2, the value of $U_{k,y^\star}$ is determined at those elements. Then we obtain

$$U_{k,y^\star}(x) = I(x; x^1, U_{k,y^\star}(x^1), \ldots, x^q, U_{k,y^\star}(x^q))$$

If transformation $\phi_{y,y^\star}$ is known, then one obtains $U_{k,y}$ for any $y \in Y_k$ from the following relation

$$U_{k,y}(x) = U_{k,y^\star}(\phi_{y,y^\star}(x)) \ . \tag{11}$$

From (10) and Lemma 1, one obtains the following result.

Lemma 3 *We have*

$$U_{k,y}(x) = I(x; \phi_{y^\star,y}(x^1), U_{k,y^\star}(x^1), \ldots, \phi_{y^\star,y}(x^q), U_{k,y^\star}(x^q)) \tag{12}$$

It proves that one needs only to specify function $y \mapsto \phi_{y^\star,y}(x^i)$ for all $i \in \{1, \ldots, q\}$, which is the aim of Section 5.

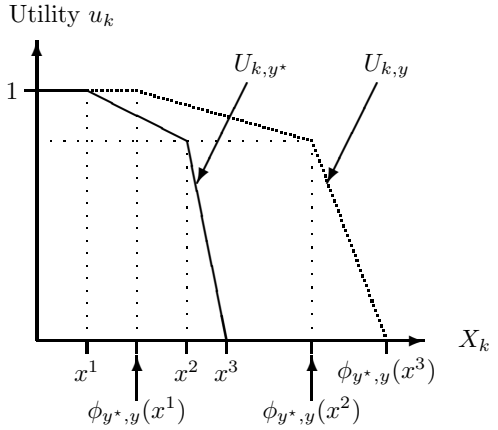


Figure 1: Fuzzy levels $N_{-m}, \ldots, N_0, \ldots, N_m$.

5 Interpolation

5.1 Approach

From Lemma 3, there remains to define $y \mapsto \phi_{y^\star,y}(x^i)$ for all $i \in \{1, \ldots, q\}$. Let us fix one $i \in \{1, \ldots, q\}$ and one $y \in Y_k$.

Function $y \mapsto \phi_{y^\star,y}(x^i)$ can be assumed to be piecewise linear. More precisely, one can assume that $\phi_{y^\star,y}(x^i)$ is known at some values $y(1), \ldots, y(r) \in Y_k$ of y, and $\phi_{y^\star,y}(x^i)$, for any y is computed by interpolation.

One can simply suppose that the points $y(1), \ldots, y(r)$ define a regular rectangular mesh in Y_k. The main advantage of this representation is that the interpolation in a rectangular mesh is simple to compute and intuitive for the DM. However, even if there are only a few elements on each contextual variable, this yields a huge number of points in the mesh. This leads to a huge number of information that the DM has to provide.

The idea is thus to consider any set $y(1), \ldots, y(r)$ of points in Y_k. We suppose that the DM has given the values of $y(1), \ldots, y(r)$ as well as the associated values $\phi_{y^\star,y(1)}(x^i) =: x^i(1)$, ..., $\phi_{y^\star,y(r)}(x^i) =: x^i(r)$. Hence the DM has to provide the following array.

Selected values on X_k	Contextual variables Y_k		
$x^i(1)$	$y_1(1)$	$\cdots$	$y_{p_k}(1)$
$x^i(2)$	$y_1(2)$	$\cdots$	$y_{p_k}(2)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$x^i(r)$	$y_1(r)$	$\cdots$	$y_{p_k}(r)$

Computing $\phi_{y^\star,y}(x^i)$ for any y from previous array is an interpolation problem. Our concern is about the choice of the interpolation. If one chooses to use a multi-dimensional interpolation on Y_k, the result becomes very hard to predict, and thus the DM will not master the consequences of the data he provides.

We define thus an interpolation that is constructed iteratively so that the DM has a better grasp on the result. This interpolation is based on an ordering of the contextual variables from the most important ones to the less important. Let τ be a permutation on P_k.

Selected values on X_k	Contextual variables Y_k		
$x^i(1)$	$y_{\tau(1)}(1)$	$\cdots$	$y_{\tau(p_k)}(1)$
$x^i(2)$	$y_{\tau(1)}(2)$	$\cdots$	$y_{\tau(p_k)}(2)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$x^i(r)$	$y_{\tau(1)}(r)$	$\cdots$	$y_{\tau(p_k)}(r)$

For $t \in \{1, \ldots, p_k\}$, we define the *Lexicographic* ordering $\preceq^t_{\text{lexico}}$ by

$$y \preceq^t_{\text{lexico}} y' \;\Leftrightarrow\; \exists s \leq t \text{ such that } \forall l \in \{1, \ldots, s\}\ y_{\tau(l)} = y'_{\tau(l)} \text{ and } y_{\tau(s)} < y'_{\tau(s)}$$

for any $y, y' \in Y_k$. Let us denote by $\prec^t_{\text{lexico}}$ be the asymmetric part of $\preceq^t_{\text{lexico}}$, and by $\sim^t_{\text{lexico}}$ the symmetric part of $\preceq^t_{\text{lexico}}$. One assumes that $y(1), \ldots, y(r)$ are ordered according to the lexicographic ordering

$$y(1) \prec^{p_k}_{\text{lexico}} y(2) \prec^{p_k}_{\text{lexico}} \cdots \prec^{p_k}_{\text{lexico}} y(r)\ .$$

The interpolation is performed one variable at a time, starting with variable $\tau(p_k)$, then $\tau(p_k - 1)$, ..., up to $\tau(1)$. Let us describe the process for variable $\tau(p_k)$. We begin by grouping the elements $y(1), \ldots, y(r)$ into clusters in which all elements have the same value on variables $y_{\tau(1)}, \ldots, y_{\tau(p_k-1)}$ within each cluster. More precisely, let us define $h_0 = 1, h_1, \ldots, h_m$ such that

$$\begin{aligned} &y(1) \sim^{p_k-1}_{\text{lexico}} y(2) \sim^{p_k-1}_{\text{lexico}} \cdots \sim^{p_k-1}_{\text{lexico}} y(h_1 - 1) \\ &\prec^{p_k-1}_{\text{lexico}} y(h_1) \sim^{p_k-1}_{\text{lexico}} \cdots \sim^{p_k-1}_{\text{lexico}} y(h_2 - 1) \\ &\prec^{p_k-1}_{\text{lexico}} y(h_2) \ \cdots\ y(h_{m-1} - 1) \\ &\prec^{p_k-1}_{\text{lexico}} y(h_{m-1}) \sim^{p_k-1}_{\text{lexico}} \cdots \sim^{p_k-1}_{\text{lexico}} y(h_m - 1)\ . \end{aligned}$$

The m clusters are thus $\{y(1), \ldots, y(h_1 - 1)\}$, $\{y(h_1), \ldots, y(h_2 - 1)\}$, ..., $\{y(h_{m-1}), \ldots, y(h_m - 1)\}$. From value $y_{\tau(p_k)}$ of the last variable in the order τ, we can transform each cluster into one single element of $Y_k \setminus Y_k^{p_k}$. More precisely, one can transform previous array in the following one, removing last variable $\tau(p_k)$

values on X_k	Contextual variables Y_k		
$\tilde{x}^i(1)$	$y_{\tau(1)}(h_0)$	$\cdots$	$y_{\tau(p_k-1)}(h_0)$
$\tilde{x}^i(2)$	$y_{\tau(1)}(h_1)$	$\cdots$	$y_{\tau(p_k-1)}(h_1)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\tilde{x}^i(m)$	$y_{\tau(1)}(h_{m-1})$	$\cdots$	$y_{\tau(p_k-1)}(h_{m-1})$

where for all $l \in \{1, \ldots, m\}$

$$\begin{aligned} \tilde{x}^i(l) = I(&y_{\tau(p_k)}; y_{\tau(p_k)}(h_{l-1}), x^i(h_{l-1}), \\ &y_{\tau(p_k)}(h_{l-1}+1), x^i(h_{l-1}+1), \ldots, \\ &y_{\tau(p_k)}(h_l - 1), x^i(h_l - 1)) \end{aligned}$$

The same process is applied to previous array in order to remove the last remaining variable in the order τ, i.e. variable $\tau(p_k - 1)$. And so on. When there remains only variable $\tau(1)$

values on X_k	Contextual variables Y_k
$\hat{x}^i(1)$	$\hat{y}_{\tau(1)}(1)$
$\hat{x}^i(2)$	$\hat{y}_{\tau(1)}(2)$
$\cdots$	$\cdots$
$\hat{x}^i(w)$	$\hat{y}_{\tau(1)}(w)$

then we obtain

$$\begin{aligned} \phi_{y^\star,y}(x^i) = I(&y_{\tau(1)}; \hat{y}_{\tau(1)}(1), \hat{x}^i(1), \hat{y}_{\tau(1)}(2), \hat{x}^i(2), \\ &\ldots, \hat{y}_{\tau(1)}(w), \hat{x}^i(w)) \end{aligned}$$

5.2 Example

To illustrate our approach, consider once more the example of shooting exercise on a simulator. Y_1 corresponds to the speed and Y_2 corresponds to the distance. The reference context is defined at fixed target at short range, i.e. $y_1^\star = 0\,m/s$ and $y_2^\star = 10\,m$.

The application of MACBETH leads to the following values of x^i and $U_{k,y^\star}(x^i)$.

Values of the utilities			
1	0.7	0.2	0
Selected values on X_k			
$10\,cm$	$15\,cm$	$20\,cm$	$30\,cm$

One notices that the smaller the value of the attribute the better.

Speed is more important than Distance. The DM provides the following values that specify $\phi_{y^\star,y}$.

Selected values on X_k				Speed	Distance
10 cm	15 cm	20 cm	30 cm	0 m/s	10 m
15 cm	20 cm	25 cm	40 cm	10 m/s	10 m
20 cm	25 cm	30 cm	45 cm	20 m/s	10 m
30 cm	35 cm	40 cm	55 cm	20 m/s	30 m
20 cm	25 cm	30 cm	45 cm	20 m/s	50 m

Note that the first line corresponds to the reference. The contextual variables are ordered in such a way that Speed comes first and Distance comes second by τ.

When shooting at a static target, the judgment does not depend on the distance. Even if aiming is more tedious for long range than short range, the instructor does not want to take this into account in his evaluation. This applies also for medium speed. However, in this case, the expert is a little more tolerant. One indeed notice that the values of the selected values are increased. Now at large speed ($20\,m/s$), the judgment now depends on the distance. At low and long range, the DM is more tolerant than for medium range. The reason is that the "defilement" speed is large at small range, which makes aiming difficult; and aiming at long range is always difficult, especially when the target is moving.

Let us explain now how the utility u_k is computed for values $x = 28\,cm$, $y_1 = 15\,m/s$ and $y_2 = 20\,m$. The removal of last contextual variable leads to the following array.

Selected values on X_k				Speed
10 cm	15 cm	20 cm	30 cm	0 m/s
15 cm	20 cm	25 cm	40 cm	10 m/s
25 cm	30 cm	35 cm	50 cm	20 m/s

As example, the last selected value for the reduced context $y_1 = 20\ m/s$ is $I(20m; 10m, 45cm; 30m, 55cm; 100m, 60cm) = 50cm$.

Repeating this on the last contextual variable, we obtain $\phi_{y^\star,y}$

Selected values on X_k			
20 cm	25 cm	30 cm	45 cm

We conclude that the utility is $I(28cm; 20cm, 1, 25cm, 0.7, 30cm, 0.2, 45cm, 0) = 0.4$.

6 Explanation of the utility function on real data

We have seen that the computation of $u_k(x, y_1, \ldots, y_{p_k}) =: v$ results from several nested interpolations. Value v taken by criterion k is not trivial to the DM. The DM is thus interested in an explanation of the reason why u_k equals v. It is completely useless to explain the precise value of v. Yet, the user wants to know why v is, for instance, rather bad, rather mean or rather good. This amounts thus to explaining why v belongs to a fuzzy set. Suppose thus that the $[0, 1]$ satisfaction scale is covered by $2\,m + 1$ fuzzy levels $N_{-m}, \ldots, N_0, \ldots, N_m$ described by the fuzzy membership functions $\mu_{N_{-m}}, \ldots, \mu_{N_m}$ (See Figure 2). Level $\mu_{N_{-m}}$ is the worse level (e.g. *very bad*), μ_{N_m} is the best level (e.g. *very good*).

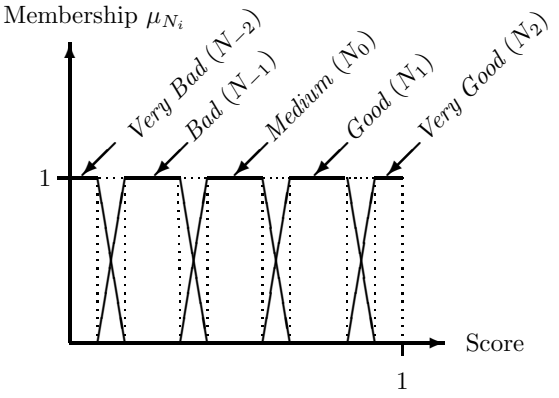


Figure 2: Fuzzy levels $N_{-m}, \ldots, N_0, \ldots, N_m$.

We begin by determining the level corresponding to value v. This is $i = \text{Argmax}_{j\in\{-m,\ldots,m\}}\mu_{N_j}(v)$.

6.1 Main argument

The main argument concerns the relation between x and u_k at fixed value of the context $y = (y_1, \ldots, y_{p_k})$. It is not possible to provide a true argumentation of the evaluation made. The explanation we propose to give consists in the interval of values of attribute k around x for which the assessment would belong to the same fuzzy level i for the same context y.

We first determine the largest interval of values $(\underline{v}, \overline{v})$ around v such that for all $t \in (\underline{v}, \overline{v})$, $\mu_{N_i}(t) = \max_{j\in\{-m,\ldots,m\}} \mu_{N_j}(t)$. Then, we determine the largest interval $(\underline{x}, \overline{x})$ such that

$$x \in (\underline{x}, \overline{x})$$

and

$$\forall x' \in (\underline{x}, \overline{x})\ ,\ u_k(x', y_1, \ldots, y_{p_k}) \in (\underline{v}, \overline{v})$$

The text that is displayed to the user is thus

> `Between` $\underline{x}$ `and` $\overline{x}$`, the value of attribute` k `is judged` N_i`.`

Note that values $\underline{x}$ and $\overline{x}$ used in the explanation do not correspond to the values that were given by the expert in the specification of the utility function.

6.2 Argument on the secondary variables

The second part of the explanation concerns the influence of the contextual variables on the evaluation. The idea is to compare the result $v := u_k(x, y)$ to what would have been obtained $v^\star = u_k(x, y^\star)$ with the reference context $y^\star$. If v is significantly larger than $v^\star$, the judgment is more tolerant in the current context y than what would have been obtained in the reference context $y^\star$. If v is significantly smaller than $v^\star$, the judgment is more intolerant in the current context than what would have been obtained in the reference context. We quantify the difference $v-v^\star$, among $2m+1$ fuzzy levels $Q_{-m}, \ldots, Q_0, \ldots, Q_m$ that cover interval $[-1, 1]$. Let μ_{Q_i} be the membership of Q_i. One can set $\mu_{Q_i}(z) = \mu_{N_i}((z + 1)/2)$ for all $z \in [-1, 1]$. We begin by determining the level corresponding to value $v - v^\star$. This is the level $g = \text{Argmax}_{j\in\{-m,\ldots,m\}}\mu_{Q_j}(v - v^\star)$.

We have the following three cases:

- If $g = 0$, then v is not significantly different from $v^\star$. Hence the context has not played an important role in the evaluation.
- If $g > 0$, the judgment is more tolerant in the current context than what would have been obtained in the reference context. Like for the main argument, the idea is to give to the user the values of the contextual variables for which the difference $v - v^\star$ would also belong to level Q_g. However, doing so means that we show to the user why the difference belongs to Q_g and not to Q_{g-1}, and why the difference belongs to Q_g and not to Q_{g+1}. The second argument is not relevant to the case $g > 0$. Hence we need to explain why the difference belongs to Q_g and not to Q_{g-1}, that is why the difference belongs to Q_g or Q_{g+1}, ..., or Q_m. More precisely, one is interested in the values y'_j of variable j for which $u_k(x, y_{-j}, y_{-j})-v^\star$ belongs to level Q_g or Q_{g+1}, ..., or Q_m. We first determine the smallest values $\underline{U_g}$ such that for all $t \geq \underline{U_g}$, $\text{Argmax}_{l\in\{-m,\ldots,m\}}\mu_{Q_l}(t) \geq g$. One looks thus for the largest interval $(\underline{y_j}, \overline{y_j})$ such that $y_j \in (\underline{y_j}, \overline{y_j})$ and
$$\forall y'_j \in (\underline{y_j}, \overline{y_j}) \ , \ u_k(x, y_j, y_{-j}) - v^\star \geq \underline{U_g}$$
- If $g < 0$, the judgment is more intolerant in the current context than what would have been obtained in the reference context. Like previous case, the idea is to give to the user the values of the contextual variables for which the difference would also belong at most to level Q_g, i.e. to levels Q_g, or $Q_{g-1}, \ldots, Q_1$. More precisely, one is interested in the values y'_j of variable j for which $u_k(x, y_{-j}, y_{-j}) - v^\star$ belongs to level Q_g or Q_{g-1}, ..., or Q_{-m}. We first determine the smallest values $\overline{U_g}$ such that for all $t \leq \underline{U_g}$, $\text{Argmax}_{l\in\{-m,\ldots,m\}}\mu_{Q_l}(t) \leq g$. One looks thus for the largest interval $(\underline{y_j}, \overline{y_j})$ such that $y_j \in (\underline{y_j}, \overline{y_j})$ and
$$\forall y'_j \in (\underline{y_j}, \overline{y_j}) \ , \ u_k(x, y_{-j}, y_{-j}) - v^\star \leq \overline{U_g}$$

All variables do not have the same influence on the utility. Thus, it is not relevant to display to the user the intervals $(\underline{y_j}, \overline{y_j})$ for all variables j. When $g \neq 0$, it makes sense to present to the DM the intervals $(\underline{y_j}, \overline{y_j})$ only for the variables $j \in P_k$ that have a significant importance in the difference $v - v^\star$.

We wish thus to assess the contribution of each variable to the difference $v - v^\star$. It is not easy to differentiate the real contribution of one variable irrespectively of the other ones since all variables vary. Let π be a permutation on P_k. Variables of P_k are considered one at a time in the order given by π. We start with the situation where the contextual variables are set to the reference values $y^\star$. Then each time a variable is considered, the value of this variable is turned to the current value y, and the difference that is obtained is defined as the contribution of this variable. For $S \subseteq P_k$, set $V(S) = u_k(x, y_S, y^\star_{-S})$. The first variable is $\pi(1)$. Its contribution is
$$C_\pi(\pi(1)) := V(\{\pi(1)\}) - V(\emptyset) \ .$$
Likewise, the contribution of the l^{th} variable is
$$\begin{aligned} C_\pi(\pi(l)) =& V(\{\pi(1), \ldots, \pi(l-1), \pi(l)\}) \\ & - V(\{\pi(1), \ldots, \pi(l-1)\}) \end{aligned}$$
One has
$$\sum_{l\in P_k} C_\pi(\pi(l)) = V(N) - V(\emptyset) = v - v^\star \ .$$
Since there is no a priori reason to chose one particular permutation π, the contribution of one variable is defined as the contribution over all possible permutations
$$C(j) = \frac{1}{2^{p_k}} \sum_\pi C_\pi(j) \ .$$
One gets
$$C(j) = \sum_{S\subseteq P_k} \frac{|S|!(p_k - |S| - 1)!}{(p_k)!} \ (V(S \cup \{j\}) - V(S))$$
This expression of $C(j)$ corresponds to the Shapley value of capacity V for variable j [10]. One has
$$\sum_{j\in P_k} C(j) = V(N) - V(\emptyset) = v - v^\star \ .$$

Hence the mean contribution of a variable is $(v - v^\star)/p_k$. The variables that have a valuable contribution to the difference $v - v^\star$ are thus the variables $j \in P_k$ such that $C(j) \geq (v - v^\star)/p_k$.

To sum-up, when $g \neq 0$, we generate the following sentence:

> `The judgment is` Q_g`-more tolerant (if` $g > 0$`)/`Q_g`-more intolerant (if` $g < 0$`) for a value on variable` j `between` $\underline{y_j}$ `and` $\overline{y_j}$ `(if` $C(j) \geq (v - v^\star)/p_k$`).`

6.3 Example

To illustrate our approach, we go back to the example given in Section 5.2.

Consider values $x = 28\,cm$, $y_1 = 15\,m/s$ and $y_2 = 20\,m$. One has $u_k = 0.4$ and the explanation that is generated is:

> `Between 26cm and 28cm, 'sighting shift' is judged 'mean' in the context "Target Speed=15 m/s" and "Distance=20 m". The judgment is more tolerant for a value of 'Target Speed' greater than 13m/s.`

Consider values $x = 15cm$, $y_1 = 15m/s$ and $y_2 = 20m$. One has $u_k = 1$ and the explanation that is generated is:

> `Under 24cm, 'sighting shift' is judged 'very good' in the context "Target Speed=15 m/s" and "Distance=20 m". The judgment is more tolerant for a value of 'Target Speed' greater than 6m/s.`

Consider values $x = 26\,cm$, $y_1 = 6\,m/s$ and $y_2 = 45\,m$. One has $u_k = 0.17$ and the explanation that is generated is:

> `Above 23cm, 'sighting shift' is judged 'very bad' in the context "Target Speed=6 m/s" and "Distance=45 m". The judgment does not depend on 'Distance'.`

References

[1] C. Baba e Costa, E. Corrêa, J.M. De Corte, J.C. Vansnick. Facilitating Bid Evaluation in Public Call for Tenders: A social-technical approach. OMEGA 30(3) (2002) 227-242.

[2] C. Bana e Costa. O Modelo de Apoio á Avaliação de Propostas nos Concursos do Metro de Porto. FER XXI 18, 1999, pp. 111-115.

[3] C. Bana e Costa, J.M. De Corte, J.C. Vansnick. On the mathematical foundations of MACBETH. In: "Multiple Criteria Decision Analysis: state of the art surveys" J. Figueira, S. Greco, M. Ehrgott eds. Springer, New York, 2005.

[4] G. Choquet, Theory of capacities, Annales de l'Institut Fourier 5 (1953) 131-295.

[5] J.C. Fodor and M. Roubens. Fuzzy preferences modeling and multicriteria decision aid. Kluwer Academic Publisher, 1994.

[6] R.L. Keeney, H. Raiffa, Decision with Multiple Objectives, Wiley, New York, 1976.

[7] D.H. Krantz, R.D. Luce, P. Suppes, A. Tversky, Foundations of measurement, vol 1: Additive and Polynomial Representations, Academic Press, San Diego, 1971.

[8] B. Roy. Decision-Aiding today: what should we expect. In: "Multi-Criteria Decision Making: advances in MCDM models, algorithms, theory and applications" T. Gal, T. Steward, T. Hanne eds. Kluwer, Dordrecht, NL, 1999.

[9] T. Saaty. Fundamentals of the Analytic Hierarchy Process. RWS publication, Pittsburgh, PA, 2000.

[10] L. Shapley. A value for n-person games. In: Contributions to the Theory of Games, Vol. II, H. Kuhn, W. Tucker (Eds.). Princeton University Press, pp. 307-317, 1953.

[11] H. Simon. Rational choice and the structure of the environment. Psychological Review 63(2) (1956) 129-138.

[12] P. Slovic, M. Finucane, E. Peters, D.G. MacGregor, The affect heuristic. in: T. Gilovitch, D. Griffin, D. Kahneman (Eds.), Heuristics and biases: the psychology of intuitive judgment, Cambridge University Press, 2002, pp. 397–420.

A Fuzzy Representation for Weights of Alternatives in AHP

Shin-ichi Ohnishi
Faculty of Engineering,
Hokkai-Gakuen University,
Sapporo, Japan
ohnishi@eli.hokkai-s-u.ac.jp

Takahiro Yamanoi
Faculty of Engineering,
Hokkai-Gakuen University,
Sapporo, Japan
yamanoi@eli.hokkai-s-u.ac.jp

Hideyuki Imai
Graduate School of Engineering,
Hokkaido University,
Sapporo, Japan
imai@main.eng.hokudai.ac.jp

Abstract

AHP (Analytic Hierarchy Process) has been widely used in decision making techniques. However, the results from AHP often lose reliability because the comparison matrix does not always have sufficient consistency. In these cases, fuzzy representation for weighting activities using results from a sensitivity analysis is useful. In this paper, we present alternative overall weights by employing some assumptions. Since an idea of less ambiguity is employed, the results show how AHP has fuzziness when the comparison matrix is not sufficiently consistent. An example is also shown in this paper.

Keywords: Decision making, AHP (Analytic Hierarchy Process), Fuzzy sets, Sensitivity analysis, Weights representation.

1 Introduction

AHP (Analytic Hierarchy Process) was proposed by Saaty T.L. in 1977 [1], [2]. The method has been popular and widely used in the domain of decision making, since it can include vagaries such as humans feelings. In addition, it can be developed to ANP (Analytic Network Process) models. However, the comparison matrix often does not have enough consistency when AHP is used since, for instance, a problem may contain too many activities for decision making. In these cases, we consider that answers from decision-makers (i.e. components of the comparison matrix) have ambiguity or fuzziness. For resolving this type of problem, fuzzy reciprocal components have been proposed as components of the data matrix in some research [10]. In this paper, we consider that weights should also have ambiguity or fuzziness. Therefore, it is necessary to represent these weights by use of fuzzy sets.

Sensitivity analysis is applied to AHP to analyze the amount the components of a pairwise comparison matrix influences the weights and consistency of a matrix. This makes it possible to show the magnitude of the fuzziness in the weights.

In previous research, we proposed a new representation for weights of activities [7][8]. For this study, a developed representation of weights of alternatives is proposed. It is represented as L-R fuzzy numbers by using the results from the sensitivity analysis. This paper encompasses methodology to represent weights by fuzzy sets. In addition, a representation of fuzziness as a result of AHP is presented when a comparison matrix does not have enough consistency.

2 Summary of AHP

2.1 Process of AHP

(Process 1) Representation of structure by a hierarchy. The problem under consideration can be represented in a hierarchical structure. The highest level of the hierarchy consists of a unique element that is the overall objective. At the lower levels, there are multiple activities (i.e. elements within a single level) with relationships among elements of the adjacent higher level to be considered. The activities are evaluated using subjective judgments of a decision maker. Elements that lie at the upper level are called parent elements while those that lie at lower level are called child elements. Alternative elements are put at the lowest level of the hierarchy

(Process 2) Paired comparison between elements at each level. A pairwise comparison matrix A is created from a decision maker's answers. Let n be the

number of elements at a certain level. The upper triangular components of the comparison matrix a_{ij} ($i< j$ = 1,…,n) are 9, 8, .. , 2, 1, 1/2, …, or 1/9. These denote intensities of importance from activity i to j. The lower triangular components a_{ji} are described with reciprocal numbers as follows

$$a_{ji} = 1/a_{ij}, \quad (1)$$

in addition, for diagonal elements, let a_{ii} = 1. The lower triangular components and diagonal elements are occasionally omitted from the written equation as they are evident if upper triangular components are shown. The decision maker should make $n(n\text{-}1)/2$ paired comparisons at a level with n elements.

(Process 3) Calculations of weight at each level. The weights of the elements, which represent grade of importance among each element, are calculated from the pairwise comparison matrix. The eigenvector that corresponds to a positive eigenvalue of the matrix is used in calculations throughout in this paper.

(Process 4) Priority of an alternative by a composition of weights. The composite weight can be calculated from the weights of one level lower. With repetition, the weights of the alternative, which are the priorities of the alternatives with respect to the overall objective, are finally found.

2.2 Consistency index

Since components of the comparison matrix are obtained by comparisons between two elements, coherent consistency is not guaranteed. In AHP, the consistency of the comparison matrix A is measured by the following consistency index (C.I.)

$$\text{C.I.} = \frac{\lambda_A - n}{n-1}, \quad (2)$$

where n is the order of matrix A, and λ_A is its maximum eigenvalue.

It should be noted that $\text{C.I.} \geq 0$ holds. And if the value of C.I. becomes smaller, then the degree of consistency becomes higher, and vice versa. The comparison matrix is consistent if the following inequality holds.

$$\text{C.I.} \leq 0.1 \quad (3)$$

3 Sensitivity Analysis of AHP

When AHP is used, the comparison matrix is often inconsistent or large differences among the overall weights of the alternatives do not appear. Thus, it is very important to investigate how the components of a pairwise comparison matrix influence the consistency or weights. Sensitivity analysis is used to analyze how results are influenced when certain variables change. Therefore, it is necessary to establish a sensitivity analysis of AHP.

In our research, a previously proposed method [7] is used to evaluate the fluctuation of the consistency index and weights when a comparison matrix is perturbed. This method is useful as it does not change the structure of the data.

Evaluating the consistency index and the weights of a perturbed comparison matrix are performed as follows.

(1) Perturbations $\varepsilon a_{ij} d_{ij}$ are imparted to component a_{ij} of a comparison matrix, and the fluctuation of the consistency index and the weight are expressed by the power series of ε.

(2) Fluctuations of the consistency index and the weights are represented by the linear combination of d_{ij}.

(3) By the coefficient of d_{ij}, it can be shown that how the component of the comparison matrix gives influence on the consistency index and the weight.

Since the pairwise comparison matrix A is a positive square matrix, the following Perron-Frobenius theorem [4] holds.

Theorem 1 (Perron – Frobenius) *For a positive square matrix A, the following holds true.*

1. *Matrix A has a positive eigenvalue. If λ_A is the largest eigenvalue then λ_A is a simple root. The positive eigenvector* $\mathbf{w}$*, corresponding to λ_A, exists. λ_A is called the Frobenius root of A.*
2. *Any positive eigenvectors of A are the constant multiples of* $\mathbf{w}$*.*
3. *The absolute value of the eigenvalues of A, except for λ_A, is smaller than λ_A.*
4. *The Frobenius root of the transposed matrix A′ is equivalent to the Frobenius root of A.*

This theorem ensures the existence of a weight vector in a pairwise comparison matrix.

From Theorem 1, the following theorem regarding a perturbed comparison matrix holds true [7].

Theorem2 *Let $A = (a_{ij})$, $i,j = 1,\ldots,n$ be a comparison matrix and let $A(\varepsilon) = A+\varepsilon D_A$, $D_A=(a_{ij}d_{ij})$ be a matrix that has been perturbed. Moreover, let λ_A be the Frobenius root of A with $\boldsymbol{w}_1$ being the corresponding eigenvector. Let $\boldsymbol{w}_2$ be the eigenvector corresponding to the Frobenius root of A', then, the Frobenius root $\lambda(\varepsilon)$ of $A(\varepsilon)$ and the corresponding eigenvector $\boldsymbol{w}_1(\varepsilon)$ can be expressed as follows*

$$\lambda(\varepsilon) = \lambda_A + \varepsilon\lambda^{(1)} + o(\varepsilon), \tag{4}$$

$$\boldsymbol{w}_1(\varepsilon) = \boldsymbol{w}_1 + \varepsilon\boldsymbol{w}^{(1)} + \boldsymbol{o}(\varepsilon), \tag{5}$$

where

$$\lambda^{(1)} = \frac{\boldsymbol{w}_2' D_A \boldsymbol{w}_1}{\boldsymbol{w}_2' \boldsymbol{w}_1}, \tag{6}$$

$\boldsymbol{w}^{(1)}$ *is an n-dimension vector that satisfies*

$$(A - \lambda_A I)\boldsymbol{w}^{(1)} = -(D_A - \lambda^{(1)} I)\boldsymbol{w}_1. \tag{7}$$

where $\boldsymbol{o}(\varepsilon)$ *denotes an n-dimension vector in which all components are* $o(\varepsilon)$.

Proof of this theorem can be found in Ohnishi's paper [7].

3.1 Sensitivity analysis of the consistency index

Regarding a fluctuation of the consistency index, the following corollary can be obtained from Theorem 2.

Corollary 1 *Using an appropriate g_{ij}, we can represent the consistency index* C.I.(ε) *of the perturbed comparison matrix as follows*

$$\text{C.I.}(\varepsilon) = \text{C.I.} + \varepsilon\sum_i^n\sum_j^n g_{ij} d_{ij} + o(\varepsilon). \tag{8}$$

(Proof)
From the definition of the consistency index (3) and (4),

$$\text{C.I.}(\varepsilon) = \text{C.I.} + \varepsilon\frac{\lambda^{(1)}}{n-1} + o(\varepsilon).$$

Let $\boldsymbol{w}_1=(w_{1i})$ and $\boldsymbol{w}_2=(w_{2i})$ from (6). $\lambda^{(1)}$ is can now be represented as

$$\lambda^{(1)} = \frac{1}{\boldsymbol{w}_2' \boldsymbol{w}_1}\sum_i^n\sum_j^n w_{2i} a_{ij} w_{1j} d_{ij},$$

therefore, the second part of the right side is expressed by a linear combination of d_{ij}. (Q.E.D)

g_{ij} in equation (8) in Corollary 1 shows the influence of comparison matrix components on the consistency.

On the other hand, since the comparison matrix $A(\varepsilon) = (a_{ij}(\varepsilon))$ is reciprocal, then $a_{ji}(\varepsilon) = 1/a_{ij}(\varepsilon)$ and becomes

$$a_{ji} + \varepsilon a_{ji} d_{ji} = \frac{1}{a_{ij}} - \varepsilon\frac{d_{ij}}{a_{ij}} + o(\varepsilon). \tag{9}$$

Here, since $a_{ji} = 1/a_{ij}$,

$$d_{ji} = -d_{ij} \tag{10}$$

is obtained. The impact on the consistency can be easily shown by use of this property.

3.2 Sensitivity analysis of weights

With regards to the fluctuation in weighs, the following corollary can also be obtained from Theorem 2.

Corollary 2 *Using an appropriate $h_{ij}^{(k)}$, we can represent the fluctuation $\boldsymbol{w}^{(1)}=(w_k^{(1)})$ of the weight (i.e. the eigenvector corresponding to the Frobenius root) as follows*

$$w_k^{(1)} = \sum_i^n\sum_j^n h_{ij}^{(k)} d_{ij}. \tag{11}$$

(Proof)
The k-th row component of the right side of (7) in Theorem 2 is represented as

$$\sum_i^n\sum_j^n \{\frac{w_{1k} w_{2i} a_{ij} w_{1j}}{\boldsymbol{w}_2' \boldsymbol{w}_1} - \delta(i,k) a_{ij} w_{1j}\} d_{ij},$$

and is expressed by a linear combination of d_{ij}. Here,$\delta(i,k)$ is Kronecker's symbol

$$\delta(i,k) = \begin{cases} 1 & (i = k), \\ 0 & (i \neq k). \end{cases}$$

In contrast, since λ_A is a simple root, Rank(A-$\lambda_A I$) = n-1. Accordingly, the weight vector is normalized as

$$\sum_k^n (w_{1k} + \varepsilon w_k^{(1)}) = \sum_k^n w_{1k} = 1,$$

then the condition is as follows.

$$\sum_{k}^{n} w_k^{(1)} = 0. \tag{12}$$

By using an elementary transformation to formula (7) in the condition above, we also can represent $w_k^{(1)}$ by linear combinations of d_{ij}. (Q.E.D)

As seen in equation (5) in Theorem 2, the component that has a great influence on weight $\boldsymbol{w}_1(\varepsilon)$ is the component which has the greatest influence on $\boldsymbol{w}^{(1)}$. $h_{ij}^{(k)}$ in equation (11) from Corollary 2 shows how the influence by the components of a comparison matrix on the weights can be calculated.

The influence can also be shown easily by use of equation (10).

4 Representation of weights using fuzzy sets

The comparison matrix often has poor consistency (i.e. 0.1<C.I.<0.2) because it encompasses several activities. In these cases, the components of a comparison matrix are considered to have fuzziness since they result from the fuzzy judgment of humans. Therefore, weights should be treated as fuzzy numbers.

To represent fuzziness of weight w_{1k}, an L-R fuzzy number is used.

4.1 L-R fuzzy number

L-R fuzzy number

$$M = (m, \alpha, \beta)_{LR}$$

is defined as fuzzy sets whose membership function is as follows.

$$\mu_M(x) = \begin{cases} R\left(\dfrac{x-m}{\beta}\right) & (x > m), \\ L\left(\dfrac{m-x}{\alpha}\right) & (x \le m). \end{cases}$$

where $L(x)$ and $R(x)$ are shape function which satisfies

(1) $L(x) = L(-x)$,

(2) $L(0) = 1$,

(3) $L(x)$ is a non increasing function

4.2 Fuzzy weights of activities

From the fluctuation of the consistency index, the multiple coefficient $g_{ij}h_{ij}^{(k)}$ in Corollary 1 and 2 is considered as the influence on a_{ij}.

Since g_{ij} is always positive, if the coefficient $h_{ij}^{(k)}$ is positive, the real weight of activity k is considered to be larger than w_{1k}. Conversely, if $h_{ij}^{(k)}$ is negative, the real weight of activity k is considered to be smaller. Therefore, the sign of $h_{ij}^{(k)}$ represents the direction of the fuzzy number spread. The absolute value $g_{ij}|h_{ij}^{(k)}|$ represents the size of the influence.

On the other hand, if C.I. becomes bigger, then the judgment becomes more fuzzy.

Consequently, multiple C.I. $g_{ij}|h_{ij}^{(k)}|$ can be regarded as a spread of a fuzzy weight $\tilde{w}_k$ concerned with a_{ij}.

Definition 1 (fuzzy weight) *Let w_{1k} be a crisp weight of activity k, and g_{ij} $|h_{ij}^{(k)}|$ denote the coefficients found in Corollary 1 and 2. If 0.1<C.I.<0.2, then a fuzzy weight $\tilde{w}_k$ is defined by*

$$\tilde{w}_k = (w_{1k}, \alpha_k, \beta_k)_{LR} \tag{13}$$

where

$$\alpha_k = \text{C.I.} \sum_{i}^{n} \sum_{j}^{n} s(-, h_{kij}) g_{ij} \, | h_{kij} \, |, \tag{14}$$

$$\beta_k = \text{C.I.} \sum_{i}^{n} \sum_{j}^{n} s(+, h_{kij}) g_{ij} \, | h_{kij} \, |, \tag{15}$$

$$s(+, h) = \begin{cases} 1, (h \ge 0) \\ 0. (h < 0) \end{cases}, \quad s(-, h) = \begin{cases} 1, (h < 0) \\ 0. (h \ge 0) \end{cases}$$

4.2 Fuzzy weights of alternatives

Using the fuzzy weights of activities defined above and local crisp weights of alternatives with respect to certain activities, we can calculate overall weights from the viewpoint of the overall objective by extension. However, the results from the operation of fuzzy numbers are frequently too ambiguous to interpret.

Fuzzy weights of activities are normalized thus their sum is 1, therefore we can avoid much ambiguity since this condition has been considered [9].

In general, operating with constraints is difficult but can be accomplished if every fuzzy membership

function is linear.

Especially for every normal triangular function with a core u_i, the constraint $\sum_i^n u_i = 1$ holds, and the order of singleton coefficients is assumed. Thus, the upper and lower limit of α -cut sets of linear sum can be easily calculated.

Let $f_t(x_k)$ be a crisp local weight of alternative *t* with respect to activity *k,* and in this paper, assume $0 \le f_t(x_1) \le f_t(x_2) \le \cdots \le f_t(x_n)$. Then, the overall weight of an alternative *t* is also the L-R fuzzy number and is represented as follows.

$$\tilde{v}_t = (v_t, l_t, r_t)_{LR}$$

where

$$v_t = \sum_k^n w_{1k} f_t(x_k),$$

$$l_t = v_t - \inf \operatorname{supp}(\tilde{v}_t), \quad r_t = \sup \operatorname{supp}(\tilde{v}_t) - v_t$$

In the above equations, inf supp, sup supp are lower and upper limits of support sets and are calculated as follows.

$$\inf \operatorname{supp}(\tilde{v}_t) = \max_j \left[\sum_{i=1}^{j-1} (w_{1i} + \beta_i) f_t(x_i) + \sum_{i=j+1}^{n} (w_{1i} - \alpha_i) f_t(x_i) + \left\{ 1 - \sum_{i=1}^{j-1} (w_{1i} + \beta_i) - \sum_{i=j+1}^{n} (w_{1i} - \alpha_i) \right\} f_t(x_j) \right]$$

$$\sup \operatorname{supp}(\tilde{v}_t) = \min_j \left[\sum_{i=1}^{j-1} (w_{1i} - \alpha_i) f_t(x_i) + \sum_{i=j+1}^{n} (w_{1i} + \beta_i) f_t(x_i) + \left\{ 1 - \sum_{i=1}^{j-1} (w_{1i} - \alpha_i) - \sum_{i=j+1}^{n} (w_{1i} + \beta_i) \right\} f_t(x_j) \right]$$

5 Example

In this section, we introduce an example of an employment selection problem, whose structure is shown in Figure 1.

Table 1 shows a comparison matrix of activities, and Table 2 shows weights of alternatives with respect to activities.

Calculated weights of activities are as follows.

$$\begin{pmatrix} \text{Salary} \\ \text{Stability} \\ \text{Prospect} \\ \text{Attractive} \\ \text{Holiday} \\ \text{Welfare} \end{pmatrix} = \begin{pmatrix} 0.04 \\ 0.41 \\ 0.26 \\ 0.16 \\ 0.04 \\ 0.09 \end{pmatrix}$$

However, they do not have much reliability, because consistency of the matrix is not so good (C.I.=0.13).

Table 3 shows a result of sensitivity analysis of consistency. The comparison value between salary and stability, or the value between holiday and welfare has much influence.

Table 4 shows fuzzy weights of activities. A weight of stability has most fuzziness.

Table 1: A comparison matrix of activities

	Salary	Stability	Prospect	Attractive	Holiday	Welfare
Salary	1	1/5	1/5	1/5	1/2	1/3
Stability		1	3	4	7	5
Prospect			1	3	6	5
Attractive				1	7	3
Holiday					1	1/5
Welfare						1

Table 2: Weights of alternatives with respect to activities.

	Company A	Company B	Company C
Salary	0.158	0.766	0.076
Stability	0.042	0.180	0.778
Prospect	0.180	0.778	0.042
Attractive	0.070	0.751	0.178
Holiday	0.157	0.249	0.594
Welfare	0.121	0.115	0.764

Table 3: Result of sensitivity analysis of consistency

	Salary	Stability	Prospect	Attractive	Holiday	Welfare
Salary	0.0000	0.3602	0.2291	0.1414	0.0847	0.1279
Stability	0.0789	0.0000	0.3254	0.2677	0.1123	0.1817
Prospect	0.1199	0.0864	0.0000	0.3052	0.1463	0.2761
Attractive	0.1907	0.1031	0.0874	0.0000	0.2715	0.2635
Holiday	0.3394	0.2620	0.1945	0.1028	0.0000	0.0782
Welfare	0.2131	0.1536	0.0977	0.1005	0.3613	0.0000

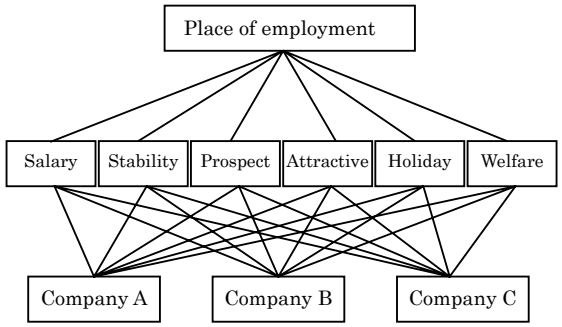


Figure 1: A hierarchical structure

Table4: fuzzy weights of activities

	Center	Spread(L)	Spread(R)
Salary	0.0381	0.0011	0.0012
Stability	0.4118	0.0085	0.0089
Prospect	0.2620	0.0059	0.0059
Attractive	0.1616	0.0026	0.0040
Holiday	0.0387	0.0013	0.0013
Welfare	0.0877	0.0022	0.0023

Finally we can find the overall alternative weights (Table5).

Table5: fuzzy weights of alternatives

	Center	Spread(L)	Spread(R)
Company A	0.0983	0.0306	0.0306
Company B	0.4484	0.1476	0.1659
Company C	0.4532	0.1872	0.1723

6. Conclusions

We proposed a representation for the overall weights of alternatives by use of fuzzy sets and the result of a sensitivity analysis for cases in which consistency of the comparison matrix is poor. Our approach shows how to represent weights, as well as how the result of AHP has fuzziness, when inconsistency exists. This was due to reduced ambiguity in the representation presented in this work compared to previous normal fuzzy operations. Moreover we show an example of the alternative fuzzy weights.

Acknowledgements

This research is partly supported by a grant from the ministry of education, sports, science, and technology to the national project of 'Advanced improvements of vision, image, speech and language information processing and the application to the technologies for the intelligent instrument and control' in the High-tech Research Center of Hokkai-Gakuen University.

References

[1] T. L. Saaty. A scaling method for priorities in hierarchical structures. J. Math. Psy., 15(3):234--281, 1977.

[2] T. L. Saaty. The Analytic Hierarchy Process. McGraw-Hill, New York, 1980.

[3] T. L. Saaty. Scaling the membership function. European J. of O.R., 25:320--329, 1986.

[4] M. Saito. An Introduction to Linear Algebra. Tokyo University Press, 1966.

[5] Y. Tanaka. Recent advance in sensitivity analysis in multivariate statistical methods. J. Japanese Soc. Comp. Stat., 7(1):1--25, 1994.

[6] K. Tone. The Game Feeling Decision Making. Nikka-giren Press, Tokyo, 1986.

[7] S. Ohnishi, H. Imai, and M. Kawaguchi. Evaluation of a stability on weights of Fuzzy Analytic Hierarchy Process using a sensitivity analysis. J. Japan Soc. for Fuzzy Theory and Sys., 9(1):140--147, 1997..

[8] S. Ohnishi, H. Imai, T. Yamanoi. Weights Representation of Analytic Hierarchy Process by use of Sensitivity Analysis, IPMU 2000 Proceedings, 2000.

[9] D. Dubois and H. Prade. Possibility Theory An Approach to Computerized Processing of Uncertainty,. Plenum Press, New York, 1988.

[10] S. Ohnishi , D. Dubois, H. Prade, T. Yamanoi, A Study on the Analytic Hierarchy Process Using a Fuzzy Reciprocal Matrix, IPMU 2006 Proceedings 2006.

Dealing with Incompleteness of Preferences in Group Decision Making Problems

Anna Pankowska
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Umultowska 87, 61-614 Poznań, Poland
aniap@amu.edu.pl

Abstract

While considering processes of decision making we often encounter the problem of incomplete information. In group decision making (GDM) problems each decision maker is supposed to provide a matrix that describes his/her preferences over the set of given options. However, we have to take into account that an expert can not be able to define his/her preferences about all the options. Usually the problem is solved by an additional phase of estimating missing values. In the article we want to suggest and discuss a totally different approach that consists in adopting GDM algorithm so as it can deal with incomplete preference matrices.

Keywords: group decision making, incomplete preference relations, Atanassov's intuitionistic fuzzy sets.

1 Introduction

Group decision making (GDM) procedures are used in processes of choosing one best option from the set of options. A central notion of these procedures is a preference matrix that models preferences of a decision maker. Possible options are compared pairwise, and to each pair a value, that represent a degree to which one option is preferred over the other, is assigned. It is usually assumed that such a value is given for every pair of options, forming a complete preference relation. However, worth considering is a case, when some of these values are missing for some reason like the lack of sufficient knowledge of an expert, impossibility of comparing some options under particular criterion, ignoring an option by an expert or simply loss of some information during the process. Algorithms of GDM should thus be able to deal with incomplete knowledge. Different approaches to that problem have been proposed in the literature, most of which try to complete somehow the incomplete preference matrices and then to apply a GDM algorithm. However, it is not always possible (and probably not always necessary) to fill all the missing values. In this article we like to suggest a different approach, namely: we leave the missing values not filled, and adopt the GDM algorithms to deal with incomplete preference relations. We examine and compare these two approaches, and discuss advantages and drawbacks of both of them.

2 Main aspects of GDM problems

First let us briefly describe the main idea of the class of GDM algorithms based on IF-sets (Atanassov's intuitionistic fuzzy sets, see [1]) with triangular norms. In the following, we focus on preference relations used in those algorithms.

The problem of group decision making consists in finding one compromise option that best suits the preferences of the whole group of decision makers. By $\boldsymbol{P} = \{p_1, ..., p_m\}$ we denote a set of $m \geq 1$ *decision makers (experts)*, and $\boldsymbol{S} = \{s_1, ..., s_n\}$ is a set of $n \geq 2$ *options*. Each expert p_k constructs his/her preference matrix $R_k = [r_{ij}^k]$ with $i, j = 1, 2, ..., n$, and $r_{ij}^k \in [0, 1]$, which represents his/her preferences about options from $\boldsymbol{S}$. Finally, Q denotes a relative linguistic quantifier of "most"-type (see [6]).

The process of GDM can be performed in two ways, as proposed by Kacprzyk in [6]: directly and indirectly.

In the direct approach a solution S_Q is obtained directly from preference matrices:

$$\{R_1, ..., R_m\} \to S_Q.$$

S_Q is then a fuzzy set of options such that a soft majority, Q experts, are not against them.

When applying the indirect approach the individual matrices $R_1, ..., R_m$ are first aggregated into one group

preference matrix:

$$\{R_1, ..., R_m\} \to R \to W_Q$$

and next the solution W_Q is constructed. It is a fuzzy set of options that are preferred to Q other options. This solution is known also as a fuzzy Q-consensus winner ([6]).

Depending on the used t-norm we divide the algorithms into another two subclasses. The first one refers to $\boldsymbol{t}$ being a nonstrict Archimedean t-norm with a negation induced by $\boldsymbol{t}$ ($\nu = \nu_{\boldsymbol{t}}$), and the second one - to $\boldsymbol{t}$ being a strict t-norm or $\boldsymbol{t} = \wedge$ with the Łukasiewicz negation $\nu = \nu_L$. This distinction is important from the point of view of GDM as it results in different semantic interpretation of a hesitation index. When using non-strict Archimedean t-norms hesitation is interpreted in classical way, but when using strict t-norm or $\boldsymbol{t} = \wedge$ hesitation index combines hesitation with fuzziness. For more details see [9]. In the article we restrict ourselves to the case of a nonstrict Archimedean t-norm and indirect approach.

2.1 Modeling preferences on the basis of IF-set theory with triangular norms

As already mentioned, each decision maker provides his/her preferences by means of a fuzzy preference relation $R_k : \boldsymbol{S} \times \boldsymbol{S} \to [0, 1]$ that can be represented by a $n \times n$ preference matrix R_k. Each element r_{ij}^k of that martix is a *degree to which an expert p_k prefers an option s_i over an option s_j*. Of course, $r_{ij}^k \in [0, 1]$; $r_{ij}^k = 0.5$ indicates indifference between s_i and s_j. Clearly, the greater the value of r_{ij}^k, the higher the preference of s_i over s_j.

It is usually assumed that $r_{ij}^k = 1 - r_{ji}^k$. However, this restriction enables to model only imprecision (vagueness) of preferences, but not uncertainty (hesitation of an expert). The theory that gives a way to incorporate uncertainty is the Atanassov's intuitionistic fuzzy set theory (IF-set theory). An IF-set $\mathcal{E}$ is a pair of fuzzy sets: $\mathcal{E} = (A, A^d)$, where A is as fuzzy set of elements that belong to $\mathcal{E}$, and A^d is dual to A and contains elements that do not belong to $\mathcal{E}$. In the classical fuzzy set theory ([13]) $A^d = A'$, where $A'(x) = 1 - A(x)$ for each x. In IF-set theory $A^d \subset A'$, so a margin of uncertainty is left. It is called *hesitation index*, and is defined as:

$$\chi_{\mathcal{E}}(x) = 1 - A(x) - A^d(x) \tag{1}$$

for each x.

As presented in [8, 9] the theory of IF-sets can be generalized by using arbitrary strong negation ν instead of the Łukasiewicz negation and arbitrary t-norm $\boldsymbol{t}$ instead of the Łukasiewicz t-norm. Thus we rewrite the main condition as $A^d \subset A^\nu$. The formula (1) will take the form:

$$\chi_{\mathcal{E}}(x) = \nu(A(x)\, \boldsymbol{t}\, ^\nu A^d(x)) = \nu(A(x))\, \boldsymbol{t}\, \nu(A^d(x)), \tag{2}$$

where $\boldsymbol{t}^\nu$ is a t-conorm defined as $a \boldsymbol{t}^\nu b = \nu(\nu(a) \boldsymbol{t} \nu(b))$.

Applying this theory to the preference modeling, we will replace the condition $r_{ij}^k = 1 - r_{ji}^k$ with more flexible one $r_{ij}^k \le 1 - r_{ji}^k$, or, in general, to $r_{ij}^k \le \nu(r_{ji}^k)$, what gives us the possibility to model hesitation of the decision maker in a flexible way. That hesitation is calculated for each expert as

$$h_{ij}^k = \nu(r_{ij}^k)\, \boldsymbol{t}\, \nu(r_{ji}^k) \tag{3}$$

and stored in a hesitation matrix $H_k = [h_{ij}^k]$.

2.2 The algorithm of GDM

Here we present the algorithm of GDM for nonstrict Archimedean t-norms and indirect approach. Functions f, f^*, g, g^* are cardinality patterns (see [12]).

Step 1 Aggregate individual preference matrices into one group preference matrix $R = [r_{ij}]$:

$$r_{ij} = \frac{1}{m} \sum_{k=1}^{m} f(r_{ij}^k).$$

A value r_{ij} represents a degree to which an option s_i is preferred over an option s_j by the whole group of decision makers.

Step 2 Calculate a degree r_i to which an option s_i is preferred over all the other options:

$$r_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} f^*(r_{ij})$$

Step 3 Construct hesitation matrices $H_k = [h_{ij}^k]$ according to (3). Next, aggregate them into one group hesitation matrix $H = [h_{ij}]$:

$$h_{ij} = \frac{1}{m} \sum_{k=1}^{m} g(h_{ij}^k).$$

Step 4 Calculate a degree h_i to which a whole group is hesitating whether to choose an option s_i or some other option:

$$h_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} g^*(h_{ij}).$$

Step 5 Compute $Q(r_i)$ and $Q(r_i \boldsymbol{t}^\circ h_i)$ for $i = 1, 2, ..., n$ that is a degree to which an option s_i is preferred over

Q other options by the group of experts, taking into account also hesitation of the group.

Step 6 The final solution is:

$$W_Q = [Q(r_1), Q(r_1 \boldsymbol{t}^\circ h_1)]/s_1 + ... \\ ... + [Q(r_n), Q(r_n \boldsymbol{t}^\circ h_n)]/s_n.$$

3 Incompleteness in GDM problems

So far we required an expert to compare every pair of options. This requirement however is not always possible to fulfil. The reasons could be varied: an expert may consider some options uninteresting for him/her for some reason, or may have only partial knowledge about some certain aspects or about a subset of options. It is important to distinguish between this situation and the situation when two options were considered indifferent (i.e. $r_{ij}^k = r_{ji}^k = 0.5$). The case of a missing value rather resembles the situation when $r_{ij}^k = 0$ and $r_{ji}^k = 0$. However, an expert should have a possibility to express his/her lack of knowledge, and must not be forced to give every value of R_k if he/she is not able to do this - such a demand could lower information quality.

The concept that allows to express the situation, when some values r_{ij}^k are missing is the concept of *incomplete fuzzy preference relation* R_k. It is defined as a fuzzy set $R_k : \boldsymbol{S} \times \boldsymbol{S} \rightarrow [0,1]$ characterized by a partial membership function. By *a partial function* we mean a function which is not defined for some of its domain.

An undefined value of a fuzzy preference relation R_k will be denoted by a symbol *, and will be called a *missing value*. A set

$$MR_k = \left\{(s_i, s_j) \in R_k : r_{ij}^k = *\right\} \tag{4}$$

is a set of all pairs of options for which preference values were not defined by an expert p_k.

From now on, to avoid any misunderstanding, the fuzzy preference relation with no missing values will be called *complete* fuzzy preference relation.

As GDM algorithms were constructed for complete fuzzy preference matrices, we have to introduce some new methods to deal with incomplete fuzzy preference matrices. There are two main possible approaches we can apply:

- complete missing values:
  - with some neutral value, like 0.5, or a mean value;
  - according to some external information, e.g. taken from preference relation of another expert(s), or by interacting with an expert;
  - according to some additional conditions - usually transitivity condition;
- do not complete missing values and adopt the algorithm of GDM so as it can operate on incomplete preference matrices.

The first approach is wildly discussed in the literature (see e.g. [2, 3, 4, 5]). We would like to point out some problems connected with it:

- Estimating missing values according to some additional property, e.g. transitivity property, is possible only under certain conditions. Moreover, as experimental studies show (see [11]), in real-life situation preference relations are inconsistent and violate transitivity condition. In such cases estimated values may not reflect the real intention of the expert.
- Interactive procedures that cooperate with an expert during the process of completing missing values, are often very hard to execute.
- When using some external knowledge, e.g. from another experts, we do not take into account the specific character of a particular expert.
- Filling missing values with a mean, random or neutral value is very simple but too rigid.
- All of these techniques introduce some additional information that was not provided in the original preference matrix, thus can disturb its meaning.

Completing missing values is particularly difficult if a degree of incompleteness of a preference matrix is high. In such a case it is very hard to complete all missing values in a reasonable way.

Therefore, we like to propose a second approach that do not change preference matrices but an algorithm itself. We belief that such an approach can form a good alternative for estimating in some cases of GDM problems. The advantages of this approach are, among others:

- We do not have to process any additional calculation to estimate missing values.
- We do not disrupt the original preference relations, and we are not loosing the information which *lack of information* also is (as we have mentioned, inserting some values in place of empty values can distort the intention of an expert, as we add some extra information that an expert didn't actually provide).

- We are able to complete the process of solving GDM problem even if we have no possibility to estimate all of the missing values.

Such an approach generates of course some new problems, such as:

- The solution of the algorithm could be in some cases incomplete - it would happen if we had not enough information about some option.
- We are not using any additional knowledge that could have been helpful.

The group of algorithms of GDM for incomplete fuzzy preference relations was developed and one of the algorithm, for nonstrict Archimedean t-norms and for indirect approach, is presented in the next subsection.

3.1 An algorithm of GDM for incomplete preference matrices

As we defined in (4), MR_k is a set of pairs of options for which an expert p_k did not define his/her preference. At the beginning all values r_{ij}^k that are missing, but for which r_{ji}^k is given and $r_{ji}^k \geq a^*$, are calculated as $\nu(r_{ji}^k)$ and are no longer considered as missing. a^* is a fixed point of a negation ν (for Łukasiewicz negation $a^* = 0.5$).

Below we present GDM algorithm for nonstrict Archimedean t-norms and indirect approach.

Step 1

$$E_{ij} = \left\{(s_i, s_j) \in R_k : r_{ij}^k = *, k = 1, 2, ..m\right\},$$

$$r_{ij} = \begin{cases} \frac{1}{m-|E_{ij}|} \sum\limits_{k=1}^{m} f(r_{ij}^k), & \text{if } e_{ij} < m \\ * & \text{otherwise.} \end{cases}$$

Step 2

$$E_i = \{(s_i, s_j) \in R : r_{ij} = *, j = 1, 2, ...n\},$$

$$r_i = \begin{cases} \frac{1}{n-1-|E_i|} \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} f^*(r_{ij}), & \text{if } e_i < n-1 \\ * & \text{otherwise.} \end{cases}$$

Step 3

$$h_{ij} = \begin{cases} \frac{1}{m-|E_{ij}|} \sum\limits_{k=1}^{m} g(h_{ij}^k), & \text{if } e_{ij} < m \\ * & \text{otherwise.} \end{cases}$$

Step 4

$$h_i = \begin{cases} \frac{1}{n-1-|E_i|} \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} g^*(h_{ij}), & \text{if } e_i < n-1 \\ * & \text{otherwise.} \end{cases}$$

Step 5 For all defined r_i and h_i compute $Q(r_i)$ and $Q(r_i \boldsymbol{t}^\circ h_i)$ for $i = 1, 2, ..., n$. Put $Q(r_i) = *$ and $Q(r_i \boldsymbol{t}^\circ h_i) = *$ otherwise.

Step 6 Finally,

$$W_Q = [Q(r_1), Q(r_1 \boldsymbol{t}^\circ h_1)]/s_1 + ... \\ ... + [Q(r_n), Q(r_n \boldsymbol{t}^\circ h_n)]/s_n.$$

It is easy to notice that the solution may be incomplete. It is also important to be aware of the fact that the solution could possibly be obtained from incomplete information. It seems useful to introduce a factor that will signify a degree of completeness for a given option. It will be determined as a ratio of number of missing values involving option s_i to all preference values involving s_i. So, we define:

$$MR_k^i = \{(s_i, s_j) \in MR_k \cup (s_j, s_i) \in MR_k, j = 1, 2, ..., n)\}$$

and the *completness factor* of an option s_i is

$$CF_i = 1 - \frac{\sum\limits_{k=1}^{m} MR_k^i}{m(2n-2)}. \tag{5}$$

Obviously, $CF_i \in [0, 1]$. $CF_i = 0$ means total incompleteness (no preferences were defined for any pair of options involving s_i), and $CF_i = 1$ means total completeness (i.e. option s_i was compared with all the other options).

4 An example of solving GDM problem with incomplete preference relations

A company X decided to introduce a new report software. Experts from three departments, p_1, p_2 and p_3, were asked to compare and assess five different reporting tools, s_1, s_2, s_3, s_4 and s_5, in order to choose and apply the best one.

As a result, the experts provided three incomplete matrices describing their preferences about given options of the software:

$$R_1 = \begin{bmatrix} 0.5 & * & 0 & * & * \\ * & 0.5 & * & * & * \\ 0.9 & * & 0.5 & * & 0.2 \\ * & * & * & 0.5 & 0.3 \\ 1 & * & 0.7 & 0.7 & 0.5 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} 0.5 & * & * & * & * \\ * & 0.5 & * & * & * \\ * & * & 0.5 & * & 0.5 \\ 0.9 & * & * & 0.5 & 0.6 \\ * & * & 0.5 & 0.4 & 0.5 \end{bmatrix},$$

$$R_3 = \begin{bmatrix} 0.5 & * & * & * & * \\ * & 0.5 & * & * & * \\ * & * & 0.5 & * & * \\ * & * & * & 0.5 & * \\ 1 & * & * & 1 & 0.5 \end{bmatrix}.$$

Each expert defined the preferences according to his/her knowledge, personal experience, and taking into account the specific needs and tasks of their departments. That is why their preferences differ to a large extend from each other and are not complete.

We have to emphasize that incompleteness of the preference matrices cannot be treated here as an aberration or a defect that has to be fixed, but is a natural property resulting from the fact that every expert is a specialist only in some part of the problem and has a very subjective point of view. Therefore, completing missing values in this case seems to be unjustified and improper. As matrices are incomplete to high degree it is also difficult or even impossible to complete them in a reasonable way.

Instead of completing missing values, we apply the version of GDM algorithm for incomplete preference matrices presented in the previous section. For simplicity, we assume that that $f = f^* = g = g^* = id$, $\boldsymbol{t} = \boldsymbol{t}_L$, $\nu = \nu_L$, $Q = id$ (setting these parameters to different values is reasonable when we try to optimize the algorithm for a specific application).

First, we make use of a property of a preference matrix saying that $r^k_{ij} \leq \nu(r^k_{ji})$, and for all missing values r^k_{ij} for which $r^k_{ji} \geq a^*$ ($a^* = 0.5$) we calculate r^k_{ij} as $1 - r^k_{ji}$.

Next, individual preference matrices are aggregated into one group preference matrix that will combine the knowledge of the experts. This matrix can be complete or incomplete, but its degree of completeness will be the same or higher then a degree of completeness of individual matrices. This feature is a result of an algorithm that puts together information from all the experts and completes missing values in one preference matrix with values from another preference matrix, if given.

The group preference matrix is presented below:

$$R = \begin{bmatrix} 0.5 & * & 0 & 0.1 & 0 \\ * & 0.5 & * & * & * \\ 0.9 & * & 0.5 & * & 0.35 \\ 0.9 & * & * & 0.5 & 0.3 \\ 1 & * & 0.6 & 0.7 & 0.5 \end{bmatrix}.$$

At the same time, we calculate individual hesitation matrices and a group hesitation matrix, which is presented below:

$$H = \begin{bmatrix} 0.5 & * & 0.1 & 0 & 0 \\ * & 0.5 & * & * & * \\ 0.1 & * & 0.5 & * & 0.05 \\ 0 & * & * & 0.5 & 0 \\ 0 & * & 0.05 & 0 & 0.5 \end{bmatrix}.$$

The last aggregation step for each option returns a degree to which this option is preferred over all other options according to the whole group of experts. Finally, we get the solution W_Q:

$$W_Q = (0.03, 0.07)/s_1 + (*)/s_2 + (0.63, 0.70)/s_3 + (0.60, 0.60)/s_4 + (0.77, 0.78)/s_5.$$

The option s_5 was thus chosen by experts to be the best one.

The solution is incomplete (there is no value assigned to option s_2) as none of the experts said anything about the reporting tool s_2. That might suggest that they do not know this tool or consider it highly improper and uninteresting for the company. It may be treated as a hint that this option should no longer be taken into consideration.

To get more information about the quality of the solution, we calculate additional factor for each option, CF_i, according to (5) (in the calculation we excluded option s_2):

$$CF_1 = 0.44,$$
$$CF_3 = 0.33,$$
$$CF_4 = 0.44,$$
$$CF_5 = 0.77.$$

The option s_5 was compared by experts with most of other options, thus a value of CF_5 is quite high. Therefore, we may conclude that the choice of this option as the best one is right and well motivated.

5 Conclusions

The article presents a new method for dealing with incompleteness of information in GDM problems. This

approach, unlike those that are usually discussed in the literature, does not fill missing values. Instead, it changes the very GDM algorithm. We pointed out some difficulties connected with estimating missing values and we shown an application of new version of algorithm for incomplete preference relations. We considered the case when the degree of incompleteness is too high to estimate or complete missing values in a reasonable way. In such cases the proposed modified algorithm is a worth considering alternative tool.

References

[1] Atanassov K., Stoeva S. (1983), *Intuitionistic fuzzy sets*, w: *Proc. Polish Symp. Interval and Fuzzy Mathematics*, Poznań, 23-26.

[2] Danan E., Ziegelmeyer A. (2004), *Are preferences incomplete? An experimental study using flexible choices*, Discussion Papers on Strategic Interaction 2004-23, Max Planck Institute of Economics, Strategic Interaction Group.

[3] Herrera-Viedma E., ChiclanaF. , Herrera F., Alonso S. (2006), *A Group Decision-Making Model with Incomplete Fuzzy Preference Relations Based on Additive Consistency*, IEEE Transactions on Systems, Man and Cybernetics, Part B, vol. 37, 176-189.

[4] Herrera-Viedma E., Chiclana F. , Herrera F., Alonso S. (2005), *Group Decision Making With Incomplete Information*, Procesos de Toma de Decisiones, Modelado y Agregacin de Preferencias. Herrera-Viedma, E. (Ed.), Granada, pp. 21-30.

[5] Herrera-Viedma E., Alonso S., Herrera F., Chiclana F., Porcel C. (2006), *A General Procedure to Estimate Missing Values for Incomplete Fuzzy Preference Relation*, 11th Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU-2006), Paris (France), pp. 1796-1802.

[6] Kacprzyk J. (1986), *Group decision making with a fuzzy linguistic majority*, Fuzzy Sets and Systems 18, 105-118.

[7] Nurmi H. (1981), *Approaches to collective decision making with fuzzy preference relations*, Fuzzy Sets and Systems 6, 249-259.

[8] Pankowska A., Wygralak M. (2004), *A General Concept of IF-Sets with Triangular Norms*, Soft Computing - Foundations and Theoretical Aspects, EXIT, 319-335

[9] Pankowska A., Wygralak M. (2006), *General IF-sets with triangular norms and their applications to group decision making*, Information Sciences, 176, 2713-2754.

[10] Szmidt E., Kacprzyk J. (1998), *Group decision making under intuitionistic fuzzy preference relations*, in: *Proc. 7th IPMU Conf.*, Paris, 172-178.

[11] Świtalski (2001), *Transitivity of fuzzy preference relationsan empirical study*, Fuzzy Sets and Systems 118, 503-508.

[12] Wygralak M. (2003), *Cardinalities of Fuzzy Sets*, Springer-Verlag, Berlin Hei-delberg.

[13] Zadeh L.A. (1965), *Fuzzy sets*, Inform. and Control 8, 338–353.

Session 10

Fuzzy Relations

Algorithms for the Construction of Digital Convex Fuzzy Hulls

Nebojša M. Ralević
Faculty of Engineering, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad,
Serbia
nralevic@uns.ns.ac.yu

Lidija Čomić
Faculty of Engineering, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad,
Serbia
comic@uns.ns.ac.yu

Abstract

We consider a digital fuzzy set placed in Oxy-plane, with support which is a set of digital points (centroids) in $\mathbb{Z}^2$. We consider two kinds of convexity of a fuzzy set, namely quasi convexity and strong convexity, and we propose two algorithms for the construction of both kinds of convex hull of a digital fuzzy set.

Keywords: fuzzy set, digital fuzzy set, quasiconvexity, convexity, DL-convex hull, digital quasiconvex fuzzy hull, digital convex fuzzy hull.

1 Introduction

Convexity is one of the basic properties of sets and shapes. Convex hull $H(S)$ of a set S is defined as the smallest convex set containing S. Convex hull is used as a tool in many different areas. In robotics, for example, it is used in collision detection, while in shape analysis, the difference between the shape and its convex hull is used as a shape descriptor.

The problem of the computation of the convex hull has been considered for a long time in many research areas, such as computational geometry, image analysis, and fuzzy set theory. In each of these areas, the problem has specific properties. In computational geometry, a finite set S of points in $\mathbb{R}^n$ is given, so points have arbitrary real coordinates [6]. In image analysis, points are restricted to have integer coordinates.The notion of the convex hull $H(I)$ of an image I depends on the way a convex set in $\mathbb{Z}^n$ is defined. Most often, T-convexity, L-convexity, and DL-convexity are used [2]. In fuzzy set theory, the notion of convexity is also not unique. Most often, quasiconvexity, (strong) convexity, and α-convexity are considered. Moreover, the universal set of a fuzzy set A can be either $\mathbb{R}^n$ or $\mathbb{Z}^n$. Here, we consider the problem of the computation of the quasiconvex hull $\mathcal{QH}(A)$ and the (strong) convex hull $\mathcal{H}(A)$ of a fuzzy set A, with support $\mathrm{supp}(A)$ which is a digital set, and we propose algorithms for computation of $\mathcal{QH}(A)$ and $\mathcal{H}(A)$ in 2D. Since a gray-scale image I may be considered as a fuzzy set with a discrete domain, our algorithms can be used to find convex hulls of 2D gray-scale images.

2 Background Notions

A set S in Euclidean space is convex if for any two points P and Q we have:

$$\text{for any } \lambda \in [0,1],\ \lambda P + (1-\lambda)Q \in S.$$

The convex hull $H(S)$ of a set S is the intersection of all convex sets that contain S:

$$H(S) = \bigcap_{i \in I} \{S_i | S \subset S_i, S_i \text{ is convex}\}.$$

A set S is convex iff $H(S) = S$.

Euclidean plane Oxy ($\mathbb{R}^2$) can be digitized using specific regular tessellation; each square is called a pixel. Set $\mathcal{D}$ of all pixels (of $\mathbb{R}^2$) is called a digital space:

$$\mathcal{D} = \{[i-0.5, i+0.5) \times [j-0.5, j+0.5) | (i,j) \in \mathbb{Z}^2\}.$$

Set R (or region) in the digital space is the union of a set of pixels, and we call it a digital set. The centroid of a pixel is the point of intersection of the two diagonals of the pixel. The union of the centroids of the pixels of R, denoted by R', may be considered as the lattice point or simply a point representation of R. Conversely, R may be called the region representation of R'.

Clearly, there exists a 1-1 correspondence δ between digital space $\mathcal{D}$ and set $\mathbb{Z}^2$, and thus between R and R'. We call δ digitalization.

Definition 1 *Digital straight line segment P_1P_2 is the set of all pixels that are intersected by Euclidean*

straight line segment p_1p_2, where p_1, p_2 are centroids (digital points) of pixels P_1, P_2.

Definition 2 *A digital region R represented by R' is called $DL-$convex if for any two digital points p_1, p_2 belonging to R', there is a digital straight line segment between P_1 and P_2 whose pixels all belong to R.*

Definition 3 *$H_{DL}(R)$, the $DL-$hull of R is the minimal $DL-$convex set that contains R.*

An algorithm for the computation of a DL-hull is given in [2].

Let us present some basic definitions and theorems related to fuzzy sets and their convexity, and extend them on digital fuzzy sets. For more details, see [3, 4, 7].

Definition 4 *A fuzzy set A defined on the universal set X is a set of ordered pairs: $A = \{(x, \mu_A(x))|x \in X\}$ where $\mu_A : X \to [0, 1]$ is the membership function of A in X.*

Definition 5 *The support of a fuzzy set A is the set*

$$\text{supp}(A) = \{x \in X|\mu_A(x) > 0\}.$$

Definition 6 *An α-cut of a fuzzy set A, for $\alpha \in [0, 1]$, is the set:*

$$A_\alpha = \{x \in X|\mu_A(x) \geq \alpha\ \}.$$

Definition 7 *A fuzzy set A is a subset of a fuzzy set B, denoted $A \subseteq B$, iff*

$$\mu_A(x) \leq \mu_B(x),$$

for all $x \in X$.

Definition 8 *A fuzzy set A of a set $\mathbb{R}^n$, given by the membership function μ_A, is called quasiconvex if*

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \min\{\mu_A(x_1), \mu_A(x_2)\}$$

holds for every $x_1, x_2 \in \text{supp}(A)$, and $\lambda \in [0, 1]$ (see [1], [8]).

Theorem 1 *A fuzzy set A is quasiconvex if and only if its α-cuts are empty or convex sets for each $\alpha \in [0, 1]$ (see [5]).*

Definition 9 *A fuzzy subset A of a set $\mathbb{R}^n$, given by the membership function μ_A, is called convex if*

$$\mu_A(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda\mu_A(x_1) + (1 - \lambda)\mu_A(x_2)$$

holds for every $x_1, x_2 \in \text{supp}(A)$, and $\lambda \in [0, 1]$.

Theorem 2 *A fuzzy set A is convex if and only if the fuzzy hypograph of A, denoted by $f.hyp(A)$, is a convex set, where*

$$f.hyp(A) = \{(x, t)|x \in \mathbb{R}^n, t \in (0, \mu_A(x)]\}$$

(see [1]).

If A is a convex fuzzy set, then for every $\alpha \in [0, 1]$, α-cut of A is convex (or empty), but the converse does not hold in general.

For example, a fuzzy subset A of $\mathbb{R}$, given by its membership function $\mu_A(x) = \frac{1}{1+x^2}$, $x \in \mathbb{R}$ is quasiconvex, but it is not convex.

Definition 10 *A (quasi) convex fuzzy hull, ($qconv(A)$) $conv(A)$ of a fuzzy set A is the smallest (with respect to relation $\subseteq$) (quasi) convex fuzzy set containing A.*

If A is a convex fuzzy set, then the convex fuzzy hull of A is equal to A.

Definition 11 *A digital fuzzy set is a fuzzy set which is defined on digital space $\mathcal{D}$, i.e., on its point representation $\mathcal{D}' = \mathbb{Z}^2$.*

Definition 12 *Let $A \subseteq \mathcal{D}$ (or $A' \subseteq \mathbb{Z}^2$), be a digital fuzzy set. A digital quasiconvex fuzzy hull $\mathcal{QH}(A)$ of A is the smallest digital fuzzy set, given by the membership function μ, such that*

$$\text{supp}(\mathcal{QH}(A)) = H_{DL}(\text{supp}(A)), \tag{1}$$

$$\mu(\lambda p_1 + (1 - \lambda)p_2) \geq \min\{\mu(p_1), \mu(p_2)\} \tag{2}$$

holds for every $p_1, p_2 \in H'_{DL}(\text{supp}(A))$, and for all $\lambda \in [0, 1]$ such that $\lambda p_1 + (1 - \lambda)p_2 \in \mathbb{Z}^2$.

Definition 13 *Let $A \subseteq \mathcal{D}$ (or $A' \subseteq \mathbb{Z}^2$), be a digital fuzzy set. A digital convex fuzzy hull $\mathcal{H}(A)$ of A is the smallest digital fuzzy set, given by the membership function μ, such that*

$$\text{supp}(\mathcal{H}(A)) = H_{DL}(\text{supp}(A)), \tag{3}$$

$$\mu(\lambda p_1 + (1 - \lambda)p_2) \geq \lambda\mu(p_1) + (1 - \lambda)\mu(p_2) \tag{4}$$

holds for every $p_1, p_2 \in H'_{DL}(\text{supp}(A))$, and for all $\lambda \in [0, 1]$ such that $\lambda p_1 + (1 - \lambda)p_2 \in \mathbb{Z}^2$.

3 Algorithms

Let A be a digital fuzzy set, with membership function μ_A. Let us construct the DL-hull $H_{DL}(\text{supp}(A))$, using the algorithm given in [2]. Let

$$H'_{DL}(\text{supp}(A)) = \mathcal{A} = \{(x_1, y_1), (x_2, y_2), ...(x_n, y_n)\},$$

$n \geq 2$, be the digitalization of $H_{DL}(\text{supp}(A))$.

For $A_k(x_k, y_k) \in \mathcal{A}$, let $\tilde{A}_k(x_k, y_k, \mu(A_k))$, where $\mu(A_k) = \mu_A(A_k) = \mu_k$. The sets $\mathcal{A}$ and $\tilde{\mathcal{A}} = \{\tilde{A}_1(x_1, y_1, \mu(A_1), ..., \tilde{A}_n(x_n, y_n, \mu(A_n)\}$, are the input sets to both our algorithms.

3.1 Algorithm for the Construction of a Digital Quasiconvex Fuzzy Hull $\mathcal{QH}(A)$

Let us present the algorithm for computing a digital quasiconvex fuzzy hull of a digital fuzzy set A. As explained above, the input to the algorithm is the set $\mathcal{A}$.

The idea of the algorithm is to process the input points iteratively, in the non-increasing order of their membership values, and to construct the set $\mathcal{A}_{\mathcal{FH}}$, by forcing α-cuts of $\mathcal{A}_{\mathcal{FH}}$ to be convex sets. Also, we shall prove that by using this algorithm, the digital quasiconvex fuzzy hull $\mathcal{QH}(A)$ of A can be obtained.

Algorithm 1 (for computing digital quasiconvex fuzzy hull):

step 0: Initialize
$\mathcal{E} = \emptyset$ ($\mathcal{E}$ is the set of extremal points),
$\mathcal{A}_H = \emptyset$ ($\mathcal{A}_H$ is the set of internal - nonextremal points).

step 1: Order set $\mathcal{A}$ by non increasing order of membership function μ and rename the points of $\mathcal{A}$ as $A_1, A_2, ..., A_n$, so that $\mu(A_1) \geq \mu(A_2) \geq ... \geq \mu(A_n)$.

step 2: Let $p = p(A_1, A_2)$ be the line determined with points A_1 and A_2. If $\mu(A_1) = \mu(A_2)$, go to step $3'$. Otherwise, (if $\mu(A_1) > \mu(A_2)$), go to step $3''$.

step $3'$: Find all points of set $p \cap \mathcal{A}$ whose membership function is equal to $\mu(A_1) = \mu(A_2)$. Find end points E_1' and E_1'' of that set, remove them from $\mathcal{A}$, and store them in $\mathcal{E}$. Remove from $\mathcal{A}$ all points $A_p \in \mathcal{A}$, which belong to line segment $E_1'E_1''$, such that $\mu(A_p) < \mu(E_1')$, and store them in $\mathcal{A}_H$, with changed membership function $\mu(A_p) = \mu(E_1') = \mu(E_1'')$ (see Figure 1). Go to step 5.

step $3''$: Remove the point A_1 from $\mathcal{A}$ and store it in $\mathcal{E}$, with unchanged membership function. Set $E_1'' \equiv E_1' \equiv A_1$ and $\mu(E_1'') = \mu(E_1') = \mu(A_1)$. Rename points of set $\mathcal{A}$ so that A_1 is the point with the largest membership function, next point is A_2, etc.

step 4: Find all points of set $p \cap \mathcal{A}$ which have equal membership function as point A_1, together with points E_1' and E_1'' (E_{i-1}' and E_{i-1}'') (together with only one of those two points, if they are equal).

Denote with E_2' and E_2'' (E_i' and E_i'') such points from all those points (if they exist), that any other point A_k from $p \cap \mathcal{A}$ either coincides with one of them, or is between them. We also take that the points of p are in the order $E_2' \preceq E_1' \preceq E_1'' \preceq E_2''$ ($E_i' \preceq E_{i-1}' \preceq E_{i-1}'' \preceq E_i''$), where the relation $\preceq$ is one of relations: $\equiv$ (congruence of points) or $-$ (be between two points).

Store the points E_2' and E_2'' (E_i' and E_i'') in set $\mathcal{E}$, if they are not already in $\mathcal{E}$. If one of them is already an extreme point, assign to it the membership function of the other point, i.e. $\mu(E_i') = \mu(E_i'')$. For example, if $E_3'' \equiv E_2'' \in \mathcal{E}$, we take that $\mu(E_3'') = \mu(E_3')$ (see Figure 1).

Remove other points $A_k \in p \cap \mathcal{A}$ from $\mathcal{A}$, and store them in $\mathcal{A}_H$.

Remove from $\mathcal{A}$ all points $A_q \in \mathcal{A}$, which belong to the line segment $E_2'E_2''$ ($E_i'E_i''$), such that $\mu(A_q) < \mu(E_2')$ ($\mu(A_q) < \mu(E_i')$), and store them in $\mathcal{A}_H$, with changed membership function $\mu(A_q) = \mu(E_2') = \mu(E_2'')$ ($\mu(A_q) = \mu(E_i') = \mu(E_i'')$).

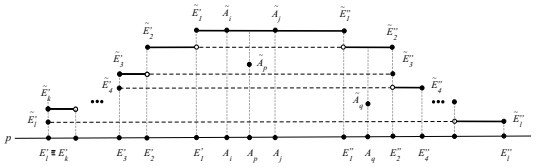


Figure 1: First steps of Algorithm 1. Points of the current α-cut are all on a line.

step 5: Rename points of set $\mathcal{A}$ so that A_1 is the point with the largest membership function, next point is A_2, etc. If $A_1 \in p$, go to step 4, where input is again A_1, and output is E_i' and E_i''. This iterative cycle is repeated as long as the previous condition is satisfied. Otherwise, denote with K_1 and K_2 the two end points from $p \cap \mathcal{E}$ (E_l' i E_l'' in Figure 1). Introduce the set $\mathcal{K}_0 = \{K_1, K_2\}$.

step 6: Join the point A_1 with points K_1 and K_2 from set $\mathcal{K}_0 \subset \mathcal{E}$. Remove from $\mathcal{A}$ all points A_k that are in the obtained triangle (together with point A_1), and store them in $\mathcal{A}_H$, with changed membership function $\mu(A_k) = \mu(A_1)$. Store in $\mathcal{E}$ the point A_1 with its membership function. This situation is illustrated in Figure 2.

Set $K_3 \equiv A_1$ and $\mu(K_3) = \mu(A_1)$ and store point K_3 in $\mathcal{K}_0$. In that way we get the set $\mathcal{K}_1 = \mathcal{K}_0 \cup \{K_3\}$. Mark the derived triangle with $\mathcal{M}(\mathcal{K}_1)$.

step 7: In the next iteration (and those after if they are needed) we take the first point (from the remain-

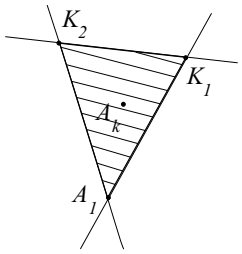


Figure 2: A_1 does not belong to the line p.

ing points, if they exist) from set $\mathcal{A}$. Because of the renaming of points of the set $\mathcal{A}$, that point is A_1.

If $\mathcal{A} = \emptyset$, then go to step 8.

Store in $\mathcal{E}$ the point A_1 with unchanged membership function, and remove A_1 from $\mathcal{A}$. If on the derived line segments or in the interior of the derived polygon in current iteration there are left some of the extreme points with the same membership function $\mu(A_1)$, remove them from set $\mathcal{E}$ and store them in set $\mathcal{A}_H$ with unchanged membership function.

From points of set $\mathcal{K}_2 = \{K_1, K_2, K_3\}$ ($\mathcal{K}_{r-1} = \{K_1, K_2, K_3, ..., K_{i-1}, K_i, K_{i+1}, ..., K_{j-1}, K_j, K_{j+1}, ..., K_m\}$) and point A_1 we make another set $\mathcal{K}_4$ ($\mathcal{K}_r$), which contains vertices of the convex hull of set $\mathcal{K}_3 \cup \{A_1\}$ ($\mathcal{K}_{r-1} \cup \{A_1\}$). In Figure 3, $\mathcal{K}_4 = \{K_1, K_2, K_3, A_1\}$, in Figures 4 and 5, $\mathcal{K}_4 = \{K_1, K_2, A_1\}$. In derived set $\mathcal{K}_4$ ($\mathcal{K}_r$) we also rename indexes and denote point A_1 with K_4 in the case like in Figure 3, or with K_3 in the cases as in Figures 4 and 5 (with K_i for some index i in the general case).

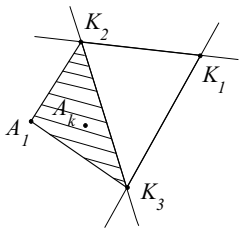


Figure 3: Triangle $\mathcal{M}(\mathcal{K}_1)$ is extended with point A_1 into quadrangle $K_1K_2A_1K_3$.

Denote with $\mathcal{M}(\mathcal{K}_4)$ ($\mathcal{M}(\mathcal{K}_r)$) polygon determined with points from $\mathcal{K}_4$ ($\mathcal{K}_r$). Remove from $\mathcal{A}$ all points A_k from $\mathcal{A} \cap (\mathcal{M}(\mathcal{K}_4) \setminus \mathcal{M}(\mathcal{K}_3))$ ($\mathcal{A} \cap (\mathcal{M}(\mathcal{K}_r) \setminus \mathcal{M}(\mathcal{K}_{r-1}))$), and store them in $\mathcal{A}_H$, with improved membership function, from $\mu(A_k)$ to $\mu(A_1)$, if that value is less than $\mu(A_1)$ (it is possible that $\mu(A_k) < \mu(A_1)$ or $\mu(A_k) = \mu(A_1)$).

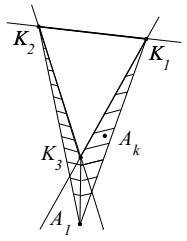


Figure 4: Triangle $\mathcal{M}(\mathcal{K}_1)$ is extended with point A_1 into triangle $K_1K_2A_1$.

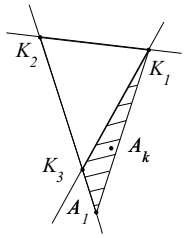


Figure 5: Triangle $\mathcal{M}(\mathcal{K}_1)$ is extended with point A_1 into triangle $K_1K_2A_1$.

If point A_1 is colinear with some two neighboring points from $\mathcal{K}_3$ ($\mathcal{K}_{r-1}$), like in Figure 5 and the middle point of that three points (in Figure 5, it is K_3) has membership function equal to $\mu(A_1)$, remove it from $\mathcal{E}$ and store it in $\mathcal{A}_H$.

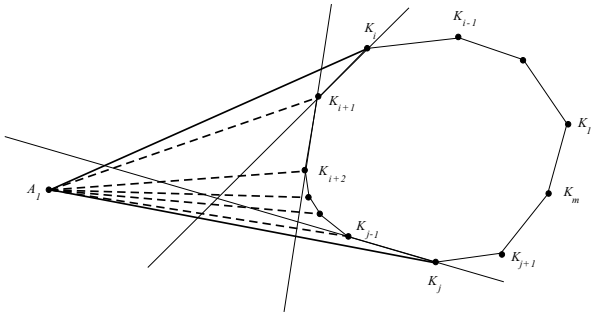


Figure 6: A_1 does not belong to the convex polygon $\mathcal{M}(\mathcal{K}_r)$. A new convex polygon is constructed.

step 8: $\mathcal{A}_{FH} = \mathcal{E} \bigcup \mathcal{A}_H$. The end.

Let us denote by $\delta^{-1}(\mathcal{A}_{FH})$ the digital fuzzy set which corresponds to a fuzzy set $\mathcal{A}_{FH}$, i.e., the digital fuzzy set whose support consists of a set of pixels $P = \delta^{-1}(p)$, where $p \in \text{supp}(\mathcal{A}_{FH})$, with the same membership function ($\mu(P) = \mu(p)$).

Theorem 3 *The digital fuzzy set $\delta^{-1}(\mathcal{A}_{FH})$, derived by the previous algorithm, is the digital quasiconvex*

fuzzy hull $\mathcal{QH}(A)$ of set A.

Proof. Any α-cut of the set $\mathcal{A}_{\mathcal{FH}}$ is by construction either a point or a line segment (this is the case when this line segment belongs to the line $p = p(A_1, A_2)$ from the algorithm), or a polygon $\mathcal{M}(\mathcal{K}_r)$ with its interior. In both cases, this α-cut is a convex set, and so by Theorem 1 the set $\mathcal{A}_{\mathcal{FH}}$ is a convex set. By construction, $\mathcal{A}_{\mathcal{FH}}$ is the smallest fuzzy set such that (2) holds. supp($\mathcal{A}_{\mathcal{FH}}$) is equal to the support of $\mathcal{A}$, which is by construction (in the preprocessing step) the digitalization of the DL-hull of the support supp(A) of the fuzzy set A. □

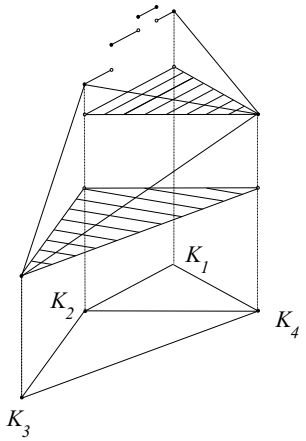


Figure 7: Illustration of the Algorithm 1.

3.2 Algorithm for the Construction of a Digital Fuzzy Convex Hull $\mathcal{H}(A)$

Let us present the algorithm for computing a digital convex fuzzy hull $\mathcal{H}(A)$ of a digital fuzzy set A. As already explained, the input to the algorithm is the set $\mathcal{A}$.

Let us denote by $\mathcal{A}^0$ the set $\mathcal{A}$ embedded in $\mathbb{R}^3$, i.e., $\mathcal{A}^0 = \{A_1(x_1, y_1, 0), ..., A_n(x_n, y_n, 0)\}$, and let $\tilde{\mathcal{A}} = \{\tilde{A}_1(x_1, y_1, \mu_1), ..., \tilde{A}_n(x_n, y_n, \mu_n)\}$, with $\mu_k = \mu(A_k)$, $k = 1, ..., n$.

The idea of the algorithm is to adapt the gift-wrapping method for the computation of the (discrete) convex hull $H(S)$ of a set S of points in $\mathbb{R}^3$. When we apply the adapted algorithm on the set $S = \tilde{\mathcal{A}} \cup \mathcal{A}^0$, we obtain its convex hull $H(S)$, and we simultaneously construct the fuzzy set $\mathcal{A}_{\mathcal{FH}}$. The adaptation is based on the fact that we do not need the bottom and the side facets of the convex hull $H(S)$ of S, and only the upper part of $H(S)$ needs to be constructed. Because of the special geometric position of the points in S, at each step of the gift-wrapping algorithm we need to take into account only the subset of input points, and also we may eliminate some points of S from any further consideration at each step of the algorithm.

Algorithm 2 (for computing digital convex fuzzy hull):

step 1: Initialize
$\mathcal{E} = \emptyset$ ($\mathcal{E}$ is the set of extremal points),
$\mathcal{L} = \emptyset$ ($\mathcal{L}$ is the list of unprocessed facets, and the unprocessed edges of the facets),
$\mathcal{A}_H = \emptyset$ ($\mathcal{A}_H$ is the set of internal - nonextremal points).

step 2: Let $\tilde{A}_1 \in \tilde{\mathcal{A}}$ be the point with maximum z coordinate, and let $\tilde{A}_2 \in \tilde{\mathcal{A}}$ be such that the angle between a vector parallel to the line $p(\tilde{A}_1, \tilde{A}_2)$, and vector $\vec{k} = (0, 0, 1)$, is minimal. Next, let $\sigma_0 = \sigma_0(\tilde{A}_1, \tilde{A}_2, \tilde{A}_3)$ be the plane determined by non-colinear points $\tilde{A}_1$, $\tilde{A}_2$, and $\tilde{A}_3$ from $\tilde{\mathcal{A}}$, such that the angle between the normal vector to σ_0, and the vector $\vec{k}$ is minimal. (We rename the points, if necessary.) Note that all points of $\tilde{\mathcal{A}}$ are either in the plane σ_0, or below it. Let us find all points $\tilde{A}_k$ from $\tilde{\mathcal{A}}$, which belong to σ_0. Let us denote by $\mathcal{S}_0$ the convex polygon, which is the convex hull of the set of such points, and by $\tilde{E}_1, \tilde{E}_2, ..., \tilde{E}_r$, the set of vertices of $\mathcal{S}_0$. (This can be done by finding the projection of such points on the Oxy plane, and using some two dimensional convex hull algorithm.)

Let us remove the points $E_1, E_2, ..., E_r$, which correspond to vertices $\tilde{E}_1, \tilde{E}_2, ..., \tilde{E}_r$ of $\mathcal{S}_0$, from $\mathcal{A}$ into the set $\mathcal{E}$ of extremal points. Each time we remove some points A_k from $\mathcal{A}$, we remove simultaneously the corresponding points $\tilde{A}_k$ from the set $\tilde{\mathcal{A}}$. A convex polygon $\mathcal{S}_0$ is a facet of the convex hull $H(S)$. We put $\mathcal{S}_0$ in the list $\mathcal{L}$ of facets of $H(S)$, together with the list $(\tilde{E}_1\tilde{E}_2, \tilde{E}_2\tilde{E}_3, \cdots, \tilde{E}_{r-1}\tilde{E}_r, \tilde{E}_r\tilde{E}_1)$ of edges of $\mathcal{S}_0$, ordered counter-clockwise. We remove from $\mathcal{A}$ the points A_k, which correspond to the remaining (non-extremal) points $\tilde{A}_k \in \sigma_0$, and put them in the set $\mathcal{A}_H$.We find points $A_l \in \text{pr}(\mathcal{S}_0)$, for which $\tilde{A}_l(x_l, y_l, \mu_l) \notin \sigma_0$. We remove the corresponding points A_l from $\mathcal{A}$ and we put them in $\mathcal{A}_H$ with new membership value $\mu(A_l)$ so that $(x_l, y_l, \mu(A_l)) \in \sigma_0$.

step 3: Let $\tilde{E}_1\tilde{E}_2$ be one edge of the polygon $\mathcal{S}_0$. Let us find a vertical plane through edge $\tilde{E}_1\tilde{E}_2$, and let $\pi_{\tilde{E}_1\tilde{E}_2}$ be the halfspace determined by that plane, which does not contain $\mathcal{S}_0$. Let $\tilde{A}_k$, be the point from the set $\tilde{\mathcal{A}}$, such that $\tilde{A}_k$ belongs to $\pi_{\tilde{E}_1\tilde{E}_2}$, and that the angle between the normal vector to the plane $\sigma_1 = r(\tilde{E}_1, \tilde{E}_2, \tilde{A}_k)$ and vector $\vec{k} = (0, 0, 1)$ is minimal. Let us find the convex hull $\mathcal{S}_1$ of the set of points of $\mathcal{B} = \tilde{\mathcal{A}} \cap \pi_{\tilde{E}_1\tilde{E}_2} \cap \sigma_1$. We remove the projections

of the vertexes of $\mathcal{S}_1$ from $\mathcal{A}$, and we put them in $\mathcal{E}$. We remove from $\mathcal{A}$ the points A_k, which correspond to the remaining (non-extremal) points $\tilde{A}_k$ of $\mathcal{S}_1$, and put them in the set $\mathcal{A}_H$.We find points $A_l \in \mathrm{pr}(\mathcal{S}_1)$, for which $\tilde{A}_l(x_l, y_l, \mu_l) \notin \sigma_1$. We remove the corresponding points A_l from $\mathcal{A}$ and we put them in $\mathcal{A}_H$ with new membership value $\mu(A_l)$ so that $(x_l, y_l, \mu(A_l)) \in \sigma_1$.

We put the convex polygon $\mathcal{S}_1$ in the list $\mathcal{L}$, together with the list of its unprocessed edges ordered counterclockwise. We remove the edge $\tilde{E}_1\tilde{E}_2$ from $\mathcal{L}$.

We repeat step 3 with each unprocessed edge of $\mathcal{S}_0$.

We iteratively process all facets which we retrieve from the list $\mathcal{L}$, by iteratively processing the unprocessed edges of such facets.

The iteration process stops when $\mathcal{A} = \emptyset$, or $\mathcal{L} = \emptyset$.

$$\mathcal{A}_{FH} = \mathcal{E} \bigcup \mathcal{A}_H.$$

The end.

Let us denote by $\delta^{-1}(\mathcal{A}_{FH})$ the digital fuzzy set which corresponds to a fuzzy set $\mathcal{A}_{FH}$, i.e., the digital fuzzy set whose support consists of a set of pixels $P = \delta^{-1}(p)$, where $p \in \mathrm{supp}(\mathcal{A}_{FH})$, with the same membership function ($\mu(P) = \mu(p)$).

Theorem 4 *The digital fuzzy set $\delta^{-1}(\mathcal{A}_{FH})$, derived by the previous algorithm, is the digital convex fuzzy hull $\mathcal{H}(A)$ of set A.*

Proof. By the preprocessing step of the Algorithm,

$$H'_{DL}(\mathrm{supp}(A)) = \mathcal{A} = \{(x_1, y_1), (x_2, y_2), ...(x_n, y_n)\},$$

so the equation 3 of Definition 13 is satisfied. The convex hull $H(\mathcal{A}^0 \cup \tilde{\mathcal{A}}) = H(\mathcal{A}^0 \cup \tilde{\mathcal{A}}_{FH})$ (where $\tilde{\mathcal{A}}_{FH} = \{(x_k, y_k, \mu_k) | (x_k, y_k) \in \mathcal{A}_{\mathcal{FH}}\}$) obtained by our modification of the gift-wrapping method is convex, so the fuzzy hypograph of $\mathcal{A}_{\mathcal{FH}}$ is convex. By Theorem 2, fuzzy set $\mathcal{A}_{\mathcal{FH}}$ (with membership function μ determined by the algorithm) is also convex, i.e., condition 4 of Definition 13 is satisfied. By construction, $\delta^{-1}(\mathcal{A}_{FH})$ is the smallest fuzzy set satisfying the conditions of Definition 13.

Thus, $\mathcal{QH}(A) = \delta^{-1}(\mathcal{A}_{FH})$. □

4 Conclusion and Future Work

We introduced the notions of a digital quasiconvex fuzzy hull $\mathcal{QH}(A)$, and of a digital convex fuzzy hull $\mathcal{H}(A)$, of a given digital fuzzy set A. $\mathcal{QH}(A)$ and $\mathcal{H}(A)$ are both defined on a set of pixels, and they both have the same support $H_{DL}(A)$. $\mathcal{QH}(A)$ is a quasiconvex set, while $\mathcal{H}(A)$ is a convex set. We proposed an algorithm for the construction of both digital fuzzy hulls.

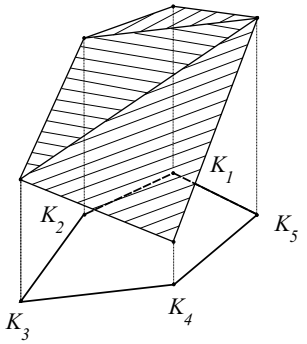


Figure 8: Illustration of the gift-wrapping algorithm.

We intend to extend this work by considering other notions of convexity of digital fuzzy sets, such as α-convexity.

References

[1] N. E. Ammar, Some Properties of Convex Fuzzy Sets and Convex Fuzzy Cones, Fuzzy Sets and Systems 106, 381-386, 1999.

[2] B. B. Chaudhuri, A. Rosenfeld, On the Computation of the Digital Convex Hull and Circular Hull of a Digital Region, Pattern Recognition, Vol. 31, No. 12, pp. 2007-2016, 1998.

[3] L. Čomić, N. M. Ralević, On Convex Hulls of Fuzzy Sets, 6^{th} International Symposium Interdisciplinary Regional Research, Hungary-Romania-Yugoslavia, Novi Sad 3-4. oktobar 2002.

[4] N. M. Ralević, V. Ćurić, M. Janev, Algorithm for computing the digital convex fuzzy hull, 4^{th} Serbian-Hungarian Joint Symposium on Intelligent Systems, September 29-30, 2006, Subotica, Serbia and Montenegro (2006), 301-309.

[5] J. G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall PTR Upper Saddle River, New Jersey, 1995.

[6] F. P. Preparata, M. I. Shamos, Computational Geometry: An Introduction, Springer Verlag, New York, 1985.

[7] N. M. Ralević, Digitals objects and convex fuzzy sets, Acta Mechanica Slovaca (accepted).

[8] X. Yang, A Note on Convex Fuzzy Sets, Fuzzy Sets and Systems 53, 117-118, 1993.

ŁUKASIEWICZ FILTERS AS CONVEX COMBINATIONS OF CRISP FILTERS

Zuzana Havranová
Dept. of Mathematics
Slovak University of Technology
Radlinského 11
813 68 Bratislava, Slovakia
zuzana@math.sk

Martin Kalina
Dept. of Mathematics
Slovak University of Technology
Radlinského 11
813 68 Bratislava, Slovakia
kalina@math.sk

Abstract

In this paper we investigate special types of Łukasiewicz filters. We show their connection to crisp filters. Further, we introduce the notion of Łukasiewicz ideal as a set-function which is in some sense complementary to Łukasiewicz filters.

Keywords: Łukasiewicz filter, Łukasiewicz ultrafilter, Łukasiewicz ideal.

In series of papers [1, 2, 4, 5] so-called generalized filters were defined and studied. Generalized filters on a set X, $\mathcal{G} : 2^X \to [0,1]$, are in fact fuzzy filters where the minimum t-norm T_M is used to get the condition

$$\mathcal{G}(A \cap B) \geq T_M(\mathcal{G}(A), \mathcal{G}(B)).$$

The authors have introduced Łukasiewicz filters in [8] (and further developed them in [6]) in changing the minimum t-norm T_M into the Łukasiewicz t-norm T_L. (For details on t-norms and/or t-conorms see [9].) Łukasiewicz filters may be used, e.g., to construct fuzzy preference relations, as it is shown in [7]. Particularly, if we have to compare a system of objects taking into account a set of criteria, X, the choice of some Łukasiewicz filter $\mathcal{F}$ on the set X says how important different criteria are and says about the ointeraction of criteria. Concerning the Łukasiewicz t-norm, from the point of view of fuzzy preference structures this is the most important case (see, e.g. [3]).

This paper is devoted to studying of a relationship between crisp filters and Łukasiewicz filters. In the last paragraph Łukasiewicz ideals are defined and a maximal Łukasiewicz ideal is constructed.

1 Notation and basic definitions

In the whole paper we will use the following notation:

- $X = \{x_1, x_2, \ldots\}$ is a countable set
- $W = \{w_1, w_2, \ldots\}$ is a weighting sequence. In general, we will assume just $0 \leq w_i \leq 1$ for each i. In partcular cases discussed in the paper some additional requirements (formula (3) below) will be put on the weghting sequence.
- S is some left-continuous t-conorm
- $\sum_j {}_S\, w_{i_j} = S\,(w_{i_1}, w_{i_2}, \ldots)$. Since S is left-continuous, this expression is well-defined.

Let us recall the basic definitions:

Definition 1 *Let $X \neq \emptyset$ be a given set. Then a set-function $F : 2^X \to \{0,1\}$ is said to be a filter on X if and only if the following hold:*

1. $F(X) = 1, F(\emptyset) = 0$
2. $F(A) \wedge F(B) \leq F(A) \cap (B)$
3. *If $A \subseteq B$ then $F(A) \leq F(B)$*

Definition 2 *Let $X \neq \emptyset$ be a given set. Then a set-function $U : 2^X \to \{0,1\}$ is said to be an ultrafilter on X if and only if the following hold:*

1. *U is a filter*
2. $\forall A \subset X \quad U(A) \vee U(A^c) = 1$

Lemma 1 *Let us denote $\mathcal{X}_F$ the system of all filters on $X \neq \emptyset$. Then $\mathcal{X}_F$ is a lattice with its bottom element*

$$F_0 = \begin{cases} 1 & \text{if } A = X \\ 0 & \text{if } A \neq X. \end{cases}$$

Ultrafilters $U : 2^X \to \{0,1\}$ are maximal elements of $\mathcal{X}_F$.

Example 1 Let us take a countable set X. Then we have the following ultrafilter $U : 2^X \to \{0,1\}$ given by:

1. $U(A) = 1$ implies A is infinite
2. $\forall A \subset X$ we have $U(A) \vee \mathcal{U}(A^c) = 1$.

Such ultrafilter is called nontrivial.

2 Łukasiewicz filters

Definition 3 *Let $X \neq \emptyset$ be a given set. Then a set-function $\mathcal{F} : 2^X \to [0,1]$ is said to be Łukasiewicz filter on X if and only if the following are satisfied:*

1. $\mathcal{F}(X) = 1, \mathcal{F}(\emptyset) = 0$
2. $(\forall A, B \subseteq X)\,(A \subseteq B \Rightarrow \mathcal{F}(A) \leq \mathcal{F}(B))$
3. $(\forall A, B \subseteq X)\,(\mathcal{F}(A \cap B) \geq T_L\,(\mathcal{F}(A), \mathcal{F}(B)))$

Definition 4 *Let S be some left-continuous t-conorm. A measure $\nu : 2^X \to [0,1]$ is said to be S-additive, if and only if for every couple $A, B \in 2^X$ of disjoint sets the following holds:*

$$\nu(A \cup B) = S\,(\nu(A), \nu(B)) \tag{1}$$

If formula (1) holds just for couples of disjoint sets $A, B \in 2^X$ such that $A \cup B \neq X$, then the measure ν will be called weakly S-additive.

Definition 5 *Let S be some left-continuous t-conorm. Denote $A = \{x_{ij}\}_j \subseteq X$. (Remind that X is an at most countable set.) Then $\mathcal{F}_S : 2^X \to [0,1]$, defined by*

$$\mathcal{F}(\mathcal{A}) = \begin{cases} 0 & if A = \emptyset \\ 1 & if A = X \\ \sum\limits_j {}_S\, w_{ij} & otherwise, \end{cases} \tag{2}$$

is said to be a (T_L, S)-filter on X if it is a Łukasiewicz filter.

Formula (2) together with properties 1 and 3 from Definition 3 imply

$$\sum_j {}_S\, w_{ij} + \sum_k {}_S\, w_{ik} \leq 1 \tag{3}$$

where $\{x_{ij}\}_j$ and $\{x_{ik}\}_k$ are some disjoint subsequences of $\{x_i\}_i$.

Definition 6 *Let $X \neq \emptyset$ be a given set, $\mathcal{F} : 2^X \to [0,1]$. Then $\mathcal{F}$ is said to be an S-additive T_L filter if is a weakly S-additive measure which is a Łukasiewicz filter.*

Example 2 Let $\mathcal{F}_1 : 2^X \to [0,1]$ be a probability measure (i.e. σ-additive), then $\mathcal{F}_1$ is a (T_L, S_L)-filter. And, of course, it is also an S_L-additive T_L-filter.

On the other hand, let us take $F_1 : 2^X \to \{0,1\}$ and $F_2 : 2^X \to \{0,1\}$ two different nontrivial ultrafilters (i.e. ultrafilters with $F_i(A) = 0$ for A finite). Then, $\mathcal{F}_2 : 2^X \to [0,1]$ defined by

$$\mathcal{F}_2(A) = 0.5F_1(A) + 0.5F_2(A)$$

is an S_L-additive T_L filter, but it is not a (T_L, S_L)-filter.

Theorem 1 *Let $S \geq S_L$ be a generated t-conorm with its additive generator $f : [0,1] \to [0,1]$. Then a set-function $\mathcal{F} : 2^X \to [0,1]$, defined by*

$$\mathcal{F}(A) = \begin{cases} 1 & if A = X \\ \sum\limits_j {}_S\, w_{i_j} & if A = \{x_{i_1}, x_{i_2}, ...\} \end{cases}$$

is a (T_L, S) filter, if $\sum f(w_i) \leq 1$.

Theorem 2 *Let $S \leq S_L$. Then a sufficient condition for $\mathcal{F} : 2^X \to [0,1]$, defined by*

$$\mathcal{F}(A) = \begin{cases} 1 & if A = X \\ \sum\limits_j {}_S\, w_{i_j} & if A = \{x_{i_1}, x_{i_2}, ...\} \end{cases}$$

to be a (T_L, S)-filter is that there exit disjoint sets A, B such that $A \cup B = X$ and $\mathcal{F}(A) \leq 0.5,\ \mathcal{F}(B) \leq 0.5$.

Remark 1 The condition $\mathcal{F}(A) \leq 0.5,\ \mathcal{F}(B) \leq 0.5$ is just sufficient. We may have, e.g. three-elements set $\{a, b, c\}$ with weights $\{0.6, 0.25, 0.2\}$ respectively. Then these weights generate a (T_L, S_P) filter.

3 Convex combinations of crisp filters

Definition 7 *$z = S(ax, by)$ is said to be an S-convex combination of x, y iff $a, b \in [0,1]$, and $b = 1 -_S a$.*

Definition 8 *$z = \sum\limits_{i=1}^{\infty} {}_S\, a_i y_i$ is said to be a σ-S-convex combination of y_i iff $a_i \in [0,1]$ and for each j the following is satisfied:*

$$a_j = 1 -_S \sum_{i \neq j} {}_S\, a_i.$$

Example 3 Let us have a sequence of weights with $\sum w_i = 1$. Further let $\mathcal{F}$ be a (T_L, S_L)-filter defined by

$$\mathcal{F}(A) = \sum_j w_{i_j} \quad \text{for } A = \left\{x_{i_j}\right\}_{j \in J}.$$

Let U_i be ultrafilters given by

$$U_i(B) = \begin{cases} 1 & \text{if } x_i \in B \\ 0 & \text{if } x_i \notin B. \end{cases}$$

Then $\mathcal{F}$ is the σ-convex combination of $\{U_i\}_i$:

$$\mathcal{F}(A) = \sum w_i U_i(A).$$

Example 4 Let us have a Yager t-conorm S, defined by

$$S = min\left\{1, \left(\sqrt{x} + \sqrt{y}\right)^2\right\}$$

and a system of weights $\{w_i\}_{=0}^{\infty}$ such that $\sum\limits_{i=1_S}^{\infty} w_i = \frac{1}{4}$ and $w_0 = \frac{1}{4} = 1 -_S \frac{1}{4}$. Let us denote

$$\begin{aligned} \mathcal{F}(A) &= \begin{cases} 1 & \text{if } A = X \\ \sum\limits_j {}_S\, w_{i_j} & \text{if } A \neq X \end{cases} \\ U_i(B) &= \begin{cases} 1 & \text{if } x_i \in B \\ 0 & \text{if } x_i \notin B \end{cases} \\ F_0 &= \begin{cases} 1 & \text{if } A = X \\ 0 & \text{if } A \neq X \end{cases} \end{aligned}$$

Then $\mathcal{F}$ is the σ-S-convex combination of $\{U_i\}_i$ and F_0:

$$\mathcal{F}(A) = S\left(\frac{1}{4}F_0(A), \sum_{i=1}^{\infty}{}_S\, w_i U_i(A)\right).$$

Theorem 3 *Each (T_L, S)-filter $\mathcal{F}$, given by*

$$\mathcal{F}(A) = \begin{cases} 1 & \text{if } A = X \\ \sum\limits_j {}_S\, w_{i_j} & \text{if} A = \{x_{i_1}, x_{i_2}, ...\} \end{cases}$$

is a σ-S-convex combination of the least filter F_0 and a sequence of trivial ultrafilters $\{U_i\}$

$$U_i(A) = \begin{cases} 1 & \text{if } x_i \in A \\ 0 & \text{if } x_i \notin A. \end{cases}$$

In this case $\mathcal{F}$ is a weakly σ-S-additive measure.

Theorem 4 *Each S-additive T_L-filter $\mathcal{F}$ is a σ-S-convex combination of the least filter F_0 and a sequence of ultrafilters (not necessarily trivial) $\{U_i\}$. In this case $\mathcal{F}$ is a weakly S-additive measure.*

Theorem 5 *Let $\mathcal{F}$ be an arbitrary Łukasiewicz filter. Then there exist a system of filters $\{F_i\}_{i\in I}$ (I might be uncountable) and a maxitive system of weights $\{a_i\}_{i\in I}$ such that $\mathcal{F}(A) = sup_i(a_i F_i(A))$.*

In general, Łukasiewicz filters are not σ-S-convex combinations of crisp filters. Łukasiewicz filters, which are not σ-S-convex combinations of crisp filters, can be constructed using transfinite construction methods (Axiom of Choice).

Classes of Łukasiewicz filters are sketched in the following Figure 1. We have four basic types of them:

- $\mathcal{F}_1$: (T_L, S)-filters, i.e. σ-S-convex combinations of trivial ultrafilters and F_0
- $\mathcal{F}_2$: S-additive T_L-filter i.e. σ-S-convex combinations of ultrafilters and F_0
- $\mathcal{F}_3$: Łukasiewicz filters which are σ-S-convex combinations of crisp filters
- $\mathcal{F}_4$: All Łukasiewicz filters.

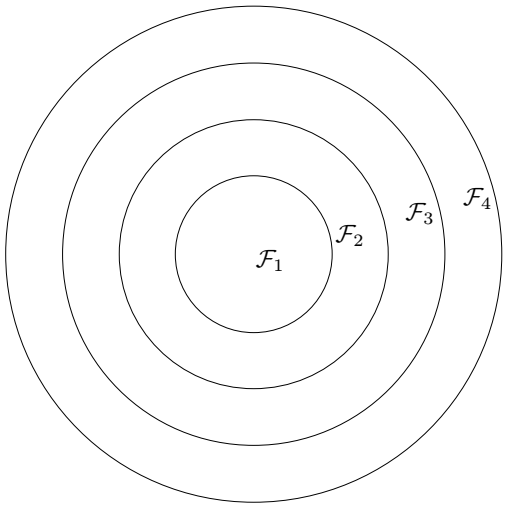


Figure 1: Types of Łukasiewicz filters

4 Łukasiewicz ideals

Definition 9 (Definition of proper ideal) *Let Z be a set ordered by $\leq$. $Y \subset Z$ is a proper ideal on $(Z, \leq)$, if:*

1. *$\forall a \in Y$ if $b \leq a$, then $b \in Y$*
2. *if $a, b \in Y$ then if $c \leq a \wedge b$, then $c \in Y$*
3. *$Y \neq Z$.*

Definition 10 *Let $Z = 2^X$. Then $Y \subset Z$ is a proper ideal on $(X, \leq)$ if:*

1. *$\forall a \in Y$, $b \subseteq a \Rightarrow b \in Y$*
2. *$a, b \in Y$, $c \subseteq a \wedge b$ implies $c \in Y$*
3. *$X \neq Y$.*

Example 5 $X = \{a, b, c, d\}$, $T = \{a, b, c\}$, $Y = 2^T$ is a proper ideal on $(X, \leq)$.

Definition 11 *Let $X \neq \emptyset$ be a given set. Then a set-function $\mathcal{I} : 2^X \to [0, 1]$ is said to be a Łukasiewicz ideal on X, if:*

1. *$\mathcal{I}(X) = 0, \mathcal{I}(\emptyset) = 1$*
2. *$\mathcal{I}(A \cup B) \geq T_L\left(\mathcal{I}(A), \mathcal{I}(B)\right)$*

3. If $A \subseteq B$ then $\mathcal{I}(A) \geq \mathcal{I}(B)$.

Definition 12 *Let $X \neq \emptyset$ be a given set. Then a set-function $\mathcal{I} : 2^X \to [0,1]$ is said to a Łukasiewicz ideal on X, if:*

1. $\exists x_0 \in X : \mathcal{I}\{x_0\} = 1$
2. $\forall A \not\subseteq \{x_0\}$ *there holds* $\mathcal{I}(A) < 1$.

Definition 13 *Let $X \neq \emptyset$ be a given set. Then a Łukasiewicz filter $\mathcal{U} : 2^X \to [0,1]$ is said to be a Łukasiewicz ultrafilter, if $\mathcal{I} = 1 - \mathcal{U}$ is a Łukasiewicz ideal.*

Example 6 Let $X = \{x_1, x_2, x_3, x_4\}$. Let us define $\mathcal{U} : 2^X \to [0,1]$ by

$$\mathcal{U}(\mathcal{A}) = \begin{cases} 0 & \text{if } A = \emptyset \\ 1/3 & \text{if } |A| = 1 \\ 1/2 & \text{if } |A| = 2 \\ 2/3 & \text{if } |A| = 3 \\ 1 & \text{if } A = X. \end{cases}$$

Then $\mathcal{U}$ is a Łukasiewicz ultrafilter.

Theorem 6 *A Łukasiewicz ideal $\mathcal{I} : 2^X \to [0,1]$ is maximal iff there exists a Łukasiewicz ultrafilter $\mathcal{U}$ such that $\mathcal{U}(A) = 1 - \mathcal{I}(A)$ for each $A \subseteq X$. $\mathcal{U}$ in that case a Łukasiewicz ultrafilter.*

5 Construction of a maximal Łukasiewicz ideal

Let $|X| = 2^n, n > 1$. We are going to construct a maximal Łukasiewicz ideal $\mathcal{I} : 2^X \to [0,1]$ such that $\mathcal{I}(A) = f(|A|)$ where $f : \{0, 1, \ldots, n\} \to [0,1]$ is a non-increasing function with $f(0) = 1$ and $f(n) = 0$. Since $1 - \mathcal{I}$ is a Łukasiexicz ultrafilter, we get the following condition for each set $A \subseteq X$:

$$\mathcal{I}(A) + \mathcal{I}(A^c) = 1.$$

From this we get

$$\frac{1}{2} \leq f(1) \leq \frac{n-1}{n}$$

and, of course, $f(n-1) = 1 - f(1)$. Similar consideration gives

$$T_L(f(1), f(1)) \leq f(2) \leq f(1)$$

and $f(n-2) = 1 - f(2)$.

We show yet two steps to make the idea of the construction clear:

A three-elements set we get as a union of one- and two-elements sets, hence

$$T_L(f(1), f(2)) \leq f(3) \leq f(2)$$

and $f(n-3) = 1 - f(3)$.

A four-elements set we get as a union of two two-elements sets, or as a union of one- and three-elements sets, hence

$$\max\{T_L(f(1), f(3)), T_L(f(2), f(2))\} \leq f(4) \leq f(3)$$

and $f(n-4) = 1 - f(4)$, and so on.

This gives a maximal ideal on the set X with $|X| = n$ such that for $A \subseteq X$, $|A| = i$ the following is satisfied

$$\mathcal{I}(A) = f(i).$$

- If we in all cases choose the maximal possible value of $f(\cdot)$, then $\mathcal{F}(A) = 1 - \mathcal{I}(A)$ is a (T_L, S_L)-filter, which in fact is the uniform probability distribution on X.
- If we in all cases choose the minimal possible value of $f(\cdot)$, then $\mathcal{F}(A) = 1 - \mathcal{I}(A)$ is a (T_L, S_M)-filter, i.e. it is a weakly maxitive measure.

6 Conclusion

In this paper we have shown the relationship between Łukasiewicz filters and crisp filters. Further, we have introduced Łukasiewicz ideals and we have constructed a maximal Łukasiewicz ideal.

Acknowledgement

This work was supported by Science and Technology Assistance Agency under the contract No. APVV-0375-06, and by the VEGA grant agency, grant number 1/3014/06.

References

[1] Burton, M.H., Muraleetharan, M., Gutiérrez García, J. (1999), Generalised filters 1, Fuzzy Sets and Systems, vol. 106, pp. 275-284.

[2] Burton, M.H., Muraleetharan, M., Gutiérrez García, J. (1999), Generalised filters 2, *Fuzzy Sets and Systems*, vol. 106, pp. 393-400.

[3] J. Fodor and M. Roubens. Fuzzy preference modeling and Multicriteria Decision Support. Kluwer, Dordrecht, 1994.

[4] Gutiérrez García, J., Mardones Pérez, I., Burton, M.H.: The relationship between various filter notions on a GL-monoid. J. Math. Anal. Appl. vol. 230 (1999) 291–302.

[5] Gutiérrez García, J., De Prada Vicente, M.A., Šostak, A.P. (2003), A unified approach to the concept of fuzzy L-uniform space, Rodabaugh,

S.E. et al. (eds.) Topological and algebraic structures in fuzzy sets. *Trends in Logic, Studia Logica Library*, vol. 20, Kluwer Acad. Publisher, Dordrecht.. pp. 81-114.

[6] Havranová, Z., Kalina, M: Weakly S-additive measures. In: Proc. of IPMU 2006, Paris (2006) pp. 630-634.

[7] Havranová, Z., Kalina, M: Fuzzy preference relations and Łukasiewicz filters. In: Proc. EUSFLAT 2007, submitted.

[8] Kalina, M.: Łukasiewicz filters and similarities. In: Proceedings of AGOP'2005. Lugano, Switzerland (2005) 57–60.

[9] E.P. Klement, R. Mesiar, E. Pap. Triangular norms. Kluwer, Dordrecht, 2000.

FUZZY PREFERENCE RELATIONS AND ŁUKASIEWICZ FILTERS

Zuzana Havranová
Dept. of Mathematics
Slovak University of Technology
Radlinského 11
813 68 Bratislava, Slovakia
zuzana@math.sk

Martin Kalina
Dept. of Mathematics
Slovak University of Technology
Radlinského 11
813 68 Bratislava, Slovakia
kalina@math.sk

Abstract

This paper is devoted to construction of fuzzy preference relations and to showing how this construction is connected with Łukasiewicz filters and T_L-evaluators. We will study also extensions of Łukasiewicz filters on subsets of some $X \neq \emptyset$ to T_L-evaluators on the system (lattice) of fuzzy functions $f : X \to [0, 1]$.

Keywords: Fuzzy preference relation, T_L-evaluator, Łukasiewicz filter.

In this paper we will show a possibility how fuzzy preference relations can be constructed using special kinds of evaluators, so-called T_L evaluators. Briefly, the idea is as follows:

We have a set of products, Ξ, and are to decide which of them suits better to our purposes. We examine a system of parameters (index set), $\mathcal{K}$. Let X_k denote possible values of the k-th parameter. We will assume that this set is linearly ordered. Using the linear order we construct a fuzzy relation $P_k : X_k \times X_k \to [0, 1]$ such that $P_k(\alpha_k, \beta_k)$ is the degree in which the value of the k-th parameter of $\alpha \in \Xi$ is preferred to or indifferent with the value of the k-th parameter of $\beta \in \Xi$.

I.e., comparing our products, we get a function (fuzzy set) defined by

$$\tilde{\alpha}(k) = P_k(\alpha_k, \beta_k),$$

which we have to evaluate by some evaluator φ. Finally, we get the fuzzy preference relation R by

$$R(\alpha,\, \beta) = \varphi(\tilde{\alpha}).$$

1 Construction of fuzzy preference structures

1.1 Fuzzification of linear order

Let $\mathcal{K}$ be an, at most countable, index set. Further, for each $k \in \mathcal{K}$ let $(X_k, \leq_k)$ be the set of values of the k-th parameter, linearly ordered by $\leq_k$. Assume $s \leq_k t$ means that the value $s \in X_k$ is preferred to $t \in X_k$. Then the fuzzy relation $P_k : X_k \times X_k \to [0, 1]$ is a fuzzification of the linear order $\leq_k$, i.e., the following properties are to be satisfied:

1. $\forall s, t \in X_k \quad s \leq_k t \;\Rightarrow\; P_k(s, t) = 1$
2. P_k is T_L-transitive.

Straightforwardly we get the following:
For all $s, t, u, v \in X_k$ if $s \leq_k t \leq_k u \leq_k v$, then

$$P_k(v, s) \leq P_k(u, t)$$

For more details on fuzzy orderings see, e.g., [1, 2]. In the following example we illustrate some possibilities how the fuzzy relation P_k might look like.

Example 1 **(a)** We may use the crisp order, i.e.

$$P_k(s, t) = \begin{cases} 1, & \text{if } s \leq_k t \\ 0 & \text{otherwise} \end{cases}$$

(b) If X_k is equipped with a difference, then we may have

$$P_k(s, t) = \begin{cases} 1, & \text{if } s \leq_k t \\ f(s - t), & \text{otherwise} \end{cases}$$

where $f : [0, \infty[\to [0, 1]$ is a convex decreasing function with $f(0) = 1$, e.g.

$$f(s - t) = \max\{0, 1 - (s - t)\}.$$

(c) If X_k is equipped with a division, then we may have

$$P_k(s, t) = \begin{cases} 1, & \text{if } s \leq_k t \\ f\left(\frac{t}{s}\right) & \text{otherwise} \end{cases}$$

where $f : [0, 1] \to [0, 1]$ is a convex increasing function with $f(1) = 1$, e.g.

$$f\left(\frac{t}{s}\right) = \left(\frac{t}{s}\right)^2.$$

1.2 Evaluation of fuzzy sets representing the comparison of products

We have n elements. Let us denote $\Xi = \{1, 2, \ldots, n\}$. The value of the k-th parameter for the i-th element of Ξ will be denoted by $\theta_i(k)$. For each pair $(i,j) \in \Xi^2$ the function $\tilde{P}_{(i,j)} : \mathcal{K} \to [0,1]$ will denote the comparison of values of the corresponding parameters, namely

$$\tilde{P}_{(i,j)}(k) = P_k\left(\theta_i(k), \theta_j(k)\right)$$

I.e., $\tilde{P}_{(i,j)}$ represents the results of comparing the pair of products (i,j) parameter by parameter. Finally, we have to evaluate somehow the partial comparisons. Let $\mathcal{M}$ be the system of all functions $f : \mathcal{K} \to [0,1]$. Then $(\mathcal{M}, \wedge, \vee, \top, \bot)$ is a complete lattice with top and bottom elements $\top$ and $\bot$, equal to constants 1 and 0, respectively. Let $\varphi : \mathcal{M} \to [0,1]$ denote the function evaluating the comparison. D. Dubois, W. Ostasiewicz, H. Prade in [4] have characterized so-called evaluators defined by:

Definition 1 *A function $\varphi : \mathcal{M} \to [0,1]$ is called a normalized evaluator on $\mathcal{M}$ iff*

1. $\varphi(\top) = 1$, $\varphi(\bot) = 0$,
2. *for all f, $g \in \mathcal{M}$, if $f \leq g$ then $\varphi(f) \leq \varphi(g)$.*

Evaluators are connected with so-called fuzzy preference relations. These have been studied already since years (see, e.g. [5]). The definition is as follows:

Definition 2 *A fuzzy relation $R : \Xi \times \Xi \to [0,1]$ is said to be a large fuzzy preference relation on Ξ, if it is reflexive and T_L-transitive. (For brevity reasons we will skip the adjective large.)*

Theorem 1 *Let $\varphi : \mathcal{M} \to [0,1]$ be a normalized evaluator on $\mathcal{M}$. Then a fuzzy relation $R : \Xi \times \Xi \to [0,1]$, defined as*

$$R(i,j) = \varphi\left(\tilde{P}_{(i,j)}\right) \quad \text{for } i,j \in \Xi \tag{1}$$

is a fuzzy preference relation on Ξ if and only if the following is satisfied for all f, $g \in \mathcal{M}$:

$$\varphi(f \wedge g) \geq T_L\left(\varphi(f), \varphi(g)\right) \tag{2}$$

Definition 3 ([3]) *A normalized evaluator $\varphi : \mathcal{M} \to [0,1]$ is said to be a T_L evaluator on $\mathcal{M}$, if it satisfies the formula (2).*

1.3 Comparability of elements of $\mathcal{K}$ and Łukasiewicz ultrafilters

In [3] it is shown that, in fact, T_L evaluators are generalizations of Łukasiewicz filters. First, recall the definition of a Łukasiewicz filter:

Definition 4 ([7]) *Let $\mathcal{K}$ be an at most countable set. A set-function $\mathcal{F} : 2^{\mathcal{K}} \to [0,1]$ is said to be a Łukasiewcz filter iff:*

1. $\mathcal{F}(\emptyset) = 0$, $\mathcal{F}(\mathcal{K}) = 1$,
2. *If $A \subseteq B \subseteq \mathcal{K}$, then $\mathcal{F}(A) \leq \mathcal{F}(B)$,*
3. *for all A, $B \in \mathcal{K}$ the following is satisfied*

$$T_L\left(\mathcal{F}(A), \mathcal{F}(B)\right) \leq \mathcal{F}(A \cap B).$$

Lemma 1 (a) *Let φ be a T_L evaluator on $\mathcal{M}$. Then there exists some Łukasiewicz filter $\mathcal{F}$ on $\mathcal{K}$ such that*

$$\forall A \subseteq \mathcal{K} \quad \varphi(1_A) = \mathcal{F}(A) \tag{3}$$

where 1_A is the characteristic function of A.
(b) *Let $\mathcal{G}$ be a Łukasiewicz filter on $\mathcal{K}$. Then there exists some T_L evaluator ψ such that*

$$\forall A \subseteq \mathcal{K} \quad \psi(1_A) = \mathcal{G}(A).$$

Definition 5 *We say that a T_L evaluator φ on $\mathcal{M}$ is na extension of a Łukasiewicz filter $\mathcal{F}$ on $\mathcal{K}$ if formula (3) is fulfilled.*

Yet, let us recall the definitions of Łukasiewicz ideal and Łukasiewicz ultrafilter form [6].

Definition 6 ([6]) *Let $\mathcal{K}$ be an at most countable set. A set-function $\mathcal{I} : 2^{\mathcal{K}} \to [0,1]$ is said to be a Łukasiewcz ideal on $\mathcal{K}$ iff:*

1. $\mathcal{I}(\emptyset) = 1$, $\mathcal{I}(\mathcal{K}) = 0$,
2. *If $A \subseteq B \subseteq \mathcal{K}$, then $\mathcal{I}(A) \geq \mathcal{I}(B)$,*
3. *for all A, $B \in \mathcal{K}$ the following is satisfied*

$$T_L\left(\mathcal{I}(A), \mathcal{I}(B)\right) \leq \mathcal{I}(A \cup B).$$

Definition 7 ([6]) *Let $\mathcal{U}$ be a Łukasiewicz filter on $\mathcal{K}$. It is said to be a Łukasiewicz ultrafilter iff $\mathcal{I}(A) = 1 - \mathcal{U}(A)$ is a Łukasiewicz ideal.*

As it will be shown next, Łukasiewicz ultrafilters are closely connected with the comparability of elements of $\mathcal{K}$.

Definition 8 *Let $\varphi : \mathcal{M} \to [0,1]$ be a T_L evaluator and $R_\varphi : \Xi \times \Xi \to [0,1]$ be a fuzzy preference relation defined by (1). Then $J_\varphi : \Xi \times \Xi \to [0,1]$ is said to be a relation of incomparability with respect to R_φ iff it is defined by:*

$$J_\varphi(i,j) = 1 - S_L\left(R_\varphi(i,j),\, R_\varphi(j,i)\right).$$

Theorem 2 *Let $\varphi : \mathcal{M} \to [0,1]$ be a T_L evaluator which is an extension of Łukasiewicz ultrafilter $\mathcal{U} : 2^{\mathcal{K}} \to [0,1]$. Then for all couples $i, j \in \Xi$ $J_\varphi(i,j) = 0$.*

Example 2 Let us have two digital cameras, c_1 and c_2, which we want to compare. We want to have a camera with at least 3× opt. zoom and at least 3 megapixels. We have a set of parameters:

Table 1: Table of parameters.

PARAMETER	c_1	c_2
price (Euro)	200	300
megapixels	4	5
opt. zoom	5×	8×
digit. zoom	3×	3×
digit. effects	good	very good

where we have the following linguistic values: {very weak, weak, average, good, very good, excellent} and we can assign then numerical values 0 up to 5.

Now, we may fuzzify the order for each parameter:

Table 2: Table of partial fuzzy order.

PARAMETER	$P_k(c_1,c_2)$	$P_k(c_2,c_1)$
price (Euro)	1	$\frac{2}{3}$
megapixels	$\frac{1}{2}$	1
opt. zoom	$\frac{2}{5}$	1
digit. zoom	1	1
digit. effects	$\frac{3}{4}$	1

To model the fuzzification of orders we have used the quotient of parameters. For megapixels and optical zoom we have 'shifted the zero' to utilise the requirement to have at least 3 megapixels and at least 3× opt. zoom (i.e. we have subtracted 3 from the values of both parameters before counting the quotient).

Now, we have to choose some Łukasiewicz filter modelling the importance of each group of parameters. We will use a special Łukasiewicz filter, so-called (T_L, S)-filter ([7]), where the t-conorm S will be the ordinal sum $\{(]0, 0.1[, S_M), (]0.1, 1[, S_L)\}$ since parameters whose weight is at most 0.1 are not very important.

I.e., the (T_L, S)-filter, $\mathcal{F} : 2^{\mathcal{K}} \to [0,1]$, is as follows:

$$\mathcal{F}(A) = \begin{cases} 1, & \text{if } A = \mathcal{K} \\ 0, & \text{if } A = \emptyset \\ S(a_1, \ldots, a_i), & \text{if } A = \{a_1, \ldots, a_i\} \end{cases}$$

Now, we have to extend the Łukasiewicz filter $\mathcal{F}$ to some T_L evaluator. We will do this using Choquet integral. We have to evaluate two functions (fuzzy sets): $f_1 = 1/1 + 2/0.5 + 3/0.4 + 4/1 + 5/\frac{3}{4}$ and $f_2 = 1/\frac{2}{3} + 2/1 + 3/1 + 4/1 + 5/1$

Table 3: Table of weights of parameters.

PARAMETER	WEIGHT
price	0.3
m-pixels	0.4
opt. zoom	0.4
digit. zoom	0.1
digit. effects	0.1

$$c\int f_1\, d\mathcal{F} = 0.605$$
$$c\int f_2\, d\mathcal{F} = 0.9$$

This means that we prefer camera c_2 to c_1 at the level 0.9 and camera c_1 to c_2 at the level 0.61.

2 Evaluators, Łukasiewicz filters and Łukasiewicz ideals

In Definition 6 Łukasiewicz ideals are defined. These correspond to so-called S_L evaluators:

Definition 9 ([3]) *A normalized evaluator $\varphi : \mathcal{M} \to [0,1]$ is an S_L evaluator if the following is satisfied:*

$$\varphi(a \vee b) \le S_L(\varphi(a), \varphi(b)).$$

Lemma 2 ([3]) *Let $\varphi : \mathcal{M} \to [0,1]$ be an S_L evaluator. Then there exists a Łukasiewicz ideal $\mathcal{I} : 2^{\mathcal{K}} \to [0,1]$ such that for each $A \in \mathcal{K}$ the following holds:*

$$\varphi(A) = 1 - \mathcal{I}(A).$$

Corollary 1 (to Lemmas 1, 2) *Let $\psi : \mathcal{M} \to [0,1]$ be a T_L and S_L evaluator. Then function $\mathcal{U} : 2^{\mathcal{K}} \to [0,1]$ defined by*

$$\mathcal{U}(A) = \psi(A) \quad \text{for } A \subseteq \mathcal{K}$$

is a Łukasiewicz ultrafilter.

Due to Corollary 1 T_L and S_L evaluators play important role in constructing fuzzy preference relations. We show now some possibilities how Łukasiewicz ultrafilters can be extended to T_L and S_L evaluators.

Theorem 3 *For each $k \in \mathcal{K}$ let $\eta_k : [0,1] \to [0,1]$ be some isotone transformation such that $\eta_k(0) = 0$ and $\eta_k(1) = 1$. Further, let $\mathcal{U} : 2^{\mathcal{K}} \to [0,1]$ be a Łukasiewicz ultrafilter. Then $\varphi : \mathcal{M} \to [0,1]$ defined by*

$$\varphi(f) = c\int \eta_k(f(k))\, d\mathcal{U} \tag{4}$$

is T_L and S_L evaluator, where $c\int$ means Choquet integral.

Theorem 4 *For each $k \in \mathcal{K}$ let $\eta_k : [0,1] \to [0,1]$ be some isotone transformation such that $\eta_k(0) = 0$ and $\eta_k(1) = 1$. Further, let $\mathcal{U} : 2^{\mathcal{K}} \to [0,1]$ be a Łukasiewicz ultrafilter. Then $\varphi : \mathcal{M} \to [0,1]$ defined by*

$$\varphi(f) = s\int \eta_k(f(k))\, d\mathcal{U} \tag{5}$$

is T_L and S_L evaluator, where $s\int$ means Sugeno integral.

The transformations η_k may be used, e.g., to make the interval $[0,1]$ discrete. In fact, the fuzzified order P_k from subsection 1.1 can be made discrete via η_k. The biggest and the smallest possible transformations are, respectively:

$$\eta_b(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \end{cases}$$

$$\eta_s(x) = \begin{cases} 1, & \text{if } x = 1 \\ 0, & \text{if } x < 1 \end{cases}$$

However, if we use identity as transformations η_k and we change the Choquet or Sugeno integrals from formulae (4) or (5) to some other fuzzy integral (general fuzzy integrals were introduced in [9]), the resulting evaluator, extending some Łukasiewicz ultrafilter $\mathcal{U}$, is not necessarily T_L and/or S_L evaluator as we will see next.

Definition 10 ([9]) *A functional $I_\mu : \mathcal{M} \to [0,1]$ is a general fuzzy integral with respect to some measure $\mu : 2^{\mathcal{K}} \to [0,1]$ iff the following is fulfilled:*

1. *$I_\mu(f) \le I(g)$ whenever $f \le g$, $f, g \in \mathcal{M}$,*
2. *$I_\mu(\tilde{c}) = c$ for any constant function $\tilde{c} \in \mathcal{M}$,*
3. *$I_\mu(1_A) = \mu(A)$ where 1_A is the characteristic function of $A \subset \mathcal{K}$.*

Theorem 5 ([9]) *Let $\mu : 2^{\mathcal{K}} \to [0,1]$ be some fuzzy measure. Then the strongest fuzzy integral with respect to μ is $I^*_\mu : \mathcal{M} \to [0,1]$ defined as:*

$$I^*_\mu(f) = \min\{\sup f, \mu(f > 0)\} \tag{6}$$

and the weakest fuzzy integral with respect to μ is $I_{*\mu} : \mathcal{M} \to [0,1]$ defined as:

$$I_{*\mu}(f) = \max\{\inf f, \mu(f = 1)\} \tag{7}$$

Theorem 6 *Let $\mathcal{F} : 2^{\mathcal{K}} \to [0,1]$ be a Łukasiewicz filter. Then $I^*_{\mathcal{F}} : \mathcal{M} \to [0,1]$, defined by formula (6) is an evaluator which is not T_L.*

Theorem 7 *Let $\mathcal{I} : 2^{\mathcal{K}} \to [0,1]$ be a Łukasiewicz ideal. Denote $\mu = 1 - \mathcal{I}$. Then $I_{*\mu} : \mathcal{M} \to [0,1]$, defined by formula (7) is an evaluator which is not S_L.*

The last assertion is trivial, however, for completeness sake we formulate the statement:

Lemma 3 *Let $\mathcal{U} : 2^{\mathcal{K}} \to [0,1]$ be some Łukasiewicz ultrafilter. Further, let $\mathcal{E}_{\mathcal{U}}$ denote the system of all fuzzy evaluators $\varphi : \mathcal{M} \to [0,1]$ extending $\mathcal{U}$. Then the system $\mathcal{E}_{\mathcal{U}}$ is convex.*

3 Conclusion

In this paper we have shown how T_L evaluators and S_L evaluators can be used to construct fuzzy preference relations. Further, we have shown under which condition the incomparability relation J is zero. We have also shown some extensions of Łukasiewicz ultrafilters which are T_L and S_L evaluators and some extensions of Łukasiewicz ultrafilters which are either not T_L or not S_L evaluators.

Acknowledgement

This work was supported by Science and Technology Assistance Agency under the contract No. APVV-0375-06, and by the VEGA grant agency, grant number 1/3014/06.

References

[1] U. Bodenhofer. A similarity-based generalisation of fuzzy orderings. Universitätsverlag R. Trauner, Linz, 1999.

[2] U. Bodenhofer. Similarity-based generalisation of fuzzy orderings preserving the classical axioms, Int. J. uncertainty, Fuzziness and Knowledge-Based Systems 8 (2000), 593-610.

[3] S. Bodjanova. T_L and S_L evaluators. In: *Proceedings of AGOP 2007*, submitted.

[4] D. Dubois, W. Ostasiewicz, H. Prade. Fuzzy sets: History and basic notations. In: Fundamentals of Fuzzy Sets, Kluwer, Dordrecht, 2000.

[5] J. Fodor and M. Roubens. Fuzzy preference modeling and Multicriteria Decision Support. Kluwer, Dordrecht, 1994.

[6] Z. Havranová, M. Kalina. Łukasiewicz filters as convex combinations. In: Proc. of EUSFLAT 2007, Ostrava, submitted.

[7] M. Kalina. Łukasiewicz fiters and similarities. In: *Proceedings of AGOP 2005*, Lugano, 2005 pp. 57-60.

[8] E.P. Klement, R. Mesiar, E. Pap. Triangular norms. Kluwer, Dordrecht, 2000.

[9] P. Struk, Extremal fuzzy integrals, Soft Computing 10 (2006), 502-505.

On a Proximity-Based Tolerant Inclusion[1]

Patrick Bosc
IRISA – ENSSAT
Technopole Anticipa
BP 80518
22305 Lannion Cedex
France
bosc@enssat.fr

Allel HadjAli
IRISA – ENSSAT
Technopole Anticipa
BP 80518
22305 Lannion Cedex
France
hadjali@enssat.fr

Olivier Pivert
IRISA – ENSSAT
Technopole Anticipa
BP 80518
22305 Lannion Cedex
France
pivert@enssat.fr

Abstract

This paper is devoted to an extension of the inclusion operator. The main idea that we suggest is to relax the arguments of the inclusion by means of proximity-based modifiers in order to obtain a more tolerant inclusion. The use of such a relaxed inclusion is illustrated in the framework of databases by the extension of the division operator.

Keywords: inclusion, tolerant inclusion, proximity relation, relational division.

1 Introduction

Several types of extensions of the inclusion have been proposed in the "fuzzy set research community" in order to either: i) define the inclusion when fuzzy sets come into play, or ii) to make the result of the inclusion more flexible, i.e., valued in [0, 1], or iii) to authorize different types of exceptions. In the first case, Zadeh [11] defined the inclusion of fuzzy sets in the following way: $E \subseteq F \Leftrightarrow \forall x \in U$, $E(x) \leq F(x)$, where U denotes the universe of discourse. In the second case, the objective is to discriminate between situations significantly different where the usual inclusion does not hold, and a solution consists in using a fuzzy implication to define the graded inclusion [1]. In the third case, two visions of exceptions have been considered so far, which lead to two types of approximate inclusion indicators. A first idea, developed in [3], consists in weakening the universal quantifier underlying the inclusion into "almost all", in the perspective of defining an *approximate* inclusion. The basic idea is to tolerate, in the evaluation of $E \subseteq F$, a certain number of exceptions (i.e., of elements of E which are not totally included in F according to a given implication), and in that sense the corresponding approximate inclusion can be called a *quantitative* one. In [4], another way of defining an approximate inclusion is presented and the idea is rather to give a central role to the intensity of the exceptions in order to define an inclusion indicator that can ignore to a certain extent the "low intensity" ones. In that sense, the operator defined can be called a *qualitative* approximate inclusion.

Here, the idea is to take into account the notion of closeness between the elements of the domain, so as to define a *proximity-based tolerant inclusion*. For instance, one may consider that a set E is included in a set F if, for every element x of E, x is present in F (classical inclusion) or if F contains an element *close* to x.

The rest of the paper is structured as follows. In section 2, we recall some basic notions related to the inclusion of fuzzy sets in the sense of Zadeh [11], as well as the way of defining proximity-based modifiers. In section 3, the principle of the proximity-based inclusion is introduced and we show how it can be applied to crisp sets as well as fuzzy sets. Section 4 illustrates the practical utility of such a tolerant inclusion in the context of database querying where a relaxation of the division operator is studied. Section 5 provides a brief comparison between the proximity-based approach and the two approximate inclusion indicators introduced in

[1] Work partially funded by the ANR Research Grant ANR-05-MMSA-0007

[3] and [4]. The last section concludes the paper. The objective and main contributions are recalled and some perspectives as to future works are discussed.

2 Preliminaries and Background

2.1 Boolean Inclusion of Fuzzy Sets

If A and B denote two crisp sets built on U, the usual way for defining the inclusion of A in B is:

$$(A \subseteq B) \Leftrightarrow (\forall x \in U, x \in A \Rightarrow x \in B) \quad (1)$$

This definition can be extended in a canonical way to two fuzzy sets A and B, which leads to:

$$(A \subseteq B) \Leftrightarrow (\forall x \in U, A(x) \rightarrow_{R\text{-}G} B(x)) \quad (2)$$

where $\rightarrow_{R\text{-}G}$ stands for Rescher-Gaines implication:

$p \rightarrow_{R\text{-}G} q = 1$ if $p \leq q$, 0 otherwise.

This view does not take into account the proximity between the elements of the universe, therefore it can happen that the result of the inclusion is false while it would be true if that notion was used.

Example 1. Let A and B be the two fuzzy sets :

$A = \{1/a + 0.6/b\}$, $B = \{1/a + 0.4/b + 0.9/c\}$.

According to formula (2), A is not included in B. However, if one has available the knowledge that the element b is very close (according to a proximity relation defined on the universe considered) to the element c (which strongly belongs to B), it may make sense to upgrade the membership degree of b to B (which corresponds to modifying B into a fuzzy set B') and then A may be included in B'. In that sense, $A \subseteq B'$ can be viewed as a relaxation of $A \subseteq B$. This is the basic idea that is developed in the following.♦

2.2 About Proximity-Based Modifiers

Let us first recall the formal definition of the concept of a *proximity relation* [9].

Definition 1. A proximity relation is a fuzzy relation R on a scalar domain U, such that for $u, v \in U$,

$R(u, u) = 1$ (*reflexivity*),

$R(u, v) = R(v, u)$ (*symmetry*).

The quantity $R(u, v)$ can be viewed as a grade of *approximate equality* of u with v.

An *absolute proximity* relation is an approximate equality relation E which can be modeled by a fuzzy relation of the form [9]:

$$E: U \times U \rightarrow [0, 1]$$

$$(u, v) \rightarrow E(u, v) = Z(u - v),$$

which only depends on the value of the difference $u - v$, and where Z, called a *tolerance indicator*, is a fuzzy interval centered in 0, such that:

i. $Z(r) = Z(-r)$, i.e., $Z = -Z$;

ii. $Z(0) = 1$;

iii. The support of Z is bounded and is of the form $[-\Omega, \Omega]$ where Ω is a positive real number.

In terms of trapezoidal membership function (t.m.f.), the parameter Z can be expressed by $(-z, z, \delta, \delta)$ with $\Omega = z + \delta$ and $[-z, z]$ represents the core of Z.

Proposition 1. Let Z_1 and Z_2 be two fuzzy intervals centered in 0 on scalar domain U. The following entailment holds:

$$Z_1 \subseteq Z_2 \Rightarrow E[Z_1] \subseteq E[Z_2].$$

The proof is straightforward. In the following, we will write $E[Z]$ to denote the absolute proximity relation E parameterized by Z. See [9] for other interesting properties of $E[Z]$.

Consider a fuzzy set F on the scalar domain U and an absolute proximity relation $E(Z)$, where Z is a tolerance indicator. The set F can be associated with a nested pair of fuzzy sets when using $E(Z)$ as a *tolerance relation*. Indeed,

i. we can build a fuzzy set $E^Z(F)$ close to F, such that $F \subseteq E^Z(F)$. This is the *dilation operation*. $E^Z(F)$ gathers the elements of F and those outside of F which are somewhat close to an element in F in the sense of E[Z].

ii. we can build a fuzzy set $E_Z(F)$ close to F, such that $E_Z(F) \subseteq F$. This is the *erosion operation*. $E_Z(F)$ gathers the elements of F such that all of their "neighbors" – i.e. those which are somewhat close to them – are in F.

These fuzzy sets can be constructed in the following way.

Dilation operation. Dilating the fuzzy set F by Z will provide a fuzzy set $E^Z(F)$ defined by:

$$E^Z(F)(s) = sup_{r\in U} \top(F(r), E[Z](s, r)) = sup_{r\in U} \top(F(r), Z(s-r)) \quad (3)$$

where $\top$ is a triangular norm.

Then, $F \subseteq E^Z(F)$ *and* $E^Z(F)$ *can be viewed as a weakened variant of* F.

Erosion Operation. Considering the meaning of $E_Z(F)$ given above, it seems natural to adopt the following definition:

$$E_Z(F)(s) = inf_{r\in U} (E[Z](s, r) \rightarrow F(r)) \quad (4)$$

where $\rightarrow$ *denotes a fuzzy implication. We will see later (in section 3.3) that the preservation of an important axiom of the inclusion constrains the choice of the fuzzy implication in (4): it has to be the S-implication based on the triangular norm used in (3), which will be denoted by* $\rightarrow_\top$.

E^Z and E_Z are fuzzy modifiers satisfying the following proposition.

Proposition 2. For F in $\mathcal{P}(U)$, we have:

$E_Z(F) \subseteq F \subseteq E^Z(F)$.

The proof is straightforward.

3 Proximity-Based Tolerant Inclusion

The basic idea is to introduce a certain tolerance into the inclusion indicator by taking into account the proximity between the elements of the domain considered. This can be done by replacing $A \subseteq B$ either by:

$A \subseteq_Z^1 B \equiv E_Z(A) \subseteq E^Z(B)$ or by:

$A \subseteq_Z^2 B \equiv (E_Z(A) \subseteq B \vee A \subseteq E^Z(B))$ or by:

$A \subseteq_Z^3 B \equiv (E_Z(A) \subseteq B \wedge A \subseteq E^Z(B))$.

Remark 1. One has:

$A \subseteq_Z^3 B \Rightarrow A \subseteq_Z^2 B \Rightarrow A \subseteq_Z^1 B$.

In the following, we focus on indicator $\subseteq_Z^3$, which acts both on A and B, and which is the most drastic way (among the three pointed out above) of relaxing the inclusion. We use the notation:

$$A \subseteq_Z B \Leftrightarrow (E_Z(A) \subseteq B \wedge A \subseteq E^Z(B)). \quad (5)$$

Remark 2. One could also consider two other relaxations, namely

$A \subseteq_Z^4 B \equiv E_Z(A) \subseteq B$ and

$A \subseteq_Z^5 B \equiv A \subseteq E^Z(B)$,

but it seems intuitively more reasonable, in general, to apply a relaxation mechanism that acts on both arguments.

To sum up, the principle is to say that A is included (with tolerance) in B iff *very(A)* is included in B and A is included in *more-or-less(B)* where the linguistic modifiers *very* and *more-or-less* are based on the notion of proximity.

3.1 Tolerant Inclusion of Crisp Sets

In the case where crisp sets are dealt with, one must use a Boolean proximity relation $E[Z]$ based on a regular interval $Z = [-\Omega, \Omega]$ centered in 0. One gets:

$E[Z](u, v)$ is true if $|u - v| \leq \Omega$, false otherwise.

Formulas (3) and (4) rewrite:

$$E^Z(F) \equiv \{s \in U \mid \exists r \in U \text{ such that } r \in F \wedge E[Z](r, s)\} \quad (6)$$

$$E_Z(F) \equiv \{s \in F \mid \forall r \in U, E[Z](r, s) \Rightarrow r \in F\} \quad (7)$$

Example 2. Let us consider the sets

$A = \{41, 59\}$ and $B = \{40, 48, 60\}$

defined on the interval [0, 100] of the integers, and the interval $Z = [-1, 1]$. We get:

$E_Z(A) = \emptyset$ and

$E^Z(B) = \{39, 40, 41, 47, 48, 49, 59, 60, 61\}$

and we have: $A \subseteq_Z B$.♦

3.2 Tolerant Inclusion of Fuzzy Sets

Here, the calculus – illustrated by the following example – is based on formulas (3) and (4).

Example 3. Let us consider the fuzzy sets:

$A = \{0.7/47, 0.9/48, 0.6/49, 1/50\}$

$B = \{0.6/41, 0.7/48, 1/49\}$

defined on the interval [0, 100] of the integers, and the fuzzy set Z represented by the t.m.f. (1, –1, 2, 2). According to (2), A is not included in B.

Using the triangular norm minimum in (3), and thus Kleene-Dienes implication in (4), we get:

$E_Z(A) = \{0.5/48, 0.5/49\}$ and

$E^Z(B) = \{0.5/39, 0.6/40, 0.6/41, 0.6/42, 0.5/43,$

0.5/46, 0.7/47, 1/48, 1/49, 1/50, 0.5/51}

$A \subseteq_Z B = (E_Z(A) \subseteq B)) \wedge (A \subseteq E^Z(B))) =$

$true \wedge true = true.$ ♦

3.3 Properties

Let us recall the three axioms valid for Boolean inclusion:

$$A \subseteq B \Leftrightarrow B_c \subseteq A_c \quad \text{(A1)}$$

where A_c (resp. B_c) denotes the complement of A (resp. B) in the universe U;

$$A \subseteq (B \cap C) \Leftrightarrow (A \subseteq B) \wedge (A \subseteq C) \quad \text{(A2)}$$

$$A \subseteq B \Leftrightarrow S(A) \subseteq S(B) \quad \text{(A3)}$$

where the set $S(A)$ is defined as $S(A)(x) = A(S(x))$ with a one-to-one mapping $s: X \to X$.

Let us now check whether these axioms remain valid when the regular inclusion is replaced by a tolerant one.

First case: crisp sets.

Axiom (A1). Do we have: $A \subseteq_Z B \Leftrightarrow B_c \subseteq_Z A_c$?

First, let us show that for a crisp set X, one has: $(E^Z(X))_c = E_Z(X_c)$.

Proof.

$(E^Z(X))_c(s) =$
$U - \{s \in U \mid \exists r \in U \text{ s.t. } r \in X \wedge E[Z](r, s)\}$
$= \{s \in U \mid \forall r \in U, r \notin X \vee \neg E[Z](r, s)\}$
$= \{s \in U \mid \forall r \in U, E[Z](r, s) \Rightarrow r \notin X\}$
$= E_Z(X_c).$ •

From this, it is straightforward to deduce that $(E_Z(X))_c = E^Z(X_c)$. Therefore, we have:

$A \subseteq_Z B \Leftrightarrow$

$(E_Z(A) \subseteq B \wedge A \subseteq E^Z(B)) \Leftrightarrow$

$B_c \subseteq (E_Z(A))_c \wedge (E^Z(B))_c \subseteq A_c \Leftrightarrow$

$B_c \subseteq E^Z(A_c) \wedge E_Z(B_c) \subseteq A_c \Leftrightarrow$

$B_c \subseteq_Z A_c$.

Axiom (A2). Do we have:

$A \subseteq_Z (B \cap C) \Leftrightarrow (A \subseteq_Z B) \wedge (A \subseteq_Z C)$?

First, let us notice that for two crisp sets B and C, one has: $E^Z(B \cap C) = E^Z(B) \cap E^Z(C)$ (the proof is straightforward). Therefore, we have:

$A \subseteq_Z (B \cap C) \Leftrightarrow$
$E_Z(A) \subseteq (B \cap C) \wedge A \subseteq E^Z(B \cap C) \Leftrightarrow$
$E_Z(A) \subseteq B \wedge E_Z(A) \subseteq C \wedge A \subseteq E^Z(B)$
$\wedge A \subseteq E^Z(C) \Leftrightarrow$
$(A \subseteq_Z B) \wedge (A \subseteq_Z C).$

As to axiom(A3), it straightforwardly holds when $\subseteq$ is replaced by $\subseteq_Z$.

Second case: fuzzy sets.

Axiom (A1). Let us prove that this axiom is preserved if the fuzzy implication used in (4) is the S-implication based on the triangular norm used in (3). Let us recall that an S-implication $\to_\top$ is associated with a triangular norm $\top$ by the following relation:

$$a \to_\top b = 1 - \top(a, 1 - b). \quad (8)$$

Thus we have:

$E^Z(X_c)(s) = sup_{r \in U} \top(E[Z](s, r), 1 - X(r))$
$= sup_{r \in U} (1 - [E[Z](s, r) \to_\top X(r))])$
$= 1 - inf_{r \in U} (E[Z](s, r) \to_\top X(r))$
$= 1 - E_Z(X)(s)$
$= (E_Z(X))_c(s).$

From this, it is straightforward to deduce that:

$(E^Z(X))_c = E_Z(X_c)$.

It follows that axiom (A1) holds (the proof is the same as in the crisp set case).

Axiom (A2). Only a weakened form of it holds when $\subseteq$ is replaced by $\subseteq_Z$ (the equivalence is replaced by an implication). First let us show that: $E^Z(B \cap C) \neq E^Z(B) \cap E^Z(C)$ in general.

Let $U = \{10, 11, 12\}$, $B = \{0.4/10, 0.7/11\}$, and $C = \{0.3/11, 0.8/12\}$. Let Z be represented by the t.m.f. (0, 0, 2, 2). Let us assume in this example that the intersection is interpreted by a minimum.

$B \cap C = \{0.3/11\}$

$E^Z(B \cap C) = \{0.3/10, 0.3/11, 0.3/12\}$

$E^Z(B) = \{0.5/10, 0.7/11, 0.5/12\}$

$E^Z(C) = \{0.3/10, 0.5/11, 0.8/12\}$

$E^Z(B) \cap E^Z(C) = \{0.3/10, 0.5/11, 0.5/12\}$.

Now, let us show that:

$E^Z(B \cap C) \subseteq E^Z(B) \cap E^Z(C)$.

Proof.

$\mu_{E^Z(B \cap C)}(x) = sup_{y \in U} \top(E[Z](x, y), \mu_{B \cap C}(y))$

$= sup_{y \in U} \top(E[Z](x, y), \top(\mu_B(y), \mu_C(y)))$

$\leq \top(sup_{y \in U} \top(E[Z](x, y), \mu_B(y)),$

$$sup_{y \in U} \top(E[Z](x, y), \mu_C(y)))$$

$$\leq \top(\mu_{E^Z}(B)(x), \mu_{E^Z}(C)(x)).\bullet$$

This result entails that:

$A \subseteq_Z (B \cap C) \Rightarrow (A \subseteq_Z B) \wedge (A \subseteq_Z C)$.

The proof is the same as in the crisp set case, except that in the second line the equivalence must be replaced by an implication.

Remark. Dubois and Prade [10] have shown that S-implications and R-implications could be merged into a single family, provided that the class of triangular norms is enlarged to non-commutative conjunction operators. One has:

$$a \rightarrow b = 1 - cnj(a, 1 - b)$$

where *cnj* is a a triangular norm $\top$ if $\rightarrow$ is an S-implication, and *cnj* is a non-commutative conjunction *cnc* if $\rightarrow$ is an R-implication. This provides another way of defining E^Z and E_Z: one could use a non-commutative conjunction in (3) and the corresponding R-implication in (4), which would also preserve axiom (A1).

4. Application to the division of relations

4.1. Reminder About the Division

The relational division, i.e., the division of relation *r* of schema *R(A, X)* by relation *s* of schema *S(B)* (where *A* and *B* are compatible subsets of attributes) is defined as follows:

$$div(r, s, A, B) = \{x \mid s \subseteq \Omega(x)\} \quad (9)$$

where $\Omega(x) = \{a \mid \langle x, a \rangle \in r\}$. In other words, an element *x* belongs to the result of the division if it is associated with *at least all* the values *a* of *B* appearing in *s*.

Example 4. Consider the relations suppliers (*s*) and census (c) of respective schemas *S(#store, chain, zipc, turnover)* and *C(city, zipc, pop)*. The query: "find the chains which have a store with a turnover greater than 0.5 k€ in every city from relation *c* whose population is over 200,000" can be formulated using a division:

div(proj(select(s, turnover > 0.5), {chain, zipc}),
proj(select(c, pop > 200,000), {zipc}),
{zipc}, {zipc})

where *select(r, cond)* is the selection of relation *r* with the condition cond and *proj(r, X)* is the projection of *r* onto the set of attributes *X*.

s

#store	chain	zipc	turnover
15	32	75000	1.2
12	32	69000	0.54
34	32	22000	0.25
26	32	13000	0.89
26	7	49000	0.37
78	7	35000	0.51

c

#city	zipc	pop
Paris	75000	2 125 800
Lyon	69000	445 274
Saint-Brieuc	22000	46 089
Marseille	13000	797 491
Angers	49000	151 322
Rennes	35000	206 194

Using the extensions of *s* and *c* above, one gets the result made of the single value <32>.♦

4.2. Tolerant Division of Crisp Relations

The tolerant inclusion introduced before can serve as a basis to define a tolerant division where the notion of proximity between values is taken into account. For instance, in the previous example, it might be useful to consider that often the stores are not in the big cities themselves but in their suburbs. Then, the zip codes in relation Supplier may be different from the ones associated with the different urban centers. Therefore, it is interesting to allow a certain tolerance on the zip code in the processing of the division (it is assumed that the difference between the zip codes reflects the distances between the cities). For instance, one may use a tolerance indicator such as *Z* = (–900, 900, 0, 0). The tolerant division is interpreted by means of formula (9) where the regular inclusion is replaced by a tolerant one – which is itself interpreted using formulas (5) and (6).

Example 5. Let us consider again the context of example 4 and let us consider the extension of relation s below. For the chain 32, it can easily be checked that the set A = {75000, 69000, 13000, 35000} is included (with tolerance) in B = {75015, 69215, 22300, 13625}. Thus the result of the tolerant division is {32} whereas it would be empty if the regular division were used.

s	#store	chain	zipc	turnover
	15	32	75015	1.2
	12	32	69215	0.54
	34	32	22300	0.25
	26	32	13625	0.89
	26	7	49000	0.37
	78	7	35830	0.51

♦

4.3. Tolerant Division of Fuzzy Relations

Now, one considers fuzzy relations, i.e., relations obtained by applying flexible conditions on relations pertaining to a regular relational database. In this context, one can envisage queries similar to that of example 4, for instance the request: "find every chain *x*, to what extent is it true that *x* has a store with an important turnover in every largely populated city present in relation *c*". This query can be algebraically expressed thanks to a division of fuzzy relations as:

div(proj(select(s, turnover is important),
{chain, zipc}),
proj(select(c, pop is large), {zipc}),
{zipc}, {zipc}).

A fuzzy relation is defined as a fuzzy subset of the Cartesian product of domains. So, a fuzzy relation r is a set of weighted tuples, denoted by μ/t where μ is the membership degree of t in r, i.e., the extent to which t complies with the concept associated with r. In the context of the previous query, the degree tied to each tuple of relation *s' = select(s, turnover is important)* expresses the extent to which it is a store with an important turnover. A flexible query is made of operations applying to fuzzy relations and whose result is a fuzzy relation as well. Such operators are obtained through a natural extension of the usual algebraic operators (see, e.g., [5, 2]).

A straightforward (Boolean) extension of the division operator to fuzzy relations can be based on formula (9) where the inclusion is defined by formula (2). As to the *tolerant* division of fuzzy relations, it consists, as in the Boolean relation case, in replacing the inclusion present in formula (9) by a tolerant one. Then, the calculus of the tolerant inclusion is based on formulas (3) and (4).

5. Related works

As mentioned in the introduction, two other approaches have been previously proposed in order to relax the inclusion by tolerating exceptions.

The operator introduced in [3], called quantitative approximate inclusion, aims at authorizing, in the evaluation of $E \subseteq F$, a certain proportion of exceptions, i.e., of elements from E which are not in F (in the sense of a fuzzy implication which can be Rescher-Gaines'one when crisp sets are dealt with). Obviously, the principle of that approach is totally different from the one described here since tolerance in the proximity-based inclusion we propose is not a matter of a *number* of exceptions.

As to the operator introduced in [4], it aims at relaxing the family of graded inclusion indicators based on R-implications so as to authorize "low intensity" exceptions. Let us recall that R-implications can be expressed:

$p \rightarrow_{R\text{-}i} q = 1$ if $p \leq q$, $f(p, q)$ otherwise

where $f(p, q)$ expresses a degree of satisfaction of the implication when the antecedent p exceeds the conclusion q. The implications of Gödel ($p \rightarrow_{Gö} q = 1$ if $p \leq q$, q otherwise), Goguen ($p \rightarrow_{Gg} q = 1$ if $p \leq q$, q/p otherwise) and Lukasiewicz ($p \rightarrow_{Lu} q = 1$ if $p \leq q$, $1 - p + q$ otherwise) are the three most used R-implications. The principle of the so-called qualitative approximate inclusion is to introduce a tolerance on the implication values, that takes into account the gap betwen the membership degrees to E and to F respectively, for a given element x. More precisely, it is considered that the situation where $\mu_E(x)$ exceeds $\mu_F(x)$ is more or less acceptable as long as the difference $\mu_E(x) - \mu_F(x)$ is in a given interval.

The proximity-based tolerant inclusion proposed here can be called qualitative too, but in a different way than that outlined above. In [4], an exception x is ignored depending on the *membership degrees* attached to x in E and F respectively whereas in the proximity-based approach it will be ignored depending on the *values* present in the sets (and on the tolerance indicator Z). A consequence of this difference is that the proximity-based approach can apply to crisp sets while the qualitative approximate inclusion described in [4] cannot. On the other hand, the proximity-based approach in its form

described here imposes that the domain be numeric. Nevertheless, in case of a non-numeric domain U, the approach can still be applied providing that a distance d be defined over U Then, E(u, v) can be defined as Z(d(u, v)) instead of Z(u – v).

As to proximity-based modifiers, the most similar work to our proposal is the recent one done by De Cock and Kerre in [8]. In that work, the authors introduce the two following fuzzy modifiers:

$$R^{\clubsuit}(A)(y) = \sup_{x \in X} C(R(x, y), A(x)),$$

$$R^{\nabla}(A)(y) = \inf_{x \in X} \mathcal{T}(R(x, y), A(x)),$$

where A is a fuzzy set on the universe X, C is a conjunctor, $\mathcal{T}$ an implicator and $R(x, y)$ a fuzzy relation. It has been shown that weakening and intensifying linguistic hedges can be modeled by such fuzzy modifiers. To do so, the authors make use of a particular suitable fuzzy relation. It is expressed by a *pseudo-metric based resemblance relation* that is reflexive and symmetric.

To provide some comparison between our approach and the one proposed by De Cock and Kerre, let us first emphasize that the motivations of these two approaches are quite different. Our starting point was mainly guided by the question: how can an *upper* and a *lower* approximation of a fuzzy set F be established using a parameterized proximity relation? It is that question that triggered the generalization of the use of a proximity relation to model fuzzy modifiers. Moreover, we can claim that our approach avoids the burden of the computation complexity since the calculus is reduced to simple fuzzy arithmetic operations. Let us also mention that the proximity relation-based approach to fuzzy modifiers seems more flexible from a user point of view. For instance, it allows for preserving either the core or the support if needed, when modifying a fuzzy set.

6. Conclusion

Various extensions of set inclusion have been proposed in the framework of fuzzy sets. In this paper, the novelty is to consider an proximity-based inclusion indicator, in order to take into account the closeness between the elements of the domain. Such an operator, based on a tolerance indicator defined over the domain considered, has been defined and its practical utility has been illustrated by the relational division of fuzzy relations. Let us notice that, beyond the database domain, the division operation plays a central role in information retrieval (and in web search engines) where the issue is to retrieve the documents that "include" all the concepts corresponding to a given set of keywords. The tolerant division we propose could thus enable to relax in an explicit way the requirement about the keywords that must appear in a document to make it more or less acceptable, providing that a semantic distance has been defined over the terms that may appear in the queries and the documents.

The tolerant inclusion considered in this paper is Boolean-valued, and one of the perspectives of this work is to extend it in order to define a graded tolerant inclusion of fuzzy sets. Then, it would be of interest to check whether the graded inclusion obtained satisfies the axioms pointed out in [7]. It would also be worth studying whether the quotient property of the division operator is preserved when one moves to a tolerant division such as that introduced in section 4.

References

[1] W. Bandler, L. Kohout, Fuzzy power sets and fuzzy implication operators, Fuzzy Sets and Systems 4 (1980) 13-30.

[2] P. Bosc, B. Buckles, F. Petry, O. Pivert, Fuzzy Databases, in: Fuzzy Sets in Approximate Reasoning and Information Systems, J. Bezdek, D. Dubois and H. Prade (Eds.), The Handbook of Fuzzy Sets Series, Kluwer Academic Publishers, 1999, pp. 403-468.

[3] P. Bosc, O. Pivert, About approximate inclusion and its axiomatization, Fuzzy Sets and Systems 157 (2006) 1438-1454.

[4] P. Bosc, O. Pivert, On a qualitative approximate inclusion – Application to the division of fuzzy relations, Proc. of the Int. Workshop on Flexible Database and Information System Technology (FlexDBIST'06), in conjunction with *DEXA'06*, Cracow, Poland, September 6, 430-434, 2006.

[5] P. Bosc, D. Dubois, O. Pivert, H. Prade, Flexible queries in relational databases – the example of the division operator, Theoretical Computer Science 171 (1997) 281-302.

[6] P. Bosc, A. Hadjali, O. Pivert, Modelling fuzzy modifiers by means of a proximity relation, Proc. of the 25th International Conference of the North American Fuzzy Information Processing

Society (NAFIPS'06), Montréal, Canada, June 3-6, 2006.

[7] C. Cornelis, C. Van der Donck, E. Kerre, Sinha-Dougherty approach to the fuzzification of set inclusion revisited, Fuzzy Sets and Systems 134 (2003) 283-296.

[8] M. De Cock, E. Kerre, Fuzzy modifiers based on fuzzy relations, Information Sciences, 160 (2004) 173-199.

[9] D. Dubois, A. HadjAli, H. Prade, Fuzzy qualitative reasoning with words, in: Computing with Words (P.P. Wang, Ed.), Vol. 3, John Wiley & Son, 2001, pp. 347-366.

[10] D. Dubois and H. Prade, A theorem on implication functions defined from triangular norms, Stochastica, vol. 8, pp. 267-279, 1984. Also in: Readings in Fuzzy Sets for Intelligent Systems (Dubois D., Prade H., Yager R.R. eds.), Morgan & Kaufmann, 1993, pp. 105-112.

[11] L.A. Zadeh, Fuzzy Sets, Information and Control 8 (1965) 371-384.

Finding close T-indistinguishability Operators to a given Proximity

L. Garmendia
Ingeniería del Software e Inteligencia Artificial
Facultad de Informática
Universidad Complutense de Madrid
C/ Profesor José García Santesmases, s/n
28040 Madrid
Spain
lgarmend@fdi.ucm.es

J. Recasens
Secció Matemàtiques i Informàtica
ETS Arquitectura del Vallès
Universitat Politècnica de Catalunya
Pere Serra 1-15
08190 Sant Cugat del Vallès
Spain
j.recasens@upc.edu

Abstract

Two ways to approximate a proximity relation R (i.e. a reflexive and symmetric fuzzy relation) by a T-transitive one where T is a continuous archimedean t-norm are given.

The first one aggregates the transitive closure $\overline{R}$ of R with a (maximal) T-transitive relation B contained in R.

The second one modifies the values of $\overline{R}$ or B to better fit them with the ones of R.

Keywords: Proximity, Transitive Closure, Opening, T-indistinguishability Operator, Aggregation Operator, Quasi Arithmetic Mean, Representation Theorem.

1 Introduction

A proximity matrix or relation on a finite universe X is a reflexive and symmetric fuzzy relation R on X. In many applications transitivity of R with respect to a t-norm T is required. In these cases, R must be replaced by a new relation E also satisfying transitivity, such relations called T-indistinguishability operators. Of course, it is desirable that E is as close as possible to R. This paper presents a couple of ways to find close transitive relations to R in a reasonable way - i.e.: easy and rapid to generate- when the t-norm is continuous archimedean.

There are of course several ways to calculate the closeness of two fuzzy relations, many of them related to some metric. In this paper we propose a way related to the natural indistinguishability operator E_T associated to T, so that the degree of closeness or similarity between two fuzzy relations R and S is calculated aggregating the similarity of their respective entries using the quasi-arithmetic mean generated by an additive generator of T.

Also the euclidean metric will be used as an alternative method to compare fuzzy relations.

Trying to find the closest E to R is very expensive. Indeed, if n is the cardinality of the universe X, the transitivity of T-indistinguishability operators can be modeled by $3\binom{n}{3}$ inequalities and they lay in the region of the $\binom{n}{2}$-dimensional space defined by them. The calculation of E becomes then a non-linear programming problem. Therefore, simpler methods to find a close E to R are desirable.

There are several algorithms to find the transitive closure $\overline{R}$ of a proximity relation R and it is well known that $\overline{R} \geq R$. There are also algorithms to find maximal T-indistinguishability operators B among the set of T-indistinguishability operators smaller or equal than R and also the Representation Theorem gives a T-indistinguishability operator $\underline{R}$ smaller or equal than R. It appears reasonable to aggregate $\overline{R}$ and B or $\underline{R}$ to obtain a new T-indistinguishability operator closer to R than $\overline{R}$, B or $\underline{R}$. This idea will be developed in Section 3.

If E is a T-indistinguishability operator, then the powers $E^{(p)}$ $p > 0$ are T-indistinguishability operators as well. This allows us to increase or decrease the values of E, since $E^{(p)} \leq E^{(q)}$ for $p \geq q$. So, we can decrease the values of the transitive closure or increase the ones of an operators smaller than R to find better approximations of it. Section 4 is devoted to this idea.

2 Preliminaries

This Section contains some results on t-norms and indistinguishability operators that will be needed later on in the paper. Besides well known definitions and theorems, the power T^n of a t-norm is generalized to irrational exponents in Definition 2.2. and given explicitly for continuous archimedean t-norms in Proposition 2.3.

Though many results remain valid for arbitrary t-

norms and especially for left continuous ones, for the sake of simplicity we will assume continuity for the t-norms throughout the paper.

Since a t-norm T is associative, we can extend it to an n-ary operation in the standard way:

$$T(x) = x$$
$$T(x_1, x_2, ...x_n) = T(x_1, T(x_2, ..x_n)).$$

In particular, $T(\overbrace{x, x, ...x}^{n\ times})$ will be denoted by $x_T^{(n)}$ or simply by $x^{(n)}$ if the t-norm is clear.

If T is continuous, the n-th root $x_T^{(\frac{1}{n})}$ of x wrt T is defined by

$$x_T^{(\frac{1}{n})} = sup\{z \in [0,1] \mid z_T^{(n)} \leq x\}$$

and for $m, n \in N$, $x_T^{(\frac{m}{n})} = \left(x_T^{(\frac{1}{n})}\right)_T^{(m)}$.

Lemma 2.1. *[8] If $k, m, n \in N$, $k, n \neq 0$ then $x_T^{(\frac{km}{kn})} = x_T^{(\frac{m}{n})}$.*

Assuming continuity for the t-norm T, the powers $x_T^{(\frac{m}{n})}$ can be extended to irrational exponents in a straightforward way.

Definition 2.2. *If $r \in R^+$ is a positive real number, let $\{a_n\}_{n\in N}$ be a sequence of rational numbers with $lim_{n\to\infty} a_n = r$. For any $x \in [0,1]$, the power $x_T^{(r)}$ is*

$$x_T^{(r)} = lim_{n\to\infty} x_T^{(a_n)}.$$

Continuity assures the existence of last limit and independence of the sequence $\{a_n\}_{n\in N}$.

Proposition 2.3. *Let T be an archimedean t-norm with additive generator t, $x \in [0,1]$ and $r \in R^+$. Then*

$$x_T^{(r)} = t^{[-1]}(rt(x)).$$

Proof. Due to continuity of t we need to prove it only for rational r.

If r is a natural number m, then trivially $x_T^{(m)} = t^{[-1]}(mt(x))$.

If $r = \frac{1}{n}$ with $n \in N$, then $x_T^{(\frac{1}{n})} = z$ with $z_T^{(n)} = x$ or $t^{[-1]}(nt(z)) = x$ and $x_T^{(\frac{1}{n})} = t^{[-1]}\left(\frac{t(x)}{n}\right)$.

For a rational number $\frac{m}{n}$,

$$x_T^{(\frac{m}{n})} = \left(x_T^{(\frac{1}{n})}\right)_T^{(m)} = t^{[-1]}\left(mt\left(x_T^{(\frac{1}{n})}\right)\right) =$$
$$t^{[-1]}\left(mt\left(t^{[-1]}\left(\frac{t(x)}{n}\right)\right)\right) = t^{[-1]}\left(\frac{m}{n}t(x)\right).$$

□

Definition 2.4. *The residuation $\vec{T}$ of a t-norm T is defined by*

$$\vec{T}(x|y) = sup\{\alpha \in [0,1] \mid T(x,\alpha) \leq y\}.$$

Definition 2.5. *The natural T-indistinguishability E_T associated to a given t-norm T is the fuzzy relation on [0,1] defined by*

$$E_T(x,y) = T(\vec{T}(x|y), \vec{T}(y|x)).$$

E_T is indeed a special kind of T-indistinguishability operator (Definition 2.6) [2] and in a logical context where T plays the role of the conjunction, E_T is interpreted as the bi-implication associated to T [5].

Definition 2.6. *Given a t-norm T, a T-indistinguishability operator E on a set X is a fuzzy relation $E : X \times X \to [0,1]$ satisfying for all $x, y, z \in X$*

1. $E(x,x) = 1$ *(Reflexivity)*
2. $E(x,y) = E(y,x)$ *(Symmetry)*
3. $T(E(x,y), E(y,z)) \leq E(x,z)$ *(T-transitivity).*

Example 2.7.

1. *If T is the Lukasiewicz t-norm, then $E_T(x,y) = 1 - |x-y|$ for all $x, y \in [0,1]$.*
2. *If T is the Product t-norm, then $E_T(x,y) = Min(\frac{x}{y}, \frac{y}{x})$ for all $x, y \in [0,1]$ where $\frac{z}{0} = 1$.*
3. *If T is the Minimum t-norm, then $E_T(x,y) = \begin{cases} Min(x,y) & \text{if } x \neq y \\ 1 & \text{otherwise.} \end{cases}$*

Theorem 2.8. *Representation Theorem [11]. Let R be a fuzzy relation on a set X and T a continuous t-norm. R is a T-indistinguishability operator if and only if there exists a family $(h_i)_{i\in I}$ of fuzzy subsets of X such that for all $x, y \in X$*

$$R(x,y) = inf_{i\in I} E_T(h_i(x), h_i(y)).$$

$(h_i)_{i\in I}$ is called a generating family of R.

In particular, given a proximity matrix or relation R on X (i.e. a reflexive and symmetric fuzzy relation), we can build the T-indistinguishability operator $\underline{R}$ generated by the set of the columns of R (i.e. the fuzzy subsets $R(x,\cdot)$, $x \in X$).

Proposition 2.9. $\underline{R} \leq R$.

Definition 2.10. *Let R be a proximity matrix or relation (i.e. a reflexive and symmetric fuzzy relation) on X and T a continuous t-norm. The T-transitive closure $\overline{R}$ of R is the smallest T-indistinguishability operator on X satisfying $R \leq \overline{R}$.*

Definition 2.11. *Let R and S be two fuzzy relations on X and T a continuous t-norm. The Sup-T product of R and S is the fuzzy relation $R \circ S$ on X defined for all $x, y \in X$ by*

$$(R \circ S)(x, y) = sup_{z \in X} T(R(x, z), S(z, y)).$$

Since the Sup-T product is associative or continuous t-norms, we can define for $n \in N$ the n^{th} power R_T^n of a fuzzy relation R:

$$R_T^n = \overbrace{R \circ ... \circ R}^{n \text{ times}}.$$

Definition 2.12. *Let R be a fuzzy relation on a set X and T a continuous t-norm. The transitive closure of R with respect to T is the fuzzy relation*

$$R_T = sup_{n \in N} R_T^n.$$

Proposition 2.13. *Let R be a proximity relation on a finite set X of cardinality n. Then*

$$R^T = sup_{s \in \{1,...,n-1\}} R_T^s.$$

3 Aggregating the transitive closure and a T-indistinguishability smaller than R

Given a proximity relation R on X, it is necessary in many cases to replace it by a T-indistinguishability operator E, since T-transitivity is required. In these cases, we want to find E close to R, where the closeness or similarity between fuzzy relations can be defined in many different ways.

Let X be a finite set of cardinality n. Ordering its elements linearly, we can view the fuzzy subsets of X as vectors: $X = \{x_1, ..., x_n\}$ and a fuzzy set h is the vector $(h(x_1), ..., h(x_n))$. A proximity relation R on X can be represented by a matrix (also called R) determined by the $\binom{n}{2}$ entries r_{ij} $1 \leq i < j \leq n$ of R above the diagonal.

Proposition 3.1. *Let $E = (e_{ij})_{i,j=1,...,n}$ be a proximity matrix on a set X of cardinality n and T a continuous archimedean t-norm with additive generator t. E is a T-indistinguishability operator if and only if for all i, j, k $1 \leq i < j < k \leq n$*

$$t(e_{ij}) + t(e_{jk}) \geq t(e_{ik})$$
$$t(e_{ij}) + t(e_{ik}) \geq t(e_{jk})$$
$$t(e_{ik}) + t(e_{jk}) \geq t(e_{ij})$$

Example 3.2. *If T is the Lukasiewicz t-norm, then we can take $t(x) = 1 - x$ and last inequalities become*

$$e_{ij} + e_{jk} - e_{ik} \leq 1$$
$$e_{ij} + e_{ik} - e_{jk} \leq 1$$
$$e_{ik} + e_{jk} - e_{ij} \leq 1$$

Example 3.3. *If T is the Product t-norm, then we can take $t(x) = -log(x)$ and last inequalities become*

$$e_{ij} \cdot e_{jk} \leq e_{ik}$$
$$e_{ij} \cdot e_{ik} \leq e_{jk}$$
$$e_{ik} \cdot e_{jk} \leq e_{ij}$$

Given a proximity matrix R, we must then search for (one of) the closest matrices E satisfying the last $3\binom{n}{3}$ inequalities which is a non-linear programming problem.

Instead of this, we propose alternative methods to obtain not the best but reasonably good approximations of proximity relations by T-indistinguishability operators.

Definition 3.4. [1], [8] *Given a continuous monotonic map $t : [0, 1] \to [-\infty, \infty]$ and p, q positive integers with $p + q = 1$, the weighted quasi-arithmetic mean $m_t^{p,q}$ generated by t and weights p and q is defined for all $x, y \in [0, 1]$ by*

$$m_t^{p,q}(x, y) = t^{-1}\left(p \cdot t(x) + q \cdot t(y)\right).$$

m_t is continuous if and only if Ran $t \neq [-\infty, \infty]$.

Proposition 3.5. *Fixed the weights p and q, the map assigning to every continuous Archimedean t-norm T with generator t the weighted mean $m_t^{p,q}$ generated by t is a bijection between the set of continuous Archimedean t-norms and the set of continuous quasi-arithmetic means with these weights.*

Proposition 3.6. *Let T be a continuous archimedean t-norm with additive generator t, $p \in [0, 1]$ and E, F two T-indistinguishability operator on X. The weighted quasi-arithmetic mean $m_t^{p,1-p}$ with weights p and $1 - p$ of E and F is a T-indistinguishability operator.*

Thanks to this last proposition, given a proximity matrix R we can calculate its transitive closure $\overline{R}$ and a smaller T-indistinguishability operator than R, for example $\underline{R}$ and find the weights $p, 1 - p$ to obtain the closest average of $\overline{R}$ and $\underline{R}$ to R.

The similarity between two fuzzy relations on X will be calculated in the following way.

Definition 3.7. *Let T be a continuous archimedean t-norm with additive generator t and R, S two fuzzy*

relations on a finite set X of cardinality n. The degree $DS(R,S)$ of similarity or closeness between R and S is defined by

$$DS(R,S) = t^{-1}\left(\frac{\sum_{1\le i,j\le n}|t(r_{ij}) - t(s_{ij})|}{n}\right).$$

Proposition 3.8. *DS is a T-indistinguishability operator on the set of fuzzy relations on X.*

Corollary 3.9. *Let $R = (r_{ij})$ be a proximity matrix on a finite set X of cardinality n, T a continuous archimedean t-norm with additive generator t, $\overline{R} = (\overline{r}_{i,j})$ its transitive closure, $\underline{R} = (\underline{r}_{ij})$ the T-indistinguishability operator obtained from R with the Representation Theorem, $p \in [0,1]$ and $m_t^{p,1-p}(\overline{R},\underline{R})$ the T-indistinguishability operator quasi-arithmetic mean of $\overline{R}$ and $\underline{R}$ with weights p and $1-p$. Then*

$$DS(R, m_t^{p,1-p}(E,F)) =$$

$$t^{-1}\left(\frac{\sum_{1\le i,j\le n}\left|p\cdot t(\overline{r}_{ij}) + (1-p)\cdot t(\underline{r}_{ij}) - t(r_{ij})\right|}{n}\right).$$

We are looking for the value (or values) of p that maximize the last equality. Since t^{-1} is a decreasing map, this is equivalent to minimize

$$\sum_{1\le i,j\le n}\left|p\cdot t(\overline{r}_{ij}) + (1-p)\cdot t(\underline{r}_{ij}) - t(r_{ij})\right|$$

and, since R is reflexive and symmetric, is equivalent to minimize

$$f(p) = \sum_{1\le i<j\le n}\left|p\cdot t(\overline{r}_{ij}) + (1-p)\cdot t(\underline{r}_{ij}) - t(r_{ij})\right|$$

Proposition 3.10. *Let $f_1,...,f_n : [0,1] \to R$ be n concave functions. Then $\sum_1^n f_i$ is a concave function.*

Proof. By definition, given two points x_1, x_2 of $[0,1]$, the segments joining their images by f_i $i = 1,..,n$ are above f_i. $\sum_1^n f_i$ will then be bellow the sum of all the segments. □

Corollary 3.11. *$f(p)$ is a concave function.*

Proof. Each summand $\left|p\cdot t(\overline{r}_{ij}) + (1-p)\cdot t(\underline{r}_{ij}) - t(r_{ij})\right|$ of f is a concave function. □

Proposition 3.12. *The set of minima of $f(p)$ consists of a single point or of a closed interval.*

Proof. f is a concave function and its graphic is a polygonal line. □

Proposition 3.13. *The computation of $m_t(\overline{R},\underline{R})$ with maximum $DS(R, m_t(\overline{R},\underline{R}))$ can be done taking $O(n^3)$ time complexity.*

Proof:

The computation of $\overline{R}$ and $\underline{R}$ can be done in $O(n^3)$ complexity time [9].

The addition (aggregation of distances) takes $O(n^2)$ time complexity.

The minimization of f(p) takes at most $O(n^2)$ time complexity.

So the most complex part of this process is the computation of $\overline{R}$ and $\underline{R}$, which still takes $O(n^3)$ complexity time.

Example 3.14. *Let X be a set of cardinality 7 and R the proximity relation given by*

$$R = \begin{pmatrix} 1 & 1 & 0.3 & 0.3 & 0.1 & 0.3 & 0.4 \\ 1 & 1 & 0.6 & 0.4 & 0.5 & 0.4 & 0.2 \\ 0.3 & 0.6 & 1 & 0.1 & 0.3 & 0.2 & 0.5 \\ 0.3 & 0.4 & 0.1 & 1 & 1 & 1 & 1 \\ 0.1 & 0.5 & 0.3 & 1 & 1 & 1 & 1 \\ 0.3 & 0.4 & 0.2 & 1 & 1 & 1 & 1 \\ 0.4 & 0.2 & 0.5 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Then, for T the Lukasiewicz t-norm,

$$\overline{R} = \begin{pmatrix} 1 & 1 & 0.6 & 0.4 & 0.5 & 0.4 & 0.4 \\ 1 & 1 & 0.6 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.6 & 0.6 & 1 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.4 & 0.5 & 0.5 & 1 & 1 & 1 & 1 \\ 0.5 & 0.5 & 0.5 & 1 & 1 & 1 & 1 \\ 0.4 & 0.5 & 0.5 & 1 & 1 & 1 & 1 \\ 0.4 & 0.5 & 0.5 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and

$$\underline{R} = \begin{pmatrix} 1 & 0.6 & 0.3 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.6 & 1 & 0.3 & 0.2 & 0.1 & 0.2 & 0.2 \\ 0.3 & 0.3 & 1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.1 & 1 & 0.8 & 0.9 & 0.6 \\ 0.1 & 0.1 & 0.1 & 0.8 & 1 & 0.8 & 0.7 \\ 0.1 & 0.2 & 0.1 & 0.9 & 0.8 & 1 & 0.7 \\ 0.1 & 0.2 & 0.1 & 0.6 & 0.7 & 0.7 & 1 \end{pmatrix}.$$

$$f(p) = |0.4p| + |0.3p - 0.3| + |0.3p - 0.1| + |0.4p - 0.4| +$$
$$|0.3p - 0.1| + |0.3p| + |0.3p| + |0.3p - 0.1| + |0.4p| + |0.3p - 0.1| +$$
$$|0.3p - 0.3| + |0.4p - 0.4| + |0.4p - 0.2| + |0.4p - 0.3| + |0.4p| +$$
$$|0.2p| + |0.1p| + |0.4p| + |0.2p| + |0.3p| + |0.3p|$$

which attains its minimum for $p = \frac{1}{3}$.

A good T-transitive approximation of R (for T the Lukasiewicz t-norm) is then

$$\begin{pmatrix} 1 & 0.733 & 0.4 & 0.2 & 0.233 & 0.2 & 0.2 \\ 0.733 & 1 & 0.4 & 0.3 & 0.233 & 0.3 & 0.3 \\ 0.4 & 0.4 & 1 & 0.233 & 0.233 & 0.233 & 0.233 \\ 0.2 & 0.3 & 0.233 & 1 & 0.867 & 0.933 & 0.733 \\ 0.233 & 0.233 & 0.233 & 0.867 & 1 & 0.867 & 0.8 \\ 0.2 & 0.3 & 0.233 & 0.933 & 0.867 & 1 & 0.8 \\ 0.2 & 0.3 & 0.233 & 0.733 & 0.8 & 0.8 & 1 \end{pmatrix}.$$

Example 3.15. *Let X be a set of cardinality 7 and R the proximity relation given by*

$$R = \begin{pmatrix} 1 & 0.5 & 0.7 & 0.7 & 0.5 & 0.7 & 0.8 \\ 0.5 & 1 & 1 & 0.8 & 0.9 & 0.8 & 0.6 \\ 0.7 & 1 & 1 & 0.5 & 0.7 & 0.6 & 0.9 \\ 0.7 & 0.8 & 0.5 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.9 & 0.7 & 0.5 & 1 & 0.5 & 0.5 \\ 0.7 & 0.8 & 0.6 & 0.5 & 0.5 & 1 & 0.5 \\ 0.8 & 0.6 & 0.9 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix}.$$

Then, for T the Product t-norm,

$$\overline{R} = \begin{pmatrix} 1 & 0.7 & 0.72 & 0.7 & 0.5 & 0.7 & 0.8 \\ 0.7 & 1 & 1 & 0.8 & 0.9 & 0.8 & 0.9 \\ 0.72 & 1 & 1 & 0.8 & 0.9 & 0.8 & 0.9 \\ 0.7 & 0.8 & 0.8 & 1 & 0.72 & 0.64 & 0.56 \\ 0.5 & 0.9 & 0.9 & 0.72 & 1 & 0.72 & 0.63 \\ 0.7 & 0.8 & 0.8 & 0.64 & 0.72 & 1 & 0.56 \\ 0.8 & 0.9 & 0.9 & 0.56 & 0.63 & 0.56 & 1 \end{pmatrix}$$

and

$$\underline{R} = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.625 & 0.5 & 0.625 & 0.714 \\ 0.5 & 1 & 0.625 & 0.5 & 0.625 & 0.555 & 0.555 \\ 0.5 & 0.625 & 1 & 0.5 & 0.555 & 0.555 & 0.6 \\ 0.625 & 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.625 & 0.555 & 0.5 & 1 & 0.5 & 0.5 \\ 0.625 & 0.555 & 0.555 & 0.5 & 0.5 & 1 & 0.5 \\ 0.714 & 0.555 & 0.6 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix}.$$

$f(p)$ attains its minimum for $p = 0.521$.

A good T-transitive approximation of R (for T the Product t-norm) is then

$$\begin{pmatrix} 1 & 0.587 & 0.595 & 0.660 & 0.5 & 0.660 & 0.754 \\ 0.587 & 1 & 0.783 & 0.626 & 0.744 & 0.662 & 0.700 \\ 0.595 & 0.783 & 1 & 0.626 & 0.700 & 0.662 & 0.729 \\ 0.660 & 0.626 & 0.626 & 1 & 0.595 & 0.563 & 0.528 \\ 0.5 & 0.744 & 0.700 & 0.595 & 1 & 0.595 & 0.559 \\ 0.660 & 0.662 & 0.662 & 0.563 & 0.595 & 1 & 0.528 \\ 0.754 & 0.700 & 0.729 & 0.528 & 0.559 & 0.528 & 1 \end{pmatrix}.$$

The degree of closeness between two fuzzy relations can also be calculated using the euclidean distance.

Definition 3.16. *Let $R = (r_{ij})$ and $S = (s_{ij})$ be two fuzzy relations on a finite set X of cardinality n. The euclidean distance D between R and S is*

$$D(R,S) = \left(\sum_{1 \le i,j \le n} (r_{ij} - s_{ij})^2 \right)^{\frac{1}{2}}$$

Corollary 3.17. *Let $R = (r_{ij})$ be a proximity matrix on a finite set X of cardinality n, T a continuous archimedean t-norm with additive generator t, $\overline{R} = (\overline{r}_{i,j})$ its transitive closure, $\underline{R} = (\underline{r}_{ij})$ the T-indistinguishability operator obtained from R with the Representation Theorem, $p \in [0,1]$ and $m_t(\overline{R}, \underline{R})$ the T-indistinguishability operator quasi-arithmetic mean of $\overline{R}$ and $\underline{R}$ with weights p and $1-p$. Then*

$$D(R, m_t(E,F)) =$$

$$\left(\sum_{1 \le i,j \le n} \left(t^{-1} \left(p \cdot t\left(\overline{r}_{ij}\right) + (1-p) \cdot t\left(\underline{r}_{ij}\right) \right) - t(r_{ij}) \right)^2 \right)^{\frac{1}{2}}.$$

Proposition 3.18. *Let T be the Lukasiewicz t-norm and R a proximity on a set X of cardinality n. The closest $m_t(\overline{R}, \underline{R})$ to R is attained for*

$$p = \frac{\sum_{1 \le i < j \le n} \left(\overline{r}_{ij} - \underline{r}_{ij}\right)\left(r_{ij} - \underline{r}_{ij}\right)}{\sum_{1 \le i < j \le n} \left(\overline{r}_{ij} - \underline{r}_{ij}\right)^2}$$

Proof. Due to symmetry and reflexivity, it is enough to minimize

$$f(p) = \sum_{1 \le i < j \le n} \left(p\left(\overline{r}_{ij} - \underline{r}_{ij}\right) + \underline{r}_{ij} - r_{ij} \right)^2.$$

$$f'(p) = 2 \sum_{1 \le i < j \le n} \left(p\left(\overline{r}_{ij} - \underline{r}_{ij}\right) + \underline{r}_{ij} - r_{ij} \right)\left(\underline{r}_{ij} - r_{ij}\right) = 0$$

and

$$p = \frac{\sum_{1 \le i < j \le n} \left(\overline{r}_{ij} - \underline{r}_{ij}\right)\left(r_{ij} - \underline{r}_{ij}\right)}{\sum_{1 \le i < j \le n} \left(\overline{r}_{ij} - \underline{r}_{ij}\right)^2}.$$

□

Example 3.19. *Let X be a set of cardinality 4 and R the proximity relation on X given by*

$$R = \begin{pmatrix} 1 & 0.8 & 0.2 & 0.4 \\ 0.8 & 1 & 0.7 & 0.1 \\ 0.2 & 0.7 & 1 & 0.6 \\ 0.4 & 0.1 & 0.6 & 1 \end{pmatrix}.$$

If T is the Lukasiewicz t-norm, the closest T-indistinguishability operator of the type $m_t(\overline{R}, \underline{R})$ (with respect to the euclidean distance) is attained for $p = 0.6388889$.

A good T-approximation of R is then

$$\begin{pmatrix} 1 & 0.6917 & 0.3917 & 0.3639 \\ 0.6917 & 1 & 0.5917 & 0.2278 \\ 0.3917 & 0.5917 & 1 & 0.5278 \\ 0.3639 & 0.2278 & 0.5278 & 1 \end{pmatrix}$$

4 Applying a homotecy to a T-indistinguishability operator

In this Section, the fact that the power of a T-indistinguishability operator is again a T-indistinguishability operator will be exploited to modify the entries of $\overline{R}$ or $\underline{R}$ to find a better approximation of R.

Proposition 4.1. *Let T be a continuous t-norm, E a T-indistinguishability operator on X and $p > 0$. Then $E^{(p)}$ is a T-indistinguishability operator.*

Example 4.2.

- *If T is a continuous archimedean t-norm with additive generator t and E a T-indistinguishability operator, then $t^{[-1]}(p \cdot t(E))$ is a T-indistinguishability operator.*
- *If T is the Lukasiewicz t-norm and E a T-indistinguishability operator, then $Max(0, 1-p+p \cdot E)$ is a T-indistinguishability operator.*
- *If T is the Product t-norm and E a T-indistinguishability operator, then E^p is a T-indistinguishability operator.*

Let $R = (r_{ij})$ be a proximity matrix on a set X of cardinality X, $p > 0$ and $E = (e_{ij})$ a T-indistinguishability operator on X with T a continuous archimedean t-norm with additive generator t. Then

$$DS(R, E^{(p)}) = t^{-1}\left(\frac{\sum_{1\le i,j\le n} |t(r_{ij}) - p \cdot t(e_{ij}))|}{n}\right).$$

To maximize the previous expression is equivalent to minimize

$$\sum_{1\le i,j\le n} |t(r_{ij}) - p \cdot t(e_{ij}))|.$$

Since R is reflexive and symmetric, this is equivalent to minimize

$$g(p) = \sum_{1\le i<j\le n} |t(r_{ij}) - p \cdot t(e_{ij}))|.$$

Again g is a sum of concave functions in $[0, 1]$ and therefore has a minimum or a close interval of minima.

Example 4.3. *Let us consider the same matrix*

$$R = \begin{pmatrix} 1 & 0.8 & 0.2 & 0.4 \\ 0.8 & 1 & 0.7 & 0.1 \\ 0.2 & 0.7 & 1 & 0.6 \\ 0.4 & 0.1 & 0.6 & 1 \end{pmatrix}.$$

Then, for T the Lukasiewicz t-norm,

$$\underline{R} = \begin{pmatrix} 1 & 0.5 & 0.2 & 0.3 \\ 0.5 & 1 & 0.4 & 0.1 \\ 0.2 & 0.4 & 1 & 0.4 \\ 0.3 & 0.1 & 0.4 & 1 \end{pmatrix}.$$

$$g(p) = |0.5 \cdot p - 0.2| + |0.8 \cdot p - 0.8| + |0.7 \cdot p - 0.6| + |0.6 \cdot p - 0.3| + |0.9 \cdot p - 0.9| + |0.6 \cdot p - 0.4|$$

which attains its minimum for $p = 0.857$.

A good approximation of R is then

$$\underline{R}^{(0.857)} = \begin{pmatrix} 1 & 0.5715 & 0.3144 & 0.4001 \\ 0.5715 & 1 & 0.4858 & 0.2287 \\ 0.3144 & 0.4858 & 1 & 0.4858 \\ 0.4001 & 0.2287 & 0.4858 & 1 \end{pmatrix}.$$

If we consider the euclidean distance between R and the power $E^{(p)}$ of a T-indistinguishability operator $E = (e_{ij})$, then

Proposition 4.4.

$$D(R, E^{(p)}) = \left(\sum_{1\le i,j\le n} \left(t^{-1}(p \cdot t(e_{ij})) - r_{ij}\right)^2\right)^{\frac{1}{2}}.$$

Example 4.5. *Continuing the last example, $D(R, \overline{R}^{(p)})$ is maximum for $p = 1.208633$ and $D(R, \underline{R}^{(p)})$ is maximum for $p = 0.821306$.*

Good approximations of R are therefore

$$\overline{R}^{(1.208633)} = \begin{pmatrix} 1 & 0.7583 & 0.3957 & 0.2748 \\ 0.7583 & 1 & 0.6374 & 0.1540 \\ 0.3957 & 0.6374 & 1 & 0.5165 \\ 0.2748 & 0.1540 & 0.5165 & 1 \end{pmatrix}.$$

and

$$\underline{R}^{(0.821306)} = \begin{pmatrix} 1 & 0.8357 & 0.5893 & 0.5072 \\ 0.8357 & 1 & 0.7536 & 0.4251 \\ 0.5893 & 0.7536 & 1 & 0.6715 \\ 0.5072 & 0.4251 & 0.6715 & 1 \end{pmatrix}.$$

5 Concluding Remarks

In this paper we have presented two ways to find good approximations of a proximity relation by T-transitive ones (T archimedean) in a reasonable computational way.

The obtained approximation R' is in general not comparable with R in the sense that neither $R' \geq R$ nor $R \geq R'$ must hold.

The simple examples show that in general these approximations are better than the transitive closure or openings of the proximity R.

The methods of the paper cannot be applied to the Minimum t-norm. Other ways to obtain similar results for this t-norm are therefore needed and the authors will work on it in forthcoming papers.

Acknowledgement

Research partially supported by DGICYT projects number TIN2006-14311 and TIN2006-06190.

References

[1] J. Aczél (1966) *Lectures on functional equations and their applications*. Academic Press. New York/London.

[2] D. Boixader, J. Jacas, J. Recasens (2000). Fuzzy Equivalence Relations: Advanced Material. In Dubois, Prade Eds. *Fundamentals of Fuzzy Sets*, Kluwer, 261-290.

[3] T. Calvo, A. Kolesárova, M. Komorníková, R. Mesiar (2002). Aggregation Operators: Properties, Classes and Construction Methods. In Mesiar, Calvo, Mayor Eds. *Aggregation Operators: New Trends and Applications*. Studies in Fuzziness and Soft Computing. Springer, 3-104.

[4] J. Fodor, M. Roubens (1995). Structure of transitive valued binary relations. *Math. Social Sci.* 30, 71–94.

[5] P. Hájek (1998) *Metamathematics of Fuzzy Logic*. Kluwer. Dordrecht.

[6] J. Jacas, J. Recasens (2003) Aggregation of T-Transitive Relations. *Int J. of Intelligent Systems* 18, 1193-1214.

[7] C. M. Ling (1965) Representation of associative functions *Publ. Math. Debrecen* 12, 189-212.

[8] E. P. Klement, R. Mesiar, E. Pap (2000). *Triangular norms*. Kluwer. Dordrecht.

[9] Naessens, H., De Meyer, H., De Baets, B., Algorithms for the Computation of T-Transitive Closures, IEEE Trans Fuzzy Systems 10 :4, 2002, 541-551.

[10] B. Schweizer, A. Sklar (1983) *Probabilistic Metric Spaces*. North-Holland. Amsterdam.

[11] L.Valverde (1985). On the structure of F-indistinguishability operators, *Fuzzy Sets and Systems* 17, 313-328.

[12] L.A.Zadeh (1971). Similarity relations and fuzzy orderings, *Information Science* 3, 177-200.

Preservation of graded properties of fuzzy relations by aggregation functions

Urszula Dudziak
Institute of Mathematics, University of Rzeszów,
35-310 Rzeszów, ul. Rejtana 16a, Poland.
e-mail: ududziak@univ.rzeszow.pl

Abstract

We study graded properties (α–properties) of fuzzy relations, which are parameterized versions of fuzzy relation properties defined by L.A. Zadeh. Namely, we take into account fuzzy relations which are: α–reflexive, α–irreflexive, α–symmetric, α–antisymmetric, α–asymmetric, α–connected, α–transitive, where $\alpha \in [0,1]$. We also pay our attention to the composed versions of these basic properties, e.g. an α–equivalence, α–orders. Using functions of n variables we consider aggregation fuzzy relation of given fuzzy relations. We give conditions for functions to preserve graded properties of fuzzy relations.

Keywords: aggregation function, fuzzy relation properties.

1 Introduction

Aggregations of fuzzy relations are important in group choice theory and multiple–criteria decision making so it is interesting to know which functions preserve properties of fuzzy relations during aggregation process. There are many works devoted to this topic (e.g. [5, 8, 9]). We consider graded properties of fuzzy relations during aggregations of finite families of these relations. These graded properties of fuzzy relations are parameterized versions of the well-known properties of fuzzy relations and they are generalizations of these properties. In order to receive more general results, we concentrate on a fixed property of fuzzy relations and give conditions on operations in $[0,1]$ (not only aggregation functions) to preserve this property in aggregation process. Next, we put examples of aggregation functions which fulfil the respective conditions. Presented results are the continuation of the ones obtained in [3, 4].

First, in Section 2 we recall definitions which will be used in the sequel. Next, we discuss the graded properties of fuzzy relations (Section 3), and we examine properties such as: reflexivity and irreflexivity, asymmetry and connectedness, symmetry and transitivity (Section 4), composed properties (Section 5).

2 Preliminaries

Now we recall some definitions which will be used in our considerations.

Definition 1 ([10]). A fuzzy relation in $X \neq \emptyset$ is an arbitrary function $R : X \times X \to [0,1]$. The family of all fuzzy relations in X is denoted by FR(X).

Definition 2 (cf. [1]). Let $n \geqslant 2$. Operation $F : [0,1]^n \to [0,1]$ is called an aggregation function if it is increasing in each variable and fulfils boundary conditions $F(0,...,0) = 0, F(1,...,1) = 1$.

With the use of n–argument functions F we aggregate given fuzzy relations $R_1, ..., R_n$.

Definition 3 (cf. [5], p. 107). Let $F : [0,1]^n \to [0,1]$, $R_1, ..., R_n \in FR(X)$. An aggregation fuzzy relation $R_F \in FR(X)$ is described by the formula

$$R_F(x,y) = F(R_1(x,y), ..., R_n(x,y)), \quad x,y \in X. \quad (1)$$

A function F preserves a property of fuzzy relations if for every $R_1, ..., R_n \in FR(X)$ having this property, R_F also has this property.

Example 1. Projections $P_k(t_1, \ldots, t_n) = t_k$, $t_1, ..., t_n \in [0,1]$, $k \in \{1, ..., n\}$ are aggregation functions and they preserve each property of fuzzy relations because for $F = P_k$ we get $R_F = R_k$.

In our further considerations we will need the following

Definition 4 ([6], p. 4). Triangular norm $T : [0,1]^2 \to [0,1]$ (triangular conorm $S : [0,1]^2 \to [0,1]$) is an arbitrary associative, commutative, increasing in both variables operation having a neutral element $e = 1$ ($e = 0$).

Examples of triangular norms (t-norms) and triangular conorms (t-conorms) and results concerning these binary operations, which are examples of aggregation functions, are given in [6]. In our considerations we will need some more notions and properties.

Definition 5 (cf. [9], Definition 2.5). Let m, $n \in \mathbb{N}$. Operation $F : [0,1]^m \to [0,1]$ dominates operation $G : [0,1]^n \to [0,1]$ ($F \gg G$), if for arbitrary matrix $[a_{ik}] = A \in [0,1]^{m\times n}$ we have

$$F(G(a_{11},...,a_{1n}),...,G(a_{m1},...,a_{mn})) \geqslant \quad (2)$$

$$G(F(a_{11},...,a_{m1}),...,F(a_{1n},...,a_{mn})).$$

Theorem 1 (cf. [9], Proposition 5.1). *Increasing in each variable function $F : [0,1]^n \to [0,1]$ dominates minimum if and only if for any $t_1,...,t_n \in [0,1]$*

$$F(t_1,...,t_n) = \min(f_1(t_1),...,f_n(t_n)), \quad (3)$$

where functions $f_k : [0,1] \to [0,1]$ are increasing, for $k = 1,...,n$.

Example 2. Here are examples of functions which fulfil the condition (3):
if $f_k(t) = t$, $k = 1,...,n$ then $F = \min$,
if for some $k \in \{1,...,n\}$, $f_k(t) = t$, $f_i(t) = 1$ for $i \neq k$ then $F = P_k$, (cf. Example 1),
if $f_k(t) = \max(1 - v_k, t)$, $v_k \in [0,1]$, $k = 1,...,n$, $\max\limits_{1\leqslant k\leqslant n} v_k = 1$ then F is the weighted minimum

$$F(t_1,...,t_n) = \min_{1\leqslant k\leqslant n} \max(1 - v_k, t_k), \quad (4)$$

where $t = (t_1,...,t_n) \in [0,1]^n$.

Lemma 1 (cf. [2], Lemma 1.1).

$$\underset{a,b\in[0,1]}{\forall} \underset{\alpha\in[0,1]}{\forall} (a \leqslant \alpha \Rightarrow b \leqslant \alpha) \Leftrightarrow b \leqslant a, \quad (5)$$

$$\underset{a,b\in[0,1]}{\forall} \underset{\alpha\in[0,1]}{\forall} (a \geqslant \alpha \Rightarrow b \geqslant \alpha) \Leftrightarrow a \leqslant b. \quad (6)$$

Remark 1. If *card* $X = n$, $X = \{x_1,...,x_n\}$, then $R \in FR(X)$ may be presented by a matrix: $R = [r_{ik}]$, where $r_{ik} = R(x_i, x_k)$, $i,k = 1,...,n$.

3 Graded properties of fuzzy relations

We will consider the following parameterized versions of fuzzy relations properties.

Definition 6 (cf. [2], p. 75). Let $\alpha \in [0,1]$. Relation $R \in FR(X)$ is:
• α–reflexive, if

$$\underset{x\in X}{\forall} R(x,x) \geqslant \alpha, \quad (7)$$

• α–irreflexive, if

$$\underset{x\in X}{\forall} R(x,x) \leqslant 1 - \alpha, \quad (8)$$

• α–symmetric, if

$$\underset{x,y\in X}{\forall} R(x,y) \geqslant 1 - \alpha \Rightarrow R(y,x) \geqslant R(x,y), \quad (9)$$

• α–asymmetric, if

$$\underset{x,y\in X}{\forall} \min(R(x,y), R(y,x)) \leqslant 1 - \alpha, \quad (10)$$

• α–antisymmetric, if

$$\underset{x,y,x\neq y\in X}{\forall} \min(R(x,y), R(y,x)) \leqslant 1 - \alpha, \quad (11)$$

• totally α–connected, if

$$\underset{x,y\in X}{\forall} \max(R(x,y), R(y,x)) \geqslant \alpha, \quad (12)$$

• α–connected, if

$$\underset{x,y,x\neq y\in X}{\forall} \max(R(x,y), R(y,x)) \geqslant \alpha, \quad (13)$$

• α–transitive, if

$$\underset{x,y,z\in X}{\forall} \min(R(x,y), R(y,z)) \geqslant 1 - \alpha \Rightarrow \quad (14)$$

$$R(x,z) \geqslant \min(R(x,y), R(y,z)).$$

Example 3. Let *card* $X = 2$. Relation $R \in FR(X)$, where

$$R = \begin{bmatrix} 0.3 & 0.5 \\ 0.7 & 0.4 \end{bmatrix}$$

is α–reflexive for $\alpha \in [0, 0.3]$, α–irreflexive for $\alpha \in [0, 0.6]$, α–symmetric for $\alpha \in [0, 0.3)$, α–asymmetric and α–antisymmetric for $\alpha \in [0, 0.5]$, totally α–connected for $\alpha \in [0, 0.3]$ and α–connected for $\alpha \in [0, 0.7]$.

It is worth mentioning that in [7] a special case of α–transitivity is considered. Namely, this is the 0.5–transitivity according to condition (14). However, the problem of preservation of this property during aggregation process is not discussed.

Example 4. Let $R \equiv c$, where $c \in [0,1]$. Relation $R \in FR(X)$ is α–transitive for arbitrary $\alpha \in [0,1]$.

Remark 2. Conditions (7) - (14) for $\alpha = 1$ become properties of fuzzy relations introduced by Zadeh [11].

More results describing α–properties (7) - (8), (10) - (13) one can find in [3], where connections between the properties of fuzzy relations and its cuts properties are discussed. Here, we give such considerations only for α–symmetry and α–transitivity, which were not discussed in [3]. We recall the following

Definition 7 (cf. [10]). Let $R \in FR(X)$, $\alpha \in [0,1]$. An α–cut of a relation R is a set of the form

$$R_\alpha = \{(x,y) : R(x,y) \geqslant \alpha\} \subset X \times X. \quad (15)$$

Theorem 2. *Let $\beta \in [0,1]$. If relation $R \in FR(X)$ is β–symmetric (β–transitive), then it is α–symmetric (α–transitive) for any $\alpha \in [0,\beta]$.*

Proof. Let $x, y \in X$,$\alpha, \beta \in [0,1]$ and $\alpha \leqslant \beta$. If relation $R \in FR(X)$ is β-symmetric and $R(x,y) \geqslant 1-\alpha$, then $R(x,y) \geqslant 1-\alpha \geqslant 1-\beta$ and by β–symmetry of R we get $R(y,x) \geqslant R(x,y)$. As a result relation R is α–symmetric for arbitrary $\alpha \in [0,\beta]$. The proof for α–transitivity is analogous. □

There are the following connections between the properties of fuzzy relations and the properties of its cuts.

Theorem 3. *Let $\alpha \in [0,1]$. If relation $R \in FR(X)$ is α–symmetric, then $R_{1-\alpha}$ is symmetric.*

Proof. Let $\alpha \in [0,1]$, $R \in FR(X)$, $x, y \in X$. If R is α–symmetric, then by (9) and (15)

$$(x,y) \in R_{1-\alpha} \Rightarrow R(x,y) \geqslant 1-\alpha \Rightarrow$$

$$R(y,x) \geqslant R(x,y) \Rightarrow R(y,x) \geqslant 1-\alpha \Rightarrow (y,x) \in R_{1-\alpha}.$$

Thus the cut $R_{1-\alpha}$ is symmetric. □

The converse theorem to Theorem 3 does not hold.

Example 5. Let *card* $X = 2$, $\alpha \in [0,1]$, $R \in FR(X)$ be the one from Example 3. The cuts R_β are symmetric for $\beta \in [0, 0.5] \cup (0.7, 1]$, so the cuts $R_{1-\alpha}$ have this property for $\alpha \geqslant 0.5$ and $\alpha < 0.3$. The fuzzy relation R is α–symmetric for $\alpha \in [0, 0.3)$, so for $\alpha = 0.5$, the cut $R_{0.5}$ is symmetric, but R is not 0.5–symmetric.

Now we will consider the property of an α–transitivity.

Theorem 4. *Let $\alpha \in [0,1]$. If relation $R \in FR(X)$ is α–transitive, then $R_{1-\alpha}$ is transitive.*

Proof. Let $\alpha \in [0,1]$, $R \in FR(X)$, $x, y, z \in X$. If R is α–transitive, then by (15) and (14) we get

$$((x,y),(y,z) \in R_{1-\alpha}) \Rightarrow$$

$$(R(x,y) \geqslant 1-\alpha, R(y,z) \geqslant 1-\alpha) \Rightarrow$$

$$\min(R(x,y), R(y,z)) \geqslant 1-\alpha \Rightarrow$$

$$R(x,z) \geqslant \min(R(x,y), R(y,z)) \Rightarrow$$

$$R(x,z) \geqslant 1-\alpha \Rightarrow (x,z) \in R_{1-\alpha},$$

so relation $R_{1-\alpha}$ is transitive. □

The converse theorem to Theorem 4 does not hold.

Example 6. Let $R \in FR(X)$, *card* $X = 3$,

$$R = \begin{bmatrix} 0.7 & 0 & 0 \\ 0.8 & 0.9 & 0 \\ 0.6 & 0.9 & 0.8 \end{bmatrix}.$$

The cuts R_β are transitive for $\beta \in [0, 0.6] \cup (0.8, 1]$, so the cuts $R_{1-\alpha}$ have this property for $\alpha \in [0, 0.2) \cup [0.4, 1]$. We see that $0.8 = \min(r_{32}, r_{21}) \geqslant 1-\alpha$, for $\alpha \in [0.4, 1]$ and $\min(r_{32}, r_{21}) = 0.8 > 0.6 = r_{31}$, so relation R is not α–transitive for $\alpha \in [0.4, 1]$.

4 Preservation of graded properties

Results concerning preservation of an α–reflexivity, an α–irreflexivity, an α–asymmetry, an α–antisymmetry and both types of an α–connectedness are given in [4]. Now symmetry and transitivity will be discussed. Presented conditions will be stated, in general, for operations F in $[0,1]$ and next adequate examples of aggregation functions will be indicated. Let us notice that condition (9) may be written in a more convenient way.

Lemma 2. *Let $\alpha \in [0,1]$. Relation $R \in FR(X)$ is α–symmetric iff*

$$\underset{x,y \in X}{\forall} R(x,y) \geqslant 1-\alpha \Rightarrow R(y,x) = R(x,y). \quad (16)$$

Proof. Let $\alpha \in [0,1]$, $x, y \in X$. If R fulfils (16) and $R(x,y) \geqslant 1-\alpha$, then by (16) we get $R(y,x) = R(x,y)$, which implies (9).

If relation R fulfils (9) and $R(x,y) \geqslant 1-\alpha$, then $R(y,x) \geqslant R(x,y)$, and as a result $R(y,x) \geqslant 1-\alpha$. Applying once again (9) to the pair (y,x) we see that $R(x,y) \geqslant R(y,x)$. Thus $R(y,x) = R(x,y)$ and relation R fulfils (16). □

Lemma 3. *Let $\alpha \in [0,1]$, $t = (t_1, ..., t_n) \in [0,1]^n$. $F : [0,1]^n \to [0,1]$ fulfils*

$$F|_{[0,1]^n \setminus [1-\alpha,1]^n} < 1-\alpha \quad (17)$$

if and only if

$$\underset{t \in [0,1]^n}{\forall} F(t_1, ..., t_n) \geqslant 1-\alpha \Rightarrow \min_{1 \leqslant k \leqslant n} t_k \geqslant 1-\alpha. \quad (18)$$

Proof. Let $\alpha \in [0,1]$. We see that the condition (18) is equivalent to the following one

$$\underset{t \in [0,1]^n}{\forall} (\underset{1 \leqslant k \leqslant n}{\exists} t_k < 1-\alpha) \Rightarrow F(t_1, ..., t_n) < 1-\alpha.$$

The above condition is equivalent to (17) which finishes the proof. □

Theorem 5. *Let $\alpha \in [0,1]$. If $F : [0,1]^n \to [0,1]$ fulfils (17), then it preserves an α–symmetry of relations $R_1, ..., R_n \in FR(X)$.*

Proof. Let $\alpha \in [0,1]$, $x, y \in X$ and relations $R_1, ..., R_n \in FR(X)$ be α–symmetric. If F fulfils (17) then using notations

$$t_k = R_k(x,y),\ k = 1,2,...,n,$$

and Lemma 3 we get

$$R_F(x,y) \geqslant 1-\alpha \Rightarrow F(t) \geqslant 1-\alpha \Rightarrow \min_{1\leqslant k\leqslant n} t_k \geqslant 1-\alpha$$

$$\Rightarrow \bigwedge_{1\leqslant k\leqslant n} t_k \geqslant 1-\alpha \Rightarrow \bigwedge_{1\leqslant k\leqslant n} R_k(x,y) \geqslant 1-\alpha.$$

By α–symmetry of relations $R_1, ..., R_n$ and Lemma 2 we obtain $R_k(x,y) = R_k(y,x)$, $k = 1, ..., n$. Thus $R_F(x,y) = R_F(y,x)$ and by Lemma 2 relation R_F is α–symmetric. □

Theorem 6. *If $F : [0,1]^n \to [0,1]$ fulfils condition $F \leqslant \min$, then it preserves an α–symmetry of fuzzy relations for arbitrary $\alpha \in [0,1]$.*

Proof. By Lemma 3 a function F fulfils (17) for arbitrary $\alpha \in [0,1]$ iff

$$\bigwedge_{\alpha\in[0,1]} \bigwedge_{t\in[0,1]^n} F(t_1, ..., t_n) \geqslant 1-\alpha \Rightarrow \min_{1\leqslant k\leqslant n} t_k \geqslant 1-\alpha.$$

So by (6) this is equivalent to the fact that $F \leqslant \min$. By Theorem 5 a function F preserves an α–symmetry of fuzzy relations for any $\alpha \in [0,1]$. □

For our further considerations we need the following statement

Lemma 4. *If $F : [0,1]^n \to [0,1]$ is increasing in each variable and has a neutral element $e = 1$, i.e.*

$$\bigwedge_{t\in[0,1]} \bigwedge_{1\leqslant k\leqslant n} F(1, \ldots, 1, t, 1, \ldots, 1) = t, \qquad (19)$$

where t is in the k-th position, then $F \leqslant \min$.

Proof. Let $t_1, ..., t_n \in [0,1]$. By assumption on F we get for any $k = 1, ..., n$

$$F(t_1, ..., t_{k-1}, t_k, t_{k+1}, ..., t_n) \leqslant F(1, ..., 1, t_k, 1, ..., 1) = t_k.$$

As a result

$$F(t_1, ..., t_n) \leqslant \min_{1\leqslant k\leqslant n} t_k,$$

so $F \leqslant \min$. □

Corollary 1. *Let $n = 2$. Every t-norm preserves an α–symmetry of fuzzy relations for arbitrary $\alpha \in [0,1]$.*

Now we pay our attention to an α–transitivity.

Theorem 7. *Let $\alpha \in [0,1]$. If increasing in each variable $F : [0,1]^n \to [0,1]$ fulfils (17) and $F \gg \min$, i.e. for any $(s_1, ..., s_n), (t_1, ..., t_n) \in [0,1]^n$*

$$F(\min(s_1,t_1), ..., \min(s_n,t_n)) \geqslant \qquad (20)$$

$$\min(F(s_1, \ldots, s_n), F(t_1, \ldots, t_n)),$$

then it preserves an α–transitivity of fuzzy relations.

Proof. Let $\alpha \in [0,1]$, $x, y, z \in X$. If relations $R_1, ..., R_n \in FR(X)$ are α–transitive and F fulfils (17) then using notations

$$s_k = R_k(x,y), \quad t_k = R_k(y,z), \quad k = 1, ..., n, \quad (21)$$

applying Lemma 3 and by the properties of the minimum we have

$$\min(R_F(x,y), R_F(y,z)) \geqslant 1-\alpha \ \Rightarrow$$

$$\min(F(s_1, ..., s_n), F(t_1, ..., t_n)) \geqslant 1-\alpha \ \Rightarrow$$

$$(F(s_1, ..., s_n) \geqslant 1-\alpha,\ F(t_1, ..., t_n) \geqslant 1-\alpha) \ \Rightarrow$$

$$(\min_{1\leqslant k\leqslant n} s_k \geqslant 1-\alpha,\ \min_{1\leqslant k\leqslant n} t_k \geqslant 1-\alpha) \ \Rightarrow$$

$$\bigwedge_{1\leqslant k\leqslant n} \min(s_k, t_k) \geqslant 1-\alpha \ \Rightarrow$$

$$\bigwedge_{1\leqslant k\leqslant n} \min(R_k(x,y), R_k(y,z)) \geqslant 1-\alpha.$$

So by α–transitivity of relations R_k we see that $R_k(x,z) \geqslant \min(R_k(x,y), R_k(y,z))$, $k = 1, ..., n$. Thus by (20), monotonicity of F and by notations (21) we have

$$\min(R_F(x,y), R_F(y,z)) =$$

$$\min(F(R_1(x,y), ..., R_n(x,y)), F(R_1(y,z), ..., R_n(y,z)))$$

$$= \min(F(s_1, ..., s_n), F(t_1, ..., t_n)) \leqslant$$

$$F(\min(s_1,t_1), ..., \min(s_n,t_n)) =$$

$$F(\min(R_1(x,y), R_1(y,z)), ..., \min(R_n(x,y), R_n(y,z)))$$

$$\leqslant F(R_1(x,z), ..., R_n(x,z)) = R_F(x,z)$$

which proves an α–transitivity of R_F. □

Using Lemma 3, similarly like for α–symmetry we receive

Theorem 8. *If increasing in each variable function $F : [0,1]^n \to [0,1]$ fulfils (20) and $F \leqslant \min$, then it preserves an α–transitivity of fuzzy relations for any $\alpha \in [0,1]$.*

Looking for functions which fulfil (20) (cf. Theorem 1, Example 2) and $F \leqslant \min$ (cf. Lemma 4) we see that the minimum, which is an example of an aggregation function, fulfils these conditions.

Corollary 2. *The minimum preserves an α–transitivity of fuzzy relations for any $\alpha \in [0,1]$.*

5 Preservation of composed graded properties

Now we will consider composed α–properties. These are such properties which are conjunctions of properties discussed in the previous sections.

Definition 8 (cf. [2], p. 77). Let $\alpha \in [0,1]$. Relation $R \in FR(X)$ is:
• *an α–tolerance*, if it is α–reflexive and α–symmetric,
• *an α–tournament*, if it is α–asymmetric and α–connected,
• *an α–equivalence*, if it is α–reflexive, α–symmetric and α–transitive,
• *a quasi–α–order*, if it is α–reflexive and α–transitive,
• *a partial α–order*, if it is α–reflexive, α–antisymmetric and α–transitive,
• *a linear quasi–α–order*, if it is an α–connected quasi–α–order,
• *a strict α–order*, if it is α–irreflexive and α–transitive,
• *a linear strict α–order*, if it is α–asymmetric, α–transitive and α–connected,
• *a linear α–order*, if it is an α–connected partial α–order.

Example 7. Let *card* $X = 2$. Relation $R \in FR(X)$, where

$$R = \begin{bmatrix} 0.5 & 0.6 \\ 0.2 & 0.5 \end{bmatrix}$$

is 0.5–asymmetric (0.5–antisymmetric) and totally 0.5–connected (0.5–connected), so it is 0.5–tournament.

Applying some results of [4] we will present functions $F : [0,1]^n \to [0,1]$ preserving composed properties of fuzzy relations. We will concentrate on the results concerning preservation of composed graded properties of fuzzy relations for any $\alpha \in [0,1]$. Although in [4] are given characterizations, the next theorems are, in main cases, only sufficient conditions, because statements for an α–symmetry and an α–transitivity are only sufficient conditions. Moreover, it will turn out that, in many cases, the described function is an n–argument minimum in $[0,1]$.

Theorem 9. *If $F = \min$, then it preserves an α–tolerance for any $\alpha \in [0,1]$.*

Proof. By [4] (Theorem 2) F preserves an α–reflexivity for any $\alpha \in [0,1]$ if and only if $F \geqslant \min$. By Theorem 6, if $F \leqslant \min$, then it preserves an α–symmetry for any $\alpha \in [0,1]$. Taking into account the definition of an α–tolerance we see that $F = \min$ preserves an α–tolerance for any $\alpha \in [0,1]$. □

Theorem 10. *Let card $X \geqslant 2$. $F : [0,1]^n \to [0,1]$ preserves an α–tournament for any $\alpha \in [0,1]$ if and only if it fulfils*

$$\bigvee_{s,t \in [0,1]^n} \min(F(s), F(t)) \leqslant \max_{1 \leqslant k \leqslant n} \min(s_k, t_k), \quad (22)$$

$$\bigvee_{s,t \in [0,1]^n} \max(F(s), F(t)) \geqslant \min_{1 \leqslant k \leqslant n} \max(s_k, t_k). \quad (23)$$

Proof. As a direct consequence of [4] (Theorems 6, 8) and definition of an α–tournament we see that F preserves this property for any $\alpha \in [0,1]$ iff it fulfils (22) (preservation of an α–asymmetry for any $\alpha \in [0,1]$) and (23) (preservation of an α–connectedness for any $\alpha \in [0,1]$). □

Theorem 11. *If $F = \min$, then it preserves an α–equivalence (a quasi–α–order, a partial α–order) for any $\alpha \in [0,1]$.*

Proof. An α–equivalence is an α–transitive relation of an α–tolerance, so by Theorem 9 and Corollary 2 we see that $F = \min$ preserves an α–equivalence for any $\alpha \in [0,1]$.

By [4] (Theorem 2), F preserves an α–reflexivity for any $\alpha \in [0,1]$ if and only if $F \geqslant \min$. By Theorem 8, Corollary 2 and by the definition of a quasi α–order we see that $F = \min$ preserves this property for any $\alpha \in [0,1]$.

A partial α–order is an α–antisymmteric relation of a quasi–α–order, so by the above results and by the fact that $F = \min$ fulfils (22), which by [4] (Theorem 6) means that it preserves an α–antisymmetry for any $\alpha \in [0,1]$, we see that $F = \min$ preserves a partial α–order for any $\alpha \in [0,1]$. □

Theorem 12. *If an increasing in each variable function $F : [0,1]^n \to [0,1]$ fulfils conditions $F \leqslant \min$ and $F \gg \min$, then it preserves a strict α–order for any $\alpha \in [0,1]$.*

Proof. By [4] (Theorem 2) a function F preserves an α–irreflexivity for any $\alpha \in [0,1]$ if and only if $F \leqslant \max$, so taking into account Theorem 8 and definition of a strict α–order we get the required statement. □

Example 8. $F = \min$ preserves a strict α–order for any $\alpha \in [0,1]$, because the minimum is increasing and $F \gg \min$ (cf. Theorem 1).

In virtue of Theorems 11, 12 and by definitions of a linear quasi–α–order, a linear α–order and a linear strict α–order and by the fact that F preserves an α–connectedness (a total α–connectedness) if and only if it fulfils (23) (cf. [4], Theorem 8) we see that $F = \min$ does not preserve these linear orders because the minimum does not fulfil (23) (see Example 9).

Example 9. Let $n = 2$. If $F = \min$, then condition (23) is of the form

$$\max(\min(s_1, s_2), \min(t_1, t_2)) \geqslant$$

$$\min(\max(s_1, t_1), \max(s_2, t_2)).$$

where $s_1, s_2, t_1, t_2 \in [0, 1]$. Let us take $s_1 = 0$, $s_2 = 0.5$, $t_1 = 0.5$, $t_2 = 0$. We see that for these numbers the converse inequality is fulfilled, because

$$0 = \max(\min(0, 0.5), \min(0.5, 0)) <$$

$$\min(\max(0, 0.5), \max(0.5, 0)) = 0.5.$$

Thus minimum does not fulfil condition (23).

References

[1] T. Calvo, A. Kolesárová, M. Komorniková, R. Mesiar, Aggregation Operators: Properties, Classes and Construction Methods, in: T. Calvo et al. (Eds.), *Aggregation Operators*. Physica-Verlag, Heildelberg, 2002, pp. 3–104.

[2] J. Drewniak, Fuzzy Relation Calculus, Silesian University, Katowice, 1989.

[3] U. Dudziak, Aggregations preserving graded properties of fuzzy relations, in: K. T. Atanassov et al. (Eds.), *Soft Computing Foundations and Theoretical Aspects*. EXIT, Warszawa, 2004, pp. 191–203.

[4] U. Dudziak, Graded properties of fuzzy relations in aggregation process, Journal of Electrical Engineering 56 (2005) (12/s) 56–58.

[5] J. Fodor, M. Roubens, Fuzzy Preference Modelling and Multicriteria Decision Support, Kluwer Acad. Publ., Dordrecht, 1994.

[6] E. P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Acad. Publ., Dordrecht, 2000.

[7] V. Peneva, I. Popchev, Aggregation of fuzzy relations, C. R. Acad. Bulgare Sci. **51**(9-10) (1998) 41–44.

[8] V. Peneva, I. Popchev, Properties of the aggregation operators related with fuzzy relations, Fuzzy Sets Syst. 139 (3) (2003) 615–633.

[9] S. Saminger, R. Mesiar, U. Bodenhofer, Domination of aggregation operators and preservation of transitivity, Int. J. Uncertain., Fuzziness, Knowl.-Based Syst. 10 Suppl. (2002) 11–35.

[10] L. A. Zadeh, Fuzzy sets, Inform. Control 8 (1965) 338–353.

[11] L.A. Zadeh, Similarity relations and fuzzy orderings, Inform. Sci. 3 (1971) 177–200.

Session 11

Uncertainty Models

Transforming probability intervals into other uncertainty models

Sébastien Destercke
IRSN (bât 702)
CE Cadarache
13115 Saint Paul Lez Durance
sebastien.destercke@irsn.fr
desterck@irit.fr

Didier Dubois
IRIT
Université Paul Sabatier
118 route de Narbonne
31062 Toulouse Cedex 9, France
dubois@irit.fr

Eric Chojnacki
IRSN (bât 702)
CE Cadarache
13115 Saint Paul Lez Durance
eric.chojnacki@irsn.fr

Abstract

Probability intervals are imprecise probability assignments over elementary events. They constitute a very convenient tool to model uncertain information : two common cases are confidence intervals on parameters of multinomial distributions built from sample data and expert opinions provided in terms of such intervals. In this paper, we study how probability intervals can be transformed into other uncertainty models such as possibility distributions, Ferson's p-boxes, random sets and Neumaier's clouds.

Keywords: Probability intervals, random sets, possibility, p-boxes, clouds.

1 Introduction

When modeling uncertainty, the theory of imprecise probabilities [17] and the so-called lower previsions formally subsume most of the existing uncertainty theories. Provided one accepts its behavioral interpretation, this theory offers a very appealing unifying and highly expressive framework. Unfortunately, computational intractability is often the price to pay for such generality and expressiveness. It is thus important to study simpler models that will often be sufficient in practice to solve a given problem tainted with uncertainty. Even when a lot of information is available (but not enough to give precise probabilities), one may want to work with approximated models (e.g. for mathematical or computational tractability considerations).

Probability intervals [1] are among these simpler models. They are easy to understand and computationally tractable. They can be, for example, confidence intervals coming from sample data or opinions given by an expert. Nevertheless, it may happen that one wishes to map the information given by probability intervals into another model, because mathematical tools proper to the latter model must be used or simply to present information in a specific way. This mapping is the object of this paper.

In section 2, we briefly recall the various formalism concerned by this paper. Section 3 then reviews some existing results relating probability intervals and random sets. The next section deals with the relations between probability intervals and p-boxes. Finally, section 5 and 6 are respectively devoted to the transformation of probability intervals into, respectively, possibility distributions and clouds (a recent model proposed by Neumaier [15]).

2 Notations and preliminaries

In the paper, we will restrict ourselves to probability families denoted $\mathcal{P}$ and defined on a finite arbitrary space X of n elements $\{x_1, \dots, x_n\}$. We will now briefly introduce the models studied in the sequel.

2.1 Lower/upper probabilities

Lower ($\underline{P}(A)$) and upper probabilities ($\overline{P}(A)$) on events are respectively defined s.t. $\underline{P}(A) = \inf_{P\in\mathcal{P}} P(A)$ and $\overline{P}(A) = \sup_{P\in\mathcal{P}} P(A)$. We have $\mathcal{P}_{\underline{P},\overline{P}}(A) = \{P | \forall A \subseteq X \text{ measurable}, \underline{P}(A) \leq P(A) \leq \overline{P}(A)\}$. Although this model is already a restriction from more general ones (lower/upper probabilities can be seen as projections of a family $\mathcal{P}$ on the subspace of events, and in general we have $\mathcal{P} \subseteq \mathcal{P}_{\underline{P},\overline{P}}$), it is still fairly general and subsumes all other models studied here.

2.2 Probability intervals

Probability intervals and their properties are extensively studied in [1]. Probability intervals are defined as lower and upper bounds of probabilities restricted to singletons x_i. They can be seen as a family of intervals $L = \{[l_i, u_i], i = 1, \dots, n\}$ defining the family $\mathcal{P}_L = \{P | l_i \leq p(x_i) \leq u_i \ \forall x_i \in X\}$.

In this paper, we will focus on reachable sets L which define non-empty families $\mathcal{P}_L$, since other ones have little interest. A set L will be called reachable if, for each x_i, we can find a probability distribution $P \in \mathcal{P}_L$ s.t. $P(x_i) = l_i$ and another one for which $P(x_i) = u_i$ (in other words, each bound can be reached by at least one distribution in the family). Non-emptiness and reachability respectively correspond to the two conditions

$$\sum_{i=1}^n l_i \leq 1 \leq \sum_{i=1}^n u_i$$

$$\sum_{j \neq i} l_j + u_i \leq 1 \text{ and } \sum_{j \neq i} u_j + l_i \geq 1 \quad \forall\, i$$

Given intervals L_i, general lower and upper probabilities can be computed through the simple formulas

$$\underline{P}(A) = \max(\sum_{x_i \in A} l_i, 1 - \sum_{x_i \notin A} u_i)$$
$$\overline{P}(A) = \min(\sum_{x_i \in A} u_i, 1 - \sum_{x_i \notin A} l_i)$$

and this lower (upper) probability is an order 2 monotone (alternate) Choquet capacity [2].

2.3 Random sets

Formally, a random set is a set-valued mapping from a (here finite) probability space to a set X. It induces lower and upper probabilities on X [5]. Here, we use mass functions [16] to represent random sets. A mass function m is defined by a mapping from the power set 2^X to the unit interval, s.t. $\sum_{E \subseteq X} m(E) = 1$ and $m(\emptyset) = 0$. A set E with positive mass is called a focal set. Two measures, a plausibility and a belief measure can be defined from this mass function:

Belief measure: $Bel(A) = \sum_{E, E \subseteq A} m(E)$

Plausibility measure: $Pl(A) = 1 - Bel(A^c)$

The set $\mathcal{P}_{Bel} = \{P | \forall A \subseteq X$ measurable, $Bel(A) \leq P(A) \leq Pl(A)\}$ is the probability family induced by the belief function.

2.4 (Generalized) P-boxes

A p-box is usually defined on the real line by a pair of cumulative distributions $[\underline{F}, \overline{F}]$, defining the probability family $\mathcal{P}_{[\underline{F}, \overline{F}]} = \{P | \underline{F}(x) \leq F(x) \leq \overline{F}(x) \quad \forall x \in \Re\}$. The notion of cumulative distribution on the real line is based on a natural ordering of numbers. In order to generalize this notion to arbitrary finite sets, we need to define a weak order relation $\leq_R$ on this space. Given $\leq_R$, an R-downset is of the form $\{x_i : x_i \leq_R x\}$, and denoted $(x]_R$. A generalized R-cumulative distribution [7] is defined as the function $F_R : X \to [0,1]$ s.t. $F_R(x) = \Pr((x]_R)$, where Pr is a probability measure on X. We can now define a generalized p-box as a pair $[\underline{F}_R(x), \overline{F}_R(x)]$ of generalized cumulative distributions defining a probability family $\mathcal{P}_{[\underline{F}_R(x), \overline{F}_R(x)]} = \{P | \forall x, \underline{F}_R(x) \leq F_R(x) \leq \overline{F}_R(x)\}$. Generalized P-boxes can also be represented by a set of constraints

$$\alpha_i \leq P(A_i) \leq \beta_i \qquad i = 1, \ldots, n \tag{1}$$

where $0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_n \leq 1$, $0 \leq \beta_1 \leq \beta_2 \leq \ldots \leq \beta_n \leq 1$ and $A_i = (x_i]_R, \forall x_i \in X$ with $x_i \leq_R x_j$ iff $i < j$ (sets A_i form a sequence of nested confidence sets $\emptyset \subset A_1 \subset A_2 \subset \ldots \subset A_n \subset X$). When $X = \Re$ and $A_i = (-\infty, x_i]$, we find back the usual definition of p-boxes.

2.5 Possibility distributions

A possibility distribution π is a mapping from X to the unit interval (hence a fuzzy set) such that $\pi(x) = 1$ for some $x \in X$. Several set-functions can be defined from them [8]:

Possibility measures: $\Pi(A) = \sup_{x \in A} \pi(x)$

Necessity measures: $N(A) = 1 - \Pi(A^c)$

Guaranteed poss. measures: $\Delta(A) = \inf_{x \in A} \pi(x)$

Possibility degrees express the extent to which an event is plausible, i.e., consistent with a possible state of the world, necessity degrees express the certainty of events and Δ-measures the extent to which all states of the world where A occurs are plausible. They apply to so-called guaranteed possibility distributions [8] generally denoted by δ.

A possibility degree can be viewed as an upper bound of a probability degree [9]. Let $\mathcal{P}_\pi = \{P, \forall A \subseteq X$ measurable, $P(A) \leq \Pi(A)\}$ be the set of probability measures encoded by π. A necessity measure is a special case of belief function when the focal sets are nested.

2.6 Clouds

Formally, a cloud is described by an Interval-Valued Fuzzy Set (IVF) s.t. $(0,1) \subseteq \cup_{x \in X} F(x) \subseteq [0,1]$, where $F(x)$ is an interval $[\delta(x), \pi(x)]$. A cloud is called *thin* when the two membership functions coincide ($\delta = \pi$). It is called *fuzzy* when the lower membership function δ is 0 everywhere. Let α_i be a sequence of α-cuts s.t. $1 = \alpha_0 > \alpha_1 > \alpha_2 > \ldots > \alpha_n > \alpha_{n+1} = 0$

with A_i the strong α-cut of π and B_j the α-cut of δ ($A_i = \{x_i, \pi(x_i) > \alpha_i\}$ and $B_i = \{x_i, \delta(x_i) \geq \alpha_i\}$). Then, a random variable x is in a cloud if it satisfies the constraints

$$P(B_i) \leq 1 - \alpha_i \leq P(A_i) \quad i = 1, \ldots, n. \tag{2}$$

with $B_i \subseteq A_i$. The probability family induced by a cloud $[\delta(x), \pi(x)]$ will be noted $\mathcal{P}_{[\delta,\pi]}$.

Moreover, the following property linking clouds and possibility distributions has been shown by Dubois and Prade [11]:

Proposition 1. *A probability family $\mathcal{P}_{[\delta,\pi]}$ described by the cloud (δ, π) is equivalent to the family $\mathcal{P}_\pi \cap \mathcal{P}_{1-\delta}$ described by the two possibility distributions π and $1 - \delta$.*

3 Probability intervals and random sets

Most of the results presented in this section can be found in [1, 6], where more details can be found. There are two main approaches to build a belief function Bel from a set L_e of probability intervals.

The first one, explored in [13] by Lemmer and Kyburg, considers probability intervals as a partial specification of a belief function and consists to find a belief function Bel_1 that extends the intervals s.t.

$$Bel_1(x_i) = l_i \text{ and } Pl_1(x_i) = u_i \ \forall i \tag{3}$$

As shown in [13], finding such a belief function is possible iff the tree following conditions hold

$$\sum_{i=1}^n l_i \leq 1 \leq \sum_{i=1}^n u_i$$

$$\sum_{j \neq i} l_j + u_i \leq 1 \text{ and } \sum_{j \neq i} u_j + l_i \geq 1 \quad \forall\, i$$

$$\sum_{i=1}^n l_i + \sum_{i=1}^n u_i \geq 2$$

where the two first conditions correspond to non-emptiness and reachability (which are always satisfied by supposition). Lemmer and Kyburg also provide a means to build one of the belief function satisfying constraints (3). Let us note that with this method, we have $Bel_1(A) \geq P(A)$, which imply $\mathcal{P}_{Bel_1} \subset \mathcal{P}_{L_e}$. Thus, The belief function Bel_1 is an inner approximation of the family $\mathcal{P}_{L_e}$.

The second approach, extensively explored by Denoeux [6], considers probability intervals L_e as some "most committed" information and try to find a conservative belief function Bel_2 s.t. $Bel_2(A) \leq \underline{P}_{L_e}(A)\ \forall A$. Since there exist a lot of such belief functions, Denoeux also proposes to find the belief function Bel_2 that maximizes a given specifity criterion, in order to keep as much information as possible (in [6], this criterion is the sum of belief degrees over events). Obviously, with this approach, we have $\mathcal{P}_{L_e} \subset \mathcal{P}_{Bel_2}$, and Bel_2 is this time an outer approximation of the family $\mathcal{P}_{L_e}$. Methods given in [6] insure that $Bel_2(x_i) = l_i$, but it can be checked from the examples in [6] that $Pl_2(x_i) > u_i$, even if the three conditions required by the first approach hold.

4 Probability intervals and (Generalized) P-boxes

Given a set L of probability intervals and a meaningful ordering relation $\leq_R$ between elements x_i, one can easily build a generalized p-box $[\underline{F}, \overline{F}]_L$ from L. Given the consecutive sets $A_i = (x_i]_R, \forall x_i \in X$ and the ordering s.t. $x_i \leq_R x_j$ iff $i < j$, lower and upper generalized cumulative distributions corresponding to L are, respectively

$$\begin{aligned} \underline{F}_R(x_i) &= \underline{P}(A_i) = \max(\textstyle\sum_{x_i \in A_i} l_j, 1 - \sum_{x_i \notin A_i} u_j) \\ \overline{F}_R(x_i) &= \overline{P}(A_i) = \min(\textstyle\sum_{x_i \in A_i} u_i, 1 - \sum_{x_i \notin A_i} l_i) \end{aligned} \tag{4}$$

Now, if we consider that this p-box is all the information we have, one can easily find back probability intervals L' from this information s.t.

$$\begin{aligned} \underline{P}'(x_i) = l'_i &= \max(0, \underline{P}(A_i) - \overline{P}(A_{i-1})) \\ \overline{P}'(x_i) = u'_i &= \overline{P}(A_i) - \underline{P}(A_{i-1}) \end{aligned}$$

and we have the following proposition

Proposition 2. *Given an initial set L of probability intervals over a space X, and given the transformations*

$$\textit{Set } L \quad \overrightarrow{p-box} \quad [\underline{F}_R(x), \overline{F}_R(x)] \quad \overrightarrow{Interval} \quad \textit{Set } L'$$

we have that $\mathcal{P}_L \subseteq \mathcal{P}_{L'}$.

Proof. Looking at equations given above, we can easily express $[l'_i, u'_i]$ in term of values l_i, u_i of the original set L, this gives us

$$\begin{aligned} l'_i \quad = \max(0, \quad & \sum_{x_i \in A_i} l_i - \sum_{x_i \in A_{i-1}} u_i, \\ & \sum_{x_i \in A_i} l_i + \sum_{x_i \notin A_{i-1}} l_i - 1, \\ & 1 - \sum_{x_i \notin A_i} u_i - \sum_{x_i \in A_{i-1}} u_i, \\ & \sum_{x_i \notin A_{i-1}} l_i - \sum_{x_i \notin A_i} u_i) \end{aligned}$$

$$= \max(0, \quad l_i + \sum_{x_i \in A_{i-1}} (l_i - u_i),$$
$$l_i + (\sum_X l_i - 1),$$
$$1 - \sum_{X \setminus x_i} u_i$$
$$l_i + \sum_{x_i \notin A_i} (l_i - u_i))$$

and, given that $u_i \geq l_i$ and that the set L is reachable and non-empty, we have that $l'_i \leq l_i$. The same procedure can be followed for the bounds u'_i, and we have $\mathcal{P}_L \subseteq \mathcal{P}_{L'}$. We also have $l'_i = l_i; u'_i = u_i \ \forall i$ only in degenerate cases of precise information or complete ignorance. □

Not surprisingly, this proposition shows that transforming probability intervals into p-boxes implies a loss of information. Let us note that the same argument holds when one wants to transform a generalized p-box $[\underline{F}, \overline{F}]$ into probability intervals, and then get back a generalized p-box $[\underline{F}', \overline{F}']$ from these intervals using equations (4) (i.e. we have $\mathcal{P}_{[\underline{F},\overline{F}]} \subset \mathcal{P}_{[\underline{F}',\overline{F}']}$).

Example 1. *Let us take the same four probability intervals as in the example given in [14], summarized in the following tabular*

	x_1	x_2	x_3	x_4
l_i	*0.10*	*0.34*	*0.25*	*0*
u_i	*0.28*	*0.56*	*0.46*	*0.08*

if we consider the order R s.t. $x_i \leq_R x_j$ iff $i \leq j$, this gives us the following generalized p-box

	$\underline{F}$	$\overline{F}$
$A_1 = \{x_1\}$	*0.10*	*0.28*
$A_2 = \{x_1, x_2\}$	*0.46*	*0.75*
$A_3 = \{x_1, x_2, x_3\}$	*0.92*	*1*
$A_4 = X$	*1*	*1*

and if we want to get back probability intervals from this sole generalized p-box, we obtain

	x_1	x_2	x_3	x_4
l'_i	*0.10*	*0.18*	*0.17*	*0*
u'_i	*0.28*	*0.65*	*0.54*	*0.08*

which is clearly less informative than the first probability intervals.

5 Probability intervals and possibility distributions

The problem of transforming a probability distribution into a quantitative possibility distribution has been adressed by many authors (see [10] for an extended discussion about the links between probabilities and quantitative possibility theory). In this paper, we will follow the same line as Masson and Denoeux in [14], where authors study the problem of transforming interval probabilities into a possibility distribution.

Concerning the transformation of a precise probability into a possibility distribution, a first consistency principle was informally stated by Zadeh [18] as: what is probable should be possible. It was later translated by Dubois and Prade [12] as the mathematical constraint

$$P(A) \leq \Pi(A) \quad \forall A \subseteq X$$

and the possibility measure Π is said to dominate the probability measure P. The transformation of a probability into a possibility then consists of choosing a possibility measure amongst those dominating P. Dubois and prade [12] then proposed to add the following constraint

$$p(x_i) \leq p(x_j) \leftrightarrow \pi(x_i) \leq \pi(x_j)$$

and to choose the least specific possibility distribution (π' is more specific than π if $\pi' \leq \pi \ \forall x$) respecting these two constraints. They showed that the solution exists and is unique. Let us consider the order on probability masses s.t.

$$p(x_1) \leq p(x_2) \leq \ldots \leq p(x_j) \leq \ldots \leq p(x_n)$$

Dubois and Prade's transformation can then be formulated as

$$\pi(x_i) = \sum_{j=1}^{i} p(x_j)$$

When working with a set $L = \{[l_i, u_i], i = 1, \ldots, n\}$ of probability intervals, the order induced on probability masses is no longer complete, and the partial order is reduced to

$$p(x_i) \leq p(x_j) \leftrightarrow u_i \leq l_j$$

and two probabilities $p(x_i), p(x_j)$ are incomparable if intervals $[l_i, u_i], [l_j, u_j]$ intersect in some way. Let us note $\mathcal{M}$ this partial order and $\mathcal{C}$ the set of its linear extensions (a linear extension $C_l \in \mathcal{C}$ is a complete order that is compatible with the partial order $\mathcal{M}$). Given this partial order, Masson and Denoeux [14] propose the following procedure to transform the set of probability intervals into a possibility distribution:

1. For each order $C_l \in \mathcal{C}$ and each element x_i, solve

$$\pi(x_i)^{C_l} = \max_{p(x_1),\ldots,p(x_n)} \sum_{\sigma_l^{-1}(j) \leq \sigma_l^{-1}(i)} p(x_j) \quad (5)$$

under the constraints

$$\begin{cases} \sum_{k=1,\ldots,n} p(x_k) = 1 \\ l_k \leq p_k \leq u_k \\ p(x_{\sigma_l(1)}) \leq p(x_{\sigma_l(2)}) \leq \ldots \leq p(x_{\sigma_l(n)}) \end{cases}$$

where σ_l is the permutation of $p(x_k)$ associated with the linear extension C_l

2. Take the distribution dominating all distributions $\pi(x_i)^{C_l}$ s.t.

$$\pi(x_i) = \max_{C_l \in \mathcal{C}} \pi(x_i)^{C_l} \quad \forall i \tag{6}$$

this procedure insures that the resulting possibility distribution π will dominate every probability distribution contained in $\mathcal{P}_L$. We thus have $\mathcal{P}_L \subset \mathcal{P}_\pi$. Let us note that this transformation can produce an important loss of information.

Example 2. *Taking back probability intervals of example 1, we have here three possible linear extensions* $C_l \in \mathcal{C}$

$$\begin{aligned} C_1 &= (L_4, L_1, L_3, L_2) \\ C_2 &= (L_4, L_1, L_2, L_3) \\ C_3 &= (L_4, L_3, L_1, L_2) \end{aligned}$$

using the above method on these three orders gives the following distribution (see [14] for more details)

	x_1	x_2	x_3	x_4
π	*0.64*	*1*	*1*	*0.08*

Let us note that, if this transformation produces the most specific possibility distribution dominating a precise probability distribution (i.e. $u_i = l_i \ \forall i$ and $\sum u_i = 1$), this is no longer the case when the probability distribution is imprecisely known through intervals ($u_i \neq l_i$). To see this, let us first recall that an upper generalized R-cumulative distribution $\overline{F}_R$ can be seen as a possibility distribution dominating a probability measure, since it is a maxitive measure (i.e. we have $\max_{x \in A} \overline{F}_R(x) \geq \Pr(A), \forall A \subseteq X$). Any R-cumulative distribution dominating a family $\mathcal{P}_L$ induced by probability intervals is thus also a possibility measure dominating this family. Now, in our example, let us consider the following order $<_R$ between the four elements: $x_4 <_R x_1 <_R x_3 <_R x_2$. From this order, we can build the following R-cumulative distribution:

	x_1	x_2	x_3	x_4
$\overline{F}_R = \pi_R$	0.36	1	0.66	0.08

which still dominates $\mathcal{P}_L$ and is more specific than the distribution built by the method given in [14]. This shows that Masson and Denoeux's method tends to give conservative bounds, and thus can result in an important loss of information, which seems hard to justify. This important loss is due to the fact that Masson and Denoeux do not consider any specific ordering relation R between the elements of X, an assumption that is made in the counter-example given above.

6 Probability intervals and clouds

Let L be a set of probability intervals and $\mathcal{P}_L$ the associated family. In this section, we introduce a method, inspired from [14], which transforms intervals L into a cloud $[\delta, \pi]_L$.

From property 1, we know that families described by clouds are equivalent to the intersection of two families described by possibility distributions. Another property of clouds is that a discrete thin cloud, up to the right transformation, can represent a precise probability distribution. Let $\pi_1(x_i) = a_i$ s.t. $a_0 = 0 < a_1 < \ldots < a_n = 1$. Consider then $\pi_2(x_i) = 1 - a_{i-1} = 1 - \delta(x_i)$. Now let us consider the cloud $[\delta = 1 - \pi_2, \pi = \pi_1]$. It has been proved by Dubois and Prade [11] that the probability family $\mathcal{P}_{[\delta,\pi]}$ induced by this cloud contains a unique probability measure P s.t. $p(x_i) = a_i - a_{i-1} \ \forall i = 1, \ldots, n$.

Given this property, two requirements that should, in our opinion, follow any transformation of probability intervals into clouds are the following:

- If intervals L describe a precise probability distribution, then the transformation should result in the corresponding thin cloud.
- The family $\mathcal{P}_{[\delta,\pi]_L}$ should be an outer approximation of $\mathcal{P}_L$ (i.e. $\mathcal{P}_L \subset \mathcal{P}_{[\delta,\pi]_L}$)

Let us consider the distribution π built with the method described in section 5 as the upper distribution of the cloud. By reversing the inequality under the summand in equation (5), we can build another distribution π_δ in the following way:

1. For each order $C_l \in \mathcal{C}$ and each element x_i, solve

$$\begin{aligned} \pi_\delta^{C_l}(x_i) &= \max_{p(x_1),\ldots,p(x_n)} \sum_{\sigma_l^{-1}(i) \leq \sigma_l^{-1}(j)} p(x_j) \\ &= 1 - \min_{p(x_1),\ldots,p(x_n)} \sum_{\sigma_l^{-1}(j) < \sigma_l^{-1}(i)} p(x_j) \\ &= 1 - \delta^{C_l}(x_i) \end{aligned}$$

with the same constraints as in section 5

2. Take the distribution dominating all distributions $\pi_\delta^{C_l}(x_i)$

$$\pi_\delta(x_i) = 1 - \delta(x_i) = \max_{C_l \in \mathcal{C}} \pi_\delta^{C_l}(x_i) \quad \forall i \tag{7}$$

And $[\delta, \pi]$ is actually a cloud ($\delta < \pi$, since $\delta(x_i)$ is a minimum computed over less intervals $L(x_i)$ than $\pi(x_i)$, which is moreover a maximum). It can be easily checked that if L is reduced to a precise probability distribution, then the above method results in the thin cloud corresponding to this distribution. We also have that $\mathcal{P}_L \subset \mathcal{P}_{\pi_\delta}$ and $\mathcal{P}_L \subset \mathcal{P}_\pi$ by construction, thus we have $\mathcal{P}_L \subset \mathcal{P}_{[\delta,\pi]_L} = \mathcal{P}_{\pi_\delta} \cap \mathcal{P}_\pi$, and our two requirements are satisfied.

Example 3. *Again, we consider the probability intervals given in example 1. From the three orders C_i given above, we can obtain the following π_δ*

C_i	$\pi_\delta(x_1)$	$\pi_\delta(x_2)$	$\pi_\delta(x_3)$	$\pi_\delta(x_4)$
1	*1*	*0.16*	*0.63*	*1*
2	*1*	*0.9*	*0.46*	*1*
3	*0.75*	*0.5*	*1*	*1*
max	*1*	*0.9*	*1*	*1*

and, finally, the following cloud

	x_1	x_2	x_3	x_4
π	*0.64*	*1*	*1*	*0.08*
δ	*0*	*0.1*	*0*	*0*

which, in this case, is only a little more informative than the upper distribution taken alone (indeed, the sole added constraint is that $p(x_2) \leq 0.9$).

As in the previous section, this method can be criticized upon the ground that an important amount of information is lost. Instead, one could take, for example, the cloud associated to the generalized p-box induced by the order $x_4 <_R x_1 <_R x_3 <_R x_2$ (since generalized p-boxes are a particular case of clouds). This would give the following distributions:

	x_1	x_2	x_3	x_4
$\overline{F}_R = \pi_R$	0.36	1	0.66	0.08
$\underline{F}_R$	0.1	1	0.44	0
δ_R	0	0.44	0.1	0

Where δ_R is $\underline{F}_R$ after a simple shift of values (the necessity of this transformation, already emphasized in [11], arises from the fact that sets A_i are strong α-cuts, while B_i are simple α-cuts). The cloud $[\delta_R, \pi_R]$ is obviously more specific than the first cloud while we still have $\mathcal{P}_L \subset \mathcal{P}_{[\delta_R,\pi_R]}$.

7 Conclusions

Probability intervals are very convenient to model uncertainty, and can be encountered in various situations. In this paper, we study how they can be transformed into other popular models of imprecise probabilities. Except for one, every method recalled or proposed here give outer approximation of the family $\mathcal{P}_L$. This corresponds to a cautious view, since there is no additional information present in the model resulting from the transformation (but some information can be lost in the transformation process).

If probability intervals are reduced to precise probability distributions, every proposed transformation result in the model corresponding to this precise probability, except for possibility distributions, which are the only model studied here that can't be seen as a generalization of classical probabilities.

Interestingly enough, most (some subclasses of clouds are not) of the studied representations in this paper are special cases of random sets, allowing one to use all the rich mathematical background of this theory as well as many computational simplifications (e.g. easiness to compute Choquet Integral).

Each of the models presented here has its own practical interest in term of expressiveness or tractability. The natural continuation of the work initiated in this paper is to extend the study made by De Campos et al. [1] to every model studied here. What becomes of random sets, possibility distributions, generalized p-boxes and clouds after fusion, marginalization, conditioning or propagation? Is it still the same kind of representation after having applied these mathematical tools, and under which assumptions? To which extends are these representations informative? Can they easily be elicited or integrated? If many results already exist for random sets and possibility distributions, few have been derived for generalized p-boxes or clouds, due to the fact that these two latter representations have only been recently proposed.

In term of behavioral interpretation of imprecise probabilities [17], conditions of non-emptiness and reachability respectively correspond to avoiding sure loss and to coherence of lower previsions. It is also interesting to note that, if we see probability intervals as a constraint satisfaction problem (CSP) [4], non-emptiness and reachability correspond to the notions of existence of a solution and of bounds consistency. Another interesting work should be to formalize links between CSP and imprecise probabilities, with the aim to study to which extent technics used in CSP can be used to solve practical problems commonly encountered when dealing with imprecise probabilities. To our knowledge, this work largely remains to be done, although CSP technics are already used to solve practical problems related to (imprecise) probabilistic reasoning, most of them being related to (credal) bayesian networks (see, for example [3]).

Acknowledgement

This paper has been supported by a grant from the

Institut de Radioprotection et de Sûreté Nucléaire (IRSN). Scientific responsibility rests with the authors.

References

[1] L. Campos, J. Huete, and S. Moral. Probability intervals : a tool for uncertain reasoning. *I. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2:167–196, 1994.

[2] G. Choquet. Theory of capacities. *Annales de l'institut Fourier*, 5:131–295, 1954.

[3] R. Dechter. Bucket-elimination : a unifying framework for probabilistic inference. In E. Horvitz and F. Jensen, editors, *Proc. 12th conf. on Uncertainty in A.I.*, pages 211–219, Portland, 1996.

[4] R. Dechter. *Constraint Processing*. Morgan Kaufman, 2003.

[5] A. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[6] T. Denoeux. Constructing belief functions from sample data using multinomial confidence regions. *I. J. of Approximate Reasoning*, 42, 2006.

[7] S. Destercke and D. Dubois. A unified view of some representations of imprecise probabilities. In J. Lawry, E. Miranda, A. Bugarin, and S. Li, editors, *Int. Conf. on Soft Methods in Probability and Statistics (SMPS)*, Advances in Soft Computing, pages 249–257, Bristol, 2006. Springer.

[8] D. Dubois, P. Hajek, and H. Prade. Knowledge-driven versus data-driven logics. *Journal of logic, Language and information*, 9:65–89, 2000.

[9] D. Dubois and H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49:65–74, 1992.

[10] D. Dubois and H. Prade. Quantitative possibility theory and its probabilistic connections. In P. Grzegorzewski, O. Hryniewicz, and M. Angeles-Gil, editors, *Soft Methods in Probability, Statistics and Data Analysis*, Advances in Soft Computing, pages 3–26. Physica Verlag, Heidelberg - Germany, 2002.

[11] D. Dubois and H. Prade. Interval-valued fuzzy sets, possibility theory and imprecise probability. In *Proceedings of International Conference in Fuzzy Logic and Technology (EUSFLAT'05)*, Barcelona, September 2005.

[12] D. Dubois, H. Prade, and S. Sandri. On possibility/probability transformations. In *Proc. of the Fourth International Fuzzy Systems Association World Congress (IFSA'91)*, pages 50–53, Brussels, Belgium, 1991.

[13] J. Lemmer and H. Kyburg. Conditions for the existence of belief functions corresponding to intervals of belief. In *Proc. 9th National Conference on A.I.*, pages 488–493, 1991.

[14] M. Masson and T. Denoeux. Inferring a possibility distribution from empirical data. *Fuzzy Sets and Systems*, 157(3):319–340, february 2006.

[15] A. Neumaier. Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 10:249–272, 2004.

[16] G. Shafer. *A mathematical Theory of Evidence*. Princeton University Press, 1976.

[17] P. Walley. *Statistical reasoning with imprecise Probabilities*. Chapman and Hall, 1991.

[18] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1:3–28, 1978.

Lattice-Valued Possibilistic Entropy Functions

Ivan Kramosil
Institute of Computer Science AS CR, v.v.i.
Pod Vodárenskou věží 2
182 07 Prague, Czech Republic
kramosil@cs.cas.cz

Abstract

Lattice-valued entropy functions defined by a lattice-valued possibilistic distribution π on a space Ω are defined as the expected value (in the sense of Sugeno integral) of the complement of the value $\pi(\omega)$ with ω ranging over Ω. The analysis is done in parallel for two alternative interpretations of the notion of complement in the complete lattice in question. Supposing that this complete lattice is completely distributive in the defined sense, the entropy values defined by independent products of finite sequences of lattice-valued possibilistic distributions are proved to be defined by the supremum value of the entropies defined by particular distributions.

Keywords: lattice-valued possibilistic distribution, possibilistically independent product of distributions, possibilistic entropy functions.

1 Introduction and Motivation

Since the earliest investigations in the field of possibilistic (possibility) theory and measures as conceived by L. A. Zadeh in 1978 [5], possibilistic measures have been considered as a qualitatively different alternative tool for uncertainty quantification and processing, if compared with those offered by additive measures and probability theory, but the notions introduced and results achieved by the classical probabilistic approach have been continually used as an inspiration for a further development of possibilistic theory and measures (for our purposes, first of all [1] is worth being explicitly mentioned).

Within the framework of standard (Shannon) information theory, based on Kolmogorov axiomatic probability theory, the notion of *entropy* plays the fundamental role, as it enables to quantify the amount of uncertainty (in the sense of randomness) contained or hidden in a probability distribution over a nonempty space of random events, in other sense, the (expected) amount of information obtained when realizing a random sample from this distribution. A great number of important and interesting results dealing with information encoding and transmission and with statistical decision making under uncertainty have been achieved when applying the notion of information. So, it seems quite reasonable to seek for an alternative modification of the notion of entropy applicable to possibilistic measures. For real-valued possibilistic measures some rather elementary ideas and results can be found in [4], the case of non-numerical and, in particular, lattice-valued possibilistic measures will be, very briefly, discussed below. Let us sketch, very shortly, the probabilistic version of the notion of entropy function (cf. [3] or any elementary textbook or monograph on information theory).

Restricting ourselves to the most simple case, let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ be a finite space and let $P : \Omega \to [0,1]$ be a probability distribution on Ω, hence, $\sum_{i=1}^{n} P(\omega_i) = 1$. *(Shannon) entropy* $I(P)$ of P is defined by

$$I(P) = \sum_{i=1}^{n} (lg(1/P(\omega_i))P(\omega_i) = -\sum_{i=1}^{n} (lg\, P(\omega_i))P(\omega_i) \tag{1.1}$$

setting $0\, lg\, 0 = 0$ and applying, as a rule, the logarithm to the base 2, so that $I(P)$ is the expected value of the random variable $1/P(\cdot)$. Within the framework of possibilistic measures the normalized Sugeno integral will play the role of expected value, so that an alternative entropy function $I^*(P)$, defined by

$$I^*(P) = \sum_{i=1}^{n} (1 - P(\omega_i))P(\omega_i) = 1 - \sum_{i=1}^{n} (P(\omega_i))^2, \tag{1.2}$$

will serve as an inspiration. Both $I(P)$ and $I^*(P)$ take their maximum values ($lg\, n$ for $I(P)$ and $1-(1/n)$ for $I^*(P)$) iff $P(\omega_i) = 1/n$ for each $i \leq n$ (the uniform probability distribution on Ω), and take their minimum value 0 for both $I(P)$ and $I^*(P)$ iff $P(\omega_{i_0}) = 1$ for one $i_0 \leq n$ (hence, $P(\omega_i) = 0$ for each $i \leq n, i \neq i_0$).

2 Lattice-Valued Possiblistic Distributions

Partially ordered set (poset) is a pair $\mathcal{T} = \langle T, \leq \rangle$, where T is a non-empty set and $\leq$ is a partial ordering on T, i.e., $\leq$ is a reflexive, antisymmetric and transitive binary relation on T (subset of $T \times T$, in set-theoretic terms). Let $\bigvee, \vee (\bigwedge, \wedge,$ resp.) denote the supremum (infimum, resp.) operation defined by $\leq$ in $\mathcal{T}$. Poset $\mathcal{T}$ is called *complete lattice*, if for each nonempty subset $A \subset T$ its supremum $\bigvee A (= \bigvee_{t \in A} t)$ and infimum $\bigwedge A (= \bigwedge_{t \in A} t)$ are defined. The *zero (unit,* resp.) element of the complete lattice $\mathcal{T} = \langle T, \leq \rangle$ is denoted by $\oslash_{\mathcal{T}}$ ($\mathbf{1}_{\mathcal{T}}$, resp.) and defined by $\oslash_{\mathcal{T}} = \bigwedge T$ ($\mathbf{1}_{\mathcal{T}} = \bigvee T$, resp.). By convention we set $\bigvee \emptyset = \oslash_{\mathcal{T}}$ and $\bigwedge \emptyset = \mathbf{1}_{\mathcal{T}}$ for the empty subset of T.

A complete lattice $\mathcal{T} = \langle T, \leq \rangle$ is called *distributive,* if the relation $s \wedge (t_1 \vee t_2) = (s \wedge t_1) \vee (s \wedge t_2)$ holds for each $s, t_1, t_2 \in T$ and $\mathcal{T}$ is called *completely distributive,* if the relation $s \wedge (\bigvee A) = \bigvee_{t \in A} (s \wedge t)$ holds for each $s \in T$ and $\emptyset \neq A \subset T$. The inequality $s \wedge (\bigvee A) \geq \bigvee_{t \in A} (s \wedge t)$ obviously holds for each poset $\mathcal{T}$, but the inverse inequality need not hold in general. E.g., the complete lattice $\mathcal{T}_0 = \langle \{t_0, t_1, t_2, t_3, t_4\}, \leq \rangle$ such that $t_0 < t_i < t_4$ holds for each $i = 1, 2, 3$, but no $t_i, t_j, i, j = 1, 2, 3, i \neq j$, are comparable by $\leq$, is not distributive. Both the most often used structures for uncertainty degrees, i.e., the unit interval of real numbers equipped by their standard ordering, as well as the space of all subsets of a nonempty set X partially ordered by set inclusion, can be easily seen to define completely distributive complete lattices.

Contrary to the operation $1 - x$ for $x \in [0, 1]$ or the set complement $X - A$ for $A \subset X$ there is no general primary operation of complement in complete lattices. Let us define, for our purposes, the $\mathcal{T}$-complement t^C for each $t \in T$, setting

$$t^C = \bigvee \{s \in T : s \wedge t = \oslash_{\mathcal{T}}\}. \tag{2.1}$$

In particular, $\oslash_{\mathcal{T}}^C = \mathbf{1}_{\mathcal{T}}$ and $\mathbf{1}_{\mathcal{T}}^C = \oslash_{\mathcal{T}}$, for $\mathcal{T} = \langle [0, 1], \leq \rangle$ we obtain $0^C = 1$ and $x^C = 0$ for each $0 < x \leq 1$, and for $\mathcal{T} = \langle \mathcal{P}(X), \subset \rangle$ $\mathcal{T}$-complement agrees with the standard set-complement operation.

Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let Ω be a nonempty space. A mapping π which takes Ω into T is called *$\mathcal{T}$-(valued) possibilistic distribution* on Ω, if $\bigvee_{\omega \in \Omega} \pi(\omega) = \mathbf{1}_{\mathcal{T}}$ holds, hence, if π is a $\mathcal{T}$-valued normalized fuzzy subset of Ω as introduced by J. A. Goguen in [2]. The *$\mathcal{T}$-possibilistic measure* Π induced by π on the power-set $\mathcal{P}(\Omega)$ is defined by $\Pi(A) = \bigvee_{\omega \in A} \pi(\omega)$, hence, $\Pi(\Omega) = \mathbf{1}_{\mathcal{T}}$ and $\Pi(\emptyset) = \oslash_{\mathcal{T}}$ by convention for the empty subset of Ω.

3 Lattice-Valued Possibilistic Entropy Functions

In what follows, we will need a mathematical tool for integration of lattice-valued mappings with respect to lattice-valued possibilistic measures.

As the first attempt, for these sakes the notion of Sugeno integral as analyzed and applied in [1] with the minimum operation on T taken as the particular case of t-norm seems to be more or less adequate and the limited scope of this contribution does not offer a space for a more detailed discussion on this matter. Nevertheless, a modification of notions and results to be introduced below to the case of a general t-norm processed by a tool of integration more powerful than the minimum-based Sugeno integral seems to be, beyond any doubts, worth being pursued in more detail when going on with the research dealing with lattice-valued possibilistic entropy functions.

Definition 3.1 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let Ω be a non-empty space, let π be a $\mathcal{T}$-possibilistic distribution on Ω with Π denoting the $\mathcal{T}$-possibilistic measure on $\mathcal{P}(\Omega)$ induced by π, let f be a mapping which takes Ω into T. The value

$$\oint f d\Pi = \bigvee_{t \in T} [t \wedge \Pi(\{\omega \in \Omega : f(\omega) \geq t\})] \tag{3.1}$$

is called the *(Sugeno) integral of the mapping f* over the space Ω, with respect to the possibilistic measure Π on $\mathcal{P}(\Omega)$ and with infimum operation $\wedge$ taken as the t-norm on T.

As a matter of fact, as analyzed and proved in [1], under our setting the value $\oint f d\Pi$ is always defined and the relation

$$\oint f d\Pi = \bigvee_{\omega \in \Omega} (f(\omega) \wedge \pi(\omega)) \tag{3.2}$$

holds. In order to recall the role of integrals as expected values of variables charged by uncertainty we will often write $Ef d\Pi$ or $Ef(\cdot) d\Pi$ instead of $\oint f d\Pi$.

When aiming to modify the definition of entropy function $I^*(P)$, given by (1.2), to the case of lattice-valued

possibilistic distributions, the problem arises how to replace the substraction operation $1 - x$. In the case of probability distribution P over finite or countable space Ω the identity $1 - P(\omega) = P_*(\Omega - \{\omega\}) = \sum_{\omega_1 \in \Omega, \omega_1 \neq \omega} P(\omega_1)$ is obvious, but in the case of possibilistic distributions (real-valued as well as lattice-valued ones in general) the relation $1 - \pi(\omega) = \Pi(\Omega - \{\omega\})$ or $(\pi(\omega))^C = \Pi(\Omega - \{\omega\})$ need not hold in general. Aiming to take in consideration both the cases in parallel, we arrive at the following definition.

Definition 3.2 Let $\mathcal{T} = \langle T, \leq \rangle$ be a complete lattice, let π be a $\mathcal{T}$-possibilistic distribution on a nonempty space Ω. (1)-*entropy* of π is denoted as $I_1(\pi)$ and defined by

$$I_1(\pi) = E(\pi(\omega))^C d\Pi = \bigvee_{\omega \in \Omega} (((\pi(\omega))^C \wedge \pi(\omega)), \quad (3.3)$$

(2)-*entropy* of π is denoted as $I_2(\pi)$ and defined by

$$I_2(\pi) = E((\Pi(\Omega - \{\omega\}))d\Pi = \bigvee_{\omega \in \Omega} ((\Pi(\Omega - \{\omega\})) \wedge \pi(\omega)). \quad (3.4)$$

Both the values $I_1(\pi)$ and $I_2(\pi)$ are evidently defined for each $\mathcal{T}$-possibilistic distribution π on Ω.

Theorem 3.1 Let $\mathcal{T} = \langle T, \leq \rangle$ be a distributive complete lattice, let π be a $\mathcal{T}$-possibilistic distribution on a nonempty space Ω. Then the inequality $I_1(\pi) \leq I_2(\pi)$ holds. In particular, if $\mathcal{T}$ is completely distributive, then $I_1(\pi) = \oslash_{\mathcal{T}}$, so that the inequality $I_1(\pi) \leq I_2(\pi)$ follows trivially.

Proof. Let $s \in T$ and $\Omega \in \Omega$ be such that $s \wedge \pi(\omega) = \oslash_{\mathcal{T}}$ holds. Due to the assumed distributivity of $\mathcal{T}$ we obtain that

$$\begin{aligned} s &= s \wedge \mathbf{1}_{\mathcal{T}} = s \wedge \left(\bigvee_{\omega_1 \in \Omega} \pi(\omega_1) \right) = \\ &= s \wedge \left(\left(\bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1) \right) \vee \pi(\omega) \right) = \\ &= \left(s \wedge \bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1) \right) \vee (s \wedge \pi(\omega)) = \\ &= s \wedge \left(\bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1) \right), \end{aligned} \quad (3.5)$$

so that the inequalities $s \leq \bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1)$ and

$$\begin{aligned} (\pi(\omega))^C &= \bigvee \{s \in T : s \wedge \pi(\omega) = \oslash_{\mathcal{T}}\} \leq \\ &\leq \bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1) = \Pi(\Omega - \{\omega\}) \end{aligned} \quad (3.6)$$

are valid. Hence, the inequality

$$\begin{aligned} I_1(\pi) &= \bigvee_{\omega \in \Omega} ((\pi(\omega))^C \wedge \pi(\omega)) \leq \\ &\leq \bigvee_{\omega \in \Omega} ((\Pi(\Omega - \{\omega\}) \wedge \pi(\omega)) = I_2(\pi) \end{aligned} \quad (3.7)$$

results. If $\mathcal{T}$ is completely distributive, then for each $t \in T$ we have

$$\begin{aligned} t^C \wedge t &= \left(\bigvee \{s \in T : s \wedge t = \oslash_{\mathcal{T}}\} \right) \wedge t = \\ &= \bigvee \{s \wedge t : s \in T, s \wedge t = \oslash_{\mathcal{T}}\} = \oslash_{\mathcal{T}}, \end{aligned} \quad (3.8)$$

so that $I_1(\pi) = \oslash_{\mathcal{T}}$ immediately follows. The assertion is proved. □

It is perhaps worth noting explicitly that for non-distributive complete lattices the assertion of Theorem 3.1 need not hold. Indeed, taking the non-distributive complete lattice $\mathcal{T}_0 = \langle \{t_0, t_1, t_2, t_3, t_4\}, \leq \rangle$, introduced in Section 2 above, setting $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\pi(\omega_i) = t_i$ for $i = 1, 2, \pi(\omega_3) = \oslash_{\tau}$, we obtain easily that π is a $\mathcal{T}_0$-possibilistic distribution on Ω, but $I_1(\pi) = \mathbf{1}_{\mathcal{T}_0} > \oslash_{\mathcal{T}_0} = I_2(\pi)$ holds.

Given a complete lattice $\mathcal{T} = \langle T, \leq \rangle$, a $\mathcal{T}$-possibilistic distribution π on Ω is called *orthogonal,* if $\pi(\omega_1) \wedge \pi(\omega_2) = \oslash_{\mathcal{T}}$ for each $\omega_1, \omega_2, \omega_1 \neq \omega_2$, from Ω. As a matter of fact, if $\mathcal{T}$ is completely distributive and π is orthogonal, then $I_2(\pi) = \oslash_{\mathcal{T}}$. Indeed, due to the complete distributivity of $\mathcal{T}$ we obtain, for each $\omega \in \Omega$, that

$$\begin{aligned} &(\Pi(\Omega - \{\omega\})) \wedge \pi(\omega) = \\ &= \left(\bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} \pi(\omega_1) \right) \wedge \pi(\omega) = \\ &= \bigvee_{\omega_1 \in \Omega, \omega_1 \neq \omega} (\pi(\omega_1) \wedge \pi(\omega)) = \oslash_{\mathcal{T}}, \end{aligned} \quad (3.9)$$

so that

$$I_2(\pi) = \bigvee_{\omega\in\Omega} ((\Pi(\Omega - \{\omega\})) \wedge \pi(\omega)) = \oslash_{\mathcal{T}} \qquad (3.10)$$

follows.

Again, both the conditions imposed on $\mathcal{T}$ and π (complete distributivity and orthogonality) can be proved to be necessary introducing appropriate counterexamples.

Theorem 3.2 Let $\mathcal{T} = \langle T, \leq\rangle$ be a complete lattice, let π be a $\mathcal{T}$-possibilistic distribution on a nonempty set Ω such that $\pi(\omega)$ takes the values $\oslash_{\mathcal{T}}$ or $\mathbf{1}_{\mathcal{T}}$ for each $\omega \in \Omega$, then $I_1(\pi) = \oslash_{\mathcal{T}}$ holds. If there is only one $\omega_0 \in \Omega$ such that $\pi(\omega_0) = \mathbf{1}_{\mathcal{T}}$ holds, then $I_2(\pi) = \oslash_{\mathcal{T}}$, otherwise $I_2(\pi) = \mathbf{1}_{\mathcal{T}}$.

Proof. As $\oslash_{\mathcal{T}}^C = \mathbf{1}_{\mathcal{T}}$ and $\mathbf{1}_{\mathcal{T}}^C = \oslash_{\mathcal{T}}$ holds for each complete lattice $\mathcal{T}$, $(\pi(\omega))^C \wedge \pi(\omega) = \oslash_{\mathcal{T}}$ results for each $\omega \in \Omega$, hence, $I_1(\pi) = \oslash_{\mathcal{T}}$ immediately follows. If $\pi(\omega_0) = \mathbf{1}_{\mathcal{T}}, \pi(\omega) = \oslash_{\mathcal{T}}$ for each $\omega \in \Omega$, $\omega \neq \omega_0$, then $\Pi(\Omega - \{\omega_0\}) = \bigvee_{\omega\in\Omega,\omega\neq\omega_0} \pi(\omega) = \oslash_{\mathcal{T}}$, so that $\Pi(\Omega - \{\omega\}) \wedge \pi(\omega) = \oslash_{\mathcal{T}}$ for every $\omega \in \Omega$ and the equality $I_2(\pi) = \oslash_{\mathcal{T}}$ follows. Otherwise, there are at least two $\omega_1, \omega_2 \in \Omega$ such that $\pi(\omega_1) = \pi(\omega_2) = \mathbf{1}_{\mathcal{T}}$, hence, for each $\omega \in \Omega$, either $\omega_1 \in \Omega - \{\omega\}$ or $\omega_2 \in \Omega - \{\omega\}$ holds, so that $\Pi(\Omega - \{\omega\}) = \mathbf{1}_{\mathcal{T}}$ for every $\omega \in \Omega$. Consequently,

$$\begin{aligned} I_2(\pi) &= \bigvee_{\omega\in\Omega} ((\Pi(\Omega - \{\omega\})) \wedge \pi(\omega)) \geq \\ &\geq ((\Pi(\Omega - \{\omega_1\})) \wedge \pi(\omega_1)) \vee \\ &\vee ((\Pi(\Omega - \{\omega_2\})) \wedge \pi(\omega_2)) = \mathbf{1}_{\mathcal{T}} \end{aligned} \qquad (3.11)$$

follows, so that the assertion is proved. □

4 Entropy Values for Independent Products of Lattice-Values Possibilistic Distributions

A well-known and important property of Shannon entropy reads that for the statistically (stochastically) independent product of probability distributions on discrete spaces the entropy of the resulting probability distribution is given by the sum of the entropy values for the particular distributions. More formally, if for both $i = 1, 2$, P_i defines a probability distribution on a nonempty discrete space Ω_i, and if $P_{12}(\omega_1, \omega_2) = P_1(\omega_1)P_2(\omega_2)$ for each $\langle\omega_1, \omega_2\rangle \in \Omega_1 \times \Omega_2$, then the equality $I(P_{12}) = I(P_1) + I(P_2)$ holds (for I^* we obtain that $I^*(P_{12}) = 1-((1-I^*(P_1))\,(1-I^*(P_2))$ holds). Let us prove a possibilistic variant of this relation.

Theorem 4.1 Let $\mathcal{T} = \langle T, \leq\rangle$ be a completely distributive complete lattice and, for both $i = 1, 2$, let π_i be a $\mathcal{T}$-possibilistic distribution on a nonempty set Ω_i. Set $\pi_{12}(\omega_1, \omega_2) = \pi_1(\omega_1) \wedge \pi_2(\omega_2)$ for each $\langle\omega_1, \omega_2\rangle \in \Omega_1 \times \Omega_2$. Then π_{12} defines a $\mathcal{T}$-possibilistic distribution on $\Omega_1 \times \Omega_2$ and, for both $j = 1, 2$, the equality $I_j(\pi_{12}) = I_j(\pi_1) \vee I_j(\pi_2)$ holds.

Proof. Applying the assumption of complete distributivity of $\mathcal{T}$, we obtain easily that

$$\begin{aligned} &\bigvee_{\langle\omega_1,\omega_2\rangle} \pi_{12}(\omega_1, \omega_2) = \\ &= \bigvee_{\omega_1\in\Omega_1} \left(\bigvee_{\omega_2\in\Omega_2} (\pi_1(\omega_1) \wedge \pi_2(\omega_2)) \right) = \\ &= \bigvee_{\omega_1\in\Omega_1} \left(\left(\bigvee_{\omega_2\in\Omega_2} \pi_2(\omega_2) \right) \wedge \pi_1(\omega_1) \right) = \\ &= \bigvee_{\omega_1\in\Omega_1} (\mathbf{1}_{\mathcal{T}} \wedge \pi_1(\omega_1)) = \mathbf{1}_{\mathcal{T}}, \end{aligned} \qquad (4.1)$$

so that π_{12} defines a $\mathcal{T}$-possibilistic distribution on $\Omega_1 \times \Omega_2$. A repeated application of the same assumption yields that

$$\begin{aligned} I_2(\pi_{12}) &= \bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} [(\Pi_{12}(\Omega_1 \times \Omega_2) - \\ &\quad -\{\langle\omega_1, \omega_2\rangle\}) \wedge \pi_{12}(\omega_1, \omega_2)] = \\ &= \bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} \{[\Pi_{12}[((\Omega_1 - \{\omega_1\}) \times \Omega_2) \cup \\ &\quad \cup(\Omega_1 \times (\Omega_2 - \{\omega_2\}))]] \wedge \pi_1(\omega_1) \wedge \pi_2(\omega_2)\} = \\ &= \bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} \{[\Pi_1(\Omega_1 - \{\omega_1\}) \vee \Pi_2(\Omega_2 - \{\omega_2\})] \wedge \\ &\quad \wedge\pi_1(\omega_1) \wedge \pi_2(\omega_2\} = \\ &= \bigvee_{\langle\omega_1,\omega_2\rangle\in\Omega_1\times\Omega_2} (\Pi_1(\Omega_1 - \{\omega_1\}) \wedge \pi_1(\omega_1) \wedge \pi_2)) \vee \\ &\vee \bigvee_{\langle\omega_1,\omega_2\rangle\Omega_1\times\Omega_2} (\Pi_2(\Omega_2 - \{\omega_2\}) \wedge \pi_1(\omega_1) \wedge \pi_2(\omega_2)) = \\ &= \bigvee_{\omega_2\in\Omega_2} \left(\left(\bigvee_{\omega_1\in\Omega_1} (\Pi_1(\Omega_1 - \{\omega_1\}) \wedge \pi_1(\omega_1) \right) \wedge \pi_2(\omega_2) \right) \vee \\ &\vee \bigvee_{\omega_1\in\Omega_1} \left(\left(\bigvee_{\omega_2\in\Omega_2} (\Pi_2(\Omega_2 - \{\omega_2\}) \wedge \pi_2(\omega_2) \right) \wedge \pi_1(\omega_1) \right) = \\ &= \bigvee_{\omega_2\in\Omega_2} (I_2(\pi_1) \wedge \pi_2(\omega_2)) \vee \bigvee_{\omega_1\in\Omega_1} (I_2(\pi_2) \wedge \pi_1(\omega_1)) = \\ &= \left(I_2(\pi_1) \wedge \left(\bigvee_{\omega_2\in\Omega_2} \pi_2(\omega_2) \right) \right) \vee \left(I_2(\pi_2) \wedge \right. \end{aligned}$$

$$\wedge\left(\bigvee_{\omega_1\in\Omega_1}\pi_1(\omega_1)\right)\Big)=$$
$$=I_2(\pi_1)\vee I_2(\pi_2). \tag{4.2}$$

For $I_1(\pi_{12})$ the same equality holds trivially, as in this case $I_1(\pi_{12})=I_1(\pi_1)=I_1(\pi_2)=\oslash_{\mathcal{T}}$ due to Theorem 3.1. The assertion is proved. □

Theorem 4.1 can be easily generalized in this way: let $\mathcal{T},\Omega_i$ and π_i be as in Theorem 4.1, but this time for $i=1,2,\ldots,n$, set $\pi^n(\omega_1,\omega_2,\ldots,\omega_n)=\bigwedge_{i=1}^n\pi_i(\omega_i)$ for each $\langle\omega_1,\omega_2,\ldots,\omega_n\rangle\in\Omega_1\times\Omega_2\times\ldots\times\Omega_n$. Then π^n is a $\mathcal{T}$-possibilistic distribution on $\mathbf{X}_{i=1}^n\Omega_i$ (as can be easily proved by induction) and $I_2(\pi^n)=\bigvee_{i=1}^n I_2(\pi_i)$ holds (for I_1 it is valid trivially). Indeed, for $n=2$ it is proved in Theorem 4.1, suppose that it is valid for $n-1$. Denoting $\mathbf{X}_{i=1}^{n-1}\Omega_i$ by Ω_1^*,Ω_n by Ω_2^*,π^{n-1} by π_1^* and π_n by π_2^*, and applying Theorem 4.1 to Ω_1^*, π_1^*,Ω_2^* and π_2^*, we obtain that

$$\begin{aligned} & I_2(\pi_n)=I_2(\pi^{n-1}\wedge\pi_n)=I_2(\pi_1^*\wedge\pi_2^*)= \\ = \;& I_2(\pi_{12}^*)=I_2(\pi_1^*)\vee I_2(\pi_2^*)= \\ = \;& \left(\bigvee_{i=1}^{n-1}I_2(\pi_i)\right)\vee I_2(\pi_n)=\bigvee_{i=1}^{n}I_2(\pi_i). \end{aligned} \tag{4.3}$$

Remark 4.1 As proved in [4], for real-valued possibilistic distributions the same equality is valid also for *infinite* sequences $\langle\Omega_1,\pi_1\rangle,\langle\Omega_2,\pi_2\rangle,\ldots$ of nonempty spaces and possibilistic distributions.

Acknowledgement

This work was partially supported by grant No. IAA100300503 of GA AS CR and by the Institutional Research Plan AV0Z10300504 "Computer Science for the Information Society: Models, Algorithms, Applications".

References

[1] G. DeCooman, Possibility theory I, II, III, *International Journal of General Systems* 25 (1997) 291-323, 325-351, 353-371.

[2] J. A. Goguen, $\mathcal{L}$-fuzzy sets, *Journal of Mathematical Analysis and Applications* 18 (1967) 145-174.

[3] A. J. Khinchin, Mathematical Foundations of Information Theory. New York, Dover Publications, 1957.

[4] I. Kramosil, Entropy functions defined by possibilistic distributions. Submitted for publication to the Proceedings of the 5th International Symposium on Imprecise Probabilities and Their Applications, Prague, 2007.

[5] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* 1 (1978) 3-28.

Revisiting Bobylev's definitions

Luis J. Rodríguez-Muñiz
Department of Statistics and OR and Mathematical Teaching
University of Oviedo (Spain)
luisj@uniovi.es

Abstract

This communication recalls some of the Bobylev's definitions of support function of a fuzzy set and distance between fuzzy sets. Their equivalence with the generalized Hausdorff metric and the support function defined by of Puri and Ralescu is analyzed.

Keywords: Bobylev's Distance, Fuzzy-valued Mapping, Fuzzy Set, Support Function.

1 Introduction

In 1985 V. N. Bobylev published two papers in which he defined a support function for a fuzzy set, a distance between fuzzy sets and a differential for fuzzy-valued mappings. Their were barely used in the posterior literature (see [3]). In part because it was published in not very known journals (one was originally written in Russian) and in part because, more or less at the same time, Puri and Ralescu definition of the generalized Hausdorff distance was introduced, and it revealed as very useful.

In the present communication, we analyze how the Bobylev's definitions matches the definition of Puri and Ralescu. The work is organized as follows: some notation and preliminaries are introduced in a first section, in the second one we recall Bobylev's definitions of support function and distance between fuzzy sets and in the last section we state our main results analyzing the relationship between Puri and Ralescu's and Bobylev's definitions.

2 Notation and Preliminaries

In this section we gather some concepts about fuzzy sets and fuzzy random variables we need for our study. We start with some set classes.

Let us denote by $\mathcal{K}(\mathbb{R}^n)$ the class of compact subsets of $\mathbb{R}^n$ and by $\mathcal{K}_c(\mathbb{R}^n)$ the convex case. We will denote by $\mathcal{F}(\mathbb{R}^n)$ denote the class of fuzzy subsets $A:\mathbb{R}^n\to[0,1]$ whose α-level sets A_α belong to $\mathcal{K}(\mathbb{R}^n)$ for $\alpha\in[0,1]$ (being for $\alpha\in(0,1]$, $A_\alpha=\{x\in\mathbb{R}^n : A(x)\geq\alpha\}$, and $A_0=\mathrm{cl}\{x\in\mathbb{R}^n : A(x)>0\}$). $\mathcal{F}_c(\mathbb{R}^n)$ will be used when $A_\alpha\in\mathcal{K}_c(\mathbb{R}^n)$, for $\alpha\in[0,1]$.

The class $\mathcal{F}_c(\mathbb{R}^n)$ can be endowed with a semilinear structure, by defining the addition and the product by Zadeh's extension principle ([11]) or, equivalently ([5], [6]), by levelwise calculation as $(A+B)_\alpha=A_\alpha+B_\alpha$ (Minkowski's addition) and $(\lambda A)_\alpha=\lambda A_\alpha$, being $A,B\in\mathcal{F}_c(\mathbb{R}^n)$, $\lambda\in\mathbb{R}$.

Being d_H the Hausdorff metric on $\mathcal{K}(\mathbb{R}^n)$, Puri and Ralescu define the generalized Hausdorff distance as

$$d_\infty(A,B)=\sup_{\alpha\in[0,1]} d_H(A_\alpha,B_\alpha)\,, \tag{1}$$

for $A,B\in\mathcal{F}(\mathbb{R}^n)$. It is well-known that $(\mathcal{F}(\mathbb{R}^n),d_\infty)$ is a complete non-separable metric space and that $(\mathcal{F}_c(\mathbb{R}^n),d_\infty)$ is a closed subspace (see [8] and [4]).

For convex fuzzy sets, Puri and Ralescu introduced the definition of support function (see [7]), by a levelwise extension of the usual definition from convex theory. Hence, being $A\in\mathcal{F}_c(\mathbb{R}^n)$ and S^{n-1} the unit sphere in $\mathbb{R}^n$, the support function of A is given by

$$s_A:[0,1]\times S^{n-1}\to\mathbb{R}\,,$$

with

$$s_A(\alpha,p)=\sup_{x\in A_\alpha}<p,x>\,,$$

being $<\cdot,\cdot>$ the inner product in $\mathbb{R}^n$. In the compact convex case, the supremum is attained.

We can underline another special case of sets in $\mathcal{F}_c(\mathbb{R}^n)$, those having their α-level sets continuous with respect to α, that is the mapping $\alpha\to A_\alpha$ is continuous with respecto to the Euclidean and the d_H metrics. We will denote this class by $\mathcal{F}_{cc}(\mathbb{R}^n)$. It is known (see [10]) that $(\mathcal{F}_{cc}(\mathbb{R}^n),d_\infty)$ is complete and

closed and it can be isometrically embedded into the Hilbert space of continuous real-valued functions from $[0,1] \times S^{n-1}$ with the $\|\cdot\|_\infty$ norm, by the embedding:

$$j : \mathcal{F}_{cc}(\mathbb{R}^n) \to \mathcal{C}([0,1] \times S^{n-1})$$

given by:

$$j(A) = s_A \,.$$

Moreover, $\mathcal{F}_{cc}(\mathbb{R}^n)$ is the greatest subclass of $\mathcal{F}_c(\mathbb{R}^n)$ that can be isometrically embedded into $\mathcal{C}([0,1] \times S^{n-1})$.

On the other hand, we have to remark that when fixing the α-level set, this type of isometrical embedding generalize the case of compact convex subsets of $\mathbb{R}^n$ (see [9]) which holds for every $A \in \mathcal{K}_c(\mathbb{R}^n)$.

3 Bobylev's definitions

Bobylev proposed a different path to reach the definition of the distance between fuzzy sets. He first defined the support function, in a different way than Puri and Ralescu, and, based on the support function, he later introduced the notion of distance.

The Bobylev support function (we will use φ for distinguishing it from the support function by Puri and Ralescu s) of a fuzzy set $A \in \mathcal{F}_c(\mathbb{R}^n)$ is given by

$$\varphi_A : B^n \to \mathbb{R} \,,$$

with

$$\varphi_A(k) = \sup_{\{x \in \mathbb{R}^n : A(x) \geq \|k\|\}} \{< k, x >\} \,,$$

where B^n is the ball in $\mathbb{R}^n$ centered at $\mathbf{0}$ with radius 1. In [1] it is demonstrated that φ_A is unique. Later, Bobylev shows the properties of support function φ_A, being the following: *i*) uppersemicontinuity, *ii*) positive homogeneity, *iii*) subadditiveness, *iv*) normality, *v*) boundedness, and *vi*) $\varphi_A(\mathbf{0}) = 0$.

We will denote by Φ^n the family of all functions $\varphi : B^n \to \mathbb{R}$ satisfying above conditions *i*) to *vi*). Bobylev proves that Φ^n is the set of all support functions of fuzzy sets in $\mathcal{F}_c(\mathbb{R}^n)$, which leads him to identify the fuzzy sets with their support functions.

Denoting Ψ^n the set of all functions $\varphi : B^n \to \mathbb{R}$ satisfying above conditions *v*) and *vi*). The set Ψ^n can be endowed with a linear structure by means of pointwise algebraic operations. Then, a norm can be defined on Ψ^n by:

$$\|\varphi\|_{\Psi^n} = \sup_{x \in B^n \setminus \{\mathbf{0}\}} \left\{ \frac{|\varphi(x)|}{\|x\|} \right\} .$$

It can be checked that Φ^n is a closed subset of Ψ^n and that it is nonseparable in the sense of norm $\|\cdot\|_{\Psi^n}$.

Therefore, by using these tools, Bobylev defines the distance between two fuzzy subsets $A, B \in \mathcal{F}_c(\mathbb{R}^n)$ in the following way:

$$d_\varphi(A,B) = \|\varphi_A - \varphi_B\|_{\Psi^n} \,. \tag{2}$$

Bobylev proves that the space $(\mathcal{F}_c(\mathbb{R}^n), d_\varphi)$ is complete and nonseparable.

Hence, we have two different approaches to the definition of distance between fuzzy sets: the so-called generalized Hausdorff distance introduced by Puri and Ralescu (d_∞) and the distance introduced by Bobylev (d_φ). Both of them generalize the Hausdorff distance in the case of compact convex sets. Also, both metric spaces $(\mathcal{F}_c(\mathbb{R}^n), d_\infty)$ and $(\mathcal{F}_c(\mathbb{R}^n), d_\varphi)$ are complete and nonseparable.

Obviously, the definitions of the support functions are not equivalent in both cases, but two question immediately arise: what is the relationship between the definitions of the support function? and are distances d_∞ and d_φ equivalent?

4 Equivalences

The first step is to analyze the relationship between both definitions of support function. The following results gather this relationship.

Proposition 4.1 *Let $A \in \mathcal{F}_c(\mathbb{R}^n)$. Given s_A its support function, in the sense of Puri and Ralescu ([7]). If we define $\gamma : B^n \to \mathbb{R}$ by*

$$\gamma(k) = \begin{cases} \|k\| \, s_A(\|k\|, \dfrac{k}{\|k\|}), & \text{if } k \neq \mathbf{0}; \\ 0, & \text{if } k = \mathbf{0} \end{cases}$$

then $\gamma = \varphi_A$, being φ_A the support function in the sense of Bobylev ([1]).

Proposition 4.2 *Let $A \in \mathcal{F}_c(\mathbb{R}^n)$. Given φ_A its support function, in the sense of Bobylev ([1]). If we define $\zeta : [0,1] \times S^{n-1} \to \mathbb{R}$ by*

$$\zeta(\alpha, p) = \begin{cases} \dfrac{1}{\alpha} \, \varphi_A(\alpha p), & \text{if } \alpha \neq 0; \\ \sup\limits_{x \in A_0} < x, p >, & \text{if } \alpha = 0; \end{cases}$$

then $\zeta = s_A$, being s_A the support function in the sense of Puri and Ralescu ([7]).

The second step is to analyze the equivalence between the two distances between fuzzy sets.

Theorem 4.3 *Let us consider the class $\mathcal{F}_c(\mathbb{R}^n)$ and the distances d_∞ and d_φ defined, respectively, as in* (1) *and* (2)*. Then it holds for every $A, B \in \mathcal{F}_c(\mathbb{R}^n)$ that:*

$$d_\infty(A,B) = d_\varphi(A,B) \,.$$

Sketch of the Proof. It makes it easier to consider that the set $\{x \in \mathbb{R}^n : A(x) \geq \|k\|\}$ corresponds to the level set $A_{\|k\|}$ for any fixed $k \in B^n$ and $A \in \mathcal{F}_c(\mathbb{R}^n)$. This allows us to rewrite the definition of Bobylev's support function φ_A.

Another important point is the treatment of the case $\mathbf{0} \in \mathbb{R}^n$ (note that the distance d_φ is defined by means of the norm $\|\cdot\|_{\Psi^n}$ that is not defined on $\mathbf{0}$). This case can be easily avoided by using properties of the α-level sets, and the definition of the distance d_∞. In fact, in the original definition of d_∞ by Puri and Ralescu ([8]), the supremum was calculated over $(0, 1]$.

The rest of the proof is obtained by using the definitions of the distance d_φ and the isometry j. By fixing the α-level set, we can use the isometry in the case of $\mathcal{K}_c(\mathbb{R}^n)$ and so we can avoid the problems with the isometry j in $\mathcal{F}_c(\mathbb{R}^n) \setminus \mathcal{F}_{cc}(\mathbb{R}^n)$. $\diamondsuit$

Acknowledgement

The research in this paper has been partially supported by the Spanish Ministry of Education and Science Grant MTM2005-02254. Its financial support is gratefully acknowledged. Author also wish to thank his colleague Antonio Martínez Abejón for his valuable comments.

References

[1] V.N. Bobylev, Support function for a fuzzy set and its characteristic properties, Akademiya Nauk SSSR. Matematicheskie Zametki 37 (1985) 507–513.

[2] V.N. Bobylev, Cauchy problem under fuzzy control, BUSEFAL 21 (1985) 117–126.

[3] P. Diamond, P. Kloeden, Metric Spaces of Fuzzy Sets. Theory and Applications, World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[4] E.P. Klement, M.L. Puri, D.A. Ralescu, Limit theorems for fuzzy random variables, Proceedings Royal Society London. Serie A 407 (1985) 171–182.

[5] M.L. Puri, D.A. Ralescu, Différentielle d'une fonction floue, Comptes Rendus Académie Sciences Paris 293-I (1981) 237–239.

[6] M.L. Puri, D.A. Ralescu, Differentials of fuzzy functions, Journal of Mathematical Analysis and its Applications 91 (1983) 552–558.

[7] M.L. Puri, D.A. Ralescu, The concept of normality for fuzzy random variables, Annals of Probability 13 (1985) 1373–1379.

[8] M.L. Puri, D.A. Ralescu, Fuzzy random variables, Journal of Mathematical Analysis and its Applications 114 (1986) 409–422.

[9] R.T. Rockafellar, Convex Analysis, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ, 1970.

[10] M.A. Rojas-Medar, R.C. Bassanezi, H. Román-Flores, A generalization of the Minkowski embedding theorem and applications, Fuzzy Sets and Systems 102 (1999) 263–269.

[11] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning I, Information Sciences 8 (1975) 199–257.

Session 12

Categories and Topology

On L-valued equalites on the L-powersets of many-valued sets

Ingrida Uljane
University of Latvia LV-1586
uljane@mail.com

Alexander Sostak
University of Latvia LV-1586
sostaks@latnet.lv

Abstract

The research in the field of the so called Fuzzy Mathematics can be conditionally devided into two mainstreams: the first one emphasizes on the study of different fuzzy structures (topological, algebraic, analytical, etc) on an ordinary set X, while for the second one the starting point are L-valued sets, that is a set equipped with some L-valued equality $E : X \times X \to L$ (L being a lattice usually with an additionally algebraic structure). The aim of this work is to discuss the problem how an L-valued equality given on a set X can be extended to the L-valued equality on the L-powerset L^X. This problem, is important, among other for the theory of L-fuzzy topological spaces in the sense of [14], [15].
Keywords: L-relations, L-valued equalities, L-valued sets.

Introduction

In our previous works [16], [17], we have introduced the concept of an L-valued L-topological space, which can be considered as a synthesis of the concept of an L-topological space in the sense of Chang-Goguen [2], [6] and the concept of a many-valued topology in the sense of Höhle [8], see also [9]. Our next aim is to introduce the concept of an L-valued L-*fuzzy* topological space, which would be analogous synthesis of the concept of an L-fuzzy topological space in the sense of the second named author [14], [15], see also [**?**] and the concept of a many-valued topology and to develop the corresponding theory. However, for realizing this plan we have an additional problem. Namely, since L-fuzzy topology on a set X is a mapping $\mathcal{T} : L^X \to L$ (and not a family $\tau \subseteq L^X$ as in case of Chang-Goguen L-topology), and since X is equipped with an L-valued equality $E : X \times X \to L$, it is natural to request some kind of extensionality for a mapping $\mathcal{T} : L^X \to L$. Therefore the problem appears how to "lift" the L-valued equality $E : X \times X \to L$ from X to to an L-valued equality on the L-powerset L^X, that is to get an L-valued equality $\mathcal{E} : L^X \times L^X \to L$.

When studying this problem we found it to be interesting by itself, and just this and some related problems are studied in this work. Besides, we found it reasonable first to study the problem of extension of an L-valued preoder type structures

$$R : X \times X \to L$$

to analogous L-valued preoder type structures

$$\mathcal{R} : L^X \times L^X \to L,$$

and then to "simmetrize" them in order to get L-valued equality on L^X.

1 Prerequisities

Let $(L, \leq, \wedge, \vee)$ be a complete lattice, i.e. $(L, \leq)$ is a partially ordered set such that for every subset $A \subset L$ the join $\bigvee A$ and the meet $\bigwedge A$ are defined. In particular, $\bigvee L =: 1$ and $\bigwedge L =: 0$ are respectively the universal upper and the universal lower bounds in L. We assume that $1 \neq 0$, i.e. L has at least two elements.

Further, let $* : L \times L \longrightarrow L$ be a binary operation on L such that

1. $\alpha * \beta = \beta * \alpha$ for all $\alpha, \beta \in L$;
2. $\alpha * (\beta * \gamma) = (\alpha * \beta) * \gamma$ for all $\alpha, \beta, \gamma \in L$;
3. $\alpha * 1 = \alpha$ and $\alpha * 0 = 0$ for all $\alpha \in L$.
4. $\alpha * (\bigvee_j \beta_j) = \bigvee_j (\alpha * \beta_j) \quad \forall \alpha \in L,$ $\forall \{\beta_j : j \in J\} \subset L$.

In what follows the 5-tuple $(L, \leq, \wedge, \vee, *)$ satisfying the above conditions will be referred to as *commutative cl-monoid* (cf. e.g. [8]).

It is well known that a further binary operation $\mapsto$: $L \times L \mapsto L$ (residuation) is defined on a commutative cl-monoid L which is connected with $*$ by Galois correspondence, that is

$$\alpha * \beta \leq \gamma \mapsto \alpha \leq \beta \mapsto \gamma \text{ for all } \alpha, \beta, \gamma \in L$$

Explicitely residuation $\mapsto$ is given by

$$\alpha \mapsto \beta = \bigvee \{\lambda \in L \mid \alpha * \lambda \leq \beta\}.$$

It is known that the following properties hold in a commutative *cl*-monoid $(L, \leq, \wedge, \vee)$ (cf e.g. [8]):

Proposition 1.1 *Let $\alpha, \beta, \gamma, \alpha_i, \beta_i$ be arbitrary elements from a commutative cl-monoid L. Then:*

1. $\left(\bigvee_{i\in\mathcal{I}} \alpha_i\right) \mapsto \beta) = \bigwedge_{i\in\mathcal{I}} (\alpha_i \mapsto \beta)$;
2. $\alpha \mapsto \left(\bigwedge_{i\in\mathcal{I}} \beta_i\right) = \bigwedge_{i\in\mathcal{I}} (\alpha \mapsto \beta_i)$;
3. *if $\alpha \leq \beta$ then $\alpha \mapsto \beta = 1$;*
4. $\alpha * \beta \leq \alpha \wedge \beta$;
5. $(\alpha \mapsto \beta) * (\beta \mapsto \gamma) \leq \alpha \mapsto \gamma$;
6. $(\alpha * \beta) \mapsto (\gamma \mapsto \delta) \geq (\alpha \mapsto \gamma) * (\beta \mapsto \delta)$.
7. $(\alpha \mapsto \beta) \wedge (\beta \mapsto \alpha) = 1 \Rightarrow \alpha = \beta$.

In what follows $L = (L, \leq, \wedge, \vee, *)$ always denotes a commutative cl-monoid.

2 L-valued preodered sets, category PROSET(L) and some related categories

Definition 2.1 *An L-valued relation (or a fuzzy relation) on a set X is a map $R : X \times X \longrightarrow L$. An L-valued relation R is called*

1. *reflexive if $R(x, x) = 1$ for all $x \in X$;*
2. *transitive, if $R(x, y) * R(y, z) \leq R(x, z)$ for all $x, y, z \in X$;*
3. *symmetric, if $R(x, y) = R(y, x)$ for all $x, y \in X$;*
4. *separated, if $R(x, y) = R(y, x) = 1$ implies that $x = y$ for all $x, y \in X$.*

A transitive L-valued relation is called an L-valued quasipreoder. A reflexive transitive L-valued relation is called an L-valued preoder. A separated L-valued preoder is called an L-valued partial order. A symmetric L-valued preorder is called an L-valued equality. The corresponding pair (X, R) will be refereed to as an L-valued quasipreodered set, L-valued preodered set, an L-valued partially ordered set, and an L-valued set resp.

If R is an L-valued preoder on a set, then given $x, y \in X$ the value $R(x, y)$ is interpreted as the degree to which x is larger than or equal to y. In case R is an L-valued equality on X, the intuitive meaning of the value $R(x, y)$ is the degree to which x and y are equal.

Remark 2.2 L-valued relations, usually in case when $L = [0, 1]$ and when $*$ is a left-semicontinuous t-norm (see e.g. [11]) were considred by many authors and they used different terminology. In particular, a fuzzy relation $R : X \times X \to [0, 1]$ satisfying (1), (2) and (3) is called *a fuzzy equality* in [8], [9] *a fuzzy equivalence* in [10], [12], or an *indistinguishability operator* [**?**]. In [3], [4], [5] a fuzzy relation $R : X \times X \longrightarrow L$ is called *a fuzzy equality* if it satisfies. all conditions (1) – (4).

Examples 2.3

1. Let $X = L$. Then by setting $R(x, y) = x \mapsto y$ we define a canonical L-valued partial oder on X and by setting $E(x, y) = R(x, y) \wedge R(y, x)$ we define a canonical L-valued separated equality on X (cf e.g. [18].)
2. Let (X, ρ) be a pseudo-quasimetric space such that $\rho(x, y) \leq 1$ for all $x, y \in X$. Then by setting $R(x, y) = 1 - \rho(x, y)$ we define an L-valued preoder on X where L is the unit interval [0,1] endowed with the Lukasiewicz conjunction $*$. Moreover, if ρ is a pseudometric, then R is an L-valued equality, and in case ρ is a metric, the L-valued equality R is separated. (cf e.g. [Ho3].)
3. Let $\mathcal{A} \subseteq L^X$ be a family of L-subsets of X. Then,by setting

$$R(\mathcal{A})(x, y) = \bigwedge_{A\in\mathcal{A}} (A(x) \mapsto A(y))$$

we obtain an L-valued preoder on X (see e.g. [**?**])

Definition 2.4 *Given L-valued (quasi)preodered sets (X, R_X) and (Y, R_Y) a mapping $f : X \to Y$ is called extensional if*

$$R_X(x_1, x_2) \leq R_Y(f(x_1), f(x_2)) \text{ for all } x_1, x_2 \in X.$$

L-valued quasi-preodered sets and extensional mappings between them form a category which will be denoted **QPROSET**(L). Its full subcategories consisting of L-valued preodered sets and L-valued sets will be denoted resp. by **PROSET**(L) and **SET**(L). To denote the subcategories of these

categories determined by separated L-valued relation we use notations **SQPROSET**(L), **SPROSET**(L) and **SSET**(L) resp. However for the category of separated L-valued partial sets **SPROSET**(L) which are separated by definition and which play a special role in our work an alternative notation **PAOSET**(L) will be also used. In the sequel our main interest here will be in categories **PROSET**(L) and **PAOSET**(L). Categories **SET**(L) and **SSET**(L) will be discussed in Section 6.

Proposition 2.5 *Let X be a set and $\mathfrak{R}(X,L)$ be the family of L-valued preoders on X. Then $\mathfrak{R}(X,L)$ is a complete lattice. Its bottom $\inf\mathfrak{R}$ is the discrete (or crisp) (L-valued) preoder*

$$R_{dis}(x,y) = \begin{cases} 1 \text{ if } x = y \text{ and} \\ 0 \text{ if } x \neq y. \end{cases}$$

The top $\sup\mathfrak{R}$ of the lattice $\mathfrak{R}(X,L)$ is the indiscrete (L-valued) preoder

$$R_{ind}(x,y) = 1 \text{ for all } x,y \in X.$$

3 L-valued preoder on the L-powerset of an L-valued preodered set

Let (X,R) be an L-valued preodered set. Our first aim is to lift the L-valued preoder R from X to the L-valued quasipreoder $\mathcal{R}$ on the L-powerset L^X of X. We do it as follows.

Given $A, C \in L^X$ we set

$$\mathcal{R}(A,C) = \bigwedge_{x,z\in X} \big(\big(R(x,z) * A(x)\big) \mapsto C(z)\big).$$

Thus we obtain an L-valued relation

$$\mathcal{R} : L^X \times L^X \to L.$$

From 1.1(6) it follows that equivalently $\mathcal{R}(A,C)$ can be defined by

$$\mathcal{R}(A,C) = \bigwedge_{x,z\in X} (R(x,z) \mapsto \big(A(x) \mapsto C(z)\big)\big).$$

Remark 3.1 The and

$$\mathcal{R} : L_E^X \times L_E^X \to L$$

, "defuzzified" meaning of the formulae

$$R(x,z) * A(x) \mapsto C(z)\big) \text{ and } R(x,z) \mapsto (A(x) \mapsto C(z))$$

can be explained as follows: If x is larger than or equal to z and x belongs to A then z belongs to C. In particular, taking $x = z$ we get $A(x) \leq C(x)$ for every $x \in X$. By verifying this condition for all $x, z \in X$ we conclude whether A is larger than or equal to C – this is the "defuzzified" meaning of the value $\mathcal{R}(A,C)$.

In case $A, C \subseteq X$

$$\mathcal{R}(A,C) = \begin{cases} 1 \text{ if } x \in A \text{ and } R(x,z) > 0 \text{ implies } z \in C \\ 0 \text{ otherwise .} \end{cases}$$

In particular, in case R is a crisp preoder $\leq$ on X, then

$$\mathcal{R}(A,C) = 1 \text{ iff } x \in A \text{ and } z \leq x \text{ implies that } z \in C$$

and $\mathcal{R}(A,C) = 0$ otherwise.

Proposition 3.2 *If $R : X \times X \to L$ is an L-valued quasioder on X, then*

$$\mathcal{R}(A,C) \geq \mathcal{R}(A,B) * \mathcal{R}(B,C) \text{ for all } A,B,C \in L^X,$$

and hence $\mathfrak{R} : L^X \times L^X \to L$ is an L-valued quasiorder on L^X.

Unfortunately, even if R is an L-valued preoder, the mapping $\mathcal{R} : L^X \times L^X \to L$ may fail to satisfy axion (1) that is $\mathcal{R}$ generally is not an L-valued preoder. To "remedy" this "defect" (which appears to be natural) one has to restrict the domain of $\mathcal{R}$ by the family of *extensional* L-subsets of X.

Here by extensionality we mean that

$$R(x,z) * A(x) \leq A(z) \text{ for all } x,z \in X,$$

see e.g. [8]. The intuitive "defuzzified" meaning of this condition is the requirement that z should belong to A whenever x belongs to A and z is less than or equal to x.

In what folows we will be usually interested in the properties of the restriction of the L-valued relation $\mathcal{R} : L^X \times L^X \to L$ to the powerset L_E^X of all extensional L-subsets of X. We use the same notation for the restriction of $\mathcal{R}$ to L_E^X that is

$$\mathcal{R} : L_E^X \times L_E^X \to L.$$

Proposition 3.3 *If $R : X \times X \to L$ is an L-valued preoder on X, then*

$$\mathcal{R}(A,A) = 1 \text{ for each } A \in L_E^X$$

and hence

$$\mathcal{R} : L_E^X \times L_E^X \to L$$

is an L-valued preoder on L_E^X.

Proposition 3.4 *The L-valued preorder $\mathcal{R} : L_E^X \times L_E^X \to L$ is separated.*

From Propositions 3.3, 3.2 and 3.4 we get

Theorem 3.5 *If $R : X \times X \to L$ is an L-valued preoder on X then $\mathcal{R} : L^X \times L^X \to L$ is an L-valued quasipreoder on the powerser L^X and an L-valued partial oder on the extensional powerset L_E^X.*

Examples 3.6 In all these examples

$$\mathcal{R} : L^X \times L^X \to L$$

is an L-valued quasipreoder on L^X induced by an L-valued preoder

$$R : X \times X \to L$$

unless specified. By α_X we denote the constant function $\alpha_X : X \to L$ with value $\alpha \in L$.

1. Let $A \in L_E^X$. Then $\mathcal{R}(A, 0_X) = \left(\bigvee_{x\in X} A(x)\right) \to 0$.
2. $\mathcal{R}(A, 1_X) = 1$ for any $A \in L^X$.
3. $\mathcal{R}(1_X, A) = 1 \to \bigwedge_{x\in X} A(x)$.
4. Given $a \in X$ let 1_a stand for the characteristic function of the set $\{a\}$. Then

$$\mathcal{R}(A, 1_a) = \bigvee_{x\neq a} (A(x)) \to 0.$$

   In particular, if $a \neq b, a, b \in X$, then $\mathcal{R}(1_a, 1_b) = 0$.

5. For every $a \in X$ we define an L-set $s_a : X \to L$ by $s_a(x) = R(a, x)$.(This is the so called singleton generated by a).

   Since

$$s_a(x) * R(x, z) = R(a, x) * R(x, z) \leq R(a, z) = s_a(z),$$

   singletons are extensional. Moreover, it is easy to notice that s_a is the smallest one of all extensional L-sets, which are larger than or equal to the L-set 1_a.

   Let $a, b \in X$. Then

$$\mathcal{R}(s_a, s_b) = \bigwedge_{x,z} (R(a, x) * R(x, z) \mapsto R(b, z)) =$$

$$\bigwedge_{z\in X} (R(a, z) \mapsto R(b, z)) \leq$$

$$\leq R(a, b) \mapsto R(b, b) = R(a, b).$$

   On the other hand, since

$$R(a, z) \geq R(a, b) * R(b, z),$$

   from Galois connection we conclude that for every $z \in X$ it holds

$$\mathcal{R}(s_a, s_b) \geq (R(b, z) \mapsto R(a, z)) \geq R(b, a),$$

   and hence in case when R is symmetric,

$$\mathcal{R}(s_a, s_b) = R(a, b).$$

   This means that by assigning to each $a \in X$ its singleton $s_a \in L_E^X$ we may identify (X, R) with the L-valued partially odered subset $(S, \mathcal{R}_S)$ of L-valued partially ordered set $(L_E^X, \mathcal{R})$ where $S = \{s_a : a \in X\}$ and $\mathcal{R}_S$ is the restriction of $\mathcal{R}$ to S.

4 Powerset functor $\Phi : \mathbf{PROSET}(L) \to \mathbf{PAOSET}(L)^{op}$

In this section we show that the construction assigning to an L-valued preodered set (X, R) its extensional powerset $(L_E^X, \mathcal{R})$ can be considered as a contravariant functor Φ from the category $\mathbf{PROSET}(L)$ into the category $\mathbf{PAOSET}(L)$ that is as a functor $\Phi : \mathbf{PROSET}L \to \mathbf{PAOSET}(L)^{op}$ and discuss some properties of this functor. We start with the following

Proposition 4.1 *Let (X, R_X), (Y, R_Y) be L-valued preodered sets and $f : X \to Y$ be an extensional mapping. Then for every $C, D \in L^Y$ it holds*

$$\mathcal{R}_X(f^{-1}(C), f^{-1}(D)) \geq \mathcal{R}_Y(C, D).$$

(Recall that the preimage of an L set is defined by $f^{-1}(C) = c \circ f$) From Proposition 4.1 and Theorem 3.5 we get

Theorem 4.2 *By assigning to each L-valued preodered set*

$$(X, R) \in Ob(\mathbf{PROSET}(L)$$

its extensional powerset $(L_E^X, \mathcal{R})$ and to each extensional mapping

$$f : (X, R_X) \to (Y, R_Y)$$

the mapping

$$f^{\leftarrow} : (L_E^Y, \mathcal{R}_X) \to (L_E^X, \mathcal{R}_Y)$$

we define a functor

$$\Phi : PROSET(L) \to PAOSET(L)^{op}.$$

(Here $f^{\leftarrow}(C) = f^{-1}(C)$ for $C \in L^Y$, cf e.g. [13])

Theorem 4.3 *Functor*

$$\Phi : \mathbf{PROSET}(L) \to \mathbf{PAOSET}(L)^{op}$$

is one-to-one on objects. The restriction Φ' of the functor Φ to $\mathbf{PAOSET}(L)$, that is the functor

$$\Phi' : \mathbf{PAOSET}(L) \to \mathbf{PAOSET}(L)^{op}$$

is an embedding.

Remark 4.4 In a similar way one can consider a functor

$$\tilde{\Phi} : \mathbf{PROSET}(L) \to \mathbf{QPROSET}(L)^{op}$$

assigning to each (X, R) the L-valued quasipreoder set $(L^X, \mathcal{R})$. The image $\tilde{\Phi}(\mathbf{PROSET}(L))$ is a subcategory of the category $\mathbf{QPROSET}(L)^{op}$. We shall not go into details of this construction here.

Remark 4.5 *Functors Φ and $\tilde{\Phi}$ are order reversing.*

Indeed , assume that R_1 and R_2 are two L-valued preoders on X and $R_1 \leq R_2$. Then for any $A, C \in L^X$

$$\mathcal{R}_1(A, C) = \bigwedge_{x,z \in X} \big(R_1(x, z) * A(x) \mapsto C(z)\big) \geq$$

$$\geq \bigwedge_{x,z \in X} \big(R_2(x, z) * A(x) \mapsto C(z)\big) = \mathcal{R}_2(A, C)$$

and hence $\mathcal{R}_1 \geq \mathcal{R}_2$.

5 Lattices $QPR(L^X)$ and $PR(L_E^X)$

Given a set X we denote by $PR(L_E^X)$ the family of all L-valued preoders $\mathcal{R}$ obtained from L-valued preoders R on X. In other words $\mathcal{S} \in PR(L_E^X)$ if and only if $(L_E^X, \mathcal{S}) \in Ob(\Phi(\mathbf{PROSET}(L))$. In a similar way $\mathcal{S} \in QPR(L^X)$ if and only if $(L^X, \mathcal{S}) \in Ob(\Phi(\mathbf{QPROSET}(L))$.

From the previous results it follows, that $QPR(L^X)$ and $PR(L_E^X)$ are bounded lattices where the largest element $\mathcal{R}_{top}$ is induced by discrete (L-valued) preoder R_{dis} on X and the smallest element $\mathcal{R}_{bot}$ is induced by indiscrete L-valued preoder R_{ind} on X. Explicitely, the smallest element is given by:

$$\mathcal{R}_{bot}(A, C) = \bigwedge_{x,z}(A(x) \mapsto C(z)).$$

Note that in case A is extensional, then

$$\mathcal{R}_{bot}(A, A) = \bigwedge_{x}(A(x) \mapsto A(x) = 1,$$

and hence $\mathcal{R}_{bot}$ is an L-valued preoder, but generally $\mathcal{R}_{bot}$ is only a quasi-preoder.

The largest element is given by

$$\mathcal{R}_{top}(A, C) = \bigwedge_{x}(A(x) \mapsto C(x)),$$

and hence $\mathcal{R}_{top}$ is the maximal L-valued preoder both in $QPR(L^X)$ and $PR(L_E^X)$.

Examples 5.1 1. Let $L = [0, 1]$ and $\wedge = *$ in $(L, \leq, \wedge, \vee, *)$, that is $(L, \leq, \wedge, \vee)$ is viewed as a Heyting algebra.
Note that in this case $a \mapsto b = \bigvee\{\lambda \in L \mid a \wedge \lambda \leq b\}$.

(a) Let $R = R_{ind}$ be the indiscrete L-valued preoder on X and $A, C \in L^X$. Then

$$\mathcal{R}(A, C) = \bigvee_{x \in X} A(x) \mapsto \bigwedge_{z \in X} C(z).$$

Hence

$$\mathcal{R}(A, B) = \begin{cases} 1 \text{ if } \sup_x A(x) \leq \inf_x C(x) \text{ and} \\ \inf_x C(x) \text{ otherwise .} \end{cases}$$

In particular, for $A, C \subseteq X$

$$\mathcal{R}(x, y) = \begin{cases} 1 \text{ if } A = \emptyset \text{ or } C = X \\ 0 \text{ otherwise .} \end{cases}$$

(Note that X and $\emptyset$ are the only extensional sets in this case.)

(b) Let $R = R_{dis}$ be the discrete L-valued preoder on X and $A, C \in L^X$. Then

$$\mathcal{R}(A, C) = \bigwedge_{x \in X} (A(x) \mapsto\ C(x)$$

Hence

$$\mathcal{R}(A, C) = \begin{cases} 1 \text{ if } A(x) \leq C(x) \forall x \in X \text{ and} \\ \inf_x\{C(x) \mid x \in X, A(x) \geq C(x) \\ \text{otherwise .} \end{cases}$$

In particular, for $A, C \subseteq X$

$$\mathcal{R}(x, y) = \begin{cases} 1 \text{ if } A \subseteq C \\ 0 \text{ if } A \not\subseteq C. \end{cases}$$

2. Let $L = [0, 1]$ and $*$ be the Łukasiewicz conjunction that is $\alpha * \beta = max\{\alpha + \beta - 1, 0\}$ for $\alpha, \beta \in [0, 1]$ and hence $(L, \leq, \wedge, \vee, *)$ is an MV-algebra.

(a) Let $R = R_{ind}$ be the indiscrete L-valued preoder on X and $A, C \in L^X$. Then

$$\mathcal{R}(A, C) = \bigwedge_{x,z \in X} \big(1 - A(x) + C(z)\big)$$

Hence

$$\mathcal{R}(A, C) = \begin{cases} 1 \text{ if } \sup_x A(x) \leq inf_x C(x) \text{ and} \\ 1 - \sup_{x \in X} A(x) + \inf_{x \in X} C(x) \\ \text{otherwise .} \end{cases}$$

(b) Let $R = R_{dis}$ be the discrete L-valued preoder on X and $A, C \in L^X$. Then

$$\mathcal{R}(A, C) = \bigwedge_{x,z \in X} (R(x, z) * A(x) \mapsto C(z)) =$$

$$= \bigwedge_{x \in X} (A(x) \mapsto\ C(x))$$

Hence

$$\mathcal{R}(A, C) = \begin{cases} 1 \text{ if } A(x) \leq C(x) \forall x \in X \text{ and} \\ \inf_x\{1 - A(x) + C(x)\} \text{ otherwise .} \end{cases}$$

3. Let $L = [0,1]$ and $*$ be the product on $[0,1]$that is $\alpha * \beta = \alpha \cdot \beta$ for $\alpha, \beta \in [0,1]$.

   (a) Let $R = R_{ind}$ be the indiscrete L-valued preoder on X and $A, C \in L^X$. Then

   $$\mathcal{R}(A,C) = \bigvee_{x\in X} A(x) \mapsto \bigwedge_{x\in X} C(x)$$

   Hence

   $$\mathcal{R}(A,C) = \begin{cases} 1 \text{ if } \sup_x A(x) \leq inf_x C(x) \text{ and} \\ \frac{\bigwedge_{x\in X} C(x)}{\bigvee_{x\in X} A(x)} \text{ otherwise .} \end{cases}$$

   (b) Let $R = R_{dis}$ be the discrete L-valued preoder on X and $A, C \in L^X$. Then

   $$\mathcal{R}(A,C) = \bigwedge_{x\in X} (A(x) \mapsto\ C(x)$$

   Hence

   $$\mathcal{R}(A,C) = \begin{cases} 1 \text{ if } A(x) \leq C(x) \forall x \in X \text{ and} \\ \frac{\bigwedge_{x\in X: A(x)\geq C(x)} C(x)}{\bigwedge_{x\in X: A(x)\geq C(x)} A(x)} \text{ otherwise .} \end{cases}$$

6 L-valued equality on the L-powerset of an L-valued set

Let X be a set and $E : X \times X \to L$ be an L-valued equality on X, that is a symmetric preoder. Refering to Section 3 by setting

$$\mathcal{R}(A,C) = \bigwedge_{x,z\in X} \big(E(x,z) \mapsto (A(x) \mapsto C(z)\big)$$

we obtain a separated L-valued preoder on L_E^X (and an L-valued quasipreoder on L^X). Further, let

$$\mathcal{E}(A,C) = \mathcal{R}(A,C) \wedge \mathcal{R}(C,A).$$

Then $\mathcal{E} : L_E^X \times L_E^X \mapsto L$ is an L-valued equality on L_E^X. Indeed, the reflexivity and symmetry of $\mathcal{E}$ are obvious. The transitivity follows from the next series of (in)equalities (see Proposition 1.1):

$$\mathcal{E}(A,B) * \mathcal{E}(B,C) =$$

$$= (\mathcal{R}(A,B) \wedge \mathcal{R}(B,A)) * (\mathcal{R}(B,C) \wedge \mathcal{R}(C,B)) \leq$$

$$\leq (\mathcal{R}(A,B) * \mathcal{R}(B,C)) \wedge (\mathcal{R}(C,B) * \mathcal{R}(B,A)) \leq$$

$$\leq \mathcal{R}(A,C) \wedge \mathcal{R}(C,A) \leq$$

$$\leq \mathcal{R}(A,C) \wedge \mathcal{R}(C,A) = \mathcal{E}(A,C).$$

Hence the pair $(L_E^X, \mathcal{E})$ is a separated L-valued set. Thus, assigning to an L-valued set (X,E) the pair $(L_E^X, \mathcal{E})$ we obtain a functor

$$\Psi : \mathbf{SET}(L) \to \mathbf{SSET}(L)^{op}.$$

One can get results about L-valued equalities on the L-powerset and the funcor Ψ analogous to the results about L-valued preoders on the L-powersets and the functor Φ discussed in sections 3,4 and 5.

The first named author acknowledges the support of the European Social Fund (ESF). The second named author acknowledges the support of the Latvian Council of Science under grant 05.1885. The authors are grateful also to the referees for reading the manuscript carefully and pointing out some minor defects.

References

[1] J. Adamek, H. Herrlich, and G. Strecker, Abstract and Concrete Categories, John Wiley & Sons, Inc. New York et al., 1990.

[2] C.L. Chang, Fuzzy topological spaces, *J. Math. Anal. Appl.*, 24 (1968), 182-190.

[3] B. De Baets, R. Mesiar, Metrics and $\mathcal{T}$-equalities, *J. Math. Anal. Appl.*, **267** (2002), 531–547.

[4] M. De Cock, E.E. Kerre, On (unsuitable) fuzzy relations to model approximate equality *Fuzzy Sets and Syst.*, **133** (2003), 137 - 153.

[5] M. Demirci, Topological properties of a class of generators of an indistinguishability operator, *Fuzzy Sets and Syst.*, **143** (2004), 413 - 426.

[6] J.A. Goguen, The fuzzy Tychonoff theorem, *J. Math. Anal. Appl.*, 43 (1973), 734 - 742.

[7] H. Herrlich, G.E. Strecker, Category Theory, Sigma Series in Pure Math., Heldermann Verlag, Berlin, 1979, 400 p.

[8] U. Höhle, M-valued sets and sheaves over integral commutative *cl*-monoids, In: *Applications of Category Theory to Fuzzy Subsets* S.E. Rodabaugh, E.P. Klement and U. Höhle eds., Kluwer, Dodrecht, Boston, 1992, pp. 33 - 72.

[9] U. Höhle, Many Valued Topology and its Applications, Kluwer Acad. Publ., 2001.

[10] P. Klawon, Castro, Similarity in fuzzy reasoning *Mathware Soft Computing* **2** (1995), 197–228.

[11] E.P. Klement, R. Messiar, E. Pap, Triangular Norms, Trends in Logic, vol. **8**, Kluwer Acad. Publ., Dordrecht, 2000.

[12] E.P. Klement, R. Messiar, E. Pap, Triangular Norms. Position paper II: general constructions and parametrized families *Fuzzy Sets and Syst.*, **145** (2003), 411 - 438.

[13] S.E. Rodabaugh, Powerset operator foundations for poslat fuzzy set theories and topologies In: *Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory , U. Höhle, S.E. Rodabaugh eds.* - Handbook Series, vol.3. Chapter 2, pp. 91 - 116, Kluwer Academic Publisher, Dordrecht, Boston. - 1999.

[14] A. Šostak, On a fuzzy topological structure, *Suppl. Rend. Circ. Matem. Palermo, Ser II,*, **11** (1985), 89-103.

[15] A. Šostak, Two decades of fuzzy topology: basic ideas, notions and results, *Russian Math. Surveys*, **44: 6** (1989), 125-186.

[16] I. Uļjane, A. Šostaks, On a category of L-valued equalities on L-sets, J. of Electrical engineering, Vol. 55, 2004, pp. 60-64.

[17] I. Uļjane, A. Šostaks, On a category of L-valued L-topological spaces, Acta Univ. Latv., vol. 2005, pp.

[18] L. Valverde, On the structure of F-indistinguishability operators *Fuzzy Sets and Syst.*, **17** (1985), 313 - 328.

[19] L.A. Zadeh, Fuzzy sets, *Inform. and Control*, 8 (1965), 338 - 353.

On fuzzification of algebraic and topological structures

Sergejs Solovjovs
Department of Mathematics, University of Latvia
19 Raina Blvd., Riga LV-1586, Latvia
sergejs@lu.lv

Abstract

Given a concrete category $(\mathbf{A}, U)$ over $\mathbf{X}$, the paper presents a fuzzification scheme of algebraic and topological structures over the category $\mathbf{X}(A)$ of $\mathbf{A}$-valued $\mathbf{X}$-objects with a fixed basis A.

Keywords: L-set, topological category, finitary quasivariety, functor-structured category.

1 Introduction

The notion of a fuzzy set introduced by Zadeh in [18] and generalized by Goguen in [5] gave rise to fuzzification of different mathematical structures. In particular, [3, 7, 9, 13] consider fuzzified versions of topological spaces and groups. All aforesaid approaches are fixed-based and use implicitly Goguen category $\mathbf{Set}(L)$ of L-sets for a suitable partially ordered set L (see, e.g., [5, 6]). A variable-basis approach over the category of semi-quantales **SQuant** (which is good enough since the categories of fuzzified structures are topological over their ground categories) is considered in [10, 11].

Given a concrete category $(\mathbf{A}, U)$ over $\mathbf{X}$, we introduced in [15] the category $\mathbf{X}(\mathbf{A})$ of $\mathbf{A}$-valued objects of the category $\mathbf{X}$ as a variable-basis generalization of Goguen category $\mathbf{Set}(L)$. Our next step will be to find a "good" fuzzification machinery over the category $\mathbf{X}(\mathbf{A})$. Following the historical move we start with the fixed-basis approach and therefore consider a subcategory $\mathbf{X}(A)$ (studied thoroughly in [17]) of the category $\mathbf{X}(\mathbf{A})$ for a fixed $\mathbf{A}$-object A. In this paper (a somewhat short version of our manuscript [14]) we present a fuzzification machinery over the category $\mathbf{X}(A)$ using the categories of generalized algebraic and topological structures $\mathbf{Alg}(T)$ and $\mathbf{Spa}(T)$ (see, e.g., [1]). As a result the categories A_τ-$\mathbf{Alg}(T)$ and A_P-$\mathbf{Spa}(T)$ of A-valued algebraic and topological structures arise. Our main concern lies in studying basic properties of these categories. In particular, we show a sufficient condition for both of them to be topological over their ground categories.

The necessary categorical background can be found in [1, 2], however it is expected from the reader to be acquainted with basic concepts of category theory. For convenience sake the first two sections recall basic facts about the categories $\mathbf{X}(\mathbf{A})$ and $\mathbf{X}(A)$.

2 The category $\mathbf{X}(\mathbf{A})$

Let $(\mathbf{A}, U)$ be a concrete category over $\mathbf{X}$ such that the following conditions are fulfilled:

(2CAT) The category $\mathbf{A}$ is a 2-category (see, e.g., [2]).

(ADJ) The functor U has a left adjoint (choose an adjoint situation $(\eta, \epsilon) : F \dashv U : \mathbf{A} \to \mathbf{X}$).

For shortness sake we write $f \overset{\tau}{\Longrightarrow} g$ instead of "there exists a 2-cell $A \overset{f}{\underset{g}{\Downarrow \tau}} B$".

Let $\mathcal{F}$ be the class of all structured arrows with domain in $\mathbf{X}$, i.e., of all triples (X, α, A) with $X \in \mathcal{O}b(\mathbf{X})$, $A \in \mathcal{O}b(\mathbf{A})$ and $X \xrightarrow{\alpha} UA \in \mathcal{M}or(\mathbf{X})$. We introduce the following relation "$\sqsubseteq$" on $\mathcal{F}$.

Definition 2.1. Let $X \underset{\beta}{\overset{\alpha}{\rightrightarrows}} UA$ be elements of $\mathcal{F}$. By (ADJ) one has two $\mathbf{A}$-morphisms $FX \underset{\overline{\beta}}{\overset{\overline{\alpha}}{\rightrightarrows}} A$ with $\overline{(\cdot)} = \epsilon_A \circ F(\cdot)$. Define $\alpha \sqsubseteq \beta$ iff $\overline{\alpha} \overset{\tau}{\Longrightarrow} \overline{\beta}$. ▪

We will use the following properties of relation "$\sqsubseteq$".

Lemma 2.2. *Let $X \xrightarrow{f} Y \underset{\beta}{\overset{\alpha}{\rightrightarrows}} UA \xrightarrow{U\varphi} UB$ be $\mathbf{X}$-morphisms. If $\alpha \sqsubseteq \beta$, then $\alpha \circ f \sqsubseteq \beta \circ f$ and $U\varphi \circ \alpha \sqsubseteq U\varphi \circ \beta$.*

Proof. $\overline{\alpha \circ f} = \epsilon_A \circ F(\alpha \circ f) = \epsilon_A \circ F\alpha \circ Ff = \overline{\alpha} \circ Ff \overset{\tau}{\Longrightarrow} \overline{\beta} \circ Ff = \overline{\beta \circ f}$ and $\overline{U\varphi \circ \alpha} = \epsilon_B \circ F(U\varphi \circ \alpha) = (\epsilon_B \circ FU\varphi) \circ F\alpha = (\varphi \circ \epsilon_A) \circ F\alpha = \varphi \circ \overline{\alpha} \overset{\tau}{\Longrightarrow} \varphi \circ \overline{\beta} = \overline{U\varphi \circ \beta}$. ∎

Lemma 2.3. *Let* $A \overset{\varphi}{\underset{\psi}{\rightrightarrows}} B$ *be* $\mathbf{A}$*-morphisms. Then* $\varphi \overset{\tau}{\Longrightarrow} \psi$ *implies* $U\varphi \sqsubseteq U\psi$.

Proof. Use the fact that $\overline{U\varphi} = \varphi \circ \epsilon_A$. ∎

With the help of Definition 2.1 we introduce the category $\mathbf{X}(\mathbf{A})$.

Definition 2.4. $\mathbf{X}(\mathbf{A})$ is the concrete category over $\mathbf{X} \times \mathbf{A}$, the objects of which (called $\mathbf{A}$*-valued* $\mathbf{X}$*-objects*) are elements of $\mathcal{F}$. Morphisms $(X, \alpha, A) \xrightarrow{(f,\varphi)} (Y, \beta, B)$ are $\mathbf{X} \times \mathbf{A}$-morphisms $(X, A) \xrightarrow{(f,\varphi)} (Y, B)$ such that $U\varphi \circ \alpha \sqsubseteq \beta \circ f$. The underlying functor to $\mathbf{X} \times \mathbf{A}$ is given by $|(X, \alpha, A) \xrightarrow{(f,\varphi)} (Y, \beta, B)| = (X, A) \xrightarrow{(f,\varphi)} (Y, B)$. ∎

We illustrate Definition 2.1 by two examples used later on in the paper (other examples can be found in [15]).

Example 2.5. Let $\mathbf{CSLat}(\bigvee)$ be the category of complete lattices and join-preserving maps; let $\mathbf{Set}$ be the category of sets and maps. One gets a construct $(\mathbf{CSLat}(\bigvee), U)$. Since free objects are just power-sets and the obvious 2-cell structure is induced by the order structure of the lattices one gets the category $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ of lattice-valued subsets of sets considered in [16]. Given two maps $X \overset{\alpha}{\underset{\beta}{\rightrightarrows}} UA$, it follows that $\alpha \sqsubseteq \beta$ iff $\alpha(x) \leqslant \beta(x)$ for every $x \in X$. ∎

Example 2.6. Let $\mathbf{SetRel}$ be the category of sets (as objects) and relations (as morphisms). The forgetful functor $\mathbf{SetRel} \xrightarrow{U} \mathbf{Set}$ is given by $U(X \xrightarrow{\rho} Y) = \mathcal{P}(X) \xrightarrow{f_\rho} \mathcal{P}(Y)$ where $\mathcal{P}(X)$ is the power-set of X and $f_\rho(S) = \{y \in Y \mid \text{there exists } x \in S \text{ such that } x\rho y\}$. Since free objects are just the same sets and the obvious 2-cell structure is induced by set-theoretic inclusion of relations one gets the category $\mathbf{Set}(\mathbf{SetRel})$. Given two maps $X \overset{\alpha}{\underset{\beta}{\rightrightarrows}} UY$, it follows that $\alpha \sqsubseteq \beta$ iff $\alpha(x) \subseteq \beta(x)$ for all $x \in X$. ∎

For the later use we recall the necessary and sufficient conditions for $\mathbf{X}(\mathbf{A})$ to be topological over $\mathbf{X} \times \mathbf{A}$. Notice that the property requires the existence of initial (resp. final) lifts. Accordingly we introduce two sets of requirements (one for each type of lifts). The first set consists of requirements `(CLX)`, `(MPX)` and `(FST)` listed below.

`(CLX)` For every $\mathbf{X} \times \mathbf{A}$-object (X, A), $(\mathbf{X}(X, UA), \sqsubseteq)$ is a complete lattice.

`(MPX)` For every $\mathbf{X}$-morphism $Y \xrightarrow{f} X$ and every $\mathbf{A}$-object A, the function $\mathbf{X}(X, UA) \xrightarrow{-\circ f} \mathbf{X}(Y, UA)$ is meet-preserving.

`(FST)` There exists a functor $\mathbf{A} \xrightarrow{(-)^*} \mathbf{X}^{op}$ ($\mathbf{X}^{op}$ is the dual category of $\mathbf{X}$) such that:

- $A^* = UA$ and $B^* = UB$;
- $1_{UA} \sqsubseteq \varphi^* \circ U\varphi$;
- $U\varphi \circ \varphi^* \sqsubseteq 1_{UB}$;
- for every $\mathbf{A}$-morphism $B \xrightarrow{\varphi} A$ and every $\mathbf{X}$-object X, $\mathbf{X}(X, UA) \xrightarrow{\varphi^* \circ -} \mathbf{X}(X, UB)$ is an order-preserving map.

The second set of requirements is in a sense a dualized version of the first one. Start with the following definition.

Definition 2.7. A category $\mathbf{C}$ is a *complete quasilattice* provided that the quotient category of $\mathbf{C}$ w.r.t. the congruence "$f \sim g$ iff $dom(f) = dom(g)$ and $cod(f) = cod(g)$" is a complete lattice considered as a category. The joins of the lattice are called *quasi-joins*. ∎

Notice that each complete quasilattice is a skeletal category, i.e., isomorphic objects are identical.

With the aforesaid definition in mind introduce the second set of requirements which contains requirements `(CQLA)`, `(QJPA)` and `(FCR)` listed below.

`(CQLA)` For every $\mathbf{X} \times \mathbf{A}$-object (X, A), $\mathbf{A}(FX, A)$ is a complete quasilattice.

`(QJPA)` For every $\mathbf{A}$-morphism $A \xrightarrow{\varphi} B$ and every $\mathbf{X}$-object X, the map $\mathbf{A}(FX, A) \xrightarrow{\varphi \circ -} \mathbf{A}(FX, B)$ is quasijoin-preserving.

`(FCR)` There exists a functor $\mathbf{X} \xrightarrow{(-)^\circ} \mathbf{A}^{op}$ such that:

- $X^\circ = FX$ and $Y^\circ = FY$;
- $Ff \circ f^\circ \overset{\tau}{\Longrightarrow} 1_{FY}$;
- $1_{FX} \overset{\upsilon}{\Longrightarrow} f^\circ \circ Ff$;
- for every $\mathbf{X}$-morphism $X \xrightarrow{f} Y$ and every $\mathbf{A}$-object A, $\mathbf{A}(FX, A) \xrightarrow{-\circ f^\circ} \mathbf{A}(FY, A)$ is a quasijoin-preserving map.

The following theorem is due to [15].

Theorem 2.8. *Suppose* `(2CAT)`, `(ADJ)` *hold. Equivalent are:*

(i) $\mathbf{X}(\mathbf{A})$ *is topological over* $\mathbf{X} \times \mathbf{A}$;

(ii) `(CLX)`, `(MPX)`, `(FST)` *hold;*

(iii) `(CQLA)`, `(QJPA)`, `(FCR)` *hold.*

Proof. For the lack of space we recall only that part of the proof which will be used later.
$(ii) \Longrightarrow (i)$: Show the existence of initial lifts. Let $\mathcal{S} = ((X, A) \xrightarrow{(f_i, \varphi_i)} |(X_i, \alpha_i, A_i)|)_{i\in I}$ be a $|-|$-structured source. Straightforward computations show that $\alpha = \bigwedge_{i\in I} (\varphi_i^* \circ \alpha_i \circ f_i)$ is the required structure.
$(iii) \Longrightarrow (i)$: Show the existence of final lifts. Let $\mathcal{S} = (|(X_i, \alpha_i, A_i)| \xrightarrow{(f_i, \varphi_i)} (X, A))_{i\in I}$ be a $|-|$-costructured sink. Straightforward computations show that $\overline{\alpha} = \bigvee_{i\in I} (\varphi_i \circ \overline{\alpha_i} \circ f_i^\circ)$ is the required structure. ▪

As an immediate consequence one gets that both categories $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ and $\mathbf{Set}(\mathbf{SetRel})$ are topological. The required functors are given in the following examples where one should recall that given a set map $X \xrightarrow{f} Y$ one has two maps (the notation is due to [12]):

- $\mathcal{P}(X) \xrightarrow{f^\rightarrow} \mathcal{P}(Y) : S \mapsto \{f(x) \in Y \mid x \in S\}$;
- $\mathcal{P}(Y) \xrightarrow{f^\leftarrow} \mathcal{P}(X) : S \mapsto \{x \in X \mid f(x) \in S\}$.

Example 2.9. In the category $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ the functor $\mathbf{CSLat}(\bigvee) \xrightarrow{(-)^*} \mathbf{Set}^{op}$ is defined by $(A \xrightarrow{\varphi} B)^* = UB \xrightarrow{\varphi^*} UA$ with $\varphi^*(b) = \bigvee \varphi^\leftarrow(\downarrow b)$. The functor $\mathbf{Set} \xrightarrow{(-)^\circ} \mathbf{CSLat}(\bigvee)^{op}$ is defined by $(X \xrightarrow{f} Y)^\circ = \mathcal{P}(Y) \xrightarrow{f^\leftarrow} \mathcal{P}(X)$. ▪

Example 2.10. In the category $\mathbf{Set}(\mathbf{SetRel})$ the functor $\mathbf{SetRel} \xrightarrow{(-)^*} \mathbf{Set}^{op}$ is defined by $(X \xrightarrow{\rho} Y)^* = \mathcal{P}(Y) \xrightarrow{f^*} \mathcal{P}(X)$ with $f^*(S) = \{x \in X \mid \{y \in Y \mid x\rho y\} \subseteq S\}$. The functor $\mathbf{Set} \xrightarrow{(-)^\circ} \mathbf{SetRel}^{op}$ is defined by $(X \xrightarrow{f} Y)^\circ = Y \xrightarrow{f^\circ} X$ with $yf^\circ x$ iff $f(x) = y$. ▪

3 The category $\mathbf{X}(A)$

Definition 3.1. Fix an object A of the category $\mathbf{A}$. The (nonfull) subcategory $\mathbf{X}(A)$ of the category $\mathbf{X}(\mathbf{A})$ is the concrete category over $\mathbf{X}$, the objects of which (called *A-valued $\mathbf{X}$-objects*) are $\mathbf{X}(\mathbf{A})$-objects (X, α, A) for shortness sake denoted by (X, α). Morphisms $(X, \alpha) \xrightarrow{f} (Y, \beta)$ are $\mathbf{X}$-morphisms $X \xrightarrow{f} Y$ such that $(X, \alpha, A) \xrightarrow{(f, 1_A)} (Y, \beta, A)$ is an $\mathbf{X}(\mathbf{A})$-morphism. The underlying functor to $\mathbf{X}$ is given by $|(X, \alpha) \xrightarrow{f} (Y, \beta)| = X \xrightarrow{f} Y$. ▪

We illustrate Definition 3.1 by the following example.

Example 3.2. Let A be a $\mathbf{CSLat}(\bigvee)$-object. The subcategory $\mathbf{Set}(A)$ of the category $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ coincides with the category of A-sets introduced by Goguen in [5]. ▪

From now on we consider an arbitrary (but fixed) subcategory $\mathbf{X}(A)$ of the category $\mathbf{X}(\mathbf{A})$. In the previous section we gave the necessary and sufficient conditions for the category $\mathbf{X}(\mathbf{A})$ to be topological over $\mathbf{X} \times \mathbf{A}$ (Theorem 2.8). It was crucial to have the functors $\mathbf{A} \xrightarrow{(-)^*} \mathbf{X}^{op}$ and $\mathbf{X} \xrightarrow{(-)^\circ} \mathbf{A}^{op}$. The situation is much simpler in case of the category $\mathbf{X}(A)$. Introduce the following weak versions of requirements `(CLX)`, `(MPX)`.

`(CLXA)` For every $\mathbf{X}$-object X, $(\mathbf{X}(X, UA), \sqsubseteq)$ is a complete lattice.

`(MPXA)` For every $\mathbf{X}$-morphism $Y \xrightarrow{f} X$, the map $\mathbf{X}(X, UA) \xrightarrow{-\circ f} \mathbf{X}(Y, UA)$ is meet-preserving.

Theorem 3.3. *Suppose* `(CLXA)`, `(MPXA)` *hold. Equivalent are:*

(i) $\mathbf{X}(A)$ *is topological over* $\mathbf{X}$*;*

(ii) `(CLXA)`, `(MPXA)` *hold.*

Proof. For the lack of space we recall only that part of the proof which will be used later.
$(ii) \Longrightarrow (i)$: Show the existence of initial lifts. Let $\mathcal{S} = (X \xrightarrow{f_i} |(X_i, \alpha_i)|)_{i\in I}$ be a $|-|$-structured source. Similar to Theorem 2.8 one shows that $\alpha = \bigwedge_{i\in I} (\alpha_i \circ f_i)$ is the desired structure. ▪

Theorems 2.8 and 3.3 imply the following result.

Corollary 3.4. *Suppose* `(2CAT)`, `(ADJ)` *hold. Equivalent are:*

(i) $\mathbf{X}(\mathbf{A})$ *is topological over* $\mathbf{X} \times \mathbf{A}$*;*

(ii) `(FST)` *holds and every subcategory* $\mathbf{X}(A)$ *of the category* $\mathbf{X}(\mathbf{A})$ *is topological over* $\mathbf{X}$*.*

By Corollary 3.4 every subcategory $\mathbf{Set}(A)$ of the category $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ or $\mathbf{Set}(\mathbf{SetRel})$ is topological.

From now on assume that the category $\mathbf{X}(\mathbf{A})$ (and thus by Corollary 3.4 every subcategory of the form $\mathbf{X}(A)$) is topological.

4 Category A_τ-$\mathbf{Alg}(T)$

This section considers a fuzzification scheme for algebraic structures over the category $\mathbf{X}(A)$. In order not to restrict ourselves to a particular class of algebras we

use the objects of the category $\mathbf{Alg}(T)$ as a sufficient generalization of the notion of abstract algebra. For convenience sake we recall its definition from [1].

Definition 4.1. Let $\mathbf{X} \xrightarrow{T} \mathbf{X}$ be a functor. $\mathbf{Alg}(T)$ is the concrete category over $\mathbf{X}$, the objects of which (called *T-algebras*) are pairs (X, x) with X an $\mathbf{X}$-object and $TX \xrightarrow{x} X$ an $\mathbf{X}$-morphism. Morphisms $(X, x) \xrightarrow{f} (X', x')$ (called *T-homomorphisms*) are $\mathbf{X}$-morphisms $X \xrightarrow{f} X'$ such that $x' \circ Tf = f \circ x$. The underlying functor to $\mathbf{X}$ is given by $|(X, x) \xrightarrow{f} (X', x')| = X \xrightarrow{f} X'$. ∎

From now on we consider an arbitrary (but fixed) category of the form $\mathbf{Alg}(T)$. The next definition introduces a fuzzification scheme of $\mathbf{Alg}(T)$ over the category $\mathbf{X}(A)$. For the sake of generality instead of the whole category $\mathbf{Alg}(T)$ we consider its subcategory $\mathbf{B}$.

Definition 4.2. Let $\mathbf{B}$ be a subcategory of the category $\mathbf{Alg}(T)$. For a fixed $\mathbf{Alg}(T)$-object (UA, τ), the category A_τ-$\mathbf{B}$ is the concrete category over $\mathbf{B}$, the objects of which (called A_τ-(T-*algebras*)) are triples (X, x, α) with (X, x) a $\mathbf{B}$-object and (X, α) an $\mathbf{X}(A)$-object such that $\tau \circ T\alpha \sqsubseteq \alpha \circ x$. Morphisms $(X, x, \alpha) \xrightarrow{f} (Y, y, \beta)$ (called A_τ-(T-*homomorphisms*)) are $\mathbf{X}$-morphisms $X \xrightarrow{f} X'$ with $(X, x) \xrightarrow{f} (Y, y)$ a $\mathbf{B}$-morphism and $(X, \alpha) \xrightarrow{f} (Y, \beta)$ an $\mathbf{X}(A)$-morphism. The underlying functor to $\mathbf{B}$ is given by $|(X, x, \alpha) \xrightarrow{f} (Y, y, \beta)| = (X, x) \xrightarrow{f} (Y, y)$. ∎

We show a sufficient condition for A_τ-$\mathbf{B}$ to be topological over $\mathbf{B}$. Introduce the following requirement:

(OPR) The map $\mathbf{X}(X, UA) \xrightarrow{\tau \circ T(-)} \mathbf{X}(TX, UA)$ is order-preserving for every $\mathbf{X}$-object X.

Theorem 4.3. *Suppose* (OPR) *holds. Then A_τ-$\mathbf{B}$ is topological over $\mathbf{B}$.*

Proof. Let $\mathcal{S} = ((X, x) \xrightarrow{f_i} |(X_i, x_i, \alpha_i)|)_{i \in I}$ be a $|-|$-structured source. Define $\alpha = \bigwedge_{i \in I} (\alpha_i \circ f_i)$. By Theorem 3.3 it is enough to show that (X, x, α) is a A_τ-$\mathbf{B}$-object. Since $\tau \circ T(\alpha_i \circ f_i) = \tau \circ T\alpha_i \circ Tf_i \sqsubseteq \alpha_i \circ x_i \circ Tf_i = \alpha_i \circ f_i \circ x$ for $i \in I$, it follows that $\tau \circ T\alpha = \tau \circ T(\bigwedge_{i \in I} (\alpha_i \circ f_i)) \sqsubseteq \bigwedge_{i \in I} \tau \circ T(\alpha_i \circ f_i) \sqsubseteq \bigwedge_{i \in I} (\alpha_i \circ f_i \circ x) = \bigwedge_{i \in I} (\alpha_i \circ f_i) \circ x = \alpha \circ x$. ∎

We are going to consider a particular category A_τ-$\mathbf{B}$ and therefore introduce the following requirement.

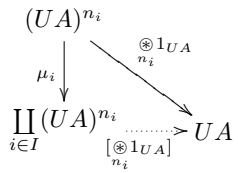
(CFP) The category $\mathbf{X}$ has coproducts as well as finite products (denoted by $((\mu_i)_{i \in I}, \coprod_{i \in I} X_i)$ and $(\prod_{i \in I} X_i, (\pi_i)_{i \in I})$ respectively).

Since $\mathbf{X}$ has finite products and $\mathbf{X}(A)$ is topological over $\mathbf{X}$, the following result holds.

Proposition 4.4. *The category $\mathbf{X}(A)$ has finite products constructed as follows. Let $((X_i, \alpha_i))_{i \in I}$ be a finite family of $\mathbf{X}(A)$-objects and let $(\prod_{i \in I} X_i, (\pi_{X_i})_{i \in I})$ be a product of the family $(X_i)_{i \in I}$ in $\mathbf{X}$. Then the source $((\prod_{i \in I} X_i, \circledast_{i \in I} \alpha_i) \xrightarrow{\pi_{X_i}} (X_i, \alpha_i))_{i \in I}$ with $\circledast_{i \in I} \alpha_i = \bigwedge_{i \in I} (\alpha_i \circ \pi_{X_i})$ is a product of the family $((X_i, \alpha_i))_{i \in I}$.*

Consider an example for Definition 4.2 (cf. [4]).

Example 4.5. Suppose (CFP) holds. Let $\Omega = (n_i)_{i \in I}$ be a family of natural numbers indexed by a set I. For $i \in I$ let $\mathbf{X} \xrightarrow{(-)^{n_i}} \mathbf{X} : X \xrightarrow{f} Y \mapsto X^{n_i} \xrightarrow{f^{n_i}} Y^{n_i}$. Let $\mathbf{X} \xrightarrow{T_\Omega} \mathbf{X} = \mathbf{X} \xrightarrow{\coprod_{i \in I} (-)^{n_i}} \mathbf{X}$ and get the category $\mathbf{Alg}(T_\Omega)$. Take the $\mathbf{Alg}(T_\Omega)$-object (UA, τ_Ω) where $TUA \xrightarrow{\tau_\Omega} UA = TUA \xrightarrow{[\circledast_{n_i} 1_{UA}]} UA$ is defined by the commutativity of the triangle



for every $i \in I$. One has the category A_{τ_Ω}-$\mathbf{Alg}(T_\Omega)$ of categorical representations of universal algebras. ∎

We want to study basic properties of Example 4.5. For our investigation we need some new notions. Start by extending the relation "$\sqsubseteq$" from elements of $\mathcal{F}$ to arbitrary U-structured sinks.

Definition 4.6. Let $\mathcal{S} = (X_i \xrightarrow{\alpha_i} UA)_{i \in I}$ and $\mathcal{T} = (X_i \xrightarrow{\beta_i} UA)_{i \in I}$ be U-structured sinks. Then $\mathcal{S} \sqsubseteq \mathcal{T}$ provided that $\alpha_i \sqsubseteq \beta_i$ for $i \in I$. ∎

With the help of Definition 4.6 introduce a modification of the notion of epi-sink.

Definition 4.7. A sink $\mathcal{S} = (X_i \xrightarrow{f_i} Y)_{i \in I}$ in $\mathbf{X}$ is said to be a $\sqsubseteq$-*epi-sink* provided that for every two elements $Y \overset{\alpha}{\underset{\beta}{\rightrightarrows}} UA$ of $\mathcal{F}$, $\alpha \circ \mathcal{S} \sqsubseteq \beta \circ \mathcal{S}$ implies $\alpha \sqsubseteq \beta$. ∎

We continue with the following notation (recall requirement (FCR)): given a sink $\mathcal{S} = (X_i \xrightarrow{f_i} X)_{i \in I}$ in $\mathbf{X}$ let $\bigvee(F\mathcal{S} \circ \mathcal{S}^\circ) = \bigvee_{i \in I} (Ff_i \circ f_i^\circ)$.

Definition 4.8. Let $\mathbf{X}(\mathbf{A})$ be topological over $\mathbf{X} \times \mathbf{A}$ and let $\mathcal{S} = (X_i \xrightarrow{f_i} X)_{i \in I}$ be a sink in $\mathbf{X}$. Say that $\mathcal{S}$ is $\sqsubseteq$-*compatible* provided that $\bigvee(F\mathcal{S} \circ \mathcal{S}^\circ) = 1_{FX}$. ∎

Straightforward computations (omitted for the lack of space) show the following proposition.

Proposition 4.9. *Suppose* (CFP) *holds and let* $\mathbf{X}(\mathbf{A})$ *be topological with* $\sqsubseteq$*-compatible coproducts. If* $\mathbf{B}$ *is a subcategory of* $\mathbf{Alg}(T_\Omega)$*, then the category* A_{τ_Ω}*-*$\mathbf{B}$ *is topological over* $\mathbf{B}$*.*

Recall from [1] that constructs that are concretely isomorphic to implicational subconstructs of $\mathbf{Alg}(\Omega)$ (which is simply $\mathbf{Alg}(T_\Omega)$ with $\mathbf{X} = \mathbf{Set}$) for some Ω are called *finitary quasivarieties*. In particular, the constructs **Vec**, R-**Mod**, **Ab**, **Grp**, **Mon**, **Sgr**, **Rng**, **Lat** of vector spaces, left R-modules, abelian groups, groups, monoids, semigroups, rings and lattices are finitary quasivarieties.

Corollary 4.10. *Let* $\mathbf{Set}(\mathbf{A})$ *be topological with* $\sqsubseteq$*-compatible coproducts. Let* A *be an* $\mathbf{A}$*-object and let* $\mathbf{B}$ *be a finitary quasivariety. Then* A_{τ_Ω}*-*$\mathbf{B}$ *is topological over* $\mathbf{B}$*.*

Corollary 4.11. *For the categories* $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ *and* $\mathbf{Set}(\mathbf{SetRel})$*, each of the following categories is topological:* A_{τ_Ω}*-***Vec***,* A_{τ_Ω}*-(*R*-***Mod***),* A_{τ_Ω}*-***Ab***,* A_{τ_Ω}*-***Grp***,* A_{τ_Ω}*-***Mon***,* A_{τ_Ω}*-***Sgr***,* A_{τ_Ω}*-***Rng***,* A_{τ_Ω}*-***Lat***.*

5 Category A_P-$\mathbf{Spa}(T)$

This section considers a fuzzification scheme for topological structures over the category $\mathbf{X}(A)$. In order not to restrict ourselves to a particular class of structures we use the objects of the *functor-structured category* $\mathbf{Spa}(T)$ as a sufficient generalization of the notion of abstract topological structure. For convenience sake we recall its definition from [1].

Definition 5.1. Let $\mathbf{X} \xrightarrow{T} \mathbf{Set}$ be a functor. $\mathbf{Spa}(T)$ is the concrete category over $\mathbf{X}$, the objects of which (called T*-spaces*) are pairs (X, α) with $\alpha \subseteq TX$. Morphisms $(X,\alpha) \xrightarrow{f} (Y,\beta)$ (called T*-maps*) are $\mathbf{X}$-morphisms $X \xrightarrow{f} Y$ such that $(T(f))^{\rightarrow}(\alpha) \subseteq \beta$. The underlying functor to $\mathbf{X}$ is given by $|(X,\alpha) \xrightarrow{f} (Y,\beta)| = X \xrightarrow{f} Y$. ∎

From now on we consider an arbitrary (but fixed) category of the form $\mathbf{Spa}(T)$. The next definition introduces a fuzzification scheme of $\mathbf{Spa}(T)$ over the category $\mathbf{X}(A)$. For the sake of generality instead of the whole category $\mathbf{Spa}(T)$ we consider its subcategory $\mathbf{B}$.

Definition 5.2. Let $\mathbf{B}$ be a subcategory of the category $\mathbf{Spa}(T)$. For a fixed $\mathbf{Spa}(T)$-object (UA, P), the category A_P-$\mathbf{B}$ is the concrete category over $\mathbf{B}$, the objects of which (called A_P-(T*-spaces*)) are triples (X, S, α) where (X, S) is a $\mathbf{B}$-object and (X,α) is an $\mathbf{X}(A)$-object such that $(X,S) \xrightarrow{\alpha} (UA, P)$ is a $\mathbf{Spa}(T)$-morphism. Morphisms $(X,S,\alpha) \xrightarrow{f} (Y,Q,\beta)$ (called A_P-(T*-maps*)) are $\mathbf{X}$-morphisms $X \xrightarrow{f} Y$ with $(X,S) \xrightarrow{f} (Y,Q)$ a $\mathbf{B}$-morphism and $(X,\alpha) \xrightarrow{f} (Y,\beta)$ an $\mathbf{X}(A)$-morphism. The underlying functor to $\mathbf{B}$ is given by $|(X,S,\alpha) \xrightarrow{f} (Y,Q,\beta)| = (X,S) \xrightarrow{f} (Y,Q)$. ∎

As in the previous section we are going to show a sufficient condition for A_P-$\mathbf{B}$ to be topological over $\mathbf{B}$. Introduce a new requirement (recall condition (CLXA)):

(MC) For every object (X,S) of the category $\mathbf{B}$, the hom-set $\mathbf{Spa}(T)((X,S),(UA,P))$ is closed under the formation of meets.

Theorem 5.3. *Suppose* (MC) *holds. Then* A_P-$\mathbf{B}$ *is topological over* $\mathbf{B}$*.*

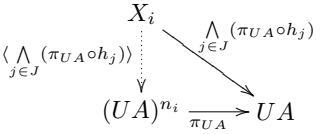
Proof. Let $\mathcal{S} = ((X,S) \xrightarrow{f_i} |(X_i, S_i, \alpha_i)|)_{i\in I}$ be a $|-|$-structured source. Define $\alpha = \bigwedge_{i\in I} (\alpha_i \circ f_i)$. It is enough to show that (X,S,α) is an A_P-$\mathbf{B}$-object and that follows from the fact that $(X,S) \xrightarrow{\alpha_i \circ f_i} (UA, P)$ is a $\mathbf{Spa}(T)$-morphism for $i \in I$ together with (MC). ∎

Consider an example of the category $\mathbf{Spa}(T)$.

Example 5.4. Suppose $\mathbf{X}$ has finite products. Let $\Omega = (n_i)_{i\in I}$ be a family of natural numbers indexed by a set I and let $(X_i)_{i\in I}$ be a family of $\mathbf{X}$-objects. For $i \in I$ let $\mathbf{X} \xrightarrow{(-)^{n_i}} \mathbf{X} : X \xrightarrow{f} Y \mapsto X^{n_i} \xrightarrow{f^{n_i}} Y^{n_i}$ and let $\mathbf{X} \xrightarrow{T_i} \mathbf{Set} = \mathbf{X} \xrightarrow{(-)^{n_i}} \mathbf{X} \xrightarrow{hom(X_i,-)} \mathbf{Set}$ where $hom(X_i, -)$ is the *covariant hom-functor*. Let $\mathbf{X} \xrightarrow{T_\Omega} \mathbf{Set} = \mathbf{X} \xrightarrow{\coprod_{i\in I} T_i} \mathbf{Set}$ and get the category $\mathbf{Spa}(T_\Omega)$. ∎

We are interested in basic properties of Example 5.4. Let $P_\Omega \subseteq (T_\Omega \circ U)(A) = \coprod_{i\in I} hom(X_i, (UA)^{n_i})$ be closed under the formation of meets in such a way that the following condition is fulfilled:

($\mathfrak{A}$) For arbitrary $i \in I$ and arbitrary subfamily $(X_i \xrightarrow{h_j} (UA)^{n_i})_{j\in J} \subseteq P_\Omega$, the $\mathbf{X}$-morphism $X_i \xrightarrow{\langle \bigwedge_{j\in J} (\pi_{UA} \circ h_j)\rangle} (UA)^{n_i}$ defined by the commutativity of the triangle



for every projection π_{UA}, belongs to P_Ω.

One gets the category A_{P_Ω}-$\mathbf{Spa}(T_\Omega)$. The next proposition shows a property of ($\mathfrak{A}$) (we omit straightforward computations again).

Proposition 5.5. *In case of the functor T_Ω,* $(\mathfrak{A})$ *implies* (MC).

Consider an interesting special case of Example 5.4.

Example 5.6. Let $\mathbf{X}(\mathbf{A})$ be $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$ and let $X_i = \{\star\}$ for $i \in I$. Then $T_i^\star(X) = hom(\{\star\}, X^{n_i}) \cong X^{n_i}$. One can easily see that $\mathbf{Spa}(T_\Omega^\star)$ is concretely isomorphic to the category $\mathbf{Rel}(\Omega)$ with objects all pairs $(X, (\rho_i)_{i\in I})$ where X is a set and $\rho_i \subseteq X^{n_i}$ for $i \in I$ and morphisms all maps $(X, (\rho_i)_{i\in I}) \xrightarrow{f} (Y, (\sigma_i)_{i\in I})$ where $(f^{n_i})^\rightarrow(\rho_i) \subseteq \sigma_i$ for $i \in I$.

Let $\mathbf{B}$ be a subcategory of $\mathbf{Spa}(T_\Omega^\star)$ (e.g., if $\Omega = (2)$ consider the category $\mathbf{Rel}$ of sets and relations, the category $\mathbf{Prost}$ of preordered sets or the category $\mathbf{Pos}$ of partially ordered sets) and let $(UA, (\rho_i)_{i\in I})$ be a $\mathbf{Spa}(T_\Omega^\star)$-object such that ρ_i is closed under the formation of meets for $i \in I$, i.e., $a_j \rho_i b_j$ for $j \in J$ imply $(\bigwedge_{j\in J} a_j)\rho_i(\bigwedge_{j\in J} b_j)$. By Proposition 5.5 $A_{P_\Omega^\star}$-$\mathbf{B}$ is topological over $\mathbf{B}$. ∎

For $\Omega = (2)$ Example 5.6 gives the category A_ρ-$\mathbf{Pos}$ of fuzzified posets (topological over $\mathbf{Pos}$). The objects of the category are triples $(X, \leqslant, \alpha)$, where $(X, \leqslant)$ is a poset, (X, α) is an A-set and, moreover, $x \leqslant y$ implies $\alpha(x) \rho\, \alpha(y)$. The fuzzification differs from the concept of *fuzzy partial order* (see, e.g., [8]).

At the end of the paper we briefly consider the case of *functor-costructured categories* of [1], i.e., categories of the form $(\mathbf{Spa}(T))^{op}$ for functors $\mathbf{X}^{op} \xrightarrow{T} \mathbf{Set}$ concrete over $\mathbf{X} = (\mathbf{X}^{op})^{op}$. In a natural way the category $A_{P^{op}}$-$\mathbf{B}$ arises (see Definition 5.2). One can easily restate Theorem 5.3 for the case of the category $A_{P^{op}}$-$\mathbf{B}$.

Let $\mathbf{Set} \xrightarrow{\mathcal{Q}} \mathbf{Set}$ be the contravariant powerset functor. By Example 22.7(2) in [1] the category $\mathbf{Top}$ of topological spaces is a full subcategory of $(\mathbf{Spa}(\mathcal{Q}))^{op}$. This motivates our next example.

Example 5.7. Suppose $\mathbf{X}(\mathbf{A})$ is $\mathbf{Set}(\mathbf{CSLat}(\bigvee))$. Let A be a $\mathbf{CSLat}(\bigvee)$-object and let $(UA, P_\downarrow)$ be a $(\mathbf{Spa}(\mathcal{Q}))^{op}$-object fulfilling the next condition:

($\mathfrak{B}$) Every $V \in P_\downarrow$ has the following properties:

(i) $\downarrow V = V$;

(ii) $\bigwedge S \in V$ implies $S \bigcap V \neq \varnothing$ for every nonempty subset $S \subseteq A$.

One gets the category $A_{P_\downarrow^{op}}$-$\mathbf{Top}$. For instance, let $\tilde{P}_\downarrow = \{A \setminus \uparrow a \,|\, a \in A\}$. Then $(UA, \tilde{P}_\downarrow)$ satisfies ($\mathfrak{B}$). The objects of the category $A_{\tilde{P}_\downarrow^{op}}$-$\mathbf{Top}$ are triples $(X, \mathcal{O}(X), \alpha)$ with $(X, \mathcal{O}(X))$ a topological space, (X, α) an A-set and $(\alpha^\leftarrow)^\rightarrow(\tilde{P}_\downarrow) \subseteq \mathcal{O}(X)$. ∎

Proposition 5.8. $A_{P_\downarrow^{op}}$-$\mathbf{Top}$ *is topological over* $\mathbf{Top}$.

Proof. Take a $\mathbf{Top}$-object $(X, \mathcal{O}(X))$ and consider a family of $(\mathbf{Spa}(\mathcal{Q}))^{op}$-morphisms $((X, \mathcal{O}(X)) \xrightarrow{\alpha_i} (UA, P_\downarrow))_{i\in I}$. Define $\alpha = \bigwedge_{i\in I} \alpha_i$ and show that $(X, \mathcal{O}(X)) \xrightarrow{\alpha} (UA, P_\downarrow)$ is a $(\mathbf{Spa}(\mathcal{Q}))^{op}$-morphism (then use Theorem 5.3). If $I = \varnothing$, then $\alpha \equiv \top := \bigvee A$ and for every $V \in P_\downarrow$, $\alpha^\leftarrow(V)$ is either X for $\top \in V$ or is $\varnothing$ otherwise. Suppose $I \neq \varnothing$ and take $V \in P_\downarrow$. Show that $\alpha^\leftarrow(V) = \bigcup_{i\in I} \alpha_i^\leftarrow(V)$ (the latter set is an element of $\mathcal{O}(X)$). If $x \in \alpha^\leftarrow(V)$, then $\bigwedge_{i\in I} \alpha_i(x) \in V$ and therefore there exists $i_0 \in I$ such that $\alpha_{i_0}(x) \in V$ by ($\mathfrak{B}$).(ii). Conversely, $x \in \bigcup_{i\in I} \alpha_i^\leftarrow(V)$ gives $i_0 \in I$ with $\alpha_{i_0}(x) \in V$ and therefore $\bigwedge_{i\in I} \alpha_i(x) \in V$ by ($\mathfrak{B}$).(i). ∎

6 Conclusion

In the paper we presented a fuzzification machinery for algebraic and topological structures. It will be the topic of our forthcoming studies to disclose its relations to other fuzzification machineries in the literature. One can see, however, at once that the aforesaid category $A_{P^{op}}$-$\mathbf{Top}$ differs from the category L-$\mathbf{Top}$ of L-topological spaces for a semi-quantale L (see, e.g., [10]) where objects are pairs (X, τ) with X a set and $\tau \subseteq L^X$ closed under $\otimes$ and arbitrary $\bigvee$. On the other hand, the objects of the category A_{τ_Ω}-$\mathbf{Grp}$ coincide with fuzzy subgroups introduced by Rosenfeld in [13]. Thus the algebraic side of our paper is more consistent with the literature.

Acknowledgement

This research was supported by the European Social Fund.

References

[1] Jiří Adámek, Horst Herrlich, and George E. Strecker, Abstract and Concrete Categories: The Joy of Cats, John Wiley & Sons, 1990.

[2] Francis Borceux, Handbook of Categorical Algebra 1: Basic Category Theory, Cambridge University Press, 1994.

[3] C. L. Chang, Fuzzy topological spaces, J. Math. Anal. Appl. 24 (1968) 182–190.

[4] Paul M. Cohn, Universal Algebra, D. Reidel Publ. Comp., 1981.

[5] J. A. Goguen, L-fuzzy sets, J. Math. Anal. Appl. 18 (1967) 145–174.

[6] J. A. Goguen, Categories of V-sets, Bull. Am. Math. Soc. 75 (1969) 622–624.

[7] U. Höhle and A. P. Šostak, Axiomatic foundations of fixed-basis fuzzy topology, in: U. Höhle, S. E. Rodabaugh (Eds.), *Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory*. Kluwer Academic, Boston, 1999, pp. 123–272.

[8] Hongliang Lai and Dexue Zhang, Fuzzy preorder and fuzzy topology, Fuzzy Sets Syst. 157 (2006) 1865–1885.

[9] R. Lowen, Fuzzy topological spaces and fuzzy compactness, J. Math. Anal. Appl. 56 (1976) 621–633.

[10] S. E. Rodabaugh, Relationship of algebraic theories to powerset theories and fuzzy topological theories for lattice-valued mathematics, to appear in Int. J. Math. Math. Sci.

[11] S. E. Rodabaugh, Categorical foundations of variable-basis fuzzy topology, in: U. Höhle, S. E. Rodabaugh (Eds.), *Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory*. Kluwer Academic, Boston, 1999, pp. 273–388.

[12] S. E. Rodabaugh, Powerset operator foundations for poslat fuzzy set theories and topologies, in: U. Höhle, S. E. Rodabaugh (Eds.), *Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory*. Kluwer Academic, Boston, 1999, pp. 91–116.

[13] Azriel Rosenfeld, Fuzzy groups, J. Math. Anal. Appl. 35 (1971) 512–517.

[14] Sergey Solovyov, On fuzzification of algebraic and topological structures, submitted to Fuzzy Sets Syst.

[15] Sergey Solovyov, Categories of lattice-valued sets as categories of arrows, Fuzzy Sets Syst. 157(6) (2006) 843–854.

[16] Sergey Solovyov, On the category Set(JCPos), Fuzzy Sets Syst. 157(3) (2006) 459–465.

[17] Sergey Solovyov, On a generalization of Goguen's category Set(L), Fuzzy Sets Syst. 158(4) (2007) 367–385.

[18] L. A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–365.

Fuzzy order relation and fuzzy ordered set category

Olga Lebedeva
Raina bulvaris 29, Riga, LV-1459, Latvia
olga.lebedeva@parex.lv

Abstract

The paper is devoted to a concept of a fuzzy ordered set and partially ordered set category fuzzyfication. The aim of this work is to define a fuzzy order relation on a usual set, on an L-valued set and on an L-set. Using the fuzzy order relation we construct fuzzy POS category and study the connection between this categories and SET(L).

Keywords: Fuzzy order relation, L-valued set, POS category.

1 Introduction

Since the first introduction of the concepts of fuzzy set theory by Lotfi A. Zadeh in the second half of 20th century, fuzzy analogues of basic concepts of classical mathematics were investigated. In this paper, we introduce the notion of a fuzzy order relation. Interests of many researchers have been directed to the study of fuzzy relations including fuzzy order relation [6]. However, as far as we know up to now there has been no paper where the categorical aspects of fuzzy order relation would be studied. This fuzzyfication generalizes the classical case, t.i., if the underlying GL-monoid consists only of two elements then the induced fuzzy order relation is equal to the classical order relation. To realize the construction of fuzzyficated POS category we need to follow the construction of usual POS category. We involve fuzzy ordered sets and define order-preserving (in the fuzzy sense) mappings between them. These sets and mappings form the L-POS category. Later we discuss the relation between L-POS category and SET(L) category. Finally we mention some results concerning a fuzzy order relation constructed on a fuzzy set.

2 Preliminaries

Definition 1 [5] *Let $(L, \leq)$ be a complete lattice. A GL-monoid is a complete infinitely distributive lattice enriched with a further monotone, commutative and associative binary operation $*$, i.e. a triple $(L, \leq, \vee, \wedge)$ such that:*

1. $\alpha * 1 = \alpha$, $\alpha * 0 = 0$ $\forall \alpha \in L$
2. $\alpha * (\vee_i \beta_i) = \vee_i (\alpha * \beta)$, $\forall \alpha \in L$ $\forall \{ \beta_i : i \in I \} \subset L$
3. $\alpha \leq \beta \Rightarrow \exists \gamma \in L : \alpha = \beta * \gamma$

Every GL-monoid is residual, i.e., there exists a binary operation " $\mapsto$ " (implication) on L satisfying the following condition:

$$\alpha * \beta \leq \gamma \Longleftrightarrow \alpha \leq (\beta \mapsto \gamma) \; \forall \alpha, \beta, \gamma \in L.$$

Explicitly the implication is given by

$$\alpha \mapsto \beta = \bigvee \{ \lambda \in L : \alpha * \lambda \leq \beta \}.$$

Let's mention some properties of GL-monoids which will be useful in the following:

(i) $\alpha \mapsto (\wedge_i \beta_i) = \wedge_i (\alpha \mapsto \beta_i)$

(ii) $\alpha * (\wedge_i \beta_i) = \wedge_i (\alpha * \beta_i)$

(iii) $(\alpha \mapsto \gamma) * (\gamma \mapsto \beta) \leq \alpha \mapsto \beta$

(iv) $\alpha \leq \beta \Rightarrow \gamma \mapsto \alpha \leq \gamma \mapsto \beta$

Definition 2 [5] *Consider an classical category C and let $\omega : Ob(C) \to L$ and $\mu : Mor(C) \to L$ be L-fuzzy subclasses of the classes of its objects and morphisms respectively. Now, an L-fuzzy category can be defined as a triple $C = (Ob(C), \omega_C, Mor(C), \mu_C, \circ)$ such that :*

1. $\mu(f) \leq \omega(X) \wedge \omega(Y)$ $\forall X, Y \in Ob(C)$ *and* $\forall f \in Mor(X, Y)$;

2. $\mu(g \circ f) \geq \mu(f) * \mu(g)$, *whenever the composition* $f \circ g$ *is defined;*
3. $\mu(e_X) = \omega(X)$, *where* $e_X : X \to X$ *is the identity morphism.*

Definition 3 [6] *By a (global) L-valued set we call a pair* (X, E) *where* X *is a set and* E *is an L-valued equality, i.e., a mapping* $E : X \times X \to L$ *such that*

1. $E(x, x) = 1$;
2. $E(x, y) = E(y, x)\ \forall x, y \in X$;
3. $E(x, y) * E(y, z) \leq E(x, z)\ \forall x, y, z \in X$.

A mapping $f : (X, E_X) \to (Y, E_Y)$ is called extensional if

$$E_X(x, x') \leq E_Y(f(x), f(x'))\ \forall x, x' \in X.$$

Global L-valued sets and extensional mappings between them form a category which will be denoted SET(L).

3 Fuzzy order relation

Let L be a fixed GL-monoid.

Definition 4 *Let* X *be a set. By an fuzzy order relation we call a fuzzy relation* $P : X \times X \to L$ *such that*

1. $P(x, x) = 1\ \forall x \in X$ - *reflexivity;*
2. $P(x, y) * P(y, z) \leq P(x, z) \forall x, y, z \in X$ - *transitivity;*
3. *If* $P(x, y) \wedge P(y, x) > 0$ *then* $x = y$ - *asymmetry.*

A pair (X, P) *is called fuzzy ordered set.*

Definition 5 *Fuzzy order relation* $P : X \times X \to L$ *is called fuzzy linear order if* $P(x, y) \vee P(y, x) > 0$ $\forall x, y \in X$. *The corresponding set* (X, P) *is called fuzzy linearly ordered set.*

Definition 6 *Let* (X, E) *be an L-valued set. By a fuzzy order relation on the L-valued set* (X, E) *we call a mapping* $P : X \times X \to L$ *such that*

1. $P(x, x) = 1\ \forall x \in X$;
2. $P(x, y) * P(y, z) \leq P(x, z)$;
3. $P(x, y) \wedge P(y, x) \leq E(x, y)$, *where* $E : X \times X \to L$ *is an L-valued equality.*

A triple (X, P, E) is called an L-valued ordered set.

If $L = \{0, 1\}$ then the above definition reduces to the classical order relation definition.

4 Fuzzyfication of POS and OS categories

Let define the category $L - POS$ - the category of fuzzy ordered sets:

$$L - POS = (Ob(L - POS), \omega, Mor(L - POS), \mu, \circ)$$

where objects is fuzzy ordered sets

$$Ob(L - POS) = \{(X, P)\}$$

where P is an fuzzy order relation in the set X and morphisms - order-preserving mappings (in the fuzzy sense)
$Mor(L - POS) = \{f : (X, P) \to (Y, P') |$
$P(x_1, x_2) \leq P'(f(x_1), F(x_2))\}$

Let prove that

$$L - POS = (Ob(L - POS), \omega, Mor(L - POS), \mu, \circ)$$

is a category.

1. First show that the composition of two morphism is a morphism : $f \in Mor(L - POS)$, $f : (X, P) \to (Y, P')$ and $g \in Mor(L - POS)$, $g : (Y, P') \to (Z, P'')$, then $g \circ f : X \to Z$ is a mapping and
$P(x_1, x_2) \leq P'(f(x_1), F(x_2)) \leq$
$\leq P''(g(f(x_1)), g(f(x_2))$.
2. Let's poof the associativity condition:
$f \in Mor(L - POS), f : (X, P) \to (Y, P')$,
$g \in Mor(L - POS),\ g : (Y, P') \to (Z, P'')$,
$h \in Mor(L - POS), h : (Z, P') \to (T, P''')$, then
$h \circ (g \circ f) = (h \circ g) \circ f$, since f, g, h are mappings.
3. There exist an identical morphism
$id_X : (X, P) \to (X, P),\ id_X(x) = x\ \forall x \in X$.
Clear, that $\forall f : (X, P) \to (Y, P')\ f \circ id_x = f$ and
$\forall g : (Z, P'') \to (X, P)\ id_x \circ g = g$.

Let us mention some properties of the category $L - POS$ structure.

Theorem 1 *A morphism* $f : (X, P) \to (Y, P')$ *is a monomorphism iff is injective.*

Theorem 2 *A morphism* $f : (X, P) \to (Y, P')$ *is an epimorphism iff is surjective.*

Proposition 1 *An ordered set* $(\varnothing, P_\varnothing)$ *is an initial object in the category* $L - POS$.

Proposition 2 *An ordered set* $(*, P_*)$ *is an final object in the category* $L - POS$.

Our next aim is to construct fuzzyfication of $L-POS$ category. Let define the category $FL-POS$ - the fuzzy analogue of POS category:

$$FL{-}POS = (Ob(FL{-}POS), \omega, Mor(FL{-}POS), \mu, \circ)$$

where objects is fuzzy ordered sets

$$Ob(FL-POS) = \{(X, P)\}$$

where P is an fuzzy order relation in the set X and morphisms - order-preserving mappings (in the fuzzy sense)
$Mor(FL-POS) = \{f : (X, P) \to (Y, P') |$
$P(x_1, x_2) \leq P'(f(x_1), F(x_2))\}$

$$\mu(f) = \inf_{x_1, x_2 \in X} (P'(f(x_1), F(x_2)) \mapsto P(x_1, x_2))$$

where $f : (X, P) \to (Y, P')$.

$$\omega((X, P)) = 1 \; \forall X \in Ob(FL - POS).$$

We continue by proving that L-fuzzy subclasses of the classes of categories objects and morphisms satisfy the necessary conditions:

1. $\mu(f) \leq \omega((X, P)) \wedge \omega((Y, P'))$
   $\forall (X, P), (Y, P') \in Ob(L - POS)$ and
   $\forall f \in Mor((X, P), (Y, P'))$, since $\omega((X, P)) = 1$ and $\omega((Y, P')) = 1$.

2. $f \in Mor(L - POS), f : (X, P) \to (Y, P')$,
   $g \in Mor(L - POS)$, $g : (Y, P') \to (Z, P'')$,
   $\mu(g \circ f) =$
   $= \inf\limits_{x_1, x_2} (P''(g(f(x_1)), g(f(x_2))) \mapsto P(x_1, x_2)) \geq$
   $\geq \inf\limits_{x_1, x_2} ((P''(g(f(x_1)), g(f(x_2))) \mapsto$
   $\mapsto P'(f(x_1), f(x_2))) * (P'(f(x_1), f(x_2)) \mapsto$
   $\mapsto P(x_1, x_2))) \geq$
   $\geq \inf\limits_{x_1, x_2} (P''(g(f(x_1)), g(f(x_2))) \mapsto$
   $\mapsto P'(f(x_1), f(x_2))) * \inf\limits_{x_1, x_2} P'(f(x_1), f(x_2)) \mapsto$
   $\mapsto P(x_1, x_2)) \geq$
   $\geq \inf\limits_{y_1, y_2} (P''(g(y_1), g(y_2)) \mapsto P'(y_1, y_2)) *$
   $* \inf\limits_{x_1, x_2} P'(f(x_1), f(x_2)) \mapsto P(x_1, x_2)) =$
   $= \mu(f) * \mu(g)$
   The result is $\mu(g \circ f) \geq \mu(f) * \mu(g)$.
   In the proof we have used the GL-monoid properties and the following statement:
   $\inf\limits_{x_1, x_2} (P''(g(f(x_1)), g(f(x_2))) \mapsto$
   $\mapsto P'(f(x_1), f(x_2))) \geq \inf\limits_{y_1, y_2} (P''(g(y_1), g(y_2)) \mapsto$
   $\mapsto P'(y_1, y_2))$
   This follows from
   $\{ \; P''(g(f(x_1)), g(f(x_2))) \; \mapsto \; P'(f(x_1), f(x_2)) \; : x_1, x_2 \in X\} \; \subseteq \{ \; P''(g(y_1), g(y_2)) \mapsto$
   $\mapsto P'(y_1, y_2) : y_1, y_2 \in Y\}$.

3. $\mu(e_x) = \inf\limits_{x_1, x_2} (P(x_1, x_2)) \mapsto P(x_1, x_2)) = 1$
   $= \omega(X)$.

The same proof works for the category

$$FL{-}POS_E = (Ob(FL{-}POS_E), \omega, Mor(FL{-}POS_E), \mu, \circ)$$

where objects are L-valued ordered sets but morphisms and mappings μ and ω as in category $FL - POS$.

In the same way we can construct the fuzzy analogue of OS category. We just have to replace fuzzy ordered sets with fuzzy linearly ordered sets.

5 Connection between $L - POS_E$ category and $SET(L)$ category

We will denote by $L - POS$ the category

$$L - POS_E = (Ob(L - POS_E), Mor(L - POS_E), \circ)$$

where

$$Ob(L - POS_E) = \{(X, P, E)\}$$

and P is a fuzzy order relation on an L-valued set (X, E).
$Mor(L - POS_E) = \{F : (X, P, E) \to (Y, P', E') |$
$P(x_1, x_2) \leq P(F(x_1), F(x_2))\}$.

Let us construct a functor F from the category $L - POS_E$ to the category $SET(L)$ as follows:
$F((X, P, E_P)) = (X, E)$, with relation $E : X \times X \to L$ satisfying $E(x, y) = P(x, y) \wedge P(y, x)$.

$$\begin{array}{ccc} (X, P, E_P) & \xrightarrow{F} & (X, E) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle F(f)} \\ (X', P', E'_{P'}) & \xrightarrow{F} & (X', E') \end{array}$$

The mapping $F(f) : (X, E) \to (X', E')$ is the same mapping as a mapping
$f : (X, P, E_P) \to (X', P', E'_P)$
$(\forall x \in X \; F(f)(x) = f(x))$.

First we have to proof that $E : X \times X \to L$ is an L-valued equality:

- $E(x, x) = P(x, x) \wedge P(x, x) = P(x, x) = 1$;
- $E(x, y) = P(x, y) \wedge P(y, x) = P(y, x) \wedge P(y, x) =$
  $= E(x, y)$;
- $E(x, y) * E(y, x) = (P(x, y) \wedge P(y, x)) *$
  $*(P(y, z) \wedge P(z, y)) = (P(x, y) * P(y, z)) \wedge$
  $\wedge (P(x, y) * P(z, y)) \wedge$
  $\wedge (P(y, x) * P(y, z)) \wedge (P(y, x) * P(z, y)) \leq$
  $\leq \; (P(x, y) \; * \; P(y, z)) \; \wedge \; (P(z, y) \; * \; P(y, x)) \; \leq$
  $P(x, z) \wedge P(z, x) = E(x, z)$.

We continue to obtain that F is indeed a functor:

- If $f : (X, P) \to (X'P')$ is a morphism in the category $L - POS$ then
  $Ff : (X, E) \to (X'E')$ is a morphism in the category $SET(L)$. It is equivalent to the statement that f is an extensional mapping
  $(E(x, y) \le E'(f(x), f(y)))$.

  Let proof the latest:

  $E(x, y) = P(x, y) \wedge P(y, x) \le$
  $\le P'(f(x), f(y)) \wedge P'(f(y), f(x)) =$
  $= E'(f(x), f(y))$. Here we used that mapping f preserves an order.

- It is easy to check that $F(f \circ g) = F(f) \circ F(g)$, where $f, g \in Mor(L{-}POS)$ and $F(id_X) = id_{F(X)}$ where $X \in Ob(L - POS)$.

Now let construct a functor G from the category $SET(L)$ to the category $L - POS_E$ imposing the following condition: $G((X, E)) = (X, P, E_P)$ where $P : X \times X \to L$ is defined by $P(x, y) = E(x, y)$ and relation E_P is equal to relation E.

$$\begin{array}{ccc} (X,E) & \xrightarrow{G} & (X,P,E_P) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle G(f)} \\ (X',E') & \xrightarrow{G} & (X',P',E'_{P'}) \end{array}$$

Mapping $G(f) : (X, P, E_P) \to (X', P', E'_{P'})$ is the same mapping as
$f : (X, E) \to (X', E')$ $(\forall x \in X\ G(f)(x) = f(x))$. Our next step is to show that $P : X \times X \to L$ is an order relation:

- $P(x, x) = E(x, x) = 1$;
- $P(x, y) * P(y, z) = E(x, y) * E(y, z) \le E(x, z) =$
  $= P(x, z)$;
- $P(x, y) \wedge P(y, x) = E(x, y) \wedge E(y, x) = E(x, y)$.

We proceed to show that G is a functor:

1. We are going to show that for the morphism $G(f)$ satisfies following
   $P(x, y) \le P'(G(f)(x), G(f)(y))$.

   Proof:

   $P(x, y) = E(x, y) \le E'(f(x), f(y)) =$
   $= P'(f(x), f(y)) = P'(G(f)(x), G(f)(y))$, which follows from the extensionality of mapping f.

2. It is easily seen that $G(f \circ g) = G(f) \circ G(g)$ whenever the composition $f \circ g$ is defined, where $f, g \in Mor(SET(L))$ and $G(id_X) = id_{G(X)}$, where $X \in Ob(SET(L))$.

Theorem 3 *The situation stated above is an adjunction, where F is a right adjoint and G is a left adjoint.*

- *We will consider a natural transformation $\eta : 1_{SET(L)} \to F \circ G$ to be defined as follows: if (X, E) is an object of the category $SET(L)$, then*
  $G((X, E)) = (X, P, E)$, where the relation P is equal to the relation E.
  $F(G((X, E))) = F((X, P, E)) = (X, E)$
  Thus we define $\eta_{(X,E)} : (X, E) \to (X, E)$ as an identity morphism.
  Hence the diagram

  $$\begin{array}{ccc} (X,E) & \xrightarrow{\eta_{(X,E)}} & F(G((X,E))) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle F(G(f))} \\ (X',E') & \xrightarrow{\eta_{(X',E')}} & F(G((X',E'))) \end{array}$$

  reduces to the diagram

  $$\begin{array}{ccc} (X,E) & \xrightarrow{id_{(X,E)}} & (X,E) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ (X',E') & \xrightarrow{id_{(X',E')}} & (X',E') \end{array}$$

  Clearly the diagram commutes.

- *Let describe a natural transformation*
  $\xi : G \circ F \to 1_{L-POS_E}$:
  If (X, P, E) is an category's $L - POS_E$ object, then $F((X, P, E)) = (X, E')$,
  where $E'(x, y) = P(x, y) \wedge P(y, x)\ \forall x, y \in X$.
  $G(F((X, P, E))) = G((X, E')) = (X, P', E')$,
  where $P'(x, y) = E'(x, y)\ \forall x, y \in X$.
  According to the above remark, we define $\xi_{(X,P,E)} : (X, P', E') \to (X, P, E)$ as an morphism, which is an identity mapping, we denote it briefly by id'_X. We continue by proving, that $P'(x, y) \le P(id'_X(x), id'_X(y))$:
  $P'(x, y) = P(x, y) \wedge P(y, x) \le P(x, y) =$
  $= P(id'_X(x), id'_X(y))$
  Thus the diagram

  $$\begin{array}{ccc} G(F((X,P,E))) & \xrightarrow{\xi_{(X,P,E)}} & (X,P,E) \\ {\scriptstyle G(F(f))}\downarrow & & \downarrow{\scriptstyle f} \\ G(F((\tilde{X},\tilde{P},\tilde{E}))) & \xrightarrow{\xi_{\tilde{X},\tilde{P},\tilde{E})}} & (\tilde{X},\tilde{P},\tilde{E}) \end{array}$$

  reduces to the diagram

  $$\begin{array}{ccc} (X,P',E') & \xrightarrow{id'_X} & (X,P,E) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ (\tilde{X},\tilde{P}',\tilde{E}') & \xrightarrow{id'_{\tilde{X}}} & (\tilde{X},\tilde{P},\tilde{E}) \end{array}$$

Obviously the diagram commutes.

The proof is completed by showing that

- $F \xrightarrow{\eta * F} F \circ G \circ F \xrightarrow{F * \xi} F = F \xrightarrow{1_F} F$ *and*
- $G \xrightarrow{G * \eta} G \circ F \circ G \xrightarrow{\xi * G} G = G \xrightarrow{1_G} G$

This follows from the definition of the natural transformations η, ξ and functors F, G.

6 Fuzzy order relation constructed on a fuzzy set

Definition 7 *Let X be a set and δ be its L-subset ($\delta(x) > 0 \ \forall x \in X$). By an order relation on a fuzzy set δ we call an L-valued equality P, such that*

1. $\delta(x) \leq P(x,x)$- *reflexivity;*
2. $P(x,y) \wedge \delta(y) \wedge P(y,z) \leq P(x,z)$- *transitivity;*
3. *If $P(x,y) \wedge P(y,x) > 0$ then $x = y$ - asymmetry.*

If $\delta(x) = 1 \ \forall x \in X$, then the above definition reduces to the definition of an order relation on an "usual" set. A triple (X, δ, P) is called fuzzy ordered L-set.

Theorem 4 *Fuzzy ordered L-sets and mappings*

$f : (X, \delta, P) \to (Y, \delta', P')$ *such that*
$\delta(x_1) \wedge P(x_1, x_2) \wedge \delta(x_2) \leq$
$\leq \delta'(f(x_1)) \wedge P'(f(x_1), f(x_2)) \wedge \delta'(f(x_2))$
$\forall x_1, x_2 \in X$ *form a category. This category weill be denoted $L - POS_L$.*
To prove this theorem we first show that the composition of two morphism is a morphism.
Let $f : (X, \delta, P) \to (Y, \delta', P')$ and
$g : (Y, \delta', P') \to (Z, \delta'', P'')$ *be mappings between corresponding fuzzy ordered L-sets and let*

$g \circ f : (X, \delta, P) \to (Z, \delta'', P'')$ *be their composition. Then* $\delta(x_1) \wedge P(x_1, x_2) \wedge \delta(x_2) \leq$
$\leq \delta'(f(x_1)) \wedge P'(f(x_1), f(x_2)) \wedge \delta'(f(x_2)) \leq$
$\leq \delta''(g(f(x_1))) \wedge P''(g(f(x_1)), g(f(x_2))) \wedge$
$\wedge \delta''(g(f(x_2)))$. *The operation of composition of mappings is obviously associative and the identical mapping*

$$id_X : (X, \delta, P) \to (X, \delta, P)$$

plays the role of the identity.

Our next step is to fuzzyfy the category $L - POS_L$. Let define the category $FL - POS_L$.

$$FL{-}POS = (Ob(FL{-}POS), \omega, Mor(FL{-}POS), \mu, \circ)$$

where objects are fuzzy ordered L-sets and morphisms -

$$Mor(FL - POS_L) = Mor(L - POS).$$

$\mu(f) = \inf_{x \in X} \delta(x) \wedge \inf_{x_1, x_2} (P'(f(x_1), f(x_2)) \leftrightarrow$
$\leftrightarrow P(x_1, x_2)) \wedge \inf_{y \in Y} \delta'(y)$.
$\omega((X, \delta, P)) = \inf_{x \in X} \delta(x) \ \forall X \in Ob(L - POS_L)$.
Mappings μ and ω satisfy all necessary conditions.

References

[1] J. A. Goguen, L-fuzzy sets, J. Math.Anal.Appl. 18 (1967), 145–174.

[2] J. A. Goguen, The fuzzy Tychonoff theorem, J. Math.Anal.Appl., 43 (1973), 737–742.

[3] H. Herrlich, G. E. Strecker, Category Theory, second edition, Heldermann Verlag, Berlin, 1979.

[4] A. Šostak, Fuzzy categories versus categories of fuzzily structured sets: Elements of the theory of fuzzy categories. In: *Mathematik-Arbeitspapiere.*, Universitat Bremen, Vol. 48, 1997, pp. 407–437.

[5] I. Uljane, A. Šostak, On a category of L-valued equalities on L-sets, Journal of Electrical Enginering 55 (2004), 60–64.

[6] I. A. Zadeh, Similarity relations and fuzzy orderings, Inform. Sci. 3 (1971), 177–200.

Monadic L-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$

Michal Holčapek
Institute for Research and Applications of Fuzzy Modeling
University of Ostrava, 30. dubna 22, 701 03 Ostrava 1, Czech Republic
michal.holcapek@osu.cz

Abstract

The aim of this paper is to introduce monadic **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$ and to investigate some of their semantic properties that are well-known in the generalized quantifier theory. Further, some relations between the semantic properties of generalized quantifiers are also studied for monadic **L**-fuzzy quantifiers. Finally, a property "to be extensional w.r.t. a given **L**-permutation equivalence" is proposed.

Keywords: Fuzzy sets, fuzzy quantifiers, fuzzy logic.

1 Introduction

Generalized quantifiers have been studied in mathematical logic since the fifties of the last century, where the pioneering work was done by A. Mostowski [10]. Mostowski's approach to generalized quantifiers based on the mappings from the power sets to the set of truth values significantly influenced further development of a generalized quantifier theory. First considerable generalization of Mostowski's approach was proposed by P. Lindström [8] and the study of generalized quantifiers in connection with natural language has been set in motion by R. Montague [9]. The first systematic investigation of natural language quantifiers is due to J. Barwise and R. Cooper [1]. A survey of works on generalized quantifiers in natural language with a substantial list of references could be found in [7, 14].

The modeling of natural language quantifiers as fuzzy quantifiers being fuzzy numbers has first been pointed out by L.A. Zadeh [16]. A fuzzy extension of Mostowski's idea was proposed by H. Thiele [15]. An investigation of expressions as *Usually* φ, φ *is probable*, *For many* x, φ etc. from the logical point of view was studied by P. Hájek [5]. Linguistic quantifiers in fuzzy logic have been also studied by V. Novák [11]. Since generalized (fuzzy) quantifiers are constructs of the second-order, it is natural to define them in the higher-order (fuzzy) logic. Novák showed a way how to establish some examples of fuzzy quantifiers in his fuzzy type theory [13]. A general theory of semi-fuzzy and fuzzy quantifiers in natural language was proposed by I. Glöckner [3,4]. His approach is based on a generalization of the Lindström's definition of generalized quantifiers of the type $\langle 1, \ldots, 1\rangle$. He proposed an axiomatic system which ensures a correctness of extensions (quantifier fuzzification mechanisms) of semi-fuzzy quantifiers to fuzzy quantifiers. A quantifier fuzzification mechanism satisfying the suggested axioms is then called the determiner fuzzification scheme (DFS) and a result of DFS is a class of plausible models of fuzzy quantifiers. He also introduced some semantic properties of fuzzy quantifiers that are well-known for generalized quantifiers and showed which properties are preserved by DFS.

The aim of this paper is to introduce monadic **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$ which may be used to describe a lot of natural language quantifiers as *many, few, about fifty, more than* etc. and to investigate some of their semantic properties. The proofs and other examples could be found in [6].

2 Preliminaries

In this paper we suppose that the structure of truth values is a *residuated lattice*, i.e. an algebra $\mathbf{L} = \langle L, \wedge, \vee, \rightarrow, \otimes, \bot, \top\rangle$ with four binary operations and two constants such that $\langle L, \wedge, \vee, \bot, \top\rangle$ is a lattice, where $\bot$ is the least element and $\top$ is the greatest element of L, $\langle L, \otimes, \top\rangle$ is a commutative monoid (i.e. $\otimes$ is associative, commutative and the identity $a \otimes \top = a$ holds for any $a \in L$) and the adjointness property is satisfied, i.e.

$$a \leq b \rightarrow c \quad \text{iff} \quad a \otimes b \leq c \tag{1}$$

holds for each $a, b, c \in L$ ($\leq$ denotes the corresponding lattice ordering). The operations $\otimes$ and $\rightarrow$ are called

multiplication and *residuum*, respectively. A residuated lattice is *complete*, if the corresponding lattice is complete. For the substantial information about residuated lattices we refer to [2].

Example 2.1 *It is easy to prove (see e.g. [2]) that* $\mathbf{L} = \langle [0,1], \min, \max, T, \to_T, 0, 1\rangle$, *where* T *is a left continuous t-norm and*

$$a \to_T b = \bigvee \{c \in [0,1] \mid T(a,c) \le b\} \qquad (2)$$

defines the residuum, is a complete residuated lattice.

Let us define the following additional operations for arbitrary $a, b \in L$ and set $\{a_i \mid i \in I\}$ of elements from L over countable (possibly empty) index set I:

$$a \leftrightarrow b = (a \to b) \wedge (b \to a) \qquad (\textit{biresiduum})$$

$$\neg a = a \to \bot \qquad (\textit{negation})$$

$$\bigotimes_{i\in I} a_i = \begin{cases} \top, & I = \emptyset, \\ \bigwedge_{K\in \mathrm{Fin}(I)} \bigotimes_{i\in K} a_i, & \text{otherwise}, \end{cases}$$

$$(\textit{countable multiplication})$$

where $\mathrm{Fin}(I)$ denotes the set of all finite subsets of I. If $|I| = n$ is a finite index set and $a_i = a$ for any $i \in I$, then we will write $a^n = \bigotimes_{i\in I} a_i$.

Let $\mathbf{L}$ be a residuated lattice and M be a universe (possibly empty). An $\mathbf{L}$-*fuzzy set* of M is a mapping $A : M \to L$, where L is the support of the residuated lattice $\mathbf{L}$. Obviously, if $M = \emptyset$, then the empty mapping is the unique $\mathbf{L}$-fuzzy set. The set of all $\mathbf{L}$-fuzzy sets of M is denoted by $\mathcal{F}_{\mathbf{L}}(M)$. An $\mathbf{L}$-fuzzy set $A \in \mathcal{F}_{\mathbf{L}}(M)$ is called *crisp*, if $A(m) \in \{\bot, \top\}$ hods for any $m \in M$. The set of all crisp $\mathbf{L}$-fuzzy sets in M is denoted by $\mathcal{P}(M)$. We say that an $\mathbf{L}$-fuzzy set A is *less or equal to* an $\mathbf{L}$-fuzzy set B and denote it by $A \le B$, if for an arbitrary $m \in M$ we have $A(m) \le B(m)$. Further, the intersection of $A, B \in \mathcal{F}_{\mathbf{L}}(M)$ is an $\mathbf{L}$-fuzzy set $C = A \cap B$ such that $C(m) = (A \cap B)(m) = A(m) \wedge B(m)$ holds for any $m \in M$ and the complement of $A \in \mathcal{F}_{\mathbf{L}}(M)$ is an $\mathbf{L}$-fuzzy set $C = \overline{A}$ such that $C(m) = \overline{A}(m) = \neg A(m)$ holds for any $m \in M$. The complement of M is denoted $\emptyset$. Note that if $M = \emptyset$, then $\mathcal{F}_{\mathbf{L}}(\emptyset) = \{\emptyset\}$ and thus trivially $\emptyset \cap \emptyset = \emptyset$ and $\overline{\emptyset} = \emptyset$. Finally, let $f : M \to M'$ be a mapping. Then $f^{\to}(A)(m) = \bigvee_{m'\in f^{-1}(m)} A(m')$ defines an $\mathbf{L}$-mapping $f^{\to} : \mathcal{F}_{\mathbf{L}}(M) \to \mathcal{F}_{\mathbf{L}}(M)$. Obviously, if f is a bijective mapping then $f^{\to}(A)(f(m)) = A(m)$ for all $m \in M$. Note that if $M = \emptyset$, then $\emptyset : \emptyset \to \emptyset$ is the unique bijective mapping between the empty sets. This mapping determines the unique mapping $\emptyset^{\to} : \mathcal{F}_{\mathbf{L}}(\emptyset) \to \mathcal{F}_{\mathbf{L}}(\emptyset)$ which assigns $\emptyset$ to $\emptyset$, i.e. $\emptyset^{\to}(\emptyset) = \emptyset$.

3 Definition and examples

Let us consider the following examples of sentences

(a) John was in love with Mary.

(b) All read the New York Times.

(c) About half employees of this firm are young.

(d) More men than women go in for sport.

In these sentences we can see that the noun phrases *John, All, About half employees of this firm, More men than women* are combined whit one place (fuzzy) predicates *was in love with Mary, read the New York Times, are young, go in for sport.*

Let M be a universe of objects as e.g. people, animals, books, etc. One place fuzzy predicates can be interpreted in M as $\mathbf{L}$-fuzzy sets of M and noun phrases in a connection with one place predicates as mappings from the set $\mathcal{F}_{\mathbf{L}}(M)$ into the support L of the residuated lattice $\mathbf{L}$ (cf. [7,10,14]). Thus the noun phrases which are combined with one place fuzzy predicates are semantically interpreted in a given universe M as mappings $Q_M : \mathcal{F}_{\mathbf{L}}(M) \to L$. They will be called $\mathbf{L}$-fuzzy quantifiers of the type $\langle 1\rangle$. If noun phrases are formed by determiners[1] as e.g. *all, about half, more than* and one or more noun phrases (according to the arity of considered determiner), then the n-place determiners may be semantically interpreted as mappings from $\mathcal{F}_{\mathbf{L}}(M)^n$ to the set of all $\mathbf{L}$-fuzzy quantifiers of the type $\langle 1\rangle$ or equivalently as mappings $Q_M : \mathcal{F}_{\mathbf{L}}(M)^{n+1} \to L$. They will be called $\mathbf{L}$-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$.

In the following text we will be interested in such types of $\mathbf{L}$-fuzzy quantifiers that are some semantic interpretation of combination of determiner and just one one place predicate.

Definition 3.1 *Let* $\mathbf{L}$ *be a residuated lattice,* M *be a universe (possibly empty) and* n *be a natural number (possibly zero). A mapping* $Q_M : \mathcal{F}_{\mathbf{L}}(M)^{n+1} \to L$ *is called the* monadic $\mathbf{L}$-fuzzy quantifier of the type $\langle 1^n, 1\rangle$ limited to M.

In natural language noun phrases (determiners) are not fixed to one universe (model). Therefore, it is reasonable to define $\mathbf{L}$-fuzzy quantifiers over a class of universes. Such quantifiers are called unlimited or global (see [7,10]).

Definition 3.2 *An* unlimited monadic $\mathbf{L}$-fuzzy quantifier of the type $\langle 1^n, 1\rangle$ *is a functional* Q *assigning to each universe* M *a monadic* $\mathbf{L}$-*fuzzy quantifier* Q_M *of the type* $\langle 1^n, 1\rangle$ *limited to* M. *The set of all unlimited* $\mathbf{L}$-*fuzzy quantifiers is denoted by* QUANT.

[1] Note that in the linguistic theory the quantifiers create a special subclass of language elements which are called, in general, determiners. Some examples of determiners which are not quantifiers are *the, this, one, first.*

Definition 3.3 *An* unlimited finite (countable) monadic **L**-fuzzy quantifier *of the type* $\langle 1^n, 1\rangle$ *is a functional* Q *assigning to each finite (countable) universe* M *a monadic* **L***-fuzzy quantifier* Q_M *of the type* $\langle 1^n, 1\rangle$ *limited to* M*. The set of all unlimited monadic* **L***-fuzzy quantifiers is denoted by* F-QUANT (C-QUANT)*.*

Example 3.1 *Let* **L** *be a complete residuated lattice. Then*

$$(\mathsf{all})_M(A) = \bigwedge_{m\in M} A(m),$$
$$(\mathsf{some})_M(A) = \bigvee_{m\in M} A(m),$$
$$(\mathsf{not\ all})_M(A) = \neg \bigwedge_{m\in M} A(m),$$
$$(\mathsf{no})_M(A) = \neg \bigvee_{m\in M} A(m),$$

where M *is an arbitrary set and* $A \in \mathcal{F}_{\mathbf{L}}(M)$*, define unlimited* **L***-fuzzy quantifiers of the type* $\langle 1\rangle$*. First two examples are the classical quantifiers* for all *and* there exists *in fuzzy logic. The others are a negation of the previous ones. Notice that* $(\mathsf{all})_\emptyset(\emptyset) = (\mathsf{no})_\emptyset(\emptyset) = \top$ *and* $(\mathsf{some})_\emptyset(\emptyset) = (\mathsf{not\ all})_\emptyset(\emptyset) = \bot$*.*

Example 3.2 *Let* **L** *be a residuated lattice. Then*

$$(\mathsf{John})_M(A) = \begin{cases} A(m_{\mathsf{John}}), & \text{if } m_{\mathsf{John}} \in M; \\ \bot, & \text{otherwise.} \end{cases}$$

where M *is an arbitrary set and* $A \in \mathcal{F}_{\mathbf{L}}(M)$*, defines unlimited* **L***-fuzzy quantifier of the type* $\langle 1\rangle$*.*

Example 3.3 *Let* **L** *be a complete residuated lattice determined by Łukasiewicz t-norm and put* $\frac{0}{0} = 1$ *and* $\sum_{m\in\emptyset} \emptyset(m) = 0$*. Then*

$$(\mathsf{nearly\ all})_M(A) = \frac{1}{|M|} \sum_{m\in M} A(m)$$
$$(\mathsf{nearly\ none})_M(A) = \frac{1}{|M|} \sum_{m\in M} \overline{A}(m)$$

where M *is an arbitrary finite set and* $A \in \mathcal{F}_{\mathbf{L}}(M)$*, define unlimited finite* **L***-fuzzy quantifiers of the type* $\langle 1\rangle$*. Note that these* **L***-fuzzy quantifiers are usually called relative fuzzy quantifiers (cf. [3, 4, 16]).*

Example 3.4 *Let* L *be a complete residuated lattice determined by a left continuous t-norm. Then*

$$(\mathsf{all})_M(A, B) = \bigwedge_{m\in M} \big(\overline{A}(m) \vee B(m)\big),$$
$$(\mathsf{some})_M(A, B) = \bigvee_{m\in M} \big(A(m) \wedge B(m)\big),$$

where M *is an arbitrary universe and* $A, B \in \mathcal{F}_{\mathbf{L}}(M)$*, define unlimited* **L***-fuzzy quantifiers of the type* $\langle 1, 1\rangle$*.*

Example 3.5 *Let* L *be a complete residuated lattice. Further, let us denote* $\mathrm{Perm}(M)$ *the set of all bijective mappings (permutations) on* M*. Then*

$$(\mathsf{more\ than})_M(A, B, C) =$$
$$\neg \bigvee_{f\in \mathrm{Perm}(M)} \bigwedge_{m\in M} ((A\cap C)(m) \to (B \cap C)(f(m))),$$
$$(\mathsf{less\ than})_M(A, B, C) =$$
$$\neg \bigvee_{f\in \mathrm{Perm}(M)} \bigwedge_{m\in M} ((B\cap C)(f(m)) \to (A \cap C)(m)),$$

where M *is an arbitrary universe and* $A, B, C \in \mathcal{F}_{\mathbf{L}}(M)$*, define unlimited* **L***-fuzzy quantifiers of the type* $\langle 1^2, 1\rangle$*.*

4 Basic semantic properties

In this part we will investigate the well-known semantic properties of generalized quantifiers for monadic **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$. Namely, we will be interested in the permutation invariance, isomorphism invariance, extension, monotonicity and conservativity. For more information about these properties for generalized quantifiers we refer to [14].

Definition 4.1 *An* **L***-fuzzy quantifier* Q *of the type* $\langle 1^n, 1\rangle$ *is* permutation invariant*, if for each universe* M *and for each bijective mapping* $f : M \to M$ *and for each* $(A_1, \ldots, B) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ *we have*

$$Q_M(A_1, \ldots, B) = Q_M(f^{\to}(A_1), \ldots, f^{\to}(B)). \quad (3)$$

The set of all permutation invariant **L***-fuzzy quantifiers is denoted by* PI

One can easily check that all mentioned examples except **L**-fuzzy quantifier John belong to PI. Obviously, the permutation invariance is a local condition. We need not to compare universes to see if it holds. In order to investigate monadic **L**-fuzzy quantifiers, if they ignore elements from a global perspective, we have to use a stronger condition than the permutation invariance which is called the isomorphism invariance (see e.g. [7, 8, 14]).

Definition 4.2 *An* **L***-fuzzy quantifier* Q *of the type* $\langle 1^n, 1\rangle$ *is* isomorphism invariant*, if for each universes* M, M' *and for each bijective mapping* $f : M \to M'$ *and for each* $(A_1, \ldots, B) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ *we have*

$$Q_M(A_1, \ldots, B) = Q_{M'}(f^{\to}(A_1), \ldots, f^{\to}(B)). \quad (4)$$

The set of all isomorphism invariant **L***-fuzzy quantifiers is denoted by* ISOM*.*

Again all mentioned examples of monadic **L**-fuzzy quantifiers except John are isomorphism invariant.

The following proposition shows that isomorphism invariance is really stronger condition than permutation invariance.

Proposition 4.1 ISOM $\subseteq$ PI.

Let us denote $A^a = \{m \in M \mid A(m) = a\}$. The following proposition generalizes the classical result on isomorphism invariant generalized quantifiers of the type $\langle 1\rangle$ (cf. [7,14]).

Theorem 4.1 *Let $Q \in$ QUANT be of the type $\langle 1\rangle$. Then $Q \in$ ISOM if and only if for each universes M, M' and for each $A \in \mathcal{F}_{\mathbf{L}}(M)$ and for each $B \in \mathcal{F}(M')$ we have if $|A^a| = |B^a|$ and $|M\backslash A^a| = |M'\backslash B^a|$ hold for all $a \in L$, then $Q_M(A) = Q_{M'}(B)$.*

The property saying that a quantifier is, in some sense, independent on the universes is called extension (see e.g. [3,4,7,8,14]).

Definition 4.3 *An* **L***-fuzzy quantifier Q of the type $\langle 1^n, 1\rangle$ satisfies extension, if for arbitrary universes M, M' with $M \subseteq M'$ and for each $(A_1, \ldots, B) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ we have*

$$Q_M(A_1, \ldots, B) = Q_{M'}(A_1, \ldots, B). \tag{5}$$

The set of all **L***-fuzzy quantifiers satisfying extension is denoted by* EXT.

The quantifier all from Example 3.1 is an example of monadic **L**-fuzzy quantifier which does not satisfy extension.

It is easy to see that simple determiners as *all, some, few, nearly all, at most k,* etc. are usually monotonic. On the other hand, there are determiners as *about half, at least five an at most five, approximately fifty,* etc. are not monotonic. The monotonicity of monadic **L**-quantifiers in some its component is naturally defined as follows (cf. [3,4,7,14]).

Definition 4.4 *Let Q be an* **L***-fuzzy quantifier of the type $\langle 1^n, 1\rangle$. We say that Q is* non-decreasing in its i-th component*, where $i = 1, \ldots, n+1$, if for any universe M and for each $(A_1, \ldots, A_i, \ldots, A_{n+1}) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ and for each $A'_i \in \mathcal{F}_{\mathbf{L}}(M)$ with $A_i \subseteq A'_i$ we have*

$$\begin{aligned} Q_M(A_1, \ldots, A_i, \ldots, A_{n+1}) \leq \\ Q_M(A_1, \ldots, A'_i, \ldots, A_{n+1}) \end{aligned} \tag{6}$$

and Q is non-increasing in its i-th component*, where $i = 1, \ldots, n+1$, if for any universe M and for each $(A_1, \ldots, A_i, \ldots, A_{n+1}) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ and for each $A'_i \in \mathcal{F}_{\mathbf{L}}(M)$ with $A'_i \subseteq A_i$ we have*

$$\begin{aligned} Q_M(A_1, \ldots, A_i, \ldots, A_{n+1}) \leq \\ Q_M(A_1, \ldots, A'_i, \ldots, A_{n+1}). \end{aligned} \tag{7}$$

An **L***-fuzzy quantifier of the type $\langle 1^n, 1\rangle$ which is non-decreasing or non-increasing in each its component is called the* non-decreasing *or* non-increasing **L**-fuzzy quantifier*, respectively.*

One can easily check that all, some, John, nearly all are non-decreasing and not all, no, nearly none are non-increasing **L**-fuzzy quantifiers of the type $\langle 1\rangle$, respectively. Further, some as the **L**-fuzzy quantifier of the type $\langle 1, 1\rangle$ is non-decreasing in both its components, however, all as the **L**-fuzzy quantifier of the type $\langle 1, 1\rangle$ is non-increasing in its first component and non-decreasing in its second component.

A property of determiners saying that we may limit our consideration of objects with the predicate property to those which also have the noun property is called conservativity.

Definition 4.5 *An* **L***-fuzzy quantifier Q of the type $\langle 1^n, 1\rangle$ is* conservative*, if for an arbitrary universe M and for each $(A_1, \ldots, A_n) \in \mathcal{F}_{\mathbf{L}}(M)^n$ and for each $B, B' \in \mathcal{F}_{\mathbf{L}}(M)$ such that $A_i \cap B = A_i \cap B'$ holds for each $i = 1, \ldots, n$ we have*

$$Q_M(A_1, \ldots, A_n, B) = Q_M(A_1, \ldots, A_n, B'). \tag{8}$$

The set of all conservative **L***-fuzzy quantifiers is denoted by* CONS.

The following proposition is analogous to that for generalized quantifiers of the type $\langle 1, 1\rangle$ (cf. [7,14])

Proposition 4.2 *Let Q be an* **L***-fuzzy quantifier of the type $\langle 1, 1\rangle$. Then $Q \in$ CONS if and only if for an arbitrary universe M and for each $A, B \in \mathcal{F}_{\mathbf{L}}(M)$ we have*

$$Q_M(A, B) = Q_M(A, A \cap B). \tag{9}$$

Let us consider the **L**-fuzzy quantifiers some and all of the type $\langle 1, 1\rangle$ from Example 3.4. One can check that some is conservative and all is not, in general, conservative.

In [3,4] Glöckner defines n-ary semi-fuzzy quantifiers as mappings $Q : \mathcal{P}(M) \to [0,1]$. The following definition of **L**-fuzzy quantifiers which are semi-fuzzy in some of its components generalizes Glöckner's idea of semi-fuzzy quantifiers.

Definition 4.6 *An* **L***-fuzzy quantifier Q of the type $\langle 1^n, 1\rangle$ is* semi-fuzzy in its i-th component*, where $i = 1, \ldots, n+1$, if for an arbitrary universe M and for each $(A_1, \ldots, A_n, A_{n+1}) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ we have*

$$\begin{aligned} Q_M(A_1, \ldots, A_i, \ldots, A_n, A_{n+1}) = \\ Q_M(A_1, \ldots, \mathrm{Supp}(A_i), \ldots, A_n, A_{n+1}). \end{aligned} \tag{10}$$

An **L***-fuzzy quantifier of the type* $\langle 1^n, 1\rangle$ *which is semi-fuzzy in each its component is called* **L**-semi-fuzzy quantifier *of the type* $\langle 1^n, 1\rangle$.

In the generalized quantifier theory there is shown that quantifiers of the type $\langle 1, 1\rangle$ belong to CONS ∩ EXT if and only if they are the relativization of quantifiers of the type $\langle 1\rangle$ (see e.g. [7,14]). An analogical equivalence is failed, in general, for **L**-fuzzy quantifiers of the type $\langle 1, 1\rangle$. The reason is that the monadic **L**-fuzzy quantifiers are limited to crisp universes and not to fuzzy universes. Nevertheless, an analogical equivalence could be proved, if we restrict ourselves on monadic **L**-fuzzy quantifiers which are semi-fuzzy in its first component.

Definition 4.7 *The relativization of an* **L***-fuzzy quantifier* Q *of the type* $\langle 1\rangle$ *is the* **L***-fuzzy quantifier* Q^{rel} *of the type* $\langle 1, 1\rangle$ *given by*

$$Q^{\mathrm{rel}}_M(A, B) = Q_{\mathrm{Supp}(A)}(A \cap B)$$

for an arbitrary universe M *and for arbitrary* $A, B \in \mathcal{F}_{\mathbf{L}}(M)$.

Theorem 4.2 *Let* Q *be an* **L***-fuzzy quantifier of the type* $\langle 1, 1\rangle$ *which is semi-fuzzy in its first component. Then* $Q \in \mathrm{CONS} \cap \mathrm{EXT}$ *if and only if there exists an* **L***-fuzzy quantifier* P *of the type* $\langle 1\rangle$ *such that* $Q = P^{\mathrm{rel}}$.

The following theorem presents the well known result in the generalized quantifier theory (see e.g. [7]) for monadic **L**-fuzzy quantifiers of the type $\langle 1, 1\rangle$.

Theorem 4.3 *Let* Q *be an* **L***-fuzzy quantifier of the type* $\langle 1, 1\rangle$ *which is semi-fuzzy in its first component. Then* $Q \in \mathrm{CONS} \cap \mathrm{EXT} \cap \mathrm{ISOM}$ *if and only if for arbitrary universes* M, M' *and for each* $A, B \in \mathcal{F}_{\mathbf{L}}(M)$ *and for each* $A', B' \in \mathcal{F}_{\mathbf{L}}(M')$ *we have if* $|(A \cap B)^a| = |(A' \cap B')^a|$ *and* $|\mathrm{Supp}(A) \setminus B^a| = |\mathrm{Supp}(A') \setminus B'^a|$ *hold for any* $a \in L$, *then* $Q_M(A, B) = Q_{M'}(A', B')$.

5 Extensional L-fuzzy quantifiers

We say that n-tuples of **L**-fuzzy sets $(A_1, \ldots, A_n)$, $(B_1, \ldots, B_n) \in \mathcal{F}_{\mathbf{L}}(M)^n$ are *permutation isomorphic*, if there exists a bijective mapping $f : M \to M$ such that $A_i(m) = B_i(f(m))$ holds for any $m \in M$ and $i = 1, \ldots, n$. It could be seen that many determiners as *all, some, no, not all, nearly all, nearly none, more than, less than,* etc. have the property to preserve a similarity of **L**-fuzzy sets. It means that for two similar **L**-fuzzy set, say from the permutation isomorphism point of view, the considered **L**-fuzzy quantifiers give similar values. This property will be called the extensionality of **L**-fuzzy quantifiers w.r.t. a given class of permutation equivalences.

Let $\equiv_M : \mathcal{F}_{\mathbf{L}}(M)^n \times \mathcal{F}_{\mathbf{L}}(M)^n \to L$ be an **L**-fuzzy relation given by $[\mathbf{A} \equiv_M \mathbf{B}] = \top$, if $\mathbf{A} = (A_1, \ldots, A_n)$, $\mathbf{B} = (B_1, \ldots, B_n)$ are permutation isomorphic, and $[\mathbf{A} \equiv_M \mathbf{B}] = \bot$, otherwise. Obviously, $\equiv_M$ is an **L**-similarity relation on $\mathcal{F}_{\mathbf{L}}(M)^n$, that is, an **L**-fuzzy relation satisfying reflexivity, symmetry and $\otimes$-transitivity (see e.g. [2,12]). The class containing $\equiv_M$ for each universe M will be denoted by $\equiv$. A generalization of the **L**-fuzzy relation $\equiv_M$ could be done as follows.

Definition 5.1 *An* **L***-fuzzy relation* $\approx_M : \mathcal{F}_{\mathbf{L}}(M)^n \times \mathcal{F}_{\mathbf{L}}(M)^n \to L$ *is called the* **L**-permutation equivalence *on* $\mathcal{F}_{\mathbf{L}}(M)^n$, *if*

$$[\mathbf{A} \equiv_M \mathbf{B}] \le [\mathbf{A} \approx_M \mathbf{B}] \tag{11}$$

$$[\mathbf{A} \approx_M \mathbf{B}] = [\mathbf{B} \approx_M \mathbf{A}] \tag{12}$$

$$[\mathbf{A} \approx_M \mathbf{B}] \le [\overline{\mathbf{A}} \approx_M \overline{\mathbf{B}}] \tag{13}$$

$$[\mathbf{A} \approx_M \mathbf{B}] \otimes [\mathbf{B} \approx_M \mathbf{C}] \le [\mathbf{A} \approx_M \mathbf{C}] \tag{14}$$

hold for arbitrary $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{F}_{\mathbf{L}}(M)^n$, *where* $\overline{\mathbf{A}} = (\overline{A}_1, \ldots, \overline{A}_n)$ *and* $\overline{\mathbf{B}} = (\overline{B}_1, \ldots, \overline{B}_n)$.

The following two theorems show examples of **L**-permutation equivalences that are a natural generalization of $\equiv_M$. Recall that $\mathrm{Perm}(M)$ denotes the set of all bijective mappings (permutations) on M.

Theorem 5.1 *Let* **L** *be a complete residuated lattice and* M *be a countable universe. Then*

$$[\mathbf{A} \approx^{\otimes}_M \mathbf{B}] = \bigvee_{f \in \mathrm{Perm}(M)} \bigotimes_{m \in M} \big((A_1(m) \leftrightarrow B_1(f(m))) \otimes \cdots \otimes (A_n(m) \leftrightarrow B_n(f(m)))\big)$$

defines an **L***-permutation equivalence* $\approx^{\otimes}_M$ *on* $\mathcal{F}_{\mathbf{L}}(M)^n$.

Theorem 5.2 *Let* **L** *be a complete residuated lattice and* M *be an arbitrary universe. Then*

$$[\mathbf{A} \approx^{\wedge}_M \mathbf{B}] = \bigvee_{f \in \mathrm{Perm}(M)} \bigwedge_{m \in M} \big((A_1(m) \leftrightarrow B_1(f(m))) \otimes \cdots \otimes (A_n(m) \leftrightarrow B_n(f(m)))\big)$$

defines an **L***-permutation equivalence* $\approx^{\wedge}_M$ *on* $\mathcal{F}_{\mathbf{L}}(M)^n$.

Now we define extensionality and weak extensionality of **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$ w.r.t. a given class of permutation equivalencies.

Definition 5.2 *Let* $\approx$ *be a class of* **L***-permutation equivalences such that for each (finite, countable) universe* M *there is a unique* $\approx_M$ *from* $\approx$ *defined on* $\mathcal{F}_{\mathbf{L}}(M)^{n+1}$. *A (finite, countable)* **L***-fuzzy quantifier* Q *of the type* $\langle 1^n, 1\rangle$ *is* extensional w.r.t. $\approx$, *if we have*

$$[(A_1, \ldots, A_n, B) \approx_M (A'_1, \ldots, A'_n, B')] \le Q_M(A_1, \ldots, A_n, B) \leftrightarrow Q_M(A'_1, \ldots, A'_n, B') \tag{15}$$

for an arbitrary (finite, countable) universe M and for each $(A_1, \ldots, A_n, B)$, $(A'_1, \ldots, A'_n, B') \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$. The set of all **L***-fuzzy quantifiers which are extensional w.r.t. $\approx$ is denoted by* EXTENS($\approx$).

Definition 5.3 *Let $\approx$ be a class of* **L***-permutation equivalences such that for each (finite, countable) universe M there is a unique $\approx_M$ from $\approx$ defined on $\mathcal{F}_{\mathbf{L}}(M)^{n+1}$. A (finite, countable)* **L***-fuzzy quantifier Q of the type $\langle 1^n, 1\rangle$ is* weak extensional w.r.t. $\approx$*, if for each (finite, countable) universe M and for each $(A_1, \ldots, A_n, B)$, $(A'_1, \ldots, A'_n, B) \in \mathcal{F}_{\mathbf{L}}(M)^{n+1}$ there exists a natural number k such that*

$$[(A_1, \ldots, A_n, B) \approx_M (A'_1, \ldots, A'_n, B')]^k \leq Q_M(A_1, \ldots, A_n, B) \leftrightarrow Q_M(A'_1, \ldots, A'_n, B'). \quad (16)$$

The set of all **L***-fuzzy quantifiers which are weak extensional w.r.t. $\approx$ is denoted by* W-EXTENS($\approx$).

It is easy to prove that the following relations are true.

Proposition 5.1 EXTENS($\approx$) $\subseteq$ W-EXTENS($\approx$) $\subseteq$ PI.

One can check that all, some, not all, no from Example 3.1 are extensional w.r.t. $\approx^{\wedge}$, where $\approx^{\wedge}$ denotes the class of all **L**-permutation equivalencies $\approx^{\wedge}_M$ defined in Theorem 5.2. Further, nearly all and nearly none from Example 3.3 are extensional w.r.t. $\approx^{\otimes}$, where $\approx^{\otimes}$ denotes the class of all **L**-permutation equivalencies $\approx^{\otimes}_M$ defined in Theorem 5.1. It is easy to show that nearly all and nearly none need not be extensional w.r.t. $\approx^{\otimes}$, if we consider a complete residuated lattice determined by a left continuous t-norm different from the Łukasiewicz t-norm.

Remark 5.1 EXTENS($\equiv$) = W-EXTENS($\equiv$) = PI.

6 Conclusion

In this paper the monadic **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$ were introduced and some of their semantic properties were investigated. The definitions of permutation invariance, isomorphism invariance, extension etc. were proposed in order to be a natural extension of the well known properties of generalized quantifiers and also to correspond to those that have been suggested by Glöckner in [3]. Further, we were interested in relations for monadic **L**-fuzzy quantifiers having some semantic properties which are well known in generalized quantifier theory and they have not been investigated for fuzzy quantifiers. Finally, new semantic property of monadic **L**-fuzzy quantifiers called extensionality w.r.t. a given **L**-permutation equivalence was suggested. In particular, monadic **L**-fuzzy quantifiers having this property assign similar values to **L**-fuzzy sets which are close from the **L**-permutation equivalence point of view.

The presented results form a basic analysis of semantic properties of monadic **L**-fuzzy quantifiers that is analogous to the conception presented in generalized quantifier theory. Moreover, it could be useful for constructions of syntax and semantics of the first or higher order fuzzy logic with monadic **L**-fuzzy quantifiers (see e.g. [5, 12, 13]).

Acknowledgement

The author wishes to acknowledge the support from the Institutional Research Plan MŠMT 6198898701.

References

[1] J. Barwise and R. Cooper, Generalized quantifiers and natural language, Linguistics and Philosophy **4** (1981), 159–219.

[2] R. Bělohlávek, Fuzzy relational systems: Foundations and principles, Kluwer Academic Publisher, New York, 2002.

[3] I. Glöckner, Fuzzy quantifiers in natural language: Semantics and computational models, Der Andere Verlag, Osnabrück, Germany, 2004.

[4] I. Glöckner, Fuzzy quantifiers: A computational theory, Springer-Verlag, Heidelberg, 2006.

[5] P. Hájek, Metamathematics of fuzzy logic, Kluwer Academic Publishers, Dordrecht, 1998.

[6] M. Holčapek. Monadic **L**-fuzzy quantifiers of the type $\langle 1^n, 1\rangle$, submitted to Fuzzy Sets and Sytems.

[7] E.L. Keenan and Westerståhl, Generalized quantifiers in linguistics and logic, in J. van Benthem and A. ter Meulen (Eds.), *Handbook of logic and Language*, Elsevier, Amsterdam, 1996, pp. 837–893.

[8] P. Lindström, First ordered predicate logic with generalized quantifiers, Theoria **32** (1966), 186–195.

[9] R. Montague, The proper treatment of quantification in ordinary english, in R. Thomason (Ed.), *Formal philosophy: selected papers of Richard Montague*, New Haven, Yale Univ. Press, 1974, pp. 247–270.

[10] A. Mostowski, On a generalization of quantifiers, Fund. Math. **44** (1957), 12–36.

[11] V. Novák, Antonyms and linguistic quantifiers in fuzzy logic, Fuzzy Sets and Systems **124** (2001), no. 3, 335–351.

[12] V. Novák, On fuzzy type theory, Fuzzy Sets and Systems **149** (2005), 235–273.

[13] V. Novák, Fuzzy logic theory of evaluating expressions and comparative quantifiers, In Proc. IMPU'06, Editions EDK, 2006, pp. 1572–1579.

[14] S. Peters and D. Westerståhl, Quantifiers in language and logic, Claredon Press, Oxford, 2006.

[15] H. Thiele, On fuzzy quantifiers, in Z. Bien and K.C. Min (Eds.), *Fuzzy Logic and its Applications, Information Sciences, and Intelligent Systems*, Kluwer Academic Publisher, Dordrecht, 1996, pp. 343–352.

[16] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Comp. Math. with Applications **9** (1983), 149–184.

HOMOMORPHISMS OF FUZZY LOGIC MODELS BASED ON SETS WITH SIMILARITIES

Jiří Močkoř
University of Ostrava
Institute for Research and Applications of Fuzzy Modeling
30. dubna 22, 701 03 Ostrava 1, Czech Republic
Jiri.Mockor@osu.cz *

Abstract

We investigate interpretations $\|\psi\|_{\mathcal{D}}$ of formulas ψ in a first order fuzzy logic in models $\mathcal{D}$ which are based on Ω-sets, i.e. sets with similarity relations with values in a complete MV-algebra Ω. These interpretations are fuzzy sets in some Ω-sets. We show that if $\varphi : \mathcal{D}_1 \to \mathcal{D}_2$ is a strong homomorphism between two models then there are also strong relationships between interpretations $\|\psi\|_{\mathcal{D}_1}$ and $\|\psi\|_{\mathcal{D}_2}$. Finally, for any model $\mathcal{D}$ based on an Ω-set (A, δ) we construct another model $S(\mathcal{D})$ based on a set $S(A, \delta)$ of all "continuous" maps from (A, δ) to $(\Omega, \leftrightarrow)$ and such that a singleton map $\{-\}$ is a strong homomorphism $\mathcal{D} \to S(\mathcal{D})$.

Keywords: Sets with similarity relations, fuzzy sets in Ω-sets, models of fuzzy logic.

1 Introduction

Let J be a first order language of fuzzy logic which consists (as classically) of a set of predicate symbols $P \in \mathcal{P}$, a set of functional symbols $f \in \mathcal{F}$ and a set of classical logical connectives $\{\wedge, \vee, \Rightarrow, \neg, \otimes\}$. Moreover J contains also a set Ω of logical constants. Then a model of a language J is

$$\mathcal{E} = (A, \{P_{\mathcal{E}} : P \in \mathcal{P}\}, \{f_{\mathcal{E}} : f \in \mathcal{F}\}),$$

where

(a) A is a set,
(b) $P_{\mathcal{E}} \underset{\sim}{\subseteq} A \times \cdots \times A$ is a fuzzy set with values in Ω,
(c) $f_{\mathcal{E}} : A \times \cdots \times A \to A$ is a map.

For a set X of variables we set $\mathcal{E}(X) = \prod_{x \in X} A = A^X$. This set then serves as a domains for interpretation of formulas and terms with free variables contained in a set X.

Supported by MSM6198898701 and grant 201/07/0191 of GAČR

Now, by well known procedure which is based on inductive principle we can obtain that for a term t and for a formula ψ in J with free variables in X,

(a) an interpretation of t in a model $\mathcal{E}$ is a map $\|t\|_{\mathcal{E}} : \mathcal{E}(X) \to A$,
(b) an interpretation of ψ in a model $\mathcal{E}$ is a fuzzy set $\|\psi\|_{\mathcal{E}} \underset{\sim}{\subseteq} \mathcal{E}(X)$ with values in Ω, i.e. $\|\psi\|_{\mathcal{E}} : \mathcal{E}(X) \to \Omega$.

It is clear that an interpretation is based significantly on a set Ω. In case that $\Omega = (L, \wedge, \vee, \otimes, \to)$ is a complete residuated lattice (which is the most important case for fuzzy set theory) then this interpretation reflects naturally relations between logical and lattice connectives. For example, if $\psi \equiv \phi \Rightarrow \sigma$ then for any $\mathbf{a} \in \mathcal{E}(X)$ we have

$$\|\psi\|_{\mathcal{E}}(\mathbf{a}) = \|\phi\|_{\mathcal{E}}(\mathbf{a}) \to \|\sigma\|_{\mathcal{E}}(\mathbf{a})$$

and analogously

$$\|(\exists x)\psi\|_{\mathcal{E}}(\mathbf{a}) = \bigvee_{b \in A} \|\psi[x/b]\|_{\mathcal{E}}(\mathbf{a}).$$

In papers [10],[12] we investigated models of first order fuzzy logic which are based on a generalization of fuzzy sets, namely on a category $\mathbf{SetF}(\Omega)$ of sets with similarity relations. Recall that a category $\mathbf{SetF}(\Omega)$ consists of objects (A, δ) (called Ω-sets), where A is a set and δ is a similarity relation, i.e. a map $\delta : A \times A \to \Omega$ such that

(a) $\delta(x, x) = 1 = 1_\Omega$,
(b) $\delta(x, y) = \delta(y, x)$,
(c) $\delta(x, y) \otimes \delta(y, z) \leq \delta(x, z)$.

A morphism $f : (A, \delta) \to (B, \gamma)$ in $\mathbf{SetF}(\Omega)$ is a map $f : A \to B$ such that $\gamma(f(x), f(y)) \geq \delta(x, y)$ for all $x, y \in A$. In case that Ω is a complete Heyting algebra this category is well known and in this case it is a topos. In a general case it can be proved that a category $\mathbf{SetF}(\Omega)$ is complete. For example, if (A_i, δ_i) are Ω-sets for $i \in I$, their product

in $\mathbf{SetF}(\Omega)$ is an Ω-set (A,δ), where $A=\prod_{i\in I}A_i$ and $\delta(\mathbf{a},\mathbf{b})=\bigwedge_{i\in I}\delta_i(a_i,b_i)$. A special object of a category $\mathbf{SetF}(\Omega)$ is $(\Omega,\leftrightarrow)$, where $\alpha\leftrightarrow\beta=(\alpha\to\beta)\wedge(\beta\to\alpha)$.

In any category $\mathbf{K}$ with Ω-sets as objects and containing special object $(\Omega,\leftrightarrow)$ a notion of fuzzy set can be also introduced as follows.

Definition 1.1 *Let $\mathbf{K}$ be a category with Ω-sets as objects. Then a fuzzy set s in an object (A,δ) in a category $\mathbf{K}$ (in symbol $s\subset_{\sim\mathbf{K}}(A,\delta)$) is a morphism*

$$s:(A,\delta)\to(\Omega,\leftrightarrow)$$

in $\mathbf{K}$, where $\alpha\leftrightarrow\beta=(\alpha\to\beta)\wedge(\beta\to\alpha)$.

A fuzzy set $s\subset_{\sim\mathbf{K}}(A,\delta)$ in a category $\mathbf{K}=\mathbf{SetF}(\Omega)$ is then a map $s:A\to\Omega$ such that $s(x)\otimes\delta(x,y)\le s(y)$ for all $x,y\in A$, i.e. s is an extensional map.

Now by using a category $\mathbf{SetF}(\Omega)$ a model theory of a language J based on sets with similarity relations can be defined.

Definition 1.2 *A model of a language J in a category $\mathbf{SetF}(\Omega)$ is*

$$\mathcal{D}=((A,\delta),\{P_{\mathcal{D}}:P\in\mathcal{P}\},\{f_{\mathcal{D}}:f\in\mathcal{F}\}),$$

where

(a) *(A,δ) is an Ω-set from a category $\mathbf{SetF}(\Omega)$,*

(b) *$P_{\mathcal{D}}\subset_{\sim\mathbf{SetF}(\Omega)}(A,\delta)\times\cdots\times(A,\delta)$,*

(c) *$f_{\mathcal{D}}:(A,\delta)\times\cdots\times(A,\delta)\to(A,\delta)$ is a morphism in a category $\mathbf{SetF}(\Omega)$.*

To define an interpretation of terms and formulas in a category $\mathbf{SetF}(\Omega)$ we use a following simple notation. Let X be a set, then a product $(A,\delta)^X$ in a category $\mathbf{SetF}(\Omega)$ is the object (A^X,δ_X), where $\delta_X(\mathbf{a},\mathbf{b})=\bigwedge_{x\in X}\delta(a_x,b_x)$ for all $\mathbf{a},\mathbf{b}\in A^X$. If $|X|=n$ then instead of $(A,\delta)^X=(A^X,\delta_X)$ we sometimes use $(A,\delta)^n=(A^n,\delta_n)$.

Let $t=f(t_1,\dots,t_n)$ be a term, where t_i are terms and f is a functional symbol. Let X be a set of variables which contains all variables of t. Then by induction principle the interpretation $\|t_i\|_{\mathcal{D}}=\|t_i\|_{\mathcal{D},X}$ is already defined as a morphism $\|t_i\|_{\mathcal{D}}:(\mathcal{D}(X),\delta_X)\to(A,\delta)$ in a category $\mathbf{SetF}(\Omega)$ and we can consider a product morphism $\prod_i\|t_i\|_{\mathcal{D}}:(\mathcal{D}(X),\delta_X)\to(A,\delta)^n$. Then we set $\|t\|^X_{\mathcal{D}}=\|t\|_{\mathcal{D}}=f_{\mathcal{D}}\circ\prod_i\|t_i\|_{\mathcal{D}}$.

To define an interpretation of formulas in a category $\mathbf{SetF}(\Omega)$ a classical interpretation in fuzzy sets has to be slightly modified. For example, an interpretation of a formula $\psi\equiv\sigma\Rightarrow\tau$ in a model $\mathcal{D}$ in a category $\mathbf{SetF}(\Omega)$ we cannot define as $\|\psi\|_{\mathcal{D}}(\mathbf{a})=\|\sigma\|_{\mathcal{D}}(\mathbf{a})\to\|\tau\|_{\mathcal{D}}(\mathbf{a})$ because $\|\psi\|_{\mathcal{D}}$ is not a fuzzy set in (A^X,δ_X), in general. Hence, the following modified definition has to be introduced.

Definition 1.3 *Let ψ be a formula and let X be a set of variables containing all free variables of ψ. An interpretation $\|\psi\|^X_{\mathcal{D}}=\|\psi\|_{\mathcal{D}}$ is then defined as follows.*

(a) *Let $\psi\equiv P(t_1,\dots,t_n)$. Then $\|t_i\|_{\mathcal{D}}:(A,\delta)^X\to(A,\delta)$ is already defined and we set $\|\psi\|_{\mathcal{D}}=P_{\mathcal{D}}\circ\prod_i\|t_i\|_{\mathcal{D}}$.*

(b) *Let $\psi\equiv(t_1=t_2)$, where t_i are terms. Then for $\mathbf{a}\in\mathcal{D}(X)$ we set*

$$\|\psi\|_{\mathcal{D}}(\mathbf{a})=$$
$$=\bigvee_{\mathbf{x}\in\mathcal{D}(X)}\delta(\|t_1\|_{\mathcal{D}}(\mathbf{x}),\|t_2\|_{\mathcal{D}}(\mathbf{x}))\otimes\delta_X(\mathbf{a},\mathbf{x}).$$

(c) *Let $\psi\equiv\sigma\Rightarrow\tau$. Then we set*

$$\|\psi\|_{\mathcal{D}}(\mathbf{a})=$$
$$=\bigvee_{\mathbf{x}\in\mathcal{D}(X)}(\|\sigma\|_{\mathcal{D}}(\mathbf{x})\to\|\tau\|_{\mathcal{D}}(\mathbf{x}))\otimes\delta_X(\mathbf{x},\mathbf{a}).$$

(d) *Let $\psi\equiv\neg\sigma$. Then we set $\|\psi\|_{\mathcal{D}}(\mathbf{a})=\|\sigma\|_{\mathcal{D}}(\mathbf{a})\to 0$.*

(e) *Let Ψ be a set of formulas with free variables in a set X. Then we set $\|\bigvee\Psi\|_{\mathcal{D}}(\mathbf{a})=\bigvee_{\psi\in\Psi}\|\psi\|_{\mathcal{D}}(\mathbf{a})$.*

(f) *Let Ψ be a set of formulas with free variables in a set X. Then we set $\|\bigwedge\Psi\|_{\mathcal{D}}(\mathbf{a})=\bigwedge_{\psi\in\Psi}\|\psi\|_{\mathcal{D}}(\mathbf{a})$.*

(g) *Let $\psi\equiv\sigma\otimes\tau$. Then we set $\|\psi\|_{\mathcal{D}}(\mathbf{a})=\|\sigma\|_{\mathcal{D}}(\mathbf{a})\otimes\|\tau\|_{\mathcal{D}}(\mathbf{a})$.*

(h) *Let $\psi\equiv(\exists x)\sigma$. Then we set*

$$\|\psi\|_{\mathcal{D}}(\mathbf{a})=\bigvee_{b\in A}\|\sigma[b/x]\|_{\mathcal{D}}(\mathbf{a}).$$

(i) *Let $\psi\equiv(\forall x)\sigma$. Then we set*

$$\|\psi\|_{\mathcal{D}}(\mathbf{a})=\bigwedge_{b\in A}\|\sigma[b/x]\|_{\mathcal{D}}(\mathbf{a}).$$

In [12] we proved the following theorem which shows that the above definition is correct.

Theorem 1.1 *Let ψ be a formula with free variables contained in a set of variables X. Then*

$$\|\psi\|_{\mathcal{D}}\subset_{\sim\mathbf{SetF}(\Omega)}\mathcal{D}(X)$$

.

In this paper we want to deal with homomorphisms between two models of a language J in a category $\mathbf{SetF}(\Omega)$. Firstly we introduce a definition of such homomorphism and then we investigate a relation between interpretations $\|\psi\|_{\mathcal{E}}$ and $\|\psi\|_{\mathcal{G}}$ in a case there exists a homomorphism $f : \mathcal{E} \to \mathcal{G}$. Finally we investigate two examples of special models of a language J in a category $\mathbf{SetF}(\Omega)$. A first model is based on Ω-sets $(F(A,\delta),\sigma)$, where $F(A,\delta)$ is a set of all fuzzy sets in a category $\mathbf{SetF}(\Omega)$ in an object (A,δ), i.e.

$$F(A,\delta) = \{s : s \subseteq_{\sim_{\mathbf{SetF}(\Omega)}} (A,\delta)\},$$

and the other model is based on Ω-sets $(C(A,\delta),\sigma)$, where $C(A,\delta)$ is a set of all "continuous" maps $(A,\delta) \to (\Omega,\leftrightarrow)$, which represent a natural generalization of fuzzy sets in (A,δ).

2 Homomorphisms of models

An interpretation of terms and formulas in a category $\mathbf{SetF}(\Omega)$ is based significantly on a set X of variables containing all free variables of corresponding formulas. The following proposition describes this relation more precisely.

Proposition 2.1 *Let t be a term and let ψ be a formula in a language J. Let Y be a set of variables containing all free variables of t and ψ and let $\mathcal{D}$ be a model of J based on Ω-set (A,δ). Let $Y \subseteq X$ and let p_{XY} be a canonical projection from $(A,\delta)^X$ to $(A,\delta)^Y$.*

(a) The following diagram commutes

$$\begin{array}{ccc} (A,\delta)^X & \xrightarrow{\|t\|^X_{\mathcal{D}}} & (A,\delta) \\ {\scriptstyle p_{XY}}\downarrow & & \| \\ (A,\delta)^Y & \xrightarrow{\|t\|^Y_{\mathcal{D}}} & (A,\delta) \end{array}$$

(b) For all $\mathbf{x} \in (A,\delta)^X$ we have

$$\|\psi\|^Y_{\mathcal{D}} \circ p_{XY}(\mathbf{x}) \geq \|\psi\|^X_{\mathcal{D}}(\mathbf{x}).$$

Proof. Since for $\mathbf{a},\mathbf{b} \in A^X$ we have

$$\delta_Y(p_{XY}(\mathbf{a}), p_{XY}(\mathbf{b})) = \bigwedge_{y\in Y} \delta(a_y,b_y) \geq$$
$$\geq \bigwedge_{x\in X} \delta(a_x,b_x) = \delta_X(\mathbf{a},\mathbf{b}),$$

it follows that p_{XY} is a morphism in $\mathbf{SetF}(\Omega)$ and the rest follows by simple computation. Let us mention that the statement (b) is equivalent to commutativity of the following diagrams, where $\|\psi\|_\alpha$ are α-cuts of a fuzzy set $\|\psi\|$ for any $\alpha \in \Omega$.

$$\begin{array}{ccc} \|\psi\|^X_\alpha & \xrightarrow{\subseteq} & (A,\delta)^X \\ {\scriptstyle p_{XY}}\downarrow & & \downarrow{\scriptstyle p_{XY}} \\ \|\psi\|^Y_\alpha & \xrightarrow{\subseteq} & (A,\delta)^Y \end{array}$$

We now introduce a slight modification of Definition 3 which will define a generalized interpretation of a formula in a model.

Definition 2.1 *Let ψ be a formula and let X be a set of variables containing all free variables of ψ. Let $\mathcal{D}$ be a model of J based on an Ω-set (A,δ) and let $B \subseteq A$. An interpretation $\|\psi\|_B = \|\psi\|^X_{B,\mathcal{D}}$ is then defined as follows.*

(a) Let $\psi \equiv (t_1 = t_2)$, where t_i are terms. Then for $\mathbf{a} \in \mathcal{D}(X)$ we set

$$\|\psi\|_B(\mathbf{a}) =$$
$$= \bigvee_{\mathbf{x}\in B^X} \delta(\|t_1\|_{\mathcal{D}}(\mathbf{x}), \|t_2\|_{\mathcal{D}}(\mathbf{x})) \otimes \delta_X(\mathbf{a},\mathbf{x}).$$

(b) Let $\psi \equiv \sigma \Rightarrow \tau$. Then we set

$$\|\psi\|_B(\mathbf{a}) =$$
$$= \bigvee_{\mathbf{x}\in B^X} (\|\sigma\|_B(\mathbf{x}) \to \|\tau\|_B(\mathbf{x})) \otimes \delta_X(\mathbf{x},\mathbf{a}).$$

(c) Let $\psi \equiv (\exists x)\sigma$. Then we set

$$\|\psi\|_B(\mathbf{a}) = \bigvee_{b\in B} \|\sigma[b/x]\|_B(\mathbf{a}).$$

(d) For other formulas ψ we set $\|\psi\|_B = \|\psi\|^X_{\mathcal{D}}$ according to Definition 3.

Instead of $\|\psi\|_{A,\mathcal{D}}$ we write simply $\|\psi\|_{\mathcal{D}} (= \|\psi\|^X_{\mathcal{D}})$.

The proof of the following proposition is the same as in Theorem 1.

Proposition 2.2 *Let ψ be a formula with free variables contained in a set X and let $\mathcal{D}$ be a model of J based on Ω-set (A,δ). Let $B \subseteq A$. Then $\|\psi\|_B \subseteq_{\sim_{\mathbf{SetF}(\Omega)}} \mathcal{D}(X)$.*

In a following definition we introduce a notion of a model homomorphism of models based on Ω-sets.

Definition 2.2 *Let $\mathcal{D}_1 = ((A,\delta),\{P_A : P \in \mathcal{P}\},\{f_A : f \in \mathcal{F}\})$, $\mathcal{D}_2 = ((B,\gamma),\{P_B : P \in \mathcal{P}\},\{f_B : f \in \mathcal{F}\})$ be models of J based on Ω-sets. Then φ is a homomorphism from $\mathcal{D}_1$ to $\mathcal{D}_2$ if $\varphi : (A,\delta) \to (B,\gamma)$ is a morphism*

in a category $\mathbf{SetF}(\Omega)$ *such that the following diagrams of morphisms from* $\mathbf{SetF}(\Omega)$ *commute.*

$$\begin{array}{ccccc} (A,\delta)^n & \xrightarrow{f_A} & (A,\delta) & , & (A,\delta)^m \xrightarrow{P_A} (\Omega,\leftrightarrow) \\ {\scriptstyle\varphi^n}\downarrow & & \downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\varphi^m} \quad \| \\ (B,\gamma)^n & \xrightarrow{f_B} & (B,\gamma) & , & (B,\gamma)^m \xrightarrow{P_B} (\Omega,\leftrightarrow) \end{array}$$

Here φ^n is the unique morphism in a category $\mathbf{SetF}(\Omega)$ such that the following diagram (with projection morphisms) commutes.

$$\begin{array}{ccc} (A,\delta) & \xrightarrow{\varphi} & (B,\gamma) \\ {\scriptstyle pr_A^n}\uparrow & & \uparrow{\scriptstyle pr_B^n} \\ (A,\delta)^n & \xrightarrow{\varphi^n} & (B,\gamma)^n \end{array}$$

Proposition 2.3 *Let* $\mathcal{D}_1 = ((A,\delta),\{P_A : P\in\mathcal{P}\},\{f_A : f\in\mathcal{F}\})$, $\mathcal{D}_2 = ((B,\gamma),\{P_B : P\in\mathcal{P}\},\{f_B : f\in\mathcal{F}\})$ *be models of* J *based on* Ω*-sets and let* $\varphi : \mathcal{D}_1 \to \mathcal{D}_2$ *be a model homomorphism. Let* t *be a term in* J *with a set of variables contained in a set* X*. Then the following diagram commutes.*

$$\begin{array}{ccc} (A,\delta)^X & \xrightarrow{\|t\|_{\mathcal{D}_1}} & (A,\delta) \\ {\scriptstyle\varphi^X}\downarrow & & \downarrow{\scriptstyle\varphi} \\ (B,\gamma)^X & \xrightarrow{\|t\|_{\mathcal{D}_2}} & (B,\gamma). \end{array}$$

Proof. Let $t = f(t_1,\dots,t_n)$. A statement then follows from commutativity of the following diagram

$$\begin{array}{ccccc} (A,\delta)^X & \xrightarrow{\prod_i \|t_i\|_{\mathcal{D}_1}} & (A,\delta)^n & \xrightarrow{f_A} & (A,\delta) \\ {\scriptstyle\varphi^X}\downarrow & & {\scriptstyle\varphi^n}\downarrow & & \downarrow{\scriptstyle\varphi} \\ (B,\gamma)^X & \xrightarrow{\prod_i \|t_i\|_{\mathcal{D}_2}} & (B,\gamma)^n & \xrightarrow{f_B} & (B,\gamma). \end{array}$$

We say that a morphism $\varphi : (A,\delta)\to(B,\gamma)$ is *strong* if

$$(\forall a,b\in A)\quad \gamma(\varphi(a),\varphi(b)) = \delta(a,b).$$

It is clear that if φ is strong then $\varphi^X : (A,\delta)^X \to (B,\gamma)^X$ is also a strong morphism.

Theorem 2.1 *Let* $\mathcal{D}_1 = ((A,\delta),\{P_A : P\in\mathcal{P}\},\{f_A : f\in\mathcal{F}\})$, $\mathcal{D}_2 = ((B,\gamma),\{P_B : P\in\mathcal{P}\},\{f_B : f\in\mathcal{F}\})$ *be models of* J *based on* Ω*-sets and let* $\varphi : \mathcal{D}_1 \to \mathcal{D}_2$ *be a model homomorphism such that* φ *is strong. Then for any formula* ψ *with a set of free variables contained in a set* X *the following diagram commutes.*

$$\begin{array}{ccc} (A,\delta)^X & \xrightarrow{\|\psi\|_{\mathcal{D}_1}} & (\Omega,\leftrightarrow) \\ {\scriptstyle\varphi^X}\downarrow & & \| \\ (B,\gamma)^X & \xrightarrow{\|\psi\|_{\varphi(A),\mathcal{D}_2}} & (\Omega,\leftrightarrow). \end{array}$$

Proof. (1) Let $\psi \equiv P(t_1,\dots,t_n)$. Then for $\mathbf{a}\in A^X$ we have

$$\|\psi\|_{\varphi(A),\mathcal{D}_2}\circ\varphi^X(\mathbf{a}) = P_B\circ\prod_i \|t_i\|_{\mathcal{D}_2}(\varphi^X(\mathbf{a})) =$$
$$= P_A(\|t_1\|_{\mathcal{D}_1}(\mathbf{a}),\dots,\|t_n\|_{\mathcal{D}_1}(\mathbf{a})) = \|\psi\|_{\mathcal{D}_1}(\mathbf{a}).$$

(2) Let $\psi \equiv t_1 = t_2$. Then for $\mathbf{a}\in A^X$ we have

$$\|\psi\|_{\varphi(A),\mathcal{D}_2}\circ\varphi^X(\mathbf{a}) =$$
$$= \bigvee_{\mathbf{x}\in A^X} \gamma(\|t_1\|_{\mathcal{D}_2}(\varphi^X(\mathbf{x})),\|t_2\|_{\mathcal{D}_2}(\varphi^X(\mathbf{x})))\otimes$$
$$\otimes\gamma_X(\varphi^X(\mathbf{a}),\varphi^X(\mathbf{x})) =$$
$$= \bigvee_{\mathbf{x}\in A^X} \delta(\|t_1\|_{\mathcal{D}_1}(\mathbf{x}),\|t_2\|_{\mathcal{D}_1}(\mathbf{x}))\otimes\delta_X(\mathbf{a},\mathbf{x}) =$$
$$= \|\psi\|_{\mathcal{D}_1}(\mathbf{a}).$$

The rest of the proof can be done similarly.

3 Models based on special Ω-sets

3.1 Models based on $S(A,\delta)$

We start with a definition of complete sets in Ω-sets and continuous maps. Recall firstly that a subset $S\subseteq A$ is called *complete* in an Ω-set (A,δ) if

$$S = \overline{S} = \{a\in A : 1 = \bigvee_{x\in S}\delta(a,x)\}.$$

Then $X\mapsto\overline{X}$ is a closure system (see [8]). To be more precise we use sometimes a notation $\overline{X}^\delta$ instead of $\overline{X}$. Then a map $\varphi : A\to B$ is called a *continuous* map from (A,δ) to (B,γ) if for any $X\subseteq A$ we have $\varphi(\overline{X}^\delta)\subseteq\overline{\varphi(X)}^\gamma$.

Lemma 3.1 *Let* $\varphi : (A,\delta)\to(B,\gamma)$ *be a morphism in a category* $\mathbf{SetF}(\Omega)$*. Then* φ *is a continuous map.*

It should be mentioned that there exist continuous maps which are not morphisms in a category $\mathbf{SetF}(\Omega)$.

Example 3.1

In fact, let Ω be a residuated lattice such that there exists an element α with a property $\alpha^2 > \alpha^3$. Let $A = \{a_1,a_2\}$ and $\delta(x,y) = 1$, iff $x = y$, otherwise $\delta(x,y) = \alpha$. Then (A,δ) is an Ω-set. Let us define a map $t : A\to\Omega$ such that $t(a_1) = \alpha, t(a_2) = \alpha^3$. Then $t(\overline{X}^\delta)\subseteq\overline{t(X)}^{\leftrightarrow}$ for any $X\subseteq A$ and t is a continuous map $(A,\delta)\to(\Omega,\leftrightarrow)$. On the other hand we have $\alpha^2 = t(a_1)\otimes\delta(a_1,a_2)\not\leq\alpha^3$ and it follows that t is not a morphism in $\mathbf{SetF}(\Omega)$.

For any Ω-set (A,δ) let $S(A,\delta)$ be the set of all continuous maps $(A,\delta)\to(\Omega,\leftrightarrow)$. Moreover, for any morphism φ :

$(A,\delta) \to (B,\gamma)$ in a category $\mathbf{SetF}(\Omega)$ we can define a map $S(\varphi) : S(A,\delta) \to S(B,\gamma)$ such that

$$(\forall t \in S(A,\delta))(\forall b \in B) \quad S(\varphi)(t)(b) = \\ = \bigvee_{x\in A} t(x) \otimes \gamma(b,\varphi(x)).$$

This definition is correct since $S(\varphi)(t) : (B,\gamma) \to (\Omega,\leftrightarrow)$ is a morphism and, hence it is continuous. Then we obtain the following proposition.

Proposition 3.1 *S is a functor from a category $\mathbf{SetF}(\Omega)$ to the category of sets.*

Let $\mathcal{D} = ((A,\delta),\{P_A : P \in \mathcal{P}\},\{f_A : f \in \mathcal{F}\})$ be a model of J. We will now construct a model

$$S(\mathcal{D}) = ((S(A,\delta),\sigma),\{P_{S(A,\delta)} : P \in \mathcal{P}\}, \\ \{f_{S(A,\delta)} : f \in \mathcal{F}\})$$

which will be based on Ω-set $(S(A,\delta),\sigma)$, where $\sigma(s,t) = \bigwedge_{x\in A} s(x) \leftrightarrow t(x)$ is a similarity relation.

Let $f \in \mathcal{F}$. Then $f_{S(A,\delta)} : (S(A,\delta),\sigma)^n \to (S(A,\delta),\sigma)$ is defined such that

$$(\forall \mathbf{s} \in S(A,\delta)^n)(\forall a \in A) \quad f_{S(A,\delta)}(\mathbf{s})(a) = \\ = \bigvee_{\mathbf{x}\in A^n} \mathbf{s}(\mathbf{x}) \otimes \delta(a, f_A(\mathbf{x})),$$

where $\mathbf{s}(\mathbf{x}) = \bigwedge_i s_i(x_i)$ for $\mathbf{s} = (s_1,\dots,s_n) \in (S(A,\delta),\sigma)^n$ and $\mathbf{x} = (x_1,\dots,x_n) \in A^n$.

Further, for $P \in \mathcal{P}$ we define $P_{S(A,\delta)} : (S(A,\delta),\sigma)^n \to (\Omega,\leftrightarrow)$ such that

$$(\forall \mathbf{s} \in S(A,\delta)^n) \quad P_{S(A,\delta)}(\mathbf{s}) = \bigvee_{\mathbf{x}\in A^n} P_A(\mathbf{x}) \otimes \mathbf{s}(\mathbf{x}).$$

Theorem 3.1 *Let $\mathcal{D} = ((A,\delta),\{P_A : P \in \mathcal{P}\},\{f_A : f \in \mathcal{F}\})$ be a model of J. Then $S(\mathcal{D}) = ((S(A,\delta),\sigma),\{P_{S(A,\delta)} : P \in \mathcal{P}\},\{f_{S(A,\delta)} : f \in \mathcal{F}\})$ is a model of J and a singleton map $\{-\} : A \to S(A,\delta)$ is a strong homomorphism $\mathcal{D} \to S(\mathcal{D})$.*

Proof. A singleton map $\{-\}$ is defined such that $\{-\}(a)(b) = \delta(a,b)$ and since $\sigma(\{-\}(a),\{-\}(b)) = \bigwedge_{x\in A} \delta(a,x) \leftrightarrow \delta(b,x) = \delta(a,b)$ it follows that $\{-\} : (A,\delta) \to (S(A,\delta),\sigma)$ is a strong morphism. We prove firstly that $f'(\mathbf{s}) = f_{S(A,\delta)}(\mathbf{s})$ is continuous for any $\mathbf{s} \in S(A,\delta)^n$. Let $X \subseteq A$ and let $a \in \overline{X}^{\delta}$. Then we have

$$\bigvee_{x\in X} f'(\mathbf{s})(x) \leftrightarrow f'(\mathbf{s})(x) = \\ \bigvee_{x\in X} ((\bigvee_{\mathbf{y}\in A^n} \mathbf{s}(\mathbf{y}) \otimes \delta(a, f_A(\mathbf{y}))) \leftrightarrow \\ \leftrightarrow (\bigvee_{\mathbf{z}\in A^n} \mathbf{s}(\mathbf{z}) \otimes \delta(x, f_A(\mathbf{z})))) \geq \\ \geq \bigvee_{x\in X} \bigwedge_{\mathbf{y}\in A^n} \mathbf{s}(\mathbf{y}) \otimes \delta(a, f_A(\mathbf{y})) \leftrightarrow \\ \leftrightarrow \mathbf{s}(\mathbf{y}) \otimes \delta(x, f_A(\mathbf{y})) \geq \\ \geq \bigvee_{x\in X} \bigwedge_{\mathbf{y}\in A^n} \delta(a, f_A(\mathbf{y})) \leftrightarrow \delta(x, f_A(\mathbf{y})) \geq \\ \geq \bigvee_{x\in X} \delta(a,x) = 1.$$

Further, $f' = f_{S(A,\delta)} : (S(A,\delta),\sigma)^n \to (S(A,\delta),\sigma)$ is a morphism in $\mathbf{SetF}(\Omega)$. In fact, for $\mathbf{s},\mathbf{t} \in S(A,\delta)^n$ we have

$$\sigma(f'(\mathbf{s}), f'(\mathbf{t})) = \\ \bigwedge_{a\in A} ((\bigvee_{\mathbf{x}\in A^n} \mathbf{s}(\mathbf{x}) \otimes \delta(a, f_A(\mathbf{x}))) \leftrightarrow \\ \leftrightarrow (\bigvee_{\mathbf{y}\in A^n} \mathbf{t}(\mathbf{y}) \otimes \delta(a, f_A(\mathbf{y})))) \geq \\ \geq \bigwedge_{a\in A} \bigwedge_{\mathbf{x}\in A^n} \mathbf{s}(\mathbf{x}) \otimes \delta(a, f_A(\mathbf{x})) \leftrightarrow \\ \leftrightarrow \mathbf{t}(\mathbf{x}) \otimes \delta(a, f_A(\mathbf{x})) \geq \sigma_n(\mathbf{s},\mathbf{t}).$$

Further, the following diagram commutes.

$$\begin{array}{ccc} (A,\delta)^n & \xrightarrow{f_A} & (A,\delta) \\ {\scriptstyle \{-\}^n}\downarrow & & \downarrow{\scriptstyle \{-\}} \\ (S(A,\delta),\sigma)^n & \xrightarrow{f_{S(A,\delta)}} & (S(A,\delta),\sigma). \end{array}$$

In fact, for $\mathbf{a} \in A^n$ and $x \in A$ we have

$$\{-\} \circ f_A(\mathbf{a})(x) = \delta(f_A(\mathbf{a}), x), \\ \delta(x, f_A(\mathbf{a})) \leq f' \circ \{-\}^n(\mathbf{a})(x) = \\ = \bigvee_{\mathbf{y}\in A^n} \delta_n(\mathbf{a},\mathbf{y}) \otimes \delta(x, f_A(\mathbf{y})) \leq \\ \leq \bigvee_{\mathbf{y}\in A^n} \delta(f_A(\mathbf{a}), f_A(\mathbf{y})) \otimes \delta(x, f_A(\mathbf{y})) \leq \delta(f_A(\mathbf{a}), x).$$

Now we prove that $P' = P_{S(A,\delta)}$ is a morphism in $\mathbf{SetF}(\Omega)$. In fact, for $\mathbf{s}, \mathbf{t} \in S(A,\delta)^n$ we have

$$\begin{aligned} P'(\mathbf{s}) \otimes \sigma_n(\mathbf{s},\mathbf{t}) \leq \\ \leq \bigvee_{\mathbf{x}\in A^n} P_A(\mathbf{x}) \otimes \mathbf{s}(\mathbf{x}) \otimes \bigwedge_{\mathbf{y}\in A^n} (\mathbf{s}(\mathbf{y}) \leftrightarrow \mathbf{t}(\mathbf{y})) \leq \\ \leq \bigvee_{\mathbf{x}\in A^n} P_A(\mathbf{x}) \otimes \mathbf{s}(\mathbf{x}) \otimes (\mathbf{s}(\mathbf{x}) \leftrightarrow \mathbf{t}(\mathbf{x})) \leq \\ \leq \bigvee_{\mathbf{x}\in A^n} P_A(\mathbf{x}) \otimes \mathbf{t}(\mathbf{x}) = P'(\mathbf{t}). \end{aligned}$$

Finally, since $P_A(\mathbf{a}) \leq P' \circ \{-\}^n(\mathbf{a}) = \bigvee_{\mathbf{x}\in A^n} P_A(\mathbf{x}) \otimes \{-\}^n(\mathbf{a})(\mathbf{x}) \leq P_A(\mathbf{a})$, we obtain that the following diagram commutes.

$$\begin{array}{ccc} (A,\delta)^n & \xrightarrow{P_A} & (\Omega,\leftrightarrow) \\ {\scriptstyle \{-\}^n}\downarrow & & \| \\ (S(A,\delta),\sigma)^n & \xrightarrow{P_{S(A,\delta)}} & (\Omega,\leftrightarrow) \end{array}$$

3.2 Models based on $F(A,\delta)$

Let $\mathcal{D} = ((A,\delta), \{P_A : P \in \mathcal{P}\}, \{f_A : f \in \mathcal{F}\})$ be a model of J. In this part we will construct an analogy of a model $S(\mathcal{D})$ from the previous subsection that will be based on a set $F(A,\delta)$ of all fuzzy sets in (A,δ), i.e. we will consider a model

$$\begin{aligned} F(\mathcal{D}) = ((F(A,\delta),\sigma), \{P_{F(A,\delta)} : P \in \mathcal{P}\}, \\ \{f_{F(A,\delta)} : f \in \mathcal{F}\}) \end{aligned}$$

where $f_{F(A,\delta)}$ and $P_{F(A,\delta)}$ are defined formally by the same rules as $P_{S(A,\delta)}$ and $f_{S(A,\delta)}$ in Section 3.1.

Theorem 3.2 *Let $\mathcal{D} = ((A,\delta), \{P_A : P \in \mathcal{P}\}, \{f_A : f \in \mathcal{F}\})$ be a model of J. Then $F(\mathcal{D}) = ((F(A,\delta),\sigma), \{P_{F(A,\delta)} : P \in \mathcal{P}\}, \{f_{F(A,\delta)} : f \in \mathcal{F}\})$ is a model of J and a singleton map $\{-\} : A \to F(A,\delta)$ is a strong homomorphism $\mathcal{D} \to F(\mathcal{D})$.*

Proof. We need only to prove that $f' = f_{F(A,\delta)} : (F(A,\delta),\sigma)^n \to (F(A,\delta),\sigma)$ is a morphism in $\mathbf{SetF}(\Omega)$ and it can be done analogously as in Theorem 3. The rest will also follow from the proof of Theorem 3.

4 Acknowledgment

This work was partially supported by the project MSM 6198898701 of the MŠMT ČR and Grant 201/07/0191 of the GA ČR.

References

[1] Goguen, J.A., L-fuzzy sets. *J.Math. Anal. Appli. 18*(1967), 145-174.

[2] Higgs, D., *A category approach to Boolean-valued set theory.* University of Waterloo (1973).

[3] Höhle, U., Presheaves over GL-monoids. *Non-Classical Logic and Their Applications to Fuzzy Subsets*, Kluwer Academic Publ., Dordrecht, New York(1995), 127-157.

[4] Höhle, U., M-Valued sets and sheaves over integral, commutative cl-monoids. *Applications of Category Theory to Fuzzy Subsets*, Kluwer Academic Publ., Dordrecht, Boston (1992), 33-72.

[5] Höhle, U., Classification of Subsheaves over GL-algebras. *Proceedings of Logic Colloquium 98 Prague*, Springer Verlag (1999).

[6] Höhle, U., Commutative, residuated l-monoids. *Non-Classical Logic and Their Applications to Fuzzy Subsets*, Kluwer Academic Publ. Dordrecht, New York (1995), 53–106.

[7] Höhle, U., Monoidal closed categories, weak topoi and generalized logics. *Fuzzy Sets and Systems* **42**(1991), 15–35.

[8] Močkoř, J., Complete subobjects of fuzzy sets over MV-algebras. *Czech.Math.J.* **129(54)**(2004), 379–392.

[9] Močkoř, J., Extensional subobjects in categories of Ω-fuzzy sets. *Czech.Math.J.* **57(132)**(2007), 631–645.

[10] Močkoř, J., Characteristic morphisms and models of fuzzy logic in a category of sets with similarities. *Theoretical Advances and Applications of Fuzzy Logic and Soft Computing* Springer-Verlag Berlin (2007),832–840.

[11] Močkoř, J., Covariant functors in categories of fuzzy sets over MV-algebras. *Advances in Fuzzy Sets and Systems.* **2(1)**(2006), 83–109.

[12] Močkoř, J., Models of fuzzy logic in a category of sets with similarity relations. *to appear in International Journal of Innnovative Computing, Information and Control*

[13] Novák,V., Perfilijeva, I., Močkoř,J., *Mathematical principles of fuzzy logic* Kluwer Academic Publishers,Boston,Dordrecht, London (1999).

Session 13

Decision Making 2

Fault Diagnosis with Progressive Symptoms Based on Multi-Agent Approach

Oleksandr Sokolov
National Aerospace University
Tchkalov str. 17, Kharkiv 61070
Ukraine
asokolov@xai.edu.ua

Michael Wagenknecht
University of Applied Sciences
Zittau/Goerlitz, IPM
Theodor-Koerner-Allee 16
Zittau 02763, Germany
m.wagenknecht@hs-zigr.de

Ulrike Gocht
University of Applied
Sciences Zittau/Goerlitz, IPM
Theodor-Koerner-Allee 16
Zittau 02763, Germany
u.gocht@hs-zigr.de

Abstract

The paper is devoted to fault diagnosis problems using fuzzy decision making.
We investigate dynamic diagnostic systems which can be represented by symptom-fault rule bases. The main question to be answered is what faults produce observable symptoms in the first moments of their appearance. To solve this task we propose to use a set of agents for making fuzzy hypothesis about symptoms and to solve inverse fuzzy relations.

Keywords: symptom-fault model, agent, fuzzy, diagnosis.

1 Symptom-fault model

We use $F = \{f_l\}_{l=1}^{N_F}$ and $S = \{s_m\}_{m=1}^{N_S}$ to denote the finite sets of all possible faults and symptoms, respectively.

There is also a multi-valued mapping

$$\psi : S \to F \qquad (1)$$

that can be presented as a ***binary diagnostic matrix*** like in Table 1 reflecting the influence of relations between elements of sets S and F as numbers from $I = \{0,1\}$.

The propagation of faults to observable symptoms follows in general physical cause-effect relationships.

Figure 1 shows that a fault in general influences events as intermediate steps which

Table 1: Example of a binary diagnostic matrix

S / F	f_1	...	f_{N_F}
s_1	1	...	0
...	0	...	1
s_{N_S}	1	...	1

then influence the measurable or observable symptoms, both by internal physical properties. The fault diagnosis proceeds the reverse way. It has to conclude from the observed symptoms to the faults [1]. This implies the inversion of the causality. We propose to consider these intermediate events between faults and symptoms as a set of agents that are responsible for observing of symptoms and making hypothesis about possible faults [2].

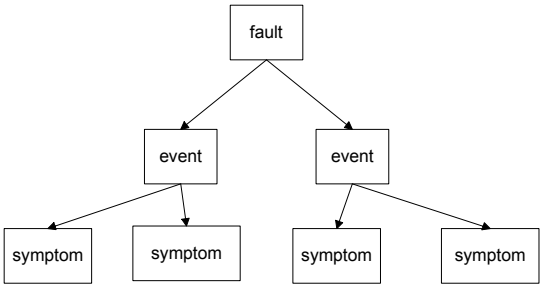


Figure 1: Fault-event-symptom diagram.

Usually the diagnostic methods solve the fault isolation problem using direct mapping ‘symptom-fault’.

The main mission of agents in diagnostic task is to perceive the symptoms on first stages of their arising and recognize caused faults.

Additional advantages of using agents are:

- spatially distributed diagnosis,
- activities of events,
- possibility to communicate [3].

We can formulate the diagnosis tasks according to the multi-agent approach [1]. The general goals of fault diagnosis are:
(1) to detect as many faults as possible;
(2) to detect the faults as soon as possible;
(3) to avoid false alarms.

Let the set of agents given by $A = \{a_i\}_{i=1}^{N_A}$, $N_A \le N_F$ and consider a map

$$\xi : A \to F \tag{2}$$

fulfilling the following conditions

$\xi(A) = F$,
$\forall i, j \ne i \in N_A \ \ \xi(a_i) \ne \xi(a_j)$,
There may exist
$i, j \ne i \in N_F \ \ \xi^{-1}(f_i) = \xi^{-1}(f_j)$

It means that each agent may be responsible for several faults but one fault can be recognized by only one agent. The way in which the set F is divided by agents from A is a state of art of the subject diagnostic area. The natural manner is to divide faults according to subsystems of diagnostic objects [5].

2 Multi-agent diagnosis model

First of all we need to give the mechanism of information of agents *A* about observed symptoms *S*. It may be a new mapping $\vartheta : S \to A$ or a general information about faults in form of logical expression obtained from a binary diagnostic matrix. We consider two tasks.
We suppose here that all agents obtain information about faults from mapping (1) (binary diagnostic matrix).

Let $S_O \subseteq S$ be a subset of symptoms observed by some management application. It reduces to the generation of an appropriate logical expression from (1) in Conjunctive Normal Form, i.e.

from $\psi|_{S_O} : S \to F$ we can get

$$\sigma = \bigvee_{i=1}^{D(S_O)} \left(\bigwedge_{j \in C_i(S_O)} f_j \right) \tag{3}$$

where $D(S_O)$ is a number of disjunctions in the formula, $C_i(S_O)$ is a set of indexes in the i^{th} conjunction expression.

Formula (3) gives the *minimum number of disjunctions* of faults that cause the set of observed symptoms. It is a necessary condition for getting S_O . Of course, the full set of possible combinations of faults (sufficient condition) σ_{full} that could be generated from (3) is expected to be essentially larger, but this is not realistic in practice.

Now we can formulate the task of ***faults refinement*** in terms of *cooperation logic*. According to [4] cooperation logic is a logic that is to enable reasoning about coalitions in multi-agent systems, and in particular, the *power* which such coalition has.

Coalition logic systems are based upon the notion of a cooperation modality: a unary modal operator, indexed by a set of agents which is used to represent the fact that this set of agents can cooperate so as to make ***true*** the state of affairs given as an argument to the operator.

In Coalition logic, for example, a formula $[1,2](p \wedge q)$ is used to express the fact that the coalition of agents $\{1,2\}$ can cooperate in such a way as to make the formula $(p \wedge q)$ *true.*

The key idea of using coalition logic in multi-agent fault diagnosis is that each agent is assumed to verify a set of faults assigned to this agent by mapping (2). The strategies (choices) available to an agent then correspond to the different possible assignments of truth of falsity to these propositions. On top of that, the ability of a coalition to bring about some state of affairs derived from fault variables that are under the overall control of the coalition.

Within such an approach we can consider the set of faults as a set of propositional variables that has *true* or *false* values.

We now introduce a model of coalition logic as

$$M = \langle A, F, Af_1, \ldots, Af_{N_A}, \theta \rangle \tag{4}$$

where

- $A = \{a_i\}_{i=1}^{N_A}$ is a finite, non-empty set of agents;
- $F = \{f_l\}_{l=1}^{N_F}$ is a finite, non-empty set of faults (propositional variables);
- $Af_1, \ldots, Af_{N_A}$ is a partition of F among the members of A, with the intended interpretation that Af_i is a subset of F representing the variables (faults) under the control of agent $a_i \in A$;
- $\theta : F \to \{true, false\}$ is a propositional valuation function, which determines the ***initial truth value*** of every propositional variable (fault).

We have also (see Section 1)

$F = \bigcup_{i=1}^{N_A} Af_i$, i.e., every fault is controlled by some agent, and $Af_i \cap Af_j = \varnothing$ for $a_i \neq a_j \in A$. That is, no fault is controlled by more than one agent.

In case of fault diagnosis tasks we suppose that initial truth value is *false* for all faults from F.

Now a *coalition* C is simply a subset of A, i.e. $C \subseteq A$. Moreover, we set $F_C = \bigcup_{a_i \in C} Af_i$.

Given a model (4) and a coalition C in M, a C*-valuation* is a function:

$$\theta_C : F_C \to \{true, false\}.$$

Thus C*-valuation* is a function that assign truth values to just the primitive propositions controlled by the members of the coalition C. In fault diagnosis tasks, a C*-valuation* is simply checking of fault existence for appropriate agents.

Let us denote by $M \oplus \theta_C$ the model which is identical to M except that the values assigned by its valuation function to propositions controlled by members of C are determined by θ_C.

Given a model $M = \langle A, F, Af_1, \ldots, Af_{N_A}, \theta \rangle$ and a formula $\sigma = \bigvee_{i=1}^{D(S_O)} \left(\bigwedge_{j \in C_i(S_O)} f_j \right)$ we write $M \models^d \sigma$ to express that σ is satisfied (or, equivalently, *true*) in M, under the 'direct' semantics. The rules defining the satisfaction relation $\models^d$ are as follows:

- $M \models^d \top$;
- $M \models^d f$ iff $\theta(f) = true$ (where $f \in F$);
- $M \models^d \neg\varphi$ iff $M \not\models^d \varphi$;
- $M \models^d \varphi \vee \psi$ iff $M \models^d \varphi$ or $M \models^d \psi$;
- $M \models^d \Diamond_C \varphi$ iff there exists a C-*valuation* θ_C such that $M \oplus \theta_C \models^d \varphi$.

The diagnostic task can be formulated now as following:

for model $M = \langle A, F, Af_1, \ldots, Af_{N_A}, \theta \rangle$ and formula $\sigma = \bigvee_{i=1}^{D(S_O)} \left(\bigwedge_{j \in C_i(S_O)} f_j \right)$ it is necessary to find a minimal (in special sense) coalition of agents $C \subseteq A$ for that $M \models^d \Diamond_C \sigma$ remains true. The goal of minimization may be

$$\arg\min_C M \models^d \Diamond_C \sigma \qquad (5a)$$

or

$$\arg\min_F M \models^d \Diamond_{C(F)} \sigma. \qquad (5b)$$

for minimum time.

Because task (5b) is more natural in fault diagnosis we try to find the minimal set of faults that cause observed symptom set S_O.

Let us suppose that all faults are equiprobable and equipollent.

Without the multi-agent approach this task is transformed to the task of checking the faults F_C caused the set S_O element-by-element and each step to test formula (3). Let $F_C = \{F_{pos}, F_{neg}, F_{rest}\}$, where F_{pos} is set of faults taking place and caused symptoms S_O; F_{neg} is the set of faults that were checked, but they do not cause symptoms; F_{rest} is a set possible faults for symptoms S_O. If checking time of one fault is t_f then the time in *successive* algorithm is

$$T_S = t_f \cdot card(F_{pos} \cup F_{neg}).$$

It is easy to see that a *multi-agent* algorithm should fulfill

$$T_{MA} \le T_S .$$

The worst case is $T_{MA} = t_f \cdot \max\limits_{i=1,..,N_A} card\left(Af_i\right)$ and we have an $Af_i = \left\{F_{pos} \cup F_{neg}\right\}$ with $T_{MA} = T_S$. The minimal time in multi-agent approach is $T_{MA} = t_f$ when the number of agents is equal to the number of faults. Thus we have $t_f \le T_{MA} \le T_S$.

The main advantage of the multi-agent approach is parallel calculation and interaction between agents to make decisions.

We consider the fault isolation task as a dynamic process in which diagnostic agents communicate with each other and exchange knowledge to solve the diagnostic problem.

For purpose of communication organization we need to set the protocol of agent's speaking.

Each agent makes decisions about the necessity of checking faults according to the current situation. The agent chooses such a fault that could make formula σ true as soon as possible.

All agents obtain a formula σ for testing.

At the first step (unit of time) agents inform each other for which faults they will minimize time. At the next step they check faults simultaneously by prior arrangement and test formula σ. If consensus is not reached (result is false) σ is reduced according to current values of checked faults and the process is repeated.

We suppose all agents to check their faults ***simultaneously*** at one step and at the next step they exchange messages about the current state of their faults. This procedure will be repeated until consensus is reached. We suppose that this procedure to be finite due to the reliable information of observed symptoms and the truthful binary diagnosis matrix.

The proposed algorithm is presented in Figure 2.

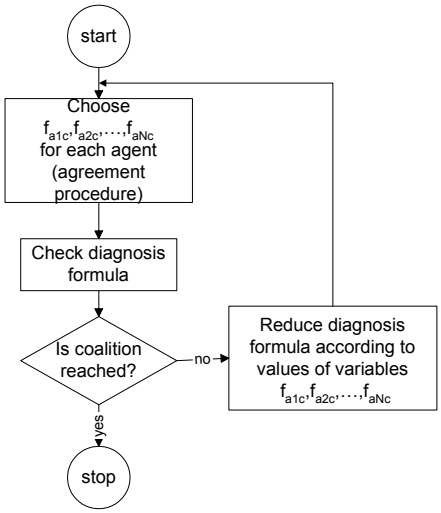


Figure 2: Algorithm of coalition reaching procedure

Let $C = \left\{a_{1_C}, \ldots, a_{N_C}\right\}$ be the set of agents responsible for σ. For each agent a_{j_C} there is the set of corresponding faults $Af_{j_C} = \left\{f_{1_{j_C}}, .., f_{N_{j_C}}\right\} \subset F$.

We use the following heuristic recurrent procedure for choosing the set of testing faults from $F_C = \left\{Af_{1_C}, \ldots Af_{N_C}\right\}$.

Let $\sigma_0 = \bigvee\limits_{i=1}^{D_0}\left(\bigwedge\limits_{j \in C_i^0} f_j\right)$ be an initial formula (3) where $D_0 = D\left(S_O\right)$ is the number of disjunctions in the formula; $C_i^0 = C_i\left(S_O\right)$ is the set of indexes in the i^{th} conjunction expression. Let $F_0 = \left\{F_{1_C}^0, \ldots F_{N_C}^0\right\} = F_C$ be the set of possible faults that caused the observed symptoms distributed on agents from C. Further we propose the procedure that performs by each agent Ag from $C = \left\{a_{1_C}, \ldots, a_{N_C}\right\}$ at the k^{th} step simultaneously.

We call this an α**-procedure**.

1. Refresh

$$\sigma_k = \bigvee_{i=1}^{D_k}\left(\bigwedge_{j \in C_i^k} f_j\right)$$

according to the result of faults checking at the previous step and abbreviate by cancellation excessive variables and false conjunctions (executed one time at each step).

2. ***If*** $F_{Ag}^{k} \neq \varnothing$ ***then*** choose one fault f_{Ag} from F_{Ag}^{k} that belongs to the conjunction in (3) with minimal length $\min card\left(C_i^k\right)$. If there are several conjuctions with the same length we apply a random choice. Set $F_{Ag}^{k+1} = F_{Ag}^{k} \setminus f_{Ag}$ ***else*** *stop* testing process for the current agent.
3. Check f_{Ag} and set $f_{Ag} = \{true, false\}$.
4. Set global state $k = k + 1$.
5. Goto 1

The proposed algorithm is converging due to the reliable set of symptoms, the truthful binary diagnosis matrix and the step-by-step reducing of the set of possible faults.
It is possible to improve the procedure not considering all agents simultaneously, but one after another. In cooperation logic such a model is called β-coalition contrary to the α-coalition for the case described above [5]. The β-coalition foresees the conditional coalition when the current decision strongly depends on the previous one.
The β **-procedure** is as follows.

1. Determine
$$\sigma_k = \bigvee_{i=1}^{D_k}\left(\bigwedge_{j\in C_i^k} f_j\right)$$
from the previous step or after modification at the current step.
2. ***If*** $F_{Ag}^{k} \neq \varnothing$ ***then*** choose one fault f_{Ag} from the set F_{Ag}^{k} that belongs to the conjunciton in (3) with minimal length $\min card\left(C_i^k\right)$. Set $F_{Ag}^{k+1} = F_{Ag}^{k} \setminus f_{Ag}$ ***else*** *stop* testing the process for the current agent.
3. ***If*** $\min card\left(C_i^k\right) = 1$ and $F_{Ag}^{k+1} = \varnothing$ ***then*** set $f_{Ag} = true$ without checking.
else check f_{Ag} and set the appropriate value $f_{Ag} = \{true, false\}$.
4. Refresh
$$\sigma_k^{\,\mathrm{mod}_{Ag}} = \bigvee_{i=1}^{D_k}\left(\bigwedge_{j\in C_i^k} f_j\right)$$
according to the value $f_{Ag} = \{true, false\}$ and abbreviate by cancelling excessive variables and false conjunctions.
5. Send to the next agent (from $C=\{a_{1_C},..,a_{N_C}\}$) the modified formula $\sigma_k^{\,\mathrm{mod}_{Ag}}$. If this agent is last in the list then set $\sigma_{k+1} = \sigma_k^{\,\mathrm{mod}_{Ag}}$.
6. After testing (4.9) by the remaining agents set global state $k = k + 1$.
7. Goto 1

Step 5 demands to set the schedule of agents for sending information. This schedule may be organized by

- a list in $C = \{a_{1_C}, \ldots, a_{N_C}\}$;
- randomly;
- in inverted range of cardinality of elements in $F_k = \{F^{k}{}_{1_C}, \ldots F^{k}{}_{N_C}\}$.

Remark.
Progressive symptoms mean that we use fuzzy logic for their evaluation and calculate the possible set of faults using fuzzy inference.

Conclusion
Due to the distributional character of fault diagnosis and necessity of taking into account different uncertainties during the diagnosis process the multi-agent methodology has been proposed. We showed that the principle of assigning the agent to subsystems in this case is also useful. Then the local knowledge base of agents and rules for communications can precise the possible diagnosis in addition to the ordinary binary or fuzzy diagnostic matrix.

As a result of this investigation we could propose engineering methods for applying multi-agent methods to fault diagnosis problems.

Acknowledgement

This research has partially been supported by the Saxon Ministry of Sciences and Art (Germany).

References

[1] R. Isermann. Fault-Diagnosis Systems. An Introduction from Fault Detection to Fault Tolerance. Springer, 2006.

[2] Soo Young Eo, Tae Suk Chang, Dongil Shin, En Sup Yoon. Cooperative problem solving in diagnostic agents for chemical processes. Computer & Chemical Engineering 24 (2000) 729 – 734.

[3] M. Wooldridge. An Introduction to Multi-Agent Systems. John Wiley&Sons, 2001.

[4] W. Hoek, M. Wooldridge. On the logic of cooperation and propositional control. Artificial intelligence 164 (2005) 81-119.

[5] B.D. Pomeroy, H.A. Spang, M.E. Dausch. Event-based architecture for diagnosis in control advisory systems. Artificial Intelligence in Engineering 5 (1990) 174-181.

Similarity of fuzzy preference structures based on metrics

Dana Hliněná
Dept. of Mathematics
FEEC Brno University of Technology
Technická 8, 616 00 Brno, Czech Rep.
hlinena@feec.vutbr.cz

Pavol Kráľ
Dept. of Quantitative Methods and Information Technology
Faculty of Economics Matej Bel University
Tajovského 10, 975 90 Banská Bystrica, Slovakia
pavol.kral@umb.sk

Abstract

We generalize the solution of a multicriterial optimization problem which has been given in [5]. They have used a comparison of the criterion fuzzy preference relations and the general fuzzy preference relation. Based on this comparison they have constructed the ordering of criteria. Our modification builds on pairwise similarity of criterion related fuzzy preference relations. We illustrate our approach on a practical example.

Keywords: Multicriteria optimization, Preference structures, Fuzzy preference structures, Similarity measures, Complementarity, Redundance.

1 Introduction to preference structures and fuzzy preference structures

A preference structure is a basic concept of preference modeling. In a classical preference structure (PS), a decision-maker makes one of three decisions for each pair (a, b) from the set $\boldsymbol{A}$ of all alternatives. His decision defines a triplet P, I, J of crisp binary relations on $\boldsymbol{A}$:

1) a is preferred to $b \Leftrightarrow (a, b) \in P$ (strict preference).
2) a and b are indifferent $\Leftrightarrow (a, b) \in I$ (indifference).
3) a and b are incomparable $\Leftrightarrow (a, b) \in J$ (incomparability).

A preference structure (PS) on a set $\boldsymbol{A}$ is a triplet (P, I, J) of binary relations on $\boldsymbol{A}$ such that

(ps1) I is reflexive, P and J are irreflexive.

(ps2) P is asymmetric, I and J are symmetric.

(ps3) $P \cap I = P \cap J = I \cap J = \emptyset$.

(ps4) $P \cup I \cup J \cup P^t = \mathbf{A} \times \mathbf{A}$ where $P^t(x, y) = P(y, x)$.

A preference structure can be characterized by the reflexive relation $R = P \cup I$ called the large preference relation. The relation R can be interpreted as

$(a, b) \in R \Leftrightarrow a$ is prefered to b or a and b are indifferent.

It can be easily proved that

$$co(R) = P^t \cup J,$$

where $coR(a, b)$ is the complement of $R(a, b)$ and

$$P = R \cap co(R^t), I = R \cap R^t, J = co(R) \cap co(R^t).$$

This allows us to construct a preference structure (P, I, J) from a reflexive binary operation R only.

Decision-makers are often uncertain or even inconsistent in their judgements. The restriction on two-valued relations has been an important drawback in their practical use. A natural demand led researchers to the introduction of a fuzzy preference structure (FPS). The original idea of using numbers between zero and one to describe the strength of links between two alternatives goes back to Menger. The introduction of fuzzy relations allowed to express degrees of preference, indifference and incomparability. Of course, the attempts to simply replace the notion used in the definition of (PS) by their fuzzy equivalents have brought some problems.

Let us recall the well-known definitions.

Definition 1 *Let X be a nonempty crisp set. A fuzzy set on $X \times X$ (i.e. a $X \times X \to [0, 1]$ mapping) is called binary fuzzy relation on X.*

Let us now repeat the most important properties of fuzzy relations.

Definition 2 *Consider a t−norm T. A binary fuzzy relation $R : X^2 \to [0,1]$ is called*

reflexive, *iff*

$$\forall x \in X : \mu_R(x,x) = 1$$

irreflexive, *iff*

$$\forall x \in X : \mu_R(x,x) = 0$$

symmetric, *iff*

$$\forall x, y \in X : \mu_R(x,y) = \mu_R(y,x)$$

T-asymmetric, *iff*

$$\forall x, y \in X : x \neq y \Rightarrow T(\mu_R(x,y), \mu_{R^{-1}}(x,y)) = 0,$$

T-transitive, *iff*

$$\forall x, y \in X : T((\mu_R(x,y), \mu_R(y,z))) \leq \mu_R(x,z).$$

Definition 3 *A unary operator $n : [0,1] \to [0,1]$ is called a negator if for any $a, b \in [0,1]$ it holds*

(i) $a < b \Rightarrow n(b) \leq n(a)$,

(ii) $n(0) = 1, n(1) = 0$.

A negator n is called *strict* if it is a permutation. A strict negator n is called *strong* if it is *involutive*, i.e. for any $x \in [0,1]; n(n(x)) = x$. It can be easily proved that for any strict negator n, its inverse n^{-1} is also strict negator, and both are continuous. One often used strong negator is $N(x) = 1 - x$.

To define (FPS) it is necessary to consider some fuzzy connectives. We shall consider a continuous De Morgan triple (T, S, N) consisting of a continuous t-norm T, continuous t-conorm S and a strong negator N such that $T(x,y) = N(S(N(x), N(y)))$. The main problem lies in the fact that the completeness condition (ps4) can be written in many forms, e.g.:

$co(P \cup P^t) = I \cup J, P = co(P^t \cup I \cup J), P \cup I = co(P^t \cup J)$.

Let (T,S,N) be De Morgan a triplet. A fuzzy preference structure (FPS) on a set $\boldsymbol{A}$ is a triplet (P, I, J) of binary fuzzy relations on $\boldsymbol{A}$ such that:

(f1) I is reflexive, P and J are irreflexive. $I(a,a) = 1$, $P(a,a) = J(a,a) = 0$

(f2) P is T-asymmetric, I and J are symmetric. $T(P(a,b), P(b,a)) = 0$

(f3) $T(P,I) = T(P,J) = T(I,J) = 0$ for all pairs of alternatives

(f4) $(\forall (a,b) \in A2) S(P, P^t, I, J) = 1$ or $N(S(P,I)) = S(P^t, J)$ or another completeness conditions.

Note that it was proved [1, 9] that reasonable constructions of fuzzy preference structure (FPS) should use a nilpotent t-norm only. Since any nilpotent t-norm (t-conorm) is isomorphic to the Lukasiewicz t-norm (t-conorm), it is enough to restrict our attention to De Morgan triple $(T_L, S_L, 1 - x)$.

2 Concept of similarity

The main aim of our contribution is to apply fuzzy preference structure in multicriteria decision making. We try to find a suitable way of comparing fuzzy preference relations. We illustrate our approach on practical examples. Specifically, we investigate some similarities and we find an appropriate similarity relation, i.e. the similarity relation which better reflects reality in our examples. In the course of our research, we have got the similarities based on metrics, and we have studied those with respect to transitivity.

In the presented paper we will generalize the approach from [5] which is in general based on a special comparison of fuzzy preference relations with respect to general preference relation. The comparison in [5] is defined in the following way

$$X \succ Y \iff \frac{|P_X \cap P_I|}{|P_X \triangle P_I|} > \frac{|P_Y \cap P_I|}{|P_Y \triangle P_I|}, \qquad (1)$$

where X, Y are criteria, P_X, P_Y are the related preference relations, P_I is a preference relation which is based on expert's global score and $\triangle$ is symmetric difference. The basic idea is evident. The greater similarity between P_X and P_I means the greater importance of a criterion. This allows us to construct the ordering of the fuzzy preference relations. The above mentioned formula leads to an ordering of criteria. But as it is presented in the following example, it is not possible to use the previous formula in all cases.

Example 1 *Let us consider six students and their grades in mathematics, physics, informatics, physical education, English, music lessons. Students are evaluated by a test score (mathematics, physics, informatics) and by a qualitative score which is given by teacher (physical education, English, music lessons). Students' names, grades and expert's global score are given in Table 1.*

We can assign a numerical score to each criterion (see Table 2).

We can also construct the following fuzzy preference relations: FP_M, FP_P, FP_I, FP_{PE}, FP_E, FP_{ML},

	mathematics	physics	informatics	physical education
Paul	99	95	100	C
James	90	90	95	B
Eve	80	75	70	A
Jane	75	65	85	C
Patty	70	60	60	B
John	60	65	80	A
	English	music lessons	expert evaluation	rank in the class
Paul	A	C	good	2
James	B	C	good	1
Eve	B	B	average	1
Jane	C	A	average	2
Patty	D	A	bad	1
John	E	D	bad	2

Table 1: Grades of the different students

	mathematics	physics	informatics	physical education
Paul	1	1	1	0.5
James	1	1	1	0.75
Eve	0.75	0.625	0.5	1
Jane	0.625	0.375	0.875	0.5
Patty	0.5	0.25	0.25	0.75
John	0.25	0.375	0.75	1
	English	music lessons	global score	
Paul	1	0.5	1	
James	0.75	0.5	0.875	
Eve	0.75	0.75	0.75	
Jane	0.5	1	0.625	
Patty	0.25	1	0.375	
John	0	0.25	0.25	

Table 2: Numerical scores on criteria

FP_{GS}, which are in a correspondence with our criteria – mathematics, physics, informatics, physical education, English, music lessons and the global score.

For example, the value of fuzzy preference in mathematics (FP_M) for Paul and James is computed from Table 2 as $FP_M(P,J) = \max\{x_M^P - x_M^J, 0\}$, where x_M^P and x_M^J are Paul's and James's mathematics score in Table 2, etc. This is one possible way of fuzzification a preference relation.

Using the formula (compare to (1))

$$X \succ Y \iff \frac{\sum\limits_{i,j} |\{fp_{x_{ij}}\} \cap \{fp_{gs_{ij}}\}|}{\sum\limits_{i,j} |\{fp_{x_{ij}}\} \triangle \{fp_{gs_{ij}}\}|} > \frac{\sum\limits_{i,j} |\{fp_{y_{ij}}\} \cap \{fp_{gs_{ij}}\}|}{\sum\limits_{i,j} |\{fp_{y_{ij}}\} \triangle \{fp_{gs_{ij}}\}|}, \quad (2)$$

where X, Y are our criteria (mathematics, physics, English, informatics, music lessons, physical education), m is a number of alternatives, $fp_{x_{ij}}, fp_{y_{ij}}$ are values of fuzzy preference structures of criterion X, Y, $fp_{gs_{ij}}$ are values of fuzzy preference relation which is based on expert's global score and $i, j \in \{1, 2 \dots m\}$. Note that our intersection $\cap$ and symmetric difference $\triangle$ are ordinary. Now we obtain the following ordering of criteria

$$mathematics \succ physics = English \succ informatics$$

$$\succ music\ lessons \succ physical\ education.$$

It is obvious that the comparison of students which is based on the above mentioned ordering is the same as the previously mentioned expert global evaluation. But such ordering generally does not allow us to solve correctly the problem of comparison of another student

FP_M	Paul	James	Eve	Jane	Patty	John
Paul	0	0	0.25	0.375	0.5	0.75
James	0	0	0.25	0.375	0.5	0.75
Eve	0	0	0	0.125	0.25	0.5
Jane	0	0	0	0	0.125	0.375
Patty	0	0	0	0	0	0.25
John	0	0	0	0	0	0

FP_P	Paul	James	Eve	Jane	Patty	John
Paul	0	0	0.375	0.625	0.75	0.625
James	0	0	0.375	0.625	0.75	0.625
Eve	0	0	0	0.25	0.375	0.25
Jane	0	0	0	0	0.125	0
Patty	0	0	0	0	0	0
John	0	0	0	0	0.125	0

FP_I	Paul	James	Eve	Jane	Patty	John
Paul	0	0	0.5	0.125	0.75	0.25
James	0	0	0.5	0.125	0.75	0.25
Eve	0	0	0	0	0.25	0
Jane	0	0	0.375	0	0.625	0.125
Patty	0	0	0	0	0	0
John	0	0	0.25	0	0.5	0

FP_{PE}	Paul	James	Eve	Jane	Patty	John
Paul	0	0	0	0	0	0
James	0.25	0	0	0.25	0	0
Eve	0.5	0.25	0	0.5	0.25	0
Jane	0	0	0	0	0	0
Patty	0.25	0	0	0.25	0	0
John	0.5	0.25	0	0.5	0.25	0

FP_E	Paul	James	Eve	Jane	Patty	John
Paul	0	0.25	0.25	0.5	0.75	1
James	0	0	0	0.25	0.5	0.75
Eve	0	0	0	0.25	0.5	0.75
Jane	0	0	0	0	0.25	0.5
Patty	0	0	0	0	0	0.25
John	0	0	0	0	0	0

FP_{ML}	Paul	James	Eve	Jane	Patty	John
Paul	0	0	0	0	0	0.25
James	0	0	0	0	0	0.25
Eve	0.25	0.25	0	0	0	0.5
Jane	0.5	0.5	0.25	0	0	0.75
Patty	0.5	0.5	0.25	0	0	0.75
John	0	0	0	0	0	0

FP_{GS}	Paul	James	Eve	Jane	Patty	John
Paul	0	0.125	0.25	0.375	0.625 0.5	0.75
James	0	0	0.125	0.25	0.5	0.625
Eve	0	0	0	0.125	0.375	0.5
Jane	0	0	0	0	0.25	0.375
Patty	0	0	0	0	0	0.125
John	0	0	0	0	0	0

Tables 3, 4, 5, 6, 7, 8, 9: Fuzzy preference relations of criteria and global score

with our students' group. Assume that we have got a new student Jacob with the following scores given in Table 10.

	mathematics	physics	informatics
Jacob	1	0.9	1
	physical education	English	music lessons
Jacob	1	0.9	1

Table 10: Jacob's numerical scores on criteria

If we compare Jacob (we use comparison based on the ordering of criteria) with the other students, we obtain that James > Jacob > Eve. But, in our opinion, such result is counterintuitive, and the correct ordering (coherent with expert evaluation) should be Jacob > James > Eve. In majority of criteria, Jacob is better than James. There is only a small difference in physics and English. Moreover, it seems that English should be more important because of the relation between scores in mathematics and physics. The problem is that Formula 1 determines only the numbers of identical cells and distinct cells in tables of relations, regardless of nature of the "difference". For instance, in some cell, the difference is between 0 and 1, while in another cell, the difference is between 0.99 and 1. Formula 1 treats these two cases as identical.

We assume that such type of problems is the result of high similarity between criterion relations which is not considered in [5]. To avoid these problems we want to modify the ordering of criteria using an appropriate similarity relation. In general the similarity of two fuzzy preference structures can be interpreted as a kind of weak equality. Two fuzzy preference structures can be considered similar because they represent criteria which are in fact an implementation of the same (general) factor. In our case this similarity will be represented by the similarity (tolerance) relation Sim. For simplicity, we restrict ourselves to a finite universe X.

Definition 4 *A binary fuzzy relation Sim on X is a similarity relation if and only if it is reflexive and symmetric.*

It is possible to assume that a similarity relation satisfies also some kind of transitivity. But in our case we assume only pairwise comparison of fuzzy preference relations and the absence of transitivity does not pose any problem. To measure similarity between fuzzy preference relations we will use the following simple similarity relation

$$Sim(FP_X, FP_Y) = 1 - \frac{\sum_{i,j(i\neq j)} |\mathrm{fp}_{x_{ij}} - \mathrm{fp}_{y_{ij}}|}{m^2 - m}, \quad (3)$$

where m is a number of alternatives, $fp_{x_{ij}}(fp_{y_{ij}})$ are values of fuzzy preference structures of criterion X (Y), and $i, j \in \{1, 2 \dots m\}$. It is easy to verify that Sim is a similarity relation. We will use the proposed similarity relation in the following way. First of all we find the ordering of criteria based on the relation given by Formula 1. Then we compute pairwise similarity Sim. If Sim is greater than some threshold α, then the criteria X, Y are redundant, and this causes the bad ordering and subsequently a bad decision (as we can see in Example 1). In our opinion there are two basic possibilities how to solve this problem – we can delete one of the redundant criteria or we change the ordering. We prefer the second one. If we have the following ordering $X > Y > Z$ and we know that $Sim(FP_X, FP_Y)$ is greater then α, we will change the ordering as follows $X > Z > Y$ so that our results can be more appropriate.

Example 2 *Assume the same set of students. Using Formula 3 we can compute the similarities between fuzzy preference relations. We obtain the following similarity table:*

Similarity	FP_M	FP_P	FP_I	FP_{PE}
FP_M	1	0.908333333	0.825	0.729166667
FP_P	0.908333333	1	0.858333333	0.729166667
FP_I	0.825	0.858333333	1	0.729166667
FP_{PE}	0.729166667	0.729166667	0.729166667	1
FP_E	0.929166667	0.879166667	0.7875	0.691666667
FP_{ML}	0.745833333	0.670833333	0.695833333	0.7875
FP_{GS}	0.966666667	0.908333333	0.833333333	0.7375
Similarity	FP_E	FP_{ML}	FP_{GS}	
FP_M	0.929166667	0.745833333	0.966666667	
FP_P	0.879166667	0.670833333	0.908333333	
FP_I	0.7875	0.695833333	0.833333333	
FP_{PE}	0.691666667	0.7875	0.7375	
FP_E	1	0.708333333	0.945833333	
FP_{ML}	0.708333333	1	0.7375	
FP_{GS}	0.945833333	0.7375	1	

Table 11: Similarity table

Using Formula 3 we obtain the following ordering of criteria

$$mathematics \succ English \succ physics \succ informatics$$

$$\succ music\ lessons = physical\ education.$$

We can see that there is high similarity between English and mathematics. So we change the ordering of criteria, which we have obtained using Formula 2. In our opinion this new ordering of criteria better reflects reality in this example.

As we have mentioned above, the absence of a transitivity does not pose any problem in general (for example if A is "near" to B and B is "near" to C A then

C do not need to be "near"). Nevertheless transitivity represents an interesting problem (at least from the theoretical point of view).

It is easy to show that our relation is not T_M-transitive, but it is T_L-transitive. Therefore we obtain that our relation is also $T-$transitive for each $T \leq T_L$. The main advantages of our similarity relation, which is based on the distance are computational simplicity and coherence with our intuition. But the proposed relation has also its drawbacks. It is not influenced by the source of (di)similarity, e.g. it does not matter whether we have $|0.9 - 0.8|$ or $|0.3 - 0.2|$. This can lead to an unappropriate decision (see the following example).

Example 3 *Let X, Y, Z be the arbitrary criteria and FP_X, FP_Y, FP_Z be the related preference relations. Let the difference between FP_X, FP_Y and FP_X, FP_Z be the same and be generated only by $|0.9 - 0.8|$ in the first case and by $|0.3 - 0.2|$ in the second case. Our similarity relation generates the ordering $Y = Z$ (with respect the criterion X). However, it seems that the first source of difference is more important and the ordering should be $Y < Z$.*

We think that such problems can be avoided in the following three ways (and their combinations). First, we can improve it by assuming metrics taking values from the unit interval. The other possibility is to use an aggregation operator to summarize the partial evaluations, e.g. the sum based on $t-$norms. The last one is to use an appropriate weighting function which is dependent on the height of values in summands. It seems that an increasing function f is the most appropriate for our purposes. We focus here our attention on the first alternative. We can redefine the proposed similarity using an arbitrary metric ϱ in the following way:

Example 4 *Let ρ be an arbitrary metric taking values from the unit interval. We can define the metric based similarity as follows*

$$Sim_\varrho(FP_X, FP_Y) = 1 - \frac{\sum_{i,j(i\neq j)} \varrho(\mathrm{fp}_{x_{ij}}, \mathrm{fp}_{y_{ij}})}{m^2 - m}. \quad (4)$$

By triangle inequality we get that Sim_ϱ is T_L- transitive, more Sim_ϱ is also $T-$transitive for each $T \leq T_L$.

Let us considere well-known discrete metric

$$\varrho(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{otherwise.} \end{cases}$$

Similarity Sim_ϱ based on this discrete metric is T_M-transitive, subsequently this similarity is $T-$ transitive for each $t-$norm T. Moreover, we can construct, for each real number $p \in]0, 1]$, a metric ϱ_p using

$$\varrho_p(x, y) = \begin{cases} p & \text{if } x \neq y, \\ 0 & \text{otherwise.} \end{cases}$$

Then similarity Sim_{ϱ_p} is T_M- transitive, too.

The latter two ways, and more properties of Sim_ϱ, will be subject of future research.

3 Conclusions

In multicriterial optimization, one of the tasks is training with input data and finding the right utility function. The first step in determining the right utility function is finding an appropriate ordering of criteria from given expert evaluation. In this step we have to solve the problem with similarity of sets, especially similarity of relations. The main aim of our contribution is to find a suitable way of comparing fuzzy preference relations and to illustrate our approach on practical examples. More, we investigate some kind of transitivity of our similarity relation Sim and we find another suitable similarity relation, an appropriate modification of Sim. We also plan to apply the above mentioned preference relations to financial analysis of companies.

Acknowledgments

Dana Hliněná has been supported by Project 1ET100300517 of the Program "Information Society" and by Project MSM0021630529 of the Ministry of Education. Pavol Kráľ has been supported by the grant VEGA 1/2002/05.

References

[1] J. Fodor and M. Roubens. *Fuzzy preference modeling and Multicriteria Decision Support*. In Kluwer, Dordrecht, 1994.

[2] B. De Baets and R. Mesiar. *Pseudo-metrics and T-equivalences*. In The Journal of Fuzzy Mathematics 5 (2), Los Angeles, 1997.

[3] M. Grabisch and M. Roubens. *An axiomatic approach of interaction in multicriteria decision making*. In 5th Eur. Congr. on Intelligent Techniques and Soft Computing (EUFIT 97), Aachen, Germany, 1997.

[4] M. Grabisch and M. Roubens. *Application of the Choquet integral in multicriteria decision making*. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, Fuzzy Measures and Integrals - Theory and Applications, pages 348–374. Physica Verlag, 2000.

[5] D. Hliněná and P. Vojtáš *A note on an example of use of fuzzy preference structures*. Submitted

[6] T. HORVÁTH, S. KRAJČI, R. LENCSES, P. VOJTÁŠ. *An ILP model for a monotone graded classification problem*. Kybernetika 40,3 (2004), 317–332.

[7] S. KRAJČI, R. LENCSES, & P. VOJTÁŠ. *A comparison of fuzzy and annotated logic programming*, Fuzzy Sets and Systems, 144 (2004), 173–192.

[8] M. ŠABO. *The history and new results in fuzzy preference structures*, In Proccedings of Aplimat, Bratislava, (2005), 531–537.

[9] B. VAN DE WALLE AND B. DE BAETS AND E. KERRE. *A comparative study of completeness conditions in fuzzy preference structures*. In Proc. IFSA'97, Prague, vol III, 74–79.

Supporting Consumer Decision by Fuzzy Revealed Preference in Online Sales

Irina Georgescu
Academy of Economic Studies,
Department of Economic Cybernetics,
P. O. Box 1–432, 014700,
Bucharest, Romania
irina.georgescu@csie.ase.ro

Xuemei Qiu
Turku Centre for Computer Science,
IAMSR, Turku, Finland
xuemei.qiu@abo.fi

Ranran Lin
Helsinki University of Technology,
Laboratory of Electromechanics,
Espoo, Finland
ranran.lin@tkk.fi

Abstract

This paper deals with a model of decision–making in two stages. In the first stage from a set of expertises one chooses the most rational ones, while in the second stage one obtains a ranking of alternatives according to different criteria. The mathematical basis of the model is the revealed preference theory for fuzzy choice functions. With them one can represent the situation when in the choice of the best options vague criteria or attributes appear. Two notions associated with a fuzzy choice function are studied: the normality coefficient used to determine the most rational expertise and the degree of dominance used in the multicriterial ranking of alternatives.

Different from other possible solutions, our model tries to correct market information asymmetry between a consumer and several producers. By interpreting the vectors that contain certain information as criteria for the customer in decision–making and by combining it with the certain information existing in the model, an optimal suggestion was illustrated by revealed preference theory and could be regarded as a recommendation for a more reliable purchase decision.

Keywords: Fuzzy choice function, Distance, Consumer decision-making, Online sales.

1 Introduction

The analysis of markets with asymmetric information represents one of the most challenging research topics in microeconomic theory [2]. Asymmetry of information appears in many situations of market interactions. We can exemplify this situation by various instances. Prospective buyers are less informed on the quality of the products than the sellers. A job applicant knows more about his working abilities than a potential employer. In such cases one party has more sufficient information than the other party. In many economic situations, such phenomena might lead to the obstruction of the transactions that are beneficial for both sides of the market.

At the same time, the Internet offers consumers and business producers a new medium to handle traditional business transaction. Through it, consumers have a broader scope to make choices. It also means more challenges not only for business to sell products via a virtual space [18], but also harder for consumers to overcome information overload and information asymmetry problems [9], [29]. Therefore, trying to give the optimal suggestion and recommendation to the consumer or to help consumers make the optimal decisions is becoming more and more important in $B-2-C$ model (Business-to-Consumer). In this paper, we propose a new method in a scenario of supporting consumers choosing a mobile device online. A new method (Fuzzy revealed preferences approach) is introduced, through which consumers' preferences would be evaluated according to multiple criteria decision–making.

Revealed preference theory was introduced by P. A. Samuelson [25] for the study of the rational behavior of a consumer. In this context, by the rationality of a consumer one understands the achievement of an optimal condition, expressed in terms of a preference relation. Later, revealed preference theory was developed in an axiomatic framework of choice functions, by the contribution of Arrow [3], Sen [26], Suzumura [28] and other economists.

In its classical form, revealed preference theory regards exact preferences and choices resulting from deterministic phenomena. There exist situations in economic and social life when the behavior of the individuals and groups takes place according to imprecise rules. Vague preferences and vague choices appear as the result of

partial information, attributes or criteria expressed indecisively or during some unfinished options.

Vague preferences are mathematically modeled by fuzzy preference relations and vague preferences by fuzzy choice functions [13]. Some authors consider that only the preferences are vague, while the result of the choices is exact (see e. g. [21] for a more detailed discussion).

Other papers study vague preferences and vague choices [4], [10], [15], [22], [30]. Specifically, papers [15], [16] attempt to develop a theory of revealed preference for fuzzy choice functions.

In this paper a decision–making model in two stages is studied:

–in the first stage, from a set of expertises one selects the one with the greatest degree of rationality;

–in the second stage, on the basis of the most rational expertise, one builds a multiple criterion decision–making problem, from which one reaches a ranking of alternatives according to criteria.

Revealed preference theory for fuzzy choice functions is the theoretical basis of the model.

The main concepts analyzed in this paper are the degree of dominance and the normality coefficient of a fuzzy choice function. With help of the normality coefficient the most rational expertise is determined, while the degree of dominance is used to rank the alternatives according to criteria. There is presented a procedure which results from the model developed in the paper.

2 Decisions with Vague Choices

In our natural world, decision–making is an interdisciplinary subject. It requires decision makers to be fully informed and demands concerns related to many factors in social activities to be computed with perfect accuracy. In business, decision–making is an important skill. Sometimes, it is unavoidable that decision makers would solve various problems by informal and highly subjective judgment to perceive and assimilate information. The conventional approaches, e.g. brainstorming, SWOT analysis, could not resolve the conflicts between competing criteria. Thus, good decision–making (rational decision–making) requires a mixture of skills, knowledge, and consensus and being decisive in acting upon the assessments made [27].

To compare assessment of alternative choices with respect to multiple conflicting evaluation aspects, one of the existing approaches is multi-criteria decision analysis (MCDA) [13], [32]. MCDA is increasingly used to take into account the relative weights of importance and the one with the highest aggregate preference index represents the best. In decision theory, commonly the optimal decision is a vague choice, which is determined by the probability and utility of potential outcomes [5]. To dynamically handle vagueness or uncertainty in the environment has been modeled with fuzzy systems [24].

In order to identity the best optimal decision and to support such cognitive process based on a certain amount of alternative options, many tools, methodologies and software are developed to help people make better decisions [24], [17]. The purpose of this work is to model a decision–making problem by concepts and results of revealed preference theory for fuzzy choice functions. After we develop the theoretical framework of fuzzy revealed preference theory, we apply the resulting procedure to solve a microeconomic model of decision–making in online shopping market.

3 Preliminaries on Fuzzy Preference Relations

On the real interval $[0,1]$ we consider the infinite operations $\bigvee_{i\in I} a_i = \sup_{i\in I} a_i$ and $\bigwedge_{i\in I} a_i = \inf_{i\in I} a_i$. In particular, for any $a, b \in [0,1]$, $a \vee b = \sup(a,b) = \max(a,b)$ and $a \wedge b = \inf(a,b) = \min(a,b)$. With respect to these operations the set $[0,1]$ gets a structure of complete distributive lattice [13].

Besides these operations, we shall consider the implication operator $\rightarrow$ defined by $a \rightarrow b = \min(1, 1-a+b)$ [13].

Let X be a non–empty set. The subsets of X are determined by bivalent properties; these subsets are called "crisp". If P is such a property, then the elements of X which verify P form a subset A of X, the others belong to the complement of A. With each subset A of X one associates its characteristic function $\chi_A : X \rightarrow \{0,1\}$ defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A. \end{cases}$$

A crisp subset A is determined by its characteristic function, therefore we can identify A with χ_A.

Real world phenomena are often described by vague attributes (e. g. "very tall", "cheap enough", "almost finished" etc.). The fuzzy set theory of Zadeh [31] describes precisely the way a vague attribute distinguishes elements of X between themselves . A fuzzy subset of X is a function $A : X \rightarrow [0,1]$; if $x \in X$ then the real number $A(x)$ called the membership degree of x in A represents the extent to which "x has the attribute A".

We shall denote by $\mathcal{P}(X)$ the family of crisp subsets of X and by $\mathcal{F}(X)$ the family of fuzzy subsets of X. By identifying a crisp subset of X with its characteristic function, we can consider $\mathcal{P}(X)$ a subset of $\mathcal{F}(X)$.

A fuzzy subset of $X^2 = X \times X$ is called a fuzzy relation on X. Accordingly, a fuzzy relation on X is a function $Q : X^2 \to [0,1]$; for any $x, y \in X$, $Q(x,y)$ is the degree to which the objects x, y are in relation Q.

We consider that X is a set of alternatives. In this case, Q can be interpreted as a preference relation on the alternatives of X; the real number $Q(x,y)$ represents the degree to which alternative x is preferred to alternative y.

Q is said to be:

reflexive if $Q(x,x) = 1$ for any $x \in X$;

total if $Q(x,y) > 0$ or $Q(y,x) > 0$ for all distinct $x, y \in X$.

4 The Rationality of Vague Choices by Revealed Preferences

In real life there exists the case when the preferences and the choices of an individual or a social group are vague. At the mathematical level the preferences are modeled by fuzzy relations and the choices are modeled by fuzzy choice functions. This situation has been emphasized in [4] and in [10].

This section is concerned with the rationality of choices by fuzzy revealed preference, extending some topics in the well-known classical revealed preference theory [3], [26]. We shall present the mathematical framework and some appropriate notions and results.

A fuzzy choice space is a pair $(X, \mathcal{B})$, where X is a universe of alternatives and $\mathcal{B}$ is a non-empty family of non-zero fuzzy subsets of X. A fuzzy choice function on $(X, \mathcal{B})$ is a function $C : \mathcal{B} \to \mathcal{F}(X)$ such that for any $S \in \mathcal{B}$, $C(S)$ is a non-zero fuzzy subset of X and $C(S) \subseteq S$.

The members of $\mathcal{B}$ are interpreted as available fuzzy sets. For any alternative x and $S \in \mathcal{B}$, the real number $S(x)$ can be viewed as the availability degree of x with respect to S. The degree $C(S)(x)$ to which x is chosen subject to S naturally belongs in the interval $[0, S(x)]$.

The members of $\mathcal{B}$ can also be interpreted as vague criteria or vague attributes. Then the real number $S(x)$ represents the degree to which the alternative x verifies the criterion S; $C(S)(x)$ expresses the degree to which the alternative x can be chosen with respect to S. These situations appear when the decision process is complex and assumes intermediary evaluations (e.g. negotiations on electronic marketplaces [19], [22]). In these cases the agent can take into account potential options which, because of the partial information are measured with a certain degree and therefore lead to fuzzy choice functions.

This modeling based on fuzzy choice functions tells the agent which are the best options for each criterion in part. Sometimes such ranking of the alternatives on criteria offers the agent richer information, that he will use in the final decision (see [30]).

Notation Suppose $X = \{x_1, \ldots, x_n\}$ and $\mathcal{B} = \{S_1, \ldots, S_m\}$. Then $\mathcal{B}$ will be represented by the *matrix of criteria*.

$$\begin{pmatrix} S_1(x_1) & S_2(x_1) & \ldots & S_m(x_1) \\ S_1(x_2) & S_2(x_2) & \ldots & S_m(x_2) \\ \ldots & & & \ldots \\ S_1(x_n) & S_2(x_n) & \ldots & S_m(x_n) \end{pmatrix}.$$

A fuzzy choice function $C : \mathcal{B} \to \mathcal{F}(X)$ will be represented by the following matrix:

$$C = \begin{pmatrix} C(S_1)(x_1) & C(S_2)(x_1) & \ldots & C(S_m)(x_1) \\ C(S_1)(x_2) & C(S_2)(x_2) & \ldots & C(S_m)(x_2) \\ \ldots & & & \ldots \\ C(S_1)(x_n) & C(S_2)(x_n) & \ldots & C(S_m)(x_n) \end{pmatrix}.$$

With any fuzzy choice function C one can associate a fuzzy preference relation R on the set X of alternatives:

$$R(x,y) = \bigvee_{S \in \mathcal{B}} (C(S)(x) \wedge S(y))$$

R is the fuzzy version of the crisp revealed preference relation defined by Samuelson [25]. $R(x,y)$ is the degree of truth of the statement "there is a criterion S such that the alternative x is chosen with respect to criterion S and the alternative y verifies S".

If Q is a fuzzy preference relation on X and $S \in \mathcal{B}$, then we define the fuzzy subset $G(S,Q)$ of X by putting $G(S,Q) = S(x) \wedge \bigwedge_{y \in X} [S(y) \to Q(x,y)]$ for all $x \in X$. The assignment $S \mapsto G(S,Q)$ defines a function $G(.,Q) : \mathcal{B} \to \mathcal{F}(X)$.

This function generalizes to fuzzy context the crisp notion of Q–greatest elements (see [28]) and allows us to introduce the rationality of fuzzy choice functions. A fuzzy choice function C is said to be rational if $C = G(.,Q)$ for some fuzzy preference relation Q on X. Thus C is recoverable from Q and it realizes the choice of Q–greatest alternatives. If $C = G(.,R)$ then we say that C is normal; in this case, C is recoverable from the fuzzy revealed preference R. The fuzzy choice function $\hat{C} = G(.,R)$ is called the image of C. Thus C is normal iff $C = \hat{C}$. By definition, $\hat{C}(S)(x) =$

$S(x) \wedge \bigwedge_{y \in X} [S(y) \to R(x,y)]$ for all $S \in \mathcal{B}$ and $x \in X$.

An act of fuzzy choice C emphasizes vague preferences on the alternatives, mathematically represented by the fuzzy revealed preference R. Then the normality means the rationality of C achieved by means of the fuzzy revealed preference R.

5 Degree of Dominance

In the previous section we have seen that by interpreting the available sets of alternatives as criteria, a fuzzy choice problem can be considered a multicriterial decision–making problem. Our goal is to determine the ranking of alternatives with respect to each criterion. The main instrument which allows us to establish such a ranking is the degree of dominance, notion defined in [16].

Let C be a fuzzy choice function on $(X, \mathcal{B})$. For all $S \in \mathcal{B}$ and $x \in X$ we define the *degree of dominance* of x in S by [16], [15]:

$$D_S^C(x) = S(x) \wedge \bigwedge_{y \in X} [C(S)(y) \to C(S)(x)].$$

If $D_S^C(x) = 1$ then we say that x is dominant in S. The degree of dominance shows to what extent an alternative has a dominant position among others as the result of the act of choice. If an alternative has the degree of dominance equal to 1, it will be said "dominant" with respect to S. If there are no dominant alternatives one can select the alternatives with the maximum degree of dominance.

This notion of degree of dominance refines the dominance in the sense of Banerjee [4]. Several notions of dominance are derived from preferences (see e.g. [21]). The above definition of degree of dominance is related to the act of choice and not to a preference relation.

If $X = \{x_1, \ldots, x_n\}$ and $\mathcal{B} = \{S_1, \ldots, S_m\}$ then we will comprise the degrees of dominance $D_{S_j}^C(x_i)$ in the matrix D^C of type $n \times m$.

$$D^C = (D_{S_j}^C(x_i))_{i=\overline{1,n}, j=\overline{1,m}}$$

Let $\Delta = \{C_1, \ldots, C_m\}$ be a finite set of fuzzy choice functions on $(X, \mathcal{B})$. We define the Δ–degree of dominance of $x \in X$ in S by:

$D_S^\Delta(x) = \frac{1}{m} \sum_{i=1}^m D_S^{C_i}(x)$.

If we interpret a fuzzy choice problem as a multiple criteria decision–making problem, then the degree of dominance allows to establish a hierarchy of alternatives for each criterion. This will enable the consumer to make an optimal decision.

The Δ–degree of dominance $D_S^\Delta(x)$ aggregates the degrees of dominance of the fuzzy choice functions of Δ. The multicriterial ranking obtained like this reflects the unified action of the fuzzy choice functions of Δ. For this model, the notion of Δ–degree of dominance is applied when the highest rationality degree is achieved by several expertises.

6 The Distance and the Normality Coefficient of Fuzzy Choice Functions

In this section we shall introduce two new notions: the distance and the normality coefficient of two fuzzy choice functions. With help of the distance the normality coefficient $Norm(C)$ of a fuzzy choice function C is defined. $Norm(C)$ evaluates the closeness between C and $\hat{C}$, by obtaining a measure of the normality of C. The normality coefficient can be used in ranking a given family of fuzzy choice functions.

Let Q_1 and Q_2 be two fuzzy preference relations on X. The distance between Q_1 and Q_2 is defined by [12]

$$d(Q_1, Q_2) = \bigvee_{x,y \in X} |Q_1(x,y) - Q_2(x,y)|.$$

The function $(Q_1, Q_2) \longmapsto d(Q_1, Q_2)$ is a metric on the set of fuzzy preference relations on X. The real number $d(Q_1, Q_2)$ appreciates the closeness of the fuzzy preference relations Q_1 and Q_2.

Analogously we will define the distance between two fuzzy choice functions.

We fix a fuzzy choice function C on a fuzzy choice space $(X, \mathcal{B})$.

Definition 6.1 *Let C_1, C_2 be two fuzzy choice functions on $(X, \mathcal{B})$. We define the* distance $d(C_1, C_2)$ *between C_1 and C_2 by*

$$d(C_1, C_2) = \bigvee_{S \in \mathcal{B}} \bigvee_{x \in X} |C_1(S)(x) - C_2(S)(x)|$$

Let $\delta \in [0,1]$. We say that C_1 and C_2 are δ–equal if $d(C_1, C_2) \leq 1 - \delta$.

We remark that the δ-equality of two fuzzy choice functions is very similar to the notion of δ–equality for fuzzy sets and fuzzy relations [7], [8], [14].

Proposition 6.2 *Let C_1, C_2, C_3 be three fuzzy choice functions on a fuzzy choice space $(X, \mathcal{B})$. Then*

(1) $d(C_1, C_2) = 0$ iff $C_1 = C_2$;

(2) $d(C_1, C_2) = d(C_2, C_1)$;

(3) $d(C_1, C_2) + d(C_2, C_3) \leq d(C_1, C_3)$.

Remark 6.3 *The previous proposition shows that the function $d(.,.)$ defined by the assignment $(C_1, C_2) \mapsto d(C_1, C_2)$ is a metric on the set of fuzzy choice functions on $(X, \mathcal{B})$. The real number $d(C_1, C_2)$ expresses how close the fuzzy choice functions C_1 and C_2 are.*

The notion of distance between two fuzzy choice functions allows us to introduce the *normality coefficient*.

Definition 6.4 *Let C be a fuzzy choice function on $(X, \mathcal{B})$. The normality coefficient $Norm(C)$ of C is defined by*

$$Norm(C) = d(C, \hat{C}) = \bigvee_{S \in \mathcal{B}} \bigvee_{x \in X} |C(S)(x) - \hat{C}(S)(x)|.$$

Remark 6.5 *By Proposition 6.2 (1), $Norm(C) = 0 \Leftrightarrow C = \hat{C} \Leftrightarrow C$ is normal.*

The normality of the fuzzy choice function C is characterized by the equality $C = \hat{C}$. By expressing the distance between C and $\hat{C}$, $Norm(C)$ is an indicator by which we can evaluate the degree of normality of the choice function C. The closer to zero $Norm(C)$ is , the better the normality of C is. If C_1, C_2 are two fuzzy choice functions on $(X, \mathcal{B})$ and $Norm(C_1) \geq Norm(C_2)$ then the normality of C_1 is worse than the normality of C_2. We have seen in Section 4 that the normality is the type of rationality defined by the revealed preference which derives from the act of choice. Then we can consider that normality coefficient is the most adequate indicator for expressing the degree of rationality of fuzzy choices.

7 A Two–Stage Decision Making Model

Several research papers [11], [33], have presented the decision–making problem in online shopping. It concludes that until now many CDSSs (consumer decision support systems) can only provide suboptimal support, since consumers need to deal with an increasing amount of information between competing alternatives. However, the simulation environment can provide more efficient comparisons and indicate further opportunities [33]. In order to prove our above theory, in this section we propose a microeconomic model of decision–making in e–commerce context. That is to simulate the fact that a consumer wants to make a purchase decision for a mobile phone online. In the current market there are many brands, hundreds of phone models therefore numerous criteria that the user needs to know and choose, before he makes a decision. In order to decide which phone model to buy, the consumer needs to classify the categories based on personal preferences: like for entertainment, for business, or just basic calling or image functions. Except the general categories, there are weight, size, connectivity, and price i.e. more than 30 criteria (functions) that need to be checked and decided. Especially, in order to evaluate the performance of the phone, the consumer needs to compare and decide many technical parameters that he may not understand much. To support the consumer to make a proper decision by avoiding adverse selection, our model tries to rank alternatives according to fuzzy choices. Furthermore, the model tries to correct the information asymmetry by separating the partial information on the traded goods (e. g. the quality of the products) from the certain information which is complete.

In the following we shall present a two–stage decision–making problem. As a result of the recommendations of online community experts, several fuzzy choice problems are obtained. In the first stage a ranking of this set of fuzzy choice problems is done by choosing those with the minimum coefficient of normality. In the second stage one obtains a ranking of alternatives according to criteria by using the degree of dominance.

A consumer wants to choose from n types of mobile phones $x_1, \ldots, x_n$. The n potential offers form the set $X = \{x_1, \ldots, x_n\}$ of alternatives. The choice of one type of mobile phones is done according to m vague criteria $P_1, \ldots P_m$.

The products are characterized by several attributes; for some attributes there is certain information (e. g. GPS (location awareness), camera (carl zeiss optics), display resolution, memory, price, etc.) and for other attributes the information is partial (e. g. the quality). Suppose the partial information is contained in the following table:

Mobile phone types	P_1	P_2	...	P_m
First type	a_{11}	a_{12}	...	a_{1m}
Second type	a_{21}	a_{22}	...	a_{2m}
...				
n–th type of phone	a_{n1}	a_{n2}	...	a_{nm}

The number $a_{ij} \in [0, 1]$, $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ describes the degree to which alternative x_i verifies criterion P_j.

In order to analyze the certain information on the products, the consumer hires l independent online community experts. Each online community expert will make an analysis and will forward the consumer the results. In his decision the consumer must choose the most adequate recommendations.

The result of the recommendation of the k–th online community expert is given in a matrix $Q_k = (q_{ij}^k)$ of dimension $n \times n$ where q_{ij}^k represents the degree to which the online community expert k considers offer x_i

at least as good as offer x_j with respect to the certain information.

Two natural conditions are imposed to the elements of Q_k:

(a) $q_{ij}^k = 1$;

(b) either $q_{ij}^k > 0$ or $q_{ji}^k > 0$ for all $i, j \in \{1, \ldots, n\}$.

This way the opinion of each online community expert is modeled as a fuzzy preference relation, reflexive and total.

Beside the results of the l online community experts, the consumer has partial information materialized in vectors $\underline{a}_j = (a_{1j}, \ldots, a_{nj})$, $j \in \{1, \ldots, m\}$.

The matrix Q_k obtained by online community expert k together with vectors $\underline{a}_j$ will lead to a fuzzy choice problem. This way, there are obtained l fuzzy choice problems according to the l online community experts.

From these fuzzy choice problems, the consumer must select those which reflect most faithfully the state of the products. This selection will be done according to a principle of rationality: the most suitable fuzzy choice problem will be that with the minimum degree of normality. There might exist several fuzzy choice problems with the minimum normality coefficient. In this case the consumer will select the problems with the minimum normality coefficient; they will be processed with the purpose of ranking the alternatives with respect to criteria.

Each selected fuzzy choice problem will be judged as a multicriterial decision–making problem: the available fuzzy sets will be interpreted as criteria. By using the Δ–degree of dominance one will rank the set of alternatives according to each criterion and globally. This will help the consumer choose the optimal alternatives.

According to the above discussion, the decision will be made in two stages:

Level a One determines the online community experts whose results offer the most rational awareness on the states of the products. Mathematically speaking, one obtains a family of fuzzy choice functions with a minimum coefficient of rationality.

Level b Once the most adequate fuzzy choice problems being chosen, one will select the optimal alternatives.

These considerations lead to the following mathematical model:

Denote by $X = \{x_1, \ldots, x_n\}$ the set of alternatives and by $\{S_1, \ldots, S_m\}$ the fuzzy subsets of X defined by

(1) $S_j(x_i) = a_{ij}$ for $i = \{1, \ldots, n\}$ and $j = \{1, \ldots, m\}$.

$S_1, \ldots, S_m$ will be the available fuzzy sets of the choice space.

The definition $S_j(x_i) = a_{ij}$ will be interpreted: the real number a_{ij} evaluates the degree to which offer x_i satisfies criterion P_j.

We denote $\mathcal{B} = \{S_1, \ldots, S_m\}$ and we consider the fuzzy choice space $(X, \mathcal{B})$. The matrix Q_k obtained by the online community expert k induces a fuzzy choice function $C_k = G(., Q_k)$ on $(X, \mathcal{B})$. This will have the following expression:

(2) $C_k(S_j)(x_i) = G(S_j, Q_k)(x_i)$

$$= S_j(x_i) \wedge \bigwedge_{u=1}^{n} [S_j(x_u) \to Q_k(x_i, x_u)]$$

$$= a_{ij} \wedge \bigwedge_{u=1}^{n} [a_{uj} \to q_{iu}^k]$$

for all $j = 1, \ldots, m$ and $i = 1, \ldots, n$. We obtain l fuzzy choice problems $\langle (X, \mathcal{B}), C_1 \rangle, \ldots, \langle (X, \mathcal{B}), C_l \rangle$.

From the normality coefficients $Norm(C_1), \ldots, Norm(C_l)$ one will determine the ones which have the smallest value.

Let Δ be the subset of $\{C_1, \ldots, C_l\}$ consisting of those C_k with $Norm(C_k)$ minimum.

For each $t \in \{1, \ldots, m\}$ one computes the Δ–degrees of dominance $D_{S_t}^{\Delta}(x_1), \ldots, D_{S_t}^{\Delta}(x_n)$ and one establishes the ranking of alternatives with respect to criterion P_t.

We denote $D(x_i) = \frac{1}{m} \sum_{j=1}^{m} D_{Sj}{}^{\Delta}(x_i)$ the aggregated degree of dominance of the alternative x_i. Then by ranking the set $\{D(x_1), \ldots, D(x_n)\}$ one obtains the global ranking of the set of alternatives.

The procedure

The input data are as follows:

n=the number of alternatives(=offers);

m=the number of criteria;

l=the number of online community experts;

$Q_1, \ldots, Q_l$=the matrices that describe the results of the recommendation, where $Q_k = (q_{ij}^k)$ for all $k = 1, \ldots, l$, $i, j = 1, \ldots, n$.

$\underline{a}_1, \ldots, \underline{a}_m$=vectors which describe the partial information

The procedure is developed in two stages:

Level a

Step 1 Compute the fuzzy sets $S_1, \ldots, S_m$:

$S_j(x_i) = a_{ij}$, $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$.

Step 2 Compute the fuzzy choice functions $C_1, \ldots, C_l$:

$$C_k(S_j)(x_i) = S_j(x_i) \wedge \bigwedge_{u=1}^{n} [a_{ij} \to q_{iu}^k]$$

for $k \in \{1, \ldots, l\}$, $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n\}$.

Step 3 Compute the fuzzy preference relations $R_1, \ldots, R_l$ associated with $C_1, \ldots, C_l$:

$$R_k(x_i, x_j) = \bigvee_{u=1}^{m} [C_k(S_u)(x_i) \wedge S_u(x_j)]$$

$$= \bigvee_{u=1}^{m} [C_k(S_u)(x_i) \wedge a_{ju}]$$

for $k \in \{1, \ldots, l\}$ and $i, j \in \{1, \ldots, n\}$.

Step 4 Compute the fuzzy choice function $\hat{C}_1, \ldots, \hat{C}_l$:

$$\hat{C}_k(S_j)(x_i) = S_j(x_i) \wedge \bigwedge_{u=1}^{n} [S_j(x_u) \to R_k(x_i, x_u)]$$

$$= a_{ij} \wedge \bigwedge_{u=1}^{n} [a_{uj} \to q_{iu}^k]$$

for $k \in \{1, \ldots, l\}$, $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n\}$.

Step 5 Compute the normality coefficients $Norm(C_1), \ldots, Norm(C_l)$:

$$Norm(C_k) = d(C_k, \hat{C}_k) = \bigvee_{j=1}^{m} \bigvee_{i=1}^{n} |C_k(S_j)(x_i) - \hat{C}_k(S_j)(x_i)|$$

for $k = 1, \ldots, l$.

Step 6 Determine the indices k for which

$Norm(C_k) = \min\{Norm(C_1), \ldots Norm(C_l)\}$.

Let K be the set of these indices and $\Delta = \{C_k | k \in K\}$.

Level b

Step 7 For each $k \in K$ compute the degrees of dominance:

$$D_{S_j}^k(x_i) = D_{S_j}^{C_k}(x_i) = S_j(x_i) \wedge \bigwedge_{u=1}^{n} [C_k(S_j)(x_u) \to C_k(S_j)(x_i)]$$

for $j = 1, \ldots, m$ and $i = 1, \ldots, n$.

Step 8 Compute the Δ–degrees of dominance:

$D_{S_j}^{\Delta}(x_i) = \frac{1}{card(K)} \sum_{k \in K} D_{S_j}^k(x_i)$

for $j = 1, \ldots m$ and $i = 1, \ldots n$.

Step 9 Ordering the set $\{D_{S_j}^{\Delta}(x_1), \ldots, D_{S_j}^{\Delta}(x_n)\}$ one obtains a ranking of alternatives with respect to criterion S_j, $j = 1, \ldots, m$.

Step 10 Compute the aggregated degrees of dominance:

$D(x_i) = \frac{1}{m} \sum_{j=1}^{m} D_{S_j}^{\Delta}(x_i)$, $i = 1, \ldots, n$.

Step 11 Ordering the set $\{D(x_1), \ldots, D(x_n)\}$ one obtains a global hierarchy of alternatives.

8 Concluding Remarks

The paper deals with a decision–making problem mathematically modeled by revealed preference theory [15]. In the proposed model, the decision–making process takes place in two–stages. At the first level, from a set of independent expertises the most rational one is chosen, in the sense of the rationality defined by fuzzy revealed preference theory. At the second level, based on the data of the best expertise, a multicriterial hierarchy of the set of alternatives is chosen.

References

[1] V Adriaenssens, B De Baets, P L M Goethals, N De Pauw Fuzzy rule–based models for decision support in ecosystem management. Science of the Total Environment 319 (2004) 1–12

[2] G A Akerlof The market for lemons: quality information and the market mechanism. Quarterly J Ec 84 (1970) 488–500

[3] K J Arrow Rational choice functions and orderings. Economica 26 (1959) 121–127

[4] A Banerjee Fuzzy choice functions, revealed preference and rationality. Fuzzy Sets and Systems 70 (1995) 31–43

[5] R E Bellman, L A Zadehg Decision–Making in a fuzzy environment. Management Science, Vol. 17, No. 4, Application Series (Dec., 1970), pp. B141–B164

[6] R Bělohlávek Fuzzy Relational Systems. Foundations and Principles. Kluwer (2002)

[7] K Y Cai δ–Equalities of fuzzy sets. Fuzzy Sets and Systems 76 (1995) 97–112

[8] K Y Cai Robustness of fuzzy reasoning and δ-equalities of fuzzy sets. IEEE Transactions on Fuzzy Systems 9 (2001) 738–750

[9] Y Cao, Y Li An intelligent fuzzy-based recommendation system for consumer electronic products. Expert Systems with Applications 33 (2007) 230–240

[10] M Dasgupta, R Deb Fuzzy choice functions. Social Choice and Welfare 8 (1991) 171–182

[11] A M Degeratu, A Rangaswamy, J N Wu Consumer choice behavior in online and traditional supermarket: The effects of brand name, price and other search attributes. International Journal of Research in Marketing 17 (2000) 55–78

[12] D Dubois, H Prade Fuzzy Sets and Systems. Theory and Applications. Academic Press, New York (1980)

[13] J Fodor, M Roubens Fuzzy Preference Modeling and Multicriteria Decision Support. Kluwer, Dordrecht (1994)

[14] I Georgescu A generalization of the Cai δ-equality of fuzzy sets, in Proceedings of the International Conference on Fuzzy Information Processing, Theory and Applications, Springer, Beijing, China, March (2003) 123–127

[15] I Georgescu Rational choice and revealed preference: a fuzzy approach. Turku Centre for Computer Science, Turku, Finland PhD Dissertation (2005)

[16] I Georgescu Degree of dominance and congruence axioms for fuzzy choice functions. Fuzzy Sets and Systems 155 (2005) 390–407

[17] C L Hwang, Y J Lai Fuzzy Multiple Objective Decision Making: Methods and Applications. Springer (1994)

[18] T Jelassi, S Leenen An E–commerce sales model for manufacturing companies: a conceptual framework and a European example. European Management Journal 21 (2003) 38–47

[19] S Klaue, K Kurbel, I Loutschko Automated negotiations on agent-based e-marketplaces: An overview. Proceedings of the 14th Bled Electronic Commerce Conference, Bled, Slovenia (2001) 508–519

[20] E P Klement, R Mesiar, E Pap Triangular Norms. Kluwer (2000)

[21] P Kulshreshtha, B Shekar Interrelationship among fuzzy preference–based choice function and significance of rationality conditions: a taxonomic and intuitive perspective. Fuzzy Sets and Systems 109 (2000) 429–445

[22] K Kurbel, I Loutschko Towards multi–agent electronic marketplaces: What is there and what is missing? The Knowledge Engineering Review 18 (2003) 33–46

[23] H Y Lin, P Y Hsu, G J Sheen A fuzzy–based decision–making procedure for data warehouse system selection. Expert Systems with Applications 32 (2007) 939–953

[24] R A Ribeiro Fuzzy multiple attribute decision making: A review and new preference elicitation techniques. Fuzzy Sets and Systems 78 (1996) 155–181

[25] P A Samuelson A note on the pure theory of consumer's behaviour. Economica 5 (1938) 61–71

[26] A K Sen Choice functions and revealed preference. Rev Ec Studies 38 (1971) 307–317

[27] H A Simon Rational decision making in business organizations. American Economic Review. American Economic Association. 69 (1979) 493–513

[28] K Suzumura Rational choice and revealed preference. Rev Ec Studies 43 (1976) 149–159

[29] G van Dijk, S Minocha, A Laing Consumers, channels and communication: Online and offline communication in service consumption. Interacting with Computers 19 (2007) 7–19

[30] B Van de Walle, S Heitsch, P Faratin Coping with one-to-many multi-criteria negotiations in electronic markets. Proceedings of the Second International Workshop on Negotiations in Electronic Markets: beyond Price Discovery, Munich, Germany (2001) 747–751

[31] L A Zadeh Similarity relations and fuzzy orderings. Information Sciences 3 (1971) 177–200

[32] M Zeleny Multiple Criteria Decision Making. McGraw-Hill, New York (1982)

[33] J Y Zhang, P Pu Performance evaluation of consumer decision support systems. International Journal of E–business Research 2 (2006) 28–45

Revealed Preference in a Fuzzy Framework

Irina Georgescu
Academy of Economic Studies
Department of Economic Cybernetics
P. O. Box 1-432, 014700
Bucharest, Romania
irina.georgescu@csie.ase.ro

Abstract

To a fuzzy choice function C we assign the indicators of revealed preference $WAFRP(C)$, $SAFRP(C)$ and the indicators of congruence $WFCA(C)$, $SFCA(C)$. These indicators measure the degree to which the fuzzy choice function C verifies the axioms of revealed preference $WAFRP$, $SAFRP$ and of congruence $WFCA$ and $SFCA$, respectively.

Keywords: Fuzzy choice function, Revealed preference axioms, Congruence axioms.

1 Introduction

The revealed preference theory was introduced by P. A. Samuelson [13] in order to express the rationality of a consumer in terms of some preference relation associated with a demand function. By [18], "the rationality of a consumer may be described by postulating that a consumer has a definite preference over all conceivable commodity bundles and that he chooses those commodity bundles that are optimal with respect to his preference subject to budgetary constraints". The foundation of this theory is built on The Weak Axiom of Consumer Behavior [13] and on The Strong Axiom of Consumer Behavior [10].

Even if this theory appeared in the context of consumers, its concepts and results can be developed in a larger context offered by the theory of choice functions [1], [12], [15], [16].

The classic theory of revealed preference is founded on two main concepts: the preference relations and the choice functions. The axioms and the main results of the theory express various connections between preferences and choices. A rational choice function is recoverable from a preference relation. In revealed preference theory there are studied various conditions of rationality, such as the axioms of revealed preference $WARP$, $SARP$ and the axioms of congruence WCA, SCA. $SARP$ and $WARP$ are axiomatized forms of the The Strong Axiom of Consumer Behavior and The Weak Axiom of Consumer Behavior. The Arrow–Sen theorem [1], [15] establishes the relationship between these conditions.

In social and economic life there exist complex processes subject to several parameters in which the preferences and the choices of the individuals and groups can manifest indecisively, without answers of type "yes or no". In this case we deal with vague preferences and vague choices. They arise from partial information, the multitude of criteria and in general from sources of human subjectivity.

The theory of fuzzy sets and fuzzy relations, initiated by Zadeh [20], represents the mathematical framework for modelling vague preferences and vague criteria [4].

In [5], [6], [7] we have tried to develop a theory of revealed preference for fuzzy choice functions. In the above mentioned papers there have been studied the axioms of revealed preference $WAFRP$, $SAFRP$ and the axioms of congruence $WFCA$, $SFCA$ in relationship with the rationality of fuzzy choice functions.

In this paper we introduce the indicators of revealed preference $WAFRP(C)$, $SAFRP(C)$ and the indicators of congruence $WFCA(C)$ and $SFCA(C)$. These indicators evaluate the degree to which the choice function C verifies the axioms $WAFRP$, $SAFRP$, $WFCA$ and $SFCA$, respectively.

The paper is organized as follows. Section 2 contains definitions and basic facts on t–norms, fuzzy sets and fuzzy preference relations. Section 3 presents an introduction to fuzzy choice functions and fuzzy preference relations associated with them. In Section 4 the degree of similarity of two fuzzy choice functions is introduced as a measure of how similar they are. The degree of similarity offers a way of expressing the rationality and the normality of a choice function. Section 5 contains the main contribution of

this paper. In this section the indicators of revealed preference $WAFRP(C)$, $SAFRP(C)$ and congruence $WFCA(C)$ and $SFCA(C)$ are studied. The main result of the paper is Theorem 2, that shows that in the presence of two hypotheses, $WAFRP(C)$ equals $SAFRP(C)$ and they can be expressed in terms of a normality degree and of the transitivity degree of R_C or $\bar{R}_C$.

2 Preliminaries

The notion of t–norm was introduced by [14]. The t–norm models some intersection operation of two fuzzy sets and the conjunction in fuzzy logic.

In this section we present some basic facts on continuous t-norms and residua. The background for these results can be found in [4], [9], [11].

A mapping $* : [0,1] \times [0,1] \to [0,1]$ is a t-norm iff it is commutative, associative, non-decreasing in each argument and $a * 1 = a$ for all $a \in [0,1]$.

With any continuous t-norm $*$ we associate its *residuum* :

$a \to b = \bigvee\{c \in [0,1] | a * c \leq b\}$.

The most well-known continuous t-norms are:

Lukasiewicz t-norm: $a *_L b = \max(0, a+b-1)$; $a \to_L b = \min(1, 1-a+b)$

minimum operator: $a *_G b = \min(a,b)$; $a \to_G b = \begin{cases} 1 & \text{if } a \leq b \\ b & \text{if } a > b \end{cases}$

Product t-norm: $a *_P b = ab$; $a \to_P b = \begin{cases} 1 & \text{if } a \leq b \\ b/a & \text{if } a > b \end{cases}$

The negation operation $\neg$ associated with $*$ is defined by

$\neg a = a \to 0 = \bigvee\{c \in [0,1] | a * c = 0\}$.

Let $*$ be a continuous t-norm.

The properties mentioned in the following two lemmas reflect the main connections between the t-norm $*$ and its residuum $\to$.

Lemma 1 *[3], [4] For any $a, b, c \in [0,1]$ the following properties hold:*

(1) $a * b \leq c \Leftrightarrow a \leq b \to c$;

(2) $a * (a \to b) = a \wedge b$;

(3) $a \leq b \Leftrightarrow a \to b = 1$;

(4) $a = 1 \to a$;

(5) $(a \to b) * (b \to c) \leq a \to c$.

Lemma 2 *[3], [4] For any $\{a_i\}_{i\in I}, \{b_i\}_{i\in I} \subseteq [0,1]$ and $a \in [0,1]$ the following properties hold:*

(1) $a \to (\bigwedge_{i\in I} a_i) = \bigwedge_{i\in I}(a \to a_i)$; *(2)* $(\bigvee_{i\in I} a_i) \to a = \bigwedge_{i\in I}(a_i \to a)$; *(3)* $\bigvee_{i\in I}(a_i \to a) \leq (\bigwedge_{i\in I} a_i) \to a$; *(4)* $(\bigvee_{i\in I} a_i) * a = \bigvee_{i\in I}(a_i * a)$; *(5)* $(\bigwedge_{i\in I} a_i) * (\bigwedge_{j\in I} b_j) \leq \bigwedge_{i,j\in I}(a_i * b_j)$.

The biresiduum associated with the continuous t-norm $*$ is defined by
$\rho(a,b) = a \leftrightarrow b = (a \to b) \wedge (b \to a)$.

Let X be a non-empty set. A *fuzzy subset* of X is a function $A : X \to [0,1]$. We denote by $\mathcal{P}(X)$ the family of crisp subsets of X and by $\mathcal{F}(X)$ the family of fuzzy subsets of X. Since a crisp subset of X is defined by its characteristic function, then we have $\mathcal{P}(X) \subseteq \mathcal{F}(X)$. For any $A, B \in \mathcal{F}(X)$, by $A \subseteq B$ we mean that $A(x) \leq B(x)$ for each $x \in X$. A fuzzy subset A of X is *non-zero* if $A(x) \neq 0$ for some $x \in X$; A is *normal* if $A(x) = 1$ for some $x \in X$.

If $x_1, \ldots, x_n \in X$ then $[x_1, \ldots, x_n]$ will denote the characteristic function of $\{x_1, \ldots, x_n\}$:

$$[x_1, \ldots, x_n](y) = \begin{cases} 1 & \text{if } y \in \{x_1, \ldots, x_n\} \\ 0 & \text{if } y \notin \{x_1, \ldots, x_n\} \end{cases}$$

For $A, B \in \mathcal{F}(X)$ let us denote

$I(A,B) = \bigwedge_{x\in X}(A(x) \to B(x))$ and $E(A,B) = \bigwedge_{x\in X}(A(x) \leftrightarrow B(x))$.

It is clear that $A \subseteq B$ iff $I(A,B) = 1$ and $A = B$ iff $E(A,B) = 1$. For any $x \in X$ we have:

$I(A,B) \leq A(x) \to B(x)$ and $E(A,B) \leq A(x) \leftrightarrow B(x)$.

$I(A,B)$ is called the *subsethood degree* of A and B and $E(A,B)$ the *degree of equality* (degree of similarity) of A and B. Intuitively $I(A,B)$ expresses the truth value of the statement "A is included in B" and $E(A,B)$ the truth value of the statement "A and B contain the same elements" (see [3], p. 82).

Let $*$ be a continuous t-norm.

A *fuzzy preference relation* Q on X is a function $Q : X^2 \to [0,1]$.

Let X be a non–empty set and Q a fuzzy relation on X. Q is said to be:

reflexive if $Q(x,x) = 1$ for all $x \in X$;

symmetric if $Q(x,y) = Q(y,x)$ for all $x, y \in X$;

$*$–transitive if $Q(x,y) * Q(y,z) \leq Q(x,z)$ for all $x, y, z \in X$;

total if $Q(x,y) > 0$ or $Q(y,x) > 0$ for all distinct $x, y \in X$;

strongly total if $Q(x,y) = 1$ or $Q(y,x) = 1$ for all distinct $x, y \in X$;

complete if $Q(x,y) > 0$ or $Q(y,x) > 0$ for all $x, y \in X$;

strongly complete if $Q(x,y) = 1$ or $Q(y,x) = 1$ for all $x, y \in X$.

We remark that Q is strongly complete iff it is reflexive and strongly total.

A fuzzy preference relation Q is said to be *regular* if it is $*$–transitive and strongly complete.

Let Q be a fuzzy relation on X. Denote by $T(Q)$ the intersection of all $*$–transitive fuzzy relations that contain Q:

$T(Q) = \bigcap\{Q' | Q \subseteq Q' \text{ and } Q' \text{ is transitive}\}$.

$T(Q)$ is called the *$*$–transitive closure* of Q. Remark that $T(Q) = Q$ iff Q is $*$–transitive.

3 Fuzzy Choice Functions

In this section we recall the definition of the fuzzy choice functions introduced in [6]. We present some fuzzy preference relations associated with a fuzzy choice function and we recall some results from [6].

A *fuzzy choice space* is a pair $(X, \mathcal{B})$ where X is a non-empty set of alternatives and $\mathcal{B}$ is a non-empty family of non-zero fuzzy subsets of X. A *fuzzy choice function* on $(X, \mathcal{B})$ is a function $C : \mathcal{B} \to \mathcal{F}(X)$ such that for any $S \in \mathcal{B}$, $C(S)$ is a non-zero fuzzy subset of X and $C(S) \subseteq S$.

Our definition of a fuzzy choice function generalizes Banerjee's [2]. In [2] the domain of a choice function is made of all non-empty finite subsets and the range is made of fuzzy subsets of X. In our approach, both the domain and the range of a choice function contain fuzzy subsets of X. If S is such an available set and $x \in X$, then $S(x)$ can be interpreted as the availability degree of x. The availability degree might be useful when the decision-maker possesses partial information on the alternative x or when a criterion limits the possibility of choosing x. Accordingly, the available sets correspond to some criteria or attributes of the alternatives in decision making processes.

The results of [6], [7] are proved provided the choice function C verifies the following hypotheses:

H_1 Every $S \in \mathcal{B}$ and $C(S)$ are normal fuzzy subsets of X;

H_2 $\mathcal{B}$ includes the fuzzy sets $[x_1, \ldots, x_n]$ for any $n \geq 1$ and $x_1, \ldots, x_n \in X$.

Since $C(S) \subseteq S$, in H_1 it suffices to assume that $C(S)$ is normal for each $S \in \mathcal{B}$. For the crisp choice functions the hypothesis H_1 is automatically fulfilled by the definition of such choice function: any S and $C(S)$ are non-empty. In the same case H_2 asserts that $\mathcal{B}$ includes all non-empty finite subsets of X, hypothesis assumed in [1], [15], [17].

Let $(X, \mathcal{B})$ be a fuzzy choice space and Q a fuzzy preference relation on X. For any $S \in \mathcal{B}$ let us define the fuzzy subsets $M(S,Q)$ and $G(S,Q)$ of X.

$$M(S,Q)(x) = S(x) * \bigwedge_{y \in X} [(S(y) * Q(y,x)) \to Q(x,y)]$$

$$G(S,Q)(x) = S(x) * \bigwedge_{y \in X} [S(y) \to Q(x,y)].$$

In the crisp case, $M(S,Q)$ represents the set of Q-maximal elements of S and $G(S,Q)$ represents the set of Q-greatest elements of S (see [16]):

$G(S,Q) = \{x \in S | (x,y) \in Q \text{ for all } y \in S\}$;

$M(S,Q) = \{x \in S | (y,x) \notin P_Q \text{ for all } y \in S\}$.

where P_Q denotes the strict preference relation associated with Q ($P_Q = \{(x,y) \in X^2 | (x,y) \in Q \text{ and } (y,x) \notin Q\}$).

In general the functions $M(.,Q) : \mathcal{B} \to \mathcal{F}(X)$ and $G(.,Q) : \mathcal{B} \to \mathcal{F}(X)$ are not fuzzy choice functions.

The functions $M(.,Q)$ and $G(.,Q)$ allow for introducing the notion of rationality of fuzzy choice functions. A fuzzy choice function C on $(X, \mathcal{B})$ is called M–rational (resp. G–rational) if $C = M(.,Q)$ (resp. $C = G(.,Q)$) for some preference relation Q on X. In case of the minimum operator $\wedge$, if Q is reflexive and strongly total, then $M(.,Q) = G(.,Q)$ [5], therefore M–rationality and G–rationality coincide.

Consider now a fuzzy choice function C on $(X, \mathcal{B})$. To C one assigns the fuzzy revealed preferences R_C, $\bar{R}_C$ and $\tilde{P}_C$ on X defined by

$$R_C(x,y) = \bigvee_{S \in \mathcal{B}} (C(S)(x) * S(y));$$

$$\bar{R}_C(x,y) = C([x,y])(x);$$

$$\tilde{P}_C(x,y) = \bigvee_{S \in \mathcal{B}} (C(S)(x) * S(y) * \neg C(S)(y))$$

for any $x, y \in X$. R_C, $\bar{R}_C$ and $\tilde{P}_C$ are fuzzy versions of some preference relations studied in classical revealed preference theory ([1], [12], [13], [15]).

Let x, y be two alternatives. By interpreting the t-

norm as a conjunction, the real number $R_C(x,y)$ is the degree of truth of the statement "there is a criterion S such that the alternative x is chosen with respect to criterion S and alternative y verifies S". $\bar{R}_C(x,y)$ represents the degree of truth of the statement "from the set $\{x,y\}$ is chosen at least the alternative x", and $\tilde{P}_C(x,y)$ is the degree of truth of the statement "there is a criterion S with respect to which x is chosen and y is rejected".

By [6], in the case of the minimum operator, $\bar{R}_C \subseteq R_C$ and R_C, $\bar{R}_C$ are strongly complete.

A fuzzy choice function C is said to be G–normal (resp. M–normal) if $C = G(.,R_C)$ (resp. $C = M(.,R_C)$) [6]. G–normality (resp. M–normality) is a special case of G–rationality (resp. M–rationality). For simplicity we shall write $\hat{C} = G(.,R_C)$. In case of the minimum operator, $C(S) \subseteq \hat{C}(S)$ for any $S \in \mathcal{B}$ [6].

Now we shall consider the following axioms of fuzzy revealed preference:

$WAFRP$ (Weak Axiom of Fuzzy Revealed Preference)

$\tilde{P}_C(x,y) \leq \neg R_C(y,x)$ for all $x,y \in X$;

$SAFRP$ (Strong Axiom of Fuzzy Revealed Preference)

$P^*{}_C(x,y) \leq \neg R_C(y,x)$ for all $x,y \in X$.

Now we shall state two axioms of congruence for fuzzy choice functions.

Let $C : \mathcal{B} \to \mathcal{F}(X)$ be a fuzzy choice function.

$WFCA$ (Weak Fuzzy Congruence Axiom)

For any $S \in \mathcal{B}$ and $x,y \in X$ the following inequality holds

$R_C(x,y) * C(S)(y) * S(x) \leq C(S)(x)$.

$SFCA$ (Strong Fuzzy Congruence Axiom)

For any $S \in \mathcal{B}$ and $x,y \in X$ the following inequality holds

$W_C(x,y) * C(S)(y) * S(x) \leq C(S)(x)$.

W_C and $P^*{}_C$ represent the transitive closures of fuzzy preference relations R_C and $\tilde{P}_C$.

Axiom $WAFRP$ says that the degree to which there exists a criterion S with respect to which alternative x is chosen, alternative y verifies S but is rejected is less or equal than the degree to which for any other criterion S' fulfilled by x, y cannot be rejected.

Axiom $SFCA$ expresses the fact that the degree to which alternative x is chosen with respect to criterion S, y verifies S and x is revealed preferred to y is less or equal than the degree to which y is chosen with respect to S.

Similar interpretations can be given to the axioms $SAFRP$ and $SFCA$.

Remark 1 *Axioms $WAFRP$, $SAFRP$, $WFCA$, $SFCA$ are fuzzy versions of axioms $WARP$, $SARP$, WCA, SCA in crisp choice function theory.*

4 Similarity of Fuzzy Choice Functions

Approximate reasoning deals with variables that are not identical, but have close behaviour becoming identifiable. Therefore we need some concepts to express situations when fuzzy sets and/or fuzzy relations are identifiable. The notion of similarity is very satisfiable for describing such situations.

The notion of similarity relation was introduced by Zadeh [19] as a generalization of the concept of (crisp) equivalence relation.

We fix a continuous t–norm $*$. Let X be a non–empty set.

A fuzzy relation Q on X is said to be a *similarity relation* if it is reflexive, symmetric and $*$–transitive. If $x,y \in X$ then $Q(x,y)$ will be called the similarity degree of x and y.

Lemma 3 *The function $E(.,.) : (\mathcal{F}(X))^2 \to [0,1]$ defined by the assignment $(A,B) \mapsto E(A,B)$ is a similarity relation on $\mathcal{F}(X)$.*

Lemma 4 *If $A,B \in \mathcal{F}(X)$ and $x \in X$ then $E(A,B) * A(x) \leq B(x)$.*

If Q_1, Q_2 are two fuzzy relations on X then the degree of similarity of Q_1 and Q_2 has the form: $E(Q_1,Q_2) = \bigwedge_{x,y \in X} (Q_1(x,y) \leftrightarrow Q_2(x,y))$.

According to Lemma 4, for any $x,y \in X$ we have $E(Q_1,Q_2) * Q_1(x,y) \leq Q_2(x,y)$.

If one interprets the fuzzy relations Q_1 and Q_2 as preferences of two agents then the real number $E(Q_1,Q_2)$ expresses how "similar" these preferences are.

The next definition introduces the degree of similarity of two fuzzy choice functions.

Definition 1 *Let C_1, C_2 be two fuzzy choice functions*

on $(X, \mathcal{B})$. *The degree of similarity* $E(C_1, C_2)$ of C_1 and C_2 is a real defined by

$$E(C_1, C_2) = \bigwedge_{x \in X} \bigwedge_{S \in \mathcal{B}} (C_1(S)(x) \leftrightarrow C_2(S)(x)).$$

Lemma 5 *[8] The assignment* $(C_1, C_2) \mapsto E(C_1, C_2)$ *defines a similarity relation on the set of fuzzy choice functions on* X.

The following three lemmas are true for the case when $*$ is the minimum operator $\wedge$.

Lemma 6 *Let* C, C' *be two fuzzy choice functions. Then for any* $S \in \mathcal{B}$ *and* $x \in X$ *we have*

(i) $E(C, C') \wedge C(S)(x) \leq C'(S)(x)$;

(ii) $E(C, C') \wedge \neg C(S)(x) \leq \neg C'(S)(x)$.

Lemma 7 *Let* C *and* C' *be two fuzzy choice functions and* $x, y \in X$. *Then*

(i) $E(C, C') \wedge R_C(x, y) \leq R_{C'}(x, y)$;

(ii) $E(C, C') \wedge \neg R_C(x, y) \leq \neg R_{C'}(x, y)$.

Lemma 8 *If* C *and* C' *are two fuzzy choice functions and* $x, y \in X$ *then* $E(C, C') \wedge \tilde{P}_C(x, y) \leq \tilde{P}_{C'}(x, y)$.

The degree of similarity $E(C_1, C_2)$ allows us to define the indicators of revealed preference and congruence from Section 5. The three lemmas from above are helpful in proving the main theorem of the paper, Theorem 2.

5 Indicators of Fuzzy Preference Relations

The properties of the fuzzy relations studied in this section (reflexivity, transitivity, etc.) throw a better light on the way (fuzzy) preference relations are connected with various alternatives.

The indicators of fuzzy preference relations introduced in this section have the following significance: instead of checking whether a fuzzy relation Q has the property P, we evaluate the degree to which Q verifies P.

Therefore, instead of saying that Q is reflexive or not we shall consider the degree of truth of the statement "Q is reflexive.". This will be a real number in the interval $[0, 1]$ and it will express "how reflexive" the fuzzy relation Q is. The indicators introduced by Definition 2 represent degrees of truth which correspond to properties of fuzzy relations such as reflexivity, symmetry, etc.

Definition 2 *Let* Q *be a fuzzy relation on* X. *We define the following indicators:*

(a) the degree of reflexivity of Q:

$$Ref(Q) = \bigwedge_{x \in X} Q(x, x);$$

(b) the degree of symmetry of Q:

$$Sym(Q) = \bigwedge_{x,y \in X} (Q(x, y) \rightarrow Q(y, x)) = \bigwedge_{x,y \in X} (Q(x, y) \leftrightarrow Q(y, x));$$

(c) the degree of $*$*–transitivity of* Q:

$$Trans(Q) = \bigwedge_{x,y,z \in X} (Q(x, y) * Q(y, z) \rightarrow Q(x, z));$$

(d) the degree of strong totality of Q:

$$ST(Q) = \bigwedge_{x \neq y} (Q(x, y) \vee Q(y, x));$$

(e) the degree of strong completeness of Q:

$$SC(Q) = \bigwedge_{x,y \in X} (Q(x, y) \vee Q(y, x)).$$

Lemma 9 *For any fuzzy relations* Q *on* X *the following assertions hold:*

(a) $Ref(Q) = 1$ *iff* Q *is reflexive;*

(b) $Sym(Q) = 1$ *iff* Q *is symmetric;*

(c) $Trans(Q) = 1$ *iff* Q *is* $*$*–transitive;*

(d) $ST(Q) = 1$ *iff* Q *is strongly total;*

(e) $SC(Q) = 1$ *iff* Q *is strongly complete.*

Proposition 1 *Let* Q *be a fuzzy relation on* X *and* $x, y, z \in X$. *Therefore*

(a) $Ref(Q) \leq Q(x, x)$;

(b) $Sym(Q) * Q(x, y) \leq Q(y, x)$;

(c) $Trans(Q) * Q(x, y) * Q(y, x) \leq Q(x, z)$;

(d) If $x \neq y$ *then* $ST(Q) \leq Q(x, y)$ *or* $ST(Q) \leq Q(y, x)$;

(e) $SC(Q) \leq Q(x, y)$ *or* $SC(Q) \leq Q(y, x)$.

The following two propositions show how the five indicators defined above are preserved by the fuzzy equality $E(.,.)$.

Proposition 2 *Let* Q_1, Q_2 *be two fuzzy relations on* X. *Then*

(1) $Ref(Q_1) * E(Q_1, Q_2) \leq Ref(Q_2)$;

(2) $ST(Q_1) * E(Q_1, Q_2) \leq ST(Q_2)$;

(3) $SC(Q_1) * E(Q_1, Q_2) \leq SC(Q_2)$;

(4) $Sym(Q_1) * E(Q_1, Q_2) \leq Sym(Q_2)$.

Proposition 3 *Suppose* $*$ *is the minimum operator* $\wedge$. *Then for any two fuzzy relations* Q_1, Q_2 *on* X *we have:*

$Trans(Q_1) \wedge E(Q_1, Q_2) \leq Trans(Q_2)$.

Proposition 4 *Let* Q *be a fuzzy relation on* X. *For any* $x, y \in X$ *we have*

$Trans(Q) * T(Q)(x, y) \leq Q(x, y)$.

Remark 2 *By Proposition 4, for all* $x, y \in X$ *we have* $Trans(Q) \leq T(Q)(x, y) \rightarrow Q(x, y) = T(Q)(x, y) \leftrightarrow Q(x, y)$ *hence*

$Trans(Q) \leq \bigwedge_{x,y \in X} (T(Q)(x, y) \leftrightarrow Q(x, y)) = E(Q, T(Q))$.

6 Revealed Preference and Congruence Indicators

The goal of this section is to define and study some indicators of the axioms of revealed preference $WAFRP$, $SAFRP$ and the axioms of congruence $WFCA$, $SFCA$. These indicators express the degree to which the axioms $WAFRP$, $SAFRP$, $WFCA$ and $SFCA$ are verified by a fuzzy choice function.

Definition 3 *For a fuzzy choice function* C *on* $(X, \mathcal{B})$ *we define the following indicators of the axioms* $WAFRP$, $SAFRP$, $WFCA$ *and* $SFCA$:

(i) $WAFRP(C) = \bigwedge_{x,y \in X} [\tilde{P}_C(x, y) \rightarrow \neg R_C(y, x)]$;

(ii) $SAFRP(C) = \bigwedge_{x,y \in X} [P^*_C(x, y) \rightarrow \neg R_C(y, x)]$;

(iii) $WFCA(C) = \bigwedge_{x,y \in X} \bigwedge_{S \in \mathcal{B}} [S(x) * C(S)(y) * R_C(x, y) \rightarrow C(S)(x)]$;

(iv) $SFCA(C) = \bigwedge_{x,y \in X} \bigwedge_{S \in \mathcal{B}} [S(x) * C(S)(y) * W_C(x, y) \rightarrow C(S)(x)]$.

Remark 3 *For a choice function* C *the following equivalences hold:*

(i) $WAFRP(C) = 1$ *iff* C *verifies* $WAFRP$;

(ii) $SAFRP(C) = 1$ *iff* C *verifies* $SAFRP$;

(iii) $WFCA(C) = 1$ *iff* C *verifies* $WFCA$;

(iv) $SFCA(C) = 1$ *iff* C *verifies* $SFCA$.

The indicator $WAFRP(C)$ expresses the degree to which the choice function C verifies $WAFRP$. Similar interpretations can be given to the other three indicators.

In this section we suppose that hypotheses H_1 and H_2 are verified.

Proposition 5 *If* C *is a fuzzy choice function on* $(X, \mathcal{B})$ *then* $WFCA(C) \leq Trans(R_C)$.

Proposition 6 *If* C *is a fuzzy choice function then* $E(C, \hat{C}) \wedge Trans(R_C) \leq WFCA(C)$.

Theorem 1 *For any fuzzy choice function* C *on* $(X, \mathcal{B})$ *we have* $E(C, \hat{C}) \leq E(R_C, \bar{R}_C)$.

The following theorem is the main result of this paper.

Theorem 2 *If* C *is a fuzzy choice function on* $(X, \mathcal{B})$ *then*

$WFCA(C) = SFCA(C) = E(C, \hat{C}) \wedge Trans(R_C) = E(C, \hat{C}) \wedge Trans(\bar{R}_C)$.

Corollary 1 *[6] For a fuzzy choice function* C *on* $(X, \mathcal{B})$ *the following assertions are equivalent:*

(1) C *verifies* $WFCA$;

(2) C *verifies* $SFCA$;

(3) C *is G–normal and* R_C *is regular;*

(4) C *is G–normal and* $\bar{R}_C$ *is regular.*

Remark 4 *The degree of similarity* $E(C, \hat{C})$ *of* C *and* $\hat{C}$ *can be regarded as a measure of the G–normality of* C. *Then Theorem 2 shows that the indicators* $WAFRP(C)$ *and* $SAFRP(C)$ *are equal and that they can be expressed in terms of the G–normality of* C *and of the degree of transitivity of* R_C *or* $\bar{R}_C$.

7 Concluding Remarks

In [6] we have proved a fuzzy version of the Arrow–Sen theorem [15] which establishes the equivalence of revealed preference axioms $WARP$, $SARP$, the congruence axioms WCA, SCA and four other conditions of rationality of choice functions.

For fuzzy choice functions, some of these equivalences of the Arrow–Sen theorem hold true for an arbitrary continuous t–norm, others for the minimum operator and others for Lukasiewicz t–norm [6].

The core of this paper is a stronger version of the part in Arrow–Sen theorem [6] that holds for the minimum operator. The equivalence of $WFCA$, $SFCA$ and other two conditions of rationality is expressed in terms of numerical indicators. These indicators enable us to compare two fuzzy choice functions with respect to each axiom. For example, for two fuzzy choice functions C_1 and C_2, if $WFCA(C_1) \leq WFCA(C_2)$ then we can say that C_2 verifies $WFCA$ to a greater extent than C_1. On this basis, any family of fuzzy choice functions can be ranked with respect to each axiom.

References

[1] K. J. Arrow. Rational Choice Functions and Orderings. Economica 26 (1959) 121–127.

[2] A. Banerjee. Fuzzy Choice Functions, Revealed Preference and Rationality. Fuzzy Sets and Systems 70 (1995) 31–43.

[3] R. Bělohlávek. Fuzzy Relational Systems. Foundations and Principles. Kluwer 2002.

[4] J. Fodor and M. Roubens. Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht 1994.

[5] I. Georgescu. Rational Choice and Revealed Preference: A Fuzzy Approach. Turku Centre for Computer Science PhD Dissertation 60 2005.

[6] I. Georgescu. On the Axioms of Revealed Preference in Fuzzy Consumer Theory. Journal of Systems Science and Systems Engineering 13 (2004) 279–296.

[7] I. Georgescu. Consistency Conditions in Fuzzy Consumers Theory. Fundamenta Informaticae 61 (2004) 223–245.

[8] I. Georgescu. Similarity of Fuzzy Choice Functions. Fuzzy Sets and Systems, in press.

[9] P. Hájek. Metamathematics of Fuzzy Logic. Kluwer 1998.

[10] H. S. Houthakker. Revealed Preference and Utility Functions. Economica 17 (1950) 159–174.

[11] E. P. Klement, R. Mesiar and E. Pap. Triangular Norms. Kluwer 2000.

[12] M. Richter. Revealed Preference Theory. Econometrica 34 (1966) 635–645.

[13] P. A. Samuelson. A Note on the Pure Theory of Consumers' Behaviour. Economica 5 (1938) 61–71.

[14] B. Schweizer and A. Sklar. Statistical Metric Spaces. Pacific Journal of Mathematics 10 (1960) 312–334.

[15] A. K. Sen. Choice Functions and Revealed Preference. Review of Economic Studies 38 (1971) 307–317.

[16] K. Suzumura. Rational Choice and Revealed Preference. Review of Economic Studies 43 (1976) 149–159.

[17] H. Uzawa. A Note on Preference and Axioms of Choice. Annals of the Institute of Statistical Mathematics 8 (1956) 35–40.

[18] H. Uzawa. Preference and Rational Choice in the Theory of Consumption. In Arrow K. J., Karlin S., Suppes P. (editors) *Mathematical Methods in the Social Sciences*. Stanford, Stanford University Press, 1959.

[19] L. A. Zadeh. Similarity Relations and Fuzzy Orderings. Information Sciences 3 (1971) 177–200.

[20] L. A. Zadeh. Fuzzy Sets. Information and Control 8 (1965) 338–353.

Session 14

Aggregation Operators in Decision Making

Application of the Fuzzy Weighted Average of Fuzzy Numbers in Decision Making Models

Ondřej Pavlačka
Department of Mathematical Analysis and Applied Mathematics, Faculty of Science, Palacký University Olomouc, tř. Svobody 26, 771 46 Olomouc, Czech Republic
onpav@centrum.cz

Jana Talašová
Department of Mathematical Analysis and Applied Mathematics, Faculty of Science, Palacký University Olomouc, tř. Svobody 26, 771 46 Olomouc, Czech Republic
talasova@inf.upol.cz

Abstract

The paper deals with the operation of fuzzy weighted average of fuzzy numbers. The operation can be applied to the aggregation of partial fuzzy evaluations in fuzzy models of multiple criteria decision making and to the computation of expected fuzzy evaluations of alternatives in discrete fuzzy-stochastic models of decision making under risk. Normalized fuzzy weights figuring in the operation have to form a special structure of fuzzy numbers; its properties will be studied and the practical procedures for setting them will be proposed. A fuzzy weighted average of fuzzy numbers with normalized fuzzy weights will be defined, an effective algorithm of its calculation will be described, and its uncertainty will be studied.

Keywords: Multiple criteria decision making, decision making under risk, normalized fuzzy weights, fuzzy weighted average, fuzzy weights of criteria, fuzzy probabilities.

1 Introduction

In multiple criteria decision making models, the weighted average is often used for aggregating partial evaluations of alternatives. Weights of criteria are understood differently in different models. The most general definition says that the weights of criteria are non-negative real numbers whose ordering expresses the importance of criteria. According to this definition, the weights mean the measurements of criteria importance that are defined on an ordinal scale. But in most of multiple criteria decision making models, the weights of criteria represent some kind of cardinal information about the importance of criteria. For example, in the model of multiple criteria evaluation that was introduced in [8] the weights of criteria represent shares of the corresponding partial objectives of evaluation in the overall one. This concept of criteria weights will be considered in this paper.

The weighted average operation is applied also in the models of decision making under risk where evaluations of alternatives depend on a finite number of states of the world whose probabilities are known. The best alternative is chosen according to expected evaluations of alternatives. They are computed by means of the weighted average operation where the probabilities of states of the world play the role of weights.

The weights of criteria as well as the probabilities of states of the world are usually set subjectively, i.e. they are more or less uncertain. In this paper, it will be shown how this kind of uncertain information can be expressed by means of tools of fuzzy sets theory. A special structure of fuzzy numbers, normalized fuzzy weights, will be introduced for expressing the uncertain normalized weights and the operation of the fuzzy weighted average of fuzzy numbers with normalized fuzzy weights will be defined. In the end, application of the fuzzy weighted average operation in decision making models will be shown.

2 Applied Notions of Fuzzy Sets Theory

In this paper, the following notation will be used. The symbol N_m denotes the index set $\{1, 2, \ldots, m\}$. A fuzzy set A on a universal set X is characterized by its membership function $A : X \to [0, 1]$. $Ker\ A$ denotes the kernel of A, i.e. $Ker\ A = \{x \in X \mid A(x) = 1\}$. For $\alpha \in (0, 1]$, A_α denotes the α-cut of A, i.e. $A_\alpha = \{x \in X \mid A(x) \geq \alpha\}$. $Supp\ A$ denotes the support of A, i.e. $Supp\ A = \{x \in X \mid A(x) > 0\}$.

A fuzzy number is a fuzzy set C on the set of all real numbers $\Re$ that fulfils the following conditions: a) $Ker\ C \neq \emptyset$, b) C_α are closed intervals for all $\alpha \in (0, 1]$, c) $Supp\ C$ is a bounded set. A fuzzy number C is

said to be defined on $[a,b]$, if $Supp\ C \subseteq [a,b]$. Let us denote $[c^1,c^4] = Cl(Supp\ C)$, $[c^2,c^3] = Ker\ C$; $Cl(Supp\ C)$ means the closure of $Supp\ C$. The real numbers $c^1 \le c^2 \le c^3 \le c^4$ are called significant values of C.

A fuzzy number C can be also characterized (see [3]) by the pair of functions $(\underline{c},\overline{c})$, both from $[0,1]$ to $\Re$, that satisfy the following property

$$\underline{c}(\alpha) \le \underline{c}(\beta) \le \overline{c}(\beta) \le \overline{c}(\alpha), \quad 0 \le \alpha \le \beta \le 1. \tag{1}$$

Then, for the membership function C of the fuzzy number $C = (\underline{c},\overline{c})$ the following holds

$$C(x) = \max\{\alpha \mid x \in [\underline{c}(\alpha), \overline{c}(\alpha)]\} \tag{2}$$

for $x \in [\underline{c}(0), \overline{c}(0)]$, and 0 elsewhere; i.e. $C_\alpha = [\underline{c}(\alpha), \overline{c}(\alpha)]$ for all $\alpha \in (0,1]$, $Cl(Supp C) = [\underline{c}(0), \overline{c}(0)]$.

If $\underline{c}(\alpha) = c = \overline{c}(\alpha)$, for all $\alpha \in [0,1]$, then C represents a real number; $C = c$, $c \in \Re$. A fuzzy number C is called symmetric with the central point c if the following holds

$$c = \frac{\underline{c}(\alpha) + \overline{c}(\alpha)}{2}, \quad \text{for } \alpha \in [0,1]. \tag{3}$$

A fuzzy number $C = (\underline{c},\overline{c})$ is called a linear fuzzy number if both $\underline{c}$ and $\overline{c}$ are linear functions. Any linear fuzzy number C is fully characterized by its significant values $c^1 \le c^2 \le c^3 \le c^4$; the functions $\underline{c}$ and $\overline{c}$ are given for all $\alpha \in [0,1]$ as follows

$$\begin{aligned} \underline{c}(\alpha) &= c^1 + \alpha(c^2 - c^1), \\ \overline{c}(\alpha) &= c^4 - \alpha(c^4 - c^3). \end{aligned} \tag{4}$$

The notation $C = \langle c^1, c^2, c^3, c^4\rangle$ is used in this paper for such a fuzzy number C. For $c^2 \ne c^3$, the linear fuzzy number C is usually referred to as a trapezoidal fuzzy number, for $c^2 = c^3$ as a triangular fuzzy number.

For any fuzzy number $C = (\underline{c},\overline{c})$, the non-increasing, non-negative function

$$\sigma_C(\alpha) = \overline{c}(\alpha) - \underline{c}(\alpha), \quad \text{for } \alpha \in [0,1], \tag{5}$$

is called a width function of C. It describes in details, for any membership degree α, the uncertainty of C. The real numbers $\sigma_C(1)$, $\sigma_C(\alpha)$ and $\sigma_C(0)$ are called the width of the kernel, the width of the α-cut and the width of the support, respectively. Obviously, $\sigma_C(\alpha) = 0$ for all $\alpha \in [0,1]$ if and only if C is a real number. It holds that

$$\int_0^1 \sigma_C(\alpha)\ d\alpha = \int_{\Re} C(x)\ dx, \tag{6}$$

where the right-hand side of the equation represents the usual aggregated measure of uncertainty of the fuzzy number C.

Calculations with fuzzy numbers are based on the extension principle (see [3]). Let $f : [a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_n,b_n] \to \Re$ be a continuous function and let $f([a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_n,b_n]) = [c,d]$. Let X_i be fuzzy numbers defined on $[a_i,b_i]$, for $i \in N_n$. Then $f(X_1,X_2,\ldots,X_n)$ is a fuzzy number Y defined on $[c,d]$ whose membership function is given as follows

$$Y(y) = \max\big\{\min\{X_1(x_1), X_2(x_2), \ldots, X_n(x_n)\} \mid y = f(x_1,x_2,\ldots,x_n)\big\} \tag{7}$$

for $y \in [c,d]$, and 0 elsewhere. Employing the $(\underline{y},\overline{y})$ representation of the fuzzy number Y, the following holds for any $\alpha \in [0,1]$

$$\begin{aligned} \underline{y}(\alpha) &= \min\{f(x_1,x_2,\ldots,x_n) \mid x_i \in X_{i\alpha},\ i \in N_n\}, \\ \overline{y}(\alpha) &= \max\{f(x_1,x_2,\ldots,x_n) \mid x_i \in X_{i\alpha},\ i \in N_n\}. \end{aligned} \tag{8}$$

In some cases (see [5] for more details), it is necessary to consider a relation R among the variables $x_1,x_2,\ldots,x_n$. Let R be a closed and convex subset of $[a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_n,b_n]$ and let $f^R : R \to [c,d]$ be the restriction of f on R. Let there exist at least one n-tuple $(x_1,x_2,\ldots,x_n) \in R$ such that $X_i(x_i) = 1$ for all $i \in N_n$. Then $f^R(X_1,X_2,\ldots,X_n)$ is a fuzzy number Y^R defined on $[c,d]$ whose membership function is given in the following way

$$Y^R(y) = \max\big\{\min\{X_1(x_1), X_2(x_2), \ldots, X_n(x_n)\} \mid y = f(x_1,x_2,\ldots,x_n),\ (x_1,x_2,\ldots,x_n) \in R\big\} \tag{9}$$

for $y \in [c,d]$, and 0 elsewhere. Employing the $(\underline{y}^R,\overline{y}^R)$ representation of the fuzzy number Y^R, the following holds for any $\alpha \in [0,1]$

$$\begin{aligned} \underline{y}^R(\alpha) &= \min\big\{f(x_1,x_2,\ldots,x_n) \mid (x_1,x_2,\ldots,x_n) \in (X_{1\alpha} \times X_{2\alpha} \times \ldots \times X_{n\alpha}) \cap R\big\}, \\ \overline{y}^R(\alpha) &= \max\big\{f(x_1,x_2,\ldots,x_n) \mid (x_1,x_2,\ldots,x_n) \in (X_{1\alpha} \times X_{2\alpha} \times \ldots \times X_{n\alpha}) \cap R\big\}. \end{aligned} \tag{10}$$

3 Normalized Fuzzy Weights

In this section, a special structure of fuzzy numbers, normalized fuzzy weights, will be introduced, and its properties will be studied. Normalized fuzzy weights represent a fuzzification of crisp normalized weights that are defined as non-negative real numbers $w_1,w_2,\ldots,w_m$ such that $\sum_{i=1}^m w_i = 1$.

Fuzzy numbers $W_1,W_2,\ldots,W_m$ defined on $[0,1]$ are called normalized fuzzy weights if for any $\alpha \in (0,1]$ and for all $i \in N_m$ the following holds:

$$\text{For any } w_i \in W_{i\alpha} \text{ there exist } w_j \in W_{j\alpha},\ j \in N_m,\ j \ne i,\ \text{such that } w_i + \sum_{j=1, j\ne i}^{m} w_j = 1. \tag{11}$$

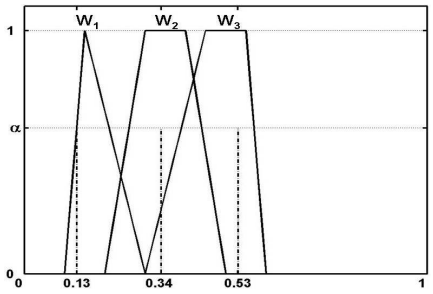


Figure 1: Example of normalized fuzzy weights.

Figure 1 illustrates an example of normalized fuzzy weights. Obviously, crisp normalized weights $w_1, w_2, \ldots, w_m$ represent a special case of normalized fuzzy weights.

From the general point of view, normalized fuzzy weights make it possible to model mathematically an uncertain division of a unit into m fractions. In decision making models, the membership degrees $W_i(w_i)$, $i \in N_m$, mean either the possibility that the share of the i-th partial objective of evaluation in the overall one is equal to w_i, or the possibility that the probability of the i-th state of the world is equal to w_i.

In [6], fuzzy numbers defined on $[0,1]$ fulfilling (11) were called *a tuple of fuzzy probabilities*. They represented the generalization of interval probabilities developed in [1], and they were applied for modelling imprecise probabilities. The structure of normalized fuzzy weights was independently introduced in [7] in order to model the uncertain weights of criteria in the method of the weighted average of partial fuzzy evaluations that was described in [8]. Later on, the normalized fuzzy weights were used to model uncertain probabilities in fuzzy-stochastic models of decision making under risk (see [10, 11]).

Verifying the condition (11) directly can be very complicated. Hence, it was proved in [6] that fuzzy numbers $W_i = (\underline{w}_i, \overline{w}_i)$, $i \in N_m$, defined on $[0,1]$ fulfil (11) if and only if for any $\alpha \in [0,1]$ and for all $i \in N_m$ the following two conditions hold

$$\overline{w}_i(\alpha) + \sum_{j=1, j\neq i}^{m} \underline{w}_j(\alpha) \leq 1, \tag{12}$$

$$\underline{w}_i(\alpha) + \sum_{j=1, j\neq i}^{m} \overline{w}_j(\alpha) \geq 1. \tag{13}$$

Linear fuzzy numbers $W_i = \langle w_i^1, w_i^2, w_i^3, w_i^4\rangle$, $i \in N_m$, defined on $[0,1]$ fulfil (11) if and only if the conditions (12) and (13) hold for $\alpha = 1$ and $\alpha = 0$, i.e. if for all $i \in N_m$ the following holds

$$w_i^3 + \sum_{j=1, j\neq i}^{m} w_j^2 \leq 1 \text{ and } w_i^2 + \sum_{j=1, j\neq i}^{m} w_j^3 \geq 1, \tag{14}$$

$$w_i^4 + \sum_{j=1, j\neq i}^{m} w_j^1 \leq 1 \text{ and } w_i^1 + \sum_{j=1, j\neq i}^{m} w_j^4 \geq 1. \tag{15}$$

The linearity of the functions $\underline{w}_i$ and $\overline{w}_i$, $i \in N_m$, ensures that the conditions (12) and (13) are fulfilled for any $\alpha \in [0,1]$ (see [6]).

For $m = 2$, we can see immediately from (12) and (13) that

$$\underline{w}_2(\alpha) = 1 - \overline{w}_1(\alpha) \text{ and } \overline{w}_2(\alpha) = 1 - \underline{w}_1(\alpha) \tag{16}$$

for all $\alpha \in [0,1]$. It means that $W_2 = 1 - W_1$.

For $i \in N_m$, let us denote by σ_{W_i} the width function of the fuzzy number $W_i = (\underline{w}_i, \overline{w}_i)$, i.e. $\sigma_{W_i}(\alpha) = \overline{w}_i(\alpha) - \underline{w}_i(\alpha)$, $\alpha \in [0,1]$. For the sake of simplicity, the conditions (12) and (13) can be, for all $\alpha \in [0,1]$, replaced by only one condition in one of the following two ways:

First, subtracting $\sum_{k=1}^{m} \underline{w}_k(\alpha)$ from (12) and (13) we obtain

$$\sigma_{W_i}(\alpha) \leq 1 - \sum_{k=1}^{m} \underline{w}_k(\alpha), \tag{17}$$

$$\sum_{j=1, j\neq i}^{m} \sigma_{W_j}(\alpha) \geq 1 - \sum_{k=1}^{m} \underline{w}_k(\alpha) \tag{18}$$

for all $i \in N_m$. Let $i^*(\alpha) \in N_m$ be an index such that

$$\sigma_{W_{i^*(\alpha)}}(\alpha) = \max_{i=1,2,\ldots,m} \{\sigma_{W_i}(\alpha)\}. \tag{19}$$

Then,

$$\sum_{\substack{j=1 \\ j\neq i^*(\alpha)}}^{m} \sigma_{W_j}(\alpha) = \min_{i=1,2,\ldots,m} \Big\{ \sum_{\substack{j=1 \\ j\neq i}}^{m} \sigma_{W_j}(\alpha) \Big\}. \tag{20}$$

Hence, the conditions (12) and (13) are equivalent to the following one

$$\sigma_{W_{i^*(\alpha)}}(\alpha) \leq 1 - \sum_{k=1}^{m} \underline{w}_k(\alpha) \leq \sum_{\substack{j=1 \\ j\neq i^*(\alpha)}}^{m} \sigma_{W_j}(\alpha). \tag{21}$$

Analogously, in the case of linear fuzzy numbers $W_i = \langle w_i^1, w_i^2, w_i^3, w_i^4\rangle$, $i \in N_m$, the following couple of conditions

$$\sigma_{W_{i^*(1)}}(1) \leq 1 - \sum_{k=1}^{m} w_k^2 \leq \sum_{j=1, j\neq i^*(1)}^{m} \sigma_{W_j}(1), \tag{22}$$

$$\sigma_{W_{i^*(0)}}(0) \leq 1 - \sum_{k=1}^{m} w_k^1 \leq \sum_{j=1, j\neq i^*(0)}^{m} \sigma_{W_j}(0) \tag{23}$$

is equivalent to (14) and (15).

Second, multiplying (12) and (13) by (-1) and adding $\sum_{k=1}^{m} \overline{w}_k(\alpha)$ to these inequalities, we can see that the conditions (12) and (13) are equivalent to the following one

$$\sigma_{W_{i^*(\alpha)}}(\alpha) \leq \sum_{k=1}^{m} \overline{w}_k(\alpha) - 1 \leq \sum_{\substack{j=1 \\ j\neq i^*(\alpha)}}^{m} \sigma_{W_j}(\alpha), \quad (24)$$

where the index $i^*(\alpha) \in N_m$ is given by (19). In the case of linear fuzzy numbers $W_i = \langle w_i^1, w_i^2, w_i^3, w_i^4\rangle$, $i \in N_m$, the following couple of conditions

$$\sigma_{W_{i^*(1)}}(1) \leq \sum_{k=1}^{m} w_k^3 - 1 \leq \sum_{j=1, j\neq i^*(1)}^{m} \sigma_{W_j}(1), \quad (25)$$

$$\sigma_{W_{i^*(0)}}(0) \leq \sum_{k=1}^{m} w_k^4 - 1 \leq \sum_{j=1, j\neq i^*(0)}^{m} \sigma_{W_j}(0) \quad (26)$$

is equivalent to (14) and (15).

At the end of this section, the case of symmetric normalized fuzzy weights will be studied. Let $w_1, w_2, \ldots, w_m$ be normalized weights and let $W_i = (\underline{w}_i, \overline{w}_i)$, $i \in N_m$, be symmetric fuzzy numbers such that for all $\alpha \in [0,1]$ the following holds

$$\underline{w}_i(\alpha) = w_i - e_i(\alpha) \text{ and } \overline{w}_i(\alpha) = w_i + e_i(\alpha), \quad (27)$$

where $e_i : [0,1] \to [0, \min\{w_i, 1-w_i\}]$, $i \in N_m$, are non-increasing functions. Then, from (21) or (24) it follows that the fuzzy numbers $W_1, W_2, \ldots, W_m$ fulfil the condition (11) if and only if for all $\alpha \in [0,1]$ the following holds

$$e_{i^*(\alpha)}(\alpha) \leq \sum_{j=1, j\neq i^*(\alpha)}^{m} e_j(\alpha), \quad (28)$$

where

$$e_{i^*(\alpha)}(\alpha) = \max_{i=1,2,\ldots,m} \{e_i(\alpha)\}. \quad (29)$$

Obviously, if e_i are linear functions for all $i \in N_m$, then (28) is fulfilled for all $\alpha \in [0,1]$ if and only if

$$e_{i^*(1)}(1) \leq \sum_{j=1, j\neq i^*(1)}^{m} e_j(1), \quad (30)$$

$$e_{i^*(0)}(0) \leq \sum_{j=1, j\neq i^*(0)}^{m} e_j(0). \quad (31)$$

4 Procedures of Setting Normalized Fuzzy Weights

In practical applications, it is useful to set normalized fuzzy weights as linear fuzzy numbers $W_i = \langle w_i^1, w_i^2, w_i^3, w_i^4\rangle$, $i \in N_m$. The interpretation is the following: $[w_i^2, w_i^3]$ represents the interval of fully possible values of the i-th weight, while outside the interval $[w_i^1, w_i^4]$ the i-th weight cannot lie.

For $m = 2$, it is sufficient to set only one weight in the form of a linear fuzzy number $W = \langle w^1, w^2, w^3, w^4\rangle$ defined on $[0,1]$; then it follows from (16) that the other is the linear fuzzy number $1 - W = \langle 1-w^4, 1-w^3, 1-w^2, 1-w^1\rangle$.

For $m > 2$, it is not so easy for an expert to set the normalized fuzzy weights directly. Therefore, two procedures of setting normalized fuzzy weights will be shown here. The first one is based on the transformation of expert's estimates of fuzzy weights into the normalized fuzzy weights. The second one is a fuzzification of a crisp estimation of normalized weights.

In the first procedure, an expert expresses his/her estimates of the weights by linear fuzzy numbers $W_i' = \langle w_i'^1, w_i'^2, w_i'^3, w_i'^4\rangle$, $i \in N_m$, whose significant values satisfy the following natural condition

$$\sum_{i=1}^{m} w_i'^1 \leq \sum_{i=1}^{m} w_i'^2 \leq 1 \leq \sum_{i=1}^{m} w_i'^3 \leq \sum_{i=1}^{m} w_i'^4. \quad (32)$$

It means that the fuzzy numbers $W_1', W_2', \ldots, W_m'$ need not fulfil the condition (11), but there must exist at least one m-tuple of normalized weights $w_1, w_2, \ldots, w_m$ such that $W_i'(w_i) = 1$, $i \in N_m$. The associated normalized fuzzy weights $W_i = \langle w_i^1, w_i^2, w_i^3, w_i^4\rangle$, $i \in N_m$, are obtained from $W_1', W_2', \ldots, W_m'$ by the following transformation

$$\begin{aligned} w_i^k &= \max\{w_i'^k, 1 - \sum_{j=1, j\neq i}^{m} w_j'^{5-k}\}, \\ w_i^l &= \min\{w_i'^l, 1 - \sum_{j=1, j\neq i}^{m} w_j'^{5-l}\}, \end{aligned} \quad (33)$$

for $k = 1, 2$, $l = 3, 4$. The transformation (33), that was originally described for interval probabilities in [1], eliminates the inconsistency of expert's estimates (see Figure 2). For any $\alpha \in (0,1]$ and for every m-tuple of crisp normalized weights $w_1, w_2, \ldots, w_m$ such that $w_i \in W_{i\alpha}'$ for all $i \in N_m$, it holds that $w_i \in W_{i\alpha}$ for all $i \in N_m$. It means that no relevant information concerning the weights was lost by the transformation.

In the second procedure, an expert sets the crisp estimation of normalized weights by real numbers $w_1, w_2, \ldots, w_m$ such that $w_i \geq 0$, $i \in N_m$, and $\sum_{i=1}^{m} w_i = 1$. Then, he/she sets coefficients k_i and s_i, $0 \leq k_i \leq s_i$, $w_i - s_i \geq 0$, $w_i + s_i \leq 1$, $i \in N_m$, that characterize the uncertainty of kernels and supports of the particular fuzzy weights. It follows from

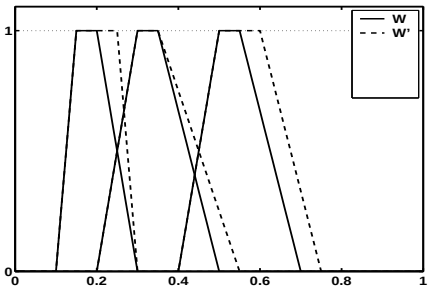


Figure 2: Transformation of expert's estimation of fuzzy weights into the normalized fuzzy weights.

(30) and (31) that the symmetric linear fuzzy numbers $W_i = \langle w_i - s_i, w_i - k_i, w_i + k_i, w_i + s_i\rangle$, $i \in N_m$, are normalized fuzzy weights if k_i and s_i, $i \in N_m$, satisfy the following condition

$$k_{i^*} \leq \sum_{i=1, i\neq i^*}^{m} k_i \text{ and } s_{j^*} \leq \sum_{j=1, j\neq j^*}^{m} s_j, \qquad (34)$$

where $k_{i^*} = \max\{k_i,\ i \in N_m\}$ and $s_{j^*} = \max\{s_j,\ j \in N_m\}$. For instance, the condition (34) is satisfied in the case of symmetric triangular fuzzy numbers $W_i = \langle w_i - s, w_i, w_i, w_i + s\rangle$, $i \in N_m$, where $w_1, w_2, \ldots, w_m$ are normalized weights, $s \geq 0$, and, for all $i \in N_m$, $w_i - s \geq 0$ and $w_i + s \leq 1$.

5 Fuzzy Weighted Averages of Fuzzy Numbers

A weighted average of real numbers $u_1, u_2, \ldots, u_m$ with weights $w_1, w_2, \ldots, w_m$ is defined by the following formula

$$u = \frac{w_1u_1 + w_2u_2 + \ldots + w_mu_m}{w_1 + w_2 + \ldots + w_m}, \qquad (35)$$

where the weights $w_1, w_2, \ldots, w_m$ are generally non-negative real numbers whose sum is different from zero. Especially, if $\sum_{i=1}^{m} w_i = 1$, i.e. if the weights $w_1, w_2, \ldots, w_m$ are normalized, the weighted average can be expressed as follows

$$u = w_1u_1 + w_2u_2 + \ldots + w_mu_m. \qquad (36)$$

In this paper, it will be shown that these formulas do not coincide in fuzzy case.

Let $W_1, W_2, \ldots, W_m$ be non-negative fuzzy numbers, i.e. $Supp\, W_i \subset [0, \infty)$ for all $i \in N_m$, and let there exist $w > 0$ and an index $i \in N_m$ such that $W_i(w) = 1$. Then a fuzzy weighted average of fuzzy numbers $U_1, U_2, \ldots, U_m$ with fuzzy weights $W_1, W_2, \ldots, W_m$ was defined in [2] as a fuzzy number U^W whose membership function is given, for each $u \in \Re$, in the following way

$$U^W(u) = \max\big\{\min\{U_1(u_1), U_2(u_2), \ldots, U_m(u_m), W_1(w_1), W_2(w_2), \ldots, W_m(w_m)\} \mid u = \tfrac{w_1u_1+w_2u_2+\ldots+w_mu_m}{w_1+w_2+\ldots+w_m}, \textstyle\sum_{i=1}^{m} w_i > 0\big\}. \qquad (37)$$

The fuzzy weighted average U^W represents a fuzzification according to (9) of the operation (35). An algorithm of computing the fuzzy weighted average U^W was described in [4].

The fuzzy weighted average U^W cannot be used, if the fuzzy weights $W_1, W_2, \ldots, W_m$ have the special interpretation described in Section 1: If they express either uncertain shares of partial objectives of evaluation in the overall one in multiple criteria decision making models or uncertain probabilities of states of the world in models of decision making under risk. In such cases, the structure of normalized fuzzy weights has to be used, and a fuzzification according to (9) of the formula (36) instead of the formula (35) has to be applied for computing the fuzzy weighted average. The problem is illustrated by the following example:

Example 1: Let S_1, S_2 and S_3 be states of the world; their uncertain probabilities be, for simplicity, defined by the intervals $P_1 = [0.1, 0.2]$, $P_2 = [0.4, 0.6]$ and $P_3 = [0.3, 0.4]$ (P_1, P_2 and P_3 express possible values of probabilities and, therefore, form the structure of normalized fuzzy weights). Let evaluations of an alternative x under the states of the world S_1, S_2, S_3 be given by the intervals $U_1 = [0.1, 0.2]$, $U_2 = [0.4, 0.5]$ and $U_3 = [0.8, 0.9]$. Then, according to (37) where W_i are replaced by P_i for $i \in \{1, 2, 3\}$, $U^W = [0.45, 0.64]$. When the minimal value of U^W, i.e. 0.45, is computed, the third weight $\frac{p_3}{p_1+p_2+p_3}$ is equal to $\frac{0.3}{0.2+0.6+0.3} = 0.27$, but $0.27 \notin P_3$. Notice that this value lies outside the expertly set range of possible probabilities of S_3; therefore, a fuzzy expected evaluation of the alternative x cannot be computed by (37).

Let fuzzy numbers $W_1, W_2, \ldots, W_m$ be normalized fuzzy weights. Then the fuzzy weighted average of fuzzy numbers $U_1, U_2, \ldots, U_m$ with normalized fuzzy weights $W_1, W_2, \ldots, W_m$ is defined as a fuzzy number U^N whose membership function is given, for each $u \in \Re$, as follows

$$U^N(u) = \max\big\{\min\{U_1(u_1), U_2(u_2), \ldots, U_m(u_m), W_1(w_1), W_2(w_2), \ldots, W_m(w_m)\} \mid u = \textstyle\sum_{i=1}^{m} w_i\, u_i,\ \sum_{i=1}^{m} w_i = 1\big\}. \qquad (38)$$

The fuzzy weighted average U^N represents the generalization of the operation that was introduced in [1] for computing the expected value of a discrete random variable with interval probabilities. The fuzzy

weighted average U^N was independently introduced in [7] for aggregating partial fuzzy evaluations of an alternative in the model of multiple criteria evaluation. The following denotation will be used for the fuzzy weighted average U^N of fuzzy numbers $U_1, U_2, \ldots, U_m$ with normalized fuzzy weights $W_1, W_2, \ldots, W_m$

$$U^N = (\mathcal{N}) \sum_{i=1}^{m} W_i \cdot U_i. \tag{39}$$

In the case of crisp normalized weights, the fuzzy weighted averages U^W and U^N coincide. However, this does not hold generally for normalized fuzzy weights. If $W_1, W_2, \ldots, W_m$ form the structure of normalized fuzzy weights, then $U^N_\alpha \subseteq U^W_\alpha$ for all $\alpha \in [0,1]$. It follows from the fact that even such $w_1, w_2, \ldots, w_m$ whose sum is not equal to one enter the calculation of U^W. For example, the fuzzy weighted average U^N that expresses the fuzzy expected evaluation of the alternative x in Example 1 is equal to $[0.46, 0.63]$.

The following algorithm of computing the fuzzy weighted average U^N is based on an algorithm that was originally developed in [1] for computing the expected value of a discrete random variable with interval probabilities. The following representation of the considered fuzzy numbers will be used: $U^N = (\underline{u}^N, \overline{u}^N)$, $U_i = (\underline{u}_i, \overline{u}_i)$ and $W_i = (\underline{w}_i, \overline{w}_i)$, $i \in N_m$.

For each $\alpha \in [0,1]$, let $\{i_k\}_{k=1}^m$ be such a permutation on an index set N_m that $\underline{u}_{i_1}(\alpha) \leq \underline{u}_{i_2}(\alpha) \leq \ldots \leq \underline{u}_{i_m}(\alpha)$. For $k \in N_m$, let us denote

$$w_{i_k}(\alpha) = 1 - \sum_{j=1}^{k-1} \overline{w}_{i_j}(\alpha) - \sum_{j=k+1}^{m} \underline{w}_{i_j}(\alpha). \tag{40}$$

Let $k^* \in N_m$ be such an index that $\underline{w}_{i_{k^*}}(\alpha) \leq w_{i_{k^*}}(\alpha) \leq \overline{w}_{i_{k^*}}(\alpha)$. Then

$$\underline{u}^N(\alpha) = \sum_{j=1}^{k^*-1} \overline{w}_{i_j}(\alpha) \cdot \underline{u}_{i_j}(\alpha) +$$

$$w_{i_{k^*}}(\alpha) \cdot \underline{u}_{i_{k^*}}(\alpha) + \sum_{j=k^*+1}^{m} \underline{w}_{i_j}(\alpha) \cdot \underline{u}_{i_j}(\alpha). \tag{41}$$

Let $\{i_h\}_{h=1}^m$ be such a permutation on an index set N_m that $\overline{u}_{i_1}(\alpha) \geq \overline{u}_{i_2}(\alpha) \geq \ldots \geq \overline{u}_{i_m}(\alpha)$. For $h \in N_m$, let us denote

$$w_{i_h}(\alpha) = 1 - \sum_{j=1}^{h-1} \overline{w}_{i_j}(\alpha) - \sum_{j=h+1}^{m} \underline{w}_{i_j}(\alpha). \tag{42}$$

Let $h^* \in N_m$ be such an index that $\underline{w}_{i_{h^*}}(\alpha) \leq w_{i_{h^*}}(\alpha) \leq \overline{w}_{i_{h^*}}(\alpha)$. Then

$$\overline{u}^N(\alpha) = \sum_{j=1}^{h^*-1} \overline{w}_{i_j}(\alpha) \cdot \overline{u}_{i_j}(\alpha) +$$

$$w_{i_{h^*}}(\alpha) \cdot \overline{u}_{i_{h^*}}(\alpha) + \sum_{j=h^*+1}^{m} \underline{w}_{i_j}(\alpha) \cdot \overline{u}_{i_j}(\alpha). \tag{43}$$

Let us consider now the case where the fuzzy numbers $U_i = (\underline{u}_i, \overline{u}_i)$ and the normalized fuzzy weights $W_i = (\underline{w}_i, \overline{w}_i)$, $i \in N_m$, are linear fuzzy numbers, i.e. the functions $\underline{u}_i$, $\overline{u}_i$ as well as $\underline{w}_i$, $\overline{w}_i$ are linear. It follows from (41) and (43) that the fuzzy weighted average $U^N = (\mathcal{N}) \sum_{i=1}^m W_i \cdot U_i$ is not a linear fuzzy number, in general.

6 The Fuzzy Weighted Average U^N in Decision Making Models

In this section, a multiple criteria decision making model and a model of decision making under risk where the fuzzy weighted average U^N is applied, will be described.

First, let us consider a problem of multiple criteria decision making, where the best of alternatives $x_1, x_2, \ldots, x_n$ is to be chosen. The alternatives are being evaluated with respect to a given objective that is partitioned into m partial objectives associated with criteria $C_1, C_2, \ldots, C_m$. Let the uncertain information about the shares of the partial objectives in the overall one be given by normalized fuzzy weights $W_1, W_2, \ldots, W_m$. Let uncertain partial fuzzy evaluations of alternatives x_i, $i \in N_n$, with respect to the criteria C_j, $j \in N_m$, be expressed by fuzzy numbers $U_{i,j}$ defined on $[0,1]$ that represent the fuzzy degrees of fulfilment of the corresponding partial objectives of evaluation. Then the overall fuzzy evaluation U_i^N of the alternative x_i, $i \in N_n$, can be expressed as follows

$$U_i^N = (\mathcal{N}) \sum_{j=1}^{m} W_j \cdot U_{i,j} \ . \tag{44}$$

The overall fuzzy evaluations U_i^N, $i \in N_n$, express the fuzzy degrees of fulfilment of the overall objective of evaluation. The best alternative is the first alternative in an ordering of the fuzzy numbers U_i, $i \in N_n$, or the closest to the ideal alternative whose evaluation is equal to 1. The overall fuzzy evaluations of alternatives can also be approximated linguistically by linearly ordered elements of a proper linguistic evaluation scale defined on $[0,1]$. For more details on metrics, ordering of fuzzy numbers and linguistic approximation see [8].

Second, let us consider a problem of decision making under risk that is described by the following fuzzy decision matrix (see Table 1), where $x_1, x_2, \ldots, x_n$ are alternatives, $S_1, S_2, \ldots, S_r$ states of the world, $P_1, P_2, \ldots, P_r$ their fuzzy probabilities, and $U_{i,k}$, $i \in N_n$, $k \in N_r$, denote fuzzy degrees in which the alter-

natives x_i, $i \in N_n$, satisfy a given decision objective if the states S_k, $k \in N_r$, occur.

Table 1: Fuzzy decision matrix.

	S_1 P_1	S_2 P_2	... ...	S_r P_r	FEU^N
x_1	$U_{1,1}$	$U_{1,2}$	...	$U_{1,r}$	FEU_1^N
x_2	$U_{2,1}$	$U_{2,2}$	...	$U_{2,r}$	FEU_2^N
...	...	...	...	...	...
x_n	$U_{n,1}$	$U_{n,2}$	...	$U_{n,r}$	FEU_n^N

For $i \in N_n$, the fuzzy numbers FEU_i^N express expected fuzzy evaluations of alternatives x_i; i.e. they are calculated according to the formula

$$FEU_i^N = (\mathcal{N}) \sum_{k=1}^{r} P_k \cdot U_{i,k} \ . \tag{45}$$

The best alternative is the alternative with the maximum expected fuzzy evaluation.

Examples of practical applications of such decision making models, namely in finance and banking, are described in [10, 11]. A similar approach can be applied also to multiple criteria decision making under risk (see [9]).

7 Width Function of the Fuzzy Weighted Average U^N

In the last section, the dependance of the uncertainty of the fuzzy weighted average U^N on given fuzzy numbers $U_1, U_2, \ldots, U_m$ and on given normalized fuzzy weights $W_1, W_2, \ldots, W_m$ will be studied.

The uncertainty of $U^N = (\underline{u}^N, \overline{u}^N)$ is described by its width function $\sigma_{U^N}(\alpha) = \overline{u}^N(\alpha) - \underline{u}^N(\alpha)$, $\alpha \in [0,1]$. Let us denote by σ_{U_i} the width functions of the fuzzy numbers $U_i = (\underline{u}_i, \overline{u}_i)$ for $i \in N_m$, i.e. $\sigma_{U_i}(\alpha) = \overline{u}_i(\alpha) - \underline{u}_i(\alpha)$, $\alpha \in [0,1]$. Let $w_\alpha^* = (w_{1\alpha}^*, w_{2\alpha}^*, \ldots, w_{m\alpha}^*)$ and $w_\alpha^{**} = (w_{1\alpha}^{**}, w_{2\alpha}^{**}, \ldots, w_{m\alpha}^{**})$ be, for each $\alpha \in [0,1]$, the couple of normalized weights such that $w_{i\alpha}^*, w_{i\alpha}^{**} \in W_{i\alpha}$, $i \in N_m$, $\sum_{i=1}^m w_{i\alpha}^* = \sum_{i=1}^m w_{i\alpha}^{**} = 1$, and

$$\underline{u}^N(\alpha) = \sum_{i=1}^{m} w_{i\alpha}^* \cdot \underline{u}_i(\alpha), \ \overline{u}^N(\alpha) = \sum_{i=1}^{m} w_{i\alpha}^{**} \cdot \overline{u}_i(\alpha). \tag{46}$$

Then, for all $\alpha \in [0,1]$, the width function of U^N is given by

$$\sigma_{U^N}(\alpha) = \sum_{i=1}^{m} w_{i\alpha}^{**} \cdot \overline{u}_i(\alpha) - \sum_{i=1}^{m} w_{i\alpha}^* \cdot \underline{u}_i(\alpha). \tag{47}$$

From (47) it can be derived the following formula

$$\sigma_{U^N}(\alpha) = \sum_{i=1}^{m} \frac{w_{i\alpha}^* + w_{i\alpha}^{**}}{2} \cdot \sigma_{U_i}(\alpha) + \sum_{i=1}^{m} (w_{i\alpha}^{**} - w_{i\alpha}^*) \cdot \frac{\underline{u}_i(\alpha) + \overline{u}_i(\alpha)}{2}. \tag{48}$$

Let us denote

$$\varepsilon_i(\alpha) = \frac{\underline{u}_i(\alpha) + \overline{u}_i(\alpha)}{2} - \frac{1}{m} \cdot \sum_{j=1}^{m} \frac{\underline{u}_j(\alpha) + \overline{u}_j(\alpha)}{2}, \tag{49}$$

then the formula (48) can be rewritten as follows

$$\sigma_{U^N}(\alpha) = \sum_{i=1}^{m} \frac{w_{i\alpha}^* + w_{i\alpha}^{**}}{2} \cdot \sigma_{U_i}(\alpha) + \sum_{i=1}^{m} (w_{i\alpha}^{**} - w_{i\alpha}^*) \cdot \varepsilon_i(\alpha). \tag{50}$$

The following two examples will illustrate how the dependance characterized by the formula (50) affects the behaviour of the decision making models described above.

Example 2: Let a pair of normalized fuzzy weights be given by the triangular fuzzy numbers $W_1 = W_2 = \langle 0, 0.5, 0.5, 1 \rangle$, and let two pairs of weighted values be given by the real numbers $u_{11} = 0.1$, $u_{12} = 0.6$, and $u_{21} = u_{22} = 0.8$. Figure 3 illustrates the fact that the more different the weighted values are, the more the uncertainty of normalized fuzzy weights affects the uncertainty of the fuzzy weighted average U^N. In the first case, where the weighted values are different, $u_{11} \neq u_{12}$, their fuzzy weighted average $U_1^N = \langle 0.1, 0.35, 0.35, 0.6 \rangle$. In the second case, where the weighted values are equal, $u_{21} = u_{22}$, their fuzzy weighted average $U_2^N = 0.8$; thus the uncertainty of the normalized weights does not affect the fuzzy weighted average at all.

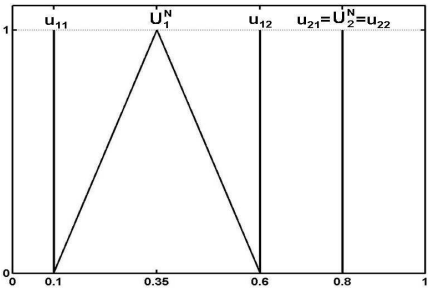


Figure 3: The dependance of the uncertainty of the fuzzy weighted average on the variability of the weighted values.

For the multiple criteria decision making model mentioned above it means the following: If the uncertainties of the partial fuzzy evaluations are approximately

the same, then the overall fuzzy evaluation of an alternative whose partial fuzzy evaluations according to the particular criteria are very different is more uncertain than the overall fuzzy evaluation of an alternative whose partial fuzzy evaluations are almost uniform. Similarly, in the discrete model of decision making under risk, the expected fuzzy evaluation of an alternative that depends strongly on states of the world is more uncertain than the expected fuzzy evaluation of an alternative that is relatively stable.

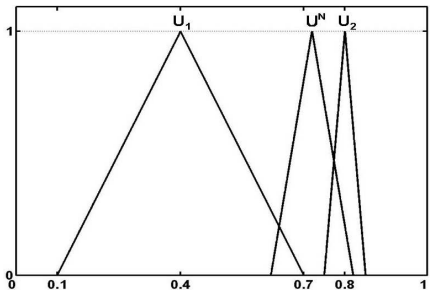


Figure 4: Fuzzy weighted average U^N with normalized weights $w_1 = 0.2$ and $w_2 = 0.8$.

Example 3: Let the weighted values be given by the triangular fuzzy numbers $U_1 = \langle 0.1, 0.4, 0.4, 0.7\rangle$ and $U_2 = \langle 0.75, 0.8, 0.8, 0.85\rangle$, and the normalized weights by the real numbers $w_1 = 0.2$ and $w_2 = 0.8$. The fuzzy weighted average of U_1 and U_2 with the normalized weights w_1 and w_2 is $U_1^N = \langle 0.62, 0.72, 0.72, 0.82\rangle$. Figure 4 shows that the uncertainty of U^N depends, apart from other things, on the size of the weights.

It means that the uncertainty of the partial fuzzy evaluations with small weights (or with small probabilities) cannot affect strongly the uncertainty of the overall fuzzy evaluation (fuzzy expected evaluation).

8 Conclusion

In decision making models based on the weighted average operation, weights of criteria as well as probabilities of states of the world are usually defined expertly. Therefore the models become more realistic if uncertainty of the weights and probabilities is taken into account.

In the paper, the uncertain weights and probabilities were expressed by normalized fuzzy weights, and the fuzzy weighted average operation was applied to computation of the aggregated fuzzy evaluations in models of multiple criteria decision making and to computation of the fuzzy expected evaluations in models of decision making under risk. It was shown that the uncertainty of the resulting fuzzy evaluations depends not only on the uncertainty of the normalized fuzzy weights and on the uncertainty of the weighted fuzzy values but also on the variance of the weighted fuzzy values.

References

[1] L. M. De Campos, J. F. Heute, S. Moral, Probability inrtervals: a tool for uncertain reasoning, Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 2 (1994) 167–196.

[2] W. M. Dong, F. S. Wong, Fuzzy weighted averages and implementation of the extension principle, Fuzzy Sets and Systems 21 (1987) 183–199.

[3] D. Dubois, E. Kerre, R. Mesiar, H. Prade, Fuzzy interval analysis, in: D. Dubois, H. Prade (Eds.), *Fundamentals of fuzzy sets*. Kluwer Academic Publishers, Boston-London-Dordrecht, 2000, pp. 483–582.

[4] Y. -Y. Guh, et al., Fuzzy weighted average: The linear programming approach via Charness and Cooper's rule, Fuzzy Sets and Systems 117 (2001) 157–160.

[5] G. J. Klir, Y. Pan, Constrained fuzzy arithmetic: Basic questions and some answers, Soft Computing 2 (1998) 100-108.

[6] Y. Pan, B. Yuan, Bayesian Inference of Fuzzy Probabilities, Int. J. General Systems 26 (1997) 73–90.

[7] O. Pavlačka, Normalized fuzzy weights (in Czech), Diploma thesis, Palacký University Olomouc, 2004.

[8] J. Talašová, Fuzzy Methods of Multiple Criteria Evaluation and Decision Making (in Czech), VUP, Olomouc 2003.

[9] J. Talašová, Fuzzy sets in decision-making under risk. (In Czech.), in: M. Kováčová (Eds.), *Proceedings of the 4th Conference Aplimat*. EX, Bratislava 2005, pp. 532–544

[10] J. Talašová, Fuzzy approach to evaluation and decision making (in Czech), in: *Proceedings of Aplimat 2006*, SUT, Bratislava 2006, pp. 221–236.

[11] J. Talašová, O. Pavlačka, Fuzzy Probability Spaces and Their Applications in Decision Making, Austrian J. of Statistics 35 (2006) 347–356.

The induced generalized OWA operator

José M. Merigó
Department of Business Administration,
University of Barcelona,
Av. Diagonal 690, 08034, Barcelona
Email: josema_merigo@hotmail.com

Anna M. Gil-Lafuente
Department of Business Administration,
University of Barcelona,
Av. Diagonal 690, 08034, Barcelona
Email: amgil@ub.edu

Abstract

We study different types of aggregation operators. We focus on the generalized OWA (GOWA) operator developed by Yager which represents a generalization to a wide range of aggregation operators. We distinguish between aggregations with a descending or with an ascending order. We introduce the induced generalized OWA (IGOWA) operator which represents an extension to the GOWA operator. It generalizes a wider range of aggregation operators as the GOWA operator is a particular case of this type of generalization. We study its main properties and some particular cases obtained with it. Finally, we develop a further generalization to the IGOWA operator by using quasi-arithmetic means.

Keywords: Generalized mean, Induced OWA operator, Aggregation operators, Quasi-arithmetic means.

1 Introduction

In the literature, we find a wide range of aggregation operators for aggregating the information. A very common aggregation method is the ordered weighted averaging (OWA) operator introduced in [17]. Since its appearance, the OWA operator has been used in a wide range of applications such as [2,30]. It provides a parameterized family of aggregation operators that includes the maximum, the minimum and the average criteria.

In [29], Yager and Filev developed an extension to the OWA operator called induced ordered weighted averaging (IOWA) operator. The difference is that the reordering step is not developed with the values of the arguments. In this case, the reordering step is induced by another mechanism such that the ordered position of the arguments depends upon the values of their associated order inducing variables. In the last years, the IOWA operator has been receiving increasing attention as it is seen in the different works developed about it such as in [5,11,16,22-23,25].

Recently, Yager [26] has developed a generalized version of the OWA operator where he introduces a reordering step in the generalized mean according to the values of the arguments. The generalized mean was introduced in [6-7] and it generalizes a wide range of mean operators such as the arithmetic mean, the geometric mean, the harmonic mean and the quadratic mean. Then, with the generalized OWA operator, Yager offered a generalization that included the particular cases found in the OWA operator such as the maximum and the minimum and the particular cases found in the generalized mean such as the geometric or the harmonic mean. More recently, Beliakov [1] has suggested a further generalization to the GOWA operator by using quasi-arithmetic means. Then, the result obtained is the Quasi-OWA operator introduced in [8]. This operator has also been studied in [2]. Note that these generalizations are different from the ones developed in [13,16].

In this paper, we suggest an induced generalized OWA (IGOWA) operator which represents an extension to the GOWA operator. In this generalization, we suggest a reordering step based on a mechanism such that the ordered position of the arguments depends upon their associated order inducing variables. Then, we can obtain a generalization that includes the special cases found in the IOWA operator such as the maximum, the minimum and the OWA operator, and the special cases found in the generalized mean such as the arithmetic or the geometric mean. We will also suggest a further generalization to the IGOWA operator by using quasi-arithmetic means. We will call it the Quasi-IOWA operator.

In order to do so, this paper is organized as follows. In Section 2, we briefly describe some aggregations operators such as the OWA operator, the IOWA operator and the generalized mean. In Section 3, we study the main concepts of the GOWA operator. In Section 4, we introduce the IGOWA operator. In

Section 5, we develop different families of IGOWA operators. In Section 6 we introduce the Quasi-IOWA operator. Finally, in Section 7, we summarize the main conclusions found in the paper.

2 Aggregation operators

In this Section we will briefly describe the OWA operator, the IOWA operator and the generalized mean.

2.1 OWA operator

The OWA operator was introduced by Yager in [17] and it provides a parameterized family of aggregation operators that include the arithmetic mean, the maximum and the minimum. It can be defined as follows.

Definition 1. An OWA operator of dimension n is a mapping $OWA:R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$OWA(a_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} w_j b_j \qquad (1)$$

where b_j is the jth largest of the a_i.

From a generalized perspective of the reordering step, we have to distinguish between the descending OWA (DOWA) operator and the ascending OWA (AOWA) operator [18]. The OWA operator is a mean or averaging operator. This is a reflection of the fact that the operator is commutative, monotonic, bounded and idempotent [17].

2.2 IOWA operator

The IOWA operator was introduced by Yager and Filev [29] and it represents an extension to the OWA operator. Its main difference is that the reordering step is not developed with the values of the arguments a_i. In this case, the reordering step is developed with order inducing variables. The IOWA operator also includes as particular cases the maximum, the minimum and the average criteria. In the following, we provide a definition of the IOWA operator as suggested in [29].

Definition 2. An IOWA operator of dimension n is a mapping $IOWA: R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$IOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle \ldots, \langle u_n,a_n\rangle) = \sum_{j=1}^{n} w_j b_j \qquad (2)$$

where b_j is the a_i value of the IOWA pair $\langle u_i,a_i\rangle$ having the jth largest u_i, u_i is the order inducing variable and a_i is the argument variable.

From a generalized perspective of the reordering step, we have to distinguish between the Descending IOWA (DIOWA) operator and the Ascending IOWA (AIOWA) operator [11]. The IOWA operator is also monotonic, bounded, idempotent and commutative [29] both for the DIOWA and the AIOWA operator.

2.3 Generalized mean

The generalized mean was introduced in [6-7] and it represents a generalization to a wide range of mean aggregations such as the arithmetic mean, the geometric mean, the harmonic mean and the quadratic mean. It can be defined as follows.

Definition 3. A generalized mean of dimension n is a mapping $GM:R^n \to R$ such that:

$$GM(a_1, a_2, \ldots, a_n) = \left(\frac{1}{n} \sum_{i=1}^{n} a_i^{\lambda} \right)^{1/\lambda} \qquad (3)$$

where a_i is the argument variable and λ is a parameter such that $\lambda \in (-\infty, \infty)$. Note that depending on the value of the parameter λ, we obtain different types of means. When $\lambda = \infty$, we obtain the maximum. When $\lambda = 1$, the arithmetic mean. When $\lambda = 0$, the geometric mean. When $\lambda = -1$, the harmonic mean. When $\lambda = 2$, the quadratic mean. When $\lambda = -\infty$, the minimum.

Another type of generalized mean is the case when the weights of the arguments are different. Then, we obtain the weighted generalized mean. It can be defined as follows.

Definition 4. A weighted generalized mean of dimension n is a mapping $WGM:R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$WGM(a_1, a_2, \ldots, a_n) = \left(\sum_{i=1}^{n} w_i a_i^{\lambda} \right)^{1/\lambda} \qquad (4)$$

where a_i is the argument variable and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

3 Generalized OWA operator

The generalized OWA (GOWA) operator was introduced by Yager in [26]. In the following, we are going to define it.

Definition 5. A GOWA operator of dimension n is a mapping $GOWA: R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$GOWA(a_1, a_2, \ldots, a_n) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \quad (5)$$

where b_j is the jth largest of the a_i, and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

From a generalized perspective of the reordering step, we have to distinguish between the descending generalized OWA (DGOWA) operator and the ascending generalized OWA (AGOWA) operator. Note that it is possible to use them in situations where the highest value is the best result and in situations where the lowest value is the best result. The DGOWA operator has the same definition than the GOWA operator. And the AGOWA operator is related to the GOWA operator by $w_j = w^*_{n+1-j}$, where w_j is the jth weight of the DGOWA (or GOWA) operator and w^*_{n+1-j} the jth weight of the AGOWA operator.

As it is demonstrated in [26], the GOWA operator is a mean operator. This is a reflection of the fact that the operator is commutative, monotonic, bounded and idempotent both for the DGOWA and the AGOWA operator. It can also be demonstrated that the GOWA operator has as special cases the maximum, the minimum, the generalized mean and the weighted generalized mean. Note that the weighted generalized mean is obtained when $j = i$, for all i and j, where j is the jth argument of the b_j and i is the ith argument of the a_i.

If we look to different values of the parameter λ, we can also obtain other special cases as the usual OWA operator [17], the ordered weighted geometric (OWG) operator [3,15], the ordered weighted harmonic averaging (OWHA) operator [26] and the ordered weighted quadratic averaging (OWQA) operator [26]. When $\lambda = 1$, with the DGOWA operator, we obtain the usual OWA operator and with the AGOWA operator, we obtain the AOWA operator. When $\lambda = 0$, we get the descending OWG (DOWG) operator and the ascending OWG (AOWG) operator. When $\lambda = -1$, we obtain the descending OWHA (DOWHA) operator and the ascending OWHA (AOWHA) operator. When $\lambda = 2$, we get the descending OWQA (DOWQA) operator and the ascending OWQA (AOWQA) operator respectively.

Another interesting issue to consider is the attitudinal character of the GOWA operator. In [26], Yager defined it as:

$$\alpha(W) = \left(\sum_{j=1}^{n} w_j \left(\frac{n-j}{n-1} \right)^{\lambda} \right)^{1/\lambda} \quad (6)$$

It can be shown that $\alpha \in [0, 1]$. The more of the weight located near the top of W, the closer α to 1 and the more of the weight located toward the bottom of W, the closer α to 0. Note that for the optimistic criteria $\alpha(W) = 1$ and for the pessimistic criteria $\alpha(W) = 0$. Here, we could also make a distinction between descending and ascending orders by using $w_j = w^*_{n+1-j}$, where w_j is the jth weight of the GOWA operator and w^*_{n+1-j} the jth weight of the AGOWA operator.

4 Induced generalized OWA operator

The induced generalized OWA (IGOWA) operator represents an extension to the GOWA operator. The main difference between them is that the reordering step of the IGOWA operator is not developed with the values of the arguments a_i. In this case, the reordering step is induced by another mechanism represented as u_i, where the ordered position of the arguments a_i depends upon the values of the order inducing variable u_i.

Definition 6. An IGOWA operator of dimension n is a mapping $IGOWA: R^n \to R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$IGOWA(\langle u_1, a_1 \rangle, \ldots, \langle u_n, a_n \rangle) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \quad (7)$$

where b_j is the a_i value of the IGOWA pair $\langle u_i, a_i \rangle$ having the jth largest u_i, u_i is the order inducing variable, a_i is the argument variable and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

From a generalized perspective of the reordering step, we have to distinguish between the descending induced generalized OWA (DIGOWA) operator and the ascending induced generalized OWA (AIGOWA) operator. Note that they can be used in situations where the highest value is the best result and in situations where the lowest value is the best result.

But in a more efficient context, it is better to use one of them for one situation and the other one for the other situation. The DIGOWA operator has the same definition than the IGOWA operator.

Definition 7. An AIGOWA operator of dimension n is a mapping $AIGOWA: R^n \rightarrow R$ that has an associated weighting vector W of dimension n such that the sum of the weights is 1 and $w_j \in [0,1]$, then:

$$AIGOWA(\langle u_1,a_1\rangle, \ldots, \langle u_n,a_n\rangle) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda}\right)^{1/\lambda} \quad (8)$$

where b_j is the a_i value of the IGOWA pair $\langle u_i,a_i\rangle$ having the jth lowest u_i, u_i is the order inducing variable, a_i is the argument variable and λ is a parameter such that $\lambda \in (-\infty, \infty)$. As we can see, the elements b_j (j= 1, 2, …, n) are ordered in an increasing way according to the values of u_i: $\langle \text{Min}\{u_i\},b_1\rangle \leq \ldots \leq \langle \text{Max}\{u_i\},b_n\rangle$.

The IGOWA operator is a mean or averaging operator. This is a reflection of the fact that the operator is commutative, monotonic, bounded and idempotent both for the DIGOWA and the AIGOWA operator. It is commutative because any permutation of the arguments has the same evaluation. That is, $IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) = IGOWA(\langle u_1,d_1\rangle, \langle u_2,d_2\rangle\ldots, \langle u_n,d_n\rangle)$, where $(d_1,\ldots, d_n)$ is any permutation of the arguments $(a_1,\ldots, a_n)$. It is monotonic because if $a_i \geq d_i$, for all a_i, then, $IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) \geq IGOWA(\langle u_1,d_1\rangle, \langle u_2,d_2\rangle\ldots, \langle u_n,d_n\rangle)$. It is bounded because the IGOWA aggregation is delimitated by the minimum and the maximum. That is, $\text{Min}\{a_i\} \leq IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) \leq \text{Max}\{a_i\}$. It is idempotent because if $a_i = a$, for all a_i, then, $IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) = a$.

An interesting issue when analysing induced aggregation operators is the problem of ties in the reordering step. In order to solve this problem, we recommend to follow the policy developed by Yager and Filev [29] where they replace each argument of the tied IOWA pairs by their average. For the GOWA operator, instead of using the arithmetic mean, we will replace each argument of the tied IGOWA pairs by its generalized mean. Then, depending on the parameter of λ, we will use a different type of mean to replace the tied arguments.

As it is explained in [29] for the IOWA operator, when studying the order inducing variable of the IGOWA operator, we should note that the values used can be drawn from a space such that the only requirement is to have a linear ordering. Then, it is possible to use different kinds of attributes for the order inducing variables that permit us, for example, to mix numbers with words in the aggregations [31]. For the IGOWA operator, this would mean that we have numerical arguments to be ordered by linguistic order inducing variables. Note that the reordering of the linguistic inducing variables can be ascendant or descendant according to the interests of the aggregation. Also note that in some situations it is possible to use the implicit lexicographic ordering associated with words such as the ordering of words in dictionaries [29].

The IGOWA operator is a generalization of the IOWA operator. Therefore, the IGOWA operator is applicable to different situations already discussed for the IOWA operator. For example, we could use it for modeling nearest neighbour rule [29], for model building [29] and for the aggregation of complex objects [23]. Other potential applications could be developed for decision making, group decision making, business decisions, etc.

5 Families of IGOWA operators

By choosing a different manifestation of the weighting vector in the IGOWA operator, we are able to obtain different types of aggregation operators. For example, we can obtain the maximum, the minimum, the generalized mean, the weighted generalized mean and the GOWA operator. Note that these results can be obtained both for the DIGOWA and the AIGOWA operators.

The maximum is obtained if $w_p = 1$ and $w_j = 0$, for all $j \neq p$, and $u_p = \text{Max}\{a_i\}$, then, $IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) = \text{Max}\{a_i\}$. The minimum is obtained if $w_p = 1$ and $w_j = 0$, for all $j \neq p$, and $u_p = \text{Min}\{a_i\}$, then, $IGOWA(\langle u_1,a_1\rangle, \langle u_2,a_2\rangle\ldots, \langle u_n,a_n\rangle) = \text{Min}\{a_i\}$. More generally, if $w_k = 1$ and $w_j = 0$, for all $j \neq k$, we get for any λ, $IGOWA(a_1, a_2,\ldots, a_n) = b_k$, where b_k is the the a_i value of the IGOWA pair $\langle u_i,a_i\rangle$ having the kth largest u_i. The generalized mean is found when $w_j = 1/n$, for all a_i. The weighted generalized mean is obtained if $u_i > u_{i+1}$, for all i, and the GOWA operator is obtained if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i.

Other families of IGOWA operators could be obtained by using a different weighting vector. For example, when $w_j = 1/m$ for $k \leq j \leq k + m - 1$ and $w_j = 0$ for $j > k + m$ and $j < k$, we are using the window-IGOWA operator that it is based on the window-OWA operator [19]. Note that k and m must be positive integers such that $k + m - 1 \leq n$. Also note that if $m = k = 1$, and the initial position of the

highest u_i is also the initial position of the highest a_i, then, the window-IGOWA is transformed in the maximum. If $m = 1$, $k = n$, and the initial position of the lowest u_i is also the initial position of the lowest a_i, then, the window-IGOWA becomes the minimum. And if $m = n$ and $k = 1$, the window-IGOWA is transformed in the generalized mean.

If $w_1 = w_n = 0$, and for all others $w_j = 1/(n - 2)$, we are using the olympic induced generalized average that it is based on the olympic average [21]. Note that if $n = 3$ or $n = 4$, the olympic induced generalized average is transformed in the IGOWA median and if $m = n - 2$ and $k = 2$, the window-IGOWA is transformed in the olympic induced generalized average. Also note that the olympic induced generalized average is transformed in the olympic generalized average if $w_p = w_q = 0$, such that $u_p = \text{Max}_i\{a_i\}$ and $u_q = \text{Min}_i\{a_i\}$, and for all others $w_j = 1/(n - 2)$.

Another type of aggregation that could be used is the E-Z IGOWA weights that it is based on the E-Z OWA weights [24]. In this case, we should distinguish between two classes. In the first class, we assign $w_j = (1/k)$ for $j = 1$ to k and $w_j = 0$ for $j > k$, and in the second class, we assign $w_j = 0$ for $j = 1$ to $n - k$ and $w_j = (1/k)$ for $j = n - k + 1$ to n. Note that the E-Z IGOWA weights becomes the E-Z GOWA weights for the first class if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i, from $j = 1$ to k. And for the second class, the E-Z IGOWA weights becomes the E-Z GOWA weights if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i, from $j = n - k + 1$ to n.

We note that the generalized median and the weighted generalized median can also be used as induced aggregation operators. For the IGOWA median, if n is odd we assign $w_{(n+1)/2} = 1$ and $w_j = 0$ for all others, and this affects the argument a_i with the $[(n + 1)/2]$th largest u_i. If n is even we assign for example, $w_{n/2} = w_{(n/2)+1} = 0.5$, and this affects the arguments with the $(n/2)$th and $[(n/2) + 1]$th largest u_i. For the weighted IGOWA median, we follow a different procedure than [20]. We select the argument a_i that has the kth largest inducing variable u_i, such that the sum of the weights from 1 to k is equal or higher than 0.5 and the sum of the weights from 1 to $k - 1$ is less than 0.5. Note that if the ordered position of u_i is the same than the ordered position of b_j such that b_j is the jth largest of a_i, then, the IGOWA median and the weighted IGOWA median become the GOWA median and the weighted GOWA median respectively.

Another interesting family is the S-IGOWA operator based on the S-OWA operator [28]. It can be divided in three classes, the "orlike", the "andlike" and the generalized S-IGOWA operator. The "orlike" S-IGOWA operator is found when $w_p = (1/n)(1 - \alpha) + \alpha$, $u_p = \text{Max}\{a_i\}$, and $w_j = (1/n)(1 - \alpha)$ for all $j \neq p$ with $\alpha \in [0, 1]$. Note that if $\alpha = 0$, we get the arithmetic mean and if $\alpha = 1$, we get the maximum. The "andlike" S-IGOWA operator is found when $w_q = (1/n)(1 - \beta) + \beta$, $u_q = \text{Min}\{a_i\}$, and $w_j = (1/n)(1 - \beta)$ for all $j \neq q$ with $\beta \in [0, 1]$. Note that in this class, if $\beta = 0$ we get the average and if $\beta = 1$, we get the minimum. Finally, the generalized S-IGOWA operator is obtained when $w_p = (1/n)(1 - (\alpha + \beta) + \alpha$, with $u_p = \text{Max}\{a_i\}$; $w_q = (1/n)(1 - (\alpha + \beta) + \beta$, with $u_q = \text{Min}\{a_i\}$; and $w_j = (1/n)(1 - (\alpha + \beta)$ for all $j \neq p,q$ where $\alpha, \beta \in [0, 1]$ and $\alpha + \beta \leq 1$. Note that if $\alpha = 0$, the generalized S-IGOWA operator becomes the "andlike" S-IGOWA operator and if $\beta = 0$, it becomes the "orlike" S-IGOWA operator.

Other families of IGOWA operators could be developed such as the weights that depend on the aggregated objects [19]. For example, we could develop the BADD-IGOWA operator that it is based on the OWA version developed in [19,28].

$$w_j = \frac{b_j^{\alpha}}{\sum_{j=1}^{n} b_j^{\alpha}} \tag{9}$$

where $\alpha \in (-\infty, \infty)$, b_j is the jth largest element of the arguments a_i. Note that the sum of the weights is 1 and $w_j \in [0,1]$. Also note that if $\alpha = 0$, we get the generalized mean and if $\alpha = \infty$, we get the maximum. Another family of IGOWA operator that depends on the aggregated objects is

$$w_j = \frac{(1 - b_j)^{\alpha}}{\sum_{j=1}^{n} (1 - b_j)^{\alpha}} \tag{10}$$

where $\alpha \in (-\infty, \infty)$, b_j is the jth largest element of the arguments a_i. Note that in this case if $\alpha = 0$, we also get the generalized mean and if $\alpha = \infty$, we get the minimum. A third family of IGOWA operator that depends on the aggregated objects is

$$w_j = \frac{(1/b_j)^{\alpha}}{\sum_{j=1}^{n} (1/b_j)^{\alpha}} \tag{11}$$

where $\alpha \in (-\infty, \infty)$, b_j is the jth largest element of the arguments a_i. In this case, we also get the generalized mean if $\alpha = 0$. If $\alpha = 1$, we obtain the harmonic mean and if $\alpha = \infty$, we get the minimum.

A very useful approach for obtaining the weights that it is also applicable for the IGOWA operator is the functional method introduced by Yager [21] for the OWA operator. It can be summarized as follows. Let f be a function f: [0, 1] $\rightarrow$ [0, 1] such that $f(0) = f(1)$ and $f(x) \geq f(y)$ for $x > y$. We call this function a basic unit interval monotonic function (BUM). Using this BUM function we obtain the IGOWA weights w_j for $j = 1$ to n as

$$w_j = f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right) \tag{12}$$

It can easily be shown that using this method, the w satisfy that the sum of the weights is 1 and $w_j \in [0,1]$.

Another family of aggregation operators that could be used in the IGOWA operator is the centered IGOWA weights. This type of operator has been suggested by Yager [27] for the OWA operator. Following the same methodology, we could say that an IGOWA operator is a centered aggregation operator if it is symmetric, strongly decaying and inclusive. It is symmetric if $w_j = w_{j+n-1}$. It is strongly decaying when $i < j \leq (n + 1)/2$ then $w_i < w_j$ and when $i > j \geq (n + 1)/2$ then $w_i < w_j$. It is inclusive if $w_j > 0$. Note that it is possible to consider a softening of the second condition by using $w_i \leq w_j$ instead of $w_i < w_j$. We shall refer to this as softly decaying centered IGOWA operator. Note that the generalized mean is an example of this particular case of centered IGOWA operator. Another particular situation of the centered IGOWA operator appears if we remove the third condition. We shall refer to it as a non-inclusive centered IGOWA operator. For this situation, we find the IGOWA median as a particular case.

A special type of centered IGOWA operator is the Gaussian IGOWA weights which follows the same methodology than the Gaussian OWA weights suggested by Xu [14]. In order to define it, we have to consider a Gaussian distribution $\eta(\mu, \sigma)$ where

$$\mu_n = \frac{1}{n}\sum_{j=1}^{n} j = \frac{n+1}{2} \tag{13}$$

$$\sigma_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(j-\mu_n)^2} \tag{14}$$

Assuming that

$$\eta(j) = \frac{1}{\sqrt{2\Pi\sigma_n}} e^{-(j-\mu_n)^2/2\sigma_n^2} \tag{15}$$

we define the IGOWA weights as

$$w_j = \frac{\eta_j}{\sum_{j=}^{n}\eta(j)} = \frac{e^{-(j-\mu_n)^2/2\sigma_n^2}}{\sum_{j=1}^{n} e^{-(j-\mu_n)^2/2\sigma_n^2}} \tag{16}$$

Note that the sum of the weights is 1 and $w_j \in [0,1]$.

If we analyze different values of the parameter λ in the IGOWA operator, we obtain another group of particular cases such as the usual IOWA operator, the induced OWG (IOWG) operator [5,16], the induced OWHA (IOWHA) operator and the induced OWQA (IOWQA) operator.

When $\lambda = 1$, the IGOWA operator becomes the IOWA operator.

$$IGOWA(\langle u_1,a_1\rangle, \ldots, \langle u_n,a_n\rangle) = \sum_{j=1}^{n} w_j b_j \tag{17}$$

From a generalized perspective of the reordering step we have to distinguish between the DIOWA operator and the AIOWA operator. In both cases, the formulation is the same with the difference that the DIOWA operator has a descending order and the AIOWA operators an ascending order.

When $\lambda = 0$, the IGOWA operator becomes the IOWG operator.

$$IGOWA(\langle u_1,a_1\rangle, \ldots, \langle u_n,a_n\rangle) = \prod_{j=1}^{n} b_j^{w_j} \tag{18}$$

With the DIGOWA operator we obtain the descending IOWG (DIOWG) operator and with the AIGOWA operator, the ascending IOWG (AIOWG) operator.

When $\lambda = -1$, we get the IOWHA operator.

$$IGOWA(\langle u_1,a_1\rangle, \ldots, \langle u_n,a_n\rangle) = \frac{1}{\sum_{j=1}^{n}\frac{w_j}{b_j}} \tag{19}$$

Note that from a generalized perspective of the reordering step we get the descending IOWHA (DIOWHA) operator and the ascending IOWHA (AIOWHA) operator.

When $\lambda = 2$, we get the IOWQA operator.

$$IGOWA(\langle u_1,a_1\rangle, \ldots, \langle u_n,a_n\rangle) = \left(\sum_{j=1}^{n} w_j b_j^2\right)^{1/2} \tag{20}$$

In this case, we obtain the descending IOWQA (DIOWQA) operator and the ascending IOWQA (AIOWQA) operator.

6 Induced Quasi-OWA operators

As it is explained in [1], a further generalization to the GOWA operator is possible by using quasi-arithmetic means [9-10,12]. Note that from a generalized perspective of the reordering step we have to distinguish between ascending and descending orderings in the Quasi-OWA operators [8]. Then, we will get the Quasi-DOWA and the Quasi-AOWA operators.

Following the same methodology as in [1], we can suggest a similar generalization to the IGOWA operator by using quasi-arithmetic means. We can call this generalization as the Quasi-IOWA operator.

$$QIOWA(\langle u_1,a_1\rangle,\ldots,\langle u_n,a_n\rangle) = g^{-1}\left(\sum_{j=1}^{n} w_j\, g\left(b_{(j)}\right)\right) \quad (21)$$

As we can see, we replace b^{λ} with a general continuous strictly monotone function *g(b)*. In this case, we have also to distinguish between ascending and descending orderings. Then, we will have to distinguish between the Quasi-DIOWA operator and the Quasi-AIOWA operator. Obviously, all the particular cases commented in the IGOWA operator are also included in this generalization such as the usual IOWA operator, the IOWG operator, the IOWHA operator and the IOWQA operator.

Another interesting issue to analyse with the Quasi-IOWA operator is the problem of ties in the aggregation. In this case, we also recommend the policy suggested by Yager and Filev [29] where they replace each argument of the tied IOWA pairs by their average. The difference for the Quasi-IOWA operator is that instead of using the arithmetic mean, we recommend to use the quasi-arithmetic mean. Then, depending on the particular type of Quasi-IOWA operator, we will replace each tied argument by its corresponding quasi-arithmetic mean.

We should also note that for the Quasi-IOWA operator, when studying the order inducing variable, it is also possible to use values drawn from a space with the only requirement of having a linear ordering. Then, it is possible to use different kinds of attributes for the order inducing variables that permit us, for example, to mix numbers with words in the aggregations.

7 Conclusions

We have studied different types of aggregation operators. First, we have reviewed some basic operators such as the OWA operator, the IOWA operator and the generalized mean. Then, we have analysed the GOWA operator introduced by Yager. We have considered its main properties and some of the particular cases found in the weighting vector *W* and in the parameter λ. The whole analysis has been developed distinguishing between aggregations with a descending or with an ascending order. With this information, we have introduced the IGOWA operator. It represents an extension to the GOWA operator with the main difference that the reordering step is developed with order inducing variables. We have briefly studied its main properties and some of the particular cases found in the weighting vector *W* and in the parameter λ such as the IOWA operator, the IOWG operator, the IOWHA operator and the IOWQA operator. Finally, we have developed the Quasi-IOWA operator which represents a generalization of the IGOWA operator.

References

[1] G. Beliakov, Learning Weights in the Generalized OWA Operators, Fuzzy Optimization and Decision Making 4 (2005) 119-130.

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

[3] F. Chiclana, F. Herrera, E. Herrera-Viedma, The ordered weighted geometric operator: Properties and application, in: Proc. 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Madrid, Spain, 2000, pp. 985-991.

[4] F. Chiclana, F. Herrera, E. Herrera-Viedma, Integrating multiplicative preference relations in a multipurpose decision-making model based on fuzzy preference relations, Fuzzy Sets and Systems 122 (2001) 277-291.

[5] F. Chiclana, F. Herrera, E. Herrera-Viedma, S. Alonso, Induced ordered weighted geometric operators and their use in the aggregation of multiplicative preference relations, International Journal of Intelligent Systems 19 (2004) 233-255.

[6] J. Dujmovic, Weighted conjunctive and disjunctive means and their application in system evaluation, Publikacije Elektro-technickog Faculteta Beograd, Serija Mate-matika i Fizika, No. 483, pp. 147-158, 1974.

[7] H. Dyckhoff, W. Pedrycz, Generalized means as a model of compensative connectives, Fuzzy Sets and Systems 14 (1984) 143-154.
[8] J. Fodor, J.L. Marichal, M. Roubens, Characterization of the ordered weighted averaging operators, IEEE Transactions on Fuzzy Systems 3 (1995) 236-240.
[9] G.H. Hardy, J.E. Littlewood, G. Pólya, Inequalities, Cambridge University Press, Cambridge, 1934.
[10] A.N. Kolmogoroff, Sur la notion de la moyenne, Atti Della Reale Accademia Nazionale Dei Lincei, 12 (1930) 388-391.
[11] J.M. Merigó, M. Casanovas, Induced and uncertain heavy OWA operators, in: Proc. 13th Congress of International Association for Fuzzy-Set Management and Economy (SIGEF), Hammamet, Tunisia, 2006, pp. 728-746.
[12] M. Nagumo, Uber eine klasse der mittelwerte, Japanese Journal of Mathematics 6 (1930) 71-79.
[13] P.A. Schaefer, H.B. Mitchell, A generalized OWA operator, International Journal of Intelligent Systems 14 (1999) 123-143.
[14] Z.S. Xu, An Overview of Methods for Determining OWA Weights, International Journal of Intelligent Systems 20 (2005) 843-865.
[15] Z.S. Xu, Q.L. Da, The Ordered Weighted Geometric Averaging Operators, International Journal of Intelligent Systems 17 (2002) 709-716.
[16] Z.S. Xu, Q.L. Da, An Overview of Operators for Aggregating Information, International Journal of Intelligent Systems 18 (2003) 953-969.
[17] R.R. Yager, On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making, IEEE Transactions on Systems, Man and Cybernetics B 18 (1988) 183-190.
[18] R.R. Yager, On generalized measures of realization in uncertain environments, Theory and Decision, 33, 1992, pp. 41-69.
[19] R.R. Yager, Families of OWA operators, Fuzzy Sets and Systems 59 (1993) 125-148.
[20] R.R. Yager, On weighted median aggregation, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2 (1994) 101-113.
[21] R.R. Yager, Quantifier Guided Aggregation Using OWA Operators, International Journal of Intelligent Systems 11 (1996) 49-73.
[22] R.R. Yager, The induced fuzzy integral aggregation operator, International Journal of Intelligent Systems 17 (2002) 1049-1065.
[23] R.R. Yager, Induced aggregation operators, Fuzzy Sets and Systems 137 (2003) 59-69.
[24] R.R. Yager, E-Z OWA weights, in: Proc. 10th International Fuzzy Systems Association (IFSA) World Congress, Istanbul, Turkey, 2003, pp. 39-42.
[25] R.R. Yager, Choquet aggregation using order inducing variables, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (2004) 69-88.
[26] R.R. Yager, Generalized OWA Aggregation Operators, Fuzzy Optimization and Decision Making 3 (2004) 93-107.
[27] R.R. Yager, Centered OWA operators, Soft Computing 11 (2007) 631-639.
[28] R.R. Yager, D.P. Filev, Parameterized “andlike” and “orlike” OWA operators. International Journal of General Systems 22 (1994) 297-316.
[29] R.R. Yager, D.P. Filev, Induced Ordered Weighted Averaging Operators, IEEE Transactions on Systems, Man and Cybernetics B 29 (1999) 141-150.
[30] R.R. Yager, J. Kacprzyk, The Ordered Weighted Averaging Operators: Theory and Applications, Kluwer Academic Publishers, Norwell, MA, 1997.
[31] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Transactions on Fuzzy Systems 4 (1996) 103-111.

Mixture operators and the additional insurance

Jana Špirková
Department of Quantitative Methods
and Information Technology
Faculty of Economics, Matej Bel University
Tajovského 10, 975 90 Banská Bystrica
Slovakia
jana.spirkova@umb.sk

Pavol Kráľ
Department of Quantitative Methods
and Information Technology
Faculty of Economics, Matej Bel University,
Tajovského 10, 975 90 Banská Bystrica
Slovakia
pavol.kral@umb.sk

Abstract

In our paper we recall some known mixture operators and their properties. We introduce sufficient conditions for weighting functions to ensure the monotonicity of the corresponding operators. We introduce new applications of mixture operators in the insurance, namely in the additional insurance.

Keywords: Mixture operator, Generalized mixture operator, Weighting function, Net annual premium.

1 Introduction

In multicriteria decision making, alternatives are characterized by score vectors describing the degree of fulfilment of chosen criteria. We suppose that each alternative $\mathbf{x}$ is characterized by a score vector $\mathbf{x} = (x_1, \ldots, x_n) \in [a, b]^n$, where $n \in N - \{1\}$ is the number of applied criteria and the decision is based on an aggregation operator $A : [a, b]^n \to [a, b]$. Commonly used anonymous (i. e., stable under permutations of score) aggregation operator is the arithmetic mean M,

$$M(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Standard incorporation of (fixed) weights (criteria importances) leads to the class of weighted arithmetic means. In this case, weights are assigned to single criteria (i. e., coordinates of score vector), independently of the actual score vector. Alternative approaches of introducing weights to the arithmetic mean aggregation link the weights and single observed score values. Arithmetic mean weighted by a weighting function $g : [a, b] \to]0, \infty[$ (also called a mixture operator), $M_g : [a, b]^n \to [a, b]$ is given by

$$M_g(x_1, \ldots, x_n) = \frac{\sum_{i=1}^{n} g(x_i) \cdot x_i}{\sum_{i=1}^{n} g(x_i)}. \tag{1}$$

Observe that due to the continuity of weighting function g, each mixture operator M_g is continuous. Moreover, each mixture operator is anonymous and idempotent(unanymous). More informations about mixture operators can be found in [1], [3], [4].

Note that sometimes different weighting functions are applied for different criteria score, thus leading to a generalized mixture operator [4], $M_{\mathbf{g}} : [a, b]^n \to [a, b]$ given by

$$M_{\mathbf{g}}(x_1, \ldots, x_n) = \frac{\sum_{i=1}^{n} g_i(x_i) \cdot x_i}{\sum_{i=1}^{n} g_i(x_i)}, \tag{2}$$

where $\mathbf{g} = (g_1, \ldots, g_n)$ is a vector of weighting functions. Obviously generalized mixture operators are continuous.

Generalized mixture operator based on the ordinal approach $M'_{\mathbf{g}} : [a, b]^n \to [a, b]$ given by

$$M'_{\mathbf{g}}(x_{(1)}, \ldots, x_{(n)}) = \frac{\sum_{i=1}^{n} g_i(x_{(i)}) \cdot x_{(i)}}{\sum_{i=1}^{n} g_i(x_{(i)})}, \tag{3}$$

where $\mathbf{g} = (g_1, \ldots, g_n)$ is a vector of weighting functions and $(x_{(1)}, \ldots, x_{(n)})$ is non - decreasing permutation of score vector, is called ordered generalized mixture operator. Ordered generalized mixture operator is generalization of OWA operators. More informations about OWA operators can be found in [9].

In our paper we recall the problem of the monotonicity of mixture operators on the interval $[a, b]$ and specially

on the interval $[0,1]$ we introduce new applications of mixture operators in additional insurance on price determination of the gross annual premium.

The paper is organized as follows. In the second section we introduce sufficient conditions for the non - decreasingness of mixture operators, namely for non - decreasing weighting functions and for non - increasing weighting functions. In the third section we introduce new applications of mixture operators and we calculate the gross annual premium for the additional insurance, specifically for time of necessary treatment, permanent consequences, permanent disability (full disability) and accidental death. In our paper we focuse only on the applications of mixture operators with non - decreasing weighting functions. Finally, some conclusions are given.

2 Monotonicity of mixture operators

2.1 Mixture operator with non - decreasing weighting function

In [4], [5] were derived the sufficient conditions for the monotonicity of mixture operators

$$M_g(x_1,\ldots,x_n)=\frac{\sum\limits_{i=1}^{n} g(x_i)\cdot x_i}{\sum\limits_{i=1}^{n} g(x_i)}$$

on the interval $[0,1]$. In [6] were derived the sufficient conditions for the monotonicity of mixture operators on the interval $[a,b]$. In the next we recall these conditions, which are written in next propositions.

Proposition 2.1. *Let $g:[a,b]\to]0,\infty[$ be a non - decreasing (smooth) weighting function, which satisfies the next condition:*

$$g(x)\geq g'(x)\cdot(b-a) \qquad (C1)$$

for all $x\in[a,b]$.
Then $M_g:[a,b]^n\to[a,b]$ is an aggregation operator for each $n\in N$, $n>1$.

Proposition 2.2. *Let $g:[a,b]\to]0,\infty[$ be a non - decreasing (smooth) weighting function, which satisfies the next condition:*

$$g(x)\geq g'(x)\cdot(b-x) \qquad (C2)$$

for all $x\in[a,b]$.
Then $M_g:[a,b]^n\to[a,b]$ is an aggregation operator for all $n\in N$, $n>1$.

Moreover, the conditions $(C1)$, $(C2)$ are enough to ensure the monotonicity of generalized mixture operators $M_{\mathbf{g}}$, as well as of ordered generalized mixture operators $M'_{\mathbf{g}}$ acting on $[a,b]$.

Proposition 2.3. *Let $g:[a,b]\to]0,\infty[$ be a non - decreasing (smooth) weighting function such that for a fixed $n\in N$, $n>1$, it satisfies the next condition:*

$$\frac{g^2(x)}{(n-1)\cdot g(b)}+g(x)\geq g'(x)\cdot(b-x) \qquad (C3)$$

for all $x\in[a,b]$.
Then $M_g:[a,b]^n\to[a,b]$ is an aggregation operator.

Proposition 2.4. *For a fixed $n\in N$, $n>1$, $i=1,\ldots,n$, let $g_i:[a,b]\to]0,\infty[$ be non - decreasing (smooth) weighting functions, which satisfy the next condition:*

$$\frac{g_i^2(x)}{\sum\limits_{j\neq i} g_j(b)}+g_i(x)\geq g_i'(x)\cdot(b-x) \qquad (C4)$$

for all $x\in[a,b]$. Then $M_{\mathbf{g}}:[a,b]^n\to[a,b]$, where $\mathbf{g}=(g_1,\ldots,g_n)$, is an aggregation operator.

Previous conditions are rewrite for the most frequent interval $[0,1]$ by the next way:

$$g(x)\geq g'(x), x\in[0,1]; \qquad (4)$$

$$g(x)\geq g'(x)\cdot(1-x), x\in[0,1]; \qquad (5)$$

$$\frac{g^2(x)}{(n-1)\cdot g(1)}+g(x)\geq g'(x)\cdot(1-x)\,, x\in[0,1]; \qquad (6)$$

$$\frac{g_i^2(x)}{\sum\limits_{j\neq i} g_j(1)}+g_i(x)\geq g_i'(x)\cdot(1-x), x\in[0,1]. \qquad (7)$$

In the next we introduce examples of weighting functions for the non - decreasing mixture operators on the interval $[0,b]$.

Example 2.1. Let $g(x):[0,b]\to]0,\infty[$ be a function given by $g(x)=ax+1-ab$, where $a\geq 0$ and $g(b)=1$. From the condition $(C1)$ and $(C2)$ we get the condition for the coefficient a in our linear function on the interval $[0,b]$, namely $a\leq\frac{1}{2b}$. For example :

- if $b=0.5$ $\quad a\in[0,1]$,
- if $b=1$ $\quad a\in[0,0.5]$,
- if $b=2$ $\quad a\in[0,0.25]$,
- if $b=3$ $\quad a\in[0,0.166667]$,
- if $b\to\infty$, then $a\to 0$, so $g(x)=1$.

Note that in the above case the second sufficient condition does not extend the choice of weighting functions. However the condition $(C3)$ for $n=2$ give us the condition: $a\in[0,\frac{2-\sqrt{2}}{b}]$, which extends the choice of weighting functions. For example:

- if $b = 0.5$ $a \in [0, 1.17157]$,
- if $b = 1$ $a \in [0, 0.585786]$,
- if $b = 2$ $a \in [0, 0.29289]$,
- if $b = 3$ $a \in [0, 0.195262]$,
- if $b \to \infty$ $a \to 0$, so $g(x) = 1$.

Observe that with greater b the choice of coefficients a is smaller.
Next example is interesting because of the fact that each of conditions $(C1)$, $(C2)$ and $(C3)$ give us a different choice of $a^{'}$s. Again observe that with greater b the choice of coefficients a is smaller.

Example 2.2. Let $g(x) : [0, b] \to]0, \infty[$ be given by $g(x) = ax^2 + 1 - ab^2$, where $a > 0$ and $b > 0$. From the condition $(C1)$ we get condition $a \in [0, \frac{1}{2b^2}]$. For example:

- if $b = 0.5$ $a \in [0, 2]$,
- if $b = 1$ $a \in [0, 0.5]$,
- if $b = 2$ $a \in [0, 0.125]$,
- if $b = 3$ $c \in [0, 0.055556]$,
- if $b \to \infty$, then $a \to 0$, so $g(x) = 1$.

The condition $(C2)$ give us $a \in [0, \frac{3}{4b^2}]$. For example:

- if $b = 0.5$ $a \in [0, 3]$,
- if $b = 1$ $a \in [0, 0.75]$,
- if $b = 2$ $a \in [0, 1875]$,
- if $b = 3$ $c \in [0, 0.083333]$,
- if $b \to \infty$ $a \to 0$, so $g(x) = 1$.

From the last condition $(C3)$ we have the condition:

$$a^2x^4 + (2a - 2a^2b^2 + 3a) \cdot x^2 - 2abx + a^2b^4 - 3ab^2 + 2 \geq 0.$$

By system MATHEMATICA we get:

- $b = 0.5$ $a \in [0; 3.220]$
- $b = 1$ $a \in [0; 0.805]$
- $b = 2$ $a \in [0; 0.201]$
- $b = 3$ $a \in [0; 0.0894]$
- $b \to \infty$ $a \to 0$, so $g(x) = 1$.

2.2 Mixture operator with non - increasing weighting function

In the next we recall conditions for the non - decreasingness of mixture operators with non - increasing weighting function on the interval $[a, b]$, which are derived in [7].

Proposition 2.5. *Let $g : [a, b] \to]0, \infty[$ be a non - increasing (smooth) weighting function, which satisfies the next condition:*

$$g(x) + g'(x) \cdot (b - a) \geq 0 \qquad (P1)$$

for all $x \in [a, b]$.
Then $M_g : [a, b]^n \to [a, b]$ is an aggregation operator for each $n \in N$, $n > 1$.

Proposition 2.6. *Let $g : [a, b] \to]0, \infty[$ be a non - increasing (smooth) weighting function, which satisfies the next condition:*

$$g(x) + g'(x) \cdot (x - a) \geq 0 \qquad (P2)$$

for all $x \in [a, b]$.
Then $M_g : [a, b]^n \to [a, b]$ is an aggregation operator for each $n \in N$, $n > 1$.

Proposition 2.7. *For a fixed $n \in N$, $n > 1$, let $g : [a, b] \to]0, \infty[$ be a non - increasing (smooth) weighting function, which satisfies the next condition:*

$$\frac{g^2(x)}{(n-1) \cdot g(a)} + g(x) + g'(x) \cdot (x - a) \geq 0 \qquad (P3)$$

for all $x \in [a, b]$.
Then $M_g : [a, b]^n \to [a, b]$ is an aggregation operator.

Proposition 2.8. *For a fixed $n \in N$, $n > 1$, $i = 1, \ldots, n$, let $g_i : [a, b] \to]0, \infty[$ be a non - increasing (smooth) weighting functions, which satisfy the next condition:*

$$\frac{g_i^2(x)}{\sum\limits_{j \neq i} g_j(a)} + g_i(x) + g_i'(x) \cdot (x - a) \geq 0 \qquad (P4)$$

for all $x \in [a, b]$.

Then $M_{\mathbf{g}} : [a, b]^n \to [a, b]$, where $\mathbf{g} = (g_1, \ldots, g_n)$, is an aggregation operator.

The conditions $(P1)$, $(P2)$, $(P3)$ and $(P4)$ are rewrite for the most frequent interval $[0, 1]$ by:

$$g(x) + g'(x) \geq 0, x \in [0, 1]; \qquad (8)$$

$$g(x) + g'(x) \cdot (x) \geq 0, x \in [0, 1]; \qquad (9)$$

$$\frac{g^2(x)}{(n-1) \cdot g(0)} + g(x) + g'(x) \cdot x \geq 0, x \in [0, 1]; \qquad (10)$$

$$\frac{g_i^2(x)}{\sum\limits_{j \neq i} g_j(0)} + g_i(x) + g_i'(x) \cdot x \geq 0, x \in [0, 1]. \qquad (11)$$

We introduce the example of the linear weighting function on the interval $[0, b]$.

Example 2.3. Let $g(x) : [0, b] \to]0, \infty[$ be given by $g(x) = ax + 1 - ab$, where $a \leq 0$, $g(b) = 1$. The condition $(P1)$ implies that

$$ax + 1 - ab + ab \geq 0,$$

whence $a \geq -\frac{1}{b}$. From the last inequality we see, that the set of coefficients a depends on the boundary conditions. For example:

- if $b = \frac{1}{3}$, then $a \in [-3, 0]$
- if $b = \frac{1}{2}$, then $a \in [-2, 0]$
- if $b = 1$, then $a \in [-1, 0]$
- if $b = 2$, then $a \in [-0.5, 0]$
- if $b = 3$, then $a \in [-0.333333, 0]$
- if $b \to \infty$, then $a \to 0$, so $g(x) = 1$.

Observe that with greater b, the set of coefficients a is smaller.

From the condition $(P2)$ we get inequality

$$ax + 1 - ab + ax \geq 0.$$

The function on the left side of this inequality is decreasing, so this function attains minimum for $x = b$. After some processing we get the condition $a \geq -\frac{1}{b}$, what is the same result as from condition $(P1)$.

The third condition $(P3)$ for $n = 2$ implies that

$$\frac{(ax + 1 - ab)^2}{1 - ab} + ax + 1 - ab + ax \geq 0.$$

After some processing we get

$$a^2x^2 - (4a^2b - 4a) \cdot x + 2a^2b^2 - 4ab + 2 \geq 0.$$

The function on the left side of the above inequality is decreasing on the interval $[0, b]$, so for x in the inequality we give $x = b$ and we get condition $a^2b^2 - 2 \leq 0$, whence $|a| \leq \frac{\sqrt{2}}{b}$. This condition give us the set of coefficients $a \in [-\frac{\sqrt{2}}{b}, 0]$.
That means, with longer input interval is smaller the choise of weighting linear functions. For example for different b we get next sets of coefficients a:

- if $b = \frac{1}{3}$, then $a \in [-4.242640; 0]$
- if $b = \frac{1}{2}$, then $a \in [-2.828427; 0]$
- if $b = 1$, then $a \in [-1.414213; 0]$
- if $b = 2$, then $a \in [-0.707106; 0]$
- if $b = 3$, then $a \in [-0.471404; 0]$
- if $b \to \infty$, then $a \to 0$, so $g(x) = 1$.

3 Applications of mixture operators in the additional insurance

The personal motivation to become the assured in a specific insurance company is influenced by many factors. One of them is the price of an insurance policy. Of course the customers of insurance companies prefer to have the least possible price. One the other hand the actuaries should include to theirs considerations also the possible risks for the insurance company. In practice the previous problem is solved in the following way. The insurance policies are decomposed into tariff groups which are as homogeneous as possible. Then there are computed global characteristics for each group using statistical data. The global characteristics are used to evaluate the gross premium. The gross premium includes for example the security charges. From the global characteristics we can compute many indicators which are used to compute the net current premium. The problem is for example that the statistical data describes often only the current year and many indicators are influenced by interest rates and so one. So we need to select the indicator carefully. We also should use the approximation of future values of indicators. To avoid the above mentioned problems we prefer to use the Lose Ratio as basic indicator. The changes of this indicator we will describe using the fuzzy sets theory. The main idea is the following. Our indicators include the possible risks, i.e. we compute the indicators using statistical data and the results are modified using and empirical expert evaluation of an actuary. The expert evaluation allows us to modify the weights of risk sources in our consideration and sometimes also it helps us to include qualitative criteria into basic indicators. In the presented paper we focus our attention only to the first problem, i.e. to the modification of weights using an expert evaluation. It is known that there exist many methods how to incorporate our opinion into the final decision. One of them is Saaty's method (see [8]). Our approach represents an alternatively look on the problem of optimal selection of weights. The previous general considerations are described in the examples in the rest of the paper.

Actuaries offer to life insurance also the additional insurance. Statistic data in general enable us to calculate for individual years and individual tariff groups the next specifications of global character:

- the number of insurance policy in the given year,
- the number of reportable event in the given year,
- the total sum insured in the given year,
- the total premium in the given year,

- the total benefits in the given year.

For more information see [2] and [9]. From the previous global data we can calculate for individual years and individual tariff groups different ratios. In our paper we focus on loss ratio(next LR), because this ratio is independent from the inflation.

Let us assume, that we have next empirical data for one risk group: for the time of necessary treatment (below TNT) we know, the total insured sum in the specific year was 175000000 SKK and involved total benefits was 2070700 SKK. For the permanent consequences (below PC) the total insured sum was 150000000 SKK in the specific year and involved total benefits was 1350000 SKK, for the permanent disability (below PD) the total insured sum in the specific year was 100000000 SKK and involved total benefits was 870000 SKK and for accidental death (below AD) the total insured sum in the specific year was 290000000 SKK and involved total benefits was 1500000 SKK. On the basis [2] loss ratio is given by

$$LR = \frac{total\ \ benefits}{total\ \ insured\ \ sum}.$$

On the basis of previous empirical data we determine for the time of necessary treatment $LR = x_1 = 0.011833$, for the permanent consequences $LR = x_2 = 0.009$, for the permanent disability $LR = x_3 = 0.0087$ and for the accidental death $LR = x_4 = 0.005172$. These values we can put as the basis of the net annual premiums(NAP):

- for AD: $NAP_{AD} = 5.172$ SKK on 1000 SKK of the insured sum,
- for PD: $NAP_{PD} = 8.700$ SKK on 1000 SKK of the insured sum,
- for PC: $NAP_{PC} = 9.000$ SKK on 1000 SKK of the insured sum,
- for TNT: $NAP_{TNT} = 11.833$ SKK on 1000 SKK of the insured sum.

In the next we try to modify these values on gross premium by mixture operators. For separately x_i (see Table 1) we use consequently different weighting functions on the interval $[0, 1]$. For example in Table 1 in first and second rows are $g_1(x_1) = 0.5x_1 + 0.5$, $g_2(x_2) = 0.4x_2 + 0.6$, $g_3(x_3) = 0.1x_3 + 0.9$ and $g_4(x_4) = 0.05x_4 + 0.95$ and we calculate generalized mixture operator $M_{\mathbf{g}} = 0.009199$, (see Table 2). By normed weights we get possible the gross annual premiums on 1000 Sk of the insured sum for separately additional insurance.

Table 1:

1.	$g_1(x_1) = 0.5x_1 + 0.5$	$g_2(x_2) = 0.4x_2 + 0.6$
	$g_3(x_3) = 0.1x_3 + 0.9$	$g_4(x_4) = 0.05x_4 + 0.95$
2.	$g_1(x_1) = 0.5x_1 + 0.5$	$g_2(x_2) = 0.45x_2 + 0.55$
	$g_3(x_3) = 0.2x_3 + 0.8$	$g_4(x_4) = 0.015x_4 + 0.985$
3.	$g_1(x_1) = 0.49x_1 + 0.51$	$g_2(x_2) = 0.4x_2 + 0.6$
	$g_3(x_3) = 0.15x_3 + 0.85$	$g_4(x_4) = 0.01x_4 + 0.99$
4.	$g_1(x_1) = 0.5x_1^2 + 0.5$	$g_2(x_2) = 0.4x_2^2 + 0.6$
	$g_3(x_3) = 0.1x_3^2 + 0.9$	$g_4(x_4) = 0.05x_4^2 + 0.95$
5.	$g_1(x_1) = 0.6x_1^2 + 0.4$	$g_2(x_2) = 0.5x_2^2 + 0.5$
	$g_3(x_3) = 0.2x_3^2 + 0.8$	$g_4(x_4) = 0.001x_4^2 + 0.999$
6.	$g_1(x_1) = 0.75x_1^2 + 0,25$	$g_2(x_2) = 0.7x_2^2 + 0,3$
	$g_3(x_3) = 0.5x_3^2 + 0.5$	$g_4(x_4) = 0.001x_4^2 + 0.999$

Table 2:

	$M_{\mathbf{g}}$		normed weights	gross annual premium
1.	0.009199	**AD**	0.169933	6.25
		PD	0.204047	7.51
		PC	0.304609	11.21
		TNT	0.321411	11.83
2.	0.009247	**AD**	0.176750	6.50
		PD	0.194802	7.17
		PC	0.281979	10.38
		TNT	0.346469	12.75
3.	0.009224	**AD**	0.173301	6.38
		PD	0.204052	7.51
		PC	0.287863	10.59
		TNT	0.334784	12.32
4.	0,009202	**AD**	0.169493	6,24
		PD	0.203396	7.48
		PC	0.305081	11.23
		TNT	0.322030	11.85
5.	0.009426	**AD**	0.148205	5.57
		PD	0.185263	6.97
		PC	0.296404	11.14
		TNT	0.370128	13.92
6.	0.009970	**AD**	0.122014	4.83
		PD	0.146431	5.80
		PC	0.244028	9.66
		TNT	0.487528	19.31

Similarly (see Table 3), we use consequently different weighting functions on the interval $[0, \frac{1}{2}]$. Similarly we calculated further decisions for the gross annual premiums (see Table 4). The availability of the gross annual premiums depends on the value of the mixture operator and on the separately weights. The choice of the optimal gross annual premiums depends on the actuary expert decision.

Table 3:

7.	$g_1(x_1) = 0.9x_1 + 0.55$	$g_2(x_2) = 0.7x_2 + 0.65$
	$g_3(x_3) = 0.5x_3 + 0.75$	$g_4(x_4) = 0.2x_4 + 0.9$
8.	$g_1(x_1) = 0.95x_1 + 0.525$	$g_2(x_2) = 0.8x_2 + 0.6$
	$g_3(x_3) = 0.6x_3 + 0.7$	$g_4(x_4) = 0.1x_4 + 0.95$
9.	$g_1(x_1) = x_1 + 0.5$	$g_2(x_2) = 0.7x_2 + 0.65$
	$g_3(x_3) = 0.4x_3 + 0.8$	$g_4(x_4) = 0.05x_4 + 0.975$
10.	$g_1(x_1) = 3x_1^2 + 0.25$	$g_2(x_2) = 2x_2^2 + 0.5$
	$g_3(x_3) = 1.5x_3^2 + 0.625$	$g_4(x_4) = 0.05x_4^2 + 0.9875$
11.	$g_1(x_1) = 3x_1^2 + 0.25$	$g_2(x_2) = 1.5x_2^2 + 0.625$
	$g_3(x_3) = x_3^2 + 0.75$	$g_4(x_4) = 0.001x_4^2 + 0.99975$
12.	$g_1(x_1) = 2.9x_1^2 + 0.275$	$g_2(x_2) = 2.5x_2^2 + 0.375$
	$g_3(x_3) = 2x_3^2 + 0.5$	$g_4(x_4) = 0.001x_4^2 + 0.99975$

On the specification of the optimal gross annual premium we have more options.

1. By the median of the mixture operators (see Table 5) we choose the optimal gross annual pre-

Table 4:

	$M_{\mathbf{g}}$		normed weights	gross annual premium
7.	0.009082	**AD**	0.193420	7.03
		PD	0.228793	8.31
		PC	0.263111	9.56
		TNT	0,314675	11.43
8.	0.009173	**AD**	0.189698	6.96
		PD	0.217279	7.97
		PC	0.252519	9.27
		TNT	0.340504	12.49
9.	0.009215	**AD**	0.171801	6.33
		PD	0.223125	8.22
		PC	0.273291	10.07
		TNT	0.331783	12.23
10.	0,009715	**AD**	0.105838	4.11
		PD	0.211672	8.23
		PC	0.264561	10.28
		TNT	0.417929	16.24
11.	0.009643	**AD**	0.095268	3.67
		PD	0.238136	9.19
		PC	0.285742	11.02
		TNT	0.380854	14.69
12.	0.009775	**AD**	0.127932	5.00
		PD	0.174492	6.82
		PC	0.232614	9.10
		TNT	0.464961	18.18

mium (GAP). Median of the gross annual premium we get as the average from the second and third global gross annual premiums 36.89 SKK and 37.00 SKK and we have the optimal gross annual premium $GAP = 36.96$ SKK on 1000 SKK of the insured sum. The gross annual premiums for individual additional insurances we calculate consequently as the average from appropriate values, so we get next the gross annual premiums:

- for AD: $GAP_{AD} = 6.47$ SKK on 1000 SKK of the insured sum,
- for PD: $GAP_{PD} = 7.37$ SKK on 1000 SKK of the insured sum,
- for PC: $GAP_{PC} = 10.53$ SKK on 1000 SKK of the insured sum,
- for TNT: $GAP_{TNT} = 12.59$ SKK on 1000 SKK of the insured sum.

2. Other approach is to choice the highest value of our mixture operator as the optimal gross annual premium. For example from sixth option (see Table 1, Table 5) we get gross annual premium $GAP = 39.48$ Sk. The gross annual premiums for individual additional insurances in this case are:

   - for AD: $GAP_{AD} = 4.82$ Sk on 1000 Sk of the insured sum,
   - for PD: $GAP_{PD} = 5.78$ Sk on 1000 Sk of the insured sum,
   - for PC: $GAP_{PC} = 9.63$ Sk on 1000 Sk of the insured sum,
   - for TNT: $GAP_{TNT} = 19.25$ Sk on 1000 Sk of the insured sum.

Typical process of comercial insurance companies is to add some safety additional charge value (about 10%) to the net annual premium. Now it is popular explicit safety additional charge value, which is calculated as specific multiple of standard deviation estimated from the last loss ratios. By our method we get comparable values.

Table 5:

order	$M_{\mathbf{g}}$
7.	0.009082
8.	0.009173
1.	0.009199
4.	0.009202
9.	0.009215
3.	0.009224
2.	0.009247
5.	0.009426
11.	0.009643
10.	0.009715
12.	0.009775
6.	0.009870

4 Conclusions

In our paper the utilization of the mixture operators is discussed, namely the utilization of different weighting functions. We use non - decreasing weighting functions to show different weights for different criteria. For the simplicity we have focused our attention on the non - decreasing weighting functions with coefficients which are specified by conditions $(C1)$ and $(C2)$. Of course, in the insurance practice it is possible to use also other conditions and non - increasing weighting functions. We have discussed especially the basic value of the net annual premiums for additional insurances - the time of necessary treatment, permanent consequences, permanent disability and accidental death. Similarly we can calculate also the net premium in the non - life insurance, for example building and contents insurance, homeowners insurance, professional indemnity insurance, etc. The values for gross premium with respect to different risk groups and different age gradation, and also the complete comparison of our approach with the classical approach, which is described in [2] will be the object of our future research.

Acknowledgement

This work was supported by grant VEGA - 1/ 3006/ 06, VEGA - 1/2002/05 and by grant APVT -20 - 003204.

References

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

[2] Cipra, T.: Insurance mathematics in praktice (in czech), Edition HZ, Praha, 1994.

[3] Pereira, R. A. M., Pasi, G.: On non - monotonic aggregation: Mixture operators. In: *Proceedings of the 4th Meeting of the EURO Working Group on Fuzzy Sets (EUROFUSE'99) and 2nd International Conference on Soft and Intelligent Computing (SIC'99)*, Budapest, Hungary, 1999.

[4] Ribeiro, R. A., Pereira, R. A. M.: Generalized mixture operators using weighting functions: A comparative study with WA and OWA, *European Journal of Operational Research*, volume 145, 2003, 329 - 342.

[5] Špirková, J.: Weighting Functions For Aggregation Operators: *3rd International Summer School On Aggregation Operators*, Universita Della Svizzera Italiana, Lugano, Switzerland, July 10 - 15, 2005, 127 - 130.

[6] Špirková, J.: Mixture and quasi - mixture operators, IPMU 2006, Paris, France, 2006, 603 - 608.

[7] Špirková, J.: Monotonicity of mixture and quasi - mixture operators, *The 13th Zittau Fuzzy Colloquium 2006*, Zittau, September 13 - 15, 2006, 212 - 219.

[8] Talašová, J.: Fuzzy methods in multicriteria decision making (in czech). Univerzity of Palacký in Olomouc, Faculty of Natural Sciences, 2003.

[9] Urbaníková, M., Vaculíková, Ľ.: Actuarial mathematics (in slovak), First Edition , STU Bratislava, 2006.

[10] Yager, R. R.: On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE Trans. Syst., Man Cybern., vol. 18, 1988, 183 - 190.

Using quasiarithmetic means in a sequential decision procedure

Miguel Ángel Ballester
Dep. d'Economia i d'Història Econòmica
Universitat Autònoma de Barcelona, Spain
mballester@pareto.uab.es

José Luis García-Lapresta
Dep. de Economía Aplicada (Matemáticas)
Universidad de Valladolid, Spain
lapresta@eco.uva.es

Abstract

We study in this paper the following sequential decision procedure. First, the members of a group show their opinions on all the members, regarding a specific attribute. Taking into account this information, quasiarithmetic means and a family of thresholds, a subgroup of individuals is selected: those members whose collective assessments (obtained through a quasiarithmetic mean) reach a specific threshold. After that, only the assessments of this qualified subgroup are taken into account emerging a new subgroup with the aggregation phase. We analyze when this recursive procedure converges providing a final subgroup of qualified members.

Keywords: Group decision making, Aggregation operators, Quasiarithmetic means, Qualification.

1 Introduction

Consider a group of experts that has to decide which of its members should participate in a concrete task or constituting a committee. In other fields of research there are works where the problem arises in choosing the members of the society satisfying a social identity (see Kasher and Rubinstein [6]), with respect to a general attribute (see Samet and Schmeidler [10]) or, in a very different context, which elements of a society should be considered to delimitate a social graph (see Laumann, Marsden and Prensky [8]). In Ballester and García-Lapresta [1] a recursive procedure for selecting a final group of individuals is introduced.

We examine in this paper the use of that sequential procedure where in each stage a subgroup is selected, and taking the opinions of this subgroup a new subgroup emerges. This tries to reflect the idea that selected members have more valuable opinions. For doing the selection, an aggregation operator and a threshold are considered over the gradual opinions.

Aggregation operators allow us to generate a collective assessment to each individual of the group taking into account the individual opinions (see Fodor and Roubens [4], Grabisch, Orlovski and Yager [5] and Calvo, Kolesárova, Komorníková and Mesiar [3], among others).

Among the large variety of aggregation operators that we can use for our sequential procedure, the case of OWA ("Ordered Weighted Averaging") operators (Yager [11]) have been widely discussed in Ballester and García-Lapresta [1]. In the present paper, we complement that study by considering a different family of operators that has attracted significant attention in the literature, the class of quasiarithmetic means.

Quasiarithmetic means were characterized by Kolmogoroff [7] and Nagumo [9]. In Bullen, Mitrinović and Vasić [2] there is an exhaustive study on means (chapters IV and VI are devoted to quasiarithmetic means). See Fodor and Roubens [4, 5.5] and Calvo, Kolesárova, Komorníková and Mesiar [3, 4.3] as well.

The paper is organized as follows. In Section 2 we consider again the basic framework for developing the group decision procedure. Section 3 briefly describes the sequential procedure subject of analysis. Section 4 includes some important families of collective evaluation mechanisms suitable for the recursive decision problem. In Section 5 we announce and describe the results of the paper. Finally, Section 6 contains some conclusions.

2 Preliminaries

Consider a finite set of individuals $N = \{1, 2, \ldots, n\}$ with $n \geq 2$. We use 2^N to denote the power set of N, i.e., the set of all the subsets of N, and $|S|$ is the cardinal of S.

A *profile* is an $n \times n$ matrix

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}$$

with values in the unit interval, where p_{ij} is the assessment with which the individual i evaluates individual j as being qualified to belong to the committee in question. The set of profiles is denoted by $\mathcal{P}$. Given a subset of individuals $S \subseteq N$, P_S denotes the $|S| \times n$ submatrix of P composed by those $i-$rows with $i \in S$. Given $j \in N$, we denote by P_S^j the j-th column vector of P_S.

Definition 2.1 *A* Committees' Evaluation Mechanism (CEM) *is a family of functions $\{v_S\}$, with $\emptyset \neq S \subseteq N$, where $v_S : \mathcal{P} \times N \longrightarrow [0,1]$ is a function that, given $P \in \mathcal{P}$, assigns a collective assessment $v_S(P,j) \in [0,1]$ to each individual $j \in N$ in such a way that $v_S(P,j) = v_S(Q,j)$ for all $P, Q \in \mathcal{P}$ satisfying $P_S^j = Q_S^j$.*

Implicit in the previous definition is the fact that the collective assessment that the subgroup S provides to individual j, $v_S(P,j)$, only depends on the individual assessments of S on the individual j.

A very natural way to transform the gradual opinion that the CEM provides into a dichotomic assessment is by means of thresholds or quotas. Intuitively, an individual is going to be qualified if the collective assessment is above a fixed quota.

Definition 2.2 *A family of values $\{\alpha_S\}$, $\emptyset \neq S \subseteq N$, and $\alpha_S \in (0,1]$ for every S, is called a* Threshold Mechanism (TM).

Given a CEM $\{v_S\}$ and a TM $\{\alpha_S\}$, the family of functions $\{V_S\}$, with $S \subseteq N$, where $V_S : \mathcal{P} \longrightarrow 2^N$ is the function that, given $P \in \mathcal{P}$, qualifies a new subgroup $V_S(P)$ as follows:

$$V_S(P) = \begin{cases} \{j \in N \mid v_S(P,j) \geq \alpha_S\}, & \text{if } S \neq \emptyset, \\ \emptyset, & \text{otherwise.} \end{cases}$$

The family of functions $\{V_S\}$ is called the Committees' Qualification Mechanism (CQM) *associated with $\{\alpha_S\}$ and $\{v_S\}$.*

Definition 2.3 *Given a CQM $\{V_S\}$, the sequence $\{S_t\}$, where $S_1 = N$ and $S_{t+1} = V_{S_t}(P)$, is called* committees' sequence. *A committees' sequence is said to be* convergent *if $\{S_t\}$ has a limit $\lim S_t$ (and it is also said that the sequence* converges *to $\lim S_t$). The CQM is said to be* convergent *if for any $P \in \mathcal{P}$ the committees' sequence generated is convergent.*

As noted in Ballester and García-Lapresta [1], since the society is a finite set, chain convergence can also be expressed as: there exists a positive integer q such that $S_t = S_q$ for every $t \geq q$.

3 An overview of the sequential procedure

We present a brief description of the sequential procedure introduced in Ballester and García-Lapresta [1]. Consider a group of individuals that have to decide which of its members are adequate for a particular task or possess an attribute or ability. The sequential procedure makes use of their opinions (the profile matrix), the social ways to determine qualification (CEM and CQM) in a sequential way (committees' sequence). The aim is to determine a final set of qualified individuals (reason for which convergence is fundamental).

1. The society endows itself of a mechanism of decision, a CQM.
2. All the opinions are stated and the profile matrix is obtained.
3. The whole society (S_1), determines a initial set of qualified members, S_2, by using its associated function V_N in the CQM.
4. The members in S_2 have now special relevance, as they are roughly considered the set of qualified members. They determine, by using the function V_{S_2} in the CQM, who should be the set of qualified members, S_3.
5. The process is iterated until a limit set (or a cycle) is attained. If the limit set is obtained, this is the set of qualified members for the given profile of opinions.

The problem is to determine mechanisms that, given any possible social opinion, determine a final set of qualified individuals (convergence). Notice that the mechanism with which a society endows itself may be useful for a large variety of situations and problems and therefore, it has to be flexible enough to answer positively to any plausible matrix of individual opinions.

4 Describing families of CEMs

In this section, we focus our attention on how general aggregation operators and classes of quasiarithmetic means can be used in the recursive procedure introduced in the previous sections.

4.1 The use of aggregation operators to construct CEMs

Typically, aggregation operators are defined for a given number of individual opinions. Notice however that for our model to apply, we have to consider collections of operators, one for each possible subgroup of individuals in society. Although the ways to afford this task are unlimited, we reconsider in this paper the intuitive two assumptions done in Ballester and García-Lapresta [1].

SE *Self-Exclusion*: When an individual had to decide on herself, she will exclude her opinion when possible (i.e., she is not the only reviewer).

CS *Cardinality-Symmetry*: When the same number of opinions have to be aggregated, the same operator will be applied (independently of the names of the reviewers).

Both of them could be eliminated without altering significatively the results, but we consider this model more natural.

4.2 Quasiarithmetic means

Definition 4.1 *Let* $\varphi : [0,1] \longrightarrow [0,1]$ *be an* order automorphism*, i.e.,* φ *is bijective and increasing. The* quasiarithmetic mean of dimension p associated with φ *is the function* $f^p_\varphi : [0,1]^p \longrightarrow [0,1]$ *defined by*

$$f^p_\varphi(a_1, \ldots, a_p) = \varphi^{-1}\left(\frac{\varphi(a_1) + \cdots + \varphi(a_p)}{p}\right).$$

It is natural to consider quasiarithmetic means that use the same order automorphism φ across dimensions. The construction of the associated CEM function $\{v_S\}$, therefore, would be, for any $\emptyset \neq S \subseteq N$:

1. The quasiarithmetic mean of dimension $|S|$, $f^{|S|}_\varphi$, applied to P^j_S, if $j \notin S$.
2. The quasiarithmetic mean of dimension $|S|-1$, $f^{|S|-1}_\varphi$, applied to $P^j_{S\setminus\{j\}}$, if $j \in S \neq \{j\}$.
3. The quasiarithmetic mean of dimension 1, f^1_φ, i.e., the identity, applied to $P^j_{\{j\}} = p_{jj}$, if $S = \{j\}$.

Notice that the average aggregator is indeed a quasiarithmetic mean simply considering the identity automorphism.

Example 4.2 The *root-power* quasiarithmetic means can be defined through the order automorphism $\varphi(x) = x^q$, with q a positive integer. The associated CEM $\{r_S\}$ is described in Table 1.

Table 1: CEM associated with the *root-power* quasiarithmetic mean

Cases	CEM: $r_S(P,j)$
1. $j \notin S$	$\left(\frac{1}{\lvert S\rvert}\sum_{i\in S} p_{ij}^q\right)^{\frac{1}{q}}$
2. $j \in S \neq \{j\}$	$\left(\frac{1}{\lvert S\rvert - 1}\sum_{i\in S\setminus\{j\}} p_{ij}^q\right)^{\frac{1}{q}}$
3. $S = \{j\}$	p_{jj}

The associated CQM $\{R_S\}$ is defined by

$$j \in R_S(P) \;\Leftrightarrow\; r_S(P,j) \geq \alpha_S.$$

Example 4.3 The *exponential* quasiarithmetic means can be defined through the order automorphism

$$\varphi(x) = \frac{e^{\beta x} - 1}{e^\beta - 1}$$

with $\beta \neq 0$. The associated CEM $\{e_S\}$ is described in Table 2.

Table 2: CEM associated with the *exponential* quasiarithmetic mean

Cases	CEM: $e_S(P,j)$
1. $j \notin S$	$\frac{1}{\beta}\ln\frac{\sum_{i\in S} e^{\beta p_{ij}}}{\lvert S\rvert}$
2. $j \in S \neq \{j\}$	$\frac{1}{\beta}\ln\frac{\sum_{i\in S\setminus\{j\}} e^{\beta p_{ij}}}{\lvert S\rvert - 1}$
3. $S = \{j\}$	p_{jj}

The associated CQM $\{E_S\}$ is defined by

$$j \in E_S(P) \;\Leftrightarrow\; e_S(P,j) \geq \alpha_S.$$

5 The results

In this section we aim to provide some sufficient conditions for the convergence of CQMs associated with quasiarithmetic means.

Definition 5.1 *Given an order automorphism* $\varphi : [0,1] \longrightarrow [0,1]$*, the function* $g_\varphi : (0,1] \longrightarrow (0,1]$ *is defined by*

$$g_\varphi(x) = \frac{\varphi(x)}{x}.$$

Lemma 5.2 *Let* φ *be an order automorphism. If the function* g_φ *is increasing, then* $x\,\varphi(y) \geq \varphi(x\,y)$ *for all* $x, y \in [0,1]$.

PROOF: First of all, notice that the result is true if $x = 0$ or $y = 0$, and $x\,y \leq y$ for all $x, y \in [0,1]$. Since g_φ is increasing, we have

$$\frac{\varphi(y)}{y} = g_\varphi(y) \geq g_\varphi(x\,y) = \frac{\varphi(x\,y)}{x\,y}$$

for all $x, y \in (0,1]$, equivalently, $\varphi(y) \geq \dfrac{\varphi(x\,y)}{x}$ and $x\,\varphi(y) \geq \varphi(x\,y)$. ∎

Proposition 5.3 *Let* $\{V_S\}$ *be a CQM associated with a quasiarithmetic mean* $\{f^p_\varphi\}$ *and a TM* $\{\alpha_S\}$*. If the following two conditions are satisfied:*

1. *The function* g_φ *is increasing,*
2. *For all non-empty sets* $U, V \subseteq N$ *such that* $U \subseteq V$ *it holds* $|U|\,\alpha_U \geq (|V| - 1)\,\alpha_V$*,*

then $\{V_S\}$ *is convergent.*

PROOF: Suppose conditions 1 and 2 hold. Given any profile P, we prove by induction that the committee sequence is decreasing, hence, convergent. Obviously, $S_2 \subseteq S_1 = N$. Suppose $S_{k+1} \subseteq S_k$ is true for any $k \in \{1, \ldots, t-1\}$ (in particular $S_t \subseteq S_{t-1}$). In order to prove $S_{t+1} \subseteq S_t$ by way of contradiction, suppose there exists an individual $j \in S_{t+1}$ such that $j \notin S_t$. Consider the greatest integer m such that $j \in S_m$, with $m \leq t-1$. Notice that this is well-defined, since $S_1 = N$. By the induction hypothesis, it must be that $S_t \subseteq S_m$ and, consequently, $S_t = S_t \setminus \{j\} \subseteq S_m \setminus \{j\}$. Therefore,

$$\frac{\sum\limits_{i \in S_m \setminus \{j\}} \varphi(p_{ij})}{|S_m| - 1} \geq \frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_m| - 1}.$$

Since φ^{-1} is also increasing, we have

$$\varphi^{-1}\left(\frac{\sum\limits_{i \in S_m \setminus \{j\}} \varphi(p_{ij})}{|S_m| - 1}\right) \geq \varphi^{-1}\left(\frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_m| - 1}\right) =$$

$$= \varphi^{-1}\left(\frac{|S_t|}{|S_m| - 1}\,\frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_t|}\right).$$

By the fact that $j \in S_{t+1} = V_{S_t}(P)$ and $j \notin S_t$, we know that

$$\varphi^{-1}\left(\frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_t|}\right) \geq \alpha_{S_t}.$$

Being φ increasing, this is equivalent to

$$\frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_t|} \geq \varphi(\alpha_{S_t}).$$

Hence, substituting:

$$\varphi^{-1}\left(\frac{|S_t|}{|S_m| - 1}\,\frac{\sum\limits_{i \in S_t} \varphi(p_{ij})}{|S_t|}\right) \geq$$

$$\geq \varphi^{-1}\left(\frac{|S_t|}{|S_m| - 1}\,\varphi(\alpha_{S_t})\right).$$

Lemma 5.2 and the fact that φ^{-1} is increasing guarantee that

$$\varphi^{-1}\left(\frac{|S_t|}{|S_m| - 1}\,\varphi(\alpha_{S_t})\right) \geq \varphi^{-1}\left(\varphi\left(\frac{|S_t|}{|S_m| - 1}\,\alpha_{S_t}\right)\right).$$

Condition 2 ensures $|S_t|\,\alpha_{S_t} \geq (|S_m| - 1)\,\alpha_{S_m}$. Then,

$$\varphi^{-1}\left(\varphi\left(\frac{|S_t|}{|S_m| - 1}\,\alpha_{S_t}\right)\right) \geq \varphi^{-1}(\varphi(\alpha_{S_m})) = \alpha_{S_m}.$$

Hence we obtain

$$\varphi^{-1}\left(\frac{\sum\limits_{i \in S_m \setminus \{j\}} \varphi(p_{ij})}{|S_m| - 1}\right) \geq \alpha_{S_m}$$

or, equivalently, $j \in V_{S_m}(P) = S_{m+1}$, leading to an absurd and concluding the proof. ∎

Corollary 5.4 *If for all non-empty sets* $U, V \subseteq N$ *such that* $U \subseteq V$ *it holds* $|U|\,\alpha_U \geq (|V| - 1)\,\alpha_V$*, then any associated CQM* $\{R_S\}$ *is convergent.*

PROOF: It is sufficient to take into account that the function g_φ, defined by

$$g_\varphi(x) = \frac{x^r}{x} = x^{r-1},$$

is increasing. ∎

Corollary 5.5 *Let $\{V_S\}$ be a CQM associated with a quasiarithmetic mean $\{f^p_\varphi\}$ and a TM $\{\alpha_S\}$. If φ is convex and for all non-empty sets $U, V \subseteq N$ such that $U \subseteq V$ it holds $|U|\,\alpha_U \geq (|V|-1)\,\alpha_V$, then $\{V_S\}$ is convergent.*

PROOF: We prove by way of contradiction that the function g_φ is increasing. Suppose that there exist $x, y \in (0,1]$ such that $x < y$ but $g_\varphi(x) > g_\varphi(y)$. In this case,

$$\frac{\varphi(x)}{x} > \frac{\varphi(y)}{y} \;\Rightarrow\; \varphi(x) > \frac{x}{y}\,\varphi(y).$$

Since $\varphi(0) = 0$ and $\dfrac{x}{y} \in (0,1)$, we have

$$\varphi\left(\frac{x}{y}\,y + \left(1 - \frac{x}{y}\right)0\right) = \varphi(x) > \frac{x}{y}\,\varphi(y) =$$

$$= \frac{x}{y}\,\varphi(y) + \left(1 - \frac{x}{y}\right)\varphi(0).$$

This fact contradicts the hypothesis that φ was convex. Thus, the function g_φ is increasing and the associated CQM $\{V_S\}$ is convergent. ∎

The previous corollary gives us a very simple result about the convergence of exponential quasiarithmetic means.

Corollary 5.6 *If for all non-empty sets $U, V \subseteq N$ such that $U \subseteq V$ it holds $|U|\,\alpha_U \geq (|V|-1)\,\alpha_V$, and $\beta > 0$, then the associated CQM $\{E_S\}$ is convergent.*

PROOF: Given Corollary 5.5, the fact that the function

$$\varphi(x) = \frac{e^{\beta x} - 1}{e^\beta - 1}$$

is convex in $(0,1]$ for any $\beta > 0$, allows us to conclude as desired. ∎

6 Concluding remarks

In this paper, we have analyzed the possibility of using different aggregation operators in a previously introduced sequential evaluation mechanism.

The main aim was to obtain results about the convergence of the associated committees' qualification mechanisms. We have provided some sufficient conditions for the convergence of mechanisms based on quasiarithmetic means, paying special attention to root-power and exponential means and those generated by convex order automorphisms.

Acknowledgement

This work is partially financed by the Spanish Ministerio de Educación y Ciencia (Projects SEJ2006-04267/ECON and SEJ2005-01481/ECON), ERDF, CREA, Junta de Castilla y León (Consejería de Educación y Cultura, Project VA040A05), Generalitat de Catalunya (grant 2005SJR00454) and Barcelona Economics-XREA.

References

[1] M. A. Ballester and J. L. García-Lapresta, A recursive group decision making procedure based on EOWA operators. *Proc. FUZZ-IEEE 2007*, London.

[2] P. S. Bullen, D. S. Mitrinović and P. M. Vasić, *Means and Their Inequalities*, Reidel, Dordrecht, 1988.

[3] T. Calvo, A. Kolesárova, M. Komorníková and R. Mesiar, Aggregation operators: Properties, classes and constructions models, in: T. Calvo, G. Mayor, R. Mesiar (Eds.), *Aggregation Operators: New Trends and Applications*, Physica-Verlag, Heidelberg, 2002, pp. 3–104.

[4] J. Fodor and M. Roubens, *Fuzzy Preference Modelling and Multicriteria Decision Support*, Kluwer Academic Publishers, Dordrecht, 1994.

[5] M. Grabisch, S. A. Orlovski and R. R. Yager, Fuzzy aggregation of numerical preferences, in: R. Slowinski (Ed.), *Fuzzy Sets in Decision Analysis, Operations and Statistics*, Kluwer Academic Publishers, Boston, 1998, pp. 31–68.

[6] A. Kasher and A. Rubinstein, On the question "Who is a J?". A social choice approach, *Logique & Analyse* 160 (1997) 385–395.

[7] A. Kolmogoroff, Sur la notion de la moyenne, *Atti della R. Academia Nazionale del Lincei. Rendiconti della Classe di Scienze Fisiche, Mathematiche e Naturali* 12 (1930) 388–391.

[8] E. Laumann, P. Marsden and D. Prensky, The boundary specification problem in network analysis, in: L. C. Freeman, D. R. White, A. K. Romney (Eds.), *Research Methods in Social Network Analysis*. George Mason University Press, Fairfax, 1989, pp. 61–87.

[9] M. Nagumo, Über eine Klasse der Mittelwerte, *Japanese Journal of Mathematics* 7 (1930) 71–79.

[10] D. Samet and D. Schmeidler, Between liberalism and democracy, *Journal of Economic Theory* 110 (2003) 213–233.

[11] R. R. Yager, Ordered weighted averaging operators in multicriteria decision making, *IEEE Transactions on Systems, Man and Cybernetics* 8 (1988) 183–190.

Author index

Achs, Á., 221/II
Adlassnig, K. P., 361/I
Agrafonova, T., 377/I
Aguiló, I., 83/I
Alcalde, C., 157/II
Alvoni, E., 185/I
Asmuss, S., 175/II
Averkin, A. N., 377/I, 387/I, 391/I, 395/I

Baczyński, M., 105/I, 111/I
Balas, M. M., 425/I
Balas, V. E., 425/I
Ballester, M. A., 479/II
Ballini, R., 231/II
Baroni, P., 281/I
Barzamini, R., 137/II, 237/II
Běhounek, L., 145/I, 153/I
Belenki, A. G., 387/I
Beliakov, G., 61/I, 67/I
Belluce, L. P., 195/II
Bobillo, F., 213/II
Bodenhofer, U., 153/I, 321/I
Bosc, P., 343/II
Botzheim, J., 361/I
Bou, F., 177/I
Boukezzoula, R., 279/II
Bradley, J., 417/I
Brinker, K., 353/I
Bronevich, A. G., 297/I
Burusco, A., 157/II

Calvo, T., 61/I
Casasnovas, J., 163/II
Chitsaz, M., 237/II
Chojnacki, E., 367/II
Cintula, P., 137/I, 153/I
Cornelis, C., 13/II
Costa, P., 231/II

Čomić, L., 325/II

De Baets, B., 47/I, 95/II
Deschrijver, G., 53/I, 13/II
Destercke, S., 367/II
Detyniecki, M., 345/I
Di Nola, A., 195/II, 207/II
Divari, M., 169/II
Drobics, M., 361/I
Drygaś, P., 185/II
Dubois, D., 367/II
Dudziak, U., 359/II
Durante, F., 185/I, 191/I
Dvořák, A., 103/II
Dyczkowski, K., 129/II

Esteva, F., 177/I, 21/II

Fermüller, Ch. G., 17/I, 437/I
Fiot, C., 329/I
Flaminio, T., 201/II
Foschi, R., 197/I
Foulloy, L., 279/II
Fuentes-González, R., 157/II

Galichet, S., 279/II
García-Lapresta, J. L., 479/II
Garmendia, L., 351/II

Gegov, A., 111/II
Georgescu, I., 437/II, 445/II
Gerla, B., 195/II
Ghiselli Ricci, R., 91/I
Gil-Lafuente, A. M., 463/II
Gispert, J., 19/I
Gobalakrishnan, N., 111/II
Gocht, U., 473/I, 425/II
Godo, L., 177/I, 21/II
Gomide, F., 231/II
Gouveia, F., 243/II
Grabisch, M., 257/I, 289/I, 313/I
Grbić, T., 221/I
Grigolia, R., 207/II
Gusev, A., 395/I

Hájek, P., 21/I
Habiballa, H., 27/II
Hadjali, A., 343/II
Hampel, R., 459/I, 463/I, 469/I
Hashemi Golpayegani, M., 63/II, 143/II, 149/II
Havranová, Z., 331/II, 337/II
Hell, M., 231/II
Hliněná, D., 431/II
Holčapek, M., 409/II
Horčík, R., 137/I
Hüllermeier, E., 353/I
Hürlimann, W., 205/I
Huňka, F., 273/II

Igrejas, G., 243/II
Imai, H., 311/II
Iyer, V., 71/II

Jaafari, A., 63/II, 143/II, 149/II
Jacas, J., 127/I
Jaworski, P., 213/I
Jayaram, B., 105/I, 111/I
Jenei, S., 23/I
Jouzdani, J., 237/II

Kacprzyk, J., 445/I
Kalganova, I., 381/I
Kalina, M., 331/II, 337/II
Kästner, W., 459/I, 463/I, 469/I
Kawaguchi, M. F., 45/II
Kerre, E. E., 13/II, 49/II
Khosravi, A., 137/II, 237/II
Klawonn, F., 321/I
Klement, E. P., 253/I
Kohout, L. J., 161/I, 429/I
Kojadinovic, I., 305/I
Kolesárová, A., 75/I
Kondo, M., 45/II
Kosiński, W., 449/I
Kosterev, V. V., 391/I
Kráľ, P., 431/II, 471/II
Kramosil, I., 375/II
Kratzsch, A., 469/I
Kroupa, T., 145/I
Křivý, I., 251/II

Labreuche, Ch., 289/I, 295/II, 303/II
Lakov, D. V., 79/II
Laurent, A., 329/I
Lebedeva, O., 403/II
Lendelová, K., 231/I
Lepskiy, A. E., 297/I
Limberg, J., 409/I
Lin, R., 437/II
Lu, J., 137/II
Lucas, C., 149/II

Maes, K. C., 47/I
Marchioni, E., 169/I
Mareš, M., 179/II
Marichal, J. L., 305/I
Marsala, C., 345/I
Mas, M., 97/I
Mayor, G., 75/I
Medina, J., 209/II

Merigó, J. M., 463/II
Mesiar, R., 75/I, 253/I
Michalíková, A., 237/I
Mihailović, B., 265/I
Minoux, M., 25/I
Miranda, P., 257/I
Miyakoshi, M., 45/II
Močkoř, J., 417/II
Molina, C., 337/I
Monserrat, M., 97/I
Montagna, F., 201/II
Motie Nasrabadi, A., 63/II, 143/II, 149/II
Murthy, G. R., 71/II

Nasibov, E., 59/II
Navara, M., 87/II
Nedović, L., 221/I
Noguera, C., 21/II
Nosková, L., 95/II
Novák, V., 103/II
Nurmi, H., 445/I

Ohnishi, S., 311/II
Ojeda-Aciego, M., 209/II
Olivas, J. A., 119/II

Pankowska, A., 317/II
Pap, E., 253/I, 265/I
Papini, P. L., 185/I
Pavlačka, O., 455/II
Pelessoni, R., 281/I
Petrík, M., 87/II
Pivert, O., 343/II
Pourmorteza, A., 63/II
Prieto, M. E., 119/II
Prokopowicz, P., 449/I

Qiu, X., 437/II
Quesada-Molina, J. J., 191/I

Radojević, D. G., 119/I
Ralević, N. M., 325/II
Recasens, J., 127/I, 351/II
Riečan, B., 227/I
Riera, J. V., 163/II
Rodríguez-Muñiz, L. J., 381/II
Rossi, G., 273/I
Ruan, D., 27/I, 49/II
Ruiz-Calviño, J., 209/II
Ruzha, V., 175/II

Sadati, N., 237/II
Salgado, P., 243/II
Salimi Khorshidi, G., 63/II, 143/II, 149/II
Saminger, S., 35/I
Sánchez, D., 337/I
Saralieva, M. R., 79/II
Sarkoci, P., 35/I
Seising, R., 401/I, 409/I
Sempi, C., 185/I
Serrano, J. M., 337/I
Shestakov, M., 395/I
Shi, Y., 49/II
Sokolov, O., 425/II
Solovjovs, S., 395/II
Soto, A., 119/II
Spizzichino, F., 197/I
Stankevich, K. V., 391/I
Stefanini, L., 259/II
Straccia, U., 213/II
Suner, J., 83/I

Šabo, M., 41/I
Škrbić, S., 247/I
Šostak, A., 387/II
Špirková, J., 471/II
Štajner-Papuga, I., 221/I
Štěpnička, M., 95/II

Takači, A., 247/I
Talašová, J., 455/II
Teisseire, M., 329/I

Tikhonov, E., 369/I
Titova, N., 377/I
Torrens, J., 83/I, 97/I
Troiano, L., 67/I
Tsoy, Y. R., 267/II
Tvrdík, J., 251/II

Úbeda-Flores, M., 191/I
Uljane, I., 387/II
Ulutagay, G., 59/II

Van Gasse, B., 13/II
Vetterlein, T., 37/II
Vicig, P., 281/I
Vila, M. A., 337/I
Vivona, D., 169/II

Wagenknecht, M., 473/I, 425/II
Watari, O., 45/II
Wygralak, M., 287/II

Xie, L., 313/I

Yamanoi, T., 311/II
Yankovskaya, A. E., 267/II

Zadeh, L. A., 29/I
Zheng, X., 137/II

Žikić-Došenović, T., 241/I