AGOP 2015

AGOP 2015

July 7–10, 2015, Katowice (Poland)

University of Silesia

Proceedings of the 8th
International Summer School
on Aggregation Operators (AGOP 2015)

Editors: Michał Baczyński, Bernard De Baets, Radko Mesiar

Michał Baczyński, Bernard De Baets, Radko Mesiar (Eds.)
Proceedings of 8th International Summer School on Aggregation Operators (AGOP 2015)

ISBN: 978–83–8012–519–3

Printing and binding:
"TOTEM.COM.PL SP. z o.o." Sp. K.
ul. Jacewska 89, 88-100 Inowrocaw, Poland

Typeset & Cover Design: Aleksandra Hankus

Contents

Preface

The extensive growth of the theory and applications of aggregation functions at the end of the 20th century has led to the creation of the working group AGOP in the framework of EUSFLAT association, and subsequently to the organization of a summer school on aggregation held in Oviedo, Spain, July 2001. After 14 years, following the past editions of AGOP in Oviedo (2001), Alcalá de Henares (2003), Lugano (2005), Ghent (2007), Palma de Mallorca (2009), Benevento (2011) and Pamplona (2013), its eighth edition is organized in Katowice by the Institute of Mathematics of University of Silesia, in cooperation with the EUSFLAT working group AGOP.

Up to the tutorials and invited lectures given by leading persons in the field of aggregation functions (tutorials: Bernard De Baets, Jean-Luc Marichal and Radko Mesiar; invited lectures: József Dombi, Balasubramaniam Jayaram and Maciej Sablik) and contributed presentations, the summer school offers space to numerous discussions and consultations of interested students and young researchers in the domain with distinguished aggregation experts.

We are grateful to all persons having contributed to the success of AGOP 2015, especially to all reviewers and all authors of submitted papers. We would like to thank all the employees of Institute of Mathematics of University of Silesia for their help in preparing this summer school. We also thank Scientific Information Centre and Academic Library (Polish acronym: CINiBA) for providing a classroom for presentations. It is our pleasant duty to acknowledge the financial support from BPSC. Finally, we also express our sincere thanks to EUSFLAT for their support and student grants.

We believe that all participants will profit from this summer school both from a scientific and a social point of view.

Katowice, July 7, 2015

Michał Baczyński
Bernard De Baets
Radko Mesiar

Organization

Chairs

M. Baczyński (Poland)
B. De Baets (Belgium)
R. Mesiar (Slovakia)

Tutorials

B. De Baets (Belgium)
J.–L. Marichal (Luxembourg)
R. Mesiar (Slovakia)

Invited talks

J. Dombi (Hungary)
B. Jayaram (India)
M. Sablik (Poland)

Program Committee

B. Bedregal (Brazil)
G. Beliakov (Australia)
H. Bustince (Spain)
T. Calvo (Spain)
J. Dombi (Hungary)
J. Drewniak (Poland)
P. Drygaś (Poland)
J. Dujmović (USA)
F. Durante (Italy)
J. Fernandez (Spain)
J. Fodor (Hungary)
J.L. Garcia-Lapresta (Spain)
M. Grabisch (France)
S. Greco (Italy)
E. Herrera-Viedma (Spain)
E. Induráin (Spain)
B. Jayaram (India)
M. Kalina (Slovakia)
E.P. Klement (Austria)
A. Kolesárová (Slovakia)
J.–L. Marichal (Luxembourg)
G. Mayor (Spain)
J. Montero (Spain)
S. Montes (Spain)
S. Massanet (Spain)
E. Pap (Serbia)
I. Perfilieva (Czech Republic)
H. Prade (France)
A. Pradera (Spain)
S. Saminger-Platz (Austria)
B. Teheux (France)
V. Torra (Spain)
J. Torrens (Spain)
R.R. Yager (USA)

Additional Referees

U. Bentkowska (Poland)
S. Bodjanova (Slovakia)
S. Graco (Italy)
P. Grzegorzewski (Poland)
V. Janiš (Slovakia)
P. Jaworski (Poland)
J.M. Merigó (Chile)
E. Pap (Serbia)
B. Pękala (Poland)
B. Riečan (Slovakia)
D. Ruiz–Aguilera (Spain)
C. Sempi (Italy)
A. Takači (Serbia)
L. Zedam (Algeria)

Local Organizing Committee

M. Baczyński
M. Gągolewski
A. Hankus
E. Kurzeja
J. Sobera
S. Wolna

Tutorials

THE ROLE OF AGGREGATION FUNCTIONS IN FUZZY RELATIONAL CALCULUS

Bernard De Baets
KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics,
Ghent University
Coupure links 653, 9000 Gent
Belgium
Bernard.DeBaets@UGent.be

Summary

In this tutorial lecture, we will consider both unipolar and bipolar fuzzy relations. In the first setting, fuzzy relations can be seen as a generalization of crisp relations. The composition of unipolar fuzzy relations is based on a conjunctive aggregation function, and is tightly linked to the notion of transitivity. In particular, we will focus on fuzzy preference relations and the related notion of an additive fuzzy preference structure.

In the second setting, fuzzy relations can be seen as a generalization of complete crisp relations and should preferably be called reciprocal relations. A notion of composition is far from evident in this setting and various notions of transitivity can be considered here, not only based on conjunctive aggregation functions. In particular, we will focus on winning probability relations and the notion of cycle-transitivity.

ISBN: 978–83–8012–519–3

GENERALIZATIONS AND VARIANTS OF ASSOCIATIVITY FOR AGGREGATION FUNCTIONS

Jean-Luc Marichal and Bruno Teheux
Mathematics Research Unit, FSTC, University of Luxembourg
6, rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg
{jean-luc.marichal,bruno.teheux}[at]uni.lu

Summary

Consider an associative operation $G: X^2 \to X$ on a set X and denote $G(a,b)$ merely by ab. By definition, we have $(ab)c = a(bc)$ for all $a,b,c \in X$ and this property enables us to define the expression abc unambiguously by setting $abc = (ab)c$. More generally, for any $a_1, a_2, \ldots, a_n \in X$, we can set

$$a_1 a_2 a_3 \cdots a_n \ = \ (\cdots((a_1 a_2)a_3)\cdots)a_n$$

and associativity shows that this expression can be computed regardless of how parentheses are inserted. This means that the identity

$$a_1 \cdots a_j \cdots a_k \cdots a_n \ = \ a_1 \cdots (a_j \cdots a_k) \cdots a_n$$

holds for any integers $1 \leqslant j \leqslant k \leqslant n$.

This latter condition has been considered in aggregation function theory to extend the classical associativity property of binary operations to variadic operations, i.e., those operations that have an indefinite arity. In this note we survey the most recent results obtained not only on this extension of associativity but also on some variants and generalizations of this property, including barycentric associativity and preassociativity.

Keywords: Associativity, Preassociativity, Barycentric associativity, Barycentric preassociativity, Variadic function, String function, Functional equation, Axiomatization.

1 Introduction

Let X denote a nonempty set, called the *alphabet,* and its elements are called *letters*. The symbol X^* stands for the set $\bigcup_{n \geqslant 0} X^n$ of all tuples on X. Its elements are called *strings* and denoted by bold roman letters $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$ If we want to stress that such an element is a letter of X, we use non-bold italic letters $x, y, z, \ldots$ We assume that X^0 has only one element; we denote it by ε and call it the *empty string.* We endow the set X^* with the concatenation operation for which the empty string ε is the neutral element. For instance, if $\mathbf{x} \in X^m$ and $y \in X$, then $\mathbf{x}y\varepsilon = \mathbf{x}y \in X^{m+1}$. For every string $\mathbf{x}$ and every integer $n \geqslant 1$, the power $\mathbf{x}^n$ stands for the string obtained by concatenating n copies of $\mathbf{x}$. By extension, we set $\mathbf{x}^0 = \varepsilon$. The *length* of a string $\mathbf{x}$ is denoted by $|\mathbf{x}|$. For instance, we have $|\varepsilon| = 0$.

Let Y be a nonempty set. Recall that, for every integer $n \geqslant 0$, a function $F: X^n \to Y$ is said to be *n-ary.* Also, a function $F: X^* \to Y$ is said to have an *indefinite arity* or to be *variadic* or *$*$-ary* (pronounced "star-ary"). A unary operation on X^* is a particular variadic function $F: X^* \to X^*$ called a *string function* over the alphabet X.

The main functional properties for variadic functions that we present and investigate in this survey are given in the following definition.

Definition 1.1. A string function $F: X^* \to X^*$ is said to be

- *associative* if, for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$, we have
$$F(\mathbf{xyz}) \ = \ F(\mathbf{x}F(\mathbf{y})\mathbf{z})\,;$$
- *barycentrically associative* (or *B-associative*) if, for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$, we have
$$F(\mathbf{xyz}) \ = \ F(\mathbf{x}F(\mathbf{y})^{|\mathbf{y}|}\mathbf{z})\,.$$

A variadic function $F: X^* \to Y$ is said to be

- *preassociative* if, for every $\mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^*$, we have
$$F(\mathbf{y}) = F(\mathbf{y}') \quad \Rightarrow \quad F(\mathbf{xyz}) = F(\mathbf{xy}'\mathbf{z})\,;$$
- *barycentrically preassociative* (or *B-preassociative*) if, for every $\mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^*$, we have
$$\left.\begin{array}{rcl} F(\mathbf{y}) & = & F(\mathbf{y}') \\ |\mathbf{y}| & = & |\mathbf{y}'| \end{array}\right\} \quad \Rightarrow \quad F(\mathbf{xyz}) \ = \ F(\mathbf{xy}'\mathbf{z})\,.$$

ISBN: 978–83–8012–519–3

For any variadic function $F: X^* \to Y$ and any integer $n \geqslant 0$, we denote by F_n the *n-ary part* of F, i.e., the restriction $F|_{X^n}$ of F to the set X^n. We also let $X^+ = X^* \smallsetminus \{\varepsilon\}$ and denote the restriction $F|_{X^+}$ of F to X^+ by F^+. The range of any function f is denoted by $\mathrm{ran}(f)$.

A variadic function $F: X^* \to Y$ is said to be

- a *variadic operation on* X (or an *operation* for short) if $\mathrm{ran}(F) \subseteq X \cup \{\varepsilon\}$.
- *standard* if $F(\mathbf{x}) = F(\varepsilon)$ if and only if $\mathbf{x} = \varepsilon$.
- ε-*standard* if $\varepsilon \in Y$ and if we have $F(\mathbf{x}) = \varepsilon$ if and only if $\mathbf{x} = \varepsilon$.

2 Associativity and variants

In this main section we investigate the properties given in Definition 1.1.

2.1 Associativity

We first discuss the associativity property for the class of variadic operations, which constitutes an important subclass of string functions. Recall that a binary operation $G: X^2 \to X$ is said to be *associative* if

$$G(G(xy)z) \;=\; G(x\,G(yz))\,, \qquad x, y, z \in X.$$

A huge number of associative binary operations have been discovered and investigated for years. They are at the root of the concepts of group and semigroup. For instance, the set intersection and union over a power set are associative binary operations. Logical connectives “and” and “or” as well as many of their fuzzy counterparts are also associative binary operations.

More recently, associative binary operations have also been studied as real or complex functions within the theories of functional equations and aggregation functions (see, e.g., [15]). Various classes of associative binary operations over real intervals can be found in [2,4,6–9,11,14–17,20–22,24,33,34].

Let us now consider an associative standard operation $F: X^* \to X \cup \{\varepsilon\}$. This operation is necessarily ε-standard and can always be constructed from an associative binary operation $G: X^2 \to X$ simply by setting $F_0 = \varepsilon$, $F_1 = \mathrm{id}_X$, $F_2 = G$, and $F_{n+1}(\mathbf{y}z) = F_2(F_n(\mathbf{y})z)$ for every $n \geqslant 2$. To give an example, from the binary operation $G: \mathbb{R}^2 \to \mathbb{R}$ defined by $G(x, y) = x + y$, we can construct the associative standard operation $F: \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$ defined by $F_n(\mathbf{x}) = \sum_{i=1}^n x_i$ for every integer $n \geqslant 1$.

This construction process immediately follows from the following important proposition.

Proposition 2.1 ([19,27,28]). *A standard operation* $F: X^* \to X \cup \{\varepsilon\}$ *is associative if and only if the following conditions hold.*

(a) $F_0(\varepsilon) = \varepsilon$, $F_1 \circ F_1 = F_1$, $F_1 \circ F_2 = F_2$.

(b) $F_2(xy) = F_2(F_1(x)y) = F_2(xF_1(y))$ *for all* $x, y \in X$.

(c) F_2 *is associative.*

(d) $F(\mathbf{y}z) = F(F(\mathbf{y})z)$ *for all* $\mathbf{y} \in X^*$ *and all* $z \in X$ *such that* $|\mathbf{y}z| \geqslant 3$.

Proposition 2.1 provides a characterization of associative standard operations $F: X^* \to X \cup \{\varepsilon\}$ in terms of conditions on their constitutive parts F_n $(n \geqslant 0)$. Conditions (a)–(c) are actually necessary and sufficient conditions on F_0, F_1, and F_2 for F to be associative, while condition (d) provides an induction property which shows that every F_n $(n \geqslant 3)$ can be constructed uniquely from F_2.

Corollary 2.2. *Any associative standard variadic operation is completely determined by its unary and binary parts.*

Let us now say some words about associativity for string functions. It is noteworthy that several data processing tasks correspond to associative string functions. For instance, the function which corresponds to sorting the letters of every string in alphabetical order is associative. Similarly, the function which consists in transforming a string of letters into upper case is also associative. In such a context, associativity is a natural property since it enables us to work locally on small pieces of data at a time.

It is to be noted that the definition of associativity remains unchanged if the length of the string $\mathbf{xz}$ is bounded by one. This observation provides an equivalent but weaker form of associativity.

Proposition 2.3 ([19,27,28]). *A function* $F: X^* \to X^*$ *is associative if and only if* $F(\mathbf{xyz}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z})$ *for any* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$ *such that* $|\mathbf{xz}| \leqslant 1$.

2.2 B-associativity

By definition, B-associativity expresses that the function value of a string does not change when replacing every letter of a substring with the value of this substring. For instance, the arithmetic mean over the set of real numbers, regarded as the ε-standard operation $F: \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$ defined as $F_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n x_i$ for every integer $n \geqslant 1$, is B-associative.

Remark 1. The name B-associativity is justified by the following geometric interpretation. Consider a set of identical homogeneous balls in $X = \mathbb{R}^n$. Each ball

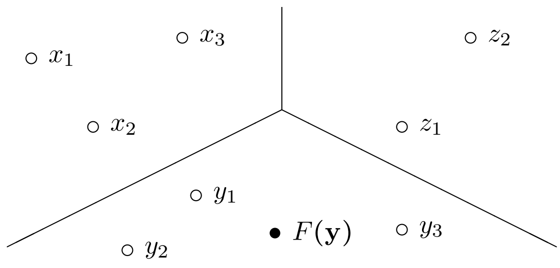


Figure 1: Barycentric associativity

is identified by the coordinates $x \in X$ of its center. Let $F: X^* \to X \cup \{\varepsilon\}$ be the ε-standard variadic operation which carries any set of balls into their barycenter. Because of the associativity-like property of the barycenter, the operation F has to satisfy the functional property of B-associativity (see Fig. 1).

A noteworthy class of B-associative variadic operations is given by the so-called quasi-arithmetic mean functions, axiomatized independently by Kolmogoroff [18] and Nagumo [32].

Definition 2.4. Let $\mathbb{I}$ be a nontrivial real interval (i.e., nonempty and not a singleton), possibly unbounded. A function $F: \mathbb{I}^* \to \mathbb{R}$ is said to be a *quasi-arithmetic mean function* if there is a continuous and strictly monotonic function $f: \mathbb{I} \to \mathbb{R}$ such that

$$F_n(\mathbf{x}) \;=\; f^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right), \qquad n \geqslant 1.$$

The following theorem gives the axiomatization by Kolmogoroff. Even though Kolmogoroff considered functions $F: \bigcup_{n\geqslant 1} \mathbb{I}^n \to \mathbb{I}$, here we have extended the domain of these functions to $\mathbb{I}^*$. Also, it has been recently proved [30] that the idempotence property of F_n (i.e., $F_n(x^n) = x$ for every $x \in \mathbb{I}$), originally stated in Kolmogoroff-Nagumo's characterization, is not needed and hence can be removed. Note also that a variant and a relaxation of Kolmogoroff-Nagumo's characterization can also be found in [12, 13, 22].

Theorem 2.5 (Kolmogoroff-Nagumo)**.** *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. A function $F: \mathbb{I}^* \to \mathbb{I}$ is B-associative and, for every integer $n \geqslant 1$, the n-ary part F_n is symmetric, continuous, and strictly increasing in each argument if and only if F is a quasi-arithmetic mean function.*

The existence of nonsymmetric B-associative operations can be illustrated by the following example, introduced in [21, p. 81] (see also [26]). For every $z \in \mathbb{R}$, the ε-standard operation $M^z: \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$ defined as

$$M^z_n(\mathbf{x}) \;=\; \frac{\sum_{i=1}^{n} z^{n-i}(1-z)^{i-1}\, x_i}{\sum_{i=1}^{n} z^{n-i}(1-z)^{i-1}}, \qquad n \geqslant 1,$$

is B-associative. Actually, one can show [25] that any B-associative ε-standard operation over $\mathbb{R}$ whose n-ary part is a nonconstant linear function for every $n \geqslant 1$ is necessarily one of the operations M^z ($z \in \mathbb{R}$). More generally, the class of B-associative polynomial ε-standard operations (i.e., such that the n-ary part is a polynomial function for every $n \geqslant 1$) over an infinite commutative integral domain D has also been characterized in [25].

2.3 Preassociativity

By definition, a function $F: X^* \to Y$ is preassociative if the function value of any string does not change when modifying any of its substring without changing its value. For instance, any ε-standard operation $F: \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$ defined by $F_n(\mathbf{x}) = f(\sum_{i=1}^{n} x_i)$ for every integer $n \geqslant 1$, where $f: \mathbb{R} \to \mathbb{R}$ is a one-to-one function, is preassociative.

The following two results clearly show that preassociativity is a generalization of associativity.

Proposition 2.6 ([19])**.** *A function $F: X^* \to X^*$ is associative if and only if it is preassociative and satisfies $F = F \circ F$.*

Proposition 2.7 ([27, 28])**.** *An ε-standard operation $F: X^* \to X \cup \{\varepsilon\}$ is associative if and only if it is preassociative and satisfies $F^+ = F_1 \circ F^+$.*

Apart from the fact that it constitutes a less stringent form of associativity, preassociativity has the remarkable feature of avoiding functional composition in its definition. Actually, Propositions 2.6 and 2.7 suggest that preassociativity is precisely the property we obtain from associativity when cleared of any functional composition. Due to this feature, preassociativity can be considered within the wider class of functions taking as inputs strings over an alphabet X and valued over a possibly different set Y. A natural and noteworthy example of a preassociative function is the mapping that outputs the length of strings.

We now show that all preassociative functions $F: X^* \to Y$ are actually strongly related to associativity, even if the set Y is different from X^*. More precisely, we give a characterization of the preassociative functions $F: X^* \to Y$ as compositions of the form $F = f \circ H$, where $H: X^* \to X^*$ is associative and $f: \mathrm{ran}(H) \to Y$ is one-to-one.

Theorem 2.8 ([19])**.** *Let $F: X^* \to Y$ be a function. The following conditions are equivalent.*

(i) F is preassociative.

(ii) There exists an associative function $H: X^* \to X^*$ and a one-to-one function $f: \mathrm{ran}(H) \to Y$ such that $F = f \circ H$.

Corollary 2.9 ([27,28]). *Let $F: X^* \to Y$ be a standard function. The following conditions are equivalent.*

(i) *F is preassociative and satisfies* $\mathrm{ran}(F_1) = \mathrm{ran}(F^+)$.

(ii) *There exists an associative ε-standard operation $H: X^* \to X \cup \{\varepsilon\}$ and a one-to-one function $f: \mathrm{ran}(H^+) \to Y$ such that $F^+ = f \circ H^+$.*

Corollary 2.9 enables us to construct preassociative functions very easily from known associative variadic operations. Just take nonempty sets X and Y, an associative ε-standard operation $H: X^* \to X \cup \{\varepsilon\}$, and a one-to-one function $f: \mathrm{ran}(H^+) \to Y$. For any $e \notin Y$, the standard function $F: X^* \to Y \cup \{e\}$ defined by $F(\varepsilon) = e$ and $F^+ = f \circ H^+$ is preassociative.

Example 2.10. Recall that the *multilinear extension* of a pseudo-Boolean function $f: \{0,1\}^n \to \mathbb{R}$ is the unique multilinear polynomial function $\mathrm{MLE}(f)$ obtained from f by linear interpolation with respect to each of the n variables. Its restriction to $\{0,1\}^n$ is the function f. Let X and Y denote the class of n-variable pseudo-Boolean functions and the class of n-variable multilinear polynomial functions, respectively. For any $e \notin Y$ and any ε-standard operation $H: X^* \to X \cup \{\varepsilon\}$, the standard function $F: X^* \to Y \cup \{e\}$ defined by $F(\varepsilon) = e$ and $F^+ = \mathrm{MLE} \circ H^+$ is preassociative.

Corollary 2.9 also enables us to produce axiomatizations of classes of preassociative functions from known axiomatizations of classes of associative functions. Let us illustrate this observation on an example. Further examples can be found in [29].

Let us recall an axiomatization of the Aczélian semigroups due to Aczél [1] (see also [7,8]).

Proposition 2.11 ([1]). *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. An operation $H: \mathbb{I}^2 \to \mathbb{I}$ is continuous, one-to-one in each argument, and associative if and only if there exists a continuous and strictly monotonic function $\varphi: \mathbb{I} \to \mathbb{J}$ such that*

$$H(x,y) = \varphi^{-1}(\varphi(x) + \varphi(y)),$$

where $\mathbb{J}$ is a real interval of one of the forms $]-\infty, b[$, $]-\infty, b]$, $]a, \infty[$, $[a, \infty[$ or $\mathbb{R} =]-\infty, \infty[$ $(b \leqslant 0 \leqslant a)$. For such an operation H, the interval $\mathbb{I}$ is necessarily open at least on one end. Moreover, φ can be chosen to be strictly increasing.

It is easy to see that there is only one associative ε-standard operation $H: \mathbb{I}^* \to \mathbb{I} \cup \{\varepsilon\}$ whose binary part coincides with the one given in Proposition 2.11. This operation is defined by

$$H_n(\mathbf{x}) = \varphi^{-1}\left(\sum_{i=1}^{n} \varphi(x_i)\right), \qquad n \geqslant 1.$$

Combining this observation with Corollary 2.9 produces the following characterization result.

Theorem 2.12 ([29]). *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. A standard function $F: \mathbb{I}^* \to \mathbb{R}$ is preassociative and unarily quasi-range-idempotent, and F_1 and F_2 are continuous and one-to-one in each argument if and only if there exist continuous and strictly monotonic functions $\varphi: \mathbb{I} \to \mathbb{J}$ and $\psi: \mathbb{J} \to \mathbb{R}$ such that*

$$F_n(\mathbf{x}) = \psi\left(\sum_{i=1}^{n} \varphi(x_i)\right), \qquad n \geqslant 1,$$

where $\mathbb{J}$ is a real interval of one of the forms $]-\infty, b[$, $]-\infty, b]$, $]a, \infty[$, $[a, \infty[$ or $\mathbb{R} =]-\infty, \infty[$ $(b \leqslant 0 \leqslant a)$. For such a function F, we have $\psi = F_1 \circ \varphi^{-1}$ and $\mathbb{I}$ is necessarily open at least on one end. Moreover, φ can be chosen to be strictly increasing.

2.4 B-preassociativity

Contrary to preassociativity, B-preassociativity recalls the associativity-like property of the barycenter and may be easily interpreted in various areas. In decision making for instance, in a sense it says that if we express an indifference when comparing two profiles, then this indifference is preserved when adding identical pieces of information to these profiles. In descriptive statistics and aggregation function theory, it says that the aggregated value of a series of numerical values remains unchanged when modifying a bundle of these values without changing their partial aggregation.

The following result is the barycentric version of Proposition 2.6 and shows that B-preassociativity is a generalization of B-associativity.

Proposition 2.13 ([30]). *A function $F: X^* \to X^*$ is B-associative if and only if it is B-preassociative and satisfies $F(\mathbf{x}) = F(F(\mathbf{x})^{|\mathbf{x}|})$ for all $\mathbf{x} \in X^*$.*

The ε-standard sum operation $F: \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$ defined as $F_n(\mathbf{x}) = \sum_{i=1}^{n} x_i$ for every $n \geqslant 1$ is an instance of B-preassociative function which is not B-associative.

A string function $F: X^* \to X^*$ is said to be *length-preserving* if $|F(\mathbf{x})| = |\mathbf{x}|$ for every $\mathbf{x} \in X^*$.

Proposition 2.14 ([31]). *Let $F: X^* \to X^*$ be a length-preserving function. Then F is associative if and only if it is B-preassociative and satisfies $F_n = F_n \circ F_n$ for every $n \geqslant 0$.*

We now show that, along with preassociative functions, all B-preassociative functions $F: X^* \to Y$ are strongly related to associativity. More precisely, B-preassociative functions can be factorized as compositions of length-preserving associative string functions with one-to-one unary maps.

Theorem 2.15 ([31]). *Let $F: X^* \to Y$ be a function. The following assertions are equivalent.*

(i) F is B-preassociative.

(ii) There exist an associative and length-preserving function $H: X^* \to X^*$ and a sequence $(f_n)_{n\geqslant 1}$ of one-to-one functions $f_n: \mathrm{ran}(H_n) \to Y$ such that $F_n = f_n \circ H_n$ for every $n \geqslant 1$.

The following corollary provides an alternative factorization result for B-preassociative functions in which the inner functions are B-associative operations. For every integer $n \geqslant 1$, the diagonal section $\delta_F: X \to Y$ of a function $F: X^n \to Y$ is defined as $\delta_F(x) = F(x^n)$.

Corollary 2.16 ([30]). *Let $F: X^* \to Y$ be a function. The following assertions are equivalent.*

(i) F is B-preassociative and satisfies $\mathrm{ran}(\delta_{F_n}) = \mathrm{ran}(F_n)$ for every $n \geqslant 1$.

(ii) There exists a B-associative ε-standard operation $H: X^* \to X \cup \{\varepsilon\}$ and a sequence $(f_n)_{n\geqslant 1}$ of one-to-one functions $f_n: \mathrm{ran}(H_n) \to Y$ such that $F_n = f_n \circ H_n$ for every $n \geqslant 1$.

Corollary 2.16 enables us to produce axiomatizations of classes of B-preassociative functions from known axiomatizations of classes of B-associative functions. Let us illustrate this observation on the class of *quasi-arithmetic pre-mean functions*.

Definition 2.17 ([30]). Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. A function $F: \mathbb{I}^* \to \mathbb{R}$ is said to be a *quasi-arithmetic pre-mean function* if there are continuous and strictly increasing functions $f: \mathbb{I} \to \mathbb{R}$ and $f_n: \mathbb{R} \to \mathbb{R}$ $(n \geqslant 1)$ such that

$$F_n(\mathbf{x}) \;=\; f_n\left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right), \qquad n \geqslant 1.$$

As expected, the class of quasi-arithmetic pre-mean functions includes all the quasi-arithmetic mean functions (just take $f_n = f^{-1}$). Actually the quasi-arithmetic mean functions are exactly those quasi-arithmetic pre-mean functions which are idempotent (i.e., such that $f_n \circ f = \mathrm{id}_{\mathbb{I}}$ for every integer $n \geqslant 1$). However, there are also many non-idempotent quasi-arithmetic pre-mean functions. Taking for instance $f_n(x) = nx$ and $f(x) = x$ over the reals $\mathbb{I} = \mathbb{R}$, we obtain the sum function. Taking $f_n(x) = \exp(nx)$ and $f(x) = \ln(x)$ over $\mathbb{I} =]0, \infty[$, we obtain the product function.

We have the following characterization of the quasi-arithmetic pre-mean functions, which generalizes Kolmogoroff-Nagumo's axiomatization of the quasi-arithmetic mean functions.

Theorem 2.18 ([30]). *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. A function $F: \mathbb{I}^* \to \mathbb{R}$ is B-preassociative and, for every $n \geqslant 1$, the function F_n is symmetric, continuous, and strictly increasing in each argument if and only if F is a quasi-arithmetic pre-mean function.*

3 Historical notes

In the framework of aggregation function theory, the associativity property for functions having an indefinite arity was introduced first for functions $F: \bigcup_{n\geqslant 1} X^n \to X$ satisfying $F_1 = \mathrm{id}_X$ (see [21, p. 24]; see also [5, p. 16], [15, p. 32], [17, p. 216] for alternative forms). Then it was introduced for ε-standard variadic operations $F: X^* \to X \cup \{\varepsilon\}$ (see [6,27,28]), and finally for string functions (see [19]).

A basic form of B-associativity was first proposed for symmetric real functions $F: \bigcup_{n\geqslant 1} \mathbb{R}^n \to \mathbb{R}$ independently by Schimmack [35], Kolmogoroff [18], and Nagumo [32]. More precisely, Schimmack introduced the condition $F(\mathbf{y}z) = F(F(\mathbf{y})^{|\mathbf{y}|}z)$ while Kolmogoroff and Nagumo considered the condition $F(\mathbf{yz}) = F(F(\mathbf{y})^{|\mathbf{y}|}\mathbf{z})$ with $|\mathbf{z}| \geqslant 1$. A more general definition appeared more recently in [3] and [21] and has then been used to characterize various classes of functions; see, e.g., [12,13,23,25,26]. The general definition of B-associativity given in Definition 1.1 appeared in [31]. For general background on B-associativity and its links with associativity, see [15, Sect. 2.3] and [30]. The B-associativity property and its different versions are known under at least three different names: *associativity of means* [10], *decomposability* [14, Sect. 5.3], and *barycentric associativity* [3, 30].

Preassociativity was introduced in [27,28] to generalize the associativity property. B-preassociativity was introduced in [30] to generalize the B-associativity property. The basic idea behind this latter definition goes back to 1931 when de Finetti [10] introduced the following associativity-like property for mean functions: for any $u \in X$ and any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$ such that $|\mathbf{xz}| \geqslant 1$ and $|\mathbf{y}| \geqslant 1$, we have $F(\mathbf{xyz}) = F(\mathbf{x}u^{|\mathbf{y}|}\mathbf{z})$ whenever $F(\mathbf{y}) = F(u^{|\mathbf{y}|})$.

Acknowledgements

This research is supported by the project F1R-MTH-PUL-15MRO3 of the University of Luxembourg.

References

[1] J. Aczél. Sur les opérations définies pour nombres réels. *Bull. Soc. Math. France*, 76:59–64, 1949.

[2] C. Alsina, M. J. Frank, and B. Schweizer. *Associative functions: triangular norms and copulas.* World Scientific, London, 2006.

[3] C. Antoine. *Les Moyennes.* Volume 3383 of *Que sais-je?* Presse Universitaires de France, Paris, 1998.

[4] B. Bacchelli. Representation of continuous associative functions. *Stochastica*, 10:13-28, 1986.

[5] T. Calvo, A. Kolesárová, M. Komorníková, and R. Mesiar. Aggregation operators: properties, classes and construction methods. In: *Aggregation operators. New trends and applications*, pages 3-104. Stud. Fuzziness Soft Comput. Vol. 97. Physica-Verlag, Heidelberg, Germany, 2002.

[6] M. Couceiro and J.-L. Marichal. Associative polynomial functions over bounded distributive lattices. *Order* 28:1-8, 2011.

[7] M. Couceiro and J.-L. Marichal. Aczélian n-ary semigroups. *Semigroup Forum* 85:81-90, 2012.

[8] R. Craigen and Zs. Páles. The associativity equation revisited. *Aeq. Math.*, 37:306–312, 1989.

[9] E. Czogała and J. Drewniak. Associative monotonic operations in fuzzy set theory. *Fuzzy Sets and Syst.*, 12:249-269, 1984.

[10] B. de Finetti. Sul concetto di media. *Giornale dell' Instituto Italiano degli Attari* 2(3):369–396, 1931.

[11] K. Domańska, An analytic description of the class of rational associative functions. *Annales Universitatis Paedagogicae Cracoviensis. Studia Mathematica*, 11:111-122, 2012.

[12] J. Fodor and J.-L. Marichal. On nonstrict means. *Aequationes Math.* 54:308–327, 1997.

[13] J. Fodor and J.-L. Marichal. Erratum to “On nonstrict means”. *Aequationes Math.* 71:318–320, 2006.

[14] J. Fodor and M. Roubens. *Fuzzy preference modelling and multicriteria decision support.* Theory and Decision Library. Series D: System Theory, Knowledge Engineering and Problem Solving. Kluwer Academic Publisher, Dordrecht, The Netherlands, 1994.

[15] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions.* Encyclopedia of Mathematics and its Applications, vol. 127. Cambridge University Press, Cambridge, 2009.

[16] E. P. Klement and R. Mesiar (Eds). *Logical, algebraic, analytic, and probabilistic aspects of triangular norms.* Elsevier Science, Amsterdam, The Netherlands, 2005.

[17] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms.* In: Trends in Logic - Studia Logica Library, vol. 8. Kluwer Academic, Dordrecht, 2000.

[18] A. N. Kolmogoroff. Sur la notion de la moyenne. (French). *Atti Accad. Naz. Lincei*, 12(6):388–391, 1930.

[19] E. Lehtonen, J.-L. Marichal, B. Teheux. Associative string functions. *Asian-European Journal of Mathematics*, 7(4):1450059 (18 pages), 2014.

[20] C.H. Ling. Representation of associative functions. *Publ. Math. Debrecen*, 12:189–212, 1965.

[21] J.-L. Marichal. *Aggregation operators for multicriteria decision aid.* PhD thesis, Department of Mathematics, University of Liège, Liège, Belgium, 1998.

[22] J.-L. Marichal. On the associativity functional equation. *Fuzzy Sets and Systems*, 114:381-389, 2000.

[23] J.-L. Marichal. On an axiomatization of the quasi-arithmetic mean values without the symmetry axiom. *Aequat. Math.*, 59:74–83, 2000.

[24] J.-L. Marichal and P. Mathonet. A description of n-ary semigroups polynomial-derived from integral domains. *Semigroup Forum*, 83:241-249, 2011.

[25] J.-L. Marichal, P. Mathonet, and J. Tomaschek. A classification of barycentrically associative polynomial functions. *Aequat. Math.*, in press, 2015.

[26] J.-L. Marichal, P. Mathonet, and E. Tousset. Characterization of some aggregation functions stable for positive linear transformations. *Fuzzy Sets and Syst.* 102:293–314, 1999.

[27] J.-L. Marichal and B. Teheux. Associative and preassociative functions. *Semigroup Forum*, 89(2):431–442, 2014.

[28] J.-L. Marichal and B. Teheux. Associative and preassociative functions (improved version). Working paper (arXiv:1309.7303v3).

[29] J.-L. Marichal and B. Teheux. Preassociative aggregation functions. *Fuzzy Sets and Systems*, 268:15–26, 2015.

[30] J.-L. Marichal and B. Teheux. Barycentrically associative and preassociative functions. *Acta Mathematica Hungarica*, 145(2):468–488, 2015.

[31] J.-L. Marichal and B. Teheux. A characterization of barycentrically preassociative functions. Working paper (arXiv:1410.0778).

[32] M. Nagumo. Über eine klasse der mittelwerte. (German). *Japanese Journ. of Math.*, 7:71–79, 1930.

[33] W. Sander. Associative aggregation operators. In: *Aggregation operators. New trends and applications*, pages 124-158. Stud. Fuzziness Soft Comput. Vol. 97. Physica-Verlag, Heidelberg, Germany, 2002.

[34] W. Sander. Some aspects of functional equations. In: *Logical, algebraic, analytic, and probabilistic aspects of triangular norms*, pages 143-187. Elsevier B.V., Amsterdam, 2005.

[35] R. Schimmack. Der Satz vom arithmetischen Mittel in axiomatischer Begründung. *Math. Ann.* 68:125–132, 1909.

AGGREGATION FUNCTIONS BASED ON INTEGRALS WITH RESPECT TO MONOTONE MEASURES

Radko Mesiar
Department of Mathematics, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 810 05 Bratislava, Slovakia
radko.mesiar@stuba.sk

Summary

We discuss and exemplify several kinds of integrals with respect to monotone measures, and in particular with respect to capacities on a finite universe N.

Keywords: Capacity, Choquet integral, Monotone measure, OWA operator, OMA operator, Sugeno integral.

1 INTRODUCTION

Integrals belong to the historical aggregation functions. Frustum of a pyramide described in the Mocow papyrus dated back to 1 850 B.C. can be considered as the earliest written trace of integrals. The exhaustion principle introduced by Eudoxus around 370 B.C. is a powerful tool also nowadays. Integrals dated till Lebesgue were based on $(\sigma-)$ additive measures, and a first integral approach, dealing with upper/lower extensions of measures - thus violating the additivity, in general - is due to Vitali [27]. Later, many kinds of integrals dealing with special or general monotone measures were considered. We restrict our considerations to finite spaces $N = \{1, ..., n\}, n \in \mathbb{N}$, and then any function $f : N \to [0, \infty]$ can be represented as an n-dimensional vector $\mathbf{x} \in [0, \infty]^n$. In several cases we will consider $\mathbf{x} \in [0, 1]^n$ only. Concerning measures, we will deal either with monotone measures $\mu : 2^N \to [0, \infty]$, $\mu(\emptyset) = 0$, $\mu(N) > 0$ and $\mu(U) \leq \mu(V)$ whenever $U \subseteq V \subseteq N$, or with normed monotone measures constraint by the boundary condition $\mu(N) = 1$. The later measures will be called capacities. For the interested scholars, we recommend monographs [2, 6, 7, 22, 28].

The tutorial is organized as follows. In its first part, after a short historical overview, we introduce and discuss basic integrals with respect to monotone measures, namely the Choquet, Shilkret and Sugeno integrals, see Section 2. Next, we discuss the universal integrals in Section 3 and decomposition integrals in Section 4. Section 5 brings several aggregation functions based on introduced integrals, including OWA, OMA and other operators. In Section 6, some generalizations are highlighted and some concluding remarks are added.

2 BASIC INTEGRALS

Definition 1. Let $(X, \mathcal{A})$ be a measurable space, μ a monotone measure on $\mathcal{A}$ and $f : X \to [0, \infty]$ a measurable function.

(i) The *Choquet integral* [10] **Ch** is given by

$$\mathbf{Ch}_\mu(f) = \int_0^\infty \mu(\{f \geq t\})\, dt, \tag{2.1}$$

where the integral on the right–hand side is the (improper) Riemann integral.

(ii) The *Shilkret integral* [23] **Sh** is given by

$$\mathbf{Sh}_\mu(f) = \sup\{\, t : \mu(\{f \geq t\}) | t \in [0, \infty]\}, \tag{2.2}$$

where the standard convention $0 \cdot \infty = \infty \cdot 0 = 0$ is used.

(iii) The *Sugeno integral* [25] **Su** is given by

$$\mathbf{Su}_\mu(f) = \sup\{\min(\, t, \mu(\{f \geq t\}))\, | t \in [0, \infty]\}. \tag{2.3}$$

Considering the finite space N and a function f in the vector form $\mathbf{x} \in [0, \infty]^n$, let $(\cdot) : N \to N$ be a permutation such that $x_{(1)} \leq ... \leq x_{(n)}$ (observe that (.) need not to be unique). Then the above integrals can be rewritten in the next form, denoting $E_{(i)} = \{(i), ..., (n)\}$:

$$\mathbf{Ch}_\mu(\mathbf{x}) = \sum_{i=1}^{n} (x_{(i)} - x_{(i-1)}) \cdot \mu(E_{(i)}) \tag{2.4}$$

ISBN: 978–83–8012–519–3

with convention $x_{(0)} = 0$;

$$\mathbf{Sh}_\mu(\mathbf{x}) = \max\left\{ x_{(i)} \cdot \mu(E_{(i)}) | i \in N \right\}, \qquad (2.5)$$

$$\mathbf{Su}_\mu(\mathbf{x}) = \max\left\{ \min(x_{(i)}, \mu(E_{(i)})) | i \in N \right\}. \qquad (2.6)$$

We have several alternative formulae for these three integrals. For example, the Lovász extension [16] based on the Möbius transform $M^\mu : 2^N \to R$,

$$M^\mu(A) = \sum_{B \subseteq A} (-1)^{\text{card}(A \setminus B)} \cdot \mu(B),$$

leads to the Chateaunef – Jaffray evaluation formula [9]

$$\mathbf{Ch}_\mu(f) = \sum_{A \subseteq N} M^\mu(A) \cdot \min\{f(x_i) \mid i \in A\}. \qquad (2.7)$$

For the Sugeno integral, we have also the next formula due to [12]:

$$\mathbf{Su}_\mu(\mathbf{x}) = \min\left\{ \max(x_{(i)}, \mu(E_{(i)})) | i \in N \right\}. \qquad (2.8)$$

Observe that all these integrals can be viewed as aggregation functions (note that the Sugeno integral $\mathbf{Su}_\mu$ is bounded by $\mu(N)$). In particular, when μ is a capacity, $\mathbf{Ch}_\mu$, $\mathbf{Sh}_\mu$, $\mathbf{Su}_\mu : [0,1]^n \to [0,1]$ are special idempotent aggregation functions characterized by the next settings of properties (axioms):

- $\mathbf{Ch}_\mu$ is comonotone additive;
- $\mathbf{Sh}_\mu$ is comonotone maxitive and positively homogeneous;
- $\mathbf{Su}_\mu$ is comonotone maxitive and min–homogeneous.

3 UNIVERSAL INTEGRALS ON $[0,1]$

Though universal integrals [14] were introduced for any measurable space $(X, \mathcal{A})$, any monotone measure μ on $\mathcal{A}$ and any $\mathcal{A}$–measurable function $f : X \to [0,\infty]$, we restrict our considerations to finite spaces N, vectors $\mathbf{x} \in [0,1]^n$ and capacities μ on N. Denote by $\mathcal{M}_\mu$ the set of all capacities on N.

Definition 2. A function $\mathcal{I} : \sum_{n\in\mathbb{N}} (\mathcal{M}_n \times [0,1]^n) \to [0,1]$ is called a $[0,1]$–*valued discrete universal integral* if it satisfies the following axioms:

(A1) $\mathcal{I}$ is non–decreasing in each component;

(A2) $\mathcal{I}(\mu, \mathbf{1}_E) = \mu(E)$ for all $N = \{1,2,...,n\}$, $\mu \in \mathcal{M}_n$, and $E \subseteq N$;

(A3) $\mathcal{I}(\mu, c \cdot \mathbf{1}_N) = c$ for all $N = \{1,2,...,n\}$, $\mu \in \mathcal{M}_n$, $E \subseteq N$ and $c \in [0,1]$;

(A4) $\mathcal{I}(\mu_1, f_1) = \mathcal{I}(\mu_2, f_2)$ for all pairs $(\mu_1, f_1) \in \mathcal{M}_{n_1} \times \mathcal{F}_{n_1}$ and $(\mu_2, f_2) \in \mathcal{M}_{n_2} \times \mathcal{F}_{n_2}$ satisfying $\mu_1(\{f_1 \geq t\}) = \mu_2(\{f_2 \geq t\})$ for each $t \in [0,1]$.

It is not difficult to check that, as a consequence of (A4), the value $\mathcal{I}(\mu, c \cdot \mathbf{1}_E)$ depends on the constant $c \in [0,1]$ and the value $\mu(E) \in [0,1]$, i.e., there is an operation $\otimes : [0,1]^2 \to [0,1]$ such that, for each $c \in [0,1]$ and each $E \subseteq \{1,2,...,n\}$ we have $\mathcal{I}(\mu, c \cdot \mathbf{1}_E) = c \otimes \mu(E)$. This operations $\otimes$ turns out to be a semicopula [1], i.e., it is non–decreasing in each component and has 1 as neutral element.

Proposition 1. Let $\mathcal{I}$ be a $[0,1]$–valued discrete universal integral. Then there exists a semicopula $\otimes$ such that we have $\mathcal{I}(\mu, c \cdot \mathbf{1}_E) = c \otimes \mu(E)$ for all $n \in \mathbb{N}, \mu \in \mathcal{M}_n$, $c \in [0,1]$, and $E \subseteq \{1,2,...,n\}$.

Obviously, **Ch**, **Sh** and **Su** are universal integrals. We recall two classes of universal integrals on $[0,1]$.

a) Copula–based $[0,1]$**–valued discrete universal integrals.**

Copulas were introduced in [24] to model the dependence structure of random vectors (for a detailed treatise see [21]). Here we restrict ourselves to two–dimensional copulas only which also can be considered as special binary aggregation functions or, more precisely, as special semicopulas.

Definition 3. A function $C : [0,1]^2 \to [0,1]$ is called a (2–dimensional) copula if it is a 2–increasing semicopula, i.e., for all $x, y, x^*, y^* \in [0,1]$ with $x \leq x^*$ and $y \leq y^*$ we have

$$C(x^*, y^*) - C(x^*, y) - C(x, y^*) + C(x, y) \geq 0.$$

Note that copulas are in a one–to–one correspondence with probability measures on the Borel subsets $\mathcal{B}([0,1]^2)$ with uniform marginals, i.e., for each copula $C : [0,1]^2 \to [0,1]$ there is a unique probability measure $P_C : \mathcal{B}([0,1]^2) \to [0,1]$ satisfying $P_C([0,a] \times [0,b]) = C(a,b)$. Based on ideas in [11, 13] (see also [14]), for a given copula C the copula–based integral $\mathcal{I}_C$ was introduced as

$$\mathcal{I}_C(\mu, \mathbf{x}) = P_C\left(\left\{(x,y) \in [0,1]^2 | y \leq \mu\left(\{i \in N | x_i \geq x\}\right)\right\}\right). \qquad (3.1)$$

Comparing with formula (2.4) one can see that for the *independence copula* Π we rediscover the Choquet integral, i.e., $\mathcal{I}_\Pi = \mathbf{Ch}$, whereas for the greatest copula $\wedge$ we obtain the Sugeno integral, i.e., $\mathcal{I}_\wedge = \mathbf{Su}$. Alternatively, one can write

$$\mathcal{I}_C(\mu, \mathbf{x}) = \sum_{i=1}^{n} \left(C(x_{(i)}, \mu(E_{(i)}) - C(x_{(i-1)}, \mu(E_{(i)})\right). \qquad (3.2)$$

b) Weakest semicopula based universal integrals.

For any semicopula $\otimes : [0,1]^2 \to [0,1]$, the weakest $[0,1]$–valued discrete universal integral $\mathcal{I}_\otimes$ related to $\otimes$ is given by

$$\mathcal{I}_{(\otimes)}(\mu, \mathbf{x}) = \max\left\{x_{(i)} \otimes \mu(E_{(i)}) | i \in N\right\}. \quad (3.3)$$

Clearly, $\mathcal{I}_{(1)} = \mathbf{Su}$ and $\mathcal{I}_{(\cdot)} = \mathbf{Sh}$.

4 DECOMPOSITION INTEGRALS

Recall that the Riemann and Lebesgue integrals are based on decomposition approaches generalized by Even and Lehrer in [4], see also [20]. Finite subsets of 2^N are called collections. Any system $\mathcal{H}$ of collections is called a decomposition system.

Definition 4. Let $\mathcal{H}$ be a decomposition system. The mappings $\mathcal{I}_\mathcal{H}$, $\mathcal{I}^\mathcal{H} : \mathcal{M}_n \times [0,1]^n \to [0,\infty[$ given by

$$\mathcal{I}_\mathcal{H}(\mu, \mathbf{x}) = \max\Big\{\textstyle\sum_{i=1}^k a_i \cdot \mu(A_i) \,|\, (A_1, ..., A_k) \in \mathcal{H},$$

$$a_1, ..., a_k \geq 0, \textstyle\sum_{i=1}^k a_i \cdot \mathbf{1}_{A_i} \leq \mathbf{x}\Big\}$$

and

$$\mathcal{I}^\mathcal{H}(\mu, \mathbf{x}) = \min\Big\{\textstyle\sum_{i=1}^k a_i \cdot \mu(A_i) \,|\, (A_1, ..., A_k) \in \mathcal{H},$$

$$a_1, ..., a_k \geq 0, \textstyle\sum_{i=1}^k a_i \cdot \mathbf{1}_{A_i} \geq \mathbf{x}\Big\}$$

are called $\mathcal{H}$–decomposition and $\mathcal{H}$–superdecomposition integrals, respectively.

Decomposition integrals can be defined for monotone measures and vectors from $[0,\infty]^n$, too.

Define:

$$\mathcal{H}_1 = \{\{A\} | \emptyset \neq A \subseteq N\},$$

$$\mathcal{H}_2 = \left\{(A_1, ..., A_n) \in (2^N)^n | (A_1, ..., A_n) \text{ is a chain}\right\},$$

$$\mathcal{H}_3 = \left\{(A_1, ..., A_n) \in (2^N)^n | A_i \cap A_j = \emptyset \text{ whenever } i \neq j\right\},$$

$$\mathcal{H}_4 = 2^{2^N}.$$

Then:

$\mathcal{I}_{\mathcal{H}_1} = \mathbf{Sh}$ is the Shilkret integral;

$\mathcal{I}^{\mathcal{H}_2} = \mathcal{I}_{\mathcal{H}_2} = \mathbf{Ch}$ is the Choquet integral;

$\mathcal{I}_{\mathcal{H}_3}$ is the PAN integral [30];

$\mathcal{I}_{\mathcal{H}_4}$ is the concave integral [15];

$\mathcal{I}^{\mathcal{H}_4}$ is the convex integral [20].

For more details, especially for a complete description of integrals which are both universal and decomposition (universal and superdecomposition) see [19].

5 AGGREGATION OPERATORS BASED ON INTEGRALS

Well known weighted arithmetic mean given by

$$W_\mathbf{w}(\mathbf{x}) = \sum_{i=1}^n w_i x_i,$$

where $\mathbf{w} = (w_1, ..., w_n) \in [0,1]^n$ is a normed weighting vector (it can be seen as a discrete probability on N) is the standard expected value of $\mathbf{x}$ with respect to the probability measure $\mu_\mathbf{w} : 2^N \to [0,1]$ given by $\mu_\mathbf{w}(E) = \sum_{i\in E} w_i$. It can be seen also as the Choquet integral, $W_\mathbf{w} = \mathbf{Ch}_{\mu_\mathbf{w}}$.

One class of the most applied aggregation functions is formed by OWA operators introduced by Yager [29],

$$\text{OWA}_\mathbf{w}(\mathbf{x}) = \sum_{i=1}^n w_i x_{(i)}. \quad (5.1)$$

Based on [5], $\text{OWA}_\mathbf{w} = \mathbf{Ch}_{\mu_{(\mathbf{w})}}$ is the Choquet integral with respect to a capacity $\mu_{(\mathbf{w})} : 2^N \to [0,1]$ given by $\mu_{(\mathbf{w})}(E) = \sum_{i=1}^{\text{card } E} w_{n-i+1}$. Observe that $\mu_{(\mathbf{w})}$ is called a symmetric capacity and it is characterized by $\mu_{(\mathbf{w})}(E) = \mu_{(\mathbf{w})}(F)$ whenever card E = card F.

Recently, OMA operators [18] were characterized as n–ary aggregation functions on $[0,1]$ which are symmetric and comonotone modular. Observe that OMA operators coincide with copula–based integrals with respect to symmetric capacities, OMA=$\mathcal{I}_C(\mu_{(\mathbf{w})}, \cdot)$.

In particular, for the greatest copula $\wedge$, i.e., considering the Sugeno integral, OWMax (Ordered Weighted Maximum) operators introduced by Dubois and Prade [3] are rediscovered.

Observe that any introduced integrals with respect to a symmetric capacity can be seen as a generalization of OWA operators.

6 CONCLUDING REMARKS

There are several fresh generalizations of integral–based aggregation functions introduced and discussed in the previous sections. We recall only superadditive and subadditive constructions generalizing decomposition and superdecomposition integrals [8, 26]. Another interesting approach generalizes formula (2.4) for the Choquet integral considering the function $C_\mu^G : [0,1]^n \to [0,\infty]$ given by

$$C_\mu^G(\mathbf{x}) = \sum_{i=1}^n G(x_{(i)} - x_{(i-1)}, \mu(E_{(i)}),$$

where $G : [0,1]^2 \to [0,1]$ is an appropriate function. Obviusly $C_\mu^\Pi = \mathbf{Ch}_\mu$ is the Choquet integral. For more details see [17].

Acknowledgements

The support of the grant VEGA 1/0420/15 is kindly announced.

References

[1] B. Bassan, F. Spizzichino (2005). Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. *Journal of Multivariate Analysis* **93 (2)**, pp. 313–339.

[2] D. Denneberg (1997). *Non-additive Measure and Integral, Second Edition.* Dordrecht, Kluwer Academic Publishers.

[3] D. Dubois, H. Prade (1985). A review of fuzzy set aggregation connectives. *Information Sciences* **36**, pp. 85–121.

[4] Y. Even, E. Lehrer (2014). Decomposition-integral: unifying Choquet and the concave integrals. *Economic Theory* **56**, pp. 33-58.

[5] M. Grabisch (1995). Fuzzy integral in multicriteria decision making. *Fuzzy Sets and Systems* **69**, pp. 279–298.

[6] M. Grabisch, T. Murofushi, M. Sugeno, and J. Kacprzyk, editors (2000). *Fuzzy Measures and Integrals. Theory and Applications* **40**, Physica Verlag, Berlin.

[7] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap (2009). *Aggregation Functions: Encyclopedia of Mathematics.* Cambridge University Press.

[8] S. Greco, R. Mesiar, F. Rindone, L. Šipeky (2015). The superadditive and the subadditive transformations of integrals and aggregation functions. Submitted.

[9] A. Chateauneuf, J.-Y. Jaffray (1989). Some Characterizations of Lower Probabilities and Other Monotone Capacities through the Use of Mobius Inversion. *Mathematical Social Sciences* **17**, pp. 263–283.

[10] G. Choquet (1954). Theory of capacities. *Ann. Inst. Fourier (Grenoble)* **5**, pp. 131–295.

[11] H. Imaoka (1997). On a Subjective Evaluation Model by a Generalized Fuzzy Integral. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **5(5)**, pp. 517–530.

[12] A. Kandel, W.-J. Byatt (1978). Fuzzy differential equations. In: Proceedings of the International Conference on Cybernetics and Society, Vol. I, II (Tokyo/Kyoto, 1978), pp. 1213-1216, IEEE, New York.

[13] E.P. Klement, R. Mesiar, E. Pap (2004). Measure-based aggregation operators. *Fuzzy Sets and Systems* **142 (1)**, pp. 3–14.

[14] E.P. Klement, R. Mesiar, E. Pap (2010). A universal integral as common frame for Choquet and Sugeno integral. *IEEE Transactions on Fuzzy Systems* **18 (1)**, art. no. 5361437, pp. 178–187.

[15] E. Lehrer (2009). A new integral for capacities. *Economic Theory* **39**, pp. 157–176

[16] L. Lovász (1983). *Submodular functions and convexity.* Mathematical Programming The State of the Art. Springer, pp. 235–257

[17] G. Lucca, J.A. Sanzi, G.P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince (2015). Pre–aggregation functions. Submitted.

[18] R. Mesiar, A. Mesiarová–Zemánková (2011). The ordered modular averages. *IEEE Transactions on Fuzzy Systems* **19 (1)**, art. no. 5580045, pp. 42–50.

[19] R. Mesiar, A. Stupňanová (2013). Decomposition integrals. *International Journal of Approximate Reasoning* **54 (8)**, pp. 1252–1259.

[20] R. Mesiar, J. Li, E. Pap (2015). Superdecomposition integrals. *Fuzzy Sets and Systems* **259**, pp. 3–11.

[21] R.B. Nelsen (2006). *An introduction to copulas.* Second Edition. Springer Series in Statistics, Springer-Verlag, New York.

[22] E. Pap, editor (2002). *Handbook of Measure Theory I, II.* Elsevier Science, Amsterdam.

[23] N. Shilkret (1971). Maxitive measure and integration. *Indag. Math.* **33**, 109-116.

[24] A. Sklar (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, pp. 229–231.

[25] M. Sugeno (1974). *Theory of fuzzy integrals and its applications.* Ph.D. thesis. Tokyo Institute of Technology.

[26] A. Šipošová, L. Šipeky, F. Rindone (2015). Super– and subadditive constructions of aggregation functions. In: Proceedings AGOP 2015, in this volume.

[27] G. Vitali (1925). Sulla definizione di integrale delle funzioni di una variabile. *Annali di Matematica Pura ed Applicata* **2 (1)**, pp. 111–121.

[28] Z. Wang, G.-J. Klir (2009). Generalized Measure Theory. *IFSR International Series on Systems Science and Engineering* **25**, Springer, New York.

[29] R.R. Yager (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *Systems, Man and Cybernetics, IEEE Transactions on* **18 (1)**, pp. 183–190.

[30] Q. Yang (1985). The Pan-integral on the Fuzzy Measure Space. *Fuzzy Mathematics* (in Chinese) **3**, pp. 107–114.

Invited Talks

ON CONSISTENT OPERATOR SYSTEMS

József Dombi
University of Szeged, Institute of Informatics
dombi@inf.u-szeged.hu

Combining several numerical values into a single value is called aggregation and the numerical function is called the aggregation function. From this, functions can be expressed as aggregation functions. Usually there use some important restrictions of the functions. Most studies that deal with aggregation function focus on a problem on how we can establish a general representation theorem for the aggregator function when it has certain explicitly defined properties. Because the universal operator does not exist, for each step of are algorithm we have to choose another one. When we have consistent system we have a set of operators and developing an algorithm seems easier to do. All these problems are invited by fuzzy theory. Now let us summarize the elements of this concept: Membership functions, intersections (conjunctive operator, t-norm), unions (disjunctive operator, t-conorm), complementary operation (negation), inclusions (implications), symmetric difference, equivalence, and so on. The operator work on membership values and other functions can also be defined. For instance, the aggregative operators (uninorms) or modifiers. When examining continuous-valued operators, the key question is how we should choose them in a consistency way. Here, consistency means that the DeMorgan law is valid and for the uninorms the self-DeMorgan law holds. Or the distributive property is valid for unary operators with conjunctive/disjunctive operators. We should also mention that the membership function should be consistent with the operators of the above systems. In our different functional equations that are used to verify the consistency of our system. We will call it the Pliant operators. The operators are examined and the membership function is developed. We give a new semantic meaning to this type of membership function and we will call them distending functions. Distending functions and the Pliant system together form a very effective tool for solving practical problems. Not only are the strict monstrously increasing operator in the focus our research, although the Łukasiewicz type operators too. We call this operator class bounded system. At the end of our talk we give some examples for approximating functions, clustering, developing decision trees and neural networks. From a scientific point of view, we present a theorem on the convergence of the sharpness measure. The better the input, the better the output.

References

[1] G. Beliakov, Gleb, A. Pradera, T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer, Berlin 2007.

[2] J. Dombi, DeMorgan systems with an infinitely many negations in the strict monotone operator case, *Information Sciences* 181:1440-1453, 2011.

[3] J. Dombi, Modalities, In: P. Melo-Pinto, P. Couto, C. Serodio, B. De Baets, *Advances in intelligent soft computing*, Workshop on fuzzy methods for knowledge-based systems. Berlin; Heidelberg: Springer 2011, pp. 52-65.

[4] J. Dombi, Pliant Operator System, In: J. Fodor et al. (eds), *Recent Advances in Intelligent Engineering Systems* (Studies in Computational Intelligence 378) Springer-Verlag, Berlin Heidelberg 2012, pp. 31-58.

[5] J. Dombi, On a certain class of aggregative operators, *Information Sciences* 245:313-328, 2013.

[6] J. Dombi, Bayes Theorem, Uninorms and Aggregating Expert Opinions, In: H. Bustince et al. (eds.), *Aggregation Functions in Theory and in Practise*, Proceedings of the 7th International Summer School on Aggregation Operators at the Public University of Navarra, Pamplona, Spain, July 16-20, 2013 (Advances in Intelligent Systems and Computing 228), Springer-Verlag, Berlin Heidelberg 2013, pp. 281-291.

ISBN: 978–83–8012–519–3

[7] J. Dombi, Unary Operators, In: N.E. Mastorakis et al. (eds.), *Advances in Applied and Pure Mathematics*, Proceedings of the 2014 International Conference on Pure Mathematics, Apllied Mathematics, Computational Methods (PMAMCM 2014), Santorini Island, Greece, July 17-21, 2014, Sofia University Press 2014, pp. 91-97.

[8] J. Dombi, O. Csiszár, Implications in bounded systems, *Information Sciences* 283:229-240, 2014.

[9] J. Dombi, O. Csiszár, The general nilpotent operator system, *Fuzzy Sets and Systems* 261:119, 2015.

[10] J. Dombi, O. Csiszár, Equivalences in bounded systems, *Fuzzy Sets and System*, submitted.

[11] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions.* Encyclopedia of Mathematics and its Applications, vol. 127. Cambridge University Press, Cambridge 2009.

FUZZY IMPLICATIONS: SOME ALGEBRAIC PERSPECTIVES

Balasubramaniam Jayaram
Deparment of Mathematics,
Indian Institute of Technology Hyderabad,
Telangana - 502 205, INDIA
jbala@iith.ac.in

Fuzzy Logic Connectives can be seen as a generalisation of the classical logic connectives. The septuple of $(\{0,1\}, \wedge, \vee, \rightarrow, \neg, 0, 1)$ corresponding to classical logic has been investigated along the following aspects: Logical, Algebraic and Applicational. Considering the septuple of $([0,1], T, S, I, N, 0, 1)$ corresponding to fuzzy logic, other than the above, one can also discuss their analytical and probabilistic aspects.

Fuzzy Implications I generalise the classical implication on $\{0,1\}$ to the extended truth value set $[0,1]$. The (i) **analytic** aspects of fuzzy implications, viz., the different generation process, families, properties, intersections among them, their characterisations and representations, functional equations they satisfy, etc., and their (ii) **applicational** aspects, viz., their role in approximate reasoning, image processing, data mining, etc., have been very well studied and documented.

The various studies on the algebraic aspects of fuzzy implications can be categorised as below.

1. Fuzzy Implications on different Algebras

In the early days of research on fuzzy logic connectives, the domain of these functions was taken to be the unit interval $[0,1]$, corresponding to the generalisation of the truth value set from the binary $\{0,1\}$. However, with the increasing awareness of their use in many practical applications and towards ensuring computational feasibility, researchers began to discretise the underlying domain. This also presented an opportunity to impose different algebras on the underlying discrete domain other than the linear order that was already available on $[0,1]$.

Among the algebraic exploration of fuzzy implications, the above approach has garnered the maximum attention.

In this talk, we will present works done on fuzzy implications from an algebraic perspective along the following lines which have been relatively unexplored and largely recent.

2. Algebras based on Fuzzy Implications

A Fuzzy Implication algebra (FI-algebra) [2] is an algebra $(X, \rightarrow, 0)$, where X is a non-empty set with a special element $0 \in X$, satisfying some axioms. We will see how FI-algebras are related to many established t-norm based logics and their consequent algebras.

3. Algebras on Fuzzy Implications

Among the many generation processes of fuzzy implications, those that generate a fuzzy implication K from a given pair of fuzzy implications (I, J) can be seen as a binary operation $\odot$ on the set of all fuzzy implications $\mathbb{I}$. Thus it is interesting to study the algebraic structures, if any, on $(\mathbb{I}, \odot)$. Many such generation methods, in particular, the point wise min and max of fuzzy implications, compositions of fuzzy implications (see [1], Chap. **6**), and the recently proposed $\circledast$-composition of fuzzy implications [3], can be viewed in this framework. An algebraic study of these structures have led to some interesting consequences. For instance, hitherto unknown representations of some families of fuzzy implications were obtained from such studies.

References

[1] Baczyński, M., Jayaram, B.: Fuzzy implications (Studies in Fuzziness and Soft Computing, vol. 231), Springer, Berlin 2008.

[2] D. Pei (2014). A survey of fuzzy implication algebras and their axiomatization. *Int. J. Approx. Reasoning* **55**, pp. 1643–1658.

[3] Vemuri, N. R., Jayaram, B. (2014). Representations through a monoid on the set of fuzzy implications. *Fuzzy Sets and Systems* **247** pp. 51–67.

ISBN: 978–83–8012–519–3

A CONVERSE TO FUBINI THEOREM AND A DESCRIPTION OF MEANS

Maciej Sablik
Institute of Mathematics, University of Silesia
40-007 Katowice, ul. Bankowa 14, Poland
maciej.sablik@us.edu.pl

Summary

We intend to point out that bisymmetry is nothing but classical Fubini theorem in a very peculiar case. We show some applications in economy, and then proceed to prove a characterization of bisymmetric means in some function spaces. In particular, we give a form of bisymmetric operations in L^1 and $C(\Omega)$.

Keywords: Mean, bisymmetry, iterated integral.

1 Introduction

By "mean" we understand a functional M, mapping the square of a real interval X into $\mathbb{R}$ and such that for for any $x, y \in X$ one has

$$\min(x, y) \leq M(x, y) \leq \max(x, y).$$

More generally, if we fix an $n \in N$, $n \geq 2$, then we can think of mean as a functional $M : X^n \longrightarrow \mathbb{R}$ satisfying

$$\min(x_1, \ldots, x_n) \leq M(x_1, \ldots, x_n) \leq \max(x_1, \ldots, x_n).$$

There is also another concept of mean coming form harmonic analysis. Let us fix a semigroup G and consider the family $B(G, \mathbb{R})$ of all real valued and bounded functions defined in G. The (invariant) mean M is a functional defined on $B(G, \mathbb{R})$ and

a) linear,

b) $M(f) \leq \sup f(G)$,

c) invariant (i.e. $M(f_a) = M(f)$ for every $a \in G$, where $f_a(x) = f(xa)$, $x \in G$).

We could look at mean values of n elements, as means defined on $B(G, \mathbb{R})$ where G stands for $\mathbb{Z}_n$. It turns out that in this case M is of the form

$$M(x_0, \ldots, x_{n-1}) = \sum_{i=0}^{n-1} w_i x_i,$$

for some weights w_i, $i \in \{0, \ldots, n-1\}$ (this fact follows from linearity (a)) and continuity being a consequence of b)). From b) we get reflexivity of M, i.e.

$$M(c \cdot \mathbf{1}) = c$$

for every $c \in \mathbb{R}$. It also follows from b) that

$$f \geq 0 \Longrightarrow M(f) \geq 0.$$

Hence $\sum_{i=0}^{n-1} w_i = 1$, and $w_i \geq 0$ for every $i \in \{0, \ldots, n-1\}$. Now, condition c) implies that all the weights are equal, i.e.

$$w_i = \frac{1}{n},\ i \in \{0, \ldots, n-1\}.$$

In other words, the only mean satisfying conditions a) – c) is the arithmetic mean.

In order to determine other, not necessarily arithmetic means, we have to skip condition c) and replace linearity (a)) by *bisymmetry.* Bisymmetry is the following property of operations. Suppose that a non-empty set X is given. In the case $n = 2$ imagine the matrix

$$\begin{bmatrix} x & y \\ u & v \end{bmatrix}$$

with entries belonging to a set X and let us aggregate lines $M(x, y)$, $M(u, v)$ and then columns $M(x, u)$, $M(y, v)$. Then aggregate the results again to obtain

$$M(M(x, y), M(u, v))$$

and

$$M(M(x, u),\ M(y, v)).$$

ISBN: 978–83–8012–519–3

Now, if the final aggregations agree for all $x, y, u, v \in X$ or the equality

$$M\left[M\left(x,y\right), M\left(u,v\right)\right] = M\left[M\left(x,u\right), M\left(y,v\right)\right]$$

holds, we say that the operation M is bisymmetric. For bisymmetric means we have the following result by J. Aczél [1]

Theorem 1 (Aczél [1], p. 281) *The quasi–arithmetic mean*

$$M(x,y) = \varphi^{-1}\left(\frac{\varphi(x)+\varphi(y)}{2}\right)$$

is the general continuous, on both sides reducible, real solution of

$$M\left[M\left(x_{11},x_{12}\right), M\left(x_{21},x_{22}\right)\right] = M\left[M\left(x_{11},x_{21}\right), M\left(x_{12},x_{22}\right)\right]$$

under the additional conditions $x,\ y,\ M(x,y) \in [a,b]$,

$$M(x,x) = x \quad \text{for all } x \in [a,b],$$

$$M(x,y) = M(y,x) \quad \text{for all } x,\ y \in [a,b].$$

(*Without symmetry M is given by*

$$M(x,y) = \varphi^{-1}\left((1-q)\varphi(x) + q\varphi(y)\right)$$

where $q \in \mathbb{R} \setminus \{0,1\}$.)

Remark 1 *Without reflexivity and symmetry M is given by*

$$M(x,y) = \varphi^{-1}\left((\alpha\varphi(x) + \beta\varphi(y) + \gamma\right)$$

for some $\alpha > 0, \beta > 0$ *and* $\gamma \in \mathbb{R}$ *such that*

$$u, v \in [a,b] \Longrightarrow \alpha u + \beta v + \gamma \in [a,b].$$

(J. Aczél [1], Gy. Maksa [6], P. Volkmann [9]).

In the case of arbitrary $n \in \mathbb{N}, n \geq 2$, we have the following result

Theorem 2 (Münnich, Maksa, Mokken [8]) *Let* $M : I^n \to I$ $(n \geq 2$ *fixed) be strictly monotone increasing in each of its arguments, continuous, and satisfying*

$$M(M(x_{11},\ldots,x_{1n}),\ldots,M(x_{n1},\ldots,x_{nn})) = M(M(x_{11},\ldots,x_{n1}),\ldots,M(x_{1n},\ldots,x_{nn}))$$

$$M(x,\ldots,x) = x, \quad x \in I.$$

Then there are strictly monotone increasing, continuous functions $\varphi : I \to \mathbb{R}$ *and constants* $\alpha_1,\ldots,\alpha_n \in (0,1)$ *such that*

$$\sum_{i=1}^{n} \alpha_i = 1$$

and

$$M(x_1,\ldots,x_n) = \varphi^{-1}\left(\sum_{i=1}^{n}\alpha_i\varphi(x_i)\right).$$

1.1 Consistent aggregation.

Define

$$\mathbf{x}_i = (x_{i1},\ldots,x_{in}),\ \mathbf{x}^k = (x_{1k},\ldots,x_{mk}),$$

for all $i \in \{1,\ldots,m\}$ and $k \in \{1,\ldots,n\}$.

Table 1: Consistent aggregation of production.

Goods / Producers	Inputs (goods and services)			(Maximal) outputs
	1	…	m	(Production functions)
P_1	$x_{1,1}$	…	$x_{m,1}$	$F_1\left(\mathbf{x^1}\right)$
⋮	⋮	…	⋮	⋮
⋮	⋮	…	⋮	⋮
⋮	⋮	…	⋮	⋮
P_n	$x_{1,n}$	…	$x_{m,n}$	$F_n\left(\mathbf{x^n}\right)$
	$G_1\left(\mathbf{x_1}\right)$	…	$G_m\left(\mathbf{x_m}\right)$	$F\left(G_1\left(\mathbf{x_1}\right),\ldots,G_m\left(\mathbf{x_m}\right)\right)$ / $G\left(F_1\left(\mathbf{x^1}\right),\ldots,F_n\left(\mathbf{x^n}\right)\right)$

$$G\left(F_1\left(x_{1,1},\ldots,x_{m,1}\right),\ldots,F_n\left(x_{1,n},\ldots,x_{m,n}\right)\right) = F\left(G_1\left(x_{1,1},\ldots,x_{1,n}\right),\ldots,G_m\left(x_{m,1},\ldots,x_{m,n}\right)\right)$$

$$G\left(F\left(1,x(\cdot,1)\right),\ldots,F\left(n,x(\cdot,n)\right)\right) = F\left(G\left(1,x(1,\cdot)\right),\ldots,G\left(m,x(m,\cdot)\right)\right).$$

1.2 Example.

Take $y : A = \{1, \dots, m\} \longrightarrow \mathbb{R}$ and assume that

$$F(y) = F(1, y) = F(2, y) = \dots = F(n, y) = \int_A y(s) ds$$

Similarly, let $B = \{1, \dots, n\}$ and let $z : B \longrightarrow \mathbb{R}$. Assume that

$$G(z) = G(1, z) = \dots = G(m, z) = \int_B z(t) dt.$$

Then the above equation becomes

$$G\left(\underbrace{F(x(\cdot, 1)), \dots, F(x(\cdot, n))}_{n}\right) = F\left(\underbrace{G(x(1, \cdot)), \dots, G(x(m, \cdot))}_{m}\right).$$

or

$$\int_B \left(\int_A x(s,t) ds\right) dt = \int_A \left(\int_B x(s,t) dt\right) ds$$

1.3 Distributions.

Let $A, B \in \mathbb{R}$, $A < B$ and denote by $\mathcal{D}(A, B)$ the family of all distribution functions $F : \mathbb{R} \longrightarrow [0, 1]$ such that

- F is constant in stretches and has only a finite number of discontinuities;
- $F(x) = \frac{1}{2}[F(x+) - F(x-)]$;
- $F(x) = 0$, $x \le A$ and $F(x) = 1$, $x \ge B$.

Note that for every $\xi \in [A, B]$ the function E_ξ given by $E_\xi(x) = \frac{1}{2}(1 + \mathrm{sgn}(x - \xi))$ belongs to $\mathcal{D}(A, B)$. The following was proved by B. de Finetti in [4] (cf. Hardy–Littlewood–Pólya [5, p. 158])

Theorem 3 (B. de Finetti, [4]) *Suppose that* $M : \mathcal{D}(A, B) \longrightarrow \mathbb{R}$ *satisfies* (a) $M(E_\xi) = \xi$, $\xi \in [A, B]$;

(b) *if* $F_1, F_2 \in \mathcal{D}(A, B)$, $F_1 \ge F_2$, *and* $F_1(x) > F_2(x)$ *for some* $x \in \mathbb{R}$, *then* $M(F_1) < M(F_2)$;

(c) *if* $F, F^*, G \in \mathcal{D}(A, B)$ *and* $M(F) = M(F^*)$, *then*

$$M(tF + (1 - t)G) = M(tF^* + (1 - t)G)$$

for $t \in (0, 1)$. *Then there is a function* ϕ, *continuous and strictly increasing in* $[A, B]$, *for which*

$$(\sharp) \qquad M(F) = \phi^{-1}\left(\int_{-\infty}^{\infty} \phi(x) dF(x)\right).$$

Conversely, if M *is defined by* ($\sharp$), *for a* ϕ *with the properties stated, then it satisfies (a), (b) and (c), so that these conditions are necessary and sufficient for the representation of* M *in the form* ($\sharp$).

2 General form of bisymmetrical mean

2.1 Definitions

Let $(\Omega, \mathcal{A}, \mu)$ be a triple consisting of a set $\Omega \neq \emptyset$, $\mathcal{A} \subset 2^\Omega$ – algebra of sets, and $\mu : \mathcal{A} \to [0, \infty]$ – an additive and nontrivial set function. Further, define $\mathcal{M} = \{f : \Omega \to \mathbb{R} : f \text{ is } \mathcal{A}\text{-measurable}\}$. If $f, g \in \mathcal{M}$ then

$$f \le g :\Longleftrightarrow \mu(\{\omega \in \Omega : f(\omega) > g(\omega)\}) = 0$$

Let us consider the following family of partitions of Ω.

$$\mathcal{P}(\Omega) = \bigcup_{n=1}^{\infty} \{(A_1, \dots, A_n) \in \mathcal{A}^n : \bigcup_{i=1}^{n} A_i = \Omega,\ i \neq j \Longrightarrow A_i \cap A_j = \emptyset\}$$

Introduce the family of simple functions.

$$\mathcal{M}_{fin} = \{f \in \mathcal{M} : f(\Omega) \text{ is finite}\} = \left\{\sum_{i=1}^{n} x_i \mathbf{1}_{A_i} : n \in \mathbb{N},\ (A_1, \dots, A_n) \in \mathcal{P}(\Omega)\right\}$$

Let $\mathcal{F} \subset \mathcal{M}$ be such that $\mathcal{M}_{fin} \subset \mathcal{F}$. We say that

- $M : \mathcal{F} \to \mathbb{R}$ is <u>reflexive</u>, iff
$$\bigwedge_{c \in \mathbb{R}} (M(c\mathbf{1}) = c).$$
- M is μ-strictly increasing, iff
$$\bigwedge_{f, g \in \mathcal{F}} (f \le g \text{ and } \mu(\{f < g\}) > 0 \Rightarrow M(f) < M(g)).$$

$$\mathcal{M}_{fin} \times \mathcal{M}_{fin} = \{\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} \mathbf{1}_{A_i \times B_j} : m, n \in \mathbb{N},\ (A_1, \dots, A_n), (B_1, \dots, B_m) \in \mathcal{P}(\Omega)\}$$

Remark 2 *If* $x \in \mathcal{M}_{fin} \times \mathcal{M}_{fin}$ *then*

$$M_{[t]}(x) := \text{"}\Omega \ni s \to M(x(s, \cdot)) \in \mathbb{R}\text{"} \in \mathcal{M}_{fin}$$

and

$$M_{[s]}(x) := \text{"}\Omega \ni t \to M(x(\cdot, t)) \in \mathbb{R}\text{"} \in \mathcal{M}_{fin}.$$

2.2 Results.

Lemma 1 *Suppose that* $(\mathcal{F}, \|\cdot\|)$ *is a normed space. If* $M : \mathcal{F} \to \mathbb{R}$ *is reflexive,* μ*-strictly increasing, continuous and*

$$M(M_{[s]}(x)) = M(M_{[t]}(x))$$

for every $x \in \mathcal{M}_{fin} \times \mathcal{M}_{fin}$ *then there exist a strictly increasing and continuous function* $\varphi : \mathbb{R} \to \mathbb{R}$ *and an additive function* $P : \mathcal{A} \to [0,1]$ *such that*

$$P(\Omega) = 1,$$

$$\bigwedge_{A \in \mathcal{A}} (P(A) > 0 \Longleftrightarrow \mu(A) > 0),$$

and

$$M(y) = \varphi^{-1}\left(\int_\Omega (\varphi \circ y) dP\right)$$

for every $y \in \mathcal{M}_{fin}$.

From the above lemma we obtain

Theorem 4 *Let* $\mu : 2^\Omega \to [0,\infty]$ *be an additive, non-trivial function. Then* $M : B(\Omega, \mathbb{R}) \to \mathbb{R}$ *is reflexive,* μ*-strictly increasing, continuous and satisfies*

$$M(M_{[s]}(x)) = M(M_{[t]}(x))$$

for every $x \in \mathcal{M}_{fin} \times \mathcal{M}_{fin}$*, if, and only if, there exist a strictly increasing and continuous function* $\varphi : \mathbb{R} \to \mathbb{R}$ *and an additive function* $P : \Omega \to [0,1]$ *such that*

$$P(\Omega) = 1,$$

$$\bigwedge_{A \subset \Omega} (P(A) > 0 \Longleftrightarrow \mu(A) > 0),$$

and

$$M(y) = \varphi^{-1}\left(\int_\Omega (\varphi \circ y) dP\right)$$

for every $y \in B(\Omega, \mathbb{R})$.

We obtain also results for integrable functions.

Theorem 5 *Let* $(\Omega, \mathcal{A}, \mu)$ *be a measure space with a finite* μ*. Then* $M : L^1(\Omega, \mathbb{R}) \to \mathbb{R}$ *is reflexive,* μ*-strictly increasing, continuous and satisfies*

$$M(M_{[s]}(x)) = M(M_{[t]}(x))$$

for every $x \in \mathcal{M}_{fin} \times \mathcal{M}_{fin}$*, if, and only if, there exist a strictly increasing and continuous function* $\varphi : \mathbb{R} \to \mathbb{R}$ *and a probability measure* $P : \mathcal{A} \to [0,1]$ *equivalent to* μ *such that*

$$M(y) = \varphi^{-1}\left(\int_\Omega (\varphi \circ y) dP\right)$$

for every $y \in L^1(\Omega, \mathbb{R})$.

2.3 Remarks about the proof.

Let m and n be positive integers. Fix partitions $(A_1, \dots, A_m) \in \mathcal{P}(\Omega_1)$ and $(B_1, \dots, B_n) \in \mathcal{P}(\Omega_2)$. Consider the function

$$x(s,t) = \sum_{i=1}^m \sum_{j=1}^n x_{ij} \mathbf{1}_{A_i \times B_j}(s,t) = \sum_{i=1}^m \left(\sum_{j=1}^n x_{ij} \mathbf{1}_{B_j}(t)\right) \mathbf{1}_{A_i}(s).$$

We have for every $i \in \{1, \dots, m\}$

$$s \in A_i \Longrightarrow x(s,\cdot) = \sum_{j=1}^n x_{ij} \mathbf{1}_{B_j}$$

whence

$$s \in A_i \Longrightarrow M\left(x(s,\cdot)\right) = M\left(\sum_{j=1}^n x_{ij} \mathbf{1}_{B_j}\right)$$

or

$$s \longrightarrow M\left(x(s,\cdot)\right) = \left(M\left(\sum_{j=1}^n x_{1j} \mathbf{1}_{B_j}\right), \dots, M\left(\sum_{j=1}^n x_{mj} \mathbf{1}_{B_j}\right)\right).$$

Similarly we obtain

$$t \longrightarrow M\left(x(\cdot,t)\right) = \left(M\left(\sum_{i=1}^m x_{i1} \mathbf{1}_{A_i}\right), \dots, M\left(\sum_{i=1}^m x_{in} \mathbf{1}_{A_i}\right)\right).$$

for every $t \in \Omega_2$.

Let I, J be real intervals and define
(1)

$$F(x_1, \dots, x_m) = M\left(\sum_{i=1}^m x_i \mathbf{1}_{A_i}\right), \ (x_1, \dots, x_m) \in I^m;$$

(2)

$$G(y_1, \dots, y_n) = M\left(\sum_{j=1}^n y_j \mathbf{1}_{B_j}\right), \ (y_1, \dots, y_n) \in J^n;$$

We easily check that

$$F\left(G(x_{11}, \dots, x_{1n}), \dots, G(x_{m1}, \dots, x_{mn})\right) = M\left(M\left(\sum_{i=1}^m \sum_{j=1}^n x_{ij} \mathbf{1}_{A_i \times B_j}\right)\right) = M\left(M\left(\sum_{i=1}^m \sum_{j=1}^n x_{ij} \mathbf{1}_{A_i \times B_j}\right)\right) = G\left(F(x_{11}, \dots, x_{m1}), \dots, F(x_{1n}, \dots, x_{mn})\right).$$

Moreover, the functions F, G are strictly increasing and continuous. The above equation is a particular case of

$$F\left(G_1(x_{11},\dots,x_{1n}),\dots,G_m(x_{m1},\dots,x_{mn})\right) = G\left(F_1(x_{11},\dots,x_{m1}),\dots,F_n(x_{1n},\dots,x_{mn})\right).$$

2.4 Maksa's theorem

Theorem 6 (Gy. Maksa, [6]) *Let $1 < n \in \mathbb{N}$, $1 < m \in \mathbb{N}$, X_{ij} be real intervals, $G_i : X_{1i} \times \cdots \times X_{mi} \to \mathbb{R}$, $G_i(X_{1i} \times \cdots \times X_{mi}) = I_i$, $F_j : X_{j1} \times \cdots \times X_{jn} \to \mathbb{R}$, $F_j(X_{j1} \times \cdots \times X_{jn}) = J_j$, G_i, $F_j \in CM$ for $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,m\}$, $G : J_1 \times \cdots \times J_m \to \mathbb{R}$, $F : I_1 \times \cdots \times I_n \to \mathbb{R}$, and G, $F \in CM$. Suppose that*

$$G(F_1(x_{11},\dots,x_{1n}),\dots,F_m(x_{m1},\dots,x_{mn})) = F(G_1(x_{11},\dots,x_{m1}),\dots,G_n(x_{1n},\dots,x_{mn}))$$

holds for all $x_{ij} \in X_{ij}$, $i \in \{1,\dots,n\}$ and $j \in \{1,\dots \dots,m\}$. Then there exist a real interval I and CM functions $\varphi : I \to \mathbb{R}$, $\alpha_i : I_i \to \mathbb{R}$, $\gamma_j : J_j \to \mathbb{R}$ and $\beta_{ji} : X_{ji} \to \mathbb{R}$, $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$ such that

$$F(z_1,\dots,z_n) = \varphi^{-1}\left(\sum_{i=1}^{n} \alpha_i(z_i)\right), \quad (z_1,\dots,z_n) \in I_1 \times \cdots \times I_n,$$

$$G(y_1,\dots,y_m) = \varphi^{-1}\left(\sum_{i=1}^{m} \gamma_j(y_j)\right), \quad (y_1,\dots,y_m) \in J_1 \times \cdots \times J_m,$$

$$F_j(x_{j1},\dots,x_{jn}) = \gamma_j^{-1}\left(\sum_{i=1}^{n} \beta_{ji}(x_{ji}\right),$$

$$G_i(x_{1i},\dots,x_{mi}) = \alpha_i^{-1}\left(\sum_{j=1}^{m} \beta_{ji}(x_{ji})\right)$$

for $x_{ji} \in X_{ji}$, $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$.

By the above we can proceed as follows.

- Fix $A \in \mathcal{A}$ and let $(A, \Omega \setminus A) \in \mathcal{P}(\Omega)$.
- Consider the function $F : I \times I \longrightarrow \mathbb{R}$ defined by
  $$F(x_1,x_2) = M\left(x_1\mathbf{1}_A + x_2\mathbf{1}_{\Omega\setminus A}\right),\ (x_1,x_2) \in I^2.$$
  F satisfies the bisymmetry equation, is continuous and strictly increasing.
- Therefore, by Theorem of Aczél or Münnich – Maksa – Mokken we get
  $$F(x_1,x_2) = \varphi^{-1}\left(\alpha_1\varphi(x_1) + \alpha_2\varphi(x_2)\right),$$
  for some strictly increasing function $\varphi : I \longrightarrow \mathbb{R}$, and some positive numbers α_i, $i = 1, 2$ summing up to 1.
- We define $P(A) := \alpha_1$. We have to check that
- P is well defined,
- P is additive.

2.5 P is well defined.

Suppose that

$$F(x,y) = \varphi^{-1}\left(\alpha_1\varphi(x) + \alpha_2\varphi(y)\right) = \psi^{-1}\left(\beta_1\psi(x) + \beta_2\psi(y)\right).$$

Denote $\gamma = \psi \circ \varphi^{-1}$. We get from the bisymmetry equation

$$\alpha_1\gamma^{-1}\left(\beta_1\gamma(u_{11}) + \beta_2\gamma(u_{12})\right) + \alpha_2\gamma^{-1}\left(\beta_1\gamma(u_{21}) + \beta_2\gamma(u_{22})\right) = \gamma^{-1}[\beta_1\gamma\left(\alpha_1u_{11} + \alpha_2u_{21}\right) + \beta_2\gamma\left(\alpha_1u_{12} + \alpha_2u_{22}\right)]$$

or, denoting $A_\gamma(u,v) = \gamma^{-1}\left(\beta_1\gamma(u) + \beta_2\gamma(u)\right)$ we obtain

$$\alpha_1A_\gamma\left(u_{11},u_{12}\right) + \alpha_2A_\gamma\left(u_{21},u_{22}\right) = A_\gamma\left(\alpha_1u_{11} + \alpha_2u_{21}, \alpha_1u_{12} + \alpha_2u_{22}\right).$$

Let $\mathbf{v} = (u_{11}, u_{12})$, $\mathbf{w} = (u_{21}, u_{22})$. Then we may write

$$\alpha_1A_\gamma(\mathbf{v}) + \alpha_2A_\gamma(\mathbf{w}) = A_\gamma(\alpha_1\mathbf{v} + \alpha_2\mathbf{w}),$$

whence

$$A_\gamma(\mathbf{v}) = <\mathbf{a}|\mathbf{v}> + b$$

for some $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ and $b \in \mathbb{R}$. In particular we get

$$s = \gamma^{-1}\left(\beta_1\gamma(s) + \beta_2\gamma(s)\right) = (a_1 + a_2)\,s + b,$$

for every $s \in \varphi(I)$. Hence obviously $a_1 + a_2 = 1$ and $b = 0$. Moreover, since A_γ is increasing, we have $a_i > 0$, $i = 1, 2$. In particular

$$\gamma(a_1s + a_2t) = \beta_1\gamma(s) + \beta_2\gamma(t),$$

and applying a trick used by Z. Daróczy and Zs. Páles in [3] we see that

$$\gamma\left(\frac{s+t}{2}\right) = \frac{\gamma(s) + \gamma(t)}{2},$$

for all s, $t \in \varphi(I)$. This yields

$$\gamma(s) = cs + d$$

for some $c > 0$ and $b \in \mathbb{R}$. Therefore, recalling definition of γ we get

$$\psi(u) = c\varphi(u) + d,$$

and subsequently

$$\varphi^{-1}\left(\alpha_1\varphi(u)+\alpha_2\varphi(v)\right)=$$
$$\psi^{-1}\left(\beta_1\psi(u)+\beta_2\psi(v)\right)=$$
$$\varphi^{-1}\left[\frac{1}{c}\left(\beta_1\left(c\varphi(u)+d\right)+\beta_2\left(c\varphi(v)+d\right)\right)-d\right]=$$
$$\varphi^{-1}\left(\beta_1\varphi(u)+\beta_2\varphi(v)\right).$$

It follows that $\alpha_i=\beta_i$, $i=1,2$, and thus the definition of $P(A)$ is correct.

2.6 P is additive.

As to the additivity of P, consider the following two partitions of Ω:

$$(A_1, A_2, \Omega\setminus(A_1\cup A_2))$$

and

$$(A_1\cup A_2, \Omega\setminus(A_1\cup A_2))$$

for some disjoint sets A_1, $A_2\in\mathcal{A}$. Then for every $s,t\in I$ we have

$$M\left(s\mathbf{1}_{A_1}+s\mathbf{1}_{A_2}+t\mathbf{1}_{\Omega\setminus(A_1\cup A_2)}\right)=$$
$$M\left(s\mathbf{1}_{A_1\cup A_2}+t\mathbf{1}_{\Omega\setminus(A_1\cup A_2)}\right),$$

which yields

$$\varphi^{-1}\left(\alpha_1\varphi(s)+\alpha_2\varphi(s)+\alpha_3\varphi(t)\right)=$$
$$\varphi^{-1}\left(\left(\alpha_1+\alpha_2\right)\varphi(s)+\alpha_3\varphi(t)\right),$$

for all $s,t\in I$. But this means that

$$P(A_1\cup A_2)=P(A_1)+P(A_2),$$

or the required additivity of P.

We have proved that

$$M(y)=\varphi^{-1}\left(\int_\Omega(\varphi\circ y)dP\right), \qquad (1)$$

for any $y\in\mathcal{M}_{fin}$. If the set $\mathcal{M}_{fin}$ is dense in a function space X (or at least any element of X can be approximated by a sequence from $\mathcal{M}_{fin}$) then obviously, due to continuity of M, φ, φ^{-1} and passing with lim under the sign of $\int$, we get formula (1) for any $y\in X$.

References

[1] J. Aczél, *Lectures on Functional Equations and their Applications.* Academic Press, New York – London, 1966.

[2] J. Aczél, Gy. Maksa and M. Taylor, *Equations of generalized bisymmetry and of consistent aggregation: Weakly surjective solutions which may be discontinuous at places.* J. Math. Anal. Appl. **214** (1997), 22–35.

[3] Z. Daróczy, Zs. Páles, *Convexity with given weight sequences.* Stochastica **11** (1987), 5–12.

[4] B. de Finetti, *Sul concetto di media.* Giornale dell'Istituto Italiano degli Attuari **2** (1931), 369–396.

[5] G.H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities.* Cambridge University Press, Cambridge, 1952.

[6] Gy. Maksa, *Solution of generalized bisymmetry type equations without surjectivity assumptions.* Aequationes Math. **57** (1999), 50–74.

[7] Gy. Maksa, *Quasisums and generalized associativity.* Aequationes Math. **69** (2005), 6–27.

[8] A. Münnich, Gy. Maksa, R. J. Mokken, *n-variable bisection.* J. Math. Psych. **44** (2000), 569–581.

[9] P. Volkmann, *Oral remark.* June 2009.

Full Papers

INCOMPARABILITY WITH RESPECT TO THE T-PARTIAL ORDER ON BOUNDED LATTICES

Emel Aşıcı
Department of Software Engineering,
Faculty of Technology,
Karadeniz Technical University,
61080 Trabzon, Turkey
emelkalin@hotmail.com

Funda Karaçal
Department of Mathematics,
Karadeniz Technical University,
61080 Trabzon, Turkey
fkaracal@yahoo.com

Summary

In this paper, we define the set of incomparable elements with respect to the T-partial order for any t-norm on a bounded lattice. By means of the T-partial order, an equivalence relation on the class of t-norms on bounded lattice is defined and this equivalence is deeply investigated. Lastly, we discuss some properties of this equivalence.

Keywords: Triangular norm; T-partial order; Bounded lattice

1 INTRODUCTION

Triangular norms were originally studied in the framework of probabilistic metric spaces [14, 15, 16] aiming at an extension of the triangle inequality.

In [12], it was defined a natural order for semigroups. Similarly, in [7], a partial order defined by means of t-norms on a bounded lattice was introduced. For any elements x, y of a bounded lattice L

$$x \preceq_T y :\Leftrightarrow T(\ell, y) = x \; for \; some \; \ell \in L,$$

where T is a t-norm. This order $\preceq_T$ is called a t-partial order of T. Moreover, the authors have investigated connections between the natural order $\leq$ on L and the T-partial order $\preceq_T$ on L.

In [7], it was obtained that $\preceq_T$ implies the natural order $\leq$ but its converse needs not be true. It was showed that a partially ordered set is not a lattice with respect to $\preceq_T$. It was determined some sets which are lattices with respect to $\preceq_T$ under some special conditions.

In [8], by means of the T-partial order, an equivalence relation on the class of t-norms is given and the equivalence classes linked to some special t-norms are characterized.
In [6], an equivalence relation on the class of the norms on $[0, 1]$ was defined. It was showed that the equivalence class of the weakest t-norm T_D on $[0, 1]$ contains a t-norm which was different from T_D.

In [1], with the help of any t-norm T on $[0, 1]$, it was obtained that the family $(T_\lambda)_{\lambda \in (0,1)}$ of t-norms on $[0, 1]$. If T was a divisible t-norm, then it was obtained that $([0, 1], \preceq_{T_\lambda})$ was a lattice.

In the present paper, we introduce the set of incomparable elements with respect to the T-partial order for any t-norm on a bounded lattice $(L, \leq, 0, 1)$. The main aim is to investigate some properties of this set. The paper is organized as follows. We shortly recall some basic notions in Section 2. In Section 3, we define the set of incomparable elements with respect to the T-partial order for any t-norm on a bounded lattice $(L, \leq, 0, 1)$ and we determine the sets of incomparable elements w.r.t. T-partial order of the infimum t-norm $T_\wedge$ and the weakest t-norm T_W. In Section 4, we define an equivalence on the class of t-norms on a bounded lattice $(L, \leq, 0, 1)$. We determine the equivalence class of the infimum t-norm $T_\wedge$ when L is a chain. Thus, we obtain that, in the case of $L = [0, 1]$, all continuous t-norms are equivalent. Although, we give some examples illustrating that left-continuous t-norms need not be equivalent, in general. We show by an example that the left-continuity of any of the t-norms in the equivalence class does not imply the left-continuity for another t-norm in the equivalence class. In [1], it was shown that "T_1 and T_2 are two t-norms on $[0, 1]$ such that for all $x \in [0, 1]$, $\mathcal{I}_{T_1}{}^{(x)} = \mathcal{I}_{T_2}{}^{(x)}$ if and only if the t-norms T_1 and T_2 are equivalent under the relation $\sim$ in (2)". In this study, by an example we show that this proposition only provides a sufficient and not a necessary condition for the relation β_L in (4).

ISBN: 978–83–8012–519–3

2 NOTATIONS, DEFINITIONS AND A REVIEW OF PREVIOUS RESULTS

Definition 2.1 ([10])**.** A triangular norm (t-norm for short) is a binary operation T on the unit interval $[0,1]$, i.e., a function $T : [0,1]^2 \to [0,1]$, such that for all $x, y, z \in [0,1]$ the following four axioms are satisfied:
$(T1)$ $T(x,y) = T(y,x)$ (commutativity)

$(T2)$ $T(x,T(y,z)) = T(T(x,y),z)$ (associativity)

$(T3)$ $T(x,y) \leq T(x,z)$ whenever $y \leq z$ (monotonicity)

$(T4)$ $T(x,1) = x$ (boundary condition)

Example 2.1 ([10])**.** The following are the four basic t-norms T_M, T_P, T_L, T_D given by, respectively:

$$T_M(x,y) = min(x,y)$$

$$T_P(x,y) = x.y$$

$$T_L(x,y) = max(x+y-1,0)$$

$$T_D(x,y) = \begin{cases} 0 & \text{if } (x,y) \in [0,1[^2 \\ min(x,y) & \text{otherwise} \end{cases}$$

Also, t-norms on a bounded lattice $(L, \leq, 0, 1)$ are defined in similar way, and then extremal t-norms $T_\wedge$ and T_W on L is defined as follows, respectively:
$T_\wedge(x,y) = x \wedge y$

$$T_W(x,y) = \begin{cases} x & \text{if } y = 1 \\ y & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Especially we obtained that $T_W = T_D$ and $T_\wedge = T_M$ for $L = [0,1]$.

Definition 2.2 ([5])**.** A t-norm T on L is divisible if the following conditions holds:

$$\forall x, y \in L \ with\ x \leq y \ there\ is\ a\ z \in L\ such\ that$$

$$x = T(y,z)$$

Proposition 2.1 ([4])**.** *Let T be a t-norm on $[0,1]$. T is divisible if and only if T is continuous.*

Definition 2.3 ([2])**.** Given a bounded lattice $(L, \leq, 0, 1)$ and $a, b \in L$, if a and b are incomparable, in this case we use the notation $a \parallel b$.

Definition 2.4 ([7])**.** Let $(L, \leq, 0, 1)$ be a bounded lattice, T be a t-norm on L. The order defined as following is called a T-partial order (triangular order) for t-norm T.

$$x \preceq_T y :\Leftrightarrow T(\ell, y) = x \ for\ some\ \ell \in L$$

Proposition 2.2 ([7])**.** *Let $(L, \leq, 0, 1)$ be a bounded lattice, T be a t-norm on L. Then the binary relation $\preceq_T$ is a partial order on L.*

Proposition 2.3 ([7])**.** *Let T be a t-norm on a bounded lattice $(L, \leq, 0, 1)$. Then, if $x \preceq_T y$ necessarily we have also $x \leq y$.*

Lemma 2.1 ([8])**.** *Let $(L, \leq, 0, 1)$ be a bounded lattice. For all t-norms on L and all $x \in L$ it holds that $0 \preceq_T x$, $x \preceq_T x$ and $x \preceq_T 1$.*

Definition 2.5 ([8])**.** Let T be a t-norm on $[0,1]$ and let K_T be defined

$$K_T = \{x \in [0,1] \mid for\ some\ y \in (0,1)$$

$$[x \leq y \ and\ x \npreceq_T y] \ or\ [y \leq x\ and\ y \npreceq_T x]\}$$

Definition 2.6 ([8])**.** Let $(L, \leq, 0, 1)$ be a given bounded lattice. Define a relation $\sim$ on the class of all t-norms on $(L, \leq, 0, 1)$ by $T_1 \sim T_2$ if and only if the T_1-partial order coincides with the T_2-partial order, that is

$$T_1 \sim T_2 :\Leftrightarrow \preceq_{T_1} = \preceq_{T_2}. \quad (1)$$

Definition 2.7 ([6])**.** Let $([0,1], \leq, 0, 1)$ be the unit interval. Define a relation β on the class of all t-norms on $[0,1]$ by $T_1 \beta T_2$ if and only if the set of incomparable elements with respect to the T_1-partial order coincides with the set of incomparable elements with respect to the T_2-partial order, that is

$$T_1 \beta T_2 :\Leftrightarrow K_{T_1} = K_{T_2}. \quad (2)$$

3 ABOUT THE SET K_T^L ON ANY BOUNDED LATTICE

In this section, we study on the set of all incomparable elements with respect to the T partial order $\preceq_T$ with some t-norm T on a bounded lattice $(L, \leq, 0, 1)$.

Definition 3.1. Let T be a t-norm on a bounded lattice $(L, \leq, 0, 1)$ and let K_T^L be defined by

$$K_T^L = \{x \in L \backslash \{0,1\} \mid for\ some\ y \in L \backslash \{0,1\}$$

$$[x < y\ implies\ x \npreceq_T y] \ or\ [y < x\ implies\ y \npreceq_T x]$$

$$or\ x \parallel y\}$$

If $L = [0,1]$, then it is trivial to see that $K_T = K_T^L$.
Note: It is obtained that $K_T^L \subseteq L \backslash \{0,1\}$ for any t-norm T by Lemma 2.1.

Proposition 3.1. *Let $(L, \leq, 0, 1)$ be a bounded lattice and T be a t-norm on L. If there exist two elements of L such that these are incomparable, then $K_T^L \neq \emptyset$.*

Although the set $K_T^L \neq \emptyset$, it need not be the case that elements in L are incomparable. Now, let us investigate the following example.

Example 3.1. Let T be a t-norm on $[0,1]$ and the family $(T_\lambda)_{\lambda\in(0,1)}$ of t-norms be given by

$$T_\lambda(x,y) = \begin{cases} 0 & , T(x,y) \le \lambda \text{ and } x,y \ne 1 \\ T(x,y) & ,\text{otherwise} \end{cases}$$

Observe that due to (Theorem 15 in [1]) the function T_λ is a t-norm. Then we have that $K_{T_\lambda} = (0,1)$, but since L is a chain all elements are comparable.

Definition 3.2. Let $(L,\le,0,1)$ be a bounded lattice. The set I_L is defined by

$$I_L = \{x \in L \mid \exists y \in L \textit{ such that } x \parallel y\}$$

Note: Due to the definition of the set K_T^L, it is obtained that $I_L \subseteq K_T^L$ for any t-norm T on L.

Lemma 3.1. *Let $(L,\le,0,1)$ be a bounded lattice. For the weakest t-norm T_W on L, $K_{T_W}^L = L\backslash\{0,1\}$.*

Proposition 3.2. *Let $(L,\le,0,1)$ be a bounded lattice. For the infimum t-norm $T_\wedge$ on L, $K_{T_\wedge}^L = I_L$.*

Remark 3.1. The converse of Proposition 3.2 is not be true. That is, T is a t-norm on L such that if $K_T^L = I_L$, then need not be $T = T_\wedge$. The previous proposition only provides a sufficient and not a necessary condition for the equality $K_T^L = I_L$, as the following example shows.

Example 3.2. Let $L = \{0,a,b,c,1\}$ and consider the order $\le$ on L as follows:

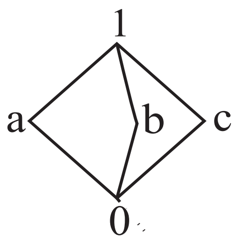


Figure 1: The order $\le$ on L

We put $T = T_W$. In this case it is trivial that $K_{T_W}^L = \{a,b,c\}$ and $I_L = \{a,b,c\}$. So it is obtained that $K_{T_W}^L = I_L$, but clearly $T_\wedge \ne T_W$.

Definition 3.3. Let T be a t-norm on $(L,\le,0,1)$ and let $\mathcal{I}_T^{L^{(c)}}$ for $c \in L$ be defined by

$$\mathcal{I}_T^{L^{(c)}} = \{x \in L\backslash\{0,1\} \mid x \textit{ is incomparable to } c \textit{ according to } \preceq_T\}$$

Lemma 3.2. *Let T be a t-norm on $(L,\le,0,1)$. Then $K_T^L = \bigcup_{x\in L} \mathcal{I}_T^{L^{(x)}}$.*

Proposition 3.3. *Let T_1 and T_2 be two t-norms on a bounded lattice $(L,\le,0,1)$. Then for all $x \in L$, $\mathcal{I}_{T_1}^{L^{(x)}} = \mathcal{I}_{T_2}^{L^{(x)}}$ if and only if the t-norms T_1 and T_2 are equivalent under $\sim$ in (2).*

4 ABOUT AN EQUIVALENCE RELATION ON THE CLASS OF T-NORMS ON ANY BOUNDED LATTICE

The above introduced the set K_T^L on any bounded lattice allows us to introduce the next equivalence relation on the class of all t-norms on $(L,\le,0,1)$.

Definition 4.1. Let $(L,\le,0,1)$ be a bounded lattice. Define a relation β_L on the class of all t-norms on $(L,\le,0,1)$ by $T_1\beta_L T_2$ if and only if

$$T_1\beta_L T_2 :\Leftrightarrow K_{T_1}^L = K_{T_2}^L \tag{3}$$

If $L = [0,1]$, then it is trivial to see that $\beta_L = \beta$, β in (3). The next result is obvious.

Lemma 4.1. *The relation β_L given in Definition 4.1 is an equivalence relation.*

Definition 4.2. For a given t-norm T on a bounded lattice $(L,\le,0,1)$, we denote by $\overline{T}$ the β_L equivalence class linked to T, i.e.,

$$\overline{T} = \{T' \mid T' \textit{ is a } t-\textit{norm on } L \textit{ and } K_T^L = K_{T'}^L\}$$

In [8], it was shown that an equivalence class of the infimum t-norm $T_\wedge$ on L under the relation $\sim$ in (2) is the set of all divisible t-norms on L. But according to the relation β_L in (4), an equivalence class of the infimum t-norm $T_\wedge$ on L is not the set of all divisible t-norms on L. To illustrate this claim we shall give the following example.

Example 4.1. Consider the bounded lattice $(L,\le,0,1)$ with $L = \{0,a,b,c,1\}$ as shown in Figure 2.

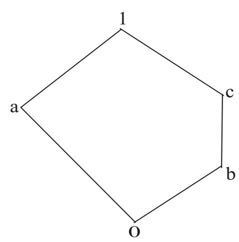


Figure 2: The order $\le$ on L

We consider $T_\wedge$ and T_W t-norms on L. It is trivial that $K_{T_\wedge}^L = \{a,b,c\}$ and $K_{T_W}^L = \{a,b,c\}$. So we have that $K_{T_\wedge}^L = K_{T_W}^L$. By the definition of the relation β_L in (4), the t-norms $T_\wedge$ and T_W are equivalent, i.e., $T_\wedge\beta_L T_W$. But the weakest t-norm T_W is not divisible t-norm on L.

Naturally, one can think when an equivalence class of the infimum t-norm $T_\wedge$ on L under the relation β_L in (4), is the set of all divisible t-norms on L. As an answer to this question, let us investigate the following Proposition.

Proposition 4.1. *If L is a chain, then an equivalence class of the infimum t-norm $T_\wedge$ on L under the relation β_L in* (4)*, is the set of all divisible t-norms on L.*

Corollary 4.1. *The equivalence class of the minimum t-norm T_M on $[0,1]$ according to the relation β_L in* (4)*, is the set of all divisible t-norms on* $[0,1]$.

Corollary 4.2. *The equivalence class of the minimum t-norm T_M on $[0,1]$ according to the relation β_L in* (4)*, is the set of all continuous t-norms on $[0,1]$ by Proposition 2.1.*

Remark 4.1. In Corollary 4.2, we have shown that any two continuous t-norms on $[0,1]=L$ are equivalent under the relation β_L in (4). Naturally, one can think whether any two left-continuous t-norms are in the same equivalence class, i.e, any two left-continuous t-norms are equivalent under the relation β_L in (4). To illustrate that two left-continuous t-norms may not be equivalent under the relation β_L in (4) we shall give the following example.

Example 4.2. Consider the t-norms on $[0,1]$ defined as follows:

$$T^{nM}(x,y)=\begin{cases}0 & ,\ if\ x+y\leq 1\\ min(x,y) & ,\text{otherwise}\end{cases}$$

and

$$T_4(x,y)=\begin{cases}min(x,y) & ,if\ max(x,y)\in(\frac{3}{4},1]\\ \frac{1}{4} & ,if\ x,y\in(\frac{1}{4},\frac{3}{4}]\\ 0 & ,\text{otherwise}\end{cases}$$

T^{nM} and T_4 are left continuous t-norms [11]. But since $K_{T^{nM}}=(0,1)$ and $K_{T_4}=(0,\frac{3}{4}]$, the t-norms T^{nM} and T_4 are not equivalent under β_L in (4).

Remark 4.2. One may ask whether any t-norm equivalent to a left continuous t-norm needs to be left-continuous, too. The following example shows that also this need not.

Example 4.3. Let $T^\star$ be a function on $[0,1]$ defined by

$$T^\star(x,y)=\begin{cases}\frac{1}{2} & ,\ if\ x,y=\frac{1}{2}\\ T^{nM}(x,y) & ,\text{otherwise}\end{cases}$$

The function $T^\star$ is a t-norm by [10]. We will show that this t-norm is equivalent to the left-continuous t-norm T^{nM}, but $T^\star$ is not left continuous t-norm.

Proposition 3.3 gives a sufficient and necessary condition for the t-norms T_1 and T_2 are equivalent under the relation $\sim$ in (2). But the following Proposition only provides a sufficient and not a necessary condition for the relation β_L in (4).

Proposition 4.2. *Let T_1 and T_2 be two t-norms on $(L,\leq,0,1)$. If for all $x\in[0,1]$, ${\mathcal{I}^L_{T_1}}^{(x)}={\mathcal{I}^L_{T_2}}^{(x)}$, then the t-norms T_1 and T_2 are equivalent under β_L in* (4).

Remark 4.3. The converse of Proposition 4.2 is not be true. Here is an example illustrating the case that need not be true.

Example 4.4. Consider the t-norm $T:[0,1]^2\to[0,1]$ defined by

$$T(x,y)=\begin{cases}\frac{xy}{2} & ,\ if\ (x,y)\in[0,1)^2\\ min(x,y) & ,\text{otherwise}\end{cases}$$

and the t-norm T_D on $[0,1]$. Then, $K_T=K_{T_D}$ by [8]. The t-norms T and T_D are equivalent under β_L in (4) and we obtained that,

a) a_1) ${\mathcal{I}_T}^{(x)}=\{y\in(0,1)\mid y\in[\frac{x}{2},2x]\ \text{ and }\ x\neq y\}$ for $x\in(0,\frac{1}{2})$
and

a_2) ${\mathcal{I}_T}^{(x)}=\{y\in(0,1)\mid y\in[\frac{x}{2},1)\ \text{ and }\ x\neq y\}$ for $x\in[\frac{1}{2},1)$

b) ${\mathcal{I}_{T_D}}^{(x)}=\{y\in(0,1)\mid x\neq y\}$ for $x\in(0,1)$

5 CONCLUSIONS

We have defined the set of incomparable elements with respect to the T-partial order for any t-norm on a bounded lattice $(L,\leq,0,1)$. Also we have introduced and studied an equivalence relation β_L in (4) defined on the class of all t-norms on L. We have shown any two continuous t-norms on $[0,1]$ are equivalent by the introduced equivalence relation. As shown by examples, all left-continuous t-norms on $[0,1]$ do not form an equivalence class in our approach. Further that we have shown when an equivalence class of the infimum t-norm $T_\wedge$ on L, is the set of all divisible t-norms on L.

References

[1] E. Aşıcı and F. Karaçal (2014). On the T-Partial order and properties. *Information Sciences* **267**, pp 323-333.

[2] G. Birkhoff (1967). Lattice Theory. *3 rd edition. Providence*

[3] B. De Baets and R. Mesiar (1999). Triangular norms on product lattices. *Fuzzy Sets and Systems* **104**, pp 61-75.

[4] B. De Baets and R. Mesiar (1999). Triangular norms on the real unit square. *Proceedings of the 1999 EUSFLAT-ESTYLF Joint Conference, Palma de Mallorca, Spain* pp 351-354.

[5] J. Casasnovasa and G. Mayor (2008). Discrete t-norms and operations on extended multisets. *Fuzzy Sets and Systems* **159**, pp 1165-1177.

[6] F. Karaçal and E. Aşıcı (2013). Some notes on T-partial order. *Journal of Inequalities and Applications* **219**

[7] F. Karaçal and M. N. Kesicioğlu (2011). A T-partial order obtained from t-norms. *Kybernetika* **47**, pp 300-314.

[8] M. N. Kesicioğlu, F. Karaçal and R. Mesiar (2015). Order-equivalent triangular norms. *Fuzzy Sets and Systems*, **268**, pp 59-71.

[9] D. Khadjiev and F. Karaçal (2012). Pairwise co-maximal elements and the classification of all ∨-distributive triangular norms of length 3. *Information Sciences* **204**, pp 36-43.

[10] E.P. Klement, R. Mesiar and E. Pap (2000). Triangular Norms. *Kluwer Academic Publishers, Dordrecht*

[11] K.C. Maes and A.Mesiarova-Zemankova (2009). Cancellativity properties for t-norms and t-subnorms. *Information Sciences* **179**, pp 135-150.

[12] H. Mitsch (1986). A natural partial order for semigroups. *Proceedings of the American Mathematical Society* **97**, pp 384-388.

[13] S. Saminger (2006). On ordinal sums of triangular norms on bounded lattices. *Fuzzy Sets and Systems* **157**, pp 1403-1413.

[14] B. Schweizer and A. Sklar (1983). Probabilistic Metric Spaces. *Elsevier, Amsterdam.*

[15] B. Schweizer and A. Sklar (1960). Statistical metric spaces. *Pacific Journal of Mathematics* **10**, pp 313-334.

[16] B. Schweizer and A. Sklar (1963). Associative functions and abstract semigroups. *Pacific Journal of Mathematics* **10**, pp 69-81.

LEARNING STABLE WEIGHTS FOR DATA OF VARYING DIMENSION

G. Beliakov, S. James
Deakin University
School of Information Technology
Faculty of Science,
Engineering and Built Environment
Burwood, VIC, Australia
{gleb,sjames}@deakin.edu.au

D. Gómez, J. T. Rodríguez, J. Montero
Complutense University
Department of Statistics and Operational Research
Faculty of Mathematics and Faculty of Statistics
Madrid, Spain
dagomez@estad.ucm.es,
{jtrodrig,monty}@mat.ucm.es

Summary

In this paper we develop a data-driven weight learning method for weighted quasi-arithmetic means where the observed data may vary in dimension.
Keywords: Aggregation functions, R-stability, linear programming, weights learning

1 INTRODUCTION

A fundamental problem in data analysis is to determine the influence of variable inputs on an observed output, however with real world data we acknowledge that sometimes the dimension of data to be processed cannot be fully fixed in advance (for example, some expected data might be lost or corrupted, but also some unexpected data might arrive).

We assume the existence of a modelling function $f : \Re^n \to \Re$, where f depends on a vector of weighting parameters $\mathbf{w}$ and $f_{\mathbf{w}}(\mathbf{x}) = y$. In the case of the classical weighted arithmetic mean, y is modeled as a linear combination of the inputs, i.e. $f(\mathbf{x}) = w_1x_1+w_2x_2+\ldots+w_nx_n$, $\sum_{i=1}^n w_i = 1$ and each weight w_i reflects the importance of the i-th contributing variable. Learning the weighting vector also allows us to predict the output for unobserved input vectors. In some situations however, rather than being dependent on n values, the dimension of $\mathbf{x}$ may vary. In practice, y may be a function of multiple input sensors, some of which are turned off or unaccessible at given times, or in multi-criteria decision making, some observable inputs may be missing.

The idea of stability (or L-stability, R-stability, LR-stability) has been proposed for aggregating data of varying dimension in a consistent way [18, 4], e.g. if we have a weighted mean given by $y = 0.6x_1 + 0.4x_2$ when $n = 2$, then it would not usually make sense if we extend this function to 3 variables with y expressed as $y = 0.1x_1 + 0.3x_2 + 0.6x_3$, as this now implies that the first variable is less important than the second. It was established in [8, 18] that for weighting vectors of n and $(n-1)$ dimensions, R-stability, i.e. stability with respect to adding a new input in the n-th position, requires:

$$w_i^n = (1 - w_n^n) \cdot w_i^{n-1}, i = 1, \ldots, n-1, \qquad (1)$$

where w_i^n denotes the i-th weight of the n-dimensional weighting vector.

In this paper, we consider the problem of learning stable weighting vectors satisfying this property from observed data.

We will set out the article as follows. We first give an overview of aggregation functions and stability, as well as some existing approaches to learning weighting vectors using linear programming techniques. In Section 3, we look at how weights can be learnt for data of different dimension, provided we have access to observed values y. In Section 4, we run some experiments to demonstrate the usefulness of the method, while in Section 5 we summarize with some discussion and directions for further research.

2 PRELIMINARIES

This contribution applies weight learning techniques in order to learn aggregation function tuples whose weights are stable. As well as providing an overview of aggregation functions, in this section we will also recall some results concerning the concept of aggregation stability and the least absolution deviation fitting method.

ISBN: 978–83–8012–519–3

2.1 AGGREGATION FUNCTIONS

Aggregation functions are core to many decision processes, providing a representative output from an n-dimensional input vector. Overviews of their properties and some fundamental results can be found in [5, 15, 21] (also see [2, 7, 9, 13]).

Definition 1 *An aggregation function* $A : [0,1]^n \to [0,1]$ *is a function increasing in each argument and satisfying* $A(0,\dots,0) = 0$ *and* $A(1,\dots,1) = 1$.

Here we are interested particularly in aggregation functions that are averaging, i.e. that their inputs are bounded such that for all $\mathbf{x} \in [0,1]^n$,

$$\min(\mathbf{x}) \le A(\mathbf{x}) \le \max(\mathbf{x}).$$

An aggregation functions that generalizes a number of important families is the quasi-arithmetic mean. We provide the following definition.

Definition 2 *For a strictly monotone continuous generating function*[1] $g : [0,1] \to [-\infty,\infty]$ *and weighting vector* $\mathbf{w}$ *such that* $w_i \in [0,1]$ *and* $\sum\limits_{i=1}^{n} w_i = 1$*, the weighted quasi-arithmetic mean is given by,*

$$QAM_{\mathbf{w}}(\mathbf{x}) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_i)\right). \qquad (2)$$

Special cases include weighted arithmetic means (WAM) $\sum_{i=1}^{n} w_i x_i$, where $g(t) = t$, weighted power means $(\sum_{i=1}^{n} w_i x_i^p)^{\frac{1}{p}}$, where $g(t) = t^p$ and weighted geometric means $G(\mathbf{x}) = \prod_{i=1}^{n} x_i^{w_i}$ if $g(t) = -\ln t$. The weights w_i are usually non-negative and sum to one.

2.2 R-STABLE WEIGHTING VECTORS

The idea of strict stability represents a kind of consistency in the aggregation process when aggregation families are defined for varying dimension. Rojas et al. proposed the following conditions in [18], using Yager's self-identity property [22] as their basis.

Definition 3 *Let* $\{A_n : [0,1]^n \to [0,1], n \in N\}$ *be a family of aggregation functions. Then it is said that:*

1. $\{A_n\}_n$ *is R-strictly stable if* $A_n(x_1, x_2, \dots, x_{n-1}, A_{n-1}(x_1, x_2, \dots, x_{n-1}))$ *coincides with* $A_{n-1}(x_1, x_2, \dots, x_{n-1})$.

2. $\{A_n\}_n$ *is L-strictly stable if* $A_n(A_{n-1}(x_1, x_2, \dots, x_{n-1}), x_1, x_2, \dots, x_{n-1})$ *coincides with* $A_{n-1}(x_1, x_2, \dots, x_{n-1})$

3. $\{A_n\}_n$ *is LR-strictly stable if both properties hold simultaneously.*

The idea that a subset of values can be replaced by their mean with no effect on the overall output for quasi-arithmetic means has been well established for almost a century, e.g. see [1] and the references contained therein. More recently, results have been established in [4, 18]. In particular, the geometric means and arithmetic means with respect to a weighting vector with equal weights are considered LR-strictly stable, as too are the maximum, minimum, and median.

For weighted versions of these operators, strict stability is dependent on the choice of weights. The basis result regarding the weights lies in the following proposition.

Proposition 1 *Let* $\mathbf{w}^n \in [0,1]^n, n \in \mathbb{N}$ *be a sequence of weighting vectors such that* $\sum\limits_{i=1}^{n} w_i^n = 1$ *holds* $\forall n \ge 2$. *Then, the family of weighted means defined by these weights is R-strictly stable if and only if the weighting property described by Eq. (1) holds.*

Analogous results hold for the additional input being included in the 1st position or j-th position [4], since in most cases it is not important where the new input is placed as long as the relationship holds between corresponding inputs as the arity is increased. It is more complicated for functions such as the ordered weighted averaging operator because we cannot generally predict where the output of A_{n-1} would be placed when the inputs are reordered. In this paper we will contain ourselves to functions which do not involve a reordering step.

2.3 LEARNING WEIGHTS USING LINEAR PROGRAMMING

In order to analyze the data set and learn weights, we can use linear programming techniques based on the minimization of the least absolute deviation (LAD) of residuals [3, 6]. In the standard case for weight learning, we have a function $f_{\mathbf{w}}$ which is dependent on $\mathbf{w}$, and a set of observed values y_k which we want the function to predict once we know its parameters. So we have

$$\text{Minimize} \quad \sum_{k=1}^{K} \left| f_{\mathbf{w}}(\mathbf{x}_k) - y_k \right|, \qquad (3)$$

subject to any desired constraints.

[1] See [5] for more information regarding the choice of generators and their construction. Where $g(0)$ or $g(1)$ approach $\pm\infty$, special care needs to be taken in calculation with the convention $0 \cdot \infty = 0$ usually adopted. Methods also exist for using non-continuous and non-strict generators.

The advantage of minimizing the least absolute deviation rather than a least-squares approach is that we convert it into a linear program, and further, the output set of weights should be less sensitive to outliers or extreme data. We represent the absolute differences between the predicted and observed output values in terms their positive and negative parts, i.e. $r_k = |f_{\mathbf{w}}(\mathbf{x})_k - y_k| = r_k^+ + r_k^-$. For each observed input/output pair $(x_{k1}, x_{k2}, ..., x_{kn}, y_k)$, one of the r_k^+, r_k^- will be zero.

The weight learning can then be performed with the objective of minimizing the residuals with the following program.

$$\begin{aligned} \underset{\mathbf{w}}{\text{Minimize}} \quad & \sum_{k=1}^{K} r_k^+ + r_k^-, & (4)\\ \text{s.t.} \quad & f_{\mathbf{w}}(\mathbf{x}_k) - r_k^+ + r_k^- = y_k, k = 1, \dots, K, \\ & w_1 + w_2 + \ldots + w_n = 1, \\ & r_k^+, r_k^- \geq 0, \\ & w_i \geq 0, i = 1, \dots, n. \end{aligned}$$

The weights for quasi-arithmetic means, including all their special cases, can be fit in the same manner with generator transformations to the observed data.

3 LEARNING STABLE WEIGHTS WITH MISSING INPUT DATA

The problem of varying dimension or missing data is common in machine learning and classification with real datasets, e.g. see [11]. New information pertaining to a particular dataset may become available, introducing new variables that are not observable for the previously collected data. In some cases, it may be found after the collection phase that some measurements were extremely unreliable, making them useless, or it may simply be that some variables are not relevant for particular instances. The two standard ways to approach such situations are:

- to make the data uniform by removing all data relating to variables that have not been measured across the dataset;
- to assign a neutral or default value for that particular variable when it is missing.

The latter case is often preferable, since otherwise we may lose a lot of information that could be useful for the task at hand, however it may not always be possible to assign a 'neutral' value. We consider the following example scenario.

Example 1 (Stable student evaluation) *Students competing for a scholarship are evaluated against 4 criteria: exam marks (40 %), interview (30%), application letter (10%) and 1 written reference (20%). Due to unforseen circumstances, however, the decision needs to be made earlier than anticipated and the reference for many students is yet to arrive. In order to be as fair as possible, students with all data available have their scores aggregated with respect to the full weighting vector* $\mathbf{w}^4 = (0.4, 0.3, 0.1, 0.2)$*, while for those students with a missing reference, a stable weighting vector is defined, consistent with Eq. (1) such that* $w_i^3 = w_i^4/(1 - 0.2)$*. This gives* $\mathbf{w}^3 = (0.5, 0.375, 0.125)$.

While the results of stability are useful for defining weighting vectors that are stable or 'consistent' across varying dimensions, we now turn to the problem of learning such families of weighting vectors when we only have the observed input/output data and don't know the relative importance of each input. Example 2 follows from Example 1.

Example 2 (Learning stable weights) *The scholarship assessment panel is unconvinced that the proportional importance allocated to the criteria properly reflects the students' potential. After the first year, they have performance data available for all the candidates, along with the data used to award the scholarships (since the late references were not used, their score data is still unavailable). An example of such data is given in Table 1.*

Table 1: Scholarship and performance scores with data missing for some students

Student	s_1	s_2	s_3	s_4	s_5
Exam (x_1)	0.5	0.6	0.2	0.5	0.7
Interview (x_2)	0.8	0.3	0.7	0.9	0.8
App. Letter (x_3)	0.4	0.6	0.1	0.3	0.7
Reference (x_4)	-	0.6	0.4	-	-
Performance (y)	0.51	0.63	0.62	0.71	0.78

The problem is now to learn both a 3- and 4-dimensional weighting vector (stable with respect to one another) from the dataset in order to approximate the importance of each criteria in assessing the academic potential of each student.

In the following section we develop our approach for addressing this problem. We note that learning paramaters for aggregating data of varying dimension has previously been considered in [16], while aggregation with missing data has been more recently considered in [19, 10, 12].

3.1 FRAMING STABLE WEIGHT LEARNING AS A BILEVEL OPTIMIZATION PROBLEM

Suppose we wish to learn a weighted quasi-arithmetic mean that best fits the data of the form in Table 1. We note that we are essentially required to learn 7 parameters, i.e. the aggregation weights that best model the output y as a function of the respective 3- and 4-dimensional input vectors. If we learn $\mathbf{w}^3$ and $\mathbf{w}^4$ separately, the weights may not satisfy Eq. (1) and therefore could not be considered stable - in this case, they would give conflicting approximations of the importance for each criterion. It would also result in fewer data with respect to the number of variables we need to learn. Lastly, the relationship in Eq. (1) is not linear with respect to the weights and therefore could not be incorporated into a simple optimization algorithm.

We therefore consider the problem as the following bilevel optimization problem.

$$\underset{\alpha,\mathbf{w}^{n-1}}{\text{Minimize}} \quad \sum_{j=1}^{J} \left| f_{\mathbf{w}}(\mathbf{x}_j) - y_j \right| + \sum_{k=1}^{K} \left| f_{\mathbf{w}}(\mathbf{x}_k) - y_k \right|$$

where $\alpha = (1 - w_n^n)$, J and K represent the number of observed data of $n-1$ and n dimensions respectively, and $\mathbf{x}_j \in [0,1]^{n-1}$ and $\mathbf{x}_k \in [0,1]^n$.

In the case of fitting quasi-arithmetic means, for fixed α, we can still minimize with respect to the same objective as in (4) and sum the residuals, however we impose the following constraints.

For $\mathbf{x}_j \in [0,1]^{n-1}$, we have

$$\left(\sum_{i=1}^{n-1} w_i g(x_{ji}) \right) - r_j^+ + r_j^- = g(y_j), \qquad (5)$$

then for $\mathbf{x}_k \in [0,1]^n$, we can use Eq. (1) and our α, giving us,

$$\alpha \left(\sum_{i=1}^{n-1} w_i g(x_{ki}) \right) + w_n g(x_{kn}) - r_k^+ + r_k^- = g(y_k). \quad (6)$$

Since α is a scalar, these constraints remain linear with respect to $\mathbf{w}^{n-1}$. We remind that w_n is obtained directly from α and hence is also a fixed constant in this step of the minimization process. We then only require the constraints such that the weights in $\mathbf{w}^{n-1}$ sum to 1 and the residuals are nonnegative.

3.2 IMPLEMENTATION

We implemented the approach in R [17], adapting existing libraries we had created for least absolute deviation fitting[2].

From a dataset with the structure given in Table 1, we first build the left hand side of the constraints array, comprising the entries $g(x_{ki})$ for each instance k and $i = 1, \dots, n-1$ and coefficients of -1 or 1 for the residuals. There is also a row for constraining the sum of weights. We note that these are not dependent on α. For each given α, we then construct the right hand side of the constraints matrix, with entries $g(y_k) - (1-\alpha) \cdot g(x_{kn})$, while the right-hand side for the weights sum is set to α. We note that while this is equivalent to implementing Eqs. (5)-(6), it means that we do not need to keep on rebuilding the left hand side for each α and this can simply be stored in memory. We used the one-dimensional Brent method for optimization which is available as part of the standard *optim* function in R and also allows the setting of lower and upper bounds (in our case $0 \le \alpha \le 1$).

4 EXPERIMENTS

Here we provide some experimental results to demonstrate the usefulness of our stable-weight learning approach.

4.1 GENERATED DATASETS

For each run of the experiment, we first set the fifth weight to a predefined value w_5 and randomly generated the remaining 4 weights along with 50 random n-dimensional input vectors. We then calculated the y values based on the weighting vector and added Guassian noise with a mean of 0 and standard deviation of 0.05. We split the data into 25 training and 25 test data and removed a number of the n-th entries from 0% to 90% in increments of 10% from both datasets.

We compared 3 potential approaches to dealing with such a dataset:

1 To remove the variable that has missing information for some entries;

2 To remove all instances with incomplete information;

3 To use the stable weight learning techniques described in the previous section so that all instances can be used in both training and testing.

We note that with the first method, the training dataset will be smaller, discounting any of the 25 training data with missing information from the learning

[2] As with the implementation of the optimization procedure described here, the experimental method is also available from http://aggregationfunctions.wordpress.com.

Table 2: Average overall error in predicting test data with increasing percentage of missing data.

% missing	method 1	method 2	method 3
0	0.141	0.045	0.044
10	0.142	0.053	0.052
20	0.145	0.065	0.064
30	0.142	0.077	0.075
40	0.143	0.086	0.088
50	0.142	0.094	0.094
60	0.143	0.111	0.107
70	0.141	0.123	0.116
80	0.146	0.144	0.134
90	0.144	0.165	0.137

Table 3: Average overall error in predicting test data with increasing weight to variable with missing entries

fixed w_5	method 1	method 2	method 3
0	0.045	0.055	0.046
0.1	0.052	0.056	0.049
0.2	0.074	0.065	0.059
0.3	0.098	0.076	0.070
0.4	0.126	0.088	0.086
0.5	0.145	0.097	0.094
0.6	0.179	0.112	0.110
0.7	0.208	0.122	0.120
0.8	0.237	0.138	0.133
0.9	0.265	0.153	0.143

algorithm. For the test data, however, we used the fitted n-dimensional vector and then determined a stable $(n-1)$-dimensional weighting vector for the vectors with missing entries.

4.2 RESULTS

The artificial set simulates a well-behaving dataset where the data vary in dimension because information about a contributing variable is missing. In this case, the variable is independent of any of the other contributing inputs and so there is no perfect solution to dealing with the data we have. The purpose of the experiments is hence to demonstrate the usefulness of the approach as a reasonable solution in this scenario.

In some contexts, however, a missing input may not necessarily represent ignorance about a contributing variable, and rather the output may depend on input vectors which differ in dimension by construction. For example, we might consider a nearest-neighbor approximation method where we include only neighbours within a given distance of the point we are approximating. In such cases, the 'missing' input is less likely to be affecting the output. Alternatively, we can think of aggregating the citations data of a journal or researcher over a given timeframe [20]. With our artificial dataset however, we chose not to generate data in this way as it would clearly bias the method of obtaining stable weighting vectors.

As would be expected, the method of leaving out the variable with missing entries has worse error as the importance of the missing variable increases. It is unaffected by changes in the number of missing data (since it treats the dataset as if all of the entries are missing). This relationship can be seen clearly in Tables 2-3.

The average error associated with removing the missing entries as opposed to stable fitting is quite similar. We could assume that the increase in error with method 2 mainly revolves around the calculation of the output associated with 4-dimensional outputs - since on average, the learned weighting vector should be close to the true vector. As the number of missing data increases, however, the training data set gets smaller, which in turn is likely to affect the accuracy. We note that in one case, with 90% of the data with a missing value and $w_5 = 0.9$, this method failed to generate a model[3]. Error for method 3, however, can arise in the actual fitting process, since the 4-dimensional training data have an output that was actually calculated using the missing variable. The advantage then, is that it is still able to use the entire training set to generate the weights. This is perhaps why it performs slightly better as the number of missing data gets larger. Comparisons of the average error can be seen in Figs. 1-2 and Tables 2-3.

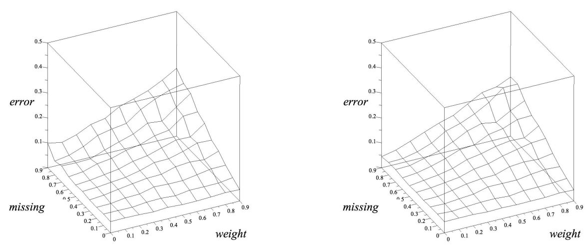


Figure 1: Average absolute error where data is fit by removing any training data with missing entries (left) and using proposed method (right) over 10 randomly generated datasets with w_5 and the percentage of x_5 entries removed.

5 CONCLUSION

We have developed and demonstrated a method for learning stable weighting vectors for datasets which have some data missing with respect to one of the variables. We conducted some experiments with artificial datasets to show that the method performs at least as reliably as other potential approaches to dealing with missing data pertaining to an independent variable. The method can obviously be extended to the problem of missing data pertaining to 2 or more variables, however with too many incomplete variables, the runtime may become impractical.

Future research for learning stable weights with miss-

[3]The error reported for this is hence the average of the 9 other tests.

ing data should take into account cases where the position of missing data cannot be imposed (see, e.g., [14], where a first attempt was proposed in terms of different variations of stability). Moreover, the search for efficient ways of dealing with missing data should be associated to the implementation with real datasets.

Acknowledgements

This research has been partially supported by the Government of Spain (grant TIN2012-32482), the Government of Madrid (CASI-CAM-CM project, grant S2013/ICE-2845)and Complutense University (research group 910149).

References

[1] J. Aczél. On mean values. *Bulletin of the American Math. Society*, 54, 1948.

[2] A. Amo, J. Montero, and E. Molina. Representation of consistent recursive rules. *European Journal of Operational Research*, 130:29–53, 2001.

[3] G. Beliakov. Construction of aggregation functions from data using linear programming. *Fuzzy Sets and Systems*, 160:65–75, 2009.

[4] G. Beliakov and S. James. Stability of weighted penalty-based aggregation functions. *Fuzzy Sets and Systems*, 226:1–18, 2013.

[5] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners.* Springer, Heidelberg, 2007.

[6] P. Bloomfield and W. Steiger. *Least Absolute Deviations. Theory, Applications and Algorithms.* Birkhauser, Boston, Basel, Stuttgart, 1983.

[7] H. Bustince, B. de Baets, J. Fernández, R. Mesiar, and J. Montero. A generalization of the migrativity property of aggregation functions. *Information Sciences*, 191:76–85, 2012.

[8] T. Calvo, G. Mayor, J. Torrens, J. Suñer, M. Mas, and M. Carbonell. Generation of weighting triangles associated with aggregation functions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(4):417–451, 2000.

[9] V. Cutello and J. Montero. Recursive connective rules. *Int. J. Intelligent Systems*, 14:3–20, 1999.

[10] J. Dujmović. Aggregation operators and observable properties of human reasoning.

[11] J. Dujmović, G. D. Tré, N. Singh, D. Tomasevich, and R. Yokoohji. Soft computing models in online real estate. In M. Jamshidi, V. Kreinovich, and J. Kacprzyk, editors, *Advance Trends in Soft Computing, WCSC 2013, Studies in Fuzziness and Soft Computing 312*, pages 77–91. Springer, 2013.

[12] J. J. Dujmović. The problem of missing data in LSP aggregation. In S. G. et al., editor, *Advances in Computational Intelligence, IPMU 2012, Part III, CCIS 299*, pages 336 – 346. Springer, Catania, Italy, 2012.

[13] D. Gómez and J. Montero. A discussion on aggregation operators. *Kybernetika*, 40:107–120, 2004.

[14] D. Gómez, K. Rojas, J. Montero, J. Rodríguez, and G. Beliakov. Consistency and stability in aggregation operators, an application to missing data problems. *Int. J. Computational Intelligence Systems*, 7:595–604, 2014.

[15] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation Functions.* Cambridge University Press, Cambridge, 2009.

[16] D. Nettleton and V. Torra. A comparison of active set method and genetic algorithm approaches for learning weighting vectors in some aggregation operators. *International Journal of Intelligent Systems*, 16(9):1069–1083, 2001.

[17] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2011. http://www.R-project.org/.

[18] K. Rojas, D. Gómez, J. Montero, and J. T. Rodríguez. Strictly stable families of aggregation operators. *Fuzzy Sets and Systems*, 228:44 – 63, 2013.

[19] K. Rojas, D. Gomez, J. Rodriguez, and J. Montero. Some properties of consistency in the families of aggregation functions. *Advances in Intelligent and Soft Computing*, 107:169–176, 2011.

[20] V. Torra and Y. Narukawa. The h-index and the number of citations: Two fuzzy integrals. *IEEE Transactions on Fuzzy Systems*, 16(3):795–797, 2008.

[21] Y. Torra and V. Narukawa. *Modeling Decisions. Information Fusion and Aggregation Operators.* Springer, 2007.

[22] R. R. Yager and A. Rybalov. Noncommutative self-identity aggregation. *Fuzzy Sets and Systems*, 85:73–82, 1997.

AGGREGATIONS OF MN-CONVEX FUNCTIONS ON COMPLETE LATTICES

Urszula Bentkowska and Józef Drewniak
University of Rzeszów
Faculty of Mathematics and Natural Sciences
ul. Pigonia 1, 35-310 Rzeszów, Poland
{ududziak,jdrewnia}@ur.edu.pl

Summary

The generalized condition of convexity was proposed by Aumann in 1933 for functions on intervals. It is convexity with respect to arbitrary binary means M and N (abbreviated to $MN-$convexity). We apply the Aumann condition with aggregation functions M, N and discuss $MN-$convex ($MN-$concave) functions on complete lattices. We will examine aggregations of such convex functions using binary aggregation functions.

Keywords: Convex function, Mean, $MN-$convexity, $MN-$concavity, Aggregation function, Lattice.

1 INTRODUCTION

The notion of convex functions was introduced in 1905 by Jensen [15]. Then the condition of convexity was generalized by many authors (cf. for example [21]). The most general condition of convexity, $MN-$convexity and $M-$convexity with respect to arbitrary means M and N, was proposed in 1933 by Aumann [2]. Recently Krassowska [18] examined $MN-$convexity for arbitrary binary operations M and N and also other authors have dealt with convexity proposed by Aumann [4]. We will consider $MN-$convexity with binary aggregation functions M and N.

For the aggregation functions defined on the intervals of the real line (cf. [7], [3], [26]) the more general setting was proposed for some types of aggregation functions, e.g. means on ordered sets [22], triangular norms on product lattices [8], triangular norms on partially ordered sets [27], OWA operators on complete lattices [20]. Specific problems for aggregation functions and their special types, were also considered, e.g. decomposability of triangular norms on product lattices [14], distributivity of triangular norms on complete lattices [16] and extensions of triangular norms on bounded lattices [24]. Aggregation functions on bounded partially ordered sets were considered by [9] and classification of such aggregation functions was presented in [17]. There was also many papers dealing with aggregation functions on some special kinds of lattices (c.f. for example [10] for uninorms in interval-valued fuzzy set theory).

The aim of this paper is to consider the notion of $MN-$convexity ($MN-$concavity) with respect to aggregation functions M, N on a complete lattice L and to discuss aggregation of such convex functions with proposing conditions guaranteeing preservation of such generalized convexity properties. Firstly, we put definition of an aggregation function and a mean and recall the most popular examples and families of means (Section 2). We also present (Section 3) some results on dominance between operations [25], [23] on partially ordered sets and lattices (cf. [11]). Then, the notion of $MN-$convex functions is presented (Section 4) and diverse operations on $MN-$convex functions are discussed with the problem of preservation of such convexity (Sections 5 and 6).

2 BINARY AGGREGATIONS

Throughout the paper we assume that $(P, \leqslant)$ is a partially ordered set (a poset for short) and $(L, \vee, \wedge)$ is a lattice (where necessary additional assumptions will be pointed out) where $\vee, \wedge$ are the supremum and infimum, respectively. In this paper, D will denote the order interval in P (independently of its type: closed, open, etc.). Concerning notions from lattice theory we follow Birkhoff [5]. In a bounded lattice L we use the notation: $\mathbf{1} = \sup L$, $\mathbf{0} = \inf L$.

In the literature there are diverse definitions of aggregation functions (cf. [7]). Aggregations on ordered

ISBN: 978–83–8012–519–3

sets were considered in [9], [17]. We follow the concept of an aggregation function presented in [13]. To simplify the notations, we consider binary aggregation functions only.

Definition 1 (cf. [13])**.** Let L be a complete lattice and $D \subset L, D \neq \emptyset$ be an interval. An aggregation function in D is a function $M : D^2 \to D$ which is isotone, i.e. for any $x_1, x_2, y_1, y_2 \in D$:

$$x_1 \leqslant y_1, \quad x_2 \leqslant y_2 \Rightarrow M(x_1, x_2) \leqslant M(y_1, y_2), \quad (1)$$

and fulfils boundary conditions

$$\inf_{x,y\in D} M(x, y) = \inf D, \quad \sup_{x,y\in D} M(x, y) = \sup D. \quad (2)$$

A function $M : D^2 \to D$ is called a mean in D if

$$x \wedge y \leqslant M(x, y) \leqslant x \vee y, \ x, y \in D. \quad (3)$$

Corollary 1. *If a function $M : D^2 \to D$ is isotone and idempotent, then property (3) is fulfilled (cf. [7], p. 10). Directly from condition (3) we see that every mean is idempotent.*

However, the Lehmer mean $M(x, y) = (x^2+y^2)/(x+y)$ in the lattice $((0,\infty), \max, \min)$ is not isotone (cf. [21], pp. 1-2). So there exist means which are not aggregation functions. And vice versa, there exist aggregation functions which are not means, for example product operation $\cdot$ in the lattice $([0, 1], \max, \min)$.

Definition 2. Let us consider lattice $(\mathbb{R}, \max, \min)$, $D \subset \mathbb{R}$. An aggregation function M is called homogeneous if $M(tx, ty) = tM(x, y)$ for $t, x, y \in D$.
An aggregation function M is called symmetric if $M(x, y) = M(y, x)$, $x, y \in D$.

Example 1 (cf. [12])**.** Projections $P_1(x, y) = x$, $P_2(x, y) = y$ for $x, y \in D$ and lattice operations $\vee$ and $\wedge$ are aggregation functions and means on any D. The most known examples of means (which are also aggregation functions) in the lattice $((0,\infty), \max, \min)$ are arithmetic mean A, geometric mean G, harmonic mean H, quadratic mean Q, logarithmic mean M_{log}, where

$$A(x, y) = \frac{x + y}{2}, \ G(x, y) = \sqrt{xy}, \ H(x, y) = \frac{2xy}{x + y},$$

$$Q(x, y) = \sqrt{\frac{x^2 + y^2}{2}}, \ M_{log}(x, y) = \begin{cases} \frac{x-y}{logx-logy}, & x \neq y \\ x, & x = y \end{cases}$$

for $x, y \in (0,\infty)$.

All the above means are symmetric. Using $\lambda \in [0, 1]$ one considers weighted means

$$W_\lambda(x, y) = \lambda x + (1 - \lambda)y, \ x, y \in (0,\infty),$$

which are nonsymmetric for $\lambda \neq \frac{1}{2}$. In particular,

$$W_1(x, y) = P_1(x, y) = x, \ W_0(x, y) = P_2(x, y) = y,$$

where $x, y \in (0,\infty)$. We may also consider weighted geometric means G_λ for $\lambda \in [0, 1]$, where

$$G_\lambda(x, y) = x^\lambda \cdot y^{(1-\lambda)}, \ x, y \in (0,\infty). \quad (4)$$

Example 2 (cf. [1], p. 287)**.** Typical examples of means in the lattice $((0,\infty), \max, \min)$ are quasi-linear means

$$M(x, y) = \varphi^{-1}(\lambda\varphi(x) + (1 - \lambda)\varphi(y)), \ \ x, y \in (0,\infty),$$

where $\lambda \in (0,\infty)$ and $\varphi : (0,\infty) \to \mathbb{R}$ is a continuous and strictly monotonic function. If $\lambda = \frac{1}{2}$, then we obtain quasi-arithmetic means. Special case of these means are power means with $\varphi(x) = x^p$ which are denoted by A_p, where $p \in \mathbb{R}$,

$$A_p(x, y) = \begin{cases} \left(\frac{x^p+y^p}{2}\right)^{\frac{1}{p}}, & p \neq 0 \\ G(x, y), & p = 0 \end{cases}, x, y \in (0,\infty). \quad (5)$$

Example 3. Addition $+$ is an aggregation function in the lattice $(\mathbb{R}, \max, \min)$. Multiplication $\cdot$ is an aggregation function in the lattices $((0,\infty), \max, \min)$ or $([0, 1], \max, \min)$.

Example 4. Let be given a lattice $([0,\infty), \max, \min)$. Since isotone functions in $D = [0,\infty)$, $f : D \to D$, form a lattice, composition of functions $K(f, g) = f \circ g$ is an aggregation function in the given lattice of isotone functions.

3 RELATION OF DOMINANCE

We may consider a relation of dominance between binary aggregations.

Definition 3 ([25], Definition 12.7.2)**.** Let $F, G : D^2 \to D$. Operation F dominates operation G ($F \gg G$), if for any $a, b, c, d \in D$

$$F(G(a, b), G(c, d)) \geqslant G(F(a, c), F(b, d)). \quad (6)$$

The standard example of dominance property gives the Minkowski inequality for $a, b, c, d \geqslant 0$ and $p \geqslant 1$ (cf. [6], p. 147):

$$(a^p + b^p)^{1/p} + (c^p + d^p)^{1/p} \geqslant ((a + c)^p + (b + d)^p)^{1/p},$$

which means that addition dominates binary operations B_p for $p \geqslant 1$, where

$$B_p(x, y) = (x^p + y^p)^{1/p}, \ x, y \in D.$$

Similarly, B_p dominates addition for $p \leqslant 1$, $p \neq 0$. In virtue of the Minkowski inequality we obtain

Lemma 1 ([11], Theorem 8)**.** *Let $p, q \in \mathbb{R}$. If $p \leqslant q$, then $A_p \gg A_q$.*

In particular

Corollary 2. *If $p \leqslant 1$, then $A_p \gg +$. If $p \geqslant 1$, then $+ \gg A_p$ (cf. Example 2) in lattice $((0,\infty), \max, \min)$.*

Proof. Let $a, b, c, d \in (0,\infty)$. Since $A_p \gg A$ for $p \leqslant 1$, we get

$$A_p(\frac{a+b}{2}, \frac{c+d}{2}) \geqslant \frac{A_p(a,c) + A_p(b,d)}{2}.$$

Moreover, means A_p are homogeneous, so

$$A_p(a+b, c+d) = 2 \cdot \frac{1}{2} \cdot A_p(a+b, c+d) =$$

$$2 \cdot A_p(\frac{a+b}{2}, \frac{c+d}{2}) \geqslant A_p(a,c) + A_p(b,d).$$

As a result $A_p \gg +$. The proof of the second statement is similar. □

Example 5. Let $a, b \in D$. If $F \equiv a, G \equiv b$ (constant operations), then

$$F \gg G \Leftrightarrow a \geqslant b.$$

Here dominance appears only in the case of comparability.

Lemma 2. *Let $F : D^2 \to D$ be isotone. Then $\wedge \gg F$, $F \gg \vee$, so for $a, b, c, d \in D \subset P$ we have, respectively*

$$F(a,b) \wedge F(c,d) \geqslant F(a \wedge c, b \wedge d), \quad (7)$$

$$F(a \vee b, c \vee d) \geqslant F(a,c) \vee F(b,d). \quad (8)$$

Proof. Let $a, b, c, d \in D \subset P$. $a \wedge c \leqslant a$ and $a \wedge c \leqslant c$. Moreover, $b \wedge d \leqslant b$ and $b \wedge d \leqslant d$. By monotonicity of F one has $F(a \wedge c, b \wedge d) \leqslant F(a,b)$ and $F(a \wedge c, b \wedge d) \leqslant F(c,d)$, so by the properties of infimum $\wedge$ condition (7) follows immediately. The second property may be proven analogically. □

Since lattice operations are isotone, we get

Corollary 3. *For lattice operations supremum $\vee$ (infimum $\wedge$) we have*

$$\wedge \gg \wedge, \quad \vee \gg \vee, \quad \wedge \gg \vee.$$

Example 6. Let us consider lattice $((0,\infty), \max, \min)$. The weighted geometric means G_λ dominate the product function $\cdot$ (cf. [23], Example 5.2). Additionally, it is easy to check that $\cdot \gg G_\lambda$, especially $\cdot \gg G$.

Example 7. It is easy to check that the weighted arithmetic means W_λ dominate $+$ and vice versa, operation $+$ dominates W_λ.

4 $MN-$CONVEXITY

We are discussing $MN-$convexity introduced by Aumann with respect to binary aggregation functions M, N defined on a poset P.

Definition 4 (cf. [2], [18])**.** A function $f : D \to D$ is called $MN-$convex with respect to aggregations M, N if

$$f(M(x,y)) \leqslant N(f(x), f(y)), \; x, y \in D.$$

Similarly, f is $MN-$concave if

$$f(M(x,y)) \geqslant N(f(x), f(y)), \; x, y \in D.$$

In the case $M = N$ we say about $M-$convexity ($M-$concavity) instead of $MM-$convexity ($MM-$concavity).

Example 8. Let $f, g : (0,\infty) \to (0,\infty)$. As particular cases of Definition 4 we get for $x, y \in (0,\infty)$ (cf. for example [21], Chapter 2):
$AA-$convexity ($AA-$concavity) known as Jensen one

$$f(\frac{x+y}{2}) \leqslant \frac{f(x)+f(y)}{2}, \; g(\frac{x+y}{2}) \geqslant \frac{g(x)+g(y)}{2},$$

$AG-$convexity ($AG-$concavity) known as logarithmic one

$$f(\frac{x+y}{2}) \leqslant \sqrt{f(x)f(y)}, \; g(\frac{x+y}{2}) \geqslant \sqrt{g(x)g(y)},$$

$GG-$convexity ($GG-$concavity) known as geometric one

$$f(\sqrt{xy}) \leqslant \sqrt{f(x)f(y)}, \; g(\sqrt{xy}) \geqslant \sqrt{g(x)g(y)},$$

$A\max-$convexity known as quasi-convexity

$$f(\frac{x+y}{2}) \leqslant \max(f(x), f(y)),$$

and $A\min-$concavity known as quasi-concavity

$$g(\frac{x+y}{2}) \geqslant \min(g(x), g(y)).$$

Example 9. Let L be an arbitrary complete lattice, $D \subset L$, $D \neq \emptyset$ be an interval. Constant functions are $MN-$convex ($MN-$concave) for arbitrary means M, N. Identity function, $I_L(x) = x$ for $x \in D$ is $M-$convex ($M-$concave) for any aggregation function M.

Theorem 1. *If M is a mean and $f : D \to D$ is monotonic (isotone or antitone), then f is $M\vee$-convex.*

Proof. If $M : D^2 \to D$ is a mean, then $x \wedge y \leqslant M(x,y) \leqslant x \vee y$. If f is antitone, then $f(x \wedge y) \geqslant f(M(x,y))$ and $f(x \wedge y) = f(x) \vee f(y)$. Finally, $f(M(x,y)) \leqslant f(x) \vee f(y)$, so f is $M\vee$-convex. Proof for an isotone function is similar. □

5 BINARY OPERATIONS ON $MN-$CONVEX FUNCTIONS

In this section we examine preservation of $MN-$convexity ($MN-$concavity) by aggregation functions.

Theorem 2. *Let $M, N, K : D^2 \to D$ be aggregation functions and $N \gg K$ ($K \gg N$). If $f, g : D \to D$ are $MN-$convex ($MN-$concave), then function $K(f, g)$ is $MN-$convex ($MN-$concave), where $K(f, g)(x) = K(f(x), g(x))$ for $x, y \in D$.*

Proof. Let $x, y \in D$. By $MN-$convexity of f, g and by assumption $N \gg K$ we obtain

$$K(f,g)(M(x,y)) = K(f(M(x,y)), g(M(x,y))) \leqslant$$
$$K(N(f(x), f(y)), N(g(x), g(y))) \leqslant$$
$$N(K(f(x), g(x)), K(f(y), g(y))) =$$
$$N(K(f,g)(x), K(f,g)(y)).$$

It means that function $K(f, g)$ is $MN-$convex. The proof for $MN-$concave functions is analogous. □

Since addition $+$ is an aggregation function in the lattice $(\mathbb{R}, \max, \min)$, then from Theorem 2 we get

Corollary 4. *Let $L = \mathbb{R}$, $D \in \{(-\infty, 0), (0, \infty), \mathbb{R}\}$ and $M, N : D^2 \to D$ be aggregation functions. If $N \gg +$ ($+ \gg N$) and $f, g : D \to D$ are $MN-$convex ($MN-$concave), then $f + g$ is $MN-$convex ($MN-$concave).*

By Theorem 2 and Lemma 2 we obtain

Corollary 5. *Let $M, N : D^2 \to D$ be aggregation functions. If $f, g : D \to D$ are $MN-$convex ($MN-$concave), then $h = f \vee g$ is $MN-$convex ($h = f \wedge g$ is $MN-$concave).*

Example 10. Power means A_p are examples of aggregation functions which dominate $+$ for $p \leqslant 1$ or are dominated by $+$ for $p \geqslant 1$ (cf. Corollary 2). So by Corollary 4 addition of functions preserves $MN-$convexity of functions f, g for arbitrary aggregation M, $N = A_p$ and $p \leqslant 1$ and addition of functions preserves $MN-$concavity of functions f, g for arbitrary aggregation M, $N = A_p$ and $p \geqslant 1$. Weighted arithmetic means W_λ dominate $+$ (cf. Example 7) and $\min \gg +$ (cf. Lemma 2) As a result, by Corollary 4, addition of functions preserves $MN-$convexity of functions f, g for arbitrary aggregation M and $N = W_\lambda$ ($N = \min$). Operation $+$ dominates W_λ (cf. Example 7) and $+ \gg \max$ (cf. Lemma 2), so it follows that by Corollary 4, addition of functions preserves $MN-$concavity of functions f, g for arbitrary aggregation M and $N = W_\lambda$ ($N = \max$).

Example 11. By Corollary 5 supremum preserves $MN-$convexity for quasi-linear means M, N and infimum preserves $MN-$concavity for quasi-linear means M, N in the lattice $(\mathbb{R}, \max, \min)$.

Since multiplication $\cdot$ is an aggregation function in the lattice $((0, \infty), \max, \min)$ or $([0, 1], \max, \min)$, then by Theorem 2 we get

Corollary 6. *Let $D \in \{[0, 1], (0, \infty)\}$, $M, N : D^2 \to D$ be aggregation functions. If $f : D \to D$ is $MN-$convex ($MN-$concave) and $N \gg \cdot$ ($\cdot \gg N$), then $f \cdot g$ is $MN-$convex ($MN-$concave).*

Example 12. Let us consider lattice $((0, \infty), \max, \min)$ or $([0, 1], \max, \min)$. The weighted geometric means G_λ dominate the product function $\cdot$ (cf. Example 6) and by Lemma 2, $\min \gg \cdot$, so by Corollary 6 product of functions $\cdot$ preserves $MN-$convexity for any aggregation function M and $N = G_\lambda$ ($N = \min$). By Lemma 2 one has $\cdot \gg \max$. Additionally, $\cdot \gg G_\lambda$ (cf. Example 6) so product of functions $\cdot$ preserves $MN-$concavity for any aggregation function M and $N = G_\lambda$ ($N = \max$).

Next example shows that lack of dominance $A \gg \cdot$ in $D = (0, \infty)$ involves lack of preservation of convexity (necessity of the condition).

Example 13. Let be given functions $f(x) = x^2$, $g(x) = \exp x$, $f, g : D \to D$. Functions f, g are convex ($M = N = A$). However, $f \cdot g$ is not convex (cf. [19], p. 124).

6 OTHER OPERATIONS ON $MN-$CONVEX FUNCTIONS

Theorem 3. *Let $L = (\mathbb{R}, \max, \min)$. Let $M, N : D^2 \to D$ be aggregation functions where N is homogeneous. If $f : D \to D$ is $MN-$convex ($MN-$concave), then αf is $MN-$convex ($MN-$concave), where $\alpha > 0$.*

Proof. Let $x, y \in D$, $(\alpha f)(x) = \alpha f(x), \alpha > 0$. By $MN-$convexity of f and by homogeneity of N we get

$$(\alpha f)(M(x,y)) = \alpha f(M(x,y)) \leqslant \alpha N(f(x), f(y)) =$$
$$N(\alpha f(x), \alpha f(y)) = N((\alpha f)(x), (\alpha f)(y)).$$

It means that αf is $MN-$convex. The proof for a $MN-$concave function is similar. □

Example 14. Power means A_p (e.g. arithmetic, geometric, quadratic, harmonic) and logarithmic mean M_{log} are homogeneous (cf. [21], p. 2). As a result, product of a function by a constant preserves $MN-$convexity ($MN-$concavity) for arbitrary aggregation M and $N = A_p$ ($N = M_{log}$).

In the sequel we will consider the inverse and the composition of $MN-$convex and $MN-$concave functions.

Theorem 4 (cf. [18], Remark 2). *If $f : D \to D$ is an isotone bijection and f is $MN-$convex ($MN-$concave), then f^{-1} is $NM-$concave ($NM-$convex). If $f : D \to D$ is an antitone bijection and f is $MN-$convex ($MN-$concave), then f^{-1} is $NM-$convex ($NM-$concave).*

Proof. Let $x, y \in D$. We will prove that if $f : D \to D$ is an isotone bijection and $MN-$concave, then f^{-1} is $NM-$convex. By assumption $f(M(x, y)) \geqslant N(f(x), f(y))$, so $M(x, y) \geqslant f^{-1}(N(f(x), f(y)))$. Since f is a bijection, there exist $a, b \in D$ such that $f(x) = a, f(y) = b$, $a, b \in D$. As a result, the previous inequality is equivalent to $M(f^{-1}(a), f^{-1}(b)) \geqslant f^{-1}(N(a, b))$, which means that f^{-1} is $NM-$convex. The rest of the properties may be justified in a similar way. □

Corollary 7. *If $f : D \to D$ is an isotone bijection, then f is $MN-$convex if and only if f^{-1} is $NM-$concave. If $f : D \to D$ is an antitone bijection, then f is $MN-$convex if and only if f^{-1} is $NM-$convex.*

Theorem 5. *Let $f : D \to D$. If f is an isotone $N-$convex ($N-$concave) function and $g : D \to D$ is $MN-$convex ($MN-$concave), then $f \circ g$ is $MN-$convex ($MN-$concave).*

Proof. Let $x, y \in D$, $(f \circ g)(x) = f(g(x))$. By $MN-$convexity of g and by $N-$convexity and monotonicity of f we obtain

$$(f \circ g)(M(x, y)) = f(g(M(x, y)) \leqslant f(N(g(x), g(y))) \leqslant$$

$$N(f(g(x)), f(g(y))) = N((f \circ g)(x), (f \circ g)(y)).$$

It means that $f \circ g$ is $MN-$convex. The proof for the other property is analogous. □

Corollary 8. *If $f : D \to D$ is an isotone $N-$convex ($N-$concave) function and $g : D \to D$ is an isotone $MN-$convex ($MN-$concave) function, then aggregation $K(f, g) = (f \circ g)$ is $MN-$convex ($MN-$concave).*

Theorem 6. *If f is an antitone $N-$convex ($N-$concave) function and $g : D \to D$ is $MN-$concave ($MN-$convex), then $f \circ g$ is $MN-$convex ($MN-$concave).*

Proof. Let $x, y \in D$, $(f \circ g)(x) = f((g(x))$. By $MN-$concavity of g and monotonicity of f we obtain

$$(f \circ g)(M(x, y)) = f(g(M(x, y))) \leqslant f(N(g(x), g(y))).$$

And by $N-$convexity of f we get

$$f(N(g(x), g(y))) \leqslant N(f(g(x)), f(g(y))) =$$

$$N((f \circ g)(x), (f \circ g)(y)),$$

so finally $f \circ g$ is $MN-$convex. The proof for the other condition is similar. □

7 CONCLUSIONS

In the paper $MN-$convex functions, where M, N are binary aggregation functions on complete lattices, were considered. Preservation of such convexity properties by binary aggregations and other typical operations on functions were discussed. Presented results coincide with the ones for standard convexity and concavity (cf. [19], Chapter V, p. 124 for addition of functions and Chapter VII for composition of functions and the converse function). Similar considerations to the ones presented in this paper my be extended to the case of n-ary aggregation functions.

Further, we are going to demonstrate applications of the presented results and provide sufficient conditions for other classes of functions in the context of preservation of the considered convexity. Moreover, other types of convexity, for example the ones considered not necessarily on complete lattices, maybe examined.

Acknowledgements

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge in Rzeszów, through Project Number RPPK.01.03.00-18-001/10.

References

[1] J. Aczél (1966). Lectures on functional equations and their applications. *Acad. Press*, New York.

[2] G. Aumann (1933). Konvexe Funktionen und die Induktion bei Ungleichungen zwischen Mittelwerten. *Sitzungsber. Bayer. Akad. Wiss., Math.-Naturwiss. Abt.* **3**, pp. 403-415.

[3] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation functions: a guide for practitioners. *Studies in Fuzziness and Soft Computing*, Springer, Berlin.

[4] B.A. Bhayo, L. Yin (2014). On the generalized convexity and concavity. *Cornell University Library* http://arxiv.org/1411.6586v1[math.CA], 24 Nov 2014.

[5] G. Birkhoff (1967). Lattice theory. *American Mathematical Society*, Providence.

[6] P.S. Bullen, D.S. Mitrinović, P.M. Vasić (1988). Means and their inequalities. *D. Reidel Publ. Com.*, Dordrecht.

[7] T. Calvo, A. Kolesárová, M. Komorniková, R. Mesiar (2002). Aggregation operators: Properties, classes and construction methods, in: T. Calvo et al. (eds.), *Aggregation Operators*, Physica-Verlag, Heildelberg, pp. 3-104.

[8] B. De Baets, R. Mesiar (1999). Triangular norms on product lattices. *Fuzzy Sets Syst.* **104**, pp. 61-76.

[9] M. Demirci (2006). Aggregation operators on partially ordered sets and their categorical foundations. *Kybernetika* **42** (3), pp. 261-277.

[10] G. Deschrijver, E. E. Kerre (2004). Uninorms in L^*-fuzzy set theory. *Fuzzy Sets Syst.* **148**, pp. 243-262.

[11] J. Drewniak, P. Drygaś, U. Dudziak (2005). Domination between multiplace operations, in: O.Hryniewicz et al. (eds.), *Issues in soft computing. Decisions and operations research*, EXIT, Warszawa, pp. 149-160.

[12] U. Dudziak (2006). Development of the mean concept. *Journal of Mathematics and Applications* **28**, pp. 25-39.

[13] M. Grabisch, J-L. Marichal, R. Mesiar, E. Pap (2011). Aggregation functions: Means. *Inform. Sciences* **181**, pp. 1-22.

[14] S. Jenei, B. De Baets (2003). On the direct decomposablity of t-norms on product lattices. *Fuzzy Sets Syst.* **139** (3), pp. 699-707.

[15] J.L.W.V. Jensen (1905). On convex functions and inequalities between mean values (Danish). *Nyt. Tidss. for Math.* **16B**, pp. 49-68 .

[16] F. Karaçal, D. Khadjiev (2005). $\vee$-distributive and infinitely $\vee$-distributive t-norms on complete lattices. *Fuzzy Sets Syst.* **151** (2), pp. 341-352.

[17] M. Komorníková, R. Mesiar (2011). Aggregation functions on bounded partially ordered sets and their classification. *Fuzzy Sets Syst.* **175**, pp. 48-56.

[18] D. Krassowska (2009). Iteration semigroups with generalized convex, concave and affine elements. *J. Inequal. Pure Appl. Math.* **10** (3), Paper No. 76, 9 p. (available on line at http://jipam.vu.edu.au).

[19] M. Kuczma (1985). An introduction to the theory of functional equations and inequalities: Cauchy's equation and Jensen's inequality. *PWN*, Katowice.

[20] I. Lizasoain, C. Moreno (2013). OWA operators defined on complete lattices. *Fuzzy Sets Syst.*, **224**, pp. 36-52.

[21] C.P. Niculescu, L.E. Persson (2006). Convex functions and their applications. A contemporary approach. *Springer Science*, New York.

[22] S. Ovchinnikov (1996). Means on ordered sets. *Math. Social Sci.* **32**, pp. 39-56.

[23] S. Saminger, R. Mesiar, U. Bodenhofer (2002). Domination of aggregation operators and preservation of transitivity. *Int. J. Uncertain., Fuzziness, Knowl.-Based Syst.* **10** Suppl., pp. 11-35.

[24] S. Saminger-Platz, E. P. Klement, R. Mesiar (2008). On extensions of triangular norms on bounded lattices. *Indagationes Mathematicae* **19** (1), pp. 135-150.

[25] B. Schweizer, A. Sklar (1983). Probabilistic metric spaces. *North-Holland*, Amsterdam.

[26] V. Torra, Y. Narukawa (2007). Modelling decisions: information fusion and aggregation operators. *Cognitive Technologies*, Springer, Berlin.

[27] D. Zhang (2005). Triangular norms on partially ordered sets. *Fuzzy Sets Syst.* **153** (2), pp. 195-209.

CONVEXITY WITH RESPECT TO AGGREGATIONS

Urszula Bentkowska and Józef Drewniak
University of Rzeszów
Faculty of Mathematics and Natural Sciences
ul. Pigonia 1, 35-310 Rzeszów, Poland
{ududziak,jdrewnia}@ur.edu.pl

Summary

The condition of $MN-$convexity was proposed by Aumann in 1933. It is convexity with respect to arbitrary binary means M and N (abbreviated to $MN-$convexity). Recently many authors have considered this notion with suitable pairs of means. We will examine and compare families of $MN-$convex and $MN-$concave functions.

Keywords: Convex function, Concave function, Mean, Aggregation, $MN-$convexity, $MN-$concavity, Geometric convexity, Logarithmic convexity, Quasi-convexity.

1 INTRODUCTION

The notion of convex functions was introduced in 1905 by Jensen [12]. Then the condition of convexity was generalized by many authors. For example, Hardy [10] introduced the condition of logarithmic convexity in 1915. In 1928, Montel [18] introduced the condition of geometric convexity (multiplicative convexity) and von Neumann [19] introduced the condition of quasi-convexity. Next, more general condition of convexity, i.e. $MN-$convexity and $M-$convexity with respect to arbitrary means M and N, was proposed in 1933 by Aumann [3].

Recently many authors have dealt with these generalizations. In particular, Niculescu [20] compared $MN-$convexity with relative convexity. Andersen et al. [2] examined inequalities implied by $MN-$convexity. Matkowski [17] examined inclusions between classes of $MM-$convexity. Krassowska [14] examined $MN-$convexity with arbitrary binary operations M and N.

The goal of this paper is a comparison of families of $MN-$convex functions for particular values of M and N, and under some relations between M and N. At first, we refer to the most popular examples and families of aggregation functions (Section 2). Next, we discuss elementary consequences of Aumann's definition (Section 3). Then, some inclusions between families of $MN-$convex functions are presented (Sections 4, 5). Finally, some open problems implied by the results presented are mentioned.

2 BINARY AGGREGATIONS

In this section and throughout the paper we will consider binary aggregation functions only.

Definition 1. Let $D \subset \mathbb{R}$ where D is an interval. An aggregation function in D (cf. [9]) is a function $M : D^2 \to D$ which is increasing, i.e. for any $x_1, x_2, y_1, y_2 \in D$:

$$x_1 \leqslant y_1, \quad x_2 \leqslant y_2 \Rightarrow M(x_1, x_2) \leqslant M(y_1, y_2), \quad (1)$$

and fulfils boundary conditions

$$\inf_{x,y \in D} M(x, y) = \inf D, \quad \sup_{x,y \in D} M(x, y) = \sup D. \quad (2)$$

A function $M : D^2 \to D$ is called a mean in D if (cf. [7, 17, 21])

$$\min(x, y) \leqslant M(x, y) \leqslant \max(x, y),\ x, y \in D. \quad (3)$$

Notice, that if a function $M : D^2 \to D$ is increasing and idempotent, i.e. $M(x, x) = x$ for all $x \in D$, then property (3) is fulfilled (cf. [6], p. 10). Directly from condition (3) we see that every mean is idempotent. However, the mean $M : D^2 \to D$, $M(x, y) = (x^2 + y^2)/(x + y)$ for $D = (0, \infty)$ fulfils (3) but is not increasing (this is one of Lehmer mean, [21], p. 2). As a result, a mean may not be an aggregation function. And vice versa, an aggregation function may not be a mean. Actually, there are many aggregation

ISBN: 978–83–8012–519–3

functions which are not means, for example multiplication $\cdot$ in $[0,1]$ or $[0,\infty]$, similarly addition $+$ in $[0,\infty]$ (families of aggregation functions which are not means one may find in [13]).

A mean M is called symmetric if $M(x,y) = M(y,x)$, $x,y \in D$. A mean is called homogeneous if for $t,x,y \in D$ it holds $M(tx,ty) = tM(x,y)$.

The most useful examples of aggregation function are means in $\mathbb{R}$, $[0,\infty]$, $(0,\infty)$ or $[0,1]$. Among them the most known are arithmetic mean A, geometric mean G, harmonic mean H, quadratic mean Q, logarithmic mean L, maximum and minimum, where for $D = (0,\infty)$

$$A(x,y) = \frac{x+y}{2},\ G(x,y) = \sqrt{xy},\ H(x,y) = \frac{2xy}{x+y},$$

$$Q(x,y) = \sqrt{\frac{x^2+y^2}{2}},\ L(x,y) = \begin{cases} \frac{x-y}{logx-logy}, & x \neq y \\ x, & x = y \end{cases}.$$

All the above means are symmetric. Using $\lambda \in [0,1]$ one considers weighted means

$$W_\lambda(x,y) = \lambda x + (1-\lambda)y,\ x,y \in D,$$

which are nonsymmetric for $\lambda \neq \frac{1}{2}$. In particular, we get projections:

$$P_1(x,y) = W_1(x,y) = x,\ P_2(x,y) = W_0(x,y) = y,$$

for $x,y \in D$. We may also consider weighted geometric means G_λ for $\lambda \in D$, where

$$G_\lambda(x,y) = x^\lambda \cdot y^{(1-\lambda)},\ x,y \in D. \tag{4}$$

Lemma 1 ([5], p. 130)**.** *The above examples of symmetric means are ordered, i.e.*

$$\min \leqslant H \leqslant G \leqslant L \leqslant A \leqslant Q \leqslant \max.$$

We use two ordered families of means.

Definition 2 ([5], p.132, 346)**.** Let $p \in \mathbb{R}$. By power means we call binary operations

$$A_p(x,y) = \begin{cases} \left(\frac{x^p+y^p}{2}\right)^{\frac{1}{p}}, & p \neq 0 \\ G(x,y), & p = 0 \end{cases},\ x,y \in D. \tag{5}$$

Binary operations given by (6) are called extended logarithmic means.

$$L_p(x,y) = \begin{cases} \left(\frac{x^{p+1}-y^{p+1}}{(p+1)(x-y)}\right)^{1/p}, & p \neq 0,-1 \\ \frac{1}{e}\left(\frac{y^y}{x^x}\right)^{1/(y-x)}, & p = 0 \\ L(x,y), & p = -1 \end{cases}, \tag{6}$$

where $x,y \in D = (0,\infty)$, $x \neq y$, $L_p(x,x) = x$, $x \in D$ and L_0 is called the intristic mean.

It is clear that $A_1 = A$, $A_{-1} = H$, $A_2 = Q$ and $A_0 = G$ is a limit case for $p \to 0$. Similarly we have $L_{-2} = G$, $L_1 = A$. Moreover

$$\lim_{p\to\infty} A_p(x,y) = \lim_{p\to\infty} L_p(x,y) = \max(x,y),$$

$$\lim_{p\to-\infty} A_p(x,y) = \lim_{p\to-\infty} L_p(x,y) = \min(x,y).$$

We also have

Lemma 2 ([11], p. 26)**.** *The power means (A_p) are ordered by index p, i.e. $p \leqslant q \Leftrightarrow A_p \leqslant A_q$.*

Lemma 3 ([5], p. 346)**.** *The extended logarithmic means (L_p) are ordered by index p, i.e. $p \leqslant q \Leftrightarrow L_p \leqslant L_q$.*

Example 1 (cf. [1], p. 287)**.** Quasi-linear means are described in the following way

$$M(x_1,x_2) = \varphi^{-1}(\lambda_1\varphi(x_1) + \lambda_2\varphi(x_2)),\ \ x_1,x_2 \in D,$$

where $D = (0,\infty)$, weights $\lambda_1,\lambda_2 \in D$ fulfil condition $\lambda_1+\lambda_2 = 1$ and $\varphi : D \to \mathbb{R}$ is a continuous and strictly monotonic function. If $\lambda_1 = \lambda_2 = \frac{1}{2}$, then we obtain quasi-arithmetic means. For $\varphi(x) = x^p$ (cf. (5)) we get power means.

3 GENERALIZED CONDITIONS OF CONVEXITY

We are discussing Aumann's definition and its elementary consequences.

Definition 3 ([3], cf. [14])**.** Let M,N be aggregation functions. A function $f : D \to D$ is called $MN-$convex ($f \in C(M,N)$) if

$$f(M(x,y)) \leqslant N(f(x),f(y)),\ x,y \in D.$$

Similarly, f is $MN-$concave ($f \in C^*(M,N)$) if

$$f(M(x,y)) \geqslant N(f(x),f(y)),\ x,y \in D.$$

In the case $N = M$ we say about $M-$convexity (concavity) instead of $MM-$convexity (concavity).

Example 2. Let $f,g : D \to D$, $D = [0,\infty]$. As particular cases of Definition 3 we get (cf. e.g. [21], Chapter 2): Jensen convexity ($f \in C(A,A)$) and concavity ($g \in C^*(A,A)$), where for $x,y \in D$

$$f(\frac{x+y}{2}) \leqslant \frac{f(x)+f(y)}{2},\ g(\frac{x+y}{2}) \geqslant \frac{g(x)+g(y)}{2},$$

logarithmic convexity ($f \in C(A,G)$) and concavity ($g \in C^*(A,G)$),

$$f(\frac{x+y}{2}) \leqslant \sqrt{f(x)f(y)},\ g(\frac{x+y}{2}) \geqslant \sqrt{g(x)g(y)},$$

geometric convexity ($f \in C(G,G)$) and concavity ($g \in C^*(G,G)$),

$$f(\sqrt{xy}) \leqslant \sqrt{f(x)f(y)},\ g(\sqrt{xy}) \geqslant \sqrt{g(x)g(y)},$$

quasi-convexity ($f \in C(A,\max)$) and quasi-concavity ($g \in C^*(A,\min)$),

$$f(\frac{x+y}{2}) \leqslant \max(f(x),f(y)),$$

$$g(\frac{x+y}{2}) \geqslant \min(g(x),g(y)).$$

Our goal is a comparison of diverse families of $MN-$convex functions with respect to commonly used means. Let ALL, INC, DEC and CST denote families of all possible, all increasing, all decreasing and all constant functions $f : D \to D$, respectively. It helps us in describing examples of function families from Definition 3.

Example 3. Let $c \in D$, $D = [0,\infty]$. We use the set inclusion '$\subset$' as an antisymmetric relation. Directly from Definition 3 we can see that the constant function $f(x) = c$, $x \in D$ is $MN-$convex and $MN-$concave for arbitrary means M, N, i.e. $CST \subset C(M,N)$, $CST \subset C^*(M,N)$. Similarly, the identity function is $MN-$convex or $MN-$concave if and only if M,N are comparable means, i.e. it is $MN-$convex for $M \leqslant N$ and $MN-$concave for $N \leqslant M$.

For the simplest five means $\min$, $\max, A, P_1, P_2$ we consider 50 function families from Definition 3, what is summarized in Table 1.

For some pairs M,N in Table 1 (M are in rows and N in columns) for $N = \min$ we get $C(M,N) = CST$, e.g. $M \in \{P_1, A, P_2, \max\}$. On the other hand $C(\min,\max) = C^*(\max,\min) = ALL$, because

$$f(\min(x,y)) \leqslant \max(f(x),f(y)),\ x,y \in D,$$

$$f(\max(x,y)) \geqslant \min(f(x),f(y)),\ x,y \in D.$$

Simultaneously, $C(\max,\min) = C^*(\min,\max) = CST$.

Means used in the conditions of quasi-convexity and quasi-concavity are not identical. Observation shows that classes $C^*(A,\max)$ and $C(A,\min)$ are trivial. They consist of constant functions, what will be explained now.

We pay attention to the family $C(A,\min)$. Let $x,y \in D$, $x < y$. If $y-x \leqslant \min(x,y)$, then $x = (2x-y+y)/2$ and $y = (x+2y-x)/2$. Thus a function $f \in C(A,\min)$ fulfils inequalities
$f(x) \leqslant \min(f(2x-y),f(y)) \leqslant f(y)$, $f(y) \leqslant \min(f(x),f(2y-x)) \leqslant f(x)$.
Therefore $f(x) = f(y) = const.$

Table 1: Extremal families of $MN-$convexity (concavity).

$C(M,N)$	min	P_1	A	P_2	max
min	INC	INC	INC	INC	ALL
P_1	CST	ALL	CST	CST	ALL
A	CST	CST	Jensen	CST	quasi
P_2	CST	CST	CST	ALL	ALL
max	CST	DEC	DEC	DEC	ALL
$C^*(M,N)$	min	P_1	A	P_2	max
min	ALL	DEC	DEC	DEC	CST
P_1	ALL	ALL	CST	CST	CST
A	quasi*	CST	Jensen*	CST	CST
P_2	ALL	CST	CST	ALL	CST
max	ALL	INC	INC	INC	INC

If $y - x > \min(x,y)$, then we can introduce the arithmetic sequence $z_k = x + kr$, $k = 0,1,\ldots,n$, where $r = (y-x)/n \leqslant \min(x,y)$. Then $z_{k+1} - z_k \leqslant \min(z_{k+1}, z_k)$ and similarly as above we obtain $f(z_k) = const. = f(x) = f(y)$. This proves that $C(A,\min) = CST$ and similarly we get $C^*(A,\max) = CST$.

Let us now consider an example of the family INC. Let $x,y \in D$, $x \leqslant y$. If $f \in C(\min,A)$, then $f(x) = f(\min(x,y)) \leqslant (f(x)+f(y))/2$. Thus $2f(x) \leqslant f(x) + f(y)$, which implies $f(x) \leqslant f(y)$ and $f \in INC$. Conversely, every increasing function belongs to $C(\min,A)$, because $\min \leqslant A$.

In a similar way we have obtained almost all the cases from Table 1. Only four cases from Table 1 denoted by: Jensen, Jensen*, quasi and quasi* coincide with cases from Example 2. These interesting families were investigated in many papers (cf. for example references in [15, 21]).

4 INCLUSIONS BETWEEN FAMILIES OF $MN-$CONVEX FUNCTIONS

Directly from Definition 3 we obtain

Lemma 4. *Let M, N_1, N_2 be arbitrary aggregation functions. If $N_1 \leqslant N_2$, then*

$$C(M,N_1) \subset C(M,N_2),\ C^*(M,N_2) \subset C^*(M,N_1).$$

Example 4. Inclusions in the above lemma need not be strict, because some families in Table 1 are equal, e.g. $C(P_1,\min) = C(P_1,P_2) = CST$. Since $\min \leqslant P_1$, then by Lemma 4 and Example 3 we get

$$CST \subset C(P_1,\min) \subset C(P_1,P_2) = CST.$$

Particular examples of strict inclusions were presented by Anderson et al. ([2], Remark 2.6, Example 2.7):

$$C(A,H) \subsetneq C(A,G) \subsetneq C(A,A),$$

$$C(G,H) \subsetneq C(G,G) \subsetneq C(G,A),$$

$$C(H,H) \subsetneq C(H,G) \subsetneq C(H,A).$$

Dually we get

$$C^*(A,A) \subsetneq C^*(A,G) \subsetneq C^*(A,H),$$

$$C^*(G,A) \subsetneq C^*(G,G) \subsetneq C^*(G,H),$$

$$C^*(H,A) \subsetneq C^*(H,G) \subsetneq C^*(H,H).$$

Because of Lemmas 1, 4 we get many inclusions between examples of families from Definition 3.

Corollary 1. *Let M be arbitrary aggregation function. Then:*

$$C(M,\min) \subset C(M,H) \subset C(M,G) \subset C(M,L)$$

$$\subset C(M,A) \subset C(M,Q) \subset C(M,\max),$$

$$C^*(M,\max) \subset C^*(M,Q) \subset C^*(M,A) \subset C^*(M,L)$$

$$\subset C^*(M,G) \subset C^*(M,H) \subset C^*(M,\min).$$

In particular for $M = A$ we have

$$CST = C(A,\min) \subset C(A,H) \subset C(A,G)$$

$$\subset C(A,L) \subset C(A,A),$$

$$CST = C^*(A,\max) \subset C^*(A,Q) \subset C^*(A,A),$$

i.e. some conditions do not generalize Jensen convexity (concavity).

Because $A_1 = A$ and $L_1 = A$, then directly from Lemmas 2, 3 we get

Theorem 1. *Let $p \in \mathbb{R}$, $D = (0,\infty)$.*

- *If $p < 1$, then we get restrictions of Jensen convexity*

$$C(A,A_p) \subset C(A,A), \quad C(A,L_p) \subset C(A,A).$$

- *If $p > 1$, then we get extensions of Jensen convexity*

$$C(A,A) \subset C(A,A_p), \quad C(A,A) \subset C(A,L_p).$$

- *If $p < 1$, then we get extensions of Jensen concavity*

$$C^*(A,A) \subset C^*(A,A_p), \quad C^*(A,A) \subset C^*(A,L_p).$$

- *If $p > 1$, then we get restrictions of Jensen concavity*

$$C^*(A,A_p) \subset C^*(A,A), \quad C^*(A,L_p) \subset C^*(A,A).$$

Similar results can be presented for other ordered families of aggregation functions.

5 CLASSES OF MONOTONE FUNCTIONS

Usually convex real functions are piecewise monotone. According to Krassowska [14] (Remark 1) we have

Lemma 5. *Let M_1, M_2, N be arbitrary aggregation functions. If $M_1 \leqslant M_2$, then*

$$DEC \cap C(M_1,N) \subset DEC \cap C(M_2,N),$$

$$DEC \cap C^*(M_2,N) \subset DEC \cap C^*(M_1,N),$$

$$INC \cap C(M_2,N) \subset INC \cap C(M_1,N),$$

$$INC \cap C^*(M_1,N) \subset INC \cap C^*(M_2,N).$$

Example 5. Inclusions in the above lemma need not be strict. From Table 1 we see that

$$CST = DEC \cap C(A,\min) = DEC \cap C(\max,\min),$$

$$INC = INC \cap C(A,\max) = INC \cap C(\min,\max),$$

$$DEC = DEC \cap C^*(\max,\min) = DEC \cap C^*(A,\min),$$

$$CST = INC \cap C^*(\min,\max) = INC \cap C^*(A,\max).$$

Using Lemmas 2, 3 in Lemma 5 we now get

Theorem 2. *Let $p \in \mathbb{R}$, $D = (0,\infty)$. If $p < 1$, then*
- $DEC \cap C(A,A_p) \subset DEC \cap C(A,A)$, $DEC \cap C(A,L_p) \subset DEC \cap C(A,A)$,
- $INC \cap C(A,A) \subset INC \cap C(A,A_p)$, $INC \cap C(A,A) \subset INC \cap C(A,L_p)$,
- $DEC \cap C^*(A,A) \subset DEC \cap C^*(A,A_p)$, $DEC \cap C^*(A,A) \subset DEC \cap C^*(A,L_p)$,
- $INC \cap C^*(A,A_p) \subset INC \cap C^*(A,A)$, $INC \cap C^*(A,L_p) \subset INC \cap C^*(A,A)$.

If $p > 1$, then
- $DEC \cap C(A,A) \subset DEC \cap C(A,A_p)$, $DEC \cap C(A,A) \subset DEC \cap C(A,L_p)$,
- $INC \cap C(A,A_p) \subset INC \cap C(A,A)$, $INC \cap C(A,L_p) \subset INC \cap C(A,A)$,
- $DEC \cap C^*(A,A_p) \subset DEC \cap C^*(A,A)$, $DEC \cap C^*(A,L_p) \subset DEC \cap C^*(A,A)$,
- $INC \cap C^*(A,A) \subset INC \cap C^*(A,A_p)$, $INC \cap C^*(A,A) \subset INC \cap C^*(A,L_p)$.

6 CONCLUSIONS

In the paper a comparison of families of $MN-$convex functions for particular aggregation functions M, N were presented. The basic consequences of Aumann definition were discussed and some inclusions between families of $MN-$convex functions were presented. Typical operations on $MN-$convex functions were investigated. The obtained results may be extended to the case of n-ary aggregation functions [4, 6, 22].

For the future work it would be interesting to consider the following tasks:
1. Characterization of pairs M, N for which $C(M,N) = CST$ or $C^*(M,N) = CST$. A certain hint for this problem provides the paper [8], where it is proved that the equation

$$f(M(x,y)) = N(f(x), f(y)),\ x,y \in D = (0,\infty)$$

has only constant solutions if one of the means M, N is quasi-arithmetic and the other one is non quasi-arithmetic.
2. Characterization of pairs M, N for which $C(M,N) = ALL$ or $C^*(M,N) = ALL$.
3. Characterization of families $C(M,N)$ and $C^*(M,N)$ for incomparable means M, N.

Moreover, it may be interesting to define $MN-$convex sets and to check how they behave under intersections, i.e. to consider if they create a system of generalized convexity defined in [16].

Acknowledgements

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge in Rzeszów, through Project Number RPPK.01.03.00-18-001/10.

References

[1] J. Aczél (1966). Lectures on functional equations and their applications. *Acad. Press*, New York.

[2] G.D. Anderson, M.K. Vamanamurthy, M. Vuorinen (2007). Generalized convexity and inequalities. *J. Math. Anal. Appl.* **335** (2), pp 1294-1308.

[3] G. Aumann (1933). Konvexe Funktionen und die Induktion bei Ungleichungen zwischen Mittelwerten. *Sitzungsber. Bayer. Akad. Wiss.*, Math.-Naturwiss. Abt. **3**, pp 403-415.

[4] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation functions: a guide for practitioners. *Studies in Fuzziness and Soft Computing*, Springer, Berlin.

[5] P.S. Bullen, D.S. Mitrinović, P.M. Vasić (1988). Means and their inequalities. *D. Reidel Publ. Com.*, Dordrecht.

[6] T. Calvo, A. Kolesárová, M. Komorniková, R. Mesiar (2002). Aggregation operators: Properties, classes and construction methods, in: T. Calvo et al. (eds.), Aggregation Operators. *Physica-Verlag*, Heildelberg, pp 3-104.

[7] A.L. Cauchy (1821). Cours d'analyse de l'Ecole Royale Polytechnique, vol. 1. *Analyse Algébraique*, Debure, Paris.

[8] R. Ger, T. Kochanek (2009). An inconsistency equation involving means. *Colloq. Math.* **115**, pp 87-99.

[9] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap (2011). Aggregation functions: Means. *Inf. Sci.* **181**, pp 1-22.

[10] G.H. Hardy (1915). The mean value of the modulus of an analytic function. *Lond. Math. Soc. Proc.* **14** (2),pp 269-277.

[11] G.H. Hardy, J.E. Littlewood, G. Polya (1952). Inequalities, *Cambridge Univ. Press*, Cambridge.

[12] J.L.W.V. Jensen (1905). On convex functions and inequalities between mean values (Danish). *Nyt. Tidss. for Math.* **16B**, pp 49-68.

[13] M. Komorníková, R. Mesiar (2011). Aggregation functions on bounded partially ordered sets and their classification. *Fuzzy Sets Syst.* **175**, pp 48-56.

[14] D. Krassowska (2009). Iteration semigroups with generalized convex, concave and affine elements. *J. Inequal. Pure Appl. Math.* **10** (3), Paper No. 76, 9 p. (available on line at http://jipam.vu.edu.au).

[15] M. Kuczma (1985). An introduction to the theory of functional equations and inequalities: Cauchy's equation and Jensen's inequality. *PWN*, Katowice.

[16] J. V. Llinares (2002). Abstract convexity. Some relations and applications. *Optimization* **51** (6), pp 797-818.

[17] J. Matkowski (2007). When M-convexity of a function implies its N-convexity for some means of M and N. *Int. Math. Forum 2*, (45-48), pp 2321-2331.

[18] P. Montel (1928). Sur les fonctions convexes et les fonctions sousharmoniques. *J. Math.* **7** (9), pp 29-60.

[19] J. von Neumann (1928). Zur Theorie der Gesellschaftsspiele. *Math. Ann.* **100**, pp 295-320.

[20] C.P. Niculescu (2003). Convexity according to means. *Math. Inequal. Appl.* **6** (4), pp 571-579.

[21] C.P. Niculescu, L.E. Persson (2006). Convex functions and their applications. A contemporary approach. *Springer Science*, New York.

[22] V. Torra, Y. Narukawa (2007). Modelling decisions: information fusion and aggregation operators. Cognitive Technologies. *Springer.*

ON INTERNAL OPERATORS AND MULTIVARIATE AGGREGATION FUNCTIONS

Humberto Bustince,
Laura De Miguel
and
Daniel Paternain
Departamento de Automática y Computación
Universidad Pública de Navarra
Pamplona, Spain
{bustince,laura.demiguel,daniel.paternain}@unavarra.es

María Jesús Campión
and
Esteban Induráin
Departamento de Matemáticas
Universidad Pública de Navarra
Pamplona, Spain
{mjesus.campion,steiner}@unavarra.es

Summary

We explore theoretical trends or research lines to discover new results concerning fusion operators defined on the Cartesian product of $n \geq 2$ copies of a nonempty given set, say S, and taking values on the same given set S. As a remarkable particular case, we pay attention to internal operators, namely those which give back (or "select") as output one of the n inputs involved, so that in this sense the fusion procedure does not give rise to new information during the process. A systematic use of this kind of results may appear when the given set is the unit interval $[0, 1]$ of the real line. Obviously, all this is closely related to aggregation of fuzzy sets of a universe U.

Among the questions to be explored, we consider n-variate operators for $n > 2$ that are not "decomposable" as a suitable iteration of bivariate (e.g.: binary) operators. Also, we study some topological, combinatorial or analytical aspects. Among them we analyze items as: a) continuity, b) possibility of defining orderings, or means or new binary relations, c) interpretation of binary operators as algebraic operations of a certain kind, d) latticial properties, e) anonimity and unanimity and/or f) characterizations of n-ary operators that have some special feature.

Keywords: Aggregation operators, Internal n-ary operators, Bivariate operators, Decomposable operators, Particular features of n-ary fusion functions.

1 INTRODUCTION AND MOTIVATION

Let X be a nonempty set. Let $n \geq 2$. A n-variate aggregation operator defined on X is a map $f : X^n \to X$.

In several contexts coming from the theory of aggregation operators it is typical to handle fusion mappings for which the output value must coincide with one of the coordinates of the vector input. To put an example concerning applications, we point out that in image processing they appear fusion algorithms where it is natural that the value of the intensity of a given pixel agrees with the value of the same pixel for some of the considered images (see e.g. [3, 13]).

This immediately induces the notion of an internal operator (see e.g. [5, 19]) or n-selector (see e.g. [15, 14, 9]). Thus, given a nonempty set S we consider mappings $f : S^n \to S$ such that for any vector input $(x_1, x_2, \ldots, x_n) \in S^n$ it holds true that the output value, namely, $f(x_1, x_2, \ldots, x_n) \in S$ coincides a fortiori with some coordinate $x_i \in \{x_1, x_2, \ldots, x_n\}$.

Among typical internal operators we could consider the following:

i) *Projections on one of the components:* Here, there exists $i \in \{1, \ldots, n\}$ such that $f(x_1, x_2, \ldots, x_n) = x_i$ for every $(x_1, \ldots, x_n) \in X^n$.

ii) *Medians, maxima and minima, percentiles:* Here, assuming that X is endowed with a linear order $\leq$, we rearrange the terms from smallest to biggest as regards the linear order, say $x_{i_1} \leq x_{i_2} \leq \ldots \leq x_{i_n}$ and then we take the median when n is an odd number, or a suitable percentile or maybe the maximum or the minimum with respect to $\leq$.

iii) *Mode:* We select the most repeated element (if any) or, say, the first element if there is a tie.

Classical situations where these kind of operators play

ISBN: 978–83–8012–519–3

an important role are related to aggregation functions of fuzzy sets. Notice that, a fuzzy set X being a map μ_X from a universe U into the unit interval $[0,1]$ of the real line, given n fuzzy subsets $X_1, \ldots X_n$ and a map $f : [0,1]^n \to [0,1]$ the composition that sends an element $u \in U$ to the element $f(\mu_{X_1}(u), \ldots, \mu_{X_n}(u)) \in [0,1]$ defines the indicator function of a new fuzzy set, so that, roughly speaking, the function f fuses the fuzzy sets $X_1, \ldots, X_n$ into a new one. For this reason, the particular case in which the former essential set S is the unit interval $[0,1]$ will deserve an special interest.

2 PARTICULAR FEATURES OF INTERNAL OPERATORS

Let X be a nonempty set. Given $n \geq 2$, let $f : X^n \to X$ stand for an internal operator defined on X. If we bear in mind the idea of "selecting" elements so that given a n-tuple $(x_1, \ldots, x_n) \in X^n$ the output $f(x_1, \ldots, x_n)$, that coincides with some x_i $(1 \leq i \leq n)$, should not depend on the order of the elements x's that appear as coordinates of the vector $(x_1, \ldots, x_n)$. In other words, for any rearrangement σ of the set $\{1, \ldots, n\}$ it should be true that $f(x_{\sigma_1}, \ldots, x_{\sigma_n}) = f(x_1, \ldots, x_n)$. This property is technically known as *anonimity*.

Also, when X is endowed with some topology τ, another typical requirement is that the aggregation operators $f : X^n \to X$ considered be *continuous* with respect to the given topology τ on X and the corresponding product topology on X^n.

Moreover, another typical property that have internal operators is technically called *unanimity*, namely, $f(x, \ldots, x) = x$ for every $x \in X$ and any internal operator $f : X^n \to X$. Obviously, this property can be defined in general –for any kind of operators– and, in that case, it does not become trivial or taken for granted.

At this stage, it is very important to remind the reader the following crucial fact: given a topological space (X, τ) the existence of a continuous, anonymous and unanimous aggregation operator[1] $F : X^n \to X$ strongly depends on the topology[2] τ with which the set X has been endowed. (See e.g. [2, 10, 11, 12, 7, 8, 6] for a further account). Thus, if X is a topological cellular complex (see e.g. [20, 7, 12]), the existence of a continuous anonymous and unanimous map on X is equivalent to the contractibility of X, so that X could be continuously deformed to a point. To put an example, if X is the unit circle S^1 endowed with the usual Euclidean topology inherited from that of the plane $\mathbb{R}^2$, even when $n = 2$ there is no continuous, unanimous and anonymous bivariate aggregation operator defined on X.

Fortunately, the unit interval $[0,1]$ of the real line, endowed with the usual Euclidean topology is indeed contractible, and consequently there exist continuous, anonymous and unanimous maps for every $n \geq 2$. An example is the arithmetic mean (which fails to be an internal operator). Another example, now internal, is the maximum mapping, where maxima are taken as regards the usual order of the real line[3].

3 DECOMPOSABILITY OF AGGREGATION OPERATORS

Assume now that $n > 2$. Let X be a nonempty set. Given an aggregation operator $f : X^n \to X$ we may wonder if f can be reached from successive composition of suitable bivariate mappings from X^2 to X. For instance, we may ask ourselves about the existence of $n-1$ bivariate mappings $g_1, g_2, \ldots, g_{n-1} : X^n \to X$ accomplishing that $f(x_1, x_2, \ldots, x_n) = g_{n-1}(x_1, g_{n-2}(x_2, g_{n-3}(x_3, \ldots g_1(x_{n-1}, x_n))) \ldots)$ holds

[1] Here, we do not mind if F is internal or not. The result is general, valid for any mapping $F : X^n \to X$.

[2] The spaces X on which a continuous, anonymous and unanimous aggregation operator $F : X^n \to X$ exists are also said *spaces with a topological n-mean*, and were studied from an abstract point of view, as a particular kind of topological spaces, by G. Aumann [2] in the 1940's. Then they reappeared in the later 1970's in contexts coming from applications of Mathematics into Economics and Social Choice, in which the classical Arrowian approach in Social Choice was substituted by a different one in the search for possibility results. In the new approach, the mathematical basis on which the problem considered leans exactly amounts to the existence of this kind of operators (namely continuous, anonymous and unanimous ones) on some suitable topological space on which the preferences are framed. In a battery of papers by Chichilnisky and Heal (see e.g. [10, 11, 12]), among others, it was proved again that some topological spaces cannot admit such operators. This is called the *Social Choice paradox* in this context. So, first some examples of those "bad" spaces were furnished –as, e.g., topological spheres–, and then a general theorem finally appeared. That theorem, namely the equivalence in suitable topological spaces between contractibility and the existence of a continuous, unanimous and anonymous operator for *any* $n > 2$ is difficult. It was proved, analyzed and generalized in several papers. In fact, this has been done by means of a wide sort of quite different mathematical techniques (namely, foliations, homotopy theory, measure theory, algebraic topology, and so on.). (See, e.g., [7, 8, 6]).

[3] This obvious example poses an important question: when trying to classify or characterize all the possible n-variate internal operators that could be defined on a nonempty set X, it is clear that some of them could be described as "taking the maximum" with respect to some linear order defined on X. But which n-variate operators could be described this way?

for every n-tuple $(x_1, \ldots, x_n) \in X^n$.

A more affordable particular case appears whenever all the mappings g_i $(1 \leq i \leq n-1)$ are the same, say g, so that f can be reached directly from g through iterated composition.

To put an example in this direction, if $g : X^2 \to X$ is a bivariate mapping such that $g(x,x) = x; g(x,y) = g(y,x)$ and $g(x, g(y,z)) = g(g(x,y),z)$ hold true for every $x,y,z \in X$, then we could define a trivariate aggregation operator $f : X^3 \to X$, with good properties as anonymity and unanimity, by just declaring that $f(x,y,z) = g(x, g(y,z))$ for every $x,y,z \in X$.

Therefore, in order to be sure that we are studying a "new theory" it would be crucial to detect, in a way, which n-variate aggregation operators cannot be decomposed as suitable composition of bivariate ones. As a matter of fact, if any n-variate operator could be reached from bivariate ones, it would happen that the classical theory of bivariate aggregation operators would be a tool good enough to interpret that "possible new" theory of n-variate operators with $n > 2$.

In this direction, several partial results are well-known (see e.g. [18, 4]). Thus, if we try to define the notion of a multivariate triangular norm $T_n : [0,1] \to [0,1]$ in order to extend to $n > 2$ the classical concept of a bivariate triangular norm $T : [0,1] \to [0,1]$ encountered in fuzzy set theory, we arrive to the fact that T_n can always we expressed as an iteration of some bivariate t-norm T. On the other hand, and looking for some result in the opposite direction, it can be proved that when $X = \{x,y,z\}$, endowed with a linear order such that $x < y < z$, then if $f : X^3 \to X$ is the median, $f(a,b,c)$ can never be expressed as $g(a, g(b,c))$ for some bivariate map $g : X^2 \to X$ and any $a,b,c \in X$. (See [4] for further details).

4 BINARY RELATIONS DEFINED BY MEANS OF INTERNAL OPERATORS

Let X be a nonempty set. Let $f : X^2 \to X$ be an internal operator, also known as a selector (see e.g. [15, 1, 14, 9]). Given $x,y \in X$ we may define a binary relation $\mathcal{R}$ on X by declaring that $x\mathcal{R}y \Leftrightarrow f(x,y) = y$, for every $x,y \in X$. It is interesting now to analyze to which kinds of binary relation $\mathcal{R}$ belongs. In particular, it is important here to say if $\mathcal{R}$ is a linear order on X. Questions of this kind have been studied in [9].

In addition $\mathcal{R}$ also defines a topology on X (see [16] for details), a subbasis of which is given by the family $\{\emptyset, X\} \bigcup \{L_x : x \in X\} \bigcup \{R_x : x \in X\}$, where $L_x = \{y \in X : y \neq x;\ y\mathcal{R}x\}$ and $R_x = \{y \in X : y \neq x;\ x\mathcal{R}y\}$ $(x \in X)$. The properties of this topology may be decisive in order to interpret and characterize the particular kinds of bivariate operator to which f belongs. (See e.g. [14, 15, 16] for further details).

Working now with a n-variate internal operator $f : X^n \to X$, but now with n strictly bigger than 2, we could try to interpret f in terms of binary relations as, for instance, the following one, denoted by $\mathcal{R}$ and given by $x\mathcal{R}y \Leftrightarrow f(x, \ldots, x, y) = y$, for every $x,y \in X$. Here the n-tuple $(x, \ldots x, y)$ has an x in the former $n-1$ positions, and y appears in the last position. Again, the case in which $\mathcal{R}$ is a linear order, as well as some topologies closely related to $\mathcal{R}$ (as in the reference [16]) could give new ideas about the properties of the internal operator f.

5 ALGEBRAIC ASPECTS OF INTERNAL OPERATORS

Let X be a nonempty set. A binary operator $f : X^2 \to X$, even in the case in which it is not an internal operator or selector), can always be interpreted as a binary operation, say $*_f$, acting on the set X, by just declaring that $x *_f y = f(x,y)$ for every $x,y \in X$. Again, the algebraic properties of $*_f$ could provide us with relevant information concerning the structure of the operator f considered. Some particular cases have been analyzed in the literature, mainly in the case in which $*_f$ acts as the join (or the meet) operation of a suitable lattice whose ground set is X, that is X can be given a (semi-)latticial structure by means of $*_f$. (See e.g. [6] for further details).

In this direction, the consideration and study, in the corresponding specialized literature, of n-variate aggregation operators with n bigger than 2, as algebraic n-ary operations on a nonempty set X is much more scarce than the analogous situation in the bivariate case. (See e.g. [18]).

Moreover, this last problem could be closely related to the aforementioned questions about decomposability of n-ary operators as iterations of a suitable bivariate one[4] (see e.g. [6, 9]).

[4]To put an example, if X is a nonempty set and $f : X^2 \to X$ is an operator that satisfies$f(x,x) = x$; $f(x,y) = f(y,x)$ and $f(x, f(y,z)) = f(f(x,y),z)$ for every $x,y,z \in X$, it can easily be extended to a n-variate operator, as a new map $F : X^n \to X$ given by $F(x_1, \ldots, x_n) = f(x_1, f(x_2, f(x_3, \ldots f(x_{n-1}, x_n) \ldots)$. This will have some simplifying properties: for instance, given $x,y \in X$ it holds that any finite sequence of applications of f in which only the elements x,y are involved finally amounts to $f(x,y) = x *_f y$. But the main question here would be to discover which n-ary aggregation operators can indeed be decomposed as an iteration of a suitable bivariate operator of this kind, or, equivalently, which n-ary

6 Functional equations related to aggregation operators

Another interesting line to be explored at this stage is that of functional equations that could satisfy a given n-variate operator $f : X^n \to X$, where X is a nonempty set and $n \geq 2$. To put an example, a bivariate operator f is said to be associative provided that it satisfies the functional equation $f(a, f(b, c)) = f(f(a, b), c)$ for every $a, b, c \in X$. Many typical operators encountered in fuzzy set theory (e.g., triangular norms and co-norms) must be associative (see e.g. [9]). Many other functional equations associated to bivariate aggregation functions on the unit interval $[0, 1]$ are often considered in this context, as, for instance, bisymmetry, permutability, migrativity or bimigrativity. (See e.g., [4, 17] for further details).

Once more, the consideration and analysis of functional equations that could be satisfied by n-ary aggregation operators (internal or not) in the case $n > 2$ would be of crucial interest in order to, by solving those functional equations, retrieve or characterize the main properties of the operators involved in the process. In this direction, we should point out that the theory of functional equations on more than two variables is quite scarce.

7 CONCLUDING REMARKS

Our intention, when preparing this note, was not that of furnishing results. Instead, we wanted to call the attention of colleagues, and potential readers, on the fact that in order to analyze well the behavior of n-ary internal operators, as well as the more general case of n-ary aggregation functions, when $n > 2$ there are many theoretical lines of research that remain quite unexplored.

Quite probably, a development in depth, in next future, of any of the sections of the present manuscript as a new piece of research could perhaps constitute by itself a whole article. And, in our opinion, any of the lines (or trends) commented here in this brief note could deserve interest enough to be explored and analyzed.

<<Learn from yesterday, live for today, hope for tomorrow. The important thing is not to stop questioning.>>

(Albert EINSTEIN, 1879 – 1955)

algebraic operations can be described as a suitable iteration of a certain binary operation on the same given set.

Acknowledgements

The authors acknowledge the partial support of the research Projects MTM2012-37894-C02-02 and TIN2013-40765-P (Spain) of the Spanish Ministry of Economy and Competitiveness, as well as the support furnished by the Research Service of the Public University of Navarre through a grant to co-author Laura de Miguel.

References

[1] A.A. Allam, M.Y. Bakeir and E.A. Abo-Tabl (2008). Some methods for generating topologies by binary relations. *Bulletin of the Malaysian Mathematical Sciences Society* bf 31(1), 35–45.

[2] G. Aumann (1943). Über Raüme mit Mittelbindungen. *Mathematische Annalen* **119**, 210–215.

[3] G. Beliakov, H. Bustince and D. Paternain (2012). Image reduction using means on discrete product lattices. *IEEE Transactions on Fuzzy Systems* **21**, 1070–1083.

[4] H. Bustince, M.J. Campión, F.J. Fernández, E. Induráin and M.D. Ugarte (2014). New trends on the permutability equation. *Aequationes Mathematicae* **88**, 211–232.

[5] H. Bustince, D. Paternain, L. De Miguel and R. Mesiar (2013). Some notions of internal operators. In *Proceedings of EUROFUSE Workshop 2013*. B. De Baets, J. Fodor and S. Montes (eds), pp. 43–48.

[6] M.J. Campión, J.C. Candeal, R. G. Catalán, J. R. De Miguel, E. Induráin and J. A. Molina (2011). Aggregation of preferences in crisp and fuzzy settings: functional equations leading to possibility results. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **19 (1)**, 89–114.

[7] J.C. Candeal and E. Induráin (1994). The Moebius strip and a social choice paradox. *Economics Letters* **45**, 407–412.

[8] J.C. Candeal and E. Induráin (1994). Medias generalizadas y aplicaciones. *Extracta Mathematicae* **9(3)**, 139–159.

[9] J.C. Candeal and E. Induráin (2012). Bivariate functional equations around associativity. *Aequationes Mathematicae* **84**, 137-155.

[10] G. Chichilnisky (1979) On fixed point theorems and social choice paradoxes. *Economics Letters* **3(4)**, 347–351.

[11] G. Chichilnisky (1982) Social aggregation rules and continuity. *Quarterly Journal of Economics* **87**, 337–352.

[12] G. Chichilnisky and G. Heal (1983) Necessary and sufficient conditions for a resolution of the social choice paradox. *Journal of Economic Theory* **31**, 68–87.

[13] M. Galar, A, Jurío, D. Paternain, C. López-Molina and H. Bustince (2013). Aggregation functions to combine RGB color channels in stereo matching. *Optics Express* **21**, 1247–1257.

[14] S. García-Ferreira, V. Gutev and T. Nogura (2008). Extensions of 2-point selections. *New Zealand Journal of Mathematics* **38**, 1–8.

[15] V. Gutev and T. Nogura (2001). Selections and order-like relations. *Applied General Topology* **2**, 205–218.

[16] E. Induráin and V. Knoblauch (2013). On topological spaces whose topology is induced by a binary relation. *Questiones Mathematicae* **36(1)**, 47–65.

[17] C. López-Molina, B. De Baets, H. Bustince, E, Induráin, A. Stupňanová and R. Mesiar (2014). Bimigrativity of binary aggregation functions. *Information Sciences* **274**, 225–235.

[18] J.-L. Marichal and P. Mathonet (2011). A description of n-ary semigroups polynomial-derived from integral domains. *Semigroup Forum* **83 (2)**, 241–249.

[19] G. Mayor and J. Martin (1999). Locally internal aggregation functions. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* **7**, 235–241.

[20] V. Rohlin and D. Fuchs (1981). *Premier cours de Topologie. Chapitres géométriques.* MIR. Moscou.

SYMMETRIC AGGREGATION OPERATORS ON COMPLETE LATTICES

Marta Cardin
Department of Economics
University Ca' Foscari of Venice
30121 Venice (Italy)
mcardin@unive.it

Summary

In many decision processes and information fusion data aggregation is required. There is often the need to aggregate data of varying dimension while aggregation operators are often considered for a fixed number of arguments. In many contexts inputs to be aggregated are of a qualitative nature and sometimes we need to evaluate objects with a scale that is not totally ordered. This paper analyzes the evaluation of sequences of ordinal input and of variable length. We consider various axioms against which different ranking methods can be compared.
Keywords: Completely distributive lattice, Aggregation functional, Sugeno integral.

1 INTRODUCTION

Aggregation operators are mathematical functions that are used to fuse information of several inputs in a single outcome (see [16] for a general background on aggregation theory). We focus on aggregation processes mapping arbitrary but finitely many argument from a set X to an object in X which is representative for the set of arguments itself. There exist a large number of different aggregation operators that differ on the assumptions on the inputs and about the information that we want to consider in the model.
There are many situations where inputs to be aggregated are qualitative and numerical values are used by convenience. Sometimes we need to evaluate objects with a scale that is not totally ordered. As the aim of this paper is to define and study a class of symmetric aggregation functionals in a purely ordinal context. We study aggregation functionals based on a complete lattices and we consider in particular the class of completely distributive lattices. A general approach to aggregation on bounded posets is considered also in [3], [12] and [20].
In [5], [6], [7] and [14] are investigated aggregation operators that consider inputs of variable length and satisfying a property called *arity-monotonicity*.
The plan of the paper is the following. In Section 2 we briefly mention some basic concepts and we provide the necessary definitions. Section 3 formulate our characterization of lattice aggregation operators while Section 4 is devoted to symmetric aggregation operators that are defined by a symmetric Sugeno integral.

2 BASIC NOTIONS AND PRELIMINARIES

In this section we give some basic notations and terminology and we introduce our framework from an axiomatic point of view. Some interpretations of our model are presented.

2.1 BACKGROUND IN LATTICE THEORY

To introduce our general framework we will need some preliminaries from lattice theory. Much of this terminology is well known and for further background in lattice theory we refer to, e.g., Birkhoff [2], Davey and Priestley [12] or Rudeanu [22]. A *lattice* is an algebraic structure $(L; \wedge, \vee)$ where L is a nonempty set and where $\wedge$ and $\vee$ are two binary operations, called *meet* and *join*, respectively, which satisfy the following axioms:

(i) for every $a \in L$, $a \vee a = a \wedge a = a$;

(ii) for every $a, b \in L$, $a \vee b = b \vee a$ and $a \wedge b = b \wedge a$;

(iii) for every $a, b, c \in L$, $a \vee (b \vee c) = (a \vee b) \vee c$ and $a \wedge (b \wedge c) = (a \wedge b) \wedge c$;

ISBN: 978–83–8012–519–3

(iv) for every $a, b \in L$, $a \wedge (a \vee b) = a$ and $a \vee (a \wedge b) = a$.

Every lattice L constitutes a partially ordered set endowed with the partial order $\leq$ such that for every $x, y \in L$, write $x \leqslant y$ if $x \wedge y = x$ or, equivalently, if $x \vee y = y$. If for every $a, b \in L$, we have $a \leqslant b$ or $b \leqslant a$, then L is said to be a *chain*. A lattice L is said to be *bounded* if it has a least and a greatest element, denoted by 0 and 1, respectively.
A lattice L is said to be *distributive*, if

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

for every $a, b, c \in L$. Clearly, every chain is distributive. A lattice L is said to be *complete* if for every $S \subseteq L$, its infimum $\bigwedge S := \bigwedge_{x \in S} x$ and supremum $\bigvee S := \bigvee_{x \in S} x$ exist. Clearly, every complete lattice is necessarily bounded.
A complete lattice L is said to be *completely distributive* if the following distributive law holds

$$\bigwedge_{i \in I} \Big(\bigvee_{j \in J} x_{ij} \Big) = \bigvee_{f \in J^I} \Big(\bigwedge_{i \in I} x_{if(i)} \Big),$$

for every doubly indexed subset $\{x_{ij} : i \in I, \, j \in J\}$ of L. Note that every complete chain (in particular, the extended real line and each product of complete chains) is completely distributive. Moreover, complete distributivity reduces to distributivity in the case of finite lattices. Throughout this paper we will assume that $(L; \wedge, \vee)$ is a bounded and completely distributive lattice. We indicate by $[n]$ the set $[n] = \{1, \ldots, n\}$. The cartesian product L^n also constitutes a lattice by defining the lattice operations componentwise. Observe that if L is bounded (distributive), then L^n is also bounded (resp. distributive).
We denote by 0_n and 1_n the least and the greatest elements, respectively, of L^n. For $a \in L$ and $x = (x_1, \ldots, x_n) \in L^n$ we define

$$a \wedge x = (x_1 \wedge a, \ldots, a \wedge x_n)$$

and

$$a \vee x = (a \vee x_1, \ldots, a \vee x_n).$$

Moreover or each $A \subseteq [n]$, we denote by $\mathbf{A}$ the *characteristic function* of A in L^n defined by $\mathbf{A}(x) = 1$ if $x \in A$ and $\mathbf{A}(x) = 0$ if $x \notin A$. We define $\mathbf{L} = \bigcup_{n \in \mathbb{N}} L^n$. Then $\mathbf{L}$ is the set of finite sequence of elements in L of any length. $\mathbf{L}$ is a lattice but not necessarily complete or bounded.

2.2 LATTICE-VALUED SUGENO INTEGRALS

The following definitions are natural extensions of the well known concepts of real-valued fuzzy measures and their associated integrals. We follow the approach proposed by Greco in [18] and more recently by Ban and Fechete in [1] for lattice-valued measures and integrals and we refer to [16]for the standard case. Let $(A, \mathcal{A})$ be a measurable space and L a lattice (if A is finite we assume that $\mathcal{A} = 2^A$).
A *fuzzy measure* on A with values in L is a function $m \colon \mathcal{A} \to L$ such that $m(X) \leq m(Y)$ whenever $X \subseteq Y$. We do not assume necessarily that $m(\emptyset) = 0$, $m(A) = 1$.
A function $f \colon A \to L$ is said to be *measurable* if the sets $\{x : f(x) \leqslant a\}$ and $\{x : f(x) \geqslant a\}$ are elements of $\mathcal{A}$ for every $a \in L$. We will use $\{f \geqslant x\}$ to indicate the weak upper level set $\{t \in L : f(t) \geqslant x\}$. We denote by $\mathcal{F}$ the set of the measurable functions $f \colon A \to L$.
We are interested in a class of integral functionals defined on a complete lattice. Following the approach in [18] we consider the functionals S_l, S_u defined by :

$$S_l(m, f) = \bigvee_{t \in L} (t \wedge m(\{f \geqslant t\})) \quad \text{and}$$

$$S_u(m, f) = \bigwedge_{t \in L} (t \vee m(\{s : f(s) \nleq t\}))$$

where $f \colon A \to L$ is a measurable function. If L is a completely distributive lattice it can be proved that $S = S_l = S_u$ (see[18]) and then functional S extends Sugeno integral to an ordinal framework and so is called the *lattice-valued Sugeno integral* of f with respect to m.
The concept of comonotonicity emerges quite naturally in many different fields such as aggregation theory, decision theory. We refer to Denneberg [13] for the definition as well as for different characterizations of comonotonicity. In this paper we consider a generalization of the notion of comonotonicity to the case of lattice-valued functions (see [4]).

If A is a non empty set and L is a lattice two function $f, g \colon A \to L$ are said to be *comonotone* if

$$\text{either} \quad \{f \geqslant t\} \supseteq \{g \geqslant t\} \quad \text{or} \quad \{g \geqslant t\} \supseteq \{f \geqslant t\}$$

for every $t \in L$. It can be prove (see [4]) that a lattice-valued Sugeno integral satisfy the following properties:

(i) (homogeneity) $S(f \wedge \mathbf{a}) = S(f) \wedge a$ for every $a \in L$ and for every $f \in \mathcal{F}$

(iii) (comonotone maxitivity): $S(f \vee g) = S(f) \vee S(g)$ if f, g are comonotone elements of $\mathcal{F}$.

where $f \colon A \to L$ is a measurable function and $\mathbf{a}$ is the constant map with value a.

2.3 OUR SETTING

We introduce some classes of aggregation operators defined on lattice L from an axiomatic point of view.
As an example we can refer to social decisions where different individuals rank alternatives as in the Arrovian framework. We are interested in different evaluating situations in which the set of voters might vary. We consider n voters that are allowed to express their opinion on the candidates, social alternatives or proposals by a ordering over the set of alternatives and we want to aggregate the individual preferences in a social ordering. Also in multicriterial choice alternatives ordered by several criteria can be ordered by the overall merit.
We can also consider the aggregation of a collection of individual "judgments" in a social "judgment" as in[12] where a "judgment" is represented by a set of sentences in some logical language.
An aggregation operator $F\colon \mathbf{L} \to L$ is said to be *nondecreasing* if, for every $x, y \in L^n$ such that $x_i \leqslant y_i$, for every $i \in A$, we have $F(x) \leqslant F(y)$. In this paper by a (*lattice*) *aggregation operator* on L we mean a nondecreasing mapping $F\colon \mathbf{L} \to L$.
We are particularly interested in certain aggregation operators that are Sugeno integrals. We also consider *symmetric* lattice aggregation operators on $\mathbf{L}$ where symmetry holds if $F(x_1, \dots, x_n) = F(x_{\pi(1)}, \dots, x_{\pi(n)})$ for every permutation π of $[n]$.
We note that symmetry is a very natural properties in our framework (see also [5], [6], [7] and [14]), in social choice symmetry is often called anonymity and it means that every individual is endowed with the same voting power as it is used in usual voting procedure.

3 SUGENO INTEGRAL BASED OPERATORS

The following proposition characterizes Sugeno integral in our framework.

Proposition 1. *If $(L; \wedge, \vee)$ is a completely distributive lattice and F is an aggregation operator $F\colon \mathbf{L} \to L$ the following conditions are equivalent :*

i) *for every $a \in L$ and $x, y \in L^n$ such that $x_i \leqslant y_i$ for every $i \in [n]$ we have that*

$$F(x \vee (a \wedge y)) = F(x) \vee (a \wedge F(y))$$

ii) *F is nondecreasing and for every $x, y \in L^n$, x,y comonotone and $a \in L$,*

$$F(a \wedge x) = a \wedge F(x) \ \text{ and } \ F(x \vee y) = F(x) \vee F(y)$$

iii) *for every $A \subseteq [n]$ there exists $m(n, A) \in L$ such that*

$$F(x) = \bigvee_{A \subseteq [n]} \left(m(n, A) \wedge \bigwedge_{i \in A} x_i \right)$$

for every $x \in L^n$

iv) *for every $A \subseteq [n]$ there exists $m'(n, A) \in L$ such that*

$$F(x) = \bigwedge_{A \subseteq [n]} \left(m'(n, A) \vee \bigvee_{i \in A} x_i \right)$$

for every $x \in L^n$.

Proof. Let $(A, \mathcal{A})$ is a measurable space and $\mathcal{F}$ the set of the measurable functions $f\colon A \to L$. By Theorem 3.2 in [18] a functional $I\mathcal{F}\colon \to L$ is a Sugeno integral if when $f, g\colon A \to L$ are two functions such that $f \leq g$ and $a \in L$

$$I\left(f \vee (a \wedge g)\right) = I(f) \vee (a \wedge I(g))\,.$$

By Theorem 3.2 in [18] if condition i) is satisfied for every $n \in \mathbb{N}$ the functional F in L^n is a Sugeno integral with respect to a monotone measure defined in $[n]$ and then condition ii) is verified. Conversely if condition ii) holds we can prove that x and $a \wedge y$ are comonotone if $x_i \geqslant y_i$ for every $i \in [n]$ and so condition ii) implies condition i). In fact if $a \geq t$ then $\{x \geqslant t\} \supseteq \{a \wedge y \geqslant t\}$ and if $a \ngeq t$ we get $\{a \wedge y \geqslant t\} = \emptyset$ and then the functions $x, a \wedge y$ are comonotone.
Hence if the functional F satisfies condition i) or ii) by Theorem 3.2 in [18] for every n there exists a fuzzy measure $m_n\colon 2^{[n]} \to L$ such that $F(x) = \bigvee_{t \in L} \left(t \wedge m_n(\{x \geqslant t\})\right)$ for every $x \in L^n$. Hence we can prove that $F(x) = \bigvee_{A \subseteq [n]} \left(m_n(A) \wedge \bigwedge_{i \in A} x_i\right)$ for every $x \in L^n$ so condition iii) is easily proved if we define $m(n, A) = m_n(A)$ for every A such that $A \subseteq [n]$.
We note that $F(x) = \bigwedge_{t \in L} \left(t \vee m_n(\{i : x_i \nleq t\})\right)$ for every $x \in L^n$ by Theorem 3.2 in [18] and that

$$\bigwedge_{t \in L} \left(t \vee m_n(\{i : x_i \nleq t\})\right) = \bigwedge_{A \subseteq [n]} \left(m([n] \setminus A) \vee \bigvee_{i \in A} x_i \right)$$

for every $x \in L^n$ as is easily seen.
Then we can prove condition iv) posing $m'(n, A) = m_n([n] \setminus A)$ for every A with $A \subseteq [n]$.
Moreover if an aggregation operator $F\colon \mathbf{L} \to L$ is defined by condition iii) or condition iv) we get that the operator F on L^n satisfies homogeneity and comonotone maxitivity conditions. □

Note that $F(0_n) = 0$ for every $n \in \mathbb{N}$ if and only if $m(n, \emptyset) = 0$ for every $n \in \mathbb{N}$.
If we assume that L is a complete chain we can prove the following result.

Proposition 2. *If L is a complete chain and F is an aggregation operator $F\colon \mathbf{L} \to L$ the following conditions are equivalent :*

i) *for every $x, y \in L^n$ such that $x_i \leqslant y_i$ for every $i \in [n]$ and $a \in L$ if $a \leqslant F(y)$ we have that $F(x \vee (a \wedge y)) = F(x) \vee a$ and if $a \geqslant F(y)$ we have that $F(x \vee (a \wedge y)) = F(y)$*

ii) *for every $A \subseteq [n]$ there exists $m(n, A) \in L$ such that*

$$F(x) = \bigvee_{A\subseteq[n]} \left(m(n, A) \wedge \bigwedge_{i\in A} x_i \right)$$

for every $x \in L^n$.

Proof. If L is a complete chain for every $x, y \in L^n$ and $a \in L$ such that $x_i \leqslant y_i$ for every $i \in [n]$ or $a \leqslant F(y)$ and then $F(x) \vee (a \wedge F(y)) = F(x) \vee a$ or $a \geqslant F(y)$ and then $F(x) \vee (a \wedge F(y)) = F(y)$. □

4 SYMMETRIC AGGREGATION OPERATORS

We introduce some definitions considering the case in which symmetry holds. If $x = (x_1, \ldots, x_n) \in L^n$ and $y = (y_1, \ldots, y_m) \in L^m$ the *concatenation* (x, y) denotes $x = (x_1, \ldots, x_n, y_1, \ldots, y_m)$.
Moreover $1x = x$ and $2x = (x, x)$ and $nx = (x, (n-1)x)$. If $(L; \wedge, \vee)$ is a bounded and completely distributive lattice and F is a symmetric aggregation operator $F\colon \mathbf{L} \to L$ we say that F is *arity-monotonic* if $F(x) \leqslant F(x, y)$ for every $x \in L^n, y \in L^m$.
If $x \in L^n$ for every $i \leq n$ we define $x_{(i)}$ by

$$x_{(i)} = \bigvee_{A\subseteq[n]|A|=i} \bigwedge_{j\in A} x_j$$

where $|A|$ is the cardinality of the set A.
Note that inn the Arrovian framework there are i voters that approve the ranking $x_{(i)}$.
If L is a bounded chain $(\cdot)$ is a permutation on $[n]$ which arranges the elements of the vector by decreasing values that is such that $x_{(1)} \geq x_{(2)} \geq, \ldots, \geq x_{(n)}$.
The following proposition consider symmetric aggregation operators. It can be proved that Sugeno integral-based symmetric aggregation operators generalize order weighted maximum to lattices setting.

Proposition 3. *If $(L; \wedge, \vee)$ is a completely distributive lattice and F is a symmetric and arity-monotonic aggregation operator $F\colon \mathbf{L} \to L$ the following conditions are equivalent :*

i) *for every $a \in L$ and $x, y \in L^n$ such that there exists a permutation π of $[n]$such that $x_{\pi(i)} \leqslant y_i$ for every $i \in [n]$ we have that*

$$F(x \vee (a \wedge y)) = F(x) \vee (a \wedge F(y))$$

ii) *F is nondecreasing and for every $x, y \in L^n$ and $a \in L$,*

$$F(a \wedge x) = a \wedge F(x) \text{ and } F(x \vee y) = F(x) \vee F(y)$$

iii) *for every $i \in \mathbb{N}$ there is an element $w_i \in L$ such that if $x \in L^n$,*

$$F(x) = \bigvee_{i\leq n} w_i \wedge x_{(i)}.$$

Moreover if $i \leq i'$ we have $w_i \leqslant w_{i'}$.

Proof. We note that if F is a symmetric aggregation operator two elements of L^n are comonotone and then if F is comonotone maxitive is maxitive i.e. for every $x, y \in L^n$ we have that $F(x \vee y) = F(x) \vee F(y)$.
By Proposition 1 if conditon i) or condition ii) is satisfied F is a Sugeno integral in L^n and then for every $n \in \mathbb{N}$ there exists a fuzzy measure $m_n\colon 2^{[n]} \to L$ such that $F(x) = \bigvee_{A\subseteq[n]} \left(m_n(A) \wedge \bigwedge_{i\in A} x_i \right)$.
Since F is a symmetric operator the measure $m_n(A)$ does not depend on n and depends only on the cardinality of the set A hence we define $w_i = m_n(A) = F(\mathbf{A}) = F(i1)$ if $|A| = i$. We get

$$\bigvee_{|A|=i} \left(w_i \wedge \bigwedge_{j\in A} x_j \right) = w_i \wedge x_{(i)}$$

and then condition iii) is easily proved. Conversely conditions iii) implies that F on L^n is non decreasing, maxitive and homogeneous. Being F arity-monotonic if $i \leq i'$ then $F(i1) \leqslant F(i'1)$ hence $w_i \leqslant w_{i'}$. □

If L is a complete chain an arity-monotonic aggregation operator $F\colon \mathbf{L} \to L$ that satisfies the conditions of Proposition 3 for every $i \in \mathbb{N}$ there is an element $w_i \in L$ such that if $x \in L^n$

$$F(x) = \bigvee_{i\leq n} w_i \wedge x_{(i)}.$$

where $x_{(1)} \geq x_{(2)} \geq, \ldots, \geq x_{(n)}$ then it is an ordered weighted maximum.

References

[1] A. Ban, I. Fechete (2007). Componentwise decomposition of some lattice-valued fuzzy integrals. *Information Sciences*,**177**, pp 1430–1440.

[2] G. Birkhoff (1967). *Lattice Theory*, 3rd edition, Colloquium Pub., Vol. **25.** American Mathematical Society, Providence, R.I.

[3] M. Cardin, M. Couceiro (2011). Invariant functionals on completely distributive lattices, *Fuzzy Sets and Systems*,**167(1)**, pp 45–56.

[4] M. Cardin (2012). A quantile approach to integration with respect to non-additive measures. In *Modeling Decisions for Artificial Intelligence*, Lecture Notes in Computer Science, **7647,** Springer, Berlin, pp. 139–148.

[5] A. Cena, M. Gagolewski (2013). OM3: Ordered maxitive, minitive, and modular aggregation operators Part I: Axiomatic analysis under arity-dependence. In: Bustince H. et al. (Eds.), *Aggregation Functions in Theory and in Practise (AISC 228)*, Springer, 93–103.

[6] A. Cena, M. Gagolewski (2013). OM3: Ordered maxitive, minitive, and modular aggregation operators Part II: A simulation study. In: Bustince H. et al. (Eds.), *Aggregation Functions in Theory and in Practise (AISC 228),* Springer, pp 105-115.

[7] A. Cena, M. Gagolewski (2015). OM3: Ordered maxitive, minitive, and modular aggregation operators axiomatic and probabilistic properties in an arity-monotonic setting.*Fuzzy Set Syst* **264**, pp 138–159.

[8] M. Couceiro, J.-L. Marichal (2009). Characterizations of discrete Sugeno integrals as lattice polynomial functions. *Proceedings of the 30th Linz Seminar on Fuzzy Set Theory (LINZ2009)*, pp 1720.

[9] M. Couceiro, J.-L. Marichal (2010). Representations and characterizations of polynomial functions on chains. *Journal of Multiple-Valued Logic and Soft Computing*, **16**, 1-2, pp 65-86.

[10] M. Couceiro, J.-L. Marichal (2010). Characterizations of discrete Sugeno integrals as polynomial functions over distributive lattices. *Fuzzy Sets and Systems* **161**, pp 694-707.

[11] T. Daniëls, E. Pacuit (2009). A General Approach to Aggregation Problems *Journal of Logic and Computation*, **19(3)**, pp 517–536, Oxford University Press.

[12] B.A. Davey, H.A. Priestley(2002). *Introduction to Lattices and Order*. Cambridge University Press, New York, second edition.

[13] D. Denneberg (1994). *Non-additive measure and integral*, Kluwer Academic Publisher, Dordrech.

[14] M. Gagolewski, P. Grzegorzewski (2010). Arity-monotonic extended aggregation operators. In: Hllermeier E., Kruse R., Hoffmann F. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems (CCIS 80)*, Springer, 693–702.

[15] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. W. Mislove, D. S. Scott (2003). *Continuous Lattices and Domains*, Cambridge University Press, Cambridge.

[16] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap (2009). *Aggregation Functions*, Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge.

[17] G. Grätzer (2003). *General Lattice Theory* Birkhäuser Verlag, Berlin.

[18] G. H. Greco (1987). Fuzzy Integral and Fuzzy Measures with Their Values in Complete Lattices, *Journal of Mathematical Analysis and Applications* **126**, pp 594–603.

[19] J.-L. Marichal(2008). Weighted lattice polynomials. *Discrete Mathematics*, **309(4)**,814–820.

[20] M. Komorníková, R. Mesiar, Aggregation Functions on Bounded Posets (2011). *35 years of Fuzzy sets Theory*, STUDFUZZ **261**, pp 3–17, Springer.

[21] G. N. Raney (1952). Completely distributive complete lattices, *Proceedings of the American Mathematical Society*, **3**, pp 677–680.

[22] S. Rudeanu(2001). *Lattice Functions and Equations.* Springer Series in Discrete Mathematics and Theoretical Computer Science. Springer-Verlag London Ltd.

AGGREGATION AND SOFT CLUSTERING OF INFORMETRIC DATA

Anna Cena
Systems Research Institute,
Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
cena@ibspan.waw.pl

Marek Gagolewski
(1) Systems Research Institute,
Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
(2) Faculty of Mathematics and Information Science,
Warsaw University of Technology
ul. Koszykowa 75, 00-661 Warsaw, Poland
gagolews@ibspan.waw.pl

Summary

The aim of this contribution is to inspect possible applications of clustering techniques computed over a set consisting of non-increasingly ordered vectors of possibly non-conforming lengths. Such data sets appear in the field of informetrics, where one needs to evaluate the quality of information items, e.g research papers, and their producers. In this paper we investigate the notion of cluster centers as an aggregated representation of all vectors from a given cluster and analyze them by means of aggregation operators.

Keywords: clustering, fuzzy clustering, c-means algorithm, distance, producers assessment problem

1 INTRODUCTION

The Producers Assessment Problem (PAP, see e.g. [7]) concerns the evaluation of a set of information resources producers according to both number and quality of their products (e.g. forum posts, research publications, etc.). More formally, this problem may be modeled by a set of vectors $\{\mathbf{x}^{(1)},\dots,\mathbf{x}^{(l)}\}$, where $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)},\dots,x_{n_i}^{(i)})$, $x_1^{(i)} \geq x_2^{(i)} \geq \dots \geq x_{n_i}^{(i)}$ with possibly $n_i \neq n_j$ for some $i \neq j$. Please note that in this model $\mathbf{x}^{(i)}$ represents the state of the i-th producer and $x_j^{(i)}$ denotes the quality assessment of his/her j-th top product. Moreover, in many real life applications, it is necessary to assume that $x_j^{(i)} \in \mathbb{I} = (-\infty,\infty)$, cf. [3] for discussion.

Usually, aggregation operators are most often used to summarize informetric data sets. However, also machine learning techniques may be applied for this very purpose. For example, in [13] some algorithms were applied on several indicators in order to obtain an automatic categorization of universities. Similarly, in [4] a data set on the 500 best world universities was divided into groups according to their various bibliometric performance indicators. Moreover, Costas, van Leeuwen, and Bordons in order to split a group of scientists into 3 clusters (top, medium, low class ones) used e.g. the h-index [9], number of publications, number of highly cited papers, median impact factor, etc, see [5].

Investigation carried out in this paper focuses on clustering techniques. In our previous work [3] we discussed problems and challenges one may encounter while performing clustering of informetric data sets. We also proposed modifications of the well-known metrics, so they can be calculated over vectors of nonconforming lengths. The obtained measures were then applied in a hierarchical clustering method. What is more, the notion of such measures allows to adapt the k-means algorithm for clustering task. In this paper we generalize the results obtained in [2]. Moreover, we focus on centroids of the clusters derived, which can be conceived as an aggregated representation of the data set.

The structure of this contribution is as follows. In the next section the definition of a metric and dissimilarity measure for vectors of nonconforming lengths is recalled. In Sec. 3 the notion of the c-means algorithm is extended so that it can be computed over PAP data sets. Next, in Sec. 4, the performance of the obtained method is investigated. Finally, Sec. 5 concludes the paper and indicates future research directions.

2 METRICS

For any $n \in \mathbb{N}$, let $\mathcal{S}_n$ denote the set of non-increasingly ordered real vectors of length n, i.e. $\mathcal{S}_n = \{(x_1,\dots,x_n) \in \mathbb{R}^n, x_1 \geq \dots \geq x_n\}$. Moreover, let $\mathcal{S}_{\leq n}$ be a set of non-increasingly ordered vectors of length at most n, that is $\mathcal{S}_{\le n} = \bigcup_{i=1}^{n} \mathcal{S}_i$. Assume that we are given l producers and $k = \max\{n_i : i = 1, 2, \ldots, l\}$. Obviously, such k is finite and well defined for each set of producers. Moreover, let $\tilde{\mathbf{x}}$ denote the vector of length k and equivalent to $\mathbf{x}$ padded with 0's, i.e. $\tilde{\mathbf{x}} = (x_1, x_2, \ldots, x_n, 0, \ldots, 0) \in \mathcal{S}_k$.

Let us now recall the definition of a class of metrics over $\mathcal{S}_{\le k}$ (see [3] for more details and a proof). Please keep in mind, that *metric* is a function $d(\mathbf{x}, \mathbf{y})$ such that $(\forall \mathbf{x}, \mathbf{y})$ (a) $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ and (b) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ and fulfills triangle inequality, i.e. $d(\mathbf{x}, \mathbf{y}) \le d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$. Moreover, in case when conditions (b) and (c) hold and $d(\mathbf{x}, \mathbf{y}) = 0 \Leftarrow \mathbf{x} = \mathbf{y}$, d is a *pseudometric.*

Theorem 1. *Let $d_M : \mathcal{S}_{\le k} \times \mathcal{S}_{\le k} \to [0, \infty)$ be such that $d_M(\mathbf{x}, \mathbf{y}) = \mu(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \nu(\mathbf{x}, \mathbf{y})$, where μ is a metric on $\mathbb{R}^k$ and ν is a pseudo-metric on $\mathcal{S}_{\le k}$. Then d_M is a metric on $\mathcal{S}_{\le k}$ if and only if for all $\mathbf{x}, \mathbf{y}$ such that $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ it holds $\nu(\mathbf{x}, \mathbf{y}) = 0 \Longrightarrow n_x = n_y$.*

Because of computational reasons, in some clustering tasks it is more convenient to consider dissimilarity measure instead of metric, i.e. a function $d(\mathbf{x}, \mathbf{y})$ such that $(\forall \mathbf{x}, \mathbf{y})$ (a) $d(\mathbf{x}, \mathbf{y}) \ge 0$, (b) $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ and (c) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$. Of course, each metric is a dissimilarity measure.

3 CLUSTERING

Cluster analysis is a machine learning technique which allows to partition the general population of objects into distinguishable – according to some criteria – groups (clusters). In other words, in the most desired partitioning scheme, the entities within each group are as similar as possible and objects in distinct groups differ as much as possible from each other (see e.g. [8]).

One approach to cluster analysis is to divide the objects into c nonempty pairwise disjoint sets $\mathcal{C} = \{C_1, C_2, \ldots, C_c\}$, $\bigcup_{i=1}^{c} C_i = \mathcal{X}$, so as to minimize the within-cluster dissimilarity – sum of dissimilarities between points in the same cluster, i.e.:

$$\mathcal{C} = \underset{\text{partition } \mathcal{C} \text{ of } \mathcal{X}}{\arg\min} \sum_{j=1}^{c} \sum_{\mathbf{x} \in C_j} d_{L_2}^2 \left(\mathbf{x}, \boldsymbol{\mu}^{(j)} \right), \qquad (1)$$

where $d_{L_2}^2$ denotes the squared Euclidean distance (a dissimilarity measure) and $\boldsymbol{\mu}^{(j)}$ is the j-th cluster centroid.

On the other hand, in fuzzy clustering, every point has a degree of belonging to each cluster, rather than belonging entirely to just one of them, see e.g. [10]. In such a case, given a set of observations $\mathcal{X} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(l)}\}$, where each $\mathbf{x}^{(i)} \in \mathbb{R}^n$, we aim to determine a fuzzy pseudopartition – a family of fuzzy subsets of $\mathcal{X}$: $\mathcal{W} = \{W_1, \ldots, W_c\}$, where $w_{ij} = W_j(\mathbf{x}^{(i)})$, such that $\sum_{j=1}^{c} W_j(\mathbf{x}^{(i)}) = 1$, describes the degree of belonging of the i-th observation to the j-th cluster. Here, we aim to find the fuzzy pseudopartitioning which minimizes

$$\underset{\text{fuzzy partition } \mathcal{W} \text{ of } \mathcal{X}}{\arg\min} \sum_{i=1}^{l} \sum_{j=1}^{c} w_{ij}^{m} d_{L_2}^2(\mathbf{x}^{(i)}, \boldsymbol{\mu}^{(j)}), \qquad (2)$$

where the $m \in R$, $m \ge 1$ is a fuzzifier, cf [1].

Fuzzy c-means algorithm. Clustering tasks can be solved using various heuristics that differ significantly in their structure. Moreover, even when an algorithm converges, the obtained minimum may only be a local minimum. Also the initial choice of weights can have a great impact on the results. Investigation carried out here concerns the fuzzy c-means algorithm (cf. [1, 10]), which may be viewed as a weighted generalization of the k-means procedure.

Let us now recall basic steps of the c-means method.

1. Set a number of clusters and randomly assign for each observation the membership degrees.
2. Until the convergence condition is met, i.e. the coefficients' change between two iterations is not greater than a given sensitivity threshold ε, repeat:
   (a) Compute the centroid for each cluster $\boldsymbol{\mu}^{(j)}$ with respect to the weighted distance.
   (b) For each point $\mathbf{x}^{(i)}$, compute its coefficients (weights, membership) w_{ij} of being in the clusters $j = 1, \ldots, c$ as $w_{ij} = \left(\sum_{u=1}^{c} \left(\frac{d_{L_2}^2(\mathbf{x}^{(i)} - \boldsymbol{\mu}^{(j)})}{d_{L_2}^2(\mathbf{x}^{(i)} - \boldsymbol{\mu}^{(u)})} \right)^{1/(m-1)} \right)^{-1}$.

Special attention should be paid when implementing the above algorithm. The main problems here may be caused by numerical errors, which can occur during the computation of weights. Please note that if $d_{L_2}^2(\mathbf{x}^{(i)} - \boldsymbol{\mu}^{(u)}) = 0$ for some $u = 1, \ldots, c$, according to the formula given above there is division by 0. In such a case, one may choose as a weight an arbitrary real number (keeping in mind that the weights must add up to 1, see e.g [10]).

3.1 DETERMINING THE WEIGHTED CENTROID

Let us focus on a squared version of an Euclidean-like metric $d_{D;pq}^2 : \mathcal{S}_{\le k} \times \mathcal{S}_{\le k}$, defined as $d_{D;pq}^2(\mathbf{x}, \mathbf{y}) = d_{L_2}^2(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + p|n_x^r - n_y^r|$, where d_{L_2} denotes the Euclidean metric in $\mathbb{R}^k$ and $p, q > 0$. Please note that this approach bases on the idea of padding input vectors with zeros. It is because in the standard arithmetic of real numbers we have $|a-0| = |0-a| = |a|$. Also note that by the fact that 0 is a distinguished value in the set of reals, the introduced metrics can be rewritten as: the distance between $\min\{n_x, n_y\}$ observations plus a norm of the remaining observations in the longer vector (which is the same as the distance to $\mathbf{0}$) plus penalty for the difference in vectors' lengths. As we see, for the purpose of the investigation carried out in this paper, we propose to use the penalty of the form $p|n_x^r - n_y^r|$. Please note that in the considered scenario 0 has a special meaning – it denotes an item with quality assessment of 0, e.g. a paper without citations. Because of that, we are interested in dissimilarity measures that can distinguish vectors, e.g., (1) and $(1, 0, 0, \ldots, 0)$ from each other. This can be done by including the mentioned penalty term: without it the dissimilarity between the exemplary vectors would be equal to 0.

Let $\mathbf{w} = (w_1, w_2, \ldots, w_l)$ denote the degree of membership of the observations to a given cluster. Our task is to find the vector $\boldsymbol{\mu}$ which minimizes

$$\arg\min_{\boldsymbol{\mu}\in\mathcal{S}} \sum_{i=1}^{l} w_i^m d_{D;pq}^2(\mathbf{x}^i, \boldsymbol{\mu}),$$

i.e. the vector which minimizes the objective function given by $F_{\mathbf{w}}(\boldsymbol{\mu}) = \sum_{i=1}^{l} w_i^m d_{D;pq}^2(\mathbf{x}^i, \boldsymbol{\mu})$.

It is easily seen that such a task is a generalization of the results presented in [2], where determining the $d_{D;pq}^2$-centroid was derived for the k-means procedure with $w_i = 1$ or $w_i = 0$ for all i.

Step 1. For $n = 1, 2, \ldots, k$ determine:

$$\boldsymbol{\mu}^{(n)} = \arg\min_{\boldsymbol{\mu}\in\mathcal{S}_n} F_{\mathbf{w}}(\boldsymbol{\mu}). \tag{3}$$

Step 2. Compute:

$$\boldsymbol{\mu} = \arg\min_{n=1,\ldots,k} F_{\mathbf{w}}(\boldsymbol{\mu}^{(n)}). \tag{4}$$

Let $n \in [k] := \{1, 2, \ldots, k\}$ be fixed and let $\mathcal{P} \subseteq 2^{[n]}$ denote the partition of a set $[n]$, such that for each $P, P' \in \mathcal{P}$ we have $P \cap P' = \emptyset$, $|P| > 0$, $\bigcup_{P\in\mathcal{P}} = [n]$ and $\{i, j\} \in P$ with $i \le j$ implies that $i+1, i+2, \ldots, j-1 \in P$. Moreover, let $P_{\{i\}}$ stand for such $P \in \mathcal{P}$ that fulfills $\{i\} \in P$. By $\mathcal{CP}[n]$ we will denote the whole class of such partitions.

Theorem 2. *For some $\mathcal{P} \in \mathcal{CP}[n]$ the vector $\mathbf{y} \in \mathbb{R}^n$ given by*

$$y_i = \frac{\sum_{f=1}^{l} \left(w_f^m \sum_{j\in P_i} \tilde{x}_j^{(f)} \right)}{|P_{\{i\}}| \sum_{f=1}^{l} w_f^m} \quad \text{for } i = 1, \ldots, n,$$

is a solution to Eq. (3) if $y_1 \ge y_2 \ge \cdots \ge y_n$ and for all $i \in [n]$ with $i \in (P_{\{i\}} \setminus \{\max P_{\{i\}}\})$ we have

$$\frac{i - \min P_{\{i\}} + 1}{|P_{\{i\}}|} \sum_{f=1}^{l} w_f^m \Big\{ \sum_{j\in P_{\{i\}}} \tilde{x}_j^{(f)} \Big\} - \sum_{f=1}^{l} w_f^m \Big\{ \sum_{j\in P_{\{i\}}, j\le i} \tilde{x}_j^{(f)} \Big\} > 0.$$

Proof. The task is to find $\min_{\mathbf{y}\in\mathbb{R}^n} F_{\mathbf{w}}(\mathbf{y})$ with respect to $n-1$ constraints of the form:

$$g_i(\mathbf{y}) \;=\; y_{i+1} - y_i \le 0 \quad \text{for } i = 1, \ldots, n-1.$$

By means of the Karush-Kuhn-Tucker (KKT) theorem (cf. [12]), we need to find $\mathbf{y}$ and $\lambda_1, \ldots, \lambda_{n-1}$ such that

$$\nabla F_{\mathbf{w}}(\mathbf{y}) + \sum_{i=1}^{n-1} \lambda_i \nabla g_i(\mathbf{y}) = 0,$$

with $\lambda_i g_i(\mathbf{y}) = 0$ and $\lambda_i \ge 0$ for $i \in [n-1]$. Note that for $h \in [n]$ we have:

$$\frac{\partial F_{\mathbf{w}}}{\partial y_h}(\mathbf{y}) = 2 \sum_{i=1}^{l} w_i^m (y_h - \tilde{x}_h^i).$$

For brevity of notation, let us assume that $\lambda_0 := 0$ and $\lambda_n := 0$. Thus, our task reduces to solving the following system of linear equations:

$$\begin{cases} 0 = 2\sum_{i=1}^{l} w_i^m (y_1 - \tilde{x}_1^i) & +\lambda_0 & -\lambda_1 \\ 0 = 2\sum_{i=1}^{l} w_i^m (y_2 - \tilde{x}_2^i) & +\lambda_1 & -\lambda_2 \\ \quad\vdots & & \\ 0 = 2\sum_{i=1}^{l} w_i^m (y_{n-1} - \tilde{x}_{n-1}^i) & +\lambda_{n-2} & -\lambda_{n-1} \\ 0 = 2\sum_{i=1}^{l} w_i^m (y_n - \tilde{x}_n^i) & +\lambda_{n-1} & -\lambda_n \\ 0 = \lambda_1 (y_2 - y_1) & & \\ \quad\vdots & & \\ 0 = \lambda_{n-1} (y_n - y_{n-1}) & & \end{cases}$$

under constraints $\lambda_1 \ge 0, \ldots, \lambda_{n-1} \ge 0$ and $y_1 \ge y_2 \ge \cdots \ge y_n$.

Thus, let us consider a solution (not necessarily feasible) that fulfills $\boldsymbol{\lambda} \ge 0$. First of all, let us take u such that $\lambda_{u-1} = \lambda_u = 0$. It immediately implies that:

$$y_u = \frac{\sum_{f=1}^{l} w_f^m \tilde{x}_u^{(f)}}{\sum_{f=1}^{l} w_f^m}.$$

On the other hand, for each u and $p \ge 2$ such that $\lambda_{u-1} = 0$, $\lambda_u > 0$, $\lambda_{u+1} > 0$, $\ldots$, $\lambda_{u+p-2} > 0$, $\lambda_{u+p-1} = 0$ we get that $y_u = \cdots = y_{u+p-1}$. More specifically, we have:

$$y_i = \frac{\sum_{f=1}^{l} w_f^m \sum_{j=u}^{u+p-1} \tilde{x}_j^{(f)}}{p \sum_{f=1}^{l} w_f^m} \quad \text{for } i = u, \ldots, u+p-1,$$

In such a case, we have that for $i = u, \dots, u+p-2$:

$$\lambda_i = 2\frac{i-u+1}{p}\sum_{f=1}^{l} w_f^m \Big(\sum_{j=u}^{u+p-1} \tilde{x}_j^{(f)}\Big) - 2\sum_{f=1}^{l} w_f^m \Big(\sum_{j=u}^{i} \tilde{x}_j^{(f)}\Big) > 0,$$

is a solution to Eq. (3). □

See Algorithm 1 for the procedure that computes the solution to the Eq. (3). Please note that this approach is a simple generalization of the algorithm included in [2], where the proof of its correctness was included. The main modification concerns the form of input data. In the scenario considered in this paper, the algorithm is applied to a list of vectors of the form $\mathbf{x}^{(j)} = w_j^m \mathbf{x}^{(j)} / \sum_{i=1}^{l} w_j^m$. It is clear to see that when weights are either 0 or 1, both procedures return the same result.

Data: A set of l vectors $\mathcal{X} \subset \mathcal{S}$ and $n \in \mathbb{N}$.
Result: $\boldsymbol{\mu}^{(n)} = \arg\min_{\boldsymbol{\mu}\in\mathcal{S}_n} F_{\mathbf{w}}(\boldsymbol{\mu})$.
Let $\tilde{\mathbf{x}}$ be such that $\bar{x}_i = \sum_{j=1}^{l} w_j^m \tilde{x}_i^{(j)} / \sum_{i=1}^{l} w_i^m$, for all $i \in [n]$;
Let $\mathcal{P} = \emptyset$;
Let $\mathbf{y} \in \mathbb{R}^n$;
for $k = 1, 2, \dots, n$ **do**
  $y_k = \bar{x}_k$;
  Let $\mathcal{P} := \mathcal{P} \cup \big\{\{k\}\big\}$; *(we have $\mathcal{P} \in \mathcal{CP}([k])$)*
  while $|\mathcal{P}| > 1$ *and* $y_{\min P^{(|\mathcal{P}|)}} > y_{\max P^{(|\mathcal{P}|-1)}}$ **do**
    $\mathcal{P} := \Big(\big(\mathcal{P} \setminus \{P^{(|\mathcal{P}|)}\}\big) \setminus \{P^{(|\mathcal{P}|-1)}\}\Big) \cup \{P^{(|\mathcal{P}|-1)} \cup P^{(|\mathcal{P}|)}\}$; *(merge $P^{(|\mathcal{P}|-1)}, P^{(|\mathcal{P}|)}$)*
    for $i \in P^{(|\mathcal{P}|)}$ **do**
      Set $y_i := \frac{1}{|P^{(|\mathcal{P}|)}|} \sum_{j \in P^{(|\mathcal{P}|)}} \bar{x}_j$;
    end
  end
end
return y;

Algorithm 1: An algorithm to determine the solution to the Eq. (3)

Remark 1. *Even though the $d^2_{D;11}$-centroid can be viewed as an aggregated representation of a set of vectors, in general the procedure given by Theorem 2 is not a $\trianglelefteq$-monotonic fusion function, where $\trianglelefteq$ is the partial ordering described in [7]. For example, let us consider $\mathcal{X} = \{\mathbf{x}^{(1)} = (10,2,1,0,0), \mathbf{x}^{(2)} = (-11), \mathbf{x}^{(3)} = (-5,-6,-10)\}$ and $\mathcal{Y} = \{\mathbf{y}^{(1)} = (10,2,1,0,0), \mathbf{y}^{(2)} = (10,-100), \mathbf{y}^{(3)} = (-5,-6,-10)\}$. It is clear to see that for each $i = 1,2,3$ we have $\mathbf{x}^{(i)} \trianglelefteq \mathbf{y}^{(i)}$. However, for the corresponding centroids we have $(-1.67,-1.67,-3) \ntrianglelefteq (5,-34.67)$. On the other hand, if all the input elements are non-negative, then $\trianglelefteq$-monotonicity always holds.*

4 EMPIRICAL ANALYSIS

Let us now consider a data set consisting of citations received by 5000 scientists[1]. The data were gathered from Elsevier's Scopus (see [6] for details).

Table 1 contains basic sample statistics of the vectors. Please note that 78% of them are only of length 1 and among them, 32% are equal to 0.

Table 1: Basic summary statistics of vectors' lengths (n), maximal value (max) and sum of all elements (sum) (Scopus data set).

	Min.	Median	Mean	Max.
n	1	1	1.62	98
max	0	3	8.96	791
sum	0	3	12.63	1211

The fuzzy c-means and k-means algorithms were applied in order to determine 6 clusters (groups). The closest crisp clustering based on weights from the fuzzy c-means algorithm, was obtained by assigning each vector to the cluster with the maximal weight. The number of common vectors in clusters obtained via the k-means and fuzzy c-means algorithm are presented in Table 2.

Table 2: Number of common vectors in clusters obtained via the k-means (columns) and the c-means (rows) algorithm.

Cluster no.	1	2	3	4	5	6
1	3472	0	0	0	0	0
2	5	1052	6	0	0	0
3	0	2	308	6	0	0
4	0	0	10	100	0	0
5	0	0	0	6	27	0
6	0	0	0	0	0	6

Figure 2 presents the step functions of citation vectors in each cluster. Corresponding centroids are marked by black o and red × for the k-means and the fuzzy c-means procedure, respectively. The agreement between these two partitioning schemes, being equal to 99%, was calculated via the Rand Index, i.e. $A/(A+D)$, where A denotes the number of all pairs of data points assigned by both partitions into the same cluster or into different clusters (both partitionings agree for all pairs A) and D denotes the number of all pairs assigned differently by both partitions (the partitions disagree for all pairs D), cf. [11]. Figure 1 presents the distribution of maximal weights per each vector. Please note that there are 188 vectors for which any weight is not greater than 0.5.

[1]The data set is available at http://cena.rexamine.com/research/.

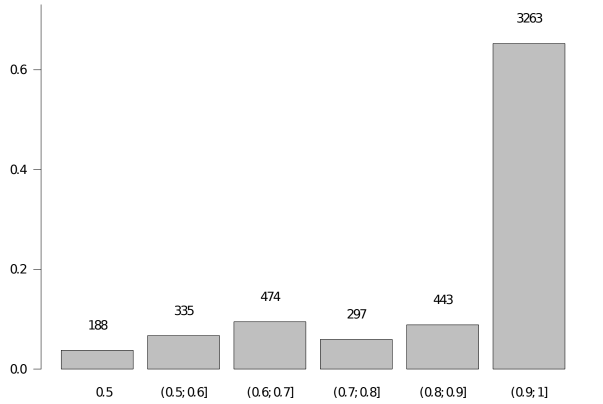


Figure 1: The distribution of cluster membership weights.

Table 3 presents the weighted sum of dissimilarities between bibliometric indices for the Scopus data set, i.e. $\sum_{j=1}^{c}\sum_{k,i=1}^{l} w_{ij}w_{kj}(x_i - x_k)^2$. We analyzed the weights determined by the fuzzy c-means algorithm, the uniformly distributed membership degrees to all clusters and weights of the form $w_{ij} = 1$ for $\mathbf{x}^{(i)}$ if the k-means algorithm assigns i-th vector to j-th cluster and $w_{ij} = 0$ otherwise. Here, we consider the Hirsch index (H), the G-index (G), length (n), maximum (max) and the arithmetic mean (mean). Please note that the weights returned by the fuzzy c-means algorithm minimizes such a sum for all the considered indexes.

Table 3: Weighted sum of dissimilarities between the Hirsch index (H), the G-index (G), length (n), Max (max) and arithmetic mean (mean) calculated for the fuzzy c-means output, uniform membership and k-means partitioning.

	c-means	uniform membership	k-means
G	951.01	2404321.83	1404.00
H	420.33	1763367.67	564.00
n	1347.73	2337926.83	1404.00
max	15865.30	27334939.66	20394.00
mean	10148.87	20259504.16	20207.30

Moreover, Table 4 presents the Hirsch index (H), G-index (G), length (n), Max operator (max) and the arithmetic mean (mean) computed for clusters centroids determined by the fuzzy c-means algorithm. We may see that the first centroid is characterized by a low H- and G-index and also a small number of publications. On the other hand, centroids corresponding to clusters 2, 3 and 4, represent researchers with a rather small number of publications (between 2 and 4) and a low H- and G-index, however, the number of citation given to their most cited paper (max) increase (12.55, 32.74, 68.77). Finally, researchers from clusters 5 and 6, represented by corresponding centroids are characterized with a high number of citations (175.51 and 907.50), and larger number of publications (10 and 24).

Table 4: The Hirsch index (H), G-index (G), length (n), maximum (max) and the arithmetic mean (mean) computed for cluster centroids determined by the fuzzy c-means algorithm.

Cl.no.	1	2	3	4	5	6
G	1.00	2.00	3.00	4.00	10.00	24.00
H	1.00	1.00	2.00	3.00	4.00	12.00
n	1.00	2.00	3.00	4.00	10.00	24.00
max	1.64	12.55	32.74	68.77	127.21	398.21
mean	1.64	7.00	12.67	20.05	17.55	37.81
sum	1.64	13.99	38.00	80.19	175.51	907.50

5 CONCLUSIONS

In this paper applications of the fuzzy c-means clustering algorithm to sets of vectors of possibly nonconforming lengths were investigated. First of all, a generalization of the procedure to compute the centroids of such sets, with respect to some dissimilarity measure tailored for vectors of unequal lengths, was provided. Moreover, the presented approach was verified by an empirical analysis on a bibliometric data set. Special attention was paid to the investigation of the clusters' centroids as an aggregated representation of a considered data set. By means of various bibliometric indexes, we evaluated the degree in which they reflect the structure of the whole data set.

Acknowledgements

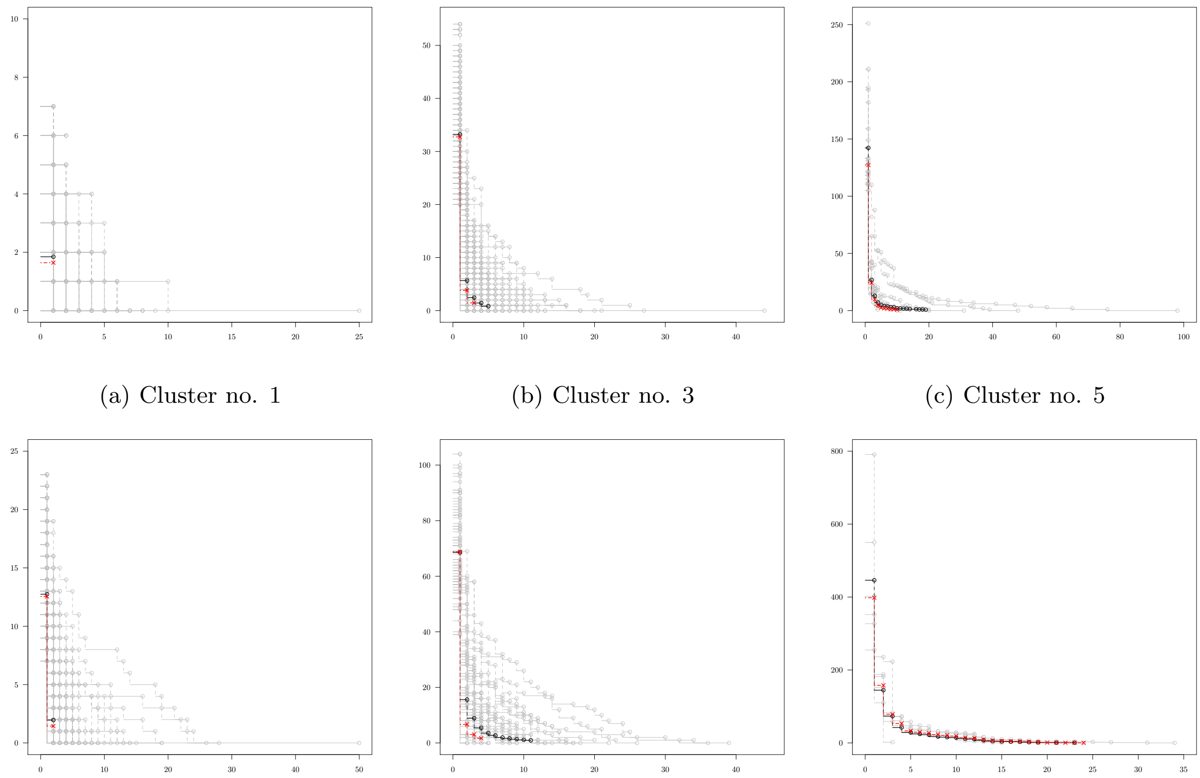
The contribution of A. Cena and M. Gagolewski was partially supported by the National Science Center, Poland, research project 2014/13/D/HS4/01700. Moreover, A. Cena would like to acknowledge the support by the European Union from resources of the European Social Fund, Project PO KL "Information technologies: Research and their interdisciplinary applications", agreement UDA-POKL.04.01.01-00-051/10-00 via the Interdisciplinary PhD Studies Program.

REFERENCES

[1] J. C. Bezdek, R. Ehrlich, and W. Full (1984). FCM: The fuzzy c-means clustering algorithm. *Computer & Geosciences* 10(2–3), pp. 191–203.

[2] A. Cena and M. Gagolewski (2015). A k-means-like algorithm for informetric data clustering. In: Proc. IFSA-EUSFLAT 2015 (in press).

[3] A. Cena, M. Gagolewski, and R. Mesiar (2015). Problems and challenges of information resources



(a) Cluster no. 1 (b) Cluster no. 3 (c) Cluster no. 5

(d) Cluster no. 2 (e) Cluster no. 4 (f) Cluster no. 6

Figure 2: Step functions depicting vectors in each cluster (Scopus) and their centroids (bold black ○ – k-means, bold red × – fuzzy c-means).

producers' clustering. *Journal of Informetrics* 9(2), pp. 273–284.

[4] Y. Cheng and N. C. Liu (2006). A first approach to the classification of the top 500 world universities by their disciplinary characteristics using scientometrics. *Scientometrics* 68(1), pp. 135–150.

[5] R. Costas, T. van Leeuwen, and M. Bordons (2001). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology* 61, pp. 1564–1581.

[6] M. Gagolewski (2011). Bibliometric impact assessment with R and the CITAN package. *Journal of Informetrics* 5(4), pp. 678–692.

[7] M. Gagolewski and P. Grzegorzewski (2011). Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning* 52(9), pp. 1312–1324.

[8] T. Hastie, R. Tibshirani, and J. Friedman (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag.

[9] J. E. Hirsch (2005). An index to quantify individual's scientific research output. *Proceedings of the National Academy of Sciences* 102(46), pp. 16569–16572.

[10] G. J. Klir and B. Yuan (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* Prentice Hall PTR.

[11] H. Lawrence and A. Phipps (1985). Comparing partitions. *Journal of Classification* 2, pp. 193–218.

[12] J. Nocedal and S. Wright (2006). *Numerical Optimization.* Springer-Verlag, New York.

[13] J. L. Ortega, E. López-Romero, and I. Fernández (2011). Multivariate approach to classify research institutes according to their outputs: The case of the CSIC's institutes. *Journal of Informetrics* 5, pp. 323–332.

MEDIAN PRESERVING AGGREGATION FUNCTIONS

Miguel Couceiro
LORIA
(CNRS - Inria G. E. - U. Lorraine)
Bat. B, Campus Sci. - B.P. 239
F-54506 Vandoeuvre-lès-Nancy
miguel.couceiro[at]inria.fr

Jean-Luc Marichal and Bruno Teheux
Mathematics Research Unit,
University of Luxembourg
6, rue Coudenhove-Kalergi
L-1359 Luxembourg
{jean-luc.marichal,bruno.teheux}[at]uni.lu

Summary

A median algebra is a ternary algebra that satisfies every equation satisfied by the median terms of distributive lattices. We present a characterization theorem for aggregation functions over conservative median algebras. In doing so, we give a characterization of conservative median algebras by means of forbidden substructures and by providing their representation as chains.

Keywords: Median algebras, Aggregation Functions, Distributive lattices.

1 INTRODUCTION AND PRELIMINARIES

Informally, an aggregation function $f\colon \mathbf{A}^n \to \mathbf{B}$ may be thought of as a mapping that preserves the structure of $\mathbf{A}$ into $\mathbf{B}$. It is common to consider that $\mathbf{B}$ is equal to $\mathbf{A}$ and is equipped with a partial order so that aggregation functions are thought of as order-preserving maps [8].

If $\mathbf{L} = \langle L, \wedge, \vee \rangle$ is a distributive lattice then the ternary term operation defined on $\mathbf{L}$ by

$$\mathbf{m}(x,y,z) = (x \vee y) \wedge (x \vee z) \wedge (y \vee z) \qquad (1.1)$$

is symmetric and self dual, and is called the *median term* over $\mathbf{L}$. If $\mathbf{L}$ is a total order, then $\mathbf{m}(a,b,c)$ is the element among a, b and c that is between the two other ones if a, b, c are mutually distinct, and is the majority element otherwise.

Median algebras are ternary algebras that were introduced in order to abstract this notion of betweenness. Formally, a median algebra is an algebra $\mathbf{A} = \langle A, \mathbf{m} \rangle$ with a single ternary operation $\mathbf{m}$ that satisfies the equations

$$\mathbf{m}(x,x,y) = x,$$
$$\mathbf{m}(x,y,z) = \mathbf{m}(y,x,z) = \mathbf{m}(y,z,x),$$
$$\mathbf{m}(\mathbf{m}(x,y,z),t,u) = \mathbf{m}(x,\mathbf{m}(y,t,u),\mathbf{m}(z,t,u)),$$

and that is called a *median operation*. In particular, every median algebra satisfies the equation

$$\mathbf{m}(x,y,\mathbf{m}(x,y,z)) \;=\; \mathbf{m}(x,y,z). \qquad (1.2)$$

Examples of median operations are given by median term operations over distributive lattices. If $\mathbf{L}$ is a distributive lattice and if $\mathbf{m}_{\mathbf{L}}$ is the operation defined on $\mathbf{L}$ by (1.1) then the algebra $\langle L, \mathbf{m}_{\mathbf{L}} \rangle$ is called the *median algebra associated with* $\mathbf{L}$. If $\mathbf{A}$ is a median algebra, the median operation is extended to $\mathbf{A}^n$ pointwise.

A median algebra $\mathbf{A} = \langle A, \mathbf{m} \rangle$ is said to be *conservative* if

$$\mathbf{m}(x,y,z) \in \{x,y,z\},$$

for every $x, y, z \in A$. It is not difficult to observe that a median algebra is conservative if and only if each of its subsets is a median subalgebra. Moreover, the median term associated with a total order is a conservative median operation. This fact was observed in §11 of [12], which presents the four element Boolean algebra as a counter-example.

The results of this paper, which were previously exposed in [5], are twofold. First, we present a description of conservative median algebras in terms of forbidden substructures (in complete analogy with Birkhoff's characterization of distributive lattices with M_5 and N_5 as forbidden substructures and Kuratowski's characterization of planar graphs in terms of forbidden minors), and that leads to a representation of conservative median algebras (with at least five elements) as chains. In fact, the only conservative median algebra that is not representable as a chain is the four element Boolean algebra.

ISBN: 978–83–8012–519–3

Second, we characterize functions $f : \mathbf{B} \to \mathbf{C}$ that satisfy the equation

$$f(\mathbf{m}(x, y, z)) \;=\; \mathbf{m}(f(x), f(y), f(z)), \qquad (1.3)$$

where $\mathbf{B}$ and $\mathbf{C}$ are finite products of (non necessarily finite) chains, as superposition of compositions of monotone maps with projection maps (Theorem 4.5). Particularized to aggregation functions $f : \mathbf{A}^n \to \mathbf{A}$, where $\mathbf{A}$ is a chain, we obtain an ARROW-like theorem: *f satisfies equation* (1.3) *if and only if it is dictatorial and monotone* (Corollary 4.6).

Throughout the paper we employ the following notation. For each positive integer n, we set $[n] = \{1, \ldots, n\}$. Algebras are denoted by bold roman capital letters $\mathbf{A}, \mathbf{B}, \mathbf{X}, \mathbf{Y} \ldots$ and their universes by italic roman capital letters $A, B, X, Y \ldots$. To simplify our presentation, we will keep the introduction of background to a minimum, and we will assume that the reader is familiar with the theory of lattices and ordered sets. We refer the reader to [7, 9] for further background. Proofs of the results presented in the fourth section are omitted because they rely on arguments involving a categorical duality that are beyond the scope of this paper. The missing proofs and details can be found in [6].

2 MEDIAN ALGEBRAS, MEDIAN SEMILATTICES AND MEDIAN GRAPHS

Median algebras have been investigated by several authors (see [4, 10] for early references on median algebras and see [2, 11] for some surveys) who illustrated the deep interactions between median algebras, order theory and graph theory.

For instance, take an element a of a median algebra $\mathbf{A}$ and consider the relation $\leq_a$ defined on A by

$$x \leq_a y \quad \iff \quad \mathbf{m}(a, x, y) = x.$$

Endowed with this relation, $\mathbf{A}$ is a $\wedge$-semilattice order with bottom element a [13]: the associated operation $\wedge$ is defined by $x \wedge y = \mathbf{m}(a, x, y)$.

Semilattices constructed in this way are called *median semilattices*, and they coincide exactly with semilattices in which every principal ideal is a distributive lattice and in which any three elements have a join whenever each pair of them is bounded above. The operation $\mathbf{m}$ on $\mathbf{A}$ can be recovered from the median semilattice order $\leq_a$ using identity (1.1) where $\wedge$ and $\vee$ are defined with respect to $\leq_a$. Semilattices associated with conservative median algebras are called *conservative median semilattices*.

Note that if a median algebra $\mathbf{A}$ contains two elements 0 and 1 such that $\mathbf{m}(0, x, 1) = x$ for every $x \in A$, then $(A, \leq_0)$ is a distributive lattice order bounded by 0 and 1, and where $x \wedge y$ and $x \vee y$ are given by $\mathbf{m}(x, y, 0)$ and $\mathbf{m}(x, y, 1)$, respectively. It is noteworthy that equations satisfied by median algebras of the form $\langle L, \mathbf{m}_\mathbf{L} \rangle$ are exactly those satisfied by median algebras. In particular, every median algebra satisfies the equation

$$\mathbf{m}(x, y, z) = \mathbf{m}\big(\mathbf{m}\big(\mathbf{m}(x, y, z), x, t\big), \mathbf{m}\big(\mathbf{m}(x, y, z), z, t\big), \mathbf{m}\big(\mathbf{m}(x, y, z), y, t\big)\big). \qquad (2.1)$$

Moreover, covering graphs (*i.e.*, undirected HASSE diagram) of median semilattices have been investigated and are, in a sense, equivalent to median graphs. Recall that a median graph is a (non necessarily finite) connected graph in which for any three vertices u, v, w there is exactly one vertex x that lies on a shortest path between u and v, on a shortest path between u and w and on a shortest path between v and w. In other words, x (the *median* of u, v and w) is the only vertex such that

$$d(u, v) = d(u, x) + d(x, v),$$
$$d(u, w) = d(u, x) + d(x, w),$$
$$d(v, w) = d(v, x) + d(x, w).$$

Every median semilattice whose intervals are finite has a median covering graph [1] and conversely, every median graph is the covering graph of a median semilattice [1, 13]. This connection is deeper: median semilattices can be characterized among the ordered sets whose bounded chains are finite and in which any two elements are bounded below as the ones whose covering graph is median [3]. For further background see, e.g., [2].

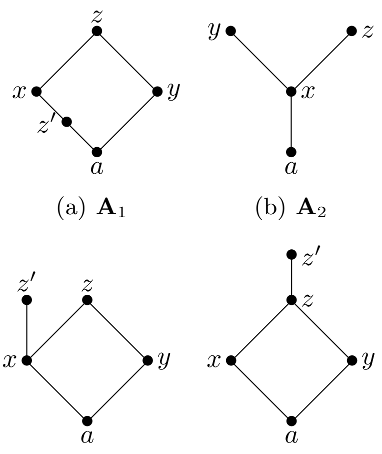
3 CHARACTERIZATIONS OF CONSERVATIVE MEDIAN ALGEBRAS

Let $\mathbf{C}_0 = \langle C_0, \leq_0, c_0 \rangle$ and $\mathbf{C}_1 = \langle C_1, \leq_1, c_1 \rangle$ be chains with bottom elements c_0 and c_1, respectively. The $\perp$-*coalesced sum* $\mathbf{C}_0 \perp \mathbf{C}_1$ of $\mathbf{C}_0$ and $\mathbf{C}_1$ is the poset obtained by amalgamating c_0 and c_1 in the disjoint union of C_0 and C_1. Formally,

$$\mathbf{C}_0 \perp \mathbf{C}_1 = \langle C_0 \sqcup C_1 \;/\!\equiv,\ \leq\ \rangle,$$

where $\sqcup$ is the disjoint union, where $\equiv$ is the equivalence generated by $\{(c_0, c_1)\}$ and where $\leq$ is defined by

$$x/\!\equiv \;\leq\; y/\!\equiv \iff (x \in \{c_0, c_1\} \text{ or } x \leq_0 y \text{ or } x \leq_1 y).$$



(a) $\mathbf{A}_1$ (b) $\mathbf{A}_2$

(c) $\mathbf{A}_3$ (d) $\mathbf{A}_4$

Figure 1: Examples of $\wedge$-semilattices that are not conservative.

Proposition 3.1. *The partially ordered sets $\mathbf{A}_1, \ldots, \mathbf{A}_4$ depicted in Fig. 1 are not conservative median semilattices.*

Proof. The poset $\mathbf{A}_1$ is a bounded lattice (also denoted by N_5 in the literature on lattice theory, e.g., in [7, 9]) that is not distributive. In $\mathbf{A}_2$ the center is equal to the median of the other three elements. The poset $\mathbf{A}_3$ contains a copy of $\mathbf{A}_2$, and $\mathbf{A}_4$ is a distributive lattice that contains a copy of the dual of $\mathbf{A}_2$ and thus it is not conservative as a median algebra. □

The following Theorem provides descriptions of conservative semilattices with at least five elements, both in terms of forbidden substructures and in the form of representations by chains. Note that any semilattice with at most four elements is conservative, but the poset depicted in Fig. 1(b).

Theorem 3.2. *Let $\mathbf{A}$ be a median algebra with $|A| \geq 5$. The following conditions are equivalent.*

(1) $\mathbf{A}$ is conservative.

(2) For every $a \in A$ the ordered set $\langle A, \leq_a\rangle$ does not contain a copy of the poset depicted in Fig. 1(b).

(3) There is an $a \in A$ and lower bounded chains $\mathbf{C}_0$ and $\mathbf{C}_1$ such that $\langle A, \leq_a\rangle$ is isomorphic to $\mathbf{C}_0 \perp \mathbf{C}_1$.

(4) For every $a \in A$, there are lower bounded chains $\mathbf{C}_0$ and $\mathbf{C}_1$ such that $\langle A, \leq_a\rangle$ is isomorphic to $\mathbf{C}_0 \perp \mathbf{C}_1$.

Proof. (1) $\implies$ (2): Follows from Proposition 3.1.

(2) $\implies$ (1): Suppose that $\mathbf{A}$ is not conservative, that is, there are $a, b, c, d \in A$ such that $d := \mathbf{m}(a, b, c) \notin \{a, b, c\}$. Clearly, a, b and c must be pairwise distinct. By (1.2), a and b are $\leq_c$-incomparable, and $d <_c a$ and $d <_c b$. Moreover, $c <_c d$ and thus $\langle\{a, b, c, d\}, \leq_c\rangle$ is a copy of A_2 in $\langle A, \leq_c\rangle$.

(1) $\implies$ (4): Let $a \in A$. First, suppose that for every $x, y \in A \setminus \{a\}$ we have $\mathbf{m}(x, y, a) \neq a$. Since $\mathbf{A}$ is conservative, for every $x, y \in A$, either $x \leq_a y$ or $y \leq_a x$. Thus $\leq_a$ is a chain with bottom element a, and we can choose $\mathbf{C}_1 = \langle A, \leq_a, a\rangle$ and $\mathbf{C}_2 = \langle\{a\}, \leq_a, a\rangle$.

Suppose now that there are $x, y \in A \setminus \{a\}$ such that $\mathbf{m}(x, y, a) = a$, that is, $x \wedge y = a$. We show that

$$z \neq a \implies \big(\mathbf{m}(x, z, a) \neq a \text{ or } \mathbf{m}(y, z, a) \neq a\big), \quad z \in A. \tag{3.1}$$

For the sake of a contradiction, suppose that $\mathbf{m}(x, z, a) = a$ and $\mathbf{m}(y, z, a) = a$ for some $z \neq a$. By equation (2.1), we have

$$\mathbf{m}(x, y, z) = \mathbf{m}\big(\mathbf{m}\big(\mathbf{m}(x, y, z), x, a\big), \mathbf{m}\big(\mathbf{m}(x, y, z), z, a\big), \mathbf{m}\big(\mathbf{m}(x, y, z), y, a\big)\big). \tag{3.2}$$

Assume that $\mathbf{m}(x, y, z) = x$. Then (3.2) is equivalent to

$$x = \mathbf{m}(x, \mathbf{m}(x, z, a), \mathbf{m}(x, y, a)) = a,$$

which yields the desired contradiction. By symmetry, we derive the same contradiction in the case $\mathbf{m}(x, y, z) \in \{y, z\}$.

We now prove that

$$z \neq a \implies \big(\mathbf{m}(x, z, a) = a \text{ or } \mathbf{m}(y, z, a) = a\big), \quad z \in A. \tag{3.3}$$

For the sake of a contradiction, suppose that $\mathbf{m}(x, z, a) \neq a$ and $\mathbf{m}(y, z, a) \neq a$ for some $z \neq a$. Since $\mathbf{m}(x, y, a) = a$ we have that $z \notin \{x, y\}$.

If $\mathbf{m}(x, z, a) = z$ and $\mathbf{m}(y, z, a) = y$, then $y \leq_a z \leq_a x$ which contradicts $x \wedge y = a$. Similarly, if $\mathbf{m}(x, z, a) = z$ and $\mathbf{m}(y, z, a) = z$, then $z \leq_a x$ and $z \leq_a y$ which also contradicts $x \wedge y = a$. The case $\mathbf{m}(x, z, a) = x$ and $\mathbf{m}(y, z, a) = z$ leads to similar contradictions.

Hence $\mathbf{m}(x, z, a) = x$ and $\mathbf{m}(y, z, a) = y$, and the $\leq_a$-median semilattice arising from the subalgebra $\mathbf{B} = \{a, x, y, z\}$ of $\mathbf{A}$ is the median semilattice associated with the four element Boolean algebra. Let $z' \in A \setminus \{a, x, y, z\}$. By (3.1) and symmetry we may assume that $\mathbf{m}(x, z', a) \in \{x, z'\}$. First, suppose that $\mathbf{m}(x, z', a) = z'$. Then $\langle\{a, x, y, z, z'\}, \leq_a\rangle$ is N_5 (Fig. 1(a)) which is not a median semilattice. Suppose then that $\mathbf{m}(x, z', a) = x$. In this case, the restriction of $\leq_a$ to $\{a, x, y, z, z'\}$ is depicted in Fig. 1(c) or 1(d), which contradicts Proposition 3.1, and the proof of (3.3) is thus complete.

Now, let $C_0 = \{z \in A \mid (x, z, a) \neq a\}$, $C_1 = \{z \in A \mid (y, z, a) \neq a\}$ and let $\mathbf{C}_0 = \langle C_0, \leq_a, a\rangle$ and $\mathbf{C}_1 = \langle C_1, \leq_a, a\rangle$. It follows from (3.1) and (3.3) that $\langle \mathbf{A}, \leq_a \rangle$ is isomorphic to $\mathbf{C}_0 \perp \mathbf{C}_1$.

(4) $\implies$ (3): Trivial.

(3) $\implies$ (1): Let $x, y, z \in \mathbf{C}_0 \perp \mathbf{C}_1$. If $x, y, z \in C_i$ for some $i \in \{0, 1\}$ then $\mathbf{m}(x, y, z) \in \{x, y, z\}$. Otherwise, if $x, y \in C_i$ and $z \notin C_i$, then $\mathbf{m}(x, y, z) \in \{x, y\}$. □

The equivalence between (3) and (1) in Proposition 3.2 gives rise to the following representation of conservative median algebras.

Theorem 3.3. *Let $\mathbf{A}$ be a median algebra with $|A| \geq 5$. Then $\mathbf{A}$ is conservative if and only if there is a totally ordered set $\mathbf{C}$ such that $\mathbf{A}$ is isomorphic to $\langle \mathbf{C}, \mathbf{m}_\mathbf{C}\rangle$.*

Proof. Sufficiency is trivial. For necessity, consider the universe of $\mathbf{C}_0 \perp \mathbf{C}_1$ in condition (3) of Proposition 3.2 endowed with $\leq$ defined by $x \leq y$ if $x \in C_1$ and $y \in C_0$ or $x, y \in C_0$ and $x \leq_0 y$ or $x, y \in C_1$ and $y \leq_1 x$. □

As stated in the next result, the totally ordered set $\mathbf{C}$ given in Theorem 3.3 is unique, up to (dual) isomorphism.

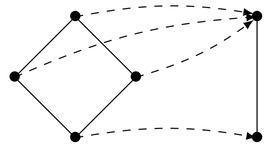
Theorem 3.4. *Let $\mathbf{A}$ be a median algebra. If $\mathbf{C}$ and $\mathbf{C}'$ are two chains such that $\mathbf{A} \cong \langle \mathbf{C}, \mathbf{m}_\mathbf{C}\rangle$ and $\mathbf{A} \cong \langle \mathbf{C}', \mathbf{m}_{\mathbf{C}'}\rangle$, then $\mathbf{C}$ is order isomorphic or dual order isomorphic to $\mathbf{C}'$.*

4 HOMOMORPHISMS BETWEEN CONSERVATIVE MEDIAN ALGEBRAS

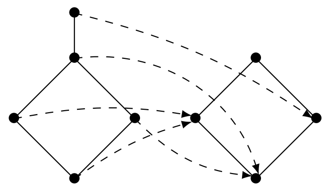
In view of Theorem 3.3 and Theorem 3.4, we introduce the following notation. Given a conservative median algebra $\mathbf{A}$ ($|A| \geq 5$), we denote a chain representation of $\mathbf{A}$ by $\mathbf{C}(\mathbf{A})$, that is, $\mathbf{C}(\mathbf{A})$ is a chain such that $\mathbf{A} \cong \langle \mathbf{C}(\mathbf{A}), \mathbf{m}_{\mathbf{C}(\mathbf{A})}\rangle$, and we denote the corresponding isomorphism by $f_\mathbf{A} : \mathbf{A} \to \langle \mathbf{C}(\mathbf{A}), \mathbf{m}_{\mathbf{C}(\mathbf{A})}\rangle$. If $f : \mathbf{A} \to \mathbf{B}$ is a map between two conservative median algebras with at least five elements, the map $f' : \mathbf{C}(\mathbf{A}) \to \mathbf{C}(\mathbf{B})$ defined as $f' = f_\mathbf{B} \circ f \circ f_\mathbf{A}^{-1}$ is said to be *induced by* f.

A function $f : \mathbf{A} \to \mathbf{B}$ between median algebras $\mathbf{A}$ and $\mathbf{B}$ is called a *median homomorphism* if it satisfies equation (1.3). We use the terminology introduced above to characterize median homomorphisms between conservative median algebras. Recall that a map between two posets is *monotone* if it is isotone or antitone.

Theorem 4.1. *Let $\mathbf{A}$ and $\mathbf{B}$ be two conservative median algebras with at least five elements. A map $f : \mathbf{A} \to \mathbf{B}$ is a median homomorphism if and only if the induced map $f' : \mathbf{C}(\mathbf{A}) \to \mathbf{C}(\mathbf{B})$ is monotone.*



(a) A monotone map which is not a median homomorphism.



(b) A median homomorphism which is not monotone.

Figure 2: Examples for Remark 4.3.

Corollary 4.2. *Let $\mathbf{C}$ and $\mathbf{C}'$ be two chains. A map $f : \mathbf{C} \to \mathbf{C}'$ is a median homomorphism if and only if it is monotone.*

Remark 4.3. Note that Corollary 4.2 only holds for chains. Indeed, Fig. 2(a) gives an example of a monotone map that is not a median homomorphism, and Fig. 2(b) gives an example of median homomorphism that is not monotone.

Since the class of conservative median algebras is clearly closed under homomorphic images, we obtain the following corollary.

Corollary 4.4. *Let $\mathbf{A}$ and $\mathbf{B}$ be two median algebras and $f : \mathbf{A} \to \mathbf{B}$. If $\mathbf{A}$ is conservative, and if $|A|, |f(A)| \geq 5$, then f is a median homomorphism if and only if $f(\mathbf{A})$ is a conservative median subalgebra of $\mathbf{B}$ and the induced map $f' : \mathbf{C}(\mathbf{A}) \to \mathbf{C}(f(\mathbf{A}))$ is monotone.*

We are actually able to lift the previous result to finite products of chains. If $f_i : A_i \to A'_i$ ($i \in [n]$) is a family of maps, let $(f_1, \ldots, f_n) : A_1 \times \cdots \times A_n \to A'_1 \times \cdots \times A'_n$ be defined by

$$(f_1, \ldots, f_n)(x_1, \ldots, x_n) := (f_1(x_1), \ldots, f_n(x_n)).$$

The following theorem characterizes median homomorphisms between finite products of chains.

Theorem 4.5. *Let $\mathbf{A} = \mathbf{C}_1 \times \cdots \times \mathbf{C}_k$ and $\mathbf{B} = \mathbf{D}_1 \times \cdots \times \mathbf{D}_n$ be two finite products of chains. Then $f : \mathbf{A} \to \mathbf{B}$ is a median homomorphism if and only if there exist $\sigma : [n] \to [k]$ and monotone maps $f_i : \mathbf{C}_{\sigma(i)} \to \mathbf{D}_i$ for $i \in [n]$ such that $f = (f_{\sigma(i)}, \ldots, f_{\sigma(n)})$.*

As an immediate consequence, it follows that aggregation functions compatible with median functions on ordinal scales are dictorial.

If $A = A_1 \times \cdots \times A_n$ and $i \in [n]$, then we denote the projection map from A onto A_i by π_i^A, or simply by π_i if there is no danger of ambiguity.

Corollary 4.6. *Let* $\mathbf{C}_1, \ldots, \mathbf{C}_n$ *and* $\mathbf{D}$ *be chains. A map* $f : \mathbf{C}_1 \times \cdots \times \mathbf{C}_n \to \mathbf{D}$ *is a median homomorphism if and only if there is a* $j \in [n]$ *and a monotone map* $g : \mathbf{C}_j \to \mathbf{D}$ *such that* $f = g \circ \pi_j$.

In the particular case of Boolean algebras (*i.e.*, powers of a two element chain), Theorem 4.5 can be restated as follows.

Corollary 4.7. *Assume that* $f : \mathbf{2}^n \to \mathbf{2}^m$ *is a map between two finite Boolean algebras.*

(1) The map f *is a median homomorphism if and only if there are* $\sigma : [m] \to ([n] \cup \{\perp\})$ *and* $\epsilon : [m] \to \{\mathrm{id}, \neg\}$ *such that*

$$f : (x_1, \ldots, x_n) \mapsto (\epsilon_1 x_{\sigma_1}, \ldots, \epsilon_m x_{\sigma_m}),$$

where $x_\perp$ *is defined as the constant map* 0.

In particular,

(2) A map $f : \mathbf{2}^n \to \mathbf{2}$ *is a median homomorphism if and only if it is a constant function, a projection map or the negation of a projection map.*

(3) A map $f : \mathbf{2}^n \to \mathbf{2}^n$ *is a median isomorphism if and only if there is a permutation* σ *of* $[n]$ *and an element* ϵ *of* $\{\mathrm{id}, \neg\}^n$ *such that* $f(x_1, \ldots, x_n) = (\epsilon_1 x_{\sigma(1)}, \ldots, \epsilon_n x_{\sigma(n)})$ *for any* $(x_1, \ldots, x_n)$ *in* A.

5 CONCLUDING REMARKS AND FURTHER RESEARCH DIRECTIONS

In this paper we have described conservative median algebras and semilattices with at least five elements in terms of forbidden configurations and have given a representation by chains. We have also characterized median-preserving maps between finite products of these algebras, showing that they are essentially determined componentwise. The next step in this line of research is to extend our results to larger classes of median algebras and their ordered counterparts.

Acknowledgements

This work was supported by the internal research project F1R-MTH-PUL-15MRO3 of the University of Luxembourg.

References

[1] S. P. Avann. Metric ternary distributive semi-lattices. *Proceedings of the American Mathematical Society*, 12:407–414, 1961.

[2] H. J. Bandelt and J. Hedlíková. Median algebras. *Discrete mathematics*, 45:1–30, 1983.

[3] H. J. Bandelt. Discrete ordered sets whose covering graphs are median. *Proceedings of the American Mathematical Society*, 91(1):6–8, 1984.

[4] G. Birkhoff and S. A. Kiss. A ternary operation in distributive lattices. *Bulletin of the American Mathematical Society*, 53:749–752, 1947.

[5] M. Couceiro and B. Teheux, An Arrow-like theorem over median algebras, *Proceedings of the DA2PL 2014 Conference*, Ecole Centrale de Paris, France, November 20–21 2014, pp. 72 – 75.

[6] M. Couceiro, J.-L. Marichal and B. Teheux, Conservative Median Algebras, `arXiv:1405.0935`.

[7] B. A. Davey and H. A. Priestley. *Introduction to lattices and order.* Cambridge University Press, New York, second edition, 2002.

[8] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions.* Encyclopedia of Mathematics and its Applications, vol. 127. Cambridge University Press, Cambridge, 2009.

[9] G. Grätzer. *General lattice theory.* Birkhäuser Verlag, Basel, second edition, 1998. New appendices by the author with B. A. Davey, R. Freese, B. Ganter, M. Greferath, P. Jipsen, H. A. Priestley, H. Rose, E. T. Schmidt, S. E. Schmidt, F. Wehrung and R. Wille.

[10] A. A. Grau. Ternary Boolean algebra. *Bulletin of the American Mathematical Society*, (May 1944):567–572, 1947.

[11] J. R. Isbell. Median algebra. *Transactions of the American Mathematical Society*, 260(2):319–362, 1980.

[12] M. Sholander. Trees, lattices, order, and betweenness. *Proceedings of the American Mathematical Society*, 3(3):369–381, 1952.

[13] M. Sholander. Medians, lattices, and trees. *Proceedings of the American Mathematical Society*, 5(5):808–812, 1954.

CONSTRUCTION OF ADMISSIBLE ORDERS FOR INTERVALS

Laura De Miguel,
Humberto Bustince
and
Javier Fernandez
Universidad Publica de Navarra,
Pamplona, Spain
{laura.demiguel,bustince,
fcojavier.fernandez}@unavarra.es

Anna Kolesárová
and
Radko Mesiar
Slovak University of Technology,
Bratislava, Slovakia
{anna.kolesarova,radko.mesiar}
@stuba.sk

Bernard De Baets
Ghent University,
Gent, Belgium
Bernard.DeBaets@UGent.be

Summary

Many theoretical and applied developments are deeply grounded in the use of linear orderings. Examples of such are Choquet and Sugeno integrals (in a theoretical context), or decision making and fuzzy classification (in an applied one). However, linear ordering becomes significantly more complicated when it applies to non-scalar or intervalar data. In this work we propose two methods to construct a particular class of linear orders for tuples of intervals, which we refer to as admissible orders. Besides, some relations between both methods are studied.

Keywords: Interval-Valued Fuzzy sets, linear order, admissible order, K_α operators.

1 INTRODUCTION

The elements in $\mathbb{R}$ and, in particular, the elements in $[0,1]$ are linearly ordered. This fact turned out to be crucial in theoretical and applied developments in the context of fuzzy set theory. When using classical fuzzy sets, and given two elements of a referential set U, their membership degree can be compared and, subsequently, ordered. However, this process becomes considerably more complicated when using extensions of fuzzy sets, examples being interval-valued fuzzy sets or set-valued fuzzy sets. In such cases, there can be elements $u_1, u_2 \in U$ whose membership degree is not comparable. Consequently, membership degrees can not be ordered, hindering the generalization of certain tools. Since the advantages of using such extensions seem to be noticeable for certain applications [2, 7, 10, 11] we have generated a method to construct linear orders for vectors of intervals in $[0,1]$. As tuples of m intervals can be represented as elements in $\mathbb{R}^{2m}$, we only consider orders which refine the partial order on $\mathbb{R}^{2m}$. We refer to these intervals as *admissible* orders. Given the relevant role of monotonicity in the constructions of our orders, aggregation functions become a useful tool in our construction methods.

The remainder of the paper is organized as follows. In Section 2 we introduce some known concepts and previous results. In Section 3 the definition of admissible orders is presented while in Section 4 two methods to construct them are introduced. Some concluding remarks and open problems are included in Section 5.

2 PRELIMINARY NOTIONS

In this paper we aim to generate linear orders in the set of m real intervals. We expect them to be useful for the comparison of membership degrees in several extensions of fuzzy sets, including interval-valued fuzzy sets and in some particular instances of set-valued fuzzy sets.

We denote $L([0,1])$ to the set of all closed subintervals of the unit interval, that is

$$L([0,1]) = \{\mathbf{x} = [\underline{x}_1, \overline{x}_1] | 0 \leq \underline{x_1} \leq \overline{x_1} \leq 1\}.$$

and $K([0,1]) \subset \mathbb{R}^2$ to the set

$$K([0,1]) = \{(\underline{x}, \overline{x}) \in [0,1]^2 | \, \underline{x} \leq \overline{x}\}.$$

Definition 2.1 ([9]) *An Interval-Valued Fuzzy Set(IVFS) A on a universe U is a mapping $A : U \to L([0,1])$, where $A(u)$ denotes the membership degree of the element u to the IVFS A.*

2.1 ON ORDERS AND PARTIALLY ORDERED SETS

Definition 2.2 *A partial order $\preceq$ on a set P is a binary relation which is reflexive, antisymmetric and transitive. If $\preceq$ is a partial order, the pair $(P, \preceq)$ is called a partially ordered set (poset).*

ISBN: 978–83–8012–519–3

Given a poset $(P, \preceq)$, and $x, y \in P$, we say that x and y are comparable if $x \preceq y$ or $y \preceq x$. Besides, we call

a) 1_P, the top of the poset, if for all $x \in P$ it holds $x \preceq 1_P$;

b) 0_P, the bottom of the poset, if for all $x \in P$ it holds $0_P \preceq x$.

Notice that, in case 1_P and 0_P exist, they are unique.

The poset $(\mathbb{R}^2, \preceq_2)$, where $\leq_2$ is given by $(p_1, p_2) \leq_2 (q_1, q_2)$ if and only if $p_1 \leq q_1$ and $p_2 \leq q_2$, induces a partial order on $L([0,1])$, $\preceq_2$, which is given by

$$\mathbf{x} \preceq_2 \mathbf{y} \text{ if and only if } \underline{x_1} \leq \underline{y_1} \text{ and } \overline{x_1} \leq \overline{y_1}. \quad (1)$$

Example 2.1 *Let the intervals* $\boldsymbol{x} = [0.05, 0.2]$, $\boldsymbol{y} = [0.1, 0.4]$ *and* $\boldsymbol{z} = [0, 0.6]$. *The partial order in Eq.* (1) *yields that* $\boldsymbol{x} \preceq_2 \boldsymbol{y}$ *but also that* $\boldsymbol{z}$ *is incomparable to both* $\boldsymbol{x}$ *and* $\boldsymbol{y}$.

Definition 2.3 *A linear order* $\leq$ *on* P *is a binary transitive, antisymmetric and total relation. Equivalently, a linear order is a partial order under which every pair of elements is comparable.*

Example 2.2 *Examples of linear orders on* $L([0,1])$ *are the lexicographic orders, given by:*

- *(lexicographic-1 order)* $\boldsymbol{x} \leq_{lex1} \boldsymbol{y}$ *if and only if* $\underline{x_1} < \underline{y_1}$ *or* $(\underline{x_1} = \underline{y_1}$ *and* $\overline{x_1} \leq \overline{y_1})$;
- *(lexicographic-2 order)* $\boldsymbol{x} \leq_{lex1} \boldsymbol{y}$ *if and only if* $\overline{x_1} < \overline{y_1}$ *or* $(\overline{x_1} = \overline{y_1}$ *and* $\underline{x_1} \leq \underline{y_1})$.

In [4], a special class of linear orders in $L([0,1])$ is defined.

Definition 2.4 ([4]) *An order* $\leq$ *on* $L([0,1])$ *is said to be admissible if it is linear and refines the partial order* $\preceq_2$ *given in Eq.* (1), *i.e., if it is a linear order satisfying that, for all* $\boldsymbol{x}, \boldsymbol{y} \in L([0,1])$ *such that* $\boldsymbol{x} \preceq_2 \boldsymbol{y}$, *it holds* $\boldsymbol{x} \leq \boldsymbol{y}$.

Since the membership degrees of IVFSs are in $L([0,1])$, these admissible orders can be used in applications dealing with IVFSs where a ranking must be calculated (see [3]). In particular the lexicographic orders presented in Example 2.2 are admissible.

2.2 AGGREGATION FUNCTIONS

Aggregation functions are a common tool to fuse and aggregate information.

Definition 2.5 *A* k*-ary aggregation function is an increasing mapping* $M : [0,1]^k \to [0,1]$ *such that* $M(0,\ldots,0) = 0$ *and* $M(1,\ldots,1) = 1$.

For further information and some generalizations of these functions see [8, 6, 1, 5]. In [4], aggregation functions were used to generate admissible orders on $L([0,1])$.

Proposition 2.1 ([4]) *Let* $B_1, B_2 : [0,1]^2 \to [0,1]$ *be two continuous aggregation functions such that, for all* $(p_1,p_2),(q_1,q_2) \in K([0,1])$, $B_1(p_1,p_2) = B_1(q_1,q_2)$ *and* $B_2(p_1,p_2) = B_2(q_1,q_2)$ *if and only if* $(p_1,p_2) = (q_1,q_2)$.

The order $\leq_{B_1,B_2}$ *on* $L([0,1])$, *given by*

$$\boldsymbol{x} \leq_{B_1,B_2} \boldsymbol{y} \text{ if and only if } B_1(\underline{x},\overline{x}) < B_1(\underline{y},\overline{y}) \text{ or } (B_1(\underline{x},\overline{x}) = B_1(\underline{y},\overline{y}) \text{ and } B_2(\underline{x},\overline{x}) \leq B_2(\underline{y},\overline{y})),$$

is an admissible order on $L([0,1])$.

In particular, we are interested in $\preceq_{B_1,B_2}$ with B_1, B_2 being two different Atanassov's operators for some $\alpha_1, \alpha_2 \in [0,1]$, i.e., aggregation functions such that $B_i(\underline{x},\overline{x}) = \underline{x} + \alpha_i(\overline{x} - \underline{x})$ with $\alpha_1 \neq \alpha_2$.

3 ADMISSIBLE ORDERS ON THE SET OF m INTERVALS

In very special situations, partial orders can completely sort the elements in a given subset, making it unnecessary the use of linear orders. However, in the great majority of the cases, there exist incomparable elements which demand the design of such orders.

Let $L^m([0,1])$ be the set of m-tuples of closed subintervals of $[0,1]$, that is

$$L^m([0,1]) = \{([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) \mid 0 \leq \underline{x}_i \leq \overline{x}_i \leq 1 \text{ for all } i \in \{1,\ldots,m\}\},$$

and let $K^{2m}([0,1])$ be the subset of $\mathbb{R}^{2m}$ given by

$$K^{2m}([0,1]) = \{(\underline{x_1},\overline{x_1},\underline{x_2},\overline{x_2},\ldots,\underline{x_m},\overline{x_m}) \in [0,1]^{2m} \mid \underline{x_i} \leq \overline{x_i} \text{ for all } i \in \{1,\ldots,m\}\}.$$

There exist a bijection

$$g : K^{2m}([0,1]) \to L^m([0,1])$$

given by

$$g((\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x^m})) = ([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]).$$

Through this bijection, the partial order on $\mathbb{R}^{2m}$,

$$(\underline{p_1},\overline{p_1},\ldots,\underline{p_m},\overline{p_m}) \preceq_{2m} (\underline{q_1},\overline{q_1},\ldots,\underline{q_m},\overline{q_m})$$

if and only if

$$\underline{p_1} \leq \underline{q_1},\ \overline{p_1} \leq \overline{q_1},\ \ldots,\ \text{ and } \overline{p_m} \leq \overline{q_m},$$

induces an equivalent partial order on $L^m([0,1])$, given by

$$([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) \preceq_{2m} ([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m])$$
$$\text{if and only if}$$
$$\underline{x_1} \leq \underline{y_1},\ \overline{x_1} \leq \overline{y_1},\ \ldots,\ \text{ and } \overline{x_m} \leq \overline{y_m}. \quad (2)$$

In this way, $(L^m([0,1]), \preceq_{2m})$ is a poset whose bottom and top are $\mathbf{0} = ([0,0],\ldots,[0,0])$ and $\mathbf{1} = ([1,1],\ldots,[1,1])$, respectively.

Definition 3.1 *An order $\leq$ on $L^m([0,1])$ is admissible if it is linear on $L^m([0,1])$ refining the order $\preceq_{2m}$ in Eq. (2).*

4 CONSTRUCTION OF ADMISSIBLE ORDERS ON $L^m([0,1])$

Since admissible orders between intervals have been already defined and constructed in [4], our first method to generate admissible orders on $L^m([0,1])$ elaborates on them.

Proposition 4.1 *Let $\preceq_{B_1,B_2}$ be an admissible order on $L([0,1])$, as in Prop. 2.1, and let $\sigma = (\sigma(1),\ldots,\sigma(m))$ be a permutation. The order $\preceq_{[\sigma,B]}$ on $L^m([0,1])$, given by*

$$([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) \preceq_{[\sigma,B]} ([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m])$$

if and only if one of the following conditions holds

- $[\underline{x}_{\sigma(1)},\overline{x}_{\sigma(1)}] \prec_{B_1,B_2} [\underline{y}_{\sigma(1)},\overline{y}_{\sigma(1)}]$
- $[\underline{x}_{\sigma(1)},\overline{x}_{\sigma(1)}] = [\underline{y}_{\sigma(1)},\overline{y}_{\sigma(1)}]$ *and* $[\underline{x}_{\sigma(2)},\overline{x}_{\sigma(2)}] \prec_{B_1,B_2} [\underline{y}_{\sigma(2)},\overline{y}_{\sigma(2)}]$

  ...
- $[\underline{x}_{\sigma(i)},\overline{x}_{\sigma(i)}] = [\underline{y}_{\sigma(i)},\overline{y}_{\sigma(i)}]$ *for all* $i \in \{1,\ldots,m-2\}$ *and* $[\underline{x}_{\sigma(m-1)},\overline{x}_{\sigma(m-1)}] \prec_{B_1,B_2} [\underline{y}_{\sigma(m-1)},\overline{y}_{\sigma(m-1)}]$
- $[\underline{x}_{\sigma(i)},\overline{x}_{\sigma(i)}] = [\underline{y}_{\sigma(i)},\overline{y}_{\sigma(i)}]$ *for all* $i \in \{1,\ldots,m-1\}$ *and* $[\underline{x}_{\sigma(m)},\overline{x}_{\sigma(m)}] \preceq_{B_1,B_2} [\underline{y}_{\sigma(m)},\overline{y}_{\sigma(m)}]$

is an admissible order.

Proof. Direct, since $\preceq_{B_1,B_2}$ is an admissible order on $L([0,1])$.

Notice that although we have taken the same admissible order to compare each interval, there is no problem in changing and comparing each interval with a different admissible order. Next, we present a more general method to construct admissible orders on $L^m([0,1])$.

Definition 4.1 *Let $A = (A_1, A_2, \ldots, A_{2m})$ be $2m$ aggregation functions $A_i : [0,1]^{2m} \to [0,1]$. The $2m$-tuple A is admissible if for all elements $([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m])$,$([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m]) \in L^m([0,1])$,*

$$A_i(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m}) = A_i(\underline{y_1},\overline{y_1},\ldots,\underline{y_m},\overline{y_m})$$

for all $i \in \{1,\ldots,2m\}$ if and only if

$$([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) = ([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m]).$$

Proposition 4.2 *Let A be an admissible $2m$-tuple of aggregation functions. An admissible order $\preceq_A$ on $L^m([0,1])$ can be defined as*

$$([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) \prec_A ([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m])$$

if and only if there is a $k \in \{1,\ldots,m\}$ such that

$$A_i(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m}) = A_i(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m})$$

for all $i \in S = \{1,\ldots,k-1\}$ and

$$A_k(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m}) < A_k(\underline{y_1},\overline{y_1},\ldots,\underline{y_m},\overline{y_m}),$$

provided that $k = 1$ induces $S = \emptyset$.

Besides,

$$([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_m,\overline{x}_m]) = ([\underline{y}_1,\overline{y}_1],\ldots,[\underline{y}_m,\overline{y}_m])$$

if and only if $\underline{x_i} = \underline{y_i}$ and $\overline{x_i} = \overline{y_i}$ for all $i \in \{1,\ldots,m\}$.

Proof. The order $\preceq_A$ refines $\leq_{2m}$, since every A_i is an aggregation function. Besides, linearity is assured since the $2m$-tuple is admissible. The transitivity follows from the transitivity of the order on $[0,1]$.

Notice that the presented method in Prop. 4.1 is a particular case of Prop. 4.2 where

$$A_{2i-1}(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m}) = B_1([\underline{x}_{\sigma(i)},\overline{x}_{\sigma(i)}])$$

and

$$A_{2i}(\underline{x_1},\overline{x_1},\ldots,\underline{x_m},\overline{x_m}) = B_2([\underline{x}_{\sigma(i)},\overline{x}_{\sigma(i)}])$$

for all $i \in \{1,\ldots,m\}$.

Example 4.1 *Lexicographic orders are admissible, and they can be constructed as in Prop. 4.2 choosing as aggregations functions the projections in $\mathbb{R}^{2m}$ (the projections can not be repeated or they are not an admissible $2m$-tuple).*

In particular, the lexicographic 1 orders given by $A_i(u_1,\ldots,u_{2m}) = \Pi_i(u_1,\ldots,u_{2m}) = u_i$ for all $i \in \{1,\ldots,2m\}$.

The main difficulty in using the second construction method (Prop. 4.2) consists of finding aggregation functions which satisfy the condition given in Eq. (4.1). However, once some aggregation functions have been found, any permutation of them generates another admissible order.

Proposition 4.3 *Let* $A = (A_1, A_2, \ldots, A_{2m})$ *be* $2m$ *aggregation functions* $A_i : [0,1]^{2m} \to [0,1]$ *given by*

$$A_i(\underline{x_1}, \overline{x_1}, \ldots, \underline{x_m}, \overline{x_m}) = a_{i1}\underline{x_1} + a_{i2}\overline{x_1} + \ldots + a_{i,2m}\overline{x_m}\,,$$

with $a_{i1} + a_{i2} + \ldots + a_{i,2m} = 1$ *and let the matrix* D*, given by*

$$D = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,2m} \\ a_{21} & a_{22} & \cdots & a_{2,2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{2m,1} & a_{2m,2} & \cdots & a_{2m,2m} \end{pmatrix},$$

be a regular matrix. Then the order $\preceq_A$*, generated by* A*, is an admissible order on* $L^m([0,1])$*.*

Proof. Let $([\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_m, \overline{x}_m])$ and $([\underline{y}_1, \overline{y}_1], \ldots, [\underline{y}_m, \overline{y}_m])$ be in $L^m([0,1])$ such that $A_i(\underline{x_1}, \overline{x_1}, \ldots, \underline{x_m}, \overline{x_m}) = A_i(\underline{y_1}, \overline{y_1}, \ldots, \underline{y_m}, \overline{y_m})$ for all $i \in \{1, \ldots, 2m\}$. Then, $D \cdot \mathbf{x} = D \cdot \mathbf{y}$ where $\cdot$ denotes the product of matrices and $\mathbf{x} = (\underline{x_1}, \overline{x_1}, \ldots, \underline{x_m}, \overline{x_m})^{\mathrm{t}}$. Due to the regularity of D, this is true if $([\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_m, \overline{x}_m]) = ([\underline{y}_1, \overline{y}_1], \ldots, [\underline{y}_m, \overline{y}_m])$ and consequently the A is an admissible $2m$-tuple.

Let D to be block matrix of the form

$$D = \left(\begin{array}{c|c|c|c} I_{11} & I_{12} & \cdots & I_{1m} \\ \hline I_{21} & I_{22} & \cdots & I_{2m} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline I_{m1} & I_{m2} & \cdots & I_{mm} \end{array}\right),$$

where each block I_{ij} is a 2×2 matrix. If in each row and in each column, there is only one block which is different from the nule matrix then, this block is

$$I_{i,k} = \begin{pmatrix} 1-\alpha_1 & \alpha_1 \\ 1-\alpha_2 & \alpha_2 \end{pmatrix}$$

for some $\alpha_1, \alpha_2 \in [0,1]$ with $\alpha_1 \neq \alpha_2$. Then the order generated as in Prop. 4.3 is the same as in Prop. 4.1 when the order on $L([0,1])$ is generated by Atanassov's operators (Prop. 2.1). This example is particularly important, as in Prop. [4] it is proven that different Atanassov's operators can yield the same linear order on intervals. Consequently, there are $2m$-tuples of different aggregation functions which generate the same linear order on $L^m([0,1])$.

5 CONCLUSIONS

In this work we have defined and generated admissible orders on $L^m([0,1])$. We have seen two different construction methods, the former producing particular instances of those by the latter. The condition we must impose to aggregation functions in order to generate admissible $2m$-tuples is quite strong.

Besides, by means of Prop. 3.8 we know that different weighted arithmetic means (aggregation functions) which generate matrices as in Eq. (4) yield the same linear order [4]. As interesting theoretical problems we open the questions:

- Apart from the weighted arithmetic means, are there any other examples of aggregation functions which generate admissible $2m$-tuples?
- Are there any other matrices D (not by blocks) that generate the same linear order?

We want to call the attention that our two construction methods are just a first approach and some others methods to construct linear (and admissible) orders are possible. However, we let a deeper study of the order for future research.

Acknowledgements

The work has been supported by projects TIN2013-40765-P and MTM2012-37894-C02-02 of the Spanish Ministry of Science and the Research Services of the Universidad Publica de Navarra. R. Mesiar and A. Kolesárová acknowledge the support of the grants VEGA 1/0419/13 and VEGA 1/0420/15.

References

[1] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation functions: a guide for practitioners. *Studies in Fuzziness and Soft Computing, Springer.*

[2] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z.S. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, B. De Baets (2015). A historical account of types of fuzzy sets and their relationships. *IEEE Trans. on Fuzzy Systems,* In press.

[3] H. Bustince, M. Galar, B. Bedregal, A. Kolesárová, R. Mesiar (2013). A new approach to interval-valued Choquet integrals and the problem of ordering in interval-valued fyzzy set applications. *IEEE Trans. on Fuzzy Systems* **21(6)**, pp 1150-1162.

[4] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar (2013). Generation of linear orders for in-

tervals by means of aggregation functions. *Fuzzy Sets and Systems* **220**, pp 69-77.

[5] T. Calvo, A. Kolesárová, M. Komorníková, R. Mesiar (2002). Aggregation operators: properties, classes and construction methods. *Studies in Fuzziness and Soft Computing, Springer.*

[6] M. Grabish, J.L. Marichal, R. Mesiar, E. Pap (2009). Aggregation Functions. *Cambridge University Press.*

[7] A. Jurio, M. Pagola, R. Mesiar, G. Beliakov, H. Bustince (2011). Image magnification using interval information. *IEEE Trans. on Image Processing* **20(11)**, pp 3112-3123.

[8] M. Komorníková, R. Mesiar (2011). Aggregation functions on bounded partially ordered sets and their classification. *Fuzzy Sets and Systems* **175(1)**, pp 48-56.

[9] R. Sambuc (1975). Fonctions and floues: application a l'aide au diagnostic en pathologie thyroidienne. *Faculté de Médecine de Marseille.*

[10] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, H. Hagras (2014). A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data, *IEEE Trans. on Fuzzy Systems,* In press.

[11] J. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, H. Bustince (2014). Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system. *Applied Soft Computing* **20**, pp. 103 - 111.

FUZZY IMPLICATIONS AND THE LAW OF O-CONDITIONALITY: THE CASE OF RESIDUAL IMPLICATIONS DERIVED FROM OVERLAP FUNCTIONS

Graaliz Pereira Dimuro
PPGCOMP, C3, FURG, Brazil
gracalizdimuro@furg.br

Benjamn Bedregal
DIMAP, UFRN, Brazil
bedregal@dimap.ufrn.br

Summary

Overlap functions are aggregation operators used in the overlap problem or when the associativity is not required. Residual implications derived from them (R_O-implications) preserve the residuation property, and any overlap function O and the respective R_O-implication form an adjoint pair, which is important in many applications. R_O-implications do not necessarily satisfy certain properties, but only weaker versions of these properties, e.g., the exchange principle. However, in general, such properties are not demanded for many applications. In this paper, we analyze the so-called law of O-Conditionality, $O(x, I(x, y)) \leq y$, for any fuzzy implication I and, in particular, for R_O-implications.

Keywords: Overlap functions, fuzzy implications, residual implications, O-Conditionality.

1 INTRODUCTION

Fuzzy implications [3] generalize the classical implication to fuzzy logic, by considering truth values varying in the unit interval $[0, 1]$ instead of in the set $\{0, 1\}$. There are different ways for obtaining this generalization [3] (e.g., S, QL, R, (U, N), RU, and D-implications [5, 6, 22, 27]). The importance of fuzzy implications in applications was discussed recently by Baczyński [2] and other authors [12].

R-implications are generalizations to $[0, 1]$ of Boolean implications defined by the identity given, for a universe set X, by $A' \cup B = (A - B)' = \bigcup\{C \subseteq X \mid (A \cap C) \subseteq B\}$, where $A, B \subseteq X$ [3, 5], where the intersection is generalized by a t-norm.

A t-norm requires the associativity and commutativity properties, which, in their turn, allow any R-implication I to satisfy the exchange principle. However, in the literature, it was shown that the associativity property is not demanded for many applications, e.g., in pairwise comparisons, image processing and mathematical morphology. [7, 12, 18]

Bustince et al. [7] introduced the *overlap functions*, a particular case of bivariate continuous aggregation operators defined by (not necessarily associative) increasing commutative bivariate functions, satisfying appropriate boundary conditions (see, e.g., [4, 9, 11]). Overlap functions have been applied in classification problems, image processing and decision making.

In [10], based on residual implicators of general conjunctions [16, 20], we introduced the R_O-implications, the residual implications derived from overlap functions O, preserving the residuation property. R_O-implications do not necessarily satisfy certain properties of R-implications, but only weaker versions of these properties. However, in general, such properties are not demanded for many applications. [10]

In this paper we analyze the law of O-Conditionality, defined by $O(x, I(x, y)) \leq y$, for any fuzzy implication I and overlap function O, and, in particular, for R_O-implications.[1] Section 2 presents basic concepts. R_O-implications are studied in Sect. 3. The O-conditionality for fuzzy implications in general is analysed in Sect. 4, and, in particular, in Sect. 5, for R_O-implications. Section 6 is the Conclusion.

2 PRELIMINARY CONCEPTS

Definition 2.1. A function $N : [0, 1]^2 \to [0, 1]$ is said to be a fuzzy negation if the following conditions hold:

(N1) N satisfies the Boundary Conditions: $N(0) = 1$ and $N(1) = 0$;

[1]The O-Conditionality is seen as a generalization of modus ponens, $x *_O (x \to y) \leq y$, which can be understood as $x \wedge_O (x \to y) \vdash y$.

ISBN: 978–83–8012–519–3

(N2) N is decreasing: if $x \leq y$ then $N(y) \leq N(x)$.

Definition 2.2. A function $A : [0,1]^n \to [0,1]$ is said to be an n-ary aggregation operator if: **(A1)** A is increasing[2] in each argument: for each $i \in \{1,\ldots,n\}$, if $x_i \leq y$, then $A(x_1,\ldots,x_n) \leq A(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)$; **(A2)** A satisfies the Boundary conditions: $A(0,\ldots,0) = 0$ and $A(1,\ldots,1) = 1$.

Definition 2.3. A bivariate function $O : [0,1]^2 \to [0,1]$ is said to be an overlap function if it satisfies the conditions: **(O1)** O is commutative; **(O2)** $O(x,y) = 0$ if and only if $xy = 0$; **(O3)** $O(x,y) = 1$ if and only if $xy = 1$; **(O4)** O is increasing; **(O5)** is continuous.

Some examples of overlap functions are [4, 9, 10, 11]:

$$\begin{aligned} O_{mM}(x,y) &= \min\{x,y\}\max\{x^2,y^2\}; \\ O^V_{mM}(x,y) &= \begin{cases} \frac{1+O_{mM}(2x-1,2y-1)}{2} & \text{if } x,y \in]0.5,1]; \\ \min\{x,y\} & \text{otherwise}; \end{cases} \\ O_p(x,y) &= x^p y^p \text{ with } p > 0 \text{ and } p \neq 1. \end{aligned}$$

An overlap function $O : [0,1]^2 \to [0,1]$ satisfies the property 1-section deflation if $O(x,1) \leq x$. **(O6)**

An overlap function $O : [0,1]^2 \to [0,1]$ satisfies the property 1-section inflation if $O(x,1) \geq x$. **(O7)**

An overlap function O satisfies **(O6)** and **(O7)** if and only if O has 1 as neutral element. Whenever an overlap function has a neutral element, then, by **(O3)**, this element is necessarily equal to 1. An overlap function O is associative if and only if O is a continuous and positive t-norm. However, O_{mM} is a non associative overlap function with 1 as neutral element.

Lemma 2.1. *An overlap function* $O : [0,1]^2 \to [0,1]$ *satisfies* **(O6)** *if and only if* $O \leq \min$.

Proof. It follows that: (⇒) If O satisfies **(O6)** then, for all $x,y \in [0,1]$, it holds that $O(x,y) \leq O(x,1) \leq x$. On the other hand, $O(x,y) = O(y,x) \leq O(y,1) \leq y$. Thus, one has that $O(x,y) \leq \min\{x,y\}$. (⇐) It is immediate. □

A function $\varphi : [0,1] \to [0,1]$ is said to be an *automorphism* if φ is bijective and increasing.

Definition 2.4. A function $I : [0,1]^2 \to [0,1]$ is a fuzzy implication if, for each $x,y,z \in [0,1]$, it holds that: **(I1)** first place antitonicity: if $x \leq y$ then $I(y,z) \leq I(x,z)$; **(I2)** second place isotonicity: if $y \leq z$ then $I(x,y) \leq I(x,z)$; **(I3)** boundary condition 1: $I(0,0) = 1$; **(I4)** boundary condition 2: $I(1,1) = 1$; **(I5)** boundary condition 3: $I(1,0) = 0$.

[2]In this paper, a increasing (decreasing) function does not need to be strictly increasing (decreasing).

There exist several properties that may be required for fuzzy implications [3]. In the following, we present some properties that are used in this paper.

Definition 2.5. A fuzzy implication $I : [0,1]^2 \to [0,1]$ satisfies: **(I6)** the ordering property if and only if $\forall x,y \in [0,1] : I(x,y) = 1 \Leftrightarrow x \leq y$; **(I7)** the strong boundary condition for 0 if and only if $\forall x \in [0,1] : x \neq 0 \Rightarrow I(x,0) = 0$; **(I8)** the exchange principle for 1 if and only if $\forall x \in [0,1] : I(x,I(y,z)) = 1 \Rightarrow I(y,I(x,z)) = 1$; **(I9)** the identity principle if and only if $\forall x \in [0,1], I(x,x) = 1$; **(I10)** the pseudo-exchange principle if and only if $\forall x,y,z \in [0,1] : I(x,z) \geq y \Leftrightarrow I(y,z) \geq x$; **(I11)** the conditional antecedent boundary if only if $\forall x,y \in [0,1] : x > y \Rightarrow I(x,y) < y$.

Definition 2.6. The natural fuzzy negation of a fuzzy implication $I : [0,1]^2 \to [0,1]$ is defined as the function $N_I : [0,1] \to [0,1]$, such that $N_I(x) = I(x,0)$.

A fuzzy implication I satisfies **(I7)** if and only if

$$N_I = N_\perp = N_\perp(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \in]0,1]. \end{cases} \quad (1)$$

3 R_O-IMPLICATIONS

In [10], we studied the class of fuzzy implications called R_O-implications, where O denotes overlap functions, in the same line of our previous work in [12]. R_O-implications is a sub-class of residual implications derived from fuzzy conjunctions [20, Theorem 2]. Let $O : [0,1]^2 \to [0,1]$ be an overlap function, and define $I_O : [0,1]^2 \to [0,1]$ by $I_O(x,y) = \max\{z \in [0,1] \mid O(x,z) \leq y\}$, for all $x,y \in [0,1]$. From [20, Theorem 1, Theorem 2], it is immediate that:

Corollary 3.1 ([10])**.** *The function I_O is a fuzzy implication, called the residual implication derived from an overlap function O, denoted by R_O-implication. O and I_O form an adjoint pair, that is, they satisfy the residuation property, given by:* $\forall x,y,u \in [0,1] : O(x,u) \leq y \Leftrightarrow I_O(x,y) \geq u$.

For an R_O-implication I_O, the overlap function O is said to be the *generator* of I_O. I_O is also called as "the residuum of O". This class of implications is related to a residuation concept from the intuitionistic logic. An example of R_O-implications is: [10]

$$I_{O^V_{mM}}(x,y) = \begin{cases} \min\left\{1, \max\left\{\frac{\sqrt{2y-1}}{2\sqrt{2x-1}}, \frac{2y-1}{2(2x-1)^2}\right\} + \frac{1}{2}\right\} \\ \qquad\qquad \text{if } x \in]0.5,1], y \in [0.5,1] \\ y \qquad\qquad \text{if } y \in [0,0.5[\text{ and } x > y \\ 1 \qquad\qquad \text{if } x \in [0,0.5] \text{ and } x \leq y. \end{cases}$$

which is generated by the overlap functions O^V_{mM}.

In [10], several properties of I_O-implications were studied, e.g., some of the following:

Proposition 3.1 ([10]). *Let $O : [0,1]^2 \to [0,1]$ be an overlap function. Then it holds that:* ***(i)*** *I_O satisfies* ***(I6)*** *if and only if O satisfies* ***(O6)*** *and* ***(O7)****;* ***(ii)*** *I_O satisfies* ***(I7)****;* ***(iii)*** *if I_O satisfies* ***(I6)*** *then I_O satisfies* ***(I8)****;* ***(iv)*** *I_O satisfies* ***(I9)*** *if and only if O satisfies* ***(O6)****;* ***(v)*** *if I_O satisfies* ***(I6)*** *then I_O satisfies* ***(I10)****;* ***(vi)*** *I_O satisfies* ***(I11)*** *if and only if $O \geq \min$.*

Proof. See [10] for the proofs of **(i)-(v)**. So, it remains to prove **(vi)**: ($\Rightarrow$) If $O \ngeq \min$, then there exist $x, y \in [0,1]$ such that $O(x,y) < \min\{x,y\}$. If $x \geq y$ then one has that $O(x,y) < y$. It holds that $I_O(x,y) = \max\{z \in [0,1] \mid O(x,z) \leq y\} > y$, which is a contradiction with **(I11)**. By the commutativity of O, if $x < y$, then $I_O(y,x) > x$, which is also a contradiction with **(I11)**. ($\Leftarrow$) Consider $O(x,y) \geq \min\{x,y\}$ and suppose that $x > y$. Then, one has that $O(x,y) \geq y$. It follows that for all $z \in [0,1]$ it holds that if $O(x,z) \leq y$ then $z \leq y$, since O is increasing. One concludes that $I_O(x,y) = \max\{z \in [0,1] \mid O(x,z) \leq y\} \leq y$. □

As a consequence, an R_O-implication satisfies the property **(I6)** if and only if O has 1 as a neutral element [10]. In [10], we presented two characterizations of R_O-implications. The first is for R_O-implications derived from the sub-class of overlap functions having 1 as neutral element, denoted by $\mathcal{O}_x$. The second characterization is concerned with R_O-implications derived from a more general sub-class of overlap functions satisfying the property **(O6)**, but not necessarily having 1 as neutral element, denoted by $\mathcal{O}_{\leq x}$.

Theorem 3.1 ([10]). *Let $I : [0,1]^2 \to [0,1]$ be a right-continuous (in both variables) fuzzy implication and consider the function $C_I : [0,1]^2 \to [0,1]$, defined by:*

$$C_I(x,y) = \min\{z \in [0,1] \mid I(x,z) \geq y\}. \quad (2)$$

(i) *If I satisfies* ***(I6)****,* ***(I7)*** *and* ***(I8)****, then C_I is an overlap function with 1 as neutral element, i.e., $C_I \in \mathcal{O}_x$;*

(ii) *If I satisfies* ***(I7)****,* ***(I9)****, and* ***(I10)*** *and and $\forall y \in [0,1) : I(1,y) < 1$, then C_I is an overlap function satisfying* ***(O6)****, i.e., $C_I \in \mathcal{O}_{\leq x}$.*

Theorem 3.2 ([10]). *Let $I : [0,1]^2 \to [0,1]$ a right-continuous fuzzy implication. Then:*

$\mathcal{O}_x$-Characterization of R_O-implications: *I satisfies the properties* ***(I6)****,* ***(I7)*** *and* ***(I8)*** *if and only if I is an R_O-implication derived from an overlap function $O \in \mathcal{O}_x$, i.e., $I = I_O$;*

$\mathcal{O}_{\leq x}$-Characterization of R_O-implications: *I satisfies the properties* ***(I7)****,* ***(I9)****,* ***(I10)*** *and $\forall y \in [0,1) : I(1,y) < 1$ if and only if I is an R_O-implication derived from an overlap function $O \in \mathcal{O}_{\leq x}$, i.e., $I = I_O$.*

The derivation of an R_O-implication by an overlap function $O \in \mathcal{O}_x$ ($O \in \mathcal{O}_{\leq x}$) is unique, that is $I_O = I_{O'} \Leftrightarrow O = O'$. [10]

Proposition 3.2. *Let $O_1, O_2 : [0,1]^2 \to [0,1]$ be overlap functions. $O_1 \leq O_2$ if and only if $I_{O_2} \leq I_{O_1}$.*

Proof. $O_1 \leq O_2 \Leftrightarrow \forall x,y,z \in [0,1] : O_2(x,z) \leq y \to O_1(x,z) \leq y \Leftrightarrow \forall x,y \in [0,1]{:}\{z{\in}[0,1] | O_2(x,z){\leq}y\} \subseteq \{z{\in}[0,1] | O_1(x,z){\leq}y\} \Leftrightarrow I_{O_2}{\leq}I_{O_1}$. □

4 FUZZY IMPLICATIONS AND THE O-CONDITIONALITY

Several axiomatizations of R-implications can be found in the literature [19]. The first set of axioms was introduced by Pedrycz [23, 24], among them the law of conditionality, given by $T(x, x \to_T y) \leq y$, where T is a t-norm and $\to_T$ is the residuum of T. In this section, we study the the law of conditionality for any fuzzy implication $I : [0,1]^2 \to [0,1]$ and overlap function $O : [0,1]^2 \to [0,1]$, which we call the law of O-conditionality.

Definition 4.1. A fuzzy implication I satisfies the law of O-conditionality for an overlap function O if and only if, for all $x, y \in [0,1]$, it holds that $O(x, I(x,y)) \leq y$. **(OC)**

In the following, we state under which conditions of I and O we have that I satisfies **(OC)** for O.

Proposition 4.1. *If a fuzzy implication I satisfies* **(I11)** *and, for all $x, y \in [0,1] : x \leq y \Rightarrow I(x,y) = 1$* **(I6$\Leftarrow$)***, then I satisfies* **(OC)** *for any overlap function $O : [0,1]^2 \to [0,1]$ satisfying* **(O6)**.

Proof. If I satisfies **(I6$\Leftarrow$)**, then, for all overlap function O satisfying **(O6)**, whenever $x \leq y$, one has that $O(x, I(x,y)) = O(x,1) \leq x \leq y$. On the other hand, since I satisfies **(I11)**, whenever $x > y$ then, for all overlap function O satisfying **(O6)**, it follows that $O(x, I(x,y)) \leq O(x,y) = O(y,x) \leq O(y,1) \leq y$. □

Proposition 4.2. *Considering a fuzzy implication I, if there exist $x, y \in [0,1]$ such that $x > y$ and $I(x,y) = 1$, then I does not satisfy* **(OC)** *for any overlap functions $O : [0,1]^2 \to [0,1]$ satisfying* **(O7)**.

Proof. Suppose that there exist $x_0, y_0 \in [0,1]$ such that $x_0 > y_0$ and $I(x_0,y_0) = 1$. Then, whenever O satisfies **(O7)**, it holds that $O(x_0, I(x_0,y_0)){=}O(x_0,1){\geq}x_0{>}y_0$. □

Proposition 4.3. *If a fuzzy implication I satisfies* **(OC)** *for some overlap function $O : [0,1]^2 \to [0,1]$, then $N_I = N_\perp$.*

Proof. If $x > 0$, then $O(x, N_I(x)) = O(x, I(x, 0)) \leq 0$, by **(OC)**. Since $x > 0$, then, by **(O2)**, it holds that $N_I(x) = 0$. Since $N_I(0) = 1$, one concludes that $N_I = N_\perp$. □

Corollary 4.1. *If a fuzzy implication I satisfies* **(OC)** *for some overlap function $O : [0,1]^2 \to [0,1]$, then $O(x, N_I(x)) = 0$.*

Proof. If $x = 0$, then it is immediate that $O(0, N_I(0)) = 0$. On the other hand, if $x > 0$, by Proposition 4.3, it holds that $O(x, N_I(x)) = O(x, N_\perp(x)) = O(x, 0) = 0$. □

Proposition 4.4. *If a fuzzy implication I satisfies* **(OC)** *for some overlap function $O_1 : [0,1]^2 \to [0,1]$, then I satisfies* **(OC)** *for any overlap function $O_2 : [0,1]^2 \to [0,1]$ such that $O_2 \leq O_1$.*

Proof. One has that $O_2(x, I(x,y)) \leq O_1(x, I(x,y)) \leq y$. □

Proposition 4.5. *If a fuzzy implication $I_1 : [0,1]^2 \to [0,1]$ satisfies* **(OC)** *for some overlap function $O : [0,1]^2 \to [0,1]$, then any fuzzy implication $I_2 : [0,1]^2 \to [0,1]$ such that $I_2 \leq I_1$ satisfies* **(OC)** *for O.*

Proof. One has that $O(x, I_2(x,y)) \leq O(x, I_1(x,y)) \leq y$. □

Proposition 4.6. *If the fuzzy implications $I_1 : [0,1]^2 \to [0,1]$ and $I_2 : [0,1]^2 \to [0,1]$ satisfy* **(OC)** *for the same overlap function $O : [0,1]^2 \to [0,1]$, then the function $(I_1 \vee I_2) : [0,1]^2 \to [0,1]$, defined by*

$$(I_1 \vee I_2)(x,y) = \max\{I_1(x,y), I_2(x,y)\},$$

is a fuzzy implication satisfying **(OC)** *for O.*

Proof. Observe that $(I_1 \vee I_2)$ is particular case of an aggregation of fuzzy implications, and then it is also a fuzzy implication. If $(I_1 \vee I_2)(x,y) = I_1(x,y)$, then it holds that $O(x, (I_1 \vee I_2)(x,y)) = O(x, I_1(x,y)) \leq y$. The proof for $(I_1 \vee I_2)(x,y) = I_2(x,y)$ is analogous. □

Now, we analyse the action of an automorphism on a fuzzy implications I satisfies ***(OC)*** for some overlap function O.

Proposition 4.7. *Let $\varphi : [0,1] \to [0,1]$ be an automorphism. If a fuzzy implication I satisfies* **(OC)** *for some overlap function $O : [0,1]^2 \to [0,1]$, then $I^\varphi : [0,1]^2 \to [0,1]$ satisfies* **(OC)** *for $O^\varphi : [0,1]^2 \to [0,1]$.*

Proof. Suppose that I satisfies **(OC)** for some overlap function O. Then, it follows that:

$$\begin{aligned} & O^\varphi(x, I^\varphi(x,y)) \\ & \quad = \varphi^{-1}(O(\varphi(x), \varphi(I^\varphi(x,y)))) \\ & \quad = \varphi^{-1}(O(\varphi(x), \varphi(\varphi^{-1}(I(\varphi(x), \varphi(y)))))) \\ & \quad = \varphi^{-1}(O(\varphi(x), I(\varphi(x), \varphi(y)))) \\ & \quad \leq \varphi^{-1}(\varphi(y)) = y. \end{aligned}$$

□

5 R_O-IMPLICATIONS AND O-CONDITIONALITY

In this section, we analyse the law of O-conditionality for R_O-implications. In particular, we state under which conditions of the overlap functions O_1 and O_2, and of the R_O-implication I_{O_2}, we have that I_{O_2} satisfies or not the law of O-conditionality for O_1 and/or O_2.

Proposition 5.1. *Let $O_1 : [0,1]^2 \to [0,1]$ be an overlap function satisfying* **(O6)** *and $O_2 : [0,1]^2 \to [0,1]$ be an overlap function such that $O_2 \geq \min$. Then, the R_O-implication $I_{O_2} : [0,1]^2 \to [0,1]$ satisfies* **(OC)** *for O_1.*

Proof. If $O_2 \geq \min$, then, by Proposition 3.1 **(vi)**, it holds that I_{O_2} satisfies **(I11)**, that is, if $x > y$ then $I_{O_2}(x,y) \leq y$. On the other hand, by Lemma 2.1, it holds that $O_1(x,y) \leq \min\{x,y\} \leq y$, and then, since O_1 is increasing, one has that $O_1(x, I_{O_2}(x,y)) \leq O_1(x,y) \leq y$. Now, if $x \leq y$, since O_1 satisfies **(O6)** and is increasing, then, $O_1(x, I_{O_2}(x,y)) \leq O_1(x,1) \leq x \leq y$. □

Corollary 5.1. *Let $O_2 : [0,1]^2 \to [0,1]$ be an overlap function such that $O_2 \geq \min$. Then $I_{O_2} : [0,1]^2 \to [0,1]$ satisfies* **(OC)** *for any t-norm $T : [0,1]^2 \to [0,1]$.*

Proposition 5.2. *Let $O_1, O_2 : [0,1]^2 \to [0,1]$ be overlap functions. If there exists $x \in [0,1]$ such that $O_2(x,1) < x$ and $O_1(x,1) > x$, then the R_O-implication $I_{O_2} : [0,1]^2 \to [0,1]$ does not satisfy* **(OC)** *for O_1.*

Proof. Consider that I_{O_2} satisfies **(OC)** for O_1. Suppose that there exists $x' \in [0,1]$ such that $O_2(x',1) < x'$ and $O_1(x',1) > x'$. It follows that $O_1(x', I_{O_2}(x',x')) = O_1(x', \sup\{z \in [0,1] \mid O_2(x',z) \leq x'\}) = O_1(x',1) > x'$. Thus, I_{O_2} does not satisfy ***(OC)*** for O_1. □

Theorem 5.1. *The R_O-implication $I_O : [0,1]^2 \to [0,1]$, derived from the overlap function $O : [0,1]^2 \to [0,1]$, satisfies* ***(OC)*** *for O.*

Proof. By Corollary 3.1, O and I_O form an adjoint pair, that is, they satisfy the residuation property: $\forall x, y, u \in [0,1] : O(x,u) \leq y \Leftrightarrow I_O(x,y) \geq u$. Consider $u = I_O(x,y)$. Then, it follows that $O(x, I_O(x,y)) \leq y$. □

Proposition 5.3. *For any overlap function $O : [0,1]^2 \to [0,1]$ satisfying **(O6)**, $I_{\min}$ satisfies **(OC)** for O.*

Proof. Suppose that $x \leq y$. Then, it holds that $O(x, I_{\min}(x,y)) = O(x,1) \leq x \leq y$. On the other hand, if $x > y$, then, by Lemma 2.1, one has that $O(x, I_{\min}(x,y)) = O(x,y) \leq \min\{x,y\} = y$. □

Corollary 5.2. *$I_{\min}$ satisfies **(OC)** for any t-norm $T : [0,1]^2 \to [0,1]$.*

Corollary 5.3. *Let $O_1, O_2 : [0,1]^2 \to [0,1]$ be overlap functions. If $O_1 \leq O_2$ then I_{O_2} satisfies **(OC)** for O_1.*

Proof. From Theorem 5.1, it holds that I_{O_1} satisfies **(OC)** for O_1. On the other hand, by Proposition 3.2, one has that if $O_1 \leq O_2$ then $I_{O_2} \leq I_{O_1}$. It follows that $O_1(x, I_{O_2}(x,y)) \leq O_1(x, I_{O_1}(x,y)) \leq y$. □

Corollary 5.4. *Let $O_1, O_2 : [0,1]^2 \to [0,1]$ be overlap functions. If $I_{O_2} \leq I_{O_1}$ then I_{O_2} satisfies **(OC)** for O_1.*

Proof. By Proposition 3.2, one has that if $I_{O_2} \leq I_{O_1}$ then $O_1 \leq O_2$. By Corollary 5.3, I_{O_2} satisfies **(OC)** for O_1. □

Proposition 5.4. *Let $I_1, I_2 : [0,1]^2 \to [0,1]$ be fuzzy implications. Whenever O_{I_2} is an overlap function, if $I_1 \leq I_2$ then I_1 satisfies* **(OC)** *for O_{I_2}.*

Proof. By Eq. (2), if $I_1 \leq I_2$, then $O_{I_2}(x, I_1(x,y)) = \inf\{z \in [0,1] \mid I_2(x,z) \geq I_1(x,y)\} \leq y$. □

6 FINAL REMARKS

Overlap function is a special kind of not necessarily associative bivariate aggregation operator used, in general, in applications involving the overlap problem and/or when the associativity property is not strongly required, as in imaging processing and decision making based on fuzzy preference relations, respectively. In those applications, there is no need the use of t-norms as the combination operator.

On the other hand, when considering fuzzy implications, there are some properties that may be not demanded for certain applications. Then, several definitions of fuzzy implications, based on weak operators, have been introduced in the literature. This is the case of the R_O-implications, which we introduced in [10]. The present paper presented an analysis of the O-Conditionality for fuzzy implications and, in particular, for R_O-implications. The awareness of the properties and restrictions can help the conscious use of overlaps functions and R_O-implications in the development of applications, allowing the use of more flexible operators.

Future theoretical work is concerned with the investigation of other kinds of fuzzy implications based on overlap functions, also in the interval-valued setting, as in [5, 8, 13]. We are also aiming at applications in the context of hybrid BDI-fuzzy [17][3] agent models, commonly used in social simulation [1, 21], where the evaluation of social values and exchanges are of a qualitative, subjective, vague nature [14, 15, 25]. Overlap functions can be used for dealing with indifference and incomparability when reasoning on the agent's fuzzy belief base, where a kind of weak preference relation may be defined. R_O-implications can be used for performing inferences, decision making and in the fuzzy control of the agents's intentions.

Acknowledgements

Supported by CNPq (Proc. No. 481283/2013-7, 306970/2013-9, 307681/2012-2).

References

[1] D. Adamatti, G. P. Dimuro, H. Coelho (eds.), Interdisciplinary Applications of Agent-Based Social Simulation and Modeling, IGI Global, Hershey, 2014.

[2] M. Baczyński, On the applications of fuzzy implication functions, in: V. E. Balas et al. (eds.), Soft Computing Applications, vol. 195 of Advances in Intel. Systems and Comp., Springer, 2013, pp. 9–10.

[3] M. Baczyński, B. Jayaram, Fuzzy Implications, Springer, Berlin, 2008.

[4] B. C. Bedregal, G. P. Dimuro, H. Bustince, E. Barrenechea, New results on overlap and grouping functions, Information Sciences 249 (2013) 148–170.

[5] B. C. Bedregal, G. P. Dimuro, R. H. S. Reiser, An approach to interval-valued R-implications and automorphisms, in: J. P. Carvalho et al. (eds.), Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference, IFSA/EUSFLAT, 2009, pp. 1–6.

[3] BDI stands for "Beliefs, Desires, Intentions", a particular cognitive agent model introduced in [26].

[6] B. C. Bedregal, G. P. Dimuro, R. H. N. Santiago, R. H. S. Reiser, On interval fuzzy S-implications, Information Sciences 180 (2010) 1373–1389.

[7] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna, Overlap functions, Nonlinear Analysis 72 (2010) 1488–1499.

[8] G. P. Dimuro, On interval fuzzy numbers, in: 2011 Workshop-School on Theoretical Computer Science, WEIT 2011, IEEE, Los Alamitos, 2011, pp. 3–8.

[9] G. P. Dimuro, B. Bedregal, Archimedean overlap functions: The ordinal sum and the cancellation, idempotency and limiting properties, Fuzzy Sets and Systems 252 (2014) 39–54.

[10] G. P. Dimuro, B. Bedregal, On residual implications derived from overlap functions, Information Sciences, 312 (2015) 78–88

[11] G. P. Dimuro, B. Bedregal, H. Bustince, M. J. Asiáin, R. Mesiar, On additive generators of overlap functions, Fuzzy Sets and Systems, DOI: http://dx.doi.org/10.1016/j.fss.2015.02.008

[12] G. P. Dimuro, B. Bedregal, R. H. N. Santiago, On (G, N)-implications derived from grouping functions, Information Sciences 279 (2014) 1–17.

[13] G. P. Dimuro, B. C. Bedregal, R. H. N. Santiago, R. H. S. Reiser, Interval additive generators of interval t-norms and interval t-conorms, Information Sciences 181 (2011) 3898–3916.

[14] G. P. Dimuro, A. C. R. Costa, L. V. Gonalves, A. Hbner, Centralized regulation of social exchanges between personality-based agents, in: P. Noriega et al. (eds.), Coordination, Organizations, Institutions, and Norms in Agent Systems II, vol. 4386 of LNCS, Springer, Berlin, 2007, pp. 338–355.

[15] G. P. Dimuro, A. R. C. Costa, L. V. Gonalves, D. Pereira, Recognizing and learning models of social exchange strategies for the regulation of social interactions in open agent societies, Journal of the Brazilian Computer Society 17 (2011) 143–161.

[16] F. Durante, E. P. Klement, R. Mesiar, C. Sempi, Conjunctors and their residual implicators: Characterizations and construction methods, Mediterranean Journal of Mathematics 4 (2007) 343–356.

[17] G. P. Farias, G. P. Dimuro, A. C. R. Costa, A fuzzy-BDI agent model for exchanges of non-economic services, based on the social exchange theory, in: 2011 Workshop-School on Theoretical Computer Science (WEIT), IEEE, Los Alamitos, 2011, pp. 134–137.

[18] J. C. Fodor, T. Keresztfalvi, Nonstandard conjunctions and implications in fuzzy logic, Intl. Journal of Approximate Reasoning 12 (1995) 69–84.

[19] E. P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publisher, Dordrecht, 2000.

[20] A. Król, Dependencies between fuzzy conjunctions and implications, in: Proceedings of EUSFLAT-LFA 2011, Atlantis Press, Amsterdam, 2011, pp. 230–237.

[21] L. F. K. Macedo, G. P. Dimuro, M. S. Aguiar, H. Coelho, An evolutionary spatial game-based approach for the self-regulation of social exchanges in mas, in: T. Schaub et al. (eds.), Proc. of ECAI 2014–21st European Confer. on Artificial Intelligence, No. 263 in Frontier in Artificial Intelligence and Applications, IOS Press, Netherlands, 2014, pp. 573–578.

[22] M. Mas, M. Monserrat, J. Torrens, A characterization of (U,N), RU, QL and D-implications derived from uninorms satisfying the law of importation, Fuzzy Sets and Systems 161 (2010) 1369–1387.

[23] W. Pedrycz, Fuzzy control and fuzzy systems, Research Studies Press 14, 1993.

[24] W. Pedrycz, On generalized fuzzy relational equations and their applications, Journal of Mathematics and Anal. Appl. 107 (1985) 520–536.

[25] D. Pereira, L. Gonalves, G. P. Dimuro, A. R. C. Costa, Towards the self-regulation of personality-based social exchange processes in multiagent systems, in: Advances in Artificial Intelligence, vol. 5249 of LNCS, Springer, Berlin, 2008, pp. 113–123.

[26] A. S. Rao, M. P. Georgeff, Modeling rational agents within a BDI-architecture, in: Proc. 2nd Intl. Conf. on Principles of Knowledge Representation and Reasoning, Morgan Kaufmann, 1991, pp. 473–484.

[27] R. H. S. Reiser, G. P. Dimuro, B. C. Bedregal, R. H. N. Santiago, Interval valued QL-implications, in: Logic, Language, Information and Comput., No. 4576 in LNCS, Springer, Berlin, 2007, pp. 307–321.

THE DISTRIBUTIVITY BETWEEN SEMI-T-OPERATORS AND UNINORMS

Paweł Drygaś
Faculty of Mathematics
and Natural Sciences,
University of Rzeszów, Poland
paweldrs@ur.edu.pl

Feng Qin
College of Mathematics
and Information Science,
Jiangxi Normal University,
330022 Nanchang, P.R. China
qinfeng923@163.com

Ewa Rak
Faculty of Mathematics
and Natural Sciences,
University of Rzeszów, Poland
ewarak@ur.edu.pl

Summary

Recently, the distributivity equations have been discussed in families of certain operations (e.g., triangular norms, conorms, uninorms and nullnorms). In this paper we present the solutions of distributivity equations between semi-t-operators and uninorms. Since we omit the assumption about the commutativity it seems appropriate to consider left and right distributivity equations separately.

Keywords: Aggregation operator, Distributivity equations, Idempotent operation, Semi-t-operator, Uninorm.

1 INTRODUCTION

The problem of distributivity was posed many years ago (cf. Aczel [1], pp. 318-319). A new topic of study in this area is mainly concerned with the distributivity between triangular norms and triangular conorms ([11], p. 17). Recently, many studies have addressed the solutions of distributivity equations for aggregation functions [3, 13], fuzzy implications [2], uninorms and nullnorms [10, 18, 23, 24], semi-nullnorms and semi-t-operators [8, 9, 27, 29], which are generalizations of triangular norms and conorms.

In this study, our aim is to obtain algebraic structures with weaker assumptions than nullnorms and t-operators fulfilling left or right distributivity equation. The characterization is interesting from a theoretical point of view and also in terms of their applications because they have proved to be useful in several fields, such as fuzzy logic framework [15], expert systems, neural networks [18], fuzzy quantifiers [15] and others e.g. [16]. This research also complement results of our previous study [8, 9].

First, we introduce weak algebraic structures (Section 2). We then recall the distributivity equations (Section 3). Next, we characterize the solutions to distributivity equations from described families (Section 4).

2 ASSOCIATIVE, MONOTONIC BINARY OPERATIONS

We start by giving some basic definitions and facts.

Definition 2.1 ([14])**.** A triangular semi-norm T is an increasing operation $T : [0,1]^2 \to [0,1]$ with neutral element 1.
A triangular semi-conorm S is an increasing operation $S : [0,1]^2 \to [0,1]$ with neutral element 0.
A triangular norm T is a commutative, associative triangular semi-norm.
A triangular conorm S is a commutative, associative triangular semi-conorm.

Definition 2.2 ([4])**.** The operation $V : [0,1]^2 \to [0,1]$ is called a nullnorm if it is commutative, associative, increasing, has a zero element $z \in [0,1]$ and satisfies

$$V(0,x) = x \quad \text{for all } x \le z, \tag{1}$$

$$V(1,x) = x \quad \text{for all } x \ge z. \tag{2}$$

By definition, the case where $z = 0$ leads back to triangular norms, whereas the case where $z = 1$ leads back to triangular conorms (cf. [14]). The next theorem shows that in other cases nullnorm is built from a triangular norm, a triangular conorm and the zero element.

Theorem 2.3 ([4])**.** *Let $z \in (0,1)$. A binary operation V is a nullnorm with zero element z if and only if a triangular norm T and a triangular conorm S exist*

ISBN: 978–83–8012–519–3

such that

$$V(x,y) = \begin{cases} zS\left(\frac{x}{z}, \frac{y}{z}\right) & \text{if } x,y \in [0,z] \\ z + (1-z)T\left(\frac{x-z}{1-z}, \frac{y-z}{1-z}\right) & \text{if } x,y \in [z,1] \\ z & \text{otherwise} \end{cases} . \quad (3)$$

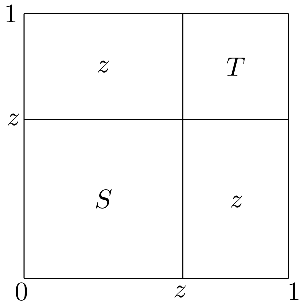


Figure 1: Structure of a nullnorm

If we omit assumptions (1) and (2) from the definition of a nullnorm, it cannot be shown that a commutative, associative, increasing binary operator V with zero element $z = 0$ or $z = 1$ behaves as a triangular norm and triangular conorm (see [6]).

In Definition 2.2, the existence of the zero element z follows from (1) and (2).

Lemma 2.4 (cf. [6])**.** *Let V be an increasing, binary operation and $z \in [0,1]$ exists such that*

$$V(0,x) = V(x,0) = x \quad \text{for all } x \le z, \qquad (4)$$

$$V(1,x) = V(x,1) = x \quad \text{for all } x \ge z, \qquad (5)$$

then $V(x,y) = z$ for $(x,y) \in [0,z]\times[z,1]\cup[z,1]\times[0,z]$, $V|_{[0,z]}$ is an increasing, binary operation with neutral element 0 and zero element z. $V|_{[z,1]}$ is an increasing, binary operation with neutral element 1 and zero element z.
Moreover, V is associative (commutative, idempotent) if and only if $V|_{[0,z]}$ and $V|_{[z,1]}$ are associative (commutative, idempotent).

More general families of operations with zero elements were examined in [19].
As we know, the structure of a nullnorm is the same as the structure of a t-operator.

Definition 2.5 ([17])**.** The operation $F : [0,1]^2 \to [0,1]$ is called a t-operator if it is commutative, associative and increasing such that

$$F(0,0) = 0, \quad F(1,1) = 1 \qquad (6)$$

and the functions F_0 and F_1 are continuous, where $F_0(x) = F(0,x)$, $F_1(x) = F(1,x)$. (7)

In this definition, the existence of the partial neutral elements (conditions (1) and (2)) follows from the continuity of the operation on the boundary of the unit square ((6) and (7)).

If we omit the commutativity condition from the definition of a nullnorm, then we obtain the operation given by (3), where the operations T and S are not necessary commutative.

Definition 2.6 ([7])**.** The operation $V : [0,1]^2 \to [0,1]$ is called a semi-nullnorm if it is associative, increasing, and has a zero element $z \in [0,1]$ that satisfies

$$V(0,x) = V(x,0) = x \quad \text{for all } x \le z, \qquad (8)$$

$$V(1,x) = V(x,1) = x \quad \text{for all } x \ge z. \qquad (9)$$

Theorem 2.7 ([7])**.** *Let $z \in (0,1)$. A binary operation V is a semi-nullnorm with zero element z if and only if V is given by* (3)*, where S is an associative triangular semi-conorm and T is an associative triangular semi-norm. The semi-nullnorm V is idempotent if and only if it is given by* (3) *with $T = \min$ and $S = \max$ (i.e., V is an idempotent nullnorm).*

This is difference with the case of t-operators. Descriptions of the family of this type of operations can be found in [26], [20] and [7].

Definition 2.8 ([7])**.** The operation $F : [0,1]^2 \to [0,1]$ is called a semi-t-operator if it is associative, increasing, and satisfies (6) such that the functions F_0, F_1, F^0, F^1 are continuous, where $F_0(x) = F(0,x)$, $F_1(x) = F(1,x)$, $F^0(x) = F(x,0)$, $F^1(x) = F(x,1)$.

Let $\mathcal{F}_{a,b}$ denote the family of all semi-t-operators such that $F(0,1) = a$, $F(1,0) = b$.

Remark 2.9. Note that in contrast to the definition of triangular semi-norms, in the above definition is assumed associativity. However, we keep the definition of the paper [8], despite the fact that it can be confusing.

Theorem 2.10 ([7])**.** *Let $F : [0,1]^2 \to [0,1]$, $F(0,1) = a$, $F(1,0) = b$. The operation $F \in \mathcal{F}_{a,b}$ if and only if an associative triangular semi-norm T and an associative triangular semi-conorm S exist such that*

$$F(x,y) = \begin{cases} aS\left(\frac{x}{a}, \frac{y}{a}\right) & \text{if } x,y \in [0,a], \\ b + (1-b)T\left(\frac{x-b}{1-b}, \frac{y-b}{1-b}\right) & \text{if } x,y \in [b,1], \\ a & \text{if } x \le a \le y, \\ b & \text{if } y \le b \le x, \\ x & \text{otherwise}, \end{cases} \qquad (10)$$

for $a \leq b$ *and*

$$F(x,y) = \begin{cases} bS\left(\frac{x}{b}, \frac{y}{b}\right) & \text{if } x, y \in [0,b], \\ a + (1-a)T\left(\frac{x-a}{1-a}, \frac{y-a}{1-a}\right) & \text{if } x, y \in [a,1], \\ a & \text{if } x \leq a \leq y, \\ b & \text{if } y \leq b \leq x, \\ y & \text{otherwise}, \end{cases} \tag{11}$$

for $b \leq a$.

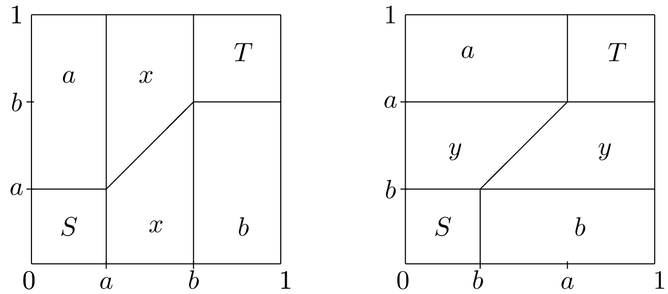


Figure 2: Structure of the operations F from Theorem 2.10 (left (10), right (11))

Definition 2.11 ([28])**.** A binary operation $U : [0,1]^2 \to [0,1]$ is called a uninorm if it is associative, commutative, increasing in each variable and there exists an element $e \in [0,1]$, called the neutral element, such that $U(e,x) = x$ for all $x \in [0,1]$.

The family of all uninorm with neutral element $e \in [0,1]$ will be denoted by $\mathcal{U}_e$

For any uninorm it is satisfied that $U(0,1) \in \{0,1\}$. A uninorm with neutral element $e = 1$ is a t-norm and a uninorm with neutral element $e = 0$ is a t-conorm. For any other value $e \in]0,1[$ the operation works as a t-norm in $[0,e]^2$, as a t-conorm in $[e,1]^2$ and its values are between minimum and maximum in the set of points $A(e)$ given by

$$A(e) = [0,e[\times]e,1] \cup]e,1] \times [0,e[,$$

as it is stated in the following theorem (see [12]).

Theorem 2.12 ([12])**.** *Let U be a uninorm with neutral element $e \in]0,1[$. Then there exists a triangular norm T_U and a triangular conorm S_U such that U is given by*

$$U(x,y) = \begin{cases} eT_U(\frac{x}{e}, \frac{y}{e}) & \text{if } (x,y) \in [0,e]^2, \\ e + (1-e)S_U(\frac{x-e}{1-e}, \frac{y-e}{1-e}) & \text{if } (x,y) \in [e,1]^2, \end{cases}$$

and $\min(x,y) \leq U(x,y) \leq \max(x,y)$ *for all* $(x,y) \in A(e)$.

Next we have the description of special classes of uninorms.

Theorem 2.13. *Let U be a uninorm with neutral element $e \in]0,1[$ and both functions $f(x) = U(x,1)$ and $h(x) = U(x,0)$ ($x \in [0,1]$) are continuous except perhaps at the point $x = e$. Then U is given by one of the following forms.*
(i) If $U(0,1) = 0$, then

$$U(x,y) = \begin{cases} eT_U(\frac{x}{e}, \frac{y}{e}) & \text{if } (x,y) \in [0,e]^2; \\ e + (1-e)S_U(\frac{x-e}{1-e}, \frac{y-e}{1-e}) & \text{if } (x,y) \in [e,1]^2; \\ \min(x,y) & \text{otherwise} \end{cases} \tag{12}$$

(ii) If $U(0,1) = 1$, then

$$U(x,y) = \begin{cases} eT_U(\frac{x}{e}, \frac{y}{e}) & \text{if } (x,y) \in [0,e]^2; \\ e + (1-e)S_U(\frac{x-e}{1-e}, \frac{y-e}{1-e}) & \text{if } (x,y) \in [e,1]^2; \\ \max(x,y) & \text{otherwise}. \end{cases} \tag{13}$$

Denote by $\mathcal{U}_{\min}$ the class of uninorms having form (12) and $\mathcal{U}_{\max}$ the class of uninorms with form (13).

Idempotent uninorms have been characterized by using terminology of Id-symmetrical functions. Let us recall some definitions about this topic, that can be found in [25].

Definition 2.14 ([25])**.** Let $g : [0,1] \to [0,1]$ be any decreasing function and let G be the graph of g, that is

$$G = \{(x, g(x)) \mid x \in [0,1]\}.$$

For any discontinuity point s of g, let us denote by s^- and s^+ the corresponding lateral limits, that are $s^- = \lim_{x \to s^-} g(x)$ and $s^+ = \lim_{x \to s^+} g(x)$.
Then, we define the *completed graph* of g, as the set $F_g = G \cup \{(0,y) \mid y > g(0)\} \cup \{(1,y) \mid y < g(1)\} \cup \{(s,y) \mid s^- \leq y \leq s^+\}$.

Definition 2.15 ([25])**.** A subset F of $[0,1]^2$ is said to be *Id-symmetrical* if for all $(x,y) \in [0,1]^2$ it holds that

$$(x,y) \in F \iff (y,x) \in F\,.$$

The above definition expresses that a subset F of $[0,1]^2$ is symmetrical with respect to the diagonal of the unit square. A similar notion of symmetry is introduced for decreasing functions (see [25]) as follows.

Definition 2.16 ([25])**.** A decreasing function $g : [0,1] \to [0,1]$ is called *Id-symmetrical* if its completed graph F_g is Id-symmetrical.

Theorem 2.17 ([25])**.** *Consider $e \in]0,1[$. U is an idempotent uninorm with neutral element e if and only if there exists a decreasing, Id-symmetrical function $g : [0,1] \to [0,1]$ with fixed point e such that U is given* *by*

$$U(x,y)=\begin{cases}\min(x,y) & \text{if } y<g(x) \text{ or}\\ & y=g(x) \text{ and } x<g(g(x)),\\ \max(x,y) & \text{if } y>g(x) \text{ or}\\ & y=g(x) \text{ and } x>g(g(x)),\\ x \text{ or } y & \text{if } y=g(x) \text{ and } x=g(g(x)),\end{cases} \quad (14)$$

being commutative on the set of points $(x,g(x))$ *such that* $x=g^2(x)$.

3 DISTRIBUTIVITY EQUATIONS

Now, we consider the distributivity equations (cf. [1], p. 318).

Definition 3.1. Let $F,U:[0,1]^2\to[0,1]$. The operation F is distributive over U if the left and right distributivity conditions are fulfilled

$$F(x,U(y,w))=U(F(x,y),F(x,w)), \quad (15)$$

$$F(U(y,w),x)=U(F(y,x),F(w,x)), \quad (16)$$

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

Lemma 3.2 (cf. [21, 22])**.** *Let* $F:X^2\to X$ *have a right (left) neutral element* e *in a subset* $\emptyset\neq Y\subset X$ *(i.e.* $\forall_{x\in Y}\ F(x,e)=x$ $(F(e,x)=x)$*). If operation* F *is left (right) distributive over operation* $U:X^2\to X$ *and satisfying* $U(e,e)=e$*, then* U *is idempotent in* Y.

Proof. Let $x\in Y\subset X$, $y,w=e\in Y\subset X$. If F is distributive over U, then $x=F(x,e)=F(x,U(e,e))=U(F(x,e),F(x,e))=U(x,x)$. The proof is similar in the case where operation F has a left neutral element. □

Corollary 3.3 ([5])**.** *If the operation* $F:[0,1]^2\to[0,1]$ *with neutral element* $e\in[0,1]$ *is distributive over operation* $U:[0,1]^2\to[0,1]$ *and satisfying* $U(e,e)=e$*, then* U *is idempotent.*

Lemma 3.4 ([21])**.** *Every increasing operation* $F:[0,1]^2\to[0,1]$ *is distributive over* max *and* min.

4 DISTRIBUTIVITY OF $F\in\mathcal{F}_{a,b}$ OVER $U\in\mathcal{U}_e$

Our main consideration concerns the distributivity between the semi-t-operator $F\in\mathcal{F}_{a,b}$ and the uninorm $U\in\mathcal{U}_e$.

Now, we start with a case where $a\le b$ and left distributivity (eq. (15)).

Lemma 4.1. *Let* $a,b,e\in[0,1]$*. If* $a\le b$ *and* $F\in\mathcal{F}_{a,b}$ *is left distributive over* $U\in\mathcal{U}_e$*, then* U *is an idempotent uninorm.*

Lemma 4.2. *Let* $a,b,e\in[0,1]$*,* $0=a\le b$*,* $F\in\mathcal{F}_{a,b}$ *and* $U\in\mathcal{U}_e$ *be a disjunctive uninorm. If* F *is left distributive over* U*, then* $e\le b$.

Theorem 4.3. *Let* $a,b,e\in[0,1]$*,* $0=a\le b$*,* $F\in\mathcal{F}_{a,b}$ *and* $U\in\mathcal{U}_e$ *be a disjunctive uninorm.* F *is left distributive over* U *if and only if* $e\le b$ *and* U *is idempotent uninorm satisfying* $U(0,y)=y$ *for all* $y>b$.

Proof. Let $0=a\le b$, $F\in\mathcal{F}_{a,b}$ be left distributive over disjunctive uninorm U. Then by Lemma 4.1 U is idempotent uninorm and by Lemma 4.2 $e\le b$. Putting $x=1$, $y=0$ and $z>b$ in (15) and using fact, that $F(1,0)=b$, $F(1,z)=z$ we have $F(1,U(0,z))=U(F(1,0),F(1,z))=U(b,z)=z$. If $U(0,z)=\min(0,z)=0$, then $z=F(1,U(0,z))=F(1,0)=b$, which is a contradiction. So, $U(0,z)=\max(0,z)=z$ for all $z>b$.
Conversely, let U be disjunctive idempotent uninorm with $U(0,y)=y$ for all $y>b$. Then $U(x,y)=\max(x,y)$ for $x\in[0,e]$, $y>b$. To prove (15) we consider following cases:
1. If $x\le b$, then $F(x,U(y,z))=x=U(x,x)=U(F(x,y),F(x,z))$.
2. Let now $x>b$
a) If $y<b$ and $z<b$, then $U(y,z)<b$ and $F(x,U(y,z))=b=U(b,b)=U(F(x,y),F(x,z))$.
b) If $y<b$ and $z>b$, then $F(x,U(y,z))=F(x,\max(y,z))=\max(F(x,y),F(x,z))=U(F(x,y),F(x,z))$.
c) If $y>b$, then distributivity we obtain, as in case 2b).
d) If $y=b$ or $z=b$, then the proof is similar as in the case 2a) or 2b).
So, F is left distributive over U. □

Lemma 4.4. *Let* $a,b,e\in[0,1]$*,* $0<a\le b$*,* $F\in\mathcal{F}_{a,b}$ *and* $U\in\mathcal{U}_e$ *be a disjunctive uninorm. If* F *is left distributive over* U*, then* $e<a$ *and* F *has the following form*

$$F(x,y)=\begin{cases}S_1(x,y) & if\ x,y\in[0,e],\\ S_2(x,y) & if\ x,y\in[e,a],\\ A(x,y) & if\ x\in[0,e],\ y\in[e,a],\\ T(x,y) & if\ x,y\in[b,1],\\ a, & if\ x\le a\le y,\\ b & if\ y\le b\le x,\\ x & if\ a\le x\le b,\\ \max(x,y) & otherwise,\end{cases} \quad (17)$$

where T *is isomorphic with an associative triangular semi-norm,* S_1 *is isomorphic with an associative triangular semi-conorm,* e *is a right neutral element of* $S_2:[e,a]^2\to[e,a]$*,* 0 *is a left neutral element of* $A:[0,e]\times[e,a]\to[e,a]$ *and where* A*,* T*,* S_1*,* S_2 *are increasing operations having common boundary values.*

Lemma 4.5. *Let $a, b, e \in [0,1]$, $0 < a \le b$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. F is left distributive over U, then $g(x) = e$ for $x \in [0,e]$ and $g(x) = 0$ for $x \in]e,1]$, i.e. $U \in \mathcal{U}_{\max}$.*

Theorem 4.6. *Let $a, b, e \in [0,1]$, $0 < a \le b$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. F is left distributive over U if and only if U is an idempotent uninorm from the class $\mathcal{U}_{\max}$, $e < a$ and F is given by (17).*

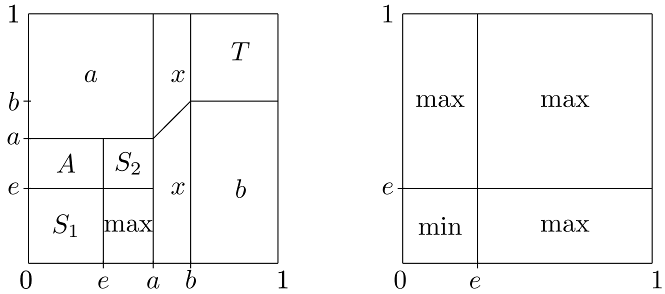


Figure 3: Structure of the operations F and U from Theorem 4.6

Similar results we obtain, if we consider the right distributivity

Lemma 4.7. *Let $a, b, e \in [0,1]$. If $b \le a$ and $F \in \mathcal{F}_{a,b}$ is right distributive over $U \in \mathcal{U}_e$, then U is an idempotent uninorm.*

Lemma 4.8. *Let $a, b, e \in [0,1]$, $0 = b \le a$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. If F is right distributive over U, then $e \le b$.*

Theorem 4.9. *Let $a, b, e \in [0,1]$, $0 = b \le a$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. F is right distributive over U if and only if $e \le a$ and U is an idempotent uninorm satisfying $U(0,y) = y$ for all $y > a$.*

Lemma 4.10. *Let $a, b, e \in [0,1]$, $0 < b \le a$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. If F is right distributive over U, then $e < b$ and F has the following form*

$$F(x,y) = \begin{cases} S_1(x,y) & if\ x,y \in [0,e], \\ S_2(x,y) & if\ x,y \in [e,a], \\ B(x,y) & if\ x \in [e,a],\ y \in [0,e], \\ T(x,y) & if\ x,y \in [b,1], \\ a, & if\ x \le a \le y, \\ b & if\ y \le b \le x, \\ y & if\ a \le y \le b, \\ \max(x,y) & otherwise, \end{cases} \quad (18)$$

where T is isomorphic with an associative triangular semi-norm, S_1 is isomorphic with an associative triangular semi-conorm, e is a right neutral element of $S_2 : [e,a]^2 \to [e,a]$, 0 is a left neutral element of $B : [e,a] \times [0,e] \to [e,a]$ and where B, T, S_1, S_2 are increasing operations having common boundary values.

Lemma 4.11. *Let $a, b, e \in [0,1]$, $0 < b \le a$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. F is right distributive over U, then $g(x) = e$ for $x \in [0,e]$ and $g(x) = 0$ for $x \in]e,1]$ i.e. $U \in \mathcal{U}_{\max}$.*

Theorem 4.12. *Let $a, b, e \in [0,1]$, $0 < b \le a$, $F \in \mathcal{F}_{a,b}$ and $U \in \mathcal{U}_e$ be a disjunctive uninorm. F is right distributive over U if and only if U is an idempotent uninorm from the class $\mathcal{U}_{\max}$, $e < b$ and F is given by (18).*

Remark 4.13. Similar results can be obtained if we consider conjunctive uninorm.

5 CONCLUSIONS

The results of the left distributivity look very similar to the right distributivity, but the left distributivity of F and U is considered when $a < b$, a right distributivity, where $b > a$. Furthermore, in the case where $a < b$ the operation F has a right neutral element in the subintervals, which add up to the unit interval, so that the left distributivity gives idempotency of uninorm (see. Lemma 3.2). The other hand, the left neutral element we obtain only on a subset of the unit interval, which gives only partial results. The next step of our work will be to complete the remaining cases, and add some counterexamples.

Acknowledgements

This study has been partially supported by research fellowship within "UR - modernity and future of region" funded from European Social Fund, Human Capital, national Cohesion Strategy (contract no. UDA-POKL.04.01.01-00-068/10-00).

References

[1] J. Aczél (1966). Lectures on Functional Equations and their Applications. *Acad. Press.* New York.

[2] M. Baczyński (2001). On a class of distributive fuzzy implications, *Internat. J. Uncertainty, Fuzzines Knowledge-Based Syst.* **9**, pp 229-238.

[3] T. Calvo (1999). On some solutions of the distributivity equation, *Fuzzy Sets and Systems* **104** pp 85-96.

[4] T. Calvo, B. De Baets, J. Fodor (2001). The functional equations of Frank and Alsina for uninorms and nullnorms, *Fuzzy Sets and Systems* **120**, pp 385-394.

[5] J. Drewniak, P. Drygaś, E. Rak (2008). Distributivity between uninorms and nullnorms, *Fuzzy Sets and Systems* **159**, pp 1646-1657.

[6] P. Drygaś (2004). A characterization of idempotent nullnorms, *Fuzzy Sets and Systems* **145**, pp 455-461.

[7] P. Drygaś (2015). Semi nullnorms and semi t-operators, *submitted.*

[8] P. Drygaś (2015). Distributivity between semi-t-operators and semi-nullnorms. *Fuzzy Sets and Systems* **264**, pp 100-109.

[9] P. Drygaś, E. Rak (2015). Distributivity equation in the class of semi-t-operators. *Fuzzy Sets and Systems*, http://dx.doi.org/10.1016/j.fss.2015.01.015.

[10] Q. Feng, Z. Bin (2005). The distributive equations for idempotent uninorms and nullnorms. *Fuzzy Sets and Systems* **155**, pp 446-458.

[11] J. Fodor, M. Roubens (1994), Fuzzy Preference Modeling and Multicriteria Decision Support, Kluwer Acad. Publ., New York.

[12] J.C. Fodor, R.R. Yager, A. Rybalov (1997). Structure of uninorms. *Internat. J. Uncertainty, Fuzzines Knowledge-Based Syst.* **5**, pp 411-427.

[13] D. Jočić, I. Štainer-Papuga (2013). Restricted distributivity for aggregation operators with absorbing element. *Fuzzy Sets and Systems* **224**, pp 23-35.

[14] E.P. Klement, R. Mesiar, E. Pap (2000), Triangular Norms. Kluwer Acad. Publ., Dordrecht.

[15] G.J. Klir, B. Yuan (1995). Fuzzy Sets and Fuzzy Logic, Theory and Application, Prentice Hall PTR, Upper Saddle River, New Jersey.

[16] A. Lundberg (2005). Variants of the distributivity equation arising in theories of utility and psychophysics. *Aequationes Math.* **69**, pp 128-145.

[17] M. Mas, G. Mayor, J. Torrens (1999), t-operators, *Internat. J. Uncertainty, Fuzzines Knowledge-Based Syst.* **7**, pp 31-50.

[18] M. Mas, G. Mayor, J. Torrens (2002). The distributivity condition for uninorms and t-operators. *Fuzzy Sets and Systems* **128**, pp 209-225.

[19] M. Mas, R. Mesiar, M. Monserat, J. Torrens (2005). Aggregation operations with annihilator. *Internat. J. Gen. Syst.* **34**, pp 1-22.

[20] J.-L. Marichal (2000), On the associativity functional equation, *Fuzzy Sets and Systems* **114**, pp 381-389.

[21] E. Rak (2005). Distributivity equation for nullnorms. *J. Electrical Engin.* **56, 12/s**, pp 53-55.

[22] E. Rak, P. Drygaś (2006), Distributivity between uninorms, *J. Electrical Engin.* **57, 7/s**, pp 35-38.

[23] D. Ruiz, J. Torrens (2003), Distributive idempotent uninorms, *Internat. J. Uncertainty, Fuzzines Knowledge-Based Syst.* **11**, pp 413-428.

[24] D. Ruiz, J. Torrens (2006). Distributivity and conditional distributivity of a uninorm and a continuous t-conorm. *IEEE Transactions on Fuzzy Systems* **14 (2)**, pp 180-190.

[25] D. Ruiz-Aguilera, J. Torrens, B. De Baets, J. Fodor (2010), Some remarks on the characterization of idempotent uninorms, in: E. Hllermeier, R. Kruse, F. Hoffmann (Eds.), *Proceedings of the 13th IPMU Conference 2010, in: Lect. Notes Comput. Sci.*, vol. 6178, Springer-Verlag, pp. 425434.

[26] W. Sander (2002), Associative aggregation operators, in: T. Calvo, G. Mayor, R. Mesiar (Eds.), Aggregation Operators, Physica-Verlag, Heidelberg, pp 124-158.

[27] Y. Su, W. Zong, H.-W. Liu, P. Xue (2015), The distributive equations for semi t-operators over uninorms, *Fuzzy Sets and Systems* submitted.

[28] R.R. Yager, A. Rybalov (1996). Uninorm aggregation operators. *Fuzzy Sets and Systems* **80**, pp 111-120.

[29] L. Yu, F. Qin (2015). Distributivity between semi-t-operators, submitted.

DISTRIBUTIVITY OF AGGREGATION OPERATORS WITH 2-NEUTRAL ELEMENTS

Paweł Drygaś and Ewa Rak
Faculty of Mathematics
and Natural Sciences,
University of Rzeszów, Poland
{paweldrs,ewarak}@ur.edu.pl

Lemnaouar Zedam
Faculty of Mathematics and Informatics,
Med Boudiaf University- Msila, Algeria
l.zedam@gmail.com

Summary

The problem of distributivity is of a great interest both for particular as well as fundamental reasons. It relates for instance to the theory of binary operators like triangular norms, conorms and their generalizations, i.e. uninorms and nullnorms. One of the recent generalizations covering both uninorms and nullnorms are 2-uninorms, which form a class of commutative, associative and increasing operators on the unit interval with an absorbing element separating two subintervals, having their own neutral elements. This paper is precisely devoted to solve the distributivity equations of such binary aggregation operators with 2-neutral element without the assumptions of associativity and commutativity. In particular, the solution for one of the possible subclasses of these operators depending on the position of its zero and neutral elements is characterized.

Keywords: Aggregation operator, Distributivity equations, Idempotent operation, 2-Semi-Uninorm.

1 INTRODUCTION

The functional equations involving aggregation operators (e.g. [6], [7], [10], [21], [22]) play an important role in theories of fuzzy sets and fuzzy logic. A new direction of investigations concerns of distributivity equation for uninorms and nullnorms and their generalizations ([5], [11], [12], [14], [15], [18], [22], [24], [25]). These special aggregation operators have been proven to be useful in many fields such as fuzzy logic, expert systems, neural networks, aggregation, utility theory and fuzzy system modeling (see [13], [17], [19], [20], [27]).

Our consideration was motivated by logical connectives which generalize the concept of 2-uninorms introduced by P. Akella in [2]. A 2-semi-uninorm belongs to the class of binary aggregation operators on the unit interval with an absorbing element separating two subintervals having their own neutral elements.

In the case of semi-nullnorms these neutral elements are respectively 0 and 1 while for 2-semi-uninorms they lie anywhere in the subintervals. Hence, in the structure of 2-semi-uninorm we have two operators isomorphic with some semi-uninorms, where in the case of semi-nullnorms these operators are isomorphic with respectively some disjunction operator and conjunction operator.

This paper is organized as follows. In Section 2, we considered the algebraic structures of semi-uninorms and semi-nullnorms. In Section 3, the concept of the class of 2-semi-uninorms and the characterization of one of its possible subclasses is introduced. Then, the functional equations of distributivity is recalled (Section 4). In the last (main) section, the solutions of distributivity equations for described subclass of 2-semi-uninorms is characterized.

2 SEMI-UNINORMS AND SEMI-NULLNORMS

We start with basic definitions and facts.

Definition 2.1 (cf. [9])**.** A (bivariate) aggregation operator is a mapping $F : [0,1]^2 \to [0,1]$ such that
i) $F(0,0) = 0$ and $F(1,1) = 1$;
ii) F is increasing with respect to both variables.

Definition 2.2 (cf. [10])**.** Let $e \in [0,1]$. By N_e we denote the family of all aggregation operators $F : [0,1]^2 \to [0,1]$ with neutral element $e \in [0,1]$ i.e.

$$F(e,x) = F(x,e) = x \quad \text{for all} \quad x \in [0,1].$$

Definition 2.3 (cf. [28])**.** An operator $F \in N_e$ is called uninorm if it is associative and commutative.

ISBN: 978–83–8012–519–3

We use the following notation $D_e = [0,e) \times (e,1] \cup (e,1] \times [0,e)$ for $e \in (0,1)$.

Theorem 2.4 ([10]). *Let $e \in [0,1]$. $F \in N_e$ if and only if*

$$F(x,y) = \begin{cases} A(x,y), & (x,y) \in [0,e]^2 \\ B(x,y), & (x,y) \in [e,1]^2 \\ C(x,y), & (x,y) \in D_e \end{cases},$$

where $A : [0,e]^2 \to [0,e]$, $B : [e,1]^2 \to [e,1]$ are aggregation operators with neutral element e and $C : D_e \to [0,1]$ is increasing and fulfils

$$\min(x,y) \le C(x,y) \le \max(x,y) \quad for \quad (x,y) \in D_e.$$

Corollary 2.5. *Operators A and B from Theorem 2.4 fulfil $0 \le A(x,y) \le \min(x,y)$, $\max(x,y) \le B(x,y) \le 1$.*

Definition 2.6 ([10]). Let $e \in [0,1]$. By $N_e^{\max}$ ($N_e^{\min}$) we denote the family of all operators $F \in N_e$ fulfilling the additional condition:

$$F(0,x) = F(x,0) = x \text{ for all } x \in (e,1]$$

$$(F(1,x) = F(x,1) = x \text{ for all } x \in [0,e)).$$

Remark 2.7. $N_0^{\min} = N_0^{\max} = N_0$ (disjunction operator [4]), $N_1^{\min} = N_1^{\max} = N_1$ (semi-copula cf. [3] or conjunction operator [4].

Theorem 2.8 ([10]). *We have*

$$F \in N_e^{\min} \Leftrightarrow F(x,y) = \begin{cases} A(x,y), & (x,y) \in [0,e]^2 \\ B(x,y), & (x,y) \in [e,1]^2 \\ \min(x,y), & (x,y) \in D_e \end{cases},$$

$$F \in N_e^{\max} \Leftrightarrow F(x,y) = \begin{cases} A(x,y), & (x,y) \in [0,e]^2 \\ B(x,y), & (x,y) \in [e,1]^2 \\ \max(x,y), & (x,y) \in D_e \end{cases},$$

where $A : [0,e]^2 \to [0,e]$, $B : [e,1]^2 \to [e,1]$ are aggregation operators with neutral element e.

Definition 2.9. An element $a \in [0,1]$ is called idempotent element of an operator $F : [0,1]^2 \to [0,1]$ if $F(a,a) = a$. The operator F is called idempotent if all elements from $[0,1]$ are idempotent.

Theorem 2.10 (cf. [8], [26]). *Let $e \in [0,1]$. Operators*

$$U_e^{\min}(x,y) = \begin{cases} \max(x,y), & (x,y) \in [e,1]^2 \\ \min(x,y) & otherwise \end{cases},$$

$$and \tag{1}$$

$$U_e^{\max}(x,y) = \begin{cases} \min(x,y), & (x,y) \in [0,e]^2 \\ \max(x,y) & otherwise \end{cases}$$

are unique idempotent uninorms in $N_e^{\min}$ and $N_e^{\max}$, respectively.

Definition 2.11 ([24]). Let $k \in [0,1]$. By Z_k we denote the family of all aggregation operators $F : [0,1]^2 \to [0,1]$ fulfilling the following conditions:
$Z1)$ $\forall_{x \in [0,k]}\ F(0,x) = F(x,0) = x$,
(neutral element $e = 0$ on $[0,k]$)
$Z2)$ $\forall_{x \in [k,1]}\ F(1,x) = F(x,1) = x$.
(neutral element $e = 1$ on $[k,1]$)

Definition 2.12 ([7]). An operator $F \in Z_k$ is called nullnorm if it is associative and commutative.

Remark 2.13. Any operation $F \in Z_k$ fulfils
$Z3)$ $\forall_{x \in [0,1]}\ F(k,x) = F(x,k) = k$. (zero element k)
More general families of operations with zero (absorbing) element are examined in [23].

Theorem 2.14 ([10]). *Let $k \in [0,1]$, $F : [0,1]^2 \to [0,1]$.*

$$F \in Z_k \Leftrightarrow F(x,y) = \begin{cases} A(x,y), & (x,y) \in [0,k]^2 \\ B(x,y), & (x,y) \in [k,1]^2 \\ k, & (x,y) \in D_k \end{cases},$$

where $A : [0,k]^2 \to [0,k]$ and $B : [k,1]^2 \to [k,1]$ are aggregation operators with respectively, neutral element 0 and 1.

3 2-SEMI-UNINORMS

Now we present the definition and some results about the class of 2-semi-uninorms $\mathbf{N}_{k(e,f)}$.

Definition 3.1. Let $k, e, f \in [0,1]$ and $e \le k \le f$. An aggregation operator $F : [0,1]^2 \to [0,1]$ having 2-neutral elements i.e. fulfilling

$$\begin{gathered} \forall_{x \le k}\ F(e,x) = F(x,e) = x \\ and \\ \forall_{x \ge k}\ F(f,x) = F(x,f) = x \end{gathered} \tag{2}$$

is called a 2-semi-uninorm.
The class of all 2-semi-uninorms we denote by $\mathbf{N}_{k(e,f)}$.

Definition 3.2 (cf. [2]). An operator $F \in \mathbf{N}_{k(e,f)}$ is called 2-uninorm if it is associative and commutative.

Remark 3.3. Any operator $F \in \mathbf{N}_{k(e,f)}$ fulfils the condition

$$\forall_{x \in [e,f]}\ F(x,k) = F(k,x) = k \tag{3}$$

by (2) and monotonicity
$k = F(e,k) \le F(x,k) \le F(f,k) = k$ and
$k = F(k,e) \le F(k,x) \le F(k,f) = k$.

Remark 3.4. • If $0 = k \le f = 1$ then 2-semi-uninorm F is a semi-copula with neutral element $f = 1$.
• If $0 = e \le k = 1$ then 2-semi-uninorm F is a disjunction operation with neutral element $e = 0$.

• If $0 = k \leq f \leq 1$ or $0 \leq e \leq k = 1$ then 2-semi-uninorm is an arbitrary semi-uninorm with neutral element f or e, respectively.
• If $e = 0$ and $f = 1$ then 2-semi-uninorm F is a semi-nullnorm (semi-t-operator) with absorbing element k.

In order to exclude the above subclasses from the class of 2-uninorms we assume that $0 \leq e \leq k \leq f \leq 1$ but $e = k$ and $k = f$ can not occur simultaneously.

Lemma 3.5. *Let $F \in \boldsymbol{N}_{k(e,f)}$ be a 2-uninorm, $0 \leq e \leq k \leq f \leq 1$. Then two mappings $F^1, F^2 \in \boldsymbol{N}_e$ defined by*

$$F^1(x,y) = \frac{F(kx,ky)}{k} \quad \text{for } x,y \in [0,1],$$

$$F^2(x,y) = \frac{F(k+(1-k)x, k+(1-k)y)}{1-k} \quad \text{for } x,y \in [0,1]$$

are semi-uninorms with neutral elements $\frac{e}{k}$ and $\frac{f-k}{1-k}$, respectively.

The proof is quite similar to the proof of Lemma 6 in [2].

Lemma 3.6. *Let $F \in \boldsymbol{N}_{k(e,f)}$ be a 2-semi-uninorm, $0 \leq e \leq k \leq f \leq 1$. Then $F(0,1) \in \{0, k, 1\}$.*

From the above lemma we can set apart one from the possible subclasses of operators in $\mathbf{N}_{k(e,f)}$ based on its zero element $F(0,1) = k$, denoted by $\mathbf{N}^k_{k(e,f)}$.

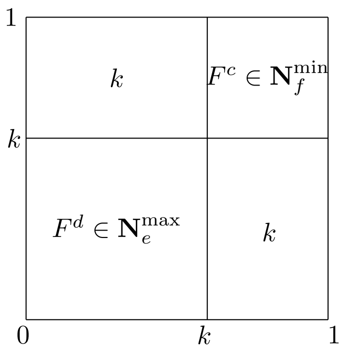


Figure 1: Structure of 2-semi-uninorm $F \in \mathbf{N}^k_{k(e,f)}$.

Representation of 2-semi-uninorm $F \in \mathbf{N}^k_{k(e,f)}$

Theorem 3.7. *Let $F \in \boldsymbol{N}_{k(e,f)}$ be a 2-semi-uninorm, $0 \leq e \leq k \leq f \leq 1$, where $F(x,0)$ and $F(x,1)$ are discontinuous at the points $e \in (0,k]$ and $f \in [k,1)$, respectively. Then $F \in \boldsymbol{N}^k_{k(e,f)}$ if and only if F has the following form (see Fig. 1)*

$$F(x,y) = \begin{cases} F^d(x,y), & (x,y) \in [0,k]^2 \\ F^c(x,y), & (x,y) \in [k,1]^2 \\ k, & (x,y) \in D_k \end{cases} \qquad (4)$$

where F^d and F^c are operators respectively isomorphic with some semi-uninorms from the class $\boldsymbol{N}_e^{\max}$ and $\boldsymbol{N}_e^{\min}$.

4 FUNCTIONAL EQUATIONS OF DISTRIBUTIVITY

We consider here the functional equations of distributivity of two binary aggregation operators. Let us remind some of the most important facts relating to this topic.

Definition 4.1 (cf. [1], p. 318)**.** Let $F, G : [0,1]^2 \to [0,1]$. We say that operator F is left distributive over G, if for all $x, y, z \in [0,1]$

$$F(x, G(y,z)) = G(F(x,y), F(x,z)). \qquad (5)$$

Operator F is right distributive over G, if for all $x, y, z \in [0,1]$

$$F(G(y,z), x) = G(F(y,x), F(z,x)). \qquad (6)$$

If equations (5) and (6) are fulfilled simultaneously (or F is commutative), we say that operation F is distributive over G.

Lemma 4.2 (cf. [24])**.** *Let $\emptyset \neq Y \subset X$, $F : X^2 \to X$, $G : Y^2 \to Y$. If aggregation operator F with neutral element $e \in Y$ is left or right distributive over aggregation operator G fulfilling $G(e,e) = e$, then G is idempotent in Y.*

Lemma 4.3 ([24])**.** *Every aggregation operator $F : [0,1]^2 \to [0,1]$ is left or right distributive over* max *and* min*.*

5 DISTRIBUTIVITY EQUATIONS BETWEEN 2-SEMI-UNINORMS $F, G \in \mathbf{N}^k_{k(e,f)}$

Now we consider the distributivity between operators $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ and $G \in \mathbf{N}^{k_2}_{k_2(e_2,f_2)}$ distinguishing both the order of their zero elements as well as their specific structures.

Theorem 5.1. *Let $k_1, k_2 \in [0,1]$ and $k_2 \leq k_1$. A 2-semi-uninorm $F \in \boldsymbol{N}^{k_1}_{k_1(e_1,f_1)}$ is left distributive over a 2-semi-uninorm $G \in \boldsymbol{N}^{k_2}_{k_2(e_2,f_2)}$ where $0 \leq e_1 \leq e_2 \leq k_2 \leq k_1 \leq f_2 \leq f_1 \leq 1$ if and only if G is idempotent (i.e. $F^c = U^{\min}_{f_2}$, $F^d = U^{\max}_{e_2}$ (cf. (1))) in G) and F is* *given by (see Fig. 2)*

$$F(x,y) = \begin{cases} A_{F^d}(x,y), & (x,y) \in [0,e_1]^2 \\ B^1_{F^d}(x,y), & (x,y) \in [e_1,k_2]^2 \\ B^2_{F^d}(x,y), & (x,y) \in [k_2,k_1]^2 \\ B^3_{F^d}(x,y), & (x,y) \in [e_1,k_2] \times [k_2,k_1] \\ F^c(x,y), & (x,y) \in [k_1,1]^2 \\ k_1, & (x,y) \in D_{k_1} \\ \max(x,y) & otherwise \end{cases}, \tag{7}$$

where $B^1_{F^d} : [e_1,k_2]^2 \to [e_1,k_2]$ *has a neutral element* e_1, $B^2_{F^d} : [k_2,k_1]^2 \to [k_2,k_1]$ *has a right neutral element* k_2, $B^3_{F^d} : [e_1,k_2] \times [k_2,k_1] \to [k_2,k_1]$ *has a left neutral element* e_1 *and* $\{A_{F^d}, B^1_{F^d}, B^2_{F^d}, B^3_{F^d}\} \in \boldsymbol{N}_e$.

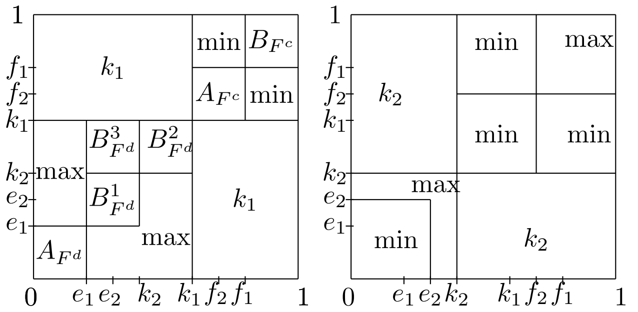


Figure 2: Structures of 2-semi-uninorms from Theorem 5.1 for $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ and $G \in \mathbf{N}^{k_2}_{k_2(e_2,f_2)}$.

Proof. Let $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ be left distributive over $G \in \mathbf{N}^{k_2}_{k_2(e_2,f_2)}$ where $k_2 \le k_1$.

First we prove that $e_1 \le e_2$ and $f_2 \le f_1$. Let us suppose that $0 \le e_2 < e_1$ and $f_1 < f_2 \le 1$. Then from increasingness $G(e_1,e_1) \ge G(e_1,e_2) = e_1 > 0$ and $G(f_1,f_f) \le G(f_1,f_2) = f_1 < 1$. Assuming $G(e_1,e_1) < k_1$, $G(f_1,f_f) > k_1$ and putting first $x = 0$, $y,z = e_1$ and subsequently $x = 1$, $y,z = f_1$ in (5), on account of assumptions of F and G, we obtain

$G(e_1,e_1) = \max(0, G(e_1,e_1)) = F(0, G(e_1,e_1)) = G(F(0,e_1), F(0,e_1)) = G(0,0) = 0$,

$G(f_1,f_1) = \min(1, G(f_1,f_1)) = F(1, G(e_1,e_1)) = G(F(1,e_1), F(1,e_1)) = 1$, which in both cases leads to a contradiction. Hence $0 \le e_1 \le e_2 \le k_2 \le k_1 \le f_2 \le f_1 \le 1$. Now we show that $G(e_1,e_1) = e_1$ and $G(f_1,f_1) = f_1$. Since $e_1 < e_2$ we have $G(e_1,e_1) \le G(e_1,e_2) = e_1$. Taking $y,z = e_1$ in (5), by (2) we get $x = \max(x, G(e_1,e_1)) = F(x, G(e_1,e_1)) = G(F(x,e_1), F(x,e_1)) = G(x,x)$ for $x \in [e_1,k_1]$.

In particular, $G(e_1,e_1) = e_1$.

Now because $f_2 < f_1$ we have $f_1 = G(f_1,f_2) \le G(f_1,f_1)$. Then putting $y,z = f_1$ in (5), by (2) we get $x = \min(x, G(f_1,f_1)) = F(x, G(f_1,f_1)) = G(F(x,f_1), F(x,f_1)) = G(x,x)$ for $x \in [k_1,f_1]$. Thus in particular, $G(f_1,f_1) = f_1$.

Now using twice Lemma 4.2 for $Y_1 = [0,k_2]$, $Y_2 = [k_2,1]$ and Theorem 2.10 we obtain that $F^c = U^{\min}_{f_1}$ and $F^d = U^{\max}_{e_1}$ in G i.e. G is the idempotent 2-semi-uninorm from the subclass $N^{k_2}_{k_2(e_2,f_2)}$.

In the next step of the proof we show that

$$F(x,k_2) = \begin{cases} k_2 & for \;\; x \in [0,k_2] \\ x & for \;\; x \in [k_2,k_1] \end{cases}. \tag{8}$$

First we prove that $F(k_2,k_2) = k_2$. Assuming $x = z = k_2$, $y = e_1$ in (5) we have $F(k_2,k_2) = F(k_2, G(e_1,k_2)) = G(F(k_2,e_1), F(k_2,k_2)) = G(k_2, F(k_2,k_2)) = k_2$. Therefore, directly by monotonicity of F for $x \in [e_1,k_2]$ we get $k_2 = F(e_1,k_2) \le F(x,k_2) \le F(k_2,k_2) = k_2$.

For $x \in [k_2,k_1]$ we have $F(x,k_2) \ge \max(x,k_2) = x$. Simultaneously, assuming in (5) $y = e_1$, $z = k_2$ from (2) and idempotency of G we obtain $F(x,k_2) = F(x, G(e_1,k_2)) = G(F(x,e_1), F(x,k_2)) = \min(x, F(x,k_2))$. Hence $F(x,k_2) \le x$ for $x \in [k_2,k_1]$.

From the above inequalities we get $F(x,k_2) = x$ for $x \in [k_2,k_1]$, which together with $F(x,k_2) = k_2$ for $x \in [0,k_2]$ proves (8).

According to Theorem 3.7 the 2-semi-uninorm F is of the form (4), where the domain of B_{F^d} must be divided to 4 parts set by k_2. On account of increasingness of F, (2), (3) and (8) we consecutively obtain:

$e_1 \le x = F(x,e_1) \le F(x,y) \le F(x,k_2) = k_2$ for $x,y \in [e_1,k_2]$;

$k_2 \le x = F(x,k_2) \le F(x,y) \le F(x,k_1) = k_1$ for $x,y \in [k_2,k_1]$;

$k_2 \le x = F(x,k_2) \le F(x,y) \le F(k_2,k_1) = k_1$ for $x \in [e_1,k_2]$, $y \in [k_2,k_1]$;

$x = F(x,e_1) \le F(x,y) \le F(x,k_2) = x$ for $x \in [k_2,k_1]$, $y \in [e_1,k_2]$, which means $F(x,y) = \max(x,y)$ for $x \in [k_2,k_1]$, $y \in [e_1,k_2]$.

Thus the restrictions $B^1_{F^d} = F|_{[e_1,k_2]^2}$, $B^2_{F^d} = F|_{[k_2,k_1]^2}$ and $B^3_{F^d} = F|_{[k_2,k_1]^2}$ are operations with the desired properties.

Conversely, let $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ be given by (7) and $G \in \mathbf{N}^{k_2}_{k_2(e_2,f_2)}$ be the idempotent 2-semi-uninorm i.e. $F^c = U^{\min}_{f_1}$, $F^d = U^{\max}_{e_1}$ (1) in G and $0 \le e_1 \le e_2 \le k_2 \le k_1 \le f_2 \le f_1 \le 1$.

Let us observe that in 2-semi-uninorm (7) $F\,|_{[0,k_2]^2}$ corresponds to $\mathbf{N}^{\max}_{e_1}$ in $[0,k_2]^2$ and $G\,|_{[0,k_2]^2}$ corresponds to $\mathbf{N}^{\max}_{e_2}$ in $[0,k_2]^2$, which are left distributive in accordance with the Theorem 4 in [25]. Moreover, using the assumption on the common boundary values $F\,|_{[k_2,1]^2}$ is increasing in suitable rectangular domains. Then by Lemma 4.3 we get (5) for $x \in [k_2,1]$ and $(y,z) \in [0,k_2]^2 \cup [k_2,1]^2$. In particular, $F\,|_{[k_1,1]^2}$ corresponding to the $\mathbf{N}^{\min}_{f_1}$ in $[k_1,1]^2$ is left distributive over $G\,|_{[k_1,1]^2}$ corresponding to the $\mathbf{N}^{\min}_{f_2}$ in $[k_1,1]^2$ on the basis of the Theorem 3 in [25].

In the rest parts of $[0,1]^2$ we have $G(y,z) = k_2$ and $L = F(x, G(y,z)) = F(x, k_2)$ is given by (8) or equal to k_1. We consider 12 cases for evaluation of the right side $R = G(F(x,y), F(x,z))$ in (5).

- If $x \le e_1$, $y \le e_1$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = k_2$ and $R = G(F(x,y), \max(x,z)) = G(F(x,y), z) = k_2$, where $F(x,y) \in [0, e_1]$.
- If $x \le e_1$, $y \le e_1$, $z \ge k_1$ then $L = F(x,k_2) = k_2$ and $R = G(F(x,y), k_1) = k_2$, where $F(x,y) \in [0,e_1]$.
- If $x \le e_1$, $e_1 \le y \le k_2$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = k_2$ and $R = G(\max(x,y), \max(x,z)) = G(y,z) = k_2$.
- If $x \le e_1$, $e_1 \le y \le k_2$, $z \ge k_1$ then $L = F(x,k_2) = k_2$ and $R = G(\max(x,y), k_1) = G(y,k_1) = k_2$.
- If $e_1 \le x \le k_2$, $y \le e_1$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = k_2$ and $R = G(\max(x,y), \max(x,z)) = G(x,z) = k_2$.
- If $e_1 \le x \le k_2$, $y \le e_1$, $z \ge k_1$ then $L = F(x,k_2) = k_2$ and $R = G(\max(x,y), k_1) = G(x,k_1) = k_2$.
- If $e_1 \le x \le k_2$, $y \le e_1$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = k_2$ and $R = G(F(x,y), \max(x,z)) = G(F(x,y), z) = k_2$, where $F(x,y) \in [e_1, k_2]$.
- If $e_1 \le x \le k_2$, $y \le e_1$, $z \ge k_1$ then $L = F(x,k_2) = k_2$ and $R = G(F(x,y), k_1) = k_2$, where $F(x,y) \in [e_1, k_2]$.
- If $k_2 \le x \le k_1$, $0 \le y \le k_2$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = x$ and $R = G(\max(x,y), S_2(x,z)) = \min(x, S_2(x,z)) = x$, where $S_2(x,z) \in [k_2, k_1]$.
- If $k_2 \le x \le k_1$, $0 \le y \le k_2$, $z \ge k_1$ then $L = F(x,k_2) = x$ and $R = G(\max(x,y), k_1) = \min(x, k_1) = x$.
- If $k_1 \le x \le 1$, $0 \le y \le k_2$, $k_2 \le z \le k_1$ then $L = F(x,k_2) = k_1$ and $R = G(k_1,k_1) = k_1$.
- If $k_1 \le x \le 1$, $0 \le y \le k_2$, $k_1 \le z \le 1$ then $L = F(x,k_2) = k_1$ and $R = G(k_1, F(x,z)) = \min(k_1, F(x,z)) = k_1$, where $F(x,y) \in [k_1, 1]$.

In all considered cases we get $L = R$, which proves (5). □

Example 5.2. *A 2-semi-uninorm $F \in \boldsymbol{N}^{\frac{3}{8}}_{\frac{3}{8}(\frac{1}{8},\frac{4}{5})}$ given by the formula*

$$F(x,y) = \begin{cases} x + y - xy, & (x,y) \in (\frac{1}{8}, \frac{3}{8}) \times (\frac{1}{4}, \frac{3}{8}) \\ \max(x,y), & (x,y) \in [\frac{1}{8}, \frac{3}{8}] \times [0, \frac{1}{4}] \cup \\ & \cup [0, \frac{1}{8}] \times [\frac{1}{8}, \frac{3}{8}] \\ 1, & (x,y) \in [\frac{4}{5}, 1]^2 \\ \frac{3}{8}, & (x,y) \in D_{\frac{3}{8}} \\ \min(x,y) & \textit{otherwise} \end{cases},$$

is left distributive over idempotent 2-semi-uninorm $G \in \boldsymbol{N}^{\frac{1}{4}}_{\frac{1}{4}(\frac{3}{16},\frac{1}{2})}$.

Next we deal with the case of distributivity between $F \in \mathbf{N}^{k_1}_{(e_1,f_1)}$ and $G \in \mathbf{N}^{k_2}_{(e_2,f_2)}$, where $k_1 < k_2$. Similarly as in Theorem 5.1 one has

Theorem 5.3. *Let $k_1, k_2 \in [0,1]$ and $k_1 < k_2$. A 2-semi-uninorm $F \in \boldsymbol{N}^{k_1}_{k_1(e_1,f_1)}$ is left distributive over a 2-semi-uninorm $G \in \boldsymbol{N}^{k_2}_{k_2(e_2,f_2)}$ where $0 \le e_1 \le e_2 \le k_1 < k_2 \le f_2 \le f_1 \le 1$ if and only if G is idempotent (i.e. $F^c = U^{\min}$, $F^d = U^{\max}$ (1) in G) and F is given by (see Fig. 3)*

$$F(x,y) = \begin{cases} F^d(x,y), & (x,y) \in [0,k_1]^2 \\ A^1_{F^c}(x,y), & (x,y) \in [k_1,k_2]^2 \\ A^2_{F^c}(x,y), & (x,y) \in [k_2,f_1]^2 \\ A^3_{F^c}(x,y), & (x,y) \in [k_1,k_2] \times [k_2,f_1] \\ B_{F^c}(x,y), & (x,y) \in [f_1,1]^2 \\ k_1, & (x,y) \in D_{k_1} \\ \min(x,y) & \textit{otherwise} \end{cases},$$

where $A^1_{F^c} : [k_1,k_2]^2 \to [k_1,k_2]$ has a left neutral element k_2, $A^2_{F^c} : [k_2,f_1]^2 \to [k_2,f_1]$ has a neutral element f_1, $A^3_{F^c} : [k_1,k_2] \times [k_2,f_1] \to [k_2,f_1]$ has a right neutral element k_2 and $\{A^1_{F^c}, A^2_{F^c}, A^3_{F^c}, B_{F^c}\} \in \boldsymbol{N}_e$.

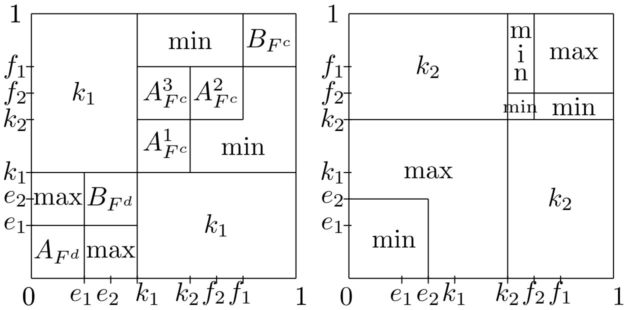


Figure 3: Structures of 2-semi-uninorms from Theorem 5.3 for $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ and $G \in \mathbf{N}^{k_2}_{k_2(e_2,f_2)}$.

Remark 5.4. If we consider the right distributivity equation (6), we obtain similar results to the previous Theorems 5.1 and 5.3 (see Fig.4).

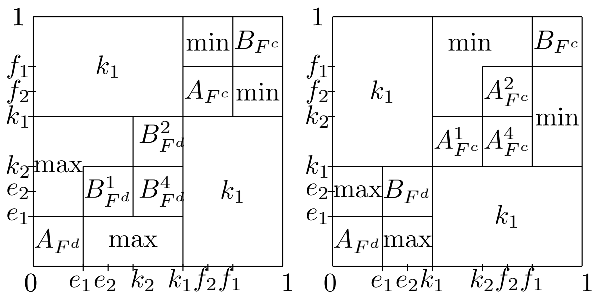


Figure 4: Structures of right distributive 2-semi-uninorms $F \in \mathbf{N}^{k_1}_{k_1(e_1,f_1)}$ (the left part corresponds to the case $k_2 \le k_1$ and the right part corresponds to the case $k_1 < k_2$).

Acknowledgements

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge No RPPK.01.03.00 − 18 − 001/10.

References

[1] J. Aczél (1966). Lectures on Functional Equations and their Applications. *Acad. Press.* New York.

[2] P. Akella (2007). Structure of n-uninorms. *Fuzzy Sets and Systems* **158**, pp 1631-1651.

[3] B. Bassan, F. Spizzichino (2005). Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. *J. Multivariate Anal.* **93**, pp 313-339.

[4] I. Batyrshin, O. Kaynak (1999). Paremetric classes of generalized conjunction and disjunction operations for fuzzy control. *IEEE Transactions on Fuzzy Systems* **7**, pp 586-596.

[5] T. Calvo (1999). On some solutions of the distributivity equation. *Fuzzy Sets and Systems* **104**, pp 85-96.

[6] T. Calvo, B. De Baets (2000). On the generalization of the absorption equation. *J. Fuzzy Math.* **8 (1)**, pp 141-149.

[7] T. Calvo, B. De Baets, J.C. Fodor (2001). The functional equations of Frank and Alsina for uninorms and nullnorms. *Fuzzy Sets and Systems* **120**, pp 385-394.

[8] B. De Baets (1999). Idempotent uninorms. *European J. Operat. Research* **118**, pp 631-642.

[9] J. Dombi (1982). Basic concepts for the theory of evaluation: The aggregative operator. *European J. Operat. Research* **10**, pp 282-293.

[10] J. Drewniak, P. Drygaś, E. Rak (2008). Distributivity equations for uninorms and nullnorms. *Fuzzy Sets and Systems* **159**, pp 1646-1657.

[11] P. Drygaś, E. Rak (2015). Distributivity equation in the class of semi-t-operators. *Fuzzy Sets and Systems*, http://dx.doi.org/10.1016/j.fss.2015.01.015.

[12] P. Drygaś, E. Rak (2015). Distributivity equation in the class of 2-uninorms. *Fuzzy Sets and Systems*, http://dx.doi.org/10.1016/j.fss.2015.02.014.

[13] D. Dubois, H. Prade (2000). Fundamentals of Fuzzy Sets. *Kluwer Acad. Publ.* Boston.

[14] Q. Feng (2004). Uninorm solutions and (or) nullnorm solutions to the modularity condition equations. *Fuzzy Sets and Systems* **148**, pp 231-242.

[15] Q. Feng, Z. Bin (2005). The distributive equations for idempotent uninorms and nullnorms. *Fuzzy Sets and Systems* **155**, pp 446-458.

[16] J. C. Fodor, R. R. Yager, A. Rybalov (1997). Structure of uninorms. *Internat. J. Uncertainty, Fuzzines Knowledge-Based Syst.* **5**, pp 411-427.

[17] D. Gabbay, G. Metcalfe (2007). Fuzzy logics based on [0, 1)-continuous uninorms. *Arch. Math. Logic* **46**, pp 425-449.

[18] D. Jočić, I. Štainer-Papuga (2013). Restricted distributivity for aggregation operators with absorbing element. *Fuzzy Sets and Systems* **224**, pp 23-35.

[19] G. J. Klir, B. Yuan (1995). Fuzzy Sets and Fuzzy Logic, Theory and Application. *Prentice Hall PTR, Upper Saddle River.* New Jersey.

[20] A. Lundberg (2005). Variants of the distributivity equation arising in theories of utility and psychophysics. *Aequationes Math.* **69**, pp 128-145.

[21] M. Mas, G. Mayor, J. Torrens (2002). The distributivity condition for uninorms and t-operators. *Fuzzy Sets and Systems* **128**, pp 209-225.

[22] M. Mas, G. Mayor, J. Torrens (2002). The modularity condition for uninorms and t-operators. *Fuzzy Sets and Systems* **126**, pp 207-218.

[23] M. Mas, R. Mesiar, M. Monserat, J. Torrens (2005). Aggregation operations with annihilator. *Internat. J. Gen. Syst.* **34**, pp 1-22.

[24] E. Rak (2005). Distributivity equation for nullnorms. *J. Electrical Engin.* **56, 12/s**, pp 53-55.

[25] E. Rak, P. Drygaś (2006). Distributivity equation between uninorms. *J. Electrical Engin.* **57, 7/s**, pp 35-38.

[26] W. Sander (2002). Associative aggregation operators.*in: T. Calvo, G. Mayor, R. Mesiar (Eds.), Aggregation Operators. Physica-Verlag.* Heidelberg, pp 124-158.

[27] R. R. Yager (2001). Uninorms in fuzzy system modeling. *Fuzzy Sets and Systems* **122**, pp 167-175.

[28] R. R. Yager, A. Rybalov (1996). Uninorm aggregation operators. *Fuzzy Sets and Systems* **80**, pp 111-120.

AN EFFICIENT ALGORITHM FOR GENERAL WEIGHTED AGGREGATION

Jozo Dujmović
Dept. of Computer Science
San Francisco State University
jozo@sfsu.edu

Summary

We propose a weighted aggregation algorithm for creating a general idempotent weighted aggregator of n variables derived from a related symmetric idempotent aggregator of two variables. This computational method, together with interpolative aggregation, can be used for the development of general idempotent logic aggregators that satisfy a variety of conditions necessary for building decision models in the area of weighted compensative logic.

Keywords: weights, aggregation, logic.

1 INTRODUCTION

Idempotent logic aggregators (ILA) are an important class of aggregators used for modeling logic operation in weighted compensative logic (WCL). In the case of adjustable andness/orness ILA support a continuous transition from the pure conjunction to the pure disjunction. In addition to internality and the adjustability of andness/orness, the ILA must support monotonicity, compensativeness, any number of inputs, and weights that describe the importance of inputs [1][2]. If these properties are supported, then ILA can be used for building decision models in WCL [3].

WCL is a seamless generalization of classic Boolean logic applicable everywhere inside the unit hypercube $[0,1]^n$. WCL and Boolean logic are identical in 2^n vertices of the unit hypercube $\{0,1\}^n$.

The fundamental logic aggregator in WCL is the Generalized Conjunction/Disjunction (GCD) [4][5]. In addition to all basic ILA properties, GCD must be capable to support hard and soft partial conjunction (HPC and SPC) and hard and soft partial disjunction (HPD and SPD). In a general case, we would also like to have independently adjustable threshold andness (border between HPC and SPC) and threshold orness (border between HPD and SPD) [3].

The list of GCD properties that are observable in human reasoning includes internality, monotonicity, compensativeness, adjustable andness and orness, weights, any number of inputs, hard and soft properties (supported and unsupported annihilators 0, 1), and adjustability of threshold andness and orness. Therefore, these are necessary mathematical properties requested in many applications, and we face an obvious question: how to make ILA that can support this long list of necessary properties?

Idempotent compensative aggregators support internality and consequently must be implemented as means [6][7]. Unfortunately, there are no means that directly support all the listed ILA properties. Even the basic adjustability of andness/orness is supported only by a few means (e.g. the power means, exponential means, generalized logarithmic means and extended/Stolarsky means [7]). To satisfy other properties we are forced to use interpolative aggregators [8].

Creating general forms of ILA is primarily a computational problem, and it needs a computational solution. If we want to use aggregators that are not weighted, or those that are defined as means of only two variables, then we need techniques for introducing weights and expanding the number of inputs beyond limitations of the original means. Naturally, mathematical literature (e.g. [1][2]) offers various solutions to this problem. Based on the work of Calvo and Mesiar [9] (see also [10]) where importance is interpreted as cardinality, in [1] we can find methods for incorporating weights into symmetric functions by repetition of arguments and a survey or research papers in this area. Methods for expanding the means of two variables (e.g., such as $L(x,y)=(x-y)/(\log x-\log y)$) to n variables are also studied in mathematical literature (e.g. [11][12]).

In WCL applications we need general asymmetric (weighted) GCD aggregators of n variables and computational methods for creating such aggregators. This paper is focused on solving this computational problem using repetition of arguments and binary trees of suitable basic binary aggregators. We propose weighted aggregation algorithms for making ILA from selected basic means, as well as efficient program implementations and the corresponding performance analysis of the proposed computational solution.

ISBN: 978–83–8012–519–3

2 WEIGHTED AGGREGATIONS BASED ON REPETITION OF ARGUMENTS AND BINARY AGGREGATION TREES

Associativity offers a way for transforming binary aggregators that use equal weights to aggregators that use multiple inputs and different weights. Let us use associative ILA based on quasi-arithmetic means (QAM) $a_1x_1\Diamond...\Diamond a_nx_n = f^{-1}(a_1 f(x_1)+...+a_n f(x_n))$, $0<a_i<1$, $i=1,...,n$, $a_1+...+a_n=1$ where $f:[0,1]\rightarrow\mathbb{R}$ is a continuous and strictly monotonic function. Using QAM we can easily verify the following basic properties of the symbolic ILA notation $y=a_1x_1\Diamond...\Diamond a_nx_n$ (note: a_ix_i is not multiplication, but a symbolic notation specifying that the weight $a_i\in]0,1[$ corresponds to the argument $x_i\in[0,1]$ in the aggregation process based on the GCD operator $\Diamond$):

$$a_1x\Diamond...\Diamond a_nx=(a_1+...+a_n)x=x,$$
$$a_1x_1\Diamond a_2x_2\Diamond a_3x_1=(a_1+a_3)x_1\Diamond a_2x_2,$$
$$a_1x_1\Diamond a_2(b_1x_2\Diamond b_2x_3)=a_1x_1\Diamond a_2b_1x_2\Diamond a_2b_2x_3,$$
$$a_1x_1\Diamond a_2(b_1x_1\Diamond b_2x_2)=(a_1+a_2b_1)x_1\Diamond a_2b_2x_2,$$
$$a_1x_1\Diamond a_2x_2\Diamond a_3x_3=$$
$$a_1x_1\Diamond(a_2+a_3)((a_2/(a_2+a_3))x_2\Diamond(a_3/(a_2+a_3))x_3)$$
.

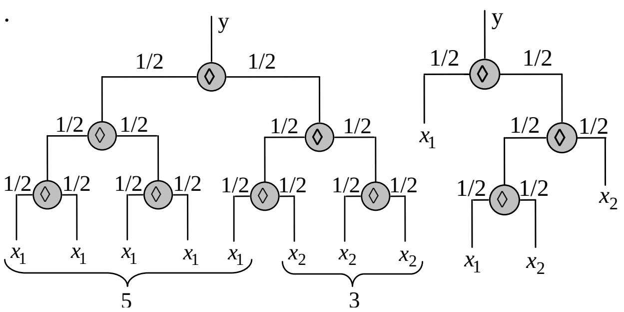
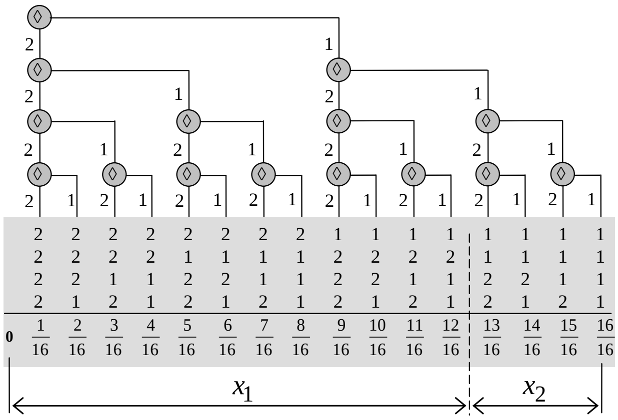


Fig. 1. Creating a weighted aggregator from aggregators that use equal weights

Suppose that we want to transform a symmetric (unweighted) binary aggregator $F(x_1,x_2)=0.5x_1\Diamond 0.5x_2$ that does not have adjustable weights, to its adjustable weighted form $y=Wx_1\Diamond(1-W)x_2$. That can be achieved by using a binary tree of symmetric aggregators $F(x_1,x_2)$ that has L levels and $N=2^L$ input arguments as exemplified in Fig. 1 for L=3, N=8, and W=5/8. To achieve the effect of weight W, the argument x_1 should be repeated K=WN=5 times, and the argument x_2 should be repeated (1-W)N=3 times. If we use idempotency to simplify the tree as shown in Fig. 1 then the resulting aggregator is the following:

$$½x_1\Diamond½(½(½x_1\Diamond½x_2)\Diamond½x_2)=½x_1\Diamond½(¼x_1\Diamond¾x_2)$$
$$=\frac{1}{2}x_1\Diamond\frac{1}{8}x_1\Diamond\frac{3}{8}x_2=\frac{5}{8}x_1\Diamond\frac{3}{8}x_2$$



1/16:	2222	$y=F(F(F(F(x_1,x_2),x_2),x_2),x_2)$
2/16:	1222	$y=F(F(F(F(x_1,x_1),x_2),x_2),x_2)=F(F(F(x_1,x_2),x_2),x_2)$
3/16:	2122	$y=F(F(F(F(x_1,x_2),x_1),x_2),x_2)$
4/16:	1122	$y=F(F(F(F(x_1,x_1),x_1),x_2),x_2)=F(F(x_1,x_2),x_2)$
5/16:	2212	$y=F(F(F(F(x_1,x_2),x_2),x_1),x_2)$
6/16:	1212	$y=F(F(F(F(x_1,x_1),x_2),x_1),x_2)=F(F(F(x_1,x_2),x_1),x_2)$
7/16:	2112	$y=F(F(F(F(x_1,x_2),x_1),x_1),x_2)$
8/16:	1112	$y=F(F(F(F(x_1,x_1),x_1),x_1),x_2)=F(x_1,x_2)$
9/16:	2221	$y=F(F(F(F(x_1,x_2),x_2),x_2),x_1)$
10/16:	1221	$y=F(F(F(F(x_1,x_1),x_2),x_2),x_1)=F(F(F(x_1,x_2),x_2),x_1)$
11/16:	2121	$y=F(F(F(F(x_1,x_2),x_1),x_2),x_1)$
12/16:	1121	$y=F(F(F(F(x_1,x_1),x_1),x_2),x_1)=F(F(x_1,x_2),x_1)$
13/16:	2211	$y=F(F(F(F(x_1,x_2),x_2),x_1),x_1)$
14/16:	1211	$y=F(F(F(F(x_1,x_1),x_2),x_1),x_1)=F(F(F(x_1,x_2),x_1),x_1)$
15/16:	2111	$y=F(F(F(F(x_1,x_2),x_1),x_1),x_1)$

Figure 2. A binary tree with 16 inputs and its weighted aggregation formulas

A general algorithm for creating weighted aggregators from symmetric unweighted aggregators can be developed using an L-level binary tree with N leaves and the technique of repetition of arguments in a way exemplified in Fig 2. We want to aggregate two input suitability degrees (x_1 and x_2) using weights W and 1-W respectively. The desired weighted aggregator of two variables is $y=Wx_1\Diamond(1-W)x_2$. We assume $0<W<1$. The weighted aggregation is achieved by using K inputs with value x_1 followed by N-K inputs with value x_2. The value of K is proportional to W and can be computed using the rounding $\lfloor WN+1/2\rfloor$. The approximate value of W and the corresponding relative error are

$$K=\lfloor WN+1/2\rfloor,\quad K\in\{1,...,N-1\}$$
$$W\approx K/N=\lfloor WN+1/2\rfloor/N,$$
$$E=100(\lfloor WN+1/2\rfloor/N-W)/W\ [\%]$$

We assume that weights cannot be 0 or 1. Using L levels of a binary tree we get $N-1$ discrete weights from the sequence $1/N, 2/N, ..., (N-1)/N$. The distance between two adjacent weights is $1/N$ and after rounding the maximum absolute error is 0.5/N yielding the maximum relative error $E_{\max}=(50/N)/W\,[\%]$. For example, if $W=0.73$ and $L=4$, $N=16$ (as in Fig. 2), then $K=\lfloor 0.73\times 16+1/2 \rfloor=\lfloor 12.18 \rfloor=12$. The obtained weight approximation is $W\approx 12/16=0.75$, the achieved error is 100(0.75-0.73)/0.73=2.74% and the maximum possible error is $E_{\max}=50/(0.73\times 16)=4.28\,\%$. The precision can be increased by using larger values of L. If L=10, then $K=\lfloor 0.73\times 1024+1/2 \rfloor=748$, the obtained weight is W≈748/1024=0.7305, and the errors are E=0.064%, $E_{\max}=0.067\%$. It is possible to use larger values of L (typically $24\le L\le 56$) and then errors become negligible.

An efficient algorithm for computation of weighted means can be derived by observing Fig. 2. In the case of example $W=0.73$ the first 12 inputs should be x_1 and the remaining 4 inputs should be x_2. Of course, due to idempotency it is not necessary to compute function in all nodes but only in nodes with different inputs. The computation can be done according to the algorithm WA2 shown in Fig. 3.

WEIGHTED AGGREGATION ALGORITHM WA2 FOR COMPUTING THE WEIGHTED AGGREGATOR $y=Wx_1\Diamond(1-W)x_2$ FROM $y=F(x_1,x_2)=0.5x_1\Diamond 0.5x_2$

(1) Compute the path index value
$B=K-1=\lfloor WN+1/2 \rfloor-1$

(2) Convert B to n-bit binary number (binary sequence) $b_n b_{n-1}...b_1$, $b_i\in\{0,1\}$:
$\left[\, b_i=B \bmod 2\in\{0,1\};\quad B:=\lfloor B/2 \rfloor \,\right], i=1,...,n$.

(3) Transform the binary sequence $b_1 b_2 ... b_n$ to the path sequence $p=p_1 p_2 ... p_n$, where $p_i=2-b_i$, $p_i\in\{1,2\}$, $i=1,...,n$

(4) Simplification based on idempotency: delete leading 1's from the sequence p so that it starts with 2 and rename the truncated sequence $t=t_1 t_2 ... t_m$, $t_1=2$, $m\le n$

(5) The resulting weighted aggregation function is
$y=F(...F(F(x_1,x_{t_1}),x_{t_2}),...,x_{t_m})$

Fig. 3. The weighted aggregation algorithm WA2 for aggregating two variables

The WA2 algorithm is exemplified in the lower part of Fig.2. For $W=0.73$, $L=4$ we have:

(1) Path index: $B=\left\lfloor 0.73\times 2^4+1/2 \right\rfloor-1=11$

(2) Conversion to the binary sequence

$$\begin{aligned} &B=11, && b_1=11 \bmod 2=1\\ &B:=\lfloor 11/2 \rfloor=5, && b_2=5 \bmod 2=1\\ &B:=\lfloor 5/2 \rfloor=2, && b_3=2 \bmod 2=0\\ &B:=\lfloor 2/2 \rfloor=1, && b_4=1 \bmod 2=1\\ &\therefore\ 11_{10}=1011_2=b_4 b_3 b_2 b_1 \end{aligned}$$

(3) The path sequence: $p=p_1 p_2 p_3 p_4=1121$

(4) Simplification based on idempotency: $t=t_1 t_2=21$

(5) The resulting weighted aggregator:
$y=F(F(x_1,x_2),x_1)=½(½x_1\Diamond½x_2)\Diamond½x_1$
$=¼x_1\Diamond¼x_2\Diamond½x_1=¾x_1\Diamond¼x_2$

If we need a better accuracy, then the same procedure can be performed with larger value of L. For example, the same algorithm for $W=0.73$, $L=10$ yields the following:

(1) Path index: $B=\left\lfloor 0.73\times 2^{10}+1/2 \right\rfloor-1=747$

(2) Conversion to the binary sequence: $747_{10}=1011101011_2=b_{10}b_9b_8b_7b_6b_5b_4b_3b_2b_1$

(3) The path sequence: $p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}=1121211121$

(4) Simplification based on idempotency: $t=t_1t_2t_3t_4t_5t_6t_7t_8=21211121$

(5) The resulting weighted aggregator:
$y=F(F(F(F(F(F(F(F(x_1,x_2),x_1),x_2),x_1),x_1),x_1),x_2),x_1)$.

The manual computation of weighted aggregators can be tedious. Fortunately, an elegant and efficient version of the WA2 algorithm can be implemented as shown in Fig. 4. Note that the time complexity of this algorithm is $O(L)$.

Let us now investigate a general case of n>1 arguments, starting with a simple case of 3 arguments $y=0.48x_1\Diamond 0.38x_2\Diamond 0.14x_3$. This asymmetric weighted aggregator should be created using the related symmetric binary aggregator $F(x_1,x_2)=0.5x_1\Diamond 0.5x_2$. The technique for creating the asymmetric aggregator consists of using a binary tree of desired symmetric aggregators and then repeating inputs in proportion to their weight, as shown in Fig. 5.

The number of levels of the binary tree is selected according to the desired accuracy of approximation of the given aggregator. In our example we use three levels and the relative importance (weight) of each input is $1/2^3$=0.125, because the weight of each leaf is the product of weights along the path from the root to the leaf. Then we distribute the attributes so that the impact of each attribute is close to the desired value of weight. If we use four inputs equal to x_1, three inputs equal to x_2 and one input equal to x_3, and apply the idempotency as shown in Fig. 5, then the desired approximation is:

$$\begin{aligned}&0.5x_1\Diamond 0.5(0.5x_2\Diamond 0.5(0.5x_2\Diamond 0.5x_3))\\&=0.5x_1\Diamond 0.5(0.5x_2\Diamond 0.25x_2\Diamond 0.25x_3)\\&=0.5x_1\Diamond 0.5(0.75x_2\Diamond 0.25x_3)\\&=0.5x_1\Diamond 0.375x_2\Diamond 0.125x_3\approx 0.48x_1\Diamond 0.38x_2\Diamond 0.14x_3\end{aligned}$$

```
// WEIGHTED AGGREGATION ALGORITHM WA2
// (the iterative case of 2 variables)
// Function F is the symmetric base
// aggregator.
// Input x1 has the weight W
// Input x2 has the weight 1-W
// L = number of binary tree levels
// Run time = O(L)

double WA2(double x1,double x2,double W,
       double(*F)(double,double),int L)
{
  int B = int(W*pow(2.,L)+0.5) - 1;
  double y = x1;
  while(L--)
  {
     y = F(y, B%2 ?  x1 : x2);
     B/=2;
  }
  return y;
}
```

Fig. 4. A function that implements the WA2 algorithm

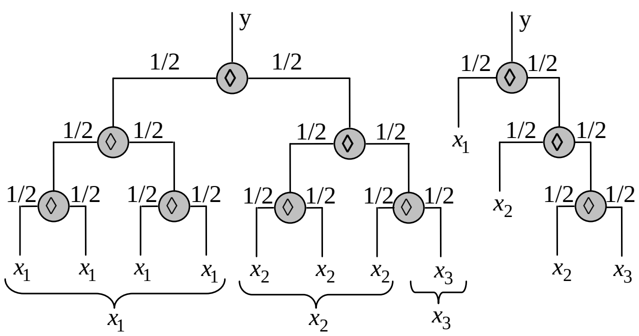


Fig. 5. Creating a weighted aggregator with multiple inputs from a binary aggregator with equal weights

The method presented in Fig. 2 can be generalized to include n variables $x_1,...,x_n$ and different weights $W_1,...,W_n$. As in the case of 2 variables, a general weighted aggregator of n variables $y=W_1x_1\Diamond...\Diamond W_nx_n$ can be derived using a related symmetric binary aggregator $F(x_1,x_2)=0.5x_1\Diamond 0.5x_2$, as shown in Fig. 6.

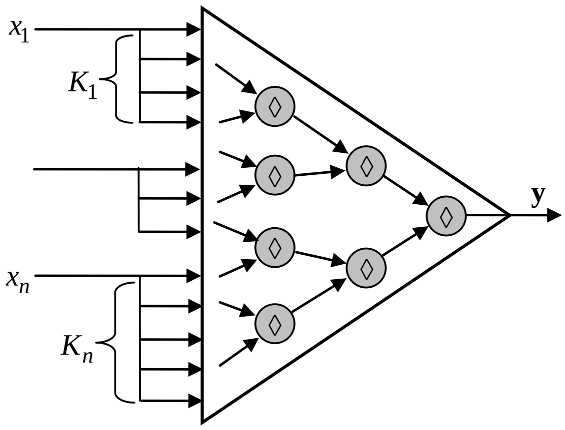


Fig. 6 A general weighted aggregation structure

This method requires idempotency, but does not need ILA associativity. The idea is that each node preserves the characteristics of the base symmetric ILA and weights are represented by the numbers of repeated inputs. So, in this case the weights are represented and adjusted as follows:

$$K_i=\lfloor W_iN+½\rfloor,\quad i=1,...,n-1,\quad K_n=N-K_1-...-K_{n-1}$$

A different approach to creating a general WAn algorithm for n>2 variables is to use the associativity of quasi-arithmetic means and to realize WAn by repeating n-1 times the WA2 algorithm as follows:

$$\begin{array}{cccccccccc}y\leftarrow & WA2 & \leftarrow & WA2 & \leftarrow & WA2 & \leftarrow & WA2 & \leftarrow & WA2\leftarrow x_n\\ & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow\\ & x_1 & & x_2 & & ... & & x_{n-2} & & x_{n-1}\end{array}$$

Indeed, the quasi-arithmetic means support the following form of associativity:

$$\begin{aligned}y&=f^{-1}\left[W_1f(x_1)+W_2f(x_2)+...+W_{n-1}f(x_{n-1})+W_nf(x_n)\right]\\&=f^{-1}\left[\begin{array}{l}W_1f(x_1)+W_2f(x_2)+...\\+(W_{n-1}+W_n)\dfrac{W_{n-1}f(x_{n-1})+W_nf(x_n)}{W_{n-1}+W_n}\end{array}\right]\\&=f^{-1}\left\{\begin{array}{l}W_1f(x_1)+W_2f(x_2)+...\\+(W_{n-1}+W_n)f\left[f^{-1}\left(\dfrac{W_{n-1}f(x_{n-1})+W_nf(x_n)}{W_{n-1}+W_n}\right)\right]\end{array}\right\}\end{aligned}$$

Using this form of QAM associativity we can write a fast O[(n-1)L] iterative WAn algorithm shown in Fig.7.

Two basic components of the performance of WA algorithms are the accuracy and speed of computation. The performance can be analyzed using as a benchmark the simple symmetric power mean of two variables $F(x_1,x_2)=(0.5x_1^r+0.5x_2^r)^{1/r}$ to implement the target weighted power mean (WPM) of n variables

$M^{[r]}(\mathbf{x};\mathbf{W}) = (W_1 x_1^r + ... + W_n x_n^r)^{1/r}$. For the analysis of accuracy of WA2, we can compute the number of accurate significant decimal digits delivered by the WA2: $D(L) = |\log|WA2(x_1, x_2, W, F, L) - M^{[r]}(\mathbf{x};\mathbf{W})|/\log 10|$. If we use uniformly distributed random values of x_1, x_2, W and r then a simulation model with 10^7 samples creates results shown in Fig. 8. The mean number of accurate significant decimal digits is a linear function of the number of levels, $D_{ave}(L) = 0.301L + 1.5233$. Thus, each level contributes one correct bit of mantissa.

```
// WEIGHTED AGGREGATION ALGORITHM WAn
// (the iterative case of n>=2 variables)
// Function F is the symmetric base
// aggregator.
// W[ ] = array of weights of inputs x[ ]
// L = number of binary tree levels
// Run time = O[(n-1)L]

double WAn(double x[],double W[],
   int n, double(*F)(double,double),int L)
{
  --n;
  double Wsum = W[n], y = x[n];
  while(n)
  { --n;
    y = WA2(x[n],y, W[n]/(W[n]+Wsum),F,L);
    Wsum += W[n];
  }
  return y;
}
```

Fig. 7 A function that implements the WAn algorithm

3 USING WAn TO EXTEND GEOMETRIC AND LOGARITHMIC MEANS

To verify the functionality of the WAn algorithm let us use two related symmetric two variable means: the geometric mean $G(x,y) = \sqrt{xy}$ and the logarithmic mean $L(x,y) = (x-y)/(\log x - \log y)$. The geometric mean is a special case of the power mean and has natural extension to weighed n-tuples: $G(x_1,...,x_n;\mathbf{W}) = x_1^{W_1}...x_n^{W_n}$. Both $G(x,y)$ and $L(x,y)$ are hard partial conjunction threshold aggregators: $G(x,y)$ defines the threshold andness of power means $\alpha_\theta(PM) = 0.667$ and $L(x,y)$ defines the threshold andness of generalized logarithmic means $\alpha_\theta(GLM) = 0.614$. In addition, both $G(x,y)$ and $L(x,y)$ are special cases of the generalized logarithmic mean and the Stolarsky mean [7]. The logarithmic mean does not have obvious extension to weighed n-tuples, but there are proposals for such extensions [11], [12]. Our WAn algorithm is one of such extensions.

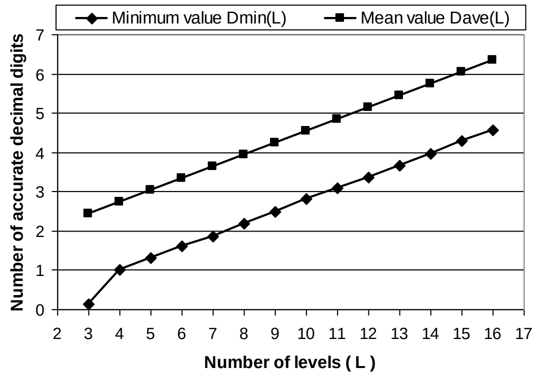


Fig. 8. The number of accurate decimal digits $D(L)$ generated by WA2

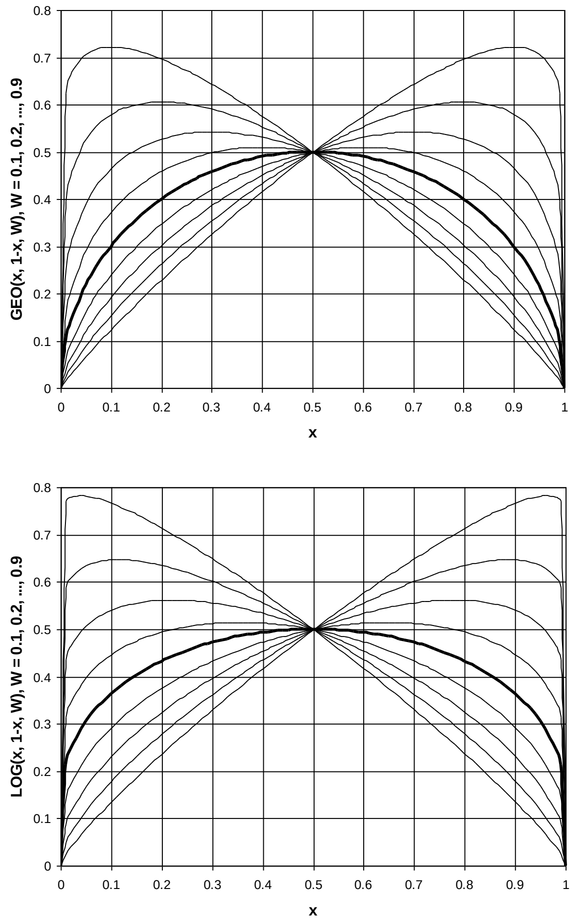


Fig. 9. Weighted geometric and weighted logarithmic means generated using the WAn algorithm with L=32

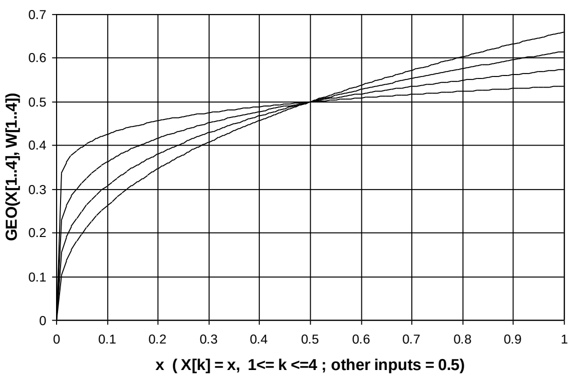
Let us now use $G(x,y)$ and $L(x,y)$ to create a weighted GCD aggregator $y = W_1 x_1 \Diamond ... \Diamond W_n x_n$. For n=2 Fig. 9 shows the weighted aggregator of two variables $y = Wx \Diamond \bar{W}\bar{x}$, $\bar{x} = 1-x$, $\bar{W} = 1-W$ for weights from 0.1 to 0.9. In the case of geometric mean, WAn generates the output that is identical to $G(x,\bar{x};W) = x^W \bar{x}^{\bar{W}}$. Of course, that is expected because WAn is based on associative properties of QAM. In contrast, the logarithmic mean is not a descendant of QAM and not associative, but the result is nevertheless very similar to the extended geometric mean showing that WAn preserves the nature of the symmetric base mean and provides a viable method for extending symmetric means to support weights.

The case n>2 can be analyzed using an aggregator with increasing weights, $y = 0.1x_1 \Diamond 0.2x_2 \Diamond 0.3x_3 \Diamond 0.4x_4$, that we can implement using $G(x, y)$ or $L(x, y)$ and WAn. The characteristics of resulting aggregators can be visualized using the following four sensitivity functions:

$$y_i(x) = 0.1x_1 \Diamond 0.2x_2 \Diamond 0.3x_3 \Diamond 0.4x_4$$

$$x_i = x,\ \ x_j = 0.5,\ j \neq i,\ \ 0 \leq x \leq 1,\ i = 1,2,3,4$$

The sensitivity functions presented in Fig. 10 show remarkable similarity and verify the usefulness of the WAn algorithm. In the case of geometric mean, WAn correctly transforms $\sqrt{xy}$ to $G(\mathbf{X}) = x_1^{0.1}x_2^{0.2}x_3^{0.3}x_4^{0.4}$.



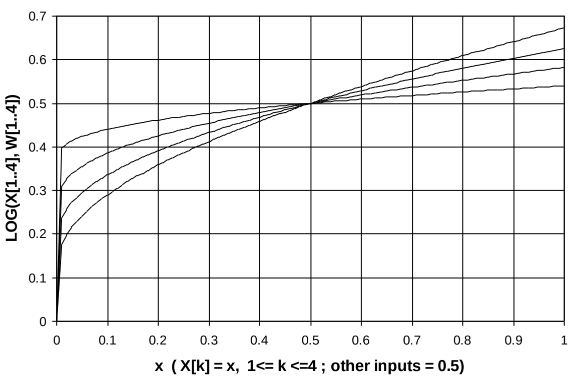


Fig. 10. Sensitivity curves for weighted geometric and logarithmic means of 4 variables generated using WAn

4 CONCLUSIONS

Each symmetric idempotent aggregator (i.e. each mean) of two variables can be expanded to weighted aggregator (weighted mean) of n variables using the WAn algorithm introduced in this paper. WAn is based on QAM associativity and for all QAM aggregators it generates correct results. However, our experiments show that WAn generates viable results also for non-associative non-QAM aggregators. To what extent is QAM associativity applicable outside the QAM world is an open question.

WAn is a general, simple and elegant iterative algorithm that can be implemented in any programming language using just a few lines of code. The accuracy of WAn, expressed as the number of significant decimal digits, is a linear function of the number of levels L in the binary aggregation tree. Each level contributes 1 accurate bit of mantissa, i.e. $\log_{10}(2) = 0.301$ significant decimal digits. Assuming a constant value of L (e.g. L=24) WAn is a fast linear algorithm with the run time T=O(n). That is a fully acceptable solution. Since all computations are done in nodes of a balanced L-level binary tree, parallel level-wise computations are also possible, but it is unlikely that further speedup might have significant practical effects.

References

[1] M. Grabisch, J-L. Marichal, R. Mesiar, and E. Pap, *Aggegation Functions*. Cambridge Univ. Press, 2009.

[2] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer 2007.

[3] J. Dujmović, "Weighted compensative logic with adjustable threshold andness and orness," IEEE TFS Vol. 23, No. 4, pp. 270-290, 2015.

[4] J. Dujmović, "Preference logic for system evaluation," IEEE Trans. on Fuzzy Systems, Vol. 15, No. 6, Dec. 2007, pp.1082-1099.

[5] J. Dujmović and H.L. Larsen, "Generalized conjunction/disjunction." *International J. Approx. Reas.* 46, 2007, pp. 423-446.

[6] C. Gini in collaboration with G. Barbensi, L. Galvani, S. Gatti, and E. Pizzetti, *Le Medie*. Unione Tipografico-Editrice Torinese, Torino, 1958.

[7] P.S. Bullen, *Handbook of means and their inequalities*, Kluwer, 2003.

[8] J. Dujmović, "Interpolative GCD Aggregators," Proc. of the 2014 IEEE World Congress on Computational Intelligence, Beijing, China July 6-11,2014, DOI: 10.1109/FUZZ-IEEE.2014.6891841, pp. 1778-1785.

[9] T. Calvo and R. Mesiar, "Fusion with Quantitative Weights." EUSFLAT 2003 Proceedings, Edited by M. Wagenknecht and R. Hampel. http://www.eusflat.org/publications_proceedings_EUSFLAT_2003.php, pp. 312-317.

[10] T. Calvo, R. Mesiar, and R. Yager, "Quantitative Weights and Aggregation." IEEE Trans. on Fuzzy Systems, Vol. 12, No. 1, Feb. 2004, pp.62-69.

[11] A. Pittenger, "The Logarithmic Mean in n Variables." The American Mathematical Monthly, Vol. 92, No. 2 (Feb. 1985), pp. 99-104.

[12] S. Simić, "An Extension of Stolarsky Means." Novi Sad J. Math., Vol. 38, No. 3, 2008, pp. 81-89.

AN INTELLIGENT AGGREGATION TECHNIQUE FOR CAPTURING HETEROGENEOUS RELATIONSHIP AMONG INPUTS

Bapi Dutta and Debashree Guha
Department of Mathematics
Indian Institute of Technology
Patna-800013, India
bapi.iitr@gmail.com
debashree@iitp.ac.in

Summary

In this paper, we provide a framework for aggregating heterogeneously interrelated inputs and introduce a family of aggregation functions to capture such kind of interrelationship in the aggregation process. We find that Bonferroni mean and its extension are special cases of this new family of aggregation functions. We also investigate its properties.

Keywords: Aggregation function, Heterogeneous Relationship, Extended Bonferroni Mean.

1 INTRODUCTION

The central issue in aggregation technique is to fuse the information that are coming from different sources, according to their interrelationship patterns and produces an output which not only be an representative of the input information but also reflects the interrelationship among aggregated arguments. Different interrelationship patterns may lead to different aggregated values for same input information and that may affect the final result of the system [3]. Therefore, taking account of interrelationships of information in aggregation process is immensely important.

Various types of aggregation functions have been introduced in the literature and their properties have also been studied extensively [7, 4]. Also, some of the aggregation functions have been designed to capture a specific relationship among the inputs or to model specific requirement of particular scenario, such as, Prioritize aggregation function [9], Power aggregation function [8], etc. However, most of the aggregation functions produce reliable outputs under the assumptions that there is no interrelationship among the aggregated arguments except few aggregation functions, such as, Choquet integral [6], Generalized Bonferroni mean [10, 5], extended Bonferroni [3] mean. The basic difference between Choquet integral and Bonferroni mean is that the former captures interrelationship based on the subjective fuzzy measure, whereas the latter relies on the direct conjunction of interrelated inputs. In this study, we will focus on direct conjunction of interrelated inputs for taking account of interrelationship among the arguments in the aggregation process.

By interrelationship, we mean a heterogeneous connection among the inputs, i.e., some of the inputs are related to a subset of rest of the inputs and others have no relation with the remaining inputs.

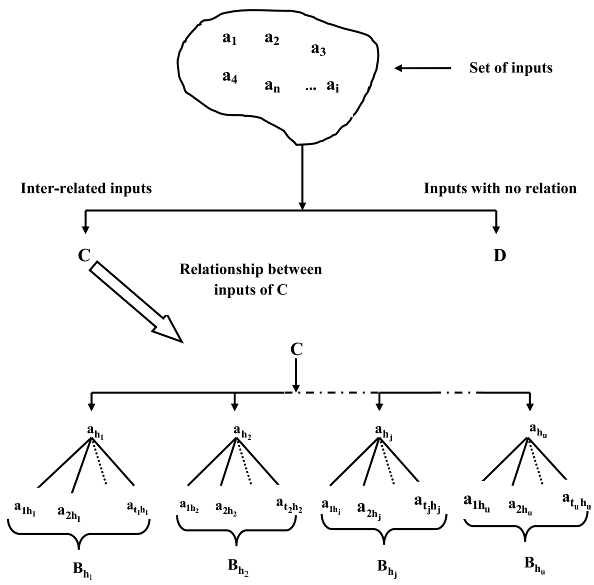


Figure 1: Heterogeneous interrelationships among inputs where $a_i - a_j$ represents that a_i is related to a_j. The attributes $\{a_{h_1}, a_{h_2}, \ldots, a_{h_u}\}$ are dependent and each $a_{h_i} \in C$ is related to a subset B_{h_i} of C..

ISBN: 978–83–8012–519–3

Let $a = (a_1, a_2, \ldots, a_n)$ be the set of inputs, where a_i's are non-negative real numbers from the interval $[0, 1]$. Then, this kind of heterogeneous relationship among the inputs can be described as in Fig. 1. Such kind of interaction among the inputs exists in many real-life problems. For instance, attributes of a multi-attribute decision making problem may be interrelated in such fashion, data sources, like different sensors may have such heterogeneous connections, at the time of forming social opinion, individuals may have this kind of connections.

In this work, we provide a general framework for capturing this kind of interrelationship among inputs in the aggregation process and define a composite aggregation function to aggregate inputs with this kind of relationship patterns. To do this, the rest of the paper is drafted as follows. In section 2, we provide a brief overview of the aggregation functions and some of its properties. Section 3, introduces a framework for describing relationship among the inputs and a composite aggregation function. Properties of new composite aggregation function are also studied in section 3. Several special cases are also analyzed.

2 PRELIMINARIES

In this section, we provide some useful concepts of aggregation functions [7], which are the basis of our proposal.

Definition 1. *An aggregation function in* $[0, 1]^n$ *is a function* $A : [0, 1]^n \to [0, 1]$ *that*

- *is non-decreasing in each argument*

  $A(\mathbf{x}) \leq A((x'))$ *for any* $\mathbf{x}, \mathbf{x}' \in [0, 1]^n$ *with* $\mathbf{x} \leq \mathbf{x}'$

- *satisfies the boundary conditions*

$$A(0, 0, \ldots, 0) = 0 \text{ and } A(1, 1, \ldots, 1) = 1$$

Definition 2. *The diagonal section of any aggregation function* $A : [0, 1]^n \to [0, 1]$ *is the unary function* $\delta_A : [0, 1] \to [0, 1]$ *defined as*

$$\delta_A(x) = A(x, x, \ldots, x) \text{ for all } x \in [0, 1]$$

Definition 3. *An aggregation function* $A : [0, 1]^2 \to [0, 1]$ *is said to have an inverse diagonal if the inverse of the diagonal section* δ_A *exists, i.e., there exists a function* $\delta_A^{-1} : [0, 1] \to [0, 1]$ *such that* $\delta_A^{-1}(\delta_A(x)) = x$

Definition 4. *An aggregation function* $A : [0, 1]^n \to [0, 1]$ *is said to be idempotent if* $\delta_A = id$*, i.e.,*

$$A(x, x, \ldots, x) = x \text{ for all } x \in [0, 1]$$

Definition 5. *Let* $A : [0, 1]^n \to [0, 1]$ *be an aggregation function, then*

- *A is conjunctive if* $A(\mathbf{x}) \leq Min(\mathbf{x})$
- *A is disjunctive if* $Max(\mathbf{x}) \leq A(\mathbf{x})$
- *A is averaging if* $Min(\mathbf{x}) \leq A(\mathbf{x}) \leq Max(\mathbf{x})$

3 NEW AGGREGATION FUNCTIONS

In this section, we define a general composite aggregation function to aggregate a set of inputs with heterogeneous relationship as described in Fig. 1.

The conventional aggregation technique is based on the common hypothesis that information is passed to an aggregation function as an ordered sequence of real numbers together with its weight information and that works pretty well in most of the aggregation functions. But sometimes we need additional information regarding inputs to fuse the data intelligently. For instance, in induced OWA operator [11], an additional information is passed to the aggregation function via inducing variables which defines the permutation among the inputs. Similarly, in order to capture the heterogeneous relationship among the inputs as described in Fig. 1 intelligently, it is needed to pass the interrelationship information to the aggregation function. For this purpose, the heterogeneous relationship among the inputs can be described as below.

Let I_i be the set of indices of inputs which are related to a_i. We have considered only one level of dependency of a_i with the elements of I_i. The mutual dependency of the elements of I_i are not accounted here. The nature of the set I_i, determines the input a_i is related to any other inputs or not. If $I_i = \emptyset$, the input a_i has no connection to any other inputs (that means a_i belongs to D as in Fig. 1). Otherwise, a_i is related to the inputs whose indices are in I_i (that means a_i belong to C as in Fig. 1 and if $i = h_1$ then the indices of the inputs of B_{h_1} are in I_i). Therefore, the input a_i with additional information I_i can describe its relationship with other inputs completely. Now, we need to group the inputs based on the fact that whether I_i is empty or not. Without loss of generality, we assume that first n_1 inputs have some sort of connection with a subset of the rest of the inputs, i.e., $I_i \neq \emptyset$ for $i = 1, 2, \ldots, n_1$ and each of the remaining inputs $(n - n_1)$ have no connection with the remaining inputs, i.e., $I_i = \emptyset$ for $i = n_1 + 1, \ldots, n$. Such kind of grouping may be viewed as pre-processing of inputs. The heterogeneous relationship information of inputs with the set of inputs can be transmitted to the aggregation function as a sequence of ordered pairs as $(\langle I_1, a_1\rangle, \langle I_2, a_2\rangle, \ldots, \langle I_n, a_n\rangle)$. The cardinality of the set I_i is denoted by $|I_i|$.

Now, our intention is to aggregate this information taking logical conjunction of interrelated inputs. The logical architecture of the aggregation system can be described as follows:

$$Agg(Agg_1(\text{interrelated inputs}), Agg_2(\text{independent inputs}))$$

where, Agg_1(interrelated inputs) implies aggregation of conjunctions of interrelated inputs, i.e.,

$$Agg_1(a_i \; AND \; Agg_3(\text{inputs which belong to } I_i))$$

Here, Agg, Agg_1,Agg_2 and Agg_3 represent aggregation of information by different aggregation functions. That is, we divide our task of aggregation into two parallel sub-tasks of aggregation on two different chunks of information of the inputs based on the interrelationship pattern of inputs. Based on this logical framework, we mathematically define the new aggregation function as follows:

Definition 6. *Let $(\langle I_1, a_1\rangle, \langle I_2, a_2\rangle, \ldots, \langle I_n, a_n\rangle)$ be the set of heterogeneously related inputs with corresponding interrelationship information and $I_i \neq \emptyset$ for $i = 1, 2, \ldots, n_1$. Let $A : [0,1]^2 \to [0,1]$, $A_1 : [0,1]^{n_1} \to [0,1]$ and $A_2 : [0,1]^{n-n_1} \to [0,1]$. Let $E_1 : [0,1]^{|I_1|} \to [0,1]$,...,$E_{n_1} : [0,1]^{|I_{n_1}|} \to [0,1]$ be the aggregation functions and $K : [0,1]^2 \to [0,1]$ be a conjunctive aggregation function. Then the function $H \equiv A(A_1(K(a_1, E_1), \ldots, K(a_{n_1}, E_{n_1})), A_2) : [0,1]^n \to [0,1]$ given by*

$$\begin{aligned} &H(\langle I_1, a_1\rangle, \langle I_2, a_2\rangle, \ldots, \langle I_n, a_n\rangle) \\ &= A(A_1(K(a_1, E_1(a_i|i \in I_1)), \ldots, K(a_{n_1}, E_{n_1}(a_i|i \in I_{n_1}))) \\ &\qquad , A_2(a_{n_1+1}, \ldots, a_n)) \quad (1) \end{aligned}$$

is a composite n-ary heterogeneously related information combining function. Here, we take the convention that aggregation of no information as zero (if $n_1 = n$, then this concerns with aggregation function A_2 and when $n_1 = 0$, this concerns with aggregation function A_1).

To clarify the definition consider the following example.

Example 1. *Let $a = (a_1, a_2, a_3, a_4, a_5, a_6)$ be the set of heterogeneously related inputs. Suppose the input a_1 is related to $\{a_2, a_3\}$, input a_2 is related to $\{a_1, a_4\}$, input a_3 is related to $\{a_1, a_4\}$, input a_4 is related to $\{a_2, a_3\}$ and inputs $\{a_5, a_5\}$ have no relation with the remaining ones. This heterogeneous relationship can be describe via I_i as follows: $I_1 = \{2, 3\}$, $I_2 = \{1, 4\}$, $I_3 = \{1, 4\}$, $I_4 = \{2, 3\}$, $I_5 = \emptyset$ and $I_6 = \emptyset$. So for the given heterogeneously inputs, the function H becomes as follows:*

$$\begin{aligned} &H(\langle I_1, a_1\rangle, \langle I_2, a_2\rangle, \langle I_3, a_3\rangle, \langle I_4, a_4\rangle, \langle I_5, a_5\rangle\langle I_6, a_6\rangle) \\ &= A(A_1(K(a_1, E_1(a_2, a_3)), K(a_2, E_2(a_1, a_4)), \\ &\quad K(a_3, E_3(a_1, a_4)), K(a_4, E_4(a_2, a_3))), A_2(a_5, a_6)) \end{aligned} \quad (2)$$

In Definition 1, conjunctive aggregation function K models the conjunction of a_i with its interrelated inputs while aggregation function A_1 provides the aggregated value of different interrelated inputs. On the other hand, A_2 aggregates the independent inputs and finally, aggregation function A produces combined aggregated values of interrelated inputs and independent inputs. By assigning different aggregation functions to do different aggregation tasks, we can model the requirements of the specific scenario. Moreover, selection of aggregation functions A, A_1, A_2 and K depend on the user's behavior towards the aggregation and specific requirement of the decision scenario. As the requirement of satisfaction levels of different interrelated inputs sets $(I_i(i = 1, 2, \ldots n_1))$ may not be same, the choice of $E_1, \ldots, E_{n_1}$ are made independently.

One of the attracting feature of H is that the nature of the aggregation function H may be determined by the nature of the aggregation functions A, A_1, A_2, K and $E_1, \ldots, E_{n_1}$.

Proposition 1. *The composite n-ary function H is an aggregation function.*

Proof. The boundary conditions $H(\langle I_1, 0\rangle, \langle I_2, 0\rangle, \ldots, \langle I_n, 0\rangle) = 0$ and $H(\langle I_1, 1\rangle, \langle I_2, 1\rangle, \ldots, \langle I_n, 1\rangle) = 1$, and the property non-decreasing in each argument of H follows from the same property possesses by the aggregation functions A, A_1, A_2, K and $E_1, \ldots, E_{n_1}$. □

Example 2. *Let $a = (a_1, a_2, a_3, a_4)$ be the set of inputs with heterogeneous interrelationships of inputs described as $I_1 = \{2, 3\}$, $I_2 = \{1\}$, $I_3 = \{1\}$ and $I_4 = \emptyset$. Let us take A, A_1 and A_2 as simple arithmetic means, the conjunctive aggregation function $K(x, y) = xy$, and aggregation functions $E_1(x, y) = x$ with E_2 and E_3 as identity. Then the aggregation H for aggregating heterogeneously related inputs H turns into*

$$\begin{aligned} &H(\langle I_1, a_1\rangle, \langle I_2, a_2\rangle, \langle I_3, a_3\rangle, \langle I_4, a_4\rangle) \\ &\quad = A(A_1(K(a_1, E_1(a_2, a_3)), K(a_2, E_2(a_1)), \\ &\qquad\qquad K(a_3, E_3(a_1))), A_2(a_4)) \\ &\quad = A(A_1(a_1a_2, a_2a_1, a_3a_1), A_2(a_4)) \\ &\quad = A((2a_1a_2 + a_1a_3)/3, a_4) \\ &\quad = \frac{(2a_1a_2 + a_1a_3)/3 + a_4}{2}. \end{aligned}$$

From the above example, one may note that H does not possess unanimity property (idempotency) even though the aggregation functions A, A_1, A_2 and $E_1,\ldots,E_{n_1}$ possess the unanimity property. That is due to the conjunctive aggregation function K which does not satisfy the unanimity property. The unanimity property is important from many applications point of view, such as, in the context of the formation of social opinion unanimity implies the sovereignty of participants. To preserve the unanimity property, we add the existence of inverse diagonal property to K and modify the Definition 6 as follows:

Definition 7. *Let $(\langle I_1,a_1\rangle,\langle I_2,a_2\rangle,\ldots,\langle I_n,a_n\rangle)$ be the set of heterogeneously related inputs with corresponding interrelationship information and $I_i\neq\emptyset$ for $i=1,2,\ldots,n_1$. . Let $A:[0,1]^2\to[0,1]$, $A_1:[0,1]^{n_1}\to[0,1]$ and $A_2:[0,1]^{n-n_1}\to[0,1]$. Let $E_1:[0,1]^{|I_1|}\to[0,1]$,…,$E_{n_1}:[0,1]^{|I_{n_1}|}\to[0,1]$ be the aggregation functions and $K:[0,1]^2\to[0,1]$ be a conjunctive aggregation function having the inverse diagonal δ_K^{-1}. Then the aggregation function $H_1\equiv A(\delta_K^{-1}(A_1(K(a_1,E_1),\ldots,K(a_{n_1},E_{n_1}))),A_2):[0,1]^n\to[0,1]$ given by*

$$\begin{aligned}&H_1(\langle I_1,a_1\rangle,\langle I_2,a_2\rangle,\ldots,\langle I_n,a_n\rangle)\\&=A(\delta_K^{-1}(A_1(K(a_1,E_1(a_i|i\in I_1)),\ldots,K(a_{n_1}\\&\qquad,E_{n_1}(a_i|i\in I_{n_1})))),A_2(a_{n_1+1},\ldots,a_n))\end{aligned}\tag{3}$$

is an composite n-ary aggregation function for combining inputs with heterogeneous relationship.

Example 3. *Consider the same set of inputs as in Example 2 with same interrelationship pattern. Now taking same set of aggregation functions as in Example 2, we have*

$$\begin{aligned}&H_1(\langle I_1,a_1\rangle,\langle I_2,a_2\rangle,\langle I_3,a_3\rangle,\langle I_4,a_4\rangle)\\&=A(\delta_K^{-1}(A_1(K(a_1,E_1(a_2,a_3)),K(a_2,E_2(a_1)),\\&\qquad\qquad K(a_3,E_3(a_1)))),A_2(a_4))\\&=A(\delta_K^{-1}A_1(a_1a_2,a_2a_1,a_3a_1),A_2(a_4))\\&=A(((2a_1a_2+a_1a_3)/3)^{1/2},a_4)\\&=\frac{((2a_1a_2+a_1a_3)/3)^{1/2}+a_4}{2}\end{aligned}$$

Clearly, H_1 possesses unanimity property as $H_1(\langle I_1,a_1\rangle,\langle I_2,a_1\rangle,\langle I_3,a_1\rangle,\langle I_4,a_1\rangle)=a_1$.

Proposition 2. *The composite n-ary aggregation function H_1, satisfies the idempotent property for any conjunctive function having inverse diagonal, if the idempotency property holds for the aggregation functions, A, A_1, A_2, and $E_1,\ldots,E_n$.*

Proof. Since the aggregation functions A_2, and $E_1,\ldots,E_n$ are idempotent

$$\begin{aligned}&H(\langle I_1,a\rangle,\langle I_2,a\rangle,\ldots,\langle I_n,a\rangle)\\&=A(\delta_K^{-1}(A_1(K(a,a)),\ldots,K(a,a)),a)\\&=A(\delta_K^{-1}(K(a,a),a)),\quad\text{as } A_1 \text{ is idempotent}\\&=A(a,a),\ \text{ as } K(a,a)=\delta_K(a) \text{ and } \delta_K^{-1}(\delta_K(a))=a\\&=a\end{aligned}$$

□

Proposition 3. *The composite n-ary aggregation function H_1 is an averaging aggregation function.*

Proof. Boundedness of H_1 directly follows from the non-decreasing and idempotent properties of H_1. Hence, H_1 is an averaging aggregation function. □

We consider some specific form of the aggregation functions A, A_1, A_2, K, and $E_1,\ldots,E_{n_1}$ in H_1 to recover some well known extension of Bonferroni mean as follows:

(i) When $A=(w_1x^p+w_2y^p)^{1/p}$ with $w_1=n_1/n$ and $w_2=(n-n_1)/n$, A_1 = arithmetic mean $E_i=(\frac{1}{|I_i|}\sum_{i\in I_i}a_i^q)^{1/q}$ for $i=1,2,\ldots,n_1$, $A_2=(\frac{n-n_1}{n}\sum_{i=n_1+1}^{n}a_i^p)^{1/p}$ and $K(x,y)=x^py^q$, we obtain extended Bonferroni mean [3]

$$\begin{aligned}&EBM^{p,q}(a_1,a_2,\ldots,a_n)\\&=\left(\frac{n_1}{n}\left(\frac{1}{n_1}\sum_{i=1}^{n_1}a_i^p\left(\frac{1}{|I_i|}\sum_{j\in I_i}a_j^q\right)\right)^{\frac{p}{p+q}}+\right.\\&\qquad\left.\frac{n-n_1}{n}\left(\frac{1}{n-n_1}\sum_{i\in I'}a_i^p\right)\right)^{\frac{1}{p}}\end{aligned}$$

(ii) When $A=w_1x+w_2y$ with $w_1=n_1/n$ and $w_2=(n-n_1)/n$, A_1 = arithmetic mean, A_2 = geometric mean, E_i = geometric mean for all i and $K(x,y)=xy$, then H_1 becomes

$$\begin{aligned}&H_1(\langle I_1,a_1\rangle,\langle I_2,a_2\rangle,\ldots,\langle I_n,a_n\rangle)\\&=\left(\frac{n_1}{n}\left(\frac{1}{n_1}\sum_{i=1}^{n_1}a_i\left(\prod_{j\in I_i}a_j\right)^{\frac{1}{|I_i|}}\right)^{1/2}+\right.\\&\qquad\left.\frac{n-n_1}{n}\left(\prod_{i\in I'}a_i\right)^{\frac{1}{n-n_1}}\right)\end{aligned}$$

(iii) When every input is related to the rest of the inputs, i.e., $n_1=n$ and $I_i=\{1,2,\ldots,i-1,i+1,\ldots,n\}$ for all i, with $A(x,y)=x$, A_1 = arithmetic mean$E_i=(\frac{1}{|I_i|}\sum_{i\in I_i}a_i^q)^{1/q}$ and $K(x,y)=$

$x^p y^q$, we obtain Bonferroni mean aggregation function [9]

$$BM^{p,q}(a_1, a_2, \ldots, a_n) = \left(\frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_i^p a_j^q \right)^{\frac{1}{p+q}}$$

(iv) When every input is related to the rest of the inputs, i.e., $n_1 = n$ and $I_i = \{1, 2, \ldots, i-1, i+1, \ldots, n\}$ for all i, with $A(x, y) = x$, $A_1 = M_1$ be an n-ary aggregation function, $E_i = M_2$ be an $n-1$-ary aggregation function for all i and K is any 2-ary conjunctive aggregation function having an inverse diagonal, we get generalized Bonferroni mean [8]

$$BM(a) = \delta_K^{-1}(M_1(K(a_1, M_2(a_i \in I_1)), \ldots, K(a_n, M_2(a_i \in I_n)))))$$

4 CONCLUSIONS

In this paper, we have developed a new family of aggregation functions for aggregating heterogeneously related inputs. The interrelationship among the inputs are captured in the aggregation process by taking direct conjunctions of "a_i AND Agg_3 (inputs related to a_i)". We have analyzed averaging behavior of this new family of aggregation functions in the light of the behavior of involved aggregation functions A, A_1, A_2, K, and $E_1, \ldots, E_{n_1}$. By choosing the aggregation functions A, A_1, A_2, K, and $E_1, \ldots, E_{n_1}$ appropriately, we can model the requirement of a specific scenario. We have also found that this new class of aggregation functions has the capability to generalize Bonferroni mean and its all extension.

Further investigation is needed to explore its modeling capability and behavior. Another issue is that how to learn such kind of interrelationship pattern from inputs. The similarity measure may be a tool for learning such interrelationship patterns. It would be also interesting to investigate the aggregation of information with variable data dimension and heterogeneously interrelated sources by the help of idea of recursive rule [12, 1]. Also, the management of sequential aggregation of interrelationships of each input may be tackled by the theory of bags [2].

Acknowledgements

The first author gratefully acknowledges the financial support provided by the Council of Scientific and Industrial Research, New Delhi, India (Award No. 09/1023(007)/2011-EMR-I)

References

[1] A. del Amo, J. Montero, E. Molina (2001). Representation of consistent recursive rules. *European Journal of Operational Research* **130**, pp 29–53.

[2] A. Kolesarova, R. Mesiar, J. Montero (2015). Sequential aggregation of bags. *Information Sciences* **294**, pp 305–314.

[3] B. Dutta, D. Guha, R. Mesiar (2015). A model based on linguistic 2-tuples for dealing with heterogeneous relationship among attributes in multi-expert decision making. *IEEE Transactions on Fuzzy Systems*, Doi:10.1109/TFUZZ.2014.2379291.

[4] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation Functions: A Guide for Practitioners, Springer-Verlag, Berlin.

[5] G. Beliakov, S. James, J. Mordelová, T. Rückschlossová, and R.R. Yager (2010). Generalized Bonferroni mean operators in multi-criteria aggregation. *Fuzzy Sets and Systems*, **161**, pp 2227–2242.

[6] G. Choquet (1954). Theory of capacities. *Annales de l'institut Fourier* **5**, pp 131–295.

[7] M. Grabisch, E. Pap, R. Mesiar, J.L. Marichal (2009). Aggregation Functions, Encyclopedia of Mathematics and its Applications **127**, Cambridge University Press, Cambridge.

[8] R.R. Yager (2001). The power average operator. *IEEE Transactions on Systems, Man, Cybernetics-Part A: Systems and Humans* **31**, pp 724–731.

[9] R.R. Yager (2008). Prioritized aggregation operators. *Internation Journal of Approximate Reasoning* **48**, pp 263–274.

[10] R.R. Yager (2009). On generalized Bonferroni mean operators for multi-criteria aggregation. *International Journal of Approximate Reasoning* **50**, pp 1279–1286.

[11] R.R. Yager, D.P. Filev (1999). Induced ordered weighted averaging operators. *IEEE Transaction on Systems, Man, and Cybernetics* **29**, pp 141–150.

[12] V. Cutello, J. Montero (1999). Recursive connective rules. *International Journal of Intelligent Systems* **14**, pp 3–20.

SOME ISSUES IN AGGREGATION OF MULTIDIMENSIONAL DATA

Marek Gagolewski
(1) Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
(2) Faculty of Mathematics and Information Science, Warsaw University of Technology
ul. Koszykowa 75, 00-661 Warsaw, Poland
gagolews@ibspan.waw.pl

Summary

The aggregation theory usually takes an interest in summarizing a predefined number of points in the real line. In many applications, like in statistics, data analysis, and mining, the notion of a mean – a nondecreasing, internal, and symmetric fusion function – plays a key role. Nevertheless, when it comes to aggregating a set of points in higher dimensional spaces, the componentwise extension of monotonicity and internality might not be the best choice. Instead, the invariance to certain classes of geometric transformations seems to be crucial in such a case.

Keywords: Aggregation, centroid, Tukey median, 1-center, 1-median, convex hull, affine invariance, orthogonalization.

1 INTRODUCTION

For fixed d, let us consider a fusion function $\mathsf{F} : (\mathbb{R}^d)^n \to \mathbb{R}^d$ that takes a set of n vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \in \mathbb{R}^d$ aggregates them into one vector in $\mathbb{R}^d$. In other words, F is such that:

$$\mathsf{F}\left(\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_d^{(1)} \end{bmatrix}, \dots, \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_d^{(n)} \end{bmatrix} \right) = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}.$$

From now on we assume that all vectors are column vectors. Hence, we may conceive F as a function acting on a $d \times n$ real matrix,

$$\mathbf{X} = [\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \mathbf{x}^{(n)}].$$

Note that in data analysis, $\mathbf{x}^{(i)}$ is often called an *observation* (and represent an object or experimental unit), whereas $x_j^{(i)}$ denotes the results of measuring the jth variable or feature (like temperature, weight, etc.) of the ith observation (e.g., a person).

Example 1. Let us take any three non-colinear points in $\mathbb{R}^2$. Even in such a simple case many ways to aggregate a triad exist in the literature, see [16]. The notion of a triangle center function, cf. [3], when rewritten in terms of vertex coordinates, leads us to a fusion function which is – among others – rotation and scale invariant (see below). Among the most well-known triangle centers we find the centroid, in-, circum-, and orthocenter. What is interesting, C. Kimberling's *Encyclopedia of Triangle Centers*[1] as of May 1, 2015 lists, names, and characterizes over 7373 such aggregation methods.

Identifying sine qua non conditions that F should fulfill in order to be useful in a particular application area is important, as the class of all fusion functions is of course too broad. The aim of this short contribution is to attract the aggregation theoreticians' attention to the multidimensional data fusion task, which is from time to time explored in the fields like computational statistics and computational geometry. It should be noted that the in-depth study of unidimensional data aggregation methods successfully led to numerous interesting results which allowed to understand many data fusion processes much better, see [1, 13]. We believe that it might also be the case when we assume that $d > 1$.

In Section 2 we recall the typical axiomatization of a *mean* (aggregation function), which – in the $d = 1$ case – is the classic object of interest of the aggregation theory. It turns out that componentwise extensions of two important properties, namely the monotonicity and internality, is not a good choice in a higher di-

[1] Available online at http://faculty.evansville.edu/ck6/encyclopedia/.

mension space. Instead of them, in Section 3 we focus our attention to invariance to particular classes of geometric transformations. We indicate, among others, a quite general way to construct an orthogonal invariant fusion function.

In Section 4 we present a few ideas on how to symmetrize multidimensional fusion functions. As we will see, it is not a trivial task, as for $d > 1$ we do not have a natural linear ordering of the input values. Finally, Section 5 concludes the paper and presents some other issues one may encounter in multidimensional data fusion tasks.

2 MULTIDIMENSIONAL EXTENSIONS OF MONOTONICITY AND INTERNALITY

As it was already noted, in the aggregation theory, we mostly focus on the $d = 1$ case. The notion of a mean (internal aggregation function) $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$, which may be used to determine the "most typical observation" among a given set of values typically requires the fulfillment of the three following properties (see [1, 13] and also [4, 18, 25]):

- symmetry, i.e. for any permutation σ of the set $\{1, \dots, n\}$ it holds:
$$\mathsf{F}(x_1^{(1)}, \dots, x_1^{(n)}) = \mathsf{F}(x_1^{(\sigma(1))}, \dots, x_1^{(\sigma(n))}),$$
- nondecreasingness, which requires that whenever $x_1^{(i)} \le x_1'^{(i)}$ for all $i = 1, \dots n$, we have:
$$\mathsf{F}(x_1^{(1)}, \dots, x_1^{(n)}) \le \mathsf{F}(x_1'^{(1)}, \dots, x_1'^{(n)}),$$
- internality, that is:
$$\mathsf{F}(x_1^{(1)}, \dots, x_1^{(n)}) \in \left[\bigwedge_{i=1}^{n} x_1^{(i)}, \bigvee_{i=1}^{n} x_1^{(i)}\right].$$

Note that, under monotonicity, internality is equivalent to idempotence, i.e., $(\forall x \in \mathbb{R})$ it holds $\mathsf{F}(x, x, \dots, x) = x$. Also please notice that the choice of the above properties, as well as the set $\mathbb{R}$ in which the vectors' elements reside is particularly sound in statistics and data analysis. From now on we assume in this paper that only these two practical domains attract our interest. This is because, e.g., a decision making or fuzzy reasoning task is of a much different nature.

Let us extend the above properties in such a way that they are valid for any d. Symmetry is the least problematic one: we may simply assume that $\mathsf{F}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \mathsf{F}(\mathbf{x}^{(\sigma(1))}, \dots, \mathbf{x}^{(\sigma(n))})$ must hold for any σ. The easiest and perhaps the most natural approach to extend the other two is by applying them in a componentwise manner.

First of all, note that the ordering structure on $\mathbb{R}$ may easily be extended to $\mathbb{R}^d$ by determining the so-called *product order*. The partial order $\le_d$ is defined so as to for $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$ we have $\mathbf{x} \le_d \mathbf{x}'$ if and only if $(\forall i \in \{1, \dots, d\})$ $x_i \le x_i'$. This leads to the concept of $\le_d$ (componentwise)-nondecreasingness. Such an approach is often used when the topic of aggregation on products of lattices/chains is explored, see e.g. [5, 7, 19].

On the other hand, componentwise internality may be defined by requiring that $\mathsf{F}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ belongs to the set:
$$\left[\bigwedge_{i=1}^{n} x_1^{(i)}, \bigvee_{i=1}^{n} x_1^{(i)}\right] \times \cdots \times \left[\bigwedge_{i=1}^{n} x_d^{(i)}, \bigvee_{i=1}^{n} x_d^{(i)}\right],$$
which is basically the bounding (hyper)rectangle of a given set of input points, compare also [27].

Here are two exemplary fusion functions that fulfill these three properties.

Example 2. The componentwise extension of the arithmetic mean,
$$\mathsf{CwMean}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_1^{(i)} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_d^{(i)} \end{bmatrix},$$
also called the *centroid* (barycenter, geometric center) of a set of points, is a point such that
$$\mathbf{y} = \arg\min_{\mathbf{y} \in \mathbb{R}^d} \sum_{i=1}^{n} \mathfrak{d}_2(\mathbf{x}^{(i)}, \mathbf{y})^2,$$
where $\mathfrak{d}_2$ is the Euclidean distance. This notion is crucial e.g. in the definition of the k-means [22] clustering algorithm.

Example 3. The componentwise extension of the sample median, Med, namely:
$$\mathsf{CwMed}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \begin{bmatrix} \mathsf{Med}\left(x_1^{(1)}, \dots, x_1^{(n)}\right) \\ \vdots \\ \mathsf{Med}\left(x_d^{(1)}, \dots, x_d^{(n)}\right) \end{bmatrix},$$
is sometimes used, cf. [30], as a robust estimate of a multidimensional probability distribution's median.

Yet, the following fusion functions are not $\le_d$-nondecreasing.

Example 4. The *Euclidean 1-center* (smallest enclosing ball, bounding sphere) problem aims at finding:

$$\mathsf{1center}_{\mathfrak{d}_2}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) = \arg\min_{\mathbf{y}\in\mathbb{R}^d} \bigvee_{i=1}^{n} \mathfrak{d}_2(\mathbf{x}^{(i)},\mathbf{y}).$$

It was first proposed by James Sylvester in 1857 [31]. Such a formulation is used in many real-world applications, see e.g. [11], like patter recognition (finding reference points), computational biology (proteine analysis), graphics (ray tracing, culling), data mining (e.g., support vector machines, high-dimensional clustering, nearest neighbor search). This is also the case of the facility location problem, which aims to seek for a location of a distribution center that minimizes the distance to a customer that is situated the farthest away. Unfortunately, there is no analytic solution to the Euclidean 1-center problem, cf. [12] for a discussion and an algorithm.

Euclidean 1-center is not componentwise monotone. Consider $n = 3$ and $d = 2$ with $\mathbf{x}^{(1)} = [1,-1]^T$, $\mathbf{x}^{(2)} = [-1,1]^T$, $\mathbf{x}^{(3)} = [-\sqrt{2},0]^T$. We have $\mathsf{1center}_{\mathfrak{d}_2}(\ldots) = [0,0]^T$. Letting $\mathbf{x}'^{(1)} = \mathbf{x}^{(1)} + [3,0]^T$ we get $\mathsf{1center}_{\mathfrak{d}_2}(\ldots) \approx [1.3,-0.5]^T \not\geq_2 [0,0]^T$.

Example 5. Ca. 1650, Evangelista Torricelli proposed a solution to a problem posed by Pierre de Fermat in the early 17th century: given three points in a plane, find the fourth point for which the sum of its distances to the three given points is as small as possible, cf. [20]. This task can be formulated for arbitrary number of points as follows. Find $\mathbf{y}$ such that:

$$\mathsf{1median}_{\mathfrak{d}_2}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) = \arg\min_{\mathbf{y}\in\mathbb{R}^d} \sum_{i=1}^{n} \mathfrak{d}_2(\mathbf{x}^{(i)},\mathbf{y}).$$

Such a point, called in the literature the *1-median*, geometric median, spatial median, L_1-median, Fermat-Weber, or Torricelli point, generalizes the concept of a one-dimensional median (assuming that the minimum is unique). In statistics, these are also known as L_1 estimators. Again, no analytic formula is known here.

Also 1-median is not componentwise monotone. Take $d = 2$, $n = 3$, and $\mathbf{x}^{(1)} = [0,0]^T$, $\mathbf{x}^{(2)} = [0,1]^T$, $\mathbf{x}^{(3)} = [1,0]^T$. We have $\mathsf{1median}_{\mathfrak{d}_2}(\ldots) \simeq [0.211,0.211]^T$. However, when we take $\mathbf{x}'^{(3)} = \mathbf{x}^{(3)} + [0,2]^T$, then we get $\mathsf{1median}_{\mathfrak{d}_2}(\ldots) = [0,1]^T \not\geq_2 [0.211,0.211]^T$.

Example 6. Tukey [32] introduced the concept of the *halfplane location depth* of $\mathbf{y}$ relative to a given set of points in $\mathbb{R}^d$. It is the smallest number of $\mathbf{x}^{(i)}$'s contained in any closed halfhyperplane with boundary line through $\mathbf{y}$. In other words:

$$\begin{aligned} &\mathrm{tdepth}_d(\mathbf{y};\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) \\ = \ &\min\nolimits_{\mathbf{u}\in\mathbb{R}^d,|\mathbf{u}|=1} |\{i : \mathbf{u}^T\mathbf{x}^{(i)} \geq \mathbf{u}^T\mathbf{y}\}|. \end{aligned}$$

As the deepest point in $d = 1$ generalizes the concept of a median, a deepest value in higher dimensions can be thought of as a multidimensional median: the center of gravity of the deepest depth region is called the *Tukey median*, TkMed. In fact, a bagplot, a bivariate version of the box-and-whisker plot, bases on such a notion [29]. For other multidimensional generalizations of the median, like the Oja or the Liu medians, please refer e.g. to [28, 30].

Tukey median is not componentwise monotone. Consider $n = 4$ and $d = 2$ with $\mathbf{x}^{(1)} = [0,0]^T$, $\mathbf{x}^{(2)} = [1,0]^T$, $\mathbf{x}^{(3)} = [1,1]^T$, and $\mathbf{x}^{(4)} = [0,1]^T$. We have $\mathsf{TkMed}(\ldots) = [0.5,0.5]^T$. Letting $\mathbf{x}'^{(4)} = \mathbf{x}^{(4)} + [1,0]^T$ we get $\mathsf{TkMed}(\ldots) = [2/3,1/3]^T \not\geq_2 [0.5,0.5]^T$.

What is more, monotonicity is not the only property that is somehow problematic. It may be observed that the above extension of internality is not a necessarily nice generalization of ordinary internality. Even though all the above-presented fusion functions fulfill it, it seems to be too weak. Let $d = 2$, $n = 3$ and consider $\mathbf{x}^{(1)} = [1,0]^T$, $\mathbf{x}^{(2)} = [0,0]^T$, $\mathbf{x}^{(3)} = [0,1]^T$. If a fusion function CwG is a componentwise extension of, e.g., $\mathsf{G}(y_1,\ldots,y_n) = \bigvee_{i=1}^n y_i$, then $\mathsf{CwG}(\mathbf{x}^{(1)},\mathbf{x}^{(2)},\mathbf{x}^{(3)}) = [1,1]^T$. Wee see that this is rather not plausible.

Remark 7. The Euclidean 1-center, 1-median, centroid, and componentwise median may be expressed as minimizers of some penalty function. A generalization of the latter two – componentwise cases – have been studied in a product lattice framework in [5].

3 TRANSLATION, SCALE, ORTHOGONAL, AND AFFINE INVARIANCE

Instead of considering monotonicity and internality, researchers in fields like computational statistics and geometry most often focus on invariances with respect to specific classes of geometrical transformations, see e.g. [9]. Namely, one might be interested in finding a fusion function F which fulfills for all input vectors:

- translation invariance: for all $\mathbf{t} \in \mathbb{R}^d$,
$$\mathsf{F}(\mathbf{x}^{(1)}+\mathbf{t},\ldots,\mathbf{x}^{(n)}+\mathbf{t}) = \mathsf{F}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) + \mathbf{t},$$
- scale invariance: for all $s \in \mathbb{R}$,
$$\mathsf{F}(s\mathbf{x}^{(1)},\ldots,s\mathbf{x}^{(n)}) = s\mathsf{F}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}),$$
- orthogonal invariance: for all orthogonal matrices $\mathbf{A} \in \mathbb{R}^{d\times d}$, i.e. matrices such that $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{I}$, or equivalently $\mathbf{A}^T = \mathbf{A}^{-1}$,
$$\mathsf{F}(\mathbf{A}\mathbf{x}^{(1)},\ldots,\mathbf{A}\mathbf{x}^{(n)}) = \mathbf{A}\mathsf{F}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}),$$

and/or

- affine invariance: for all matrices $\mathbf{A} \in \mathbb{R}^{d\times d}$ of full rank and all $\mathbf{t} \in \mathbb{R}^d$,

$$\mathsf{F}(\mathbf{A}\mathbf{x}^{(1)}+\mathbf{t},\ldots,\mathbf{A}\mathbf{x}^{(n)}+\mathbf{t}) = \mathbf{A}\mathsf{F}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)})+\mathbf{t}.$$

Note that orthogonal invariance implies invariance to all possible rotations of input points (and reflections against the axes). On the other hand, affine invariance implies translation, scale, and orthogonal invariance. Moreover, it also covers the case of scaling the coordinates differently in each direction. It is quite useful, as in the practice of data analysis, one often standardizes the variables:

$$x_i \mapsto \frac{x_i - \bar{x}_i}{\mathsf{sd}_i},$$

where $\bar{x}_i$ and sd_i denotes the arithmetic mean and standard deviation, respectively, with respect to the ith coordinate, $i = 1,\ldots,d$. Also, some machine learning methods (like the Principal Component Analysis) assume that the data points may freely be rotated.

It can be noted that, e.g., the centroid and the Tukey median (as well as the Oja and Liu medians) are affine invariant. On the other hand, e.g., the componentwise median is not even orthogonal invariant, see Table 1. In fact, from [13, Proposition 2.116] it follows that the only continuous, symmetric, and componentwise fusion function is CwMean.

Table 1: Exemplary fusion functions and the properties they fulfill: M – componentwise monotonicity, T – translation, S – scale, O – orthogonal, and A – affine invariance.

Function	M	T	S	O	A
CwMean	✓	✓	✓	✓	✓
CwMed	✓	✓	✓		
TkMed		✓	✓	✓	✓
1center		✓	✓	✓	
1median		✓		✓	
OrMed		✓	✓	✓	

3.1 ORTHOGONALIZATION

Orthomedian [14] is a nice example of orthogonalization of the componentwise median. It is defined as an averaged median of all orthogonally transformed input data sets. As the group $\mathcal{O}(d)$ of orthogonal $d \times d$ matrices is compact, we may define:

$$\begin{aligned}&\mathsf{OrMed}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)})\\ &= \int_{\mathcal{O}(d)} \mathbf{A}^{-1}\mathsf{CwMed}(\{\mathbf{A}\mathbf{x}^{(i)}\}_{i=1,\ldots,n})\,d\mathbf{A}.\end{aligned}$$

This fusion function is orthogonal invariant (by construction). Interestingly, it is no more $\le_d$-nondecreasing, so this new property is introduced at some cost.

The above construction is general and can be applied on any fusion function CwG that is a componentwise extension of $\mathsf{G} : \mathbb{R}^n \to \mathbb{R}$. However, it is not easy to compute numerically (e.g., Monte Carlo methods may be used for this purpose, see [8, Sec. 3] for a random uniform – with respect to the Haar measure, see [8] – orthogonal matrix generation algorithm). Thus, here we propose another approach which is valid if $(\forall x_i)$ $\mathsf{G}(x_1,\ldots,x_n) = -\mathsf{G}(-x_1,\ldots,-x_n)$.

Let $\mathbf{X} = [\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \cdots\ \mathbf{x}^{(n)}]$ and assume that: $\mathbf{X}_c = \mathbf{X} - \mathsf{CwMean}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)})$ is a centered version of $\mathbf{X}$. Let us consider the singular value decomposition of $\mathbf{X}_c^T = \mathbf{U}\mathbf{D}\mathbf{V}^T$, where $\mathbf{U}$ is an $n\times n$ orthogonal matrix, $\mathbf{D}$ is a $n \times d$ diagonal matrix, and $\mathbf{V}$ is a $d \times d$ orthogonal matrix. The eigenvectors $\mathbf{v}^{(i)}$ are called *principal component* directions of $\mathbf{X}_c$, see e.g. [15, Sec. 3.4 and Sec. 14.5]. The first principal component direction $\mathbf{v}^{(1)}$ has the property that $\mathbf{z}^{(1)} = \mathbf{X}_c^T\mathbf{v}^{(1)}$ is of the largest sample variance, d_1^2/n among all normalized linear combinations of $\mathbf{X}_c$'s rows. Subsequent principal components have maximum variance subject to being orthogonal to the earlier ones. Thus,

$$\begin{aligned}&\mathsf{OrG}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)})\\ =&\mathbf{V}^{-1^T}\mathsf{CwG}\left(\mathbf{V}^T(\mathbf{X} - \mathsf{CwMean}(\mathbf{X}))\right) + \mathsf{CwMean}(\mathbf{X})\end{aligned}$$

is surely orthogonal and translation invariant.

As for the monotonicity, if G is nondecreasing, then OrG is nondecreasing with respect to the direction that has the maximal variance (and other directions that orthogonal to it and also maximize the remaining variance). Also, if G is internal, then G fulfills the bounding box-based internality.

3.2 A NOTE ON INTERNALITY

Let us recall the notion of the convex hull $\mathrm{CH}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)})$ of a finite set of points. It is the smallest convex set (polytope) that includes all the provided points. Equivalently, it is the set of all convex combinations of $\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}$:

$$\mathrm{CH}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) = \Big\{ \textstyle\sum_{i=1}^n w_i\mathbf{x}^{(i)} :$$
$$\text{for all vectors } \mathbf{w} \ge \mathbf{0} \text{ with } \textstyle\sum_{i=1}^n w_i = 1\Big\}.$$

Having in mind that for $d = 1$ the convex hull is a real interval, the definition of internality may be extended to $d \ge 2$ by using this very notion, which seems far more adequate. Here, we could require that

$$\mathsf{F}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}) \in \mathrm{CH}(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}).$$

Interestingly, it might be shown that if $\mathsf{F} : (\mathbb{R}^d)^n \to \mathbb{R}^d$ is rotation and translation invariant, then the bounding-box based internality and the convex hull-based one coincide. This is because the convex hull is invariant to rotations and translations and that it is a subset of the bounding box. Moreover, the convex hull may be expressed as the intersection of appropriate halfspaces [10]. The points may always be rotated so that any convex hull's face is aligned within the axes. Then the hyperplane that includes such a face coincides with the hyperplane including the bounding box's face.

4 ORDERED FUSION FUNCTIONS

Given a non-symmetric unidimensional function, one may easily symmetrize it by referring to the notion of an order statistic, i.e. the ith smallest value among a set of input elements. It is because, cf. [13, Thm. 2.34], $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ is symmetric if and only if there exists a function $\mathsf{G} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathsf{F}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \mathsf{G}(\mathbf{x}^{(\sigma(1))}, \dots, \mathbf{x}^{(\sigma(n))}),$$

where σ is an ordering permutation the input values. In such a way, e.g., a weighted arithmetic mean becomes the OWA operator. Such a construction is only valid, however, in the $d = 1$ case, as here a natural linear order $\le$ is defined.

If $d > 1$, then it is not easy to determine which values are "small" or "large", especially if we allow a set of points to be orthogonally transformed.

For this purpose, one may order the input values with respect to increasing distances from a fixed point, e.g., the set's componentwise mean. More elaborate approaches may base on the concept of the so-called data depth. For instance, the affine invariant *Oja depth* [26] (or the simplical volume depth) for any given $\mathbf{y} \in \mathbb{R}^d$ is given by:

$$\begin{aligned} & \operatorname{odepth}(\mathbf{y}; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \\ = & \frac{1}{1 + \sum_{(i_1,\dots,i_d)} \operatorname{volume}(\operatorname{CH}(\mathbf{y}, \mathbf{x}^{(i_1)}, \dots, \mathbf{x}^{(i_d)}))}. \end{aligned}$$

Other affine invariant data depth measures include the already mentioned Tukey depth, simplical depth [21]:

$$\begin{aligned} & \operatorname{sdepth}(\mathbf{y}; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \\ = & \left|\left\{(i_1, \dots, i_{d+1}) : \mathbf{y} \in \operatorname{CH}(\mathbf{x}^{(i_1)}, \dots, \mathbf{x}^{(i_{d+1})})\right\}\right| \end{aligned}$$

or the zonoid data depth [9]:

$$\begin{aligned} & \operatorname{zdepth}(\mathbf{y}; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \\ = & \sup\{\alpha \in [0, 1] : \mathbf{y} \in D_\alpha(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})\}, \end{aligned}$$

where D_α is the *α-trimmed region* of the empirical distribution generated by $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$, i.e.:

$$\begin{aligned} & D_\alpha(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \\ = & \left\{\sum_{i=1}^n \lambda_i \mathbf{x}^{(i)} : \sum_{i=1}^n \lambda_i = 1, (\forall i) \alpha\lambda_i \le 1/n\right\}. \end{aligned}$$

With them, the points $\mathbf{x}^{(i)}$, $i = 1, \dots, n$, may be ordered with respect to their decreasing depths. In other words, we may make use of a permutation σ of $\{1, \dots, n\}$ such that $\sigma(i) \le \sigma(j)$ implies that for $i < j$:

$$\operatorname{d}(\mathbf{x}^{(\sigma(i))}; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \le \operatorname{d}(\mathbf{x}^{(\sigma(j))}; \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}),$$

where d is some data depth measure.

Having an ordered version of the input set of points, one may easily define, e.g., multidimensional versions of trimmed or winsorized means, cf. [24].

5 CONCLUSIONS

We briefly explored a few issues in aggregation of multidimensional data. We recalled some interesting fusion functions used in computational statistics and geometry.

It turns out that some of the main ideas of unidimensional aggregation cannot be simply extended to the multidimensional case. The most important one concerns the componentwise monotonicity. Notably, recently the necessity of the notion of monotonicity is being put into question in the classical framework too, see e.g. [2, 6]. In the $d = 1$ case it seems quite natural and moreover it simplifies the way the analytic results are derived. Thus, further studies concerning this notion for $d > 1$ shall be conducted.

Instead of focusing on monotonicity, it seems that researchers in computational statistics and geometry rather focus on invariances to particular classes of geometric transformations. In this contribution, for example, we presented a simple way to guarantee translation and orthogonal invariance. Also note that a recent contribution on measures of dispersion of multidimensional data [17] (some of them base on the notion of multidistances' [23] minimizers) also focuses on translation and rotation invariance.

We additionally noted that some other concepts, like internality or a fusion function's symmetrization also need some more elaborate approaches.

Another problem with multidimensional fusion functions is that many of the tools encountered in the literature cannot be expressed by analytic formulas. This drastically complicates the theoretical studies on them. Even though it is possible to characterize somehow all unidimensional fusion functions that fulfill some of the properties discussed here, things definitely get more complicated in higher dimensions.

REFERENCES

[1] G. Beliakov, A. Pradera, and T. Calvo (2007). *Aggregation functions: A guide for practitioners.* Springer-Verlag.

[2] G. Beliakov, T. Calvo, and T. Wilkin (2014). Three types of monotonicity of averaging functions. *Knowledge-Based Systems* **72**, pp. 114–122.

[3] O. Bottema (1981). Het begrip "merkwaardig" met betrekking tot punten in de driehoeksmeetkunde. *Nieuw Tijdschr. Wisk.* **69**, pp. 2–7.

[4] P. Bullen (2003). *Handbook of means and their inequalities.* Springer Science, Dordrecht.

[5] H. Bustince, E. Barrenechea, T. Calvo, S. James, and G. Beliakov (2014). Consensus in multi-expert decision making problems using penalty functions defined over a cartesian product of lattices. *Information Fusion* **17**, pp. 56–64.

[6] H. Bustince, J. Fernandez, A. Kolesárová, and R. Mesiar (2014). Fusion functions and directional monotonicity. *Communications in Computer and Information Science* **444**, pp. 262–268.

[7] B. De Baets and R. Mesiar (1999). Triangular norms on product lattices. *Fuzzy Sets and Systems* **104**, pp. 61–75.

[8] P. Diaconis and M. Shahshahani (1987). The subgroup algorithm for generating uniform random variables. *Probability In Engineering And Information Sciences* **1**, pp. 15–32.

[9] R. Dyckerhoff, G. Koshevoy, and K. Mosler (1996). Zonoid data depth: Theory and computation. In A. Prat et al. (Eds.), *Proc. COMPSTAT 1996*, Heidelberg, Physica-Verlag, pp. 235–240.

[10] H. Edelsbrunner (1987). *Algorithms in Combinatorial Geometry.* Springer-Verlag, Heidelberg.

[11] B. Gärtner (1999). Fast and robust smallest enclosing balls. *Lecture Notes in Computer Science* **1643**, pp. 325–338.

[12] B. Gärtner and S. Schönherr (2000). An efficient, exact, and generic quadratic programming solver for geometric optimization. In *Proc. 16th ACM Symposium on Computational Geometry*, pp. 110–118.

[13] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap (2009). *Aggregation functions.* Cambridge University Press.

[14] R. Grübel (1996). Orthogonalization of multivariate location estimators: The orthomedian. *The Annals of Statistics* **24**, pp. 1457–1473.

[15] T. Hastie, R. Tibshirani, and J. Friedman (2013). *The Elements of Statistical Learning.* Springer.

[16] C. Kimberling (1998). Triangle centers and central triangles. *Congressus Numerantium* **129**, pp. 1–295.

[17] A. Kołacz and P. Grzegorzewski (2015). Measures of dispersion for multidimensional data, submitted paper.

[18] A. Kolmogorov (1930). Sur la notion de la moyenne. *Atti della R. Academia nazionale dei Lincei* **12**, pp. 388–391.

[19] M. Komorníková and R. Mesiar (2011). Aggregation functions on bounded partially ordered sets and their classification. *Fuzzy Sets and Systems* **175**, pp. 48–56.

[20] J. Krarup and S. Vajda (1997). On Torricelli's geometrical solution to a problem of Fermat. *IMA Journal of Management Mathematics* **8**, pp. 215–223.

[21] R. Y. Liu (1990). On a notion of data depth based on random simplices. *Annals of Statistics* **18**, pp. 405–414.

[22] J. B. MacQueen (1967). Some methods for classification and analysis of multivariate observations. In *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, vol. 1, Berkeley, University of California Press, pp. 281–297.

[23] J. Martin and G. Mayor (2011). Multi-argument distances. *Fuzzy Sets and Systems* **167**, pp. 92–100.

[24] J.-C. Massé (2009). Multivariate trimmed means based on the Tukey depth. *Journal of Statistical Planning and Interference* **139**, pp. 366–384.

[25] M. Nagumo (1930). Über eine Klasse der Mittelwerte. *Japanese Journal of Mathematics* **7**, pp. 71–79.

[26] H. Oja (1983). Descriptive statistics for multivariate distributions. *Statistics & Probability Letters* **1**, pp. 327–332.

[27] S. Ovchinnikov (1996). Means on ordered sets. *Mathematical Social Sciences* **32**, pp. 39–56.

[28] P. J. Rousseeuw and A. Struyf (2004). Computation of robust statistics: Depth, median, and related measures. In J. E. Goodman and J. ORourke (Eds.), *The Handbook of Discrete and Computational Geometry*, Chapman & Hall/CRC, Boca Raton, pp. 1279–1292.

[29] P. J. Rousseeuw, I. Ruts, and J. W. Tukey (1999). The bagplot: A bivariate boxplot. *The American Statistician* **53**, pp. 382–387.

[30] C. G. Small (1990). A survey of multidimensional medians. *International Statistical Review* **58**, pp. 263–277.

[31] J. J. Sylvester (1857). A question in the geometry of situation. *Quarterly Journal of Mathematics* **1**, p. 79.

[32] J. W. Tukey (1974). Mathematics and the picturing of data. *Proc. Intl. Congress of Mathematicians*, pp. 523–531.

ON A CLASS OF UNINORMS WHICH ARE NOT LOCALLY INTERNAL ON THE BOUNDARY

Dana Hliněná
Dept. of Mathematics
FEEC, BUT Brno
Technická 8
Cz-616 00 Brno
Czech Rep.
hlinena@feec.vutbr.cz

Martin Kalina
Dept. of Mathematics
Slovak Univ. of Technology
Faculty of Civil Eng.
Radlinského 11
Sk-810 05 Bratislava
Slovakia
kalina@math.sk

Pavol Kráľ
Dept. of Quant. Methods and Inf. Systems
Matej Bel University
Faculty of Economics
Tajovského 10
Sk-975 90 Banská Bystrica
Slovakia
pavol.kral@umb.sk

Summary

The aim of this paper is to analyze the behaviour of uninorms with a given neutral element $e \in]0,1[$, which contain a zoomed-out uninorm in a square $]a,b[^2$ $(a < e < b)$ and which are constant on $[0,a[^2$, $]b,1]^2$, $]0,a[\times]b,1[$ and $]b,1[\times]0,a[$. Some illustrative examples of such uninorms are provided.

Keywords: Uninorm, Conjunctive uninorm, Not locally internal uninorm on the boundary, Pre-order induced by uninorm.

1 PRELIMINARIES

In 1996 Yager and Rybalov [11] proposed uninorms as a natural generalisation of both t-norms and t-conorms (for details on t-norms and their duals, t-conorms, see, e.g., [8]). Since that time researchers study properties of several distinguished families of uninorms. In the paper we focus on a particular class of uninorms which are not locally internal on the boundary. These uninorms, with a given neutral element $e \in]0,1[$, contain a zoomed-out uninorm in a square $]a,b[^2$ $(a < e < b)$ and are constant on $[0,a[^2$, $]b,1]^2$, $]0,a[\times]b,1[$ and $]b,1[\times]0,a[$. We study their behaviour in the rectangles $[0,a]\times[a,b]$, $[a,b]\times[0,a]$, $[a,b]\times[b,1]$ and $[b,1]\times[a,b]$. Moreover, we provide some illustrative examples of such uninorms.

Definition 1 ([11])**.** *An associative, commutative and increasing operation $U : [0,1]^2 \to [0,1]$ is called a uninorm, if there exists $e \in [0,1]$, called the neutral element of U, such that*

$$U(x,e) = x \quad \text{for all } x \in [0,1].$$

In the theory of fuzzy measures and integrals with respect to fuzzy measures, uninorms play the role of pseudo-multiplication [9].

An overview of basic properties of uninorms can be found in [1]. Because of lack of space we provide only a very brief introduction to uninorms.

A uninorm U is said to be *conjunctive* if $U(x,0) = 0$, and U is said to be disjunctive if $U(1,x) = 1$, for all $x \in [0,1]$.

A uninorm U is called *representable* if it can be written in the form

$$U(x,y) = g^{-1}(g(x) + g(y)),$$

where $g : [0,1] \to [-\infty,\infty]$ is a continuous strictly increasing function with $g(0) = -\infty$ and $g(1) = \infty$. Note that for every generator g there exist two different uninorms depending on convention we take: $\infty - \infty = \infty$, or $\infty - \infty = -\infty$. In the former case we get a disjunctive uninorm, in the latter case a conjunctive uninorm.

Representable uninorms are "almost continuous", i.e., they are continuous everywhere on $[0,1]^2$ except of points $(0,1)$ and $(1,0)$ (see, e.g., [3]).

A class of uninorms continuous on $]0,1[^2$ was characterized by Hu and Li [5] and later a mistake in that paper was corrected by Drygaś [2].

Conjunctive and disjunctive uninorms are dual to each other in the following way

$$U_d(x,y) = 1 - U_c(1-x, 1-y),$$

where U_c is an arbitrary conjunctive uninorm and U_d its dual disjunctive uninorm. Assuming U_c has a neutral element e, the neutral element of U_d is $1-e$.

For an arbitrary uninorm U and arbitrary $(x,y) \in]0,e[\times]e,1] \cup]e,1]\times]0,e[$ we have

$$\min\{x,y\} \le U(x,y) \le \max\{x,y\}. \qquad (1)$$

We say that a uninorm U contains a "*zoomed-out*" uninorm U_z in a square $]a,b[^2$ for $0 \le a < e < b \le 1$

(where $a \neq 0$ and/or $b \neq 1$), if

$$U(x,y) = h^{-1}(U_z(h(x),h(y))) \quad \text{for } x,y \in]a,b[\quad (2)$$

and

$$h(x) = \begin{cases} \frac{x-a}{e-a}e & \text{for } x \in]a,e], \\ \frac{x-e}{b-e}(1-e)+e & \text{for } x \in]e,b[. \end{cases}$$

If a uninorm U contains a zoomed-out uninorm U_z in $]a,b[^2$ (given by formula (2)), we will denote

$$U_s = U \upharpoonright]a,b[^2. \quad (3)$$

In [7] a partial order $\preceq_T$ induced by a t-norm T was introduced and some properties of $\preceq_T$ were studied. In [4] a relation $\preceq_U$ was introduced as a generalization of $\preceq_T$ for uninorms.

Definition 2 ([4])**.** *Let U be a uninorm. By $\preceq_U$ we denote the following relation*

$$x \preceq_U y \quad \text{if there exists } \ell \in [0,1] \text{ such that } U(y,\ell) = x\,.$$

Associativity of U implies transitivity of $\preceq_U$. The existence of a neutral element e implies reflexivity of $\preceq_U$. However, anti-symmetry of $\preceq_U$ is rather problematic.

Since for representable uninorm U and for arbitrary $x \in]0,1[$ and $y \in [0,1]$ there exists ℓ_y such that $U(x,\ell_y) = y$, the relation $\preceq_U$ is, for representable uninorms, not anti-symmetric. There exist also other types of uninorms (e.g., all uninorms which are continuous on $]0,1[^2$) for which $\preceq_U$ is not anti-symmetric (see also [4]).

Lemma 1 ([4])**.** *Let U be a uninorm. The relation $\preceq_U$ is a pre-order.*

We introduce a relation $\sim_U$.

Definition 3. *Let U be a uninorm. We say that $x,y \in [0,1]$ are U-indifferent if*

$$x \preceq_U y \quad \text{and} \quad y \preceq_U x\,.$$

If x,y are U-indifferent, we write $x \sim_U y$.

Finally, we will need the following technical notion.

Definition 4. *Let U be a uninorm. Assume that $a \in [0,1]$ is fixed. We say that U has a-divisors if there exist $x,y \in [0,1]$, $x \neq a$, $y \neq a$, such that $U(x,y) = a$.*

Definition 5. *Let U be a uninorm. We say that U is locally internal on the boundary if for all $x \in [0,1]$ we have*

$$U(1,x) \in \{1,x\}, \quad U(0,x) \in \{0,x\}.$$

2 MAIN RESULTS

Assume that U contains a zoomed-out uninorm in $]a,b[^2$ (we remind that U_s is the notation for the zoomed-out part) and $e \in]a,b[$ is its neutral element of U. Our aim is to construct a uninorm $U : [0,1]^2 \to [0,1]$ such that

$$U(x,y) = \begin{cases} U_s(x,y) & \text{for } (x,y) \in]a,b[^2, \\ 1 & \text{for } (x,y) \in]b,1]^2, \\ 0 & \text{for } (x,y) \in [0,a[^2, \\ & \text{ and if } \min\{x,y\} = 0, \\ a & \text{for } (x,y) \in]0,a[\times]b,1] \\ & \text{ and for } (x,y) \in]b,1] \times]0,a[\,, \end{cases}$$

is fulfilled in case U is conjunctive. If U is disjunctive we change its values for b if $(x,y) \in [0,a[\times]b,1[\cup]b,1[\times [0,a[$ and for 1 if $\max\{x,y\} = 1$. This means that uninorms we are interested in, are not locally internal on the boundary.

We are going to study the behaviour of U in the rectangle $[0,a] \times [a,b]$ (because of commutativity of U we get immediately the rectangle $[a,b] \times [0,a]$).

The behaviour of U in the rectangles $[a,b] \times [b,1]$ and $[b,1] \times [a,b]$ we get by duality.

For arbitrary $x \in]0,a[$ we denote by $U_x : [a,b] \to [0,b]$ the partial function $U_x(y) = U(x,y)$, and for arbitrary $y \in]a,b[$ we denote by $U^y : [0,a] \to [0,b]$ the partial function $U^y(x) = U(x,y)$.

Remark 1.

- If we construct a conjunctive uninorm then we can assume $\text{rng}(U_x) \subset [0,a]$ as well as $\text{rng}(U^y) \subset [0,a]$ for all $x \in]0,a[$ and all $y \in]a,b[$. But in general we have $\text{rng}(U_x) \subset [0,b]$ and $\text{rng}(U^y) \subset [0,b]$.
- If U_s has a-divisors, then necessarily $U(x,y) = a$ for $(x,y) \in \{a\} \times [a,e] \cup [a,e] \times \{a\}$. Similarly, if U_s has b-divisors, then necessarily $U(x,y) = b$ for $(x,y) \in \{b\} \times [e,b] \cup [e,b] \times \{b\}$. This may lead to some problems which we are not going to solve in this paper. We show one example of a conjunctive uninorm where the zoomed-out uninorm U_s has a-divisors.

Let us observe some properties of the partial functions U_x and U^y.

Lemma 2. *Let $y_1, y_2 \in]a,b[$. If $y_1 \sim_{U_s} y_2$ then*

$$\text{rng}(U^{y_1}) = \text{rng}(U^{y_2}).$$

As a direct corollary to Lemma 2 and the U_s-indifference relation we get the following assertion.

Proposition 1. *Let $y_1, y_2 \in]a,b[$ be such that $y_1 \sim_{U_s} y_2$. Assume that $x_1 \in]0,a[$ is a discontinuity point of U^{y_1} with lateral limits*

$$\lim_{x \to x_{1-}} U^{y_1}(x) = x_\ell, \quad \lim_{x \to x_{1+}} U^{y_1}(x) = x_r.$$

Then there exists $x_2 \in]0,a[$ which is a discontinuity point of U^{y_2} and has the same lateral limits at x_2 as U^{y_1} has at x_1.

Proposition 2. *Let $y_1, y_2 \in]a,b[$ be such that $y_2 \preceq_{U_s} y_1$. Assume that $I \subset]0,a[$ is an interval of constantness of U^{y_1}. Then there exists an interval J such that $I \subset J \subset]0,a[$ and J is an interval of constantness of U^{y_2}.*

Directly by Propositions 1 and 2 we get the following.

Corollary 1. *Let $y_1 \in]a,b[$ be such that $y_1 \sim_{U_s} e$. Then U^{y_1} is a bijection of $[0,a]$ onto itself.*

Now, we will analyze partial functions U_x.

Proposition 3. *Let $x_1, x_2 \in]0,a[$ be arbitrarily chosen. Further assume that $x_2 \in \text{rng}(U_{x_1})$. Then the partial function U_{x_2} is uniquely given by U_{x_1}.*

Just to illustrate Proposition 3, we show the computation of a partial function U_{x_2} given U_{x_1}, where $x_2 \in \text{rng}(U_{x_1})$: assume that $x_2 = U(x_1, y_1)$. Then

$$U(x_2, y_2) = U\left(x_1, U(y_1, y_2)\right). \tag{4}$$

Proposition 4. *Let $x_1 \in]0,a[$ be arbitrarily chosen and $y_1 \in]a,b[$ be a discontinuity point of U_{x_1}. Denote*

$$\lim_{z \to y_{1-}} U_{x_1}(z) = y_\ell, \quad \lim_{z \to y_{1+}} U_{x_1}(z) = y_r.$$

Then the following holds.
(a) *If $y_r < e$ and $T_{U_s} = U_s \upharpoonright]a,e]^2$ is continuous and without idempotent elements on $]a,e[$, then values $U_x(y)$ are uniquely defined for all $x \in]\lim_{z \to a_+} U_{x_1}(z), x_1]$ and $y \leq e$. Moreover, the partial functions U^y are constant on $[y_\ell, y_r]$ for $y \in]a,e[$.*
(b) *If $y_\ell > e$ and $S_{U_s} = U_s \upharpoonright [e,b[^2$ is continuous and without idempotent elements on $]e,b[$, then values $U_x(y)$ are uniquely defined for all $x \in [x_1, \lim_{z \to b_-} U_{x_1}(z)[$ and $y \geq e$. Moreover, the partial functions U^y are constant on $[y_\ell, y_r]$ for $y \in]e,b[$.*

Proposition 5. *Let us take $x_1 \in]a,b[$ and the corresponding partial function U_{x_1}. Assume that U_{x_1} has an interval of constantness I. Then:*
(a) *if $U_{x_1}(y) = x_2 < x_1$ for $y \in I$ then the partial functions U_x are constant on I for all $x \in \text{rng}\, U_{x_1}$, $x \leq x_1$,*
(b) *if $U_{x_1}(y) = x_2 > x_1$ for $y \in I$ then the partial functions U_x are constant on I for all $x \in \text{rng}\, U_{x_1}$, $x \geq x_1$.*

Proposition 6. *Assume that for all $y \in]a,b[$ the partial functions U^y are bijections of $[0,a]$ onto itself. Further assume that for no $x \in]0,a[$ U_x is continuous and strictly increasing. Then for every $x \in]0,a[$ one of the following properties is satisfied.*
(a) *U_x is constant.*
(b) *U_x has countably many intervals of constantness and these intervals induce a partition of $]a,b[$ into a countable system of equivalence classes $\{\mathcal{E}_i\}_{i \in \mathbb{N}}$ such that for all $i, j \in \mathbb{N}$ there exists $k \in \mathbb{N}$ fulfilling formula such that for all $i, j \in \mathcal{I}$ there exists $k \in \mathcal{I}$ fulfilling*

$$(\forall y_1 \in \mathcal{E}_i)(\forall y_2 \in \mathcal{E}_j)(U_s(y_1, y_2) \in \mathcal{E}_k). \tag{5}$$

Proposition 7. *Assume that there exists $x_0 \in]0,a[$ such that U_{x_0} is continuous and strictly increasing. Denote $U_{x_0}(a) = x_1$ and $U_{x_0}(b) = x_2 \leq a$. Then:*
(a) *if U_s is continuous on $]a,b[^2$ then U is continuous on $[x_1, x_2] \times]a,b[$ and on $]a,b[\times [x_1, x_2]$,*
(b) *if U_s is strictly increasing on $]a,b[^2$ then U is strictly increasing on $]x_1, x_2[\times]a,b[$ and on $]a,b[\times]x_1, x_2[$.*

Lemma 3. *Let U_s be representable. Assume that there exists $x_0 \in]0,a[$ such that U_{x_0} is continuous and strictly increasing fulfilling $U_{x_0}(a) = 0$ and $U_{x_0}(b) = a$. Then for all $x \in]0,a[$ the partial functions U_x are continuous and strictly increasing fulfilling $U_x(a) = 0$ and $U_x(b) = a$.*

3 ILLUSTRATIVE EXAMPLES

In all examples we will assume that the neutral element of constructed uninorms is $e = \frac{1}{2}$.

Example 1. In [4] we constructed a conjunctive uninorm which has a zoomed-out representable uninorm U_r on $]\frac{1}{4}, \frac{3}{4}[^2$. On the rectangle $[0, \frac{1}{4}[\times [\frac{1}{4}, \frac{3}{4}]$ the values of U_1 are given by the partial function $U_{\frac{1}{8}}(z) = \dfrac{z - \frac{1}{4}}{2}$. The explicit formula for the uninorm U_1 is the following

$$U_1(x,y) = \begin{cases} 0 & \text{if } \min\{x,y\} = 0 \\ & \text{ or if } \max\{x,y\} \leq \frac{1}{4}, \\ 1 & \text{if } \min\{x,y\} \geq \frac{3}{4}, \\ \frac{1}{4} & \text{if } 0 < \min\{x,y\} \leq \frac{1}{4} \\ & \text{ and if } \max\{x,y\} \geq \frac{3}{4}, \\ & \text{or if } \min\{x,y\} = \frac{1}{4} \\ & \text{ and } \max\{x,y\} > \frac{1}{4}, \\ U_r(x,y) & \text{if } (x,y) \in]\frac{1}{4}, \frac{3}{4}[^2, \\ \max\{x,y\} & \text{if } \frac{1}{4} < \min\{x,y\} < \frac{3}{4} \\ & \text{ and } \max\{x,y\} \geq \frac{3}{4}, \end{cases}$$

and values on $]0, \frac{1}{4}[\times]\frac{1}{4}, \frac{3}{4}[$ and $]\frac{1}{4}, \frac{3}{4}[\times]0, \frac{1}{4}[$ are given by formula (4) and by the partial function $U_{\frac{1}{8}}$.

The uninorm U_1 and its level-set functions of levels $\frac{1}{16}, \frac{1}{8}, \frac{3}{16}$ are sketched on Fig. 1.

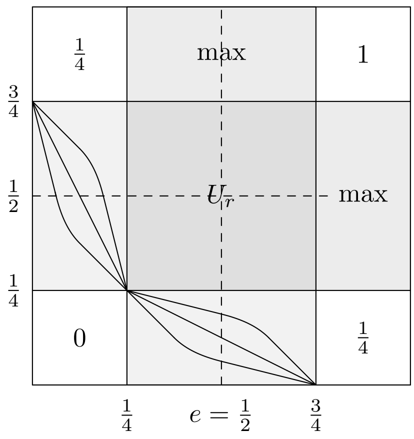


Figure 1: Uninorm U_1

Example 2. Let U_s be a uninorm zoomed-out into $]\frac{1}{4}, \frac{3}{4}[^2$ and continuous on $]\frac{1}{4}, \frac{3}{4}[^2$. Particularly assume that

$$U_s(x,y) = \begin{cases} \min\{x,y\} & \text{if } \min\{x,y\} \le \frac{3}{8}, \\ U_r(x,y) & \text{otherwise,} \end{cases}$$

where U_r is a representable uninorm zoomed-out into the square $]\frac{3}{8}, \frac{3}{4}[^2$. Then the following function $U_2 : [0,1]^2 \to [0,1]$ is a uninorm:

$$U_2(x,y) = \begin{cases} U_s(x,y) & \text{if } (x,y) \in]\frac{1}{4}, \frac{3}{4}[^2, \\ 0 & \text{if } \min\{x,y\} \le \frac{1}{4} \\ & \quad \text{or if } \min\{x,y\} = 0, \\ \frac{1}{4} & \text{if } \min\{x,y\} \le \frac{1}{4} \\ & \quad \text{and } \max\{x,y\} \ge \frac{3}{4}, \\ & \text{or if } \min\{x,y\} = \frac{1}{4} \\ & \quad \text{and } \max\{x,y\} \in]\frac{1}{4}, \frac{3}{4}[, \\ 1 & \text{if } \max\{x,y\} \ge \frac{3}{4}, \\ \frac{3}{4} & \text{if } \min\{x,y\} \in]\frac{1}{4}, \frac{3}{4}[, \\ & \quad \text{and } \max\{x,y\} = \frac{3}{4}, \end{cases}$$

and for $(x,y) \in]0, \frac{1}{4}[\times]\frac{1}{4}, \frac{3}{4}[\cup]\frac{1}{4}, \frac{3}{4}[\times]0, \frac{1}{4}[$ the values of U_2 are defined by $U_{\frac{1}{8}}(y) = \dfrac{y - \frac{1}{4}}{2}$ and by formula (4), and for $(x,y) \in]\frac{3}{4}, 1[\times]\frac{1}{4}, \frac{3}{4}[\cup]\frac{1}{4}, \frac{3}{4}[\times]\frac{3}{4}, 1[$ the values of U_2 are defined by $U_{\frac{7}{8}}(y) = \dfrac{y + \frac{5}{4}}{2}$ and by formula (4). The uninorm U_2 and its level-set functions of levels $\frac{1}{32}, \frac{1}{16}, \frac{25}{32}$ and $\frac{13}{16}$ ('L'-shaped), and of levels $\frac{3}{32}, \frac{1}{8}, \frac{3}{16}, \frac{27}{32}, \frac{7}{8}$ and $\frac{15}{16}$ are sketched on Fig. 2.

Example 3. In this example we first construct by 'paving' a zoomed-out uninorm $U_s :]0.2, 0.8[^2$ which will be strictly increasing on $]0.2, 0.8[^2$ (see [6]) and

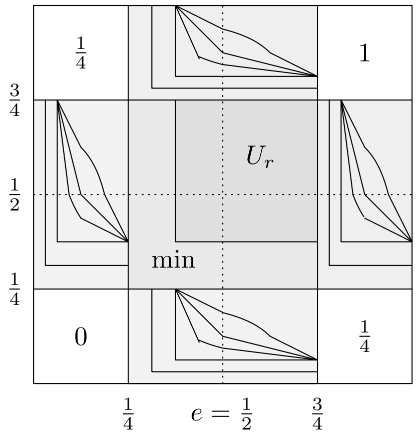


Figure 2: Uninorm U_2

then we will proceed with defining values of a uninorm $U_3 : [0,1]^2$ outside of the square $]0.2, 0.8[^2$.

First, we split the interval $]0.2, 0.8[$ into countably many sub-intervals in the following way:
we set $I_1 =]0.5, 0.6]$, $I_2 =]0.6, 0.65]$, $I_3 =]0.65, 0.7]$ and for $k \in \{4, 5, 6, \dots\}$ we set

$$I_k = \left]0.8 - 0.1 \cdot 2^{-k+4}, 0.8 - 0.1 \cdot 2^{-k+3}\right].$$

Further, $I_0 =]0.4, 0.5]$, $I_{-1} =]0.35, 0.4]$, $I_{-2} =]0.3, 0.35]$ and for $k \in \{-3, -4, -5, \dots\}$ we set

$$I_k = \left]0.2 + 0.1 \cdot 2^{k+2}, 0.2 + 0.1 \cdot 2^{k+3}\right].$$

Let $\varphi_k : I_k \to]0, 1]$ be increasing bijections for all $k \in \mathbb{Z}$. Then we define

$$U_s(x,y) = \varphi_{i+j}^{-1}\left(\varphi_i(x) \cdot \varphi_j(y)\right), \tag{6}$$

where $x \in I_i$ and $y \in I_j$.

Now, we can proceed by defining vales of U_4 on $]0, 0.2] \times]0.2, 0.8[$ (and on $]0.2, 0.8[\times]0, 0.2]$). Again, we split the interval $]0, 0.2[$ into countably many sub-intervals:

$$J_k = \begin{cases}]0.05, 0.15] & \text{if } k = 0, \\ ]0.2 - 0.1 \cdot 2^{-k}, 0.2 - 0.1 \cdot 2^{-k-1}] & \text{if } k > 0, \\ ]0.1 \cdot 2^{k-1}, 0.1 \cdot 2^{k}] & \text{if } k < 0. \end{cases}$$

By $\psi_{m,n} : J_m \to J_n$ we denote increasing bijections of J_m onto J_n where $\psi_{m,m}$ are identities and $\psi_{m,n} \circ \psi_{n,k} = \psi_{m,k}$. We set

$$U_J(x,y) = \psi_{m,n+m}(x) \quad \text{for } x \in J_m \text{ and } y \in I_n.$$

Similarly we split the interval $]0.8, 1[$ into countably many sub-intervals:

$$L_k = \begin{cases}]0.85, 0.95] & \text{if } k = 0, \\ ]1 - 0.1 \cdot 2^{-k}, 1 - 0.1 \cdot 2^{-k-1}] & \text{if } k > 0, \\ ]0.8 + 0.1 \cdot 2^{k-1}, 0.8 + 0.1 \cdot 2^{k}] & \text{if } k < 0. \end{cases}$$

By $\chi_{m,n} : L_m \to L_n$ we denote increasing bijections of L_m onto L_n where $\chi_{m,m}$ are identities and $\chi_{m,n} \circ \chi_{n,k} = \chi_{m,k}$. We set

$$U_L(x,y) = \chi_{m,n+m}(x) \quad \text{for } x \in L_m \text{ and } y \in I_n.$$

We are now ready to define the uninorm U_3:

$$U_3(x,y) = \begin{cases} U_s(x,y) & \text{if } (x,y) \in]0.2, 0.8[\,^2, \\ 0 & \text{if } \max\{x,y\} \le 0.2 \\ & \text{ or if } \min\{x,y\} = 0, \\ 1 & \text{if } \min\{x,y\} \ge 0.8 \\ & \text{ or if } \max\{x,y\} = 1 \\ & \text{ and } \min\{x,y\} > 0.2, \\ 0.2 & \text{if } \min\{x,y\} \le 0.2 \\ & \text{ and } \max\{x,y\} \ge 0.8, \\ & \text{or if } \min\{x,y\} = 0.2 \\ & \text{ and } \max\{x,y\} \in]0.2, 0.8[, \\ 0.8 & \text{if } \max\{x,y\} 0.8 \\ & \text{ and } \min\{x,y\} \in]0.2, 0.8[, \\ U_J(x,y) & \text{if } (x,y) \in \tilde{J}, \\ U_L(x,y) & \text{if } (x,y) \in \tilde{L}, \end{cases}$$

where $\tilde{J} =]0,0.2] \times]0.2,0.8[\cup]0.2,0.8[\times]0,0.2]$ and $\tilde{L} =]0.8,1[\times]0.2,0.8[\cup]0.2,0.8[\times]0.8,1[$.

In the following example we show the discussed construction in case that the underlying t-norm of the uninorm U_s has 0-divisors.

Example 4. Let U_s be a uninorm zoomed-out into $[0.35, 0.8]^2$ such that

$$U_s(x,y) = \begin{cases} \max\{x+y-\frac{1}{2}, 0.35\} & \text{if } (x,y) \in [0.35,0.5]^2, \\ \max\{x,y\} & \text{if } (x,y) \in]0.5,0.8]^2, \\ \min\{x,y\} & \text{otherwise.} \end{cases}$$

Since $U_s(x,y) = 0.35$ for $(x,y) \in [0.35,0.8]^2$ such that $x + y - \frac{1}{2} \le 0.35$, we have that $U_s(0.35,0.35) = 0.35$ to keep the associativity of U_s. On the other hand, if we set $U_4(x,y) = 0.35$ for all $x \in]0,0.35]$ and all $y \in]0.8,1]$, we get

$$\begin{aligned} 0.35 &= U_4(U_4(x,y), 0.35) = U_4(x, U_4(y, 0.35)) \\ &= U_4(x, 0.35) \le x \end{aligned}$$

and hence the associativity is violated. For this reason we define function $U_4 : [0,1]^2 \to [0,1]$ (except of rectangles $[0,0.35[\times [0.35,0.8] \cup [0.35,0.8] \times [0,0.35[$) by

$$U_4(x,y) = \begin{cases} U_s(x,y) & \text{if } (x,y) \in [0.35,0.8]^2, \\ 0 & \text{if } \max\{x,y\} < 0.35, \\ & \text{ or if } \min\{x,y\} = 0, \\ 0.3 & \text{if } 0 < \min\{x,y\} < 0.3 \\ & \text{ and } \max\{x,y\} > 0.8, \\ \min\{x,y\} & \text{if } \min\{x,y\} \in [0.35, 0.5[\\ & \text{ and } \max\{x,y\} > 0.8, \\ & \text{or if } \min\{x,y\} \in]0.3, 0.35[\\ & \text{and } \max\{x,y\} \ge 0.35, \\ \max\{x,y\} & \text{if } \min\{x,y\} \in [0.5, 0.8] \\ & \text{ and } \max\{x,y\} > 0.8, \\ 1 & \text{if } (x,y) \in]0.8,1]^2. \end{cases}$$

Further we introduce the following partial function $U_{0.15} : [0.35, 0.8] \to [0, 0.3]$

$$U_{0.15}(y) = \begin{cases} 0 & \text{if } y < 0.4, \\ \frac{y}{2} - 0.1 & \text{if } y \in [0.4, 0.8]. \end{cases}$$

The partial function $U_{0.15}$ is discontinuous at 0.4 with the following lateral limits

$$\lim_{z \to 0.4_-} U_{0.15}(z) = 0, \quad \lim_{z \to 0.4_+} U_{0.15}(z) = 0.1.$$

The underlying t-norm T_{U_s} is continuous and without idempotent elements on $]0.35, 0, 5[$. This implies (Proposition 4) that U_4 is uniquely defined on $[0,0.15] \times [0.35,0.5] \cap [0.35,0.5] \times [0,0.15]$. But is still undefined on $]0,0.1[\times]0.5,0.8] \cup]0.5,0.8] \times]0,0.1[$.

In these two rectangles, we may define U_4 as

$$U_4(x,y) = \begin{cases} \min\{x,y\} & \text{if } x \in]0,0,1[\\ & \text{ and } y < 0.8 - 3x, \\ & \text{or if } y \in]0,0,1[\\ & \text{ and } x < 0.8 - 3y, \\ 0.1 & \text{if } x \in]0, 0.1[\\ & \text{ and } y \ge 0.8 - 3x, \\ & \text{or if } y \in]0,0,1[\\ & \text{ and } x \ge 0.8 - 3y. \end{cases}$$

The uninorm U_4 is sketched on Fig. 3. To make this figure better readable we give explicit borders of particular regions for $(x,y) \in [0,0.8] \times [0,0.35[$ and for $(x,y) \in [0,0.35[\times [0,0.8]$.

- $U_4(x,y) = 0$ if $\min\{x,y\} = 0$ or if (x,y) is an inner point of the polygon with vertices $(0,0)$, $(0.5,0)$, $(0.5,0.1)$, $(0.4,0.1)$, $(0.4,0.3)$, $(0.35,0.3)$, $(0.35,0.35)$, $(0.3,0.35)$ $(0.3,0.4)$, $(0.1,0.4)$, $(0.1,0.5)$, $(0,0.5)$.

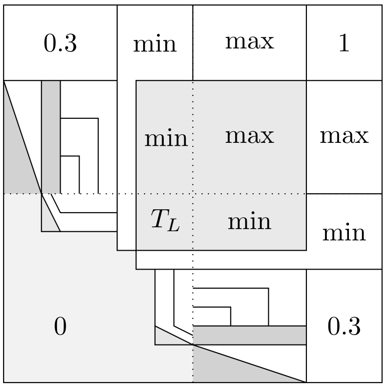


Figure 3: Uninorm U_4

- $U_4(x, y) = 0.1$ if (x, y) is a point of the triangle with vertices $(0.1, 0.5)$, $(0.1, 0.4)$, $(0.15, 0.4)$, respectively $(0.5, 0.1)$, $(0.4, 0.1)$, $(0.4, 0.15)$, or if (x, y) lies on the segment with endpoints $(0.4, 0.15)$, $(0.4, 0.3)$, respectively with endpoints $(0.15, 0.4)$, $(0.3, 0.4)$, or if (x, y) is a point of the triangles with vertices $(0, 0.8)$, $(0.1, 0.5)$, $(0.1, 0.8)$ and $(0.8, 0)$, $(0.8, 0.1)$, $(0.5, 0.1)$ (except of points $(0.8, 0)$ and $(0.8, 0)$ where the value is equal to 0), and on segments with endpoints $(0.1, 0.5)$, $(0.15, 0.4)$,, $(0.15, 0.4)$, $(0.3, 0.4)$, $(0.4, 0.3)$, $(0.4, 0.15)$, $(0.4, 0.15)$, $(0.5, 0.1)$, but except of at points $(0.3, 0.4)$ and $(0.4, 0.3)$ where the value is equal to 0.3.

- $U_4(x, y) = \min\{x, y\}$ if $(x, y) \in [0.1, 0.15] \times [0.3, 0.8] \cup [0.3, 0.8] \times [0.1, 0.15]$ and if (x, y) is an inner point of the triangles with vertices $(0, 0.8)$, $(0, 0.5)$, $(0.1, 0.5)$, and $(0.8, 0)$, $(0.5, 0)$, $(0.5, 0.1)$ (the dark-grey areas on Fig. 3), or if $(x, y) \in]0.3, 0.35[\times [0.35, 0.8] \cup [0.35, 0.8] \times]0.3, 0.35[$.

- The values of U_4 vary within the interval $[0.15, 0.3]$ on the area $]0.15, 0.3] \times [0.5, 0.8] \cup [0.5, 0.8] \times]0.15, 0.3]$ with 'L'-shaped level-set functions, as sketched on Fig. 3.

- On the triangular areas with vertices $(0.1, 0.5)$, $(0.15, 0.5)$, $(0.15, 0.4)$ and $(0.5, 0.1)$, $(0.5, 0.15)$, $(0.4, 0.15)$ the values of U_4 vary within the interval $[0.1, 0.15]$, with level-set functions parallel with the segment with endpoints $(0.1, 0.5)$, $(0.15, 0.4)$, respectively parallel with the segment with endpoints $(0.5, 0.1)$, $(0.4, 0.15)$.

- On rectangles $]0.15, 0.3[\times]0.4, 0.5[$ and $]0.4, 0.5[\times]0.15, 0.3[$ the values of U_4 vary within the interval $]0.1, 0.15[$, with level-set functions parallel with the segment with endpoints $(0.15, 0.4)$, $(0.3, 0.4)$, respectively parallel with the segment with endpoints $(0.4, 0.15)$, $(0.4, 0.3)$.

Acknowledgements

Martin Kalina has been supported from the VEGA grant agency, grant number 2/0059/12, and from the Science and Technology Assistance Agency under contract No. APVV-14-0013.

Pavol Král̆ has been supported from the VEGA grant agency, grant number 1/0647/14.

This paper has been supported by the Project: Mobility - enhancing research, science and education at the Matej Bel University, ITMS code: 26110230082, under the Operational Program Education co-financed by the European Social Fund.

References

[1] T. Calvo, A. Kolesárová, M., Komorníková, R., Mesiar (2002). Aggregation operators: Properties, classes and construction methods. *In: T. Calvo, G. Mayor, R. Mesiar (Eds.) Aggregation operators*, Physica-Verlag Heidelberg, pp. 3-104.

[2] P. Drygaś (2007). On the structure of continuous uninorms. *Kybernetika* **43**, pp. 183-196.

[3] J. Fodor, B. De Baets (2012). A single-point characterization of representable uninorms. *Fuzzy Sets and Systems* **202**, pp. 89-99.

[4] D. Hliněná, M. Kalina, P. Král̆ (2014). Pre-orders and orders generated by conjunctive uninorms. *15th Int. Conference IPMU 2014, Proceedings, Part III, Montpellier, France*, Springer, Heidelberg, pp. 307–316

[5] S. Hu, Z. Li (2001). The structure of continuous uni-norms. *Fuzzy Sets and Systems* **124**, pp. 43-52.

[6] M. Kalina, P. Král̆ (2015). Construction of commutative and associative operations by paving. *Proceeding of IFSA-EUSFLAT Conference 2015*, Atlantis-Press, Gijon, in press.

[7] F. Karaçal,M. N. Kesicioğlu (2011). A t-partial order obtained from t-norms. *Kybernetika* **47**, pp 300-314.

[8] E. P. Klement, R. Mesiar, and E. Pap (2000). Triangular Norms. *Springer*, 1st edition.

[9] R. Mesiar (1995). Choquet-like integrals. *J. Math. Anal. Appl.* **194**, pp. 477-488.

[10] B. Schweizer and A. Sklar (1983). Probabilistic Metric Spaces, *North Holland*, New York.

[11] R. R. Yager and A. Rybalov (1996). Uninorm aggregation operators. *Fuzzy Sets and Systems* **80**, pp. 111-120.

DIFFERENT KINDS OF GENERALIZED CONVEXITY RELATED TO AGGREGATION OF FUZZY SETS

Esteban Induráin
Department of Mathematics
Public University of Navarre, Spain
steiner@unavarra.es

Vladimír Janiš
Department of Mathematics
Faculty of Natural Sciences
Matej Bel University
Slovak Republic
vladimir.janis@umb.sk

Susana Montes
Department of Statistics and O.R.
University of Oviedo, Spain
montes@uniovi.es

Summary

We analyze the existence of fuzzy sets of a universe that are convex with respect to certain particular classes of fusion operators that merge two fuzzy sets. In addition, we study aggregation operators that preserve various classes of generalized convexity on fuzzy sets.

We study the existence of such sets with respect to different classes of aggregation operators (the corresponding functions F), and preserving F-convexity under aggregation of fuzzy sets. Among those typical classes, triangular norms T will be analyzed, giving rise to the concept of norm convexity or T-convexity, as a particular case of F-convexity.

Other different kinds of generalized convexities will also be discussed as a by-product.

Keywords: Fuzzy sets, generalized convexities, fusion operators, triangular norms, real line.

1 INTRODUCTION

Convexity, as one of the most important notions in geometry, has been studied thoroughly from different points of view and has been generalized in different ways. One of the most important generalizations is based on its crucial property, namely, convexity is preserved under set intersection. Based on that property, systems of subsets of a given set, that define a structure called a "generalized convexity" have been defined and studied in depth in [7].

Our aim is to study convexity for fuzzy sets keeping in mind the classical geometrical interpretation of convex sets in an Euclidean space (e.g. the real plane, the real space, etc.), where for each pair of points of a convex set the whole line segment that joins them also belongs to that set. However, for fuzzy sets we have to specify the notion of the membership. The unifying idea, in our considerations and approach in this manuscript, will be the fact that the grade of membership for the points on a line segment that joints two points depends on the grade of memberships of those two boundary points.

We will also pay attention to the problem of characterizing the existence of such fuzzy sets, depending on functions of two variables given a priori. Furthermore, we will also analyze conditions under which suitable functions create a generalized convexity. We will formulate our main results not only for the intersection of fuzzy sets, but for an arbitrary aggregation of fuzzy sets.

2 PREVIOUS CONCEPTS AND RESULTS ABOUT DIFFERENT KINDS OF CONVEXITIES

We start by recalling the standard definition of a fuzzy set.

Definition 1 ([13]). Let X be a nonempty set, usually called the *universe*. A **fuzzy set** A in X is defined by means of a map $\mu_A : X \to [0, 1]$. The map μ_A is said to be the *membership function* of A.

The *support* of A is the crisp set $Supp(A) = \{t \in X : \mu_A(t) \neq 0\} \subseteq X$, whereas the *kernel* of A is the crisp set $Ker(A) = \{t \in X : \mu_A(t) = 1\} \subseteq X$. The fuzzy set A is said to be *normal* provided that it has nonempty kernel.

Given $\alpha \in (0, 1]$, the crisp subset of X defined by $A_\alpha = \{t \in X : \mu_A(t) \geq \alpha\}$ is said to be the *α-cut (level set)* of the fuzzy set A.

In the literature that deals with fuzzy sets, perhaps the

ISBN: 978–83–8012–519–3

most common definition for the concept of a convexity (as a matter of fact, usually called "quasi-convexity") has been introduced in [1] as follows:

Definition 2. Let X be a linear space. A fuzzy subset A of the universe X is said to be **quasi-convex** if for all $x, y \in X,\ \lambda \in [0,1]$ it holds true that

$$\mu_A(\lambda x + (1-\lambda)y) \geq \min\{\mu_A(x), \mu_A(y)\},$$

where μ_A stands here for the membership function of the fuzzy set A.

In models arising in fuzzy logic, the minimum represents the classical conjunction. From this point of view Definition 2 can be read as the statement – if x **and** y are in the fuzzy set A then any point between them is also in A. However, if we use a different model for the conjunction, then the connective **and** is represented by some triangular norm. This leads to the notion of *T-convexity* or *convexity with respect to a triangular norm T*, discussed later.

Another similar concept may by inspired by ideas from [12]. Here the notion of a weakly convex fuzzy set has been defined in the following way:

Definition 3. A fuzzy subset A of a linear space X is said to be **weakly quasi-convex** if for all x, y that belong to the support of μ_A there exists $\lambda \in (0,1)$ such that $\mu_A(\lambda x + (1-\lambda)y) \geq \min\{\mu_A(x), \mu_A(y)\}$.

The condition of weak quasi-convexity is mild, not too restrictive (see [6]). However, its underlying idea can be developed further. Roughly speaking, we keep in mind that the value of the membership function at an "inner point" may depend on the values that it takes at the "endpoints". This can be interpreted, defined and/or understood in a more general way than the (more restrictive) one introduced in Definition 2 and Definition 3.

As it has already been mentioned, we will deal with systems preserving convexity, thus we recall the definition of a generalized convexity from [7].

Definition 4. A system $\mathcal{C}$ of subsets of the universe X for which $\emptyset$ and X belong to $\mathcal{C}$ and $\mathcal{C}$ is closed under arbitrary intersections is a **generalized convexity** on X.

As we will work with fuzzy sets, the system of fuzzy subsets of the universe fulfilling the properties from Definition 4 we will also denote as a generalized convexity on X.

3 CONVEXITIES WITH RESPECT TO TRIANGULAR NORMS AND AGGREGATION OPERATORS IN TWO VARIABLES

The notion of quasi-convexity has been widely studied and applied. However, it could still be too restrictive in several situations, especially in frameworks arising from fuzzy logic. In those contexts, it is typical to find models in which a triangular norm (a t-norm) other than the minimum is used. By this reason, the notion of convexity with respect to triangular norms (or T-convexity, for short) was launched in [5], as follows:

Definition 5. Let X be a linear space and let T be a t-norm. A fuzzy set A of the universe X is said to be **T-convex** if for all $x, y \in X,\ \lambda \in [0,1]$ it holds that $\mu_A(\lambda x + (1-\lambda)y) \geq T(\mu_A(x), \mu_A(y))$.

Notice that here, the triangular norm T has assumed the role of the minimum. In Definition 5, T could be any triangular norm. Nevertheless, both concepts, namely quasi-convexity and T-convexity are close and deeply related. They coincide in the case of normal fuzzy sets as it is proven in the next proposition.

Proposition 1. *Let X be a linear space. Let A be a fuzzy set of the universe X. If A is quasi-convex, then it is T-convex for any t-norm T. Moreover, if A is normal, the converse also holds true.*

Proof. Since the minimum t-norm is the biggest triangular norm, we have that $T \leq \min$ holds for any t-norm T. Therefore, quasi-convexity trivially implies T-convexity, for any triangular norm T. Conversely, if A is normal, then there exists at least one element $a \in X$ such that $\mu_A(a) = 1$. Thus, for any pair $(x, y) \in X \times X$ and any linear combination of its coordinates x, y, that is, $\lambda x + (1-\lambda)y$, it follows that $\lambda x + (1-\lambda)y \leq a$ or, alternatively, $\lambda x + (1-\lambda)y \geq a$. Let us suppose that $x \leq y$. Then we have

$$\mu_A(\lambda x + (1-\lambda)y) \geq T\{\mu_A(x), \mu_A(a)\}$$

or

$$\mu_A(\lambda x + (1-\lambda)y) \geq T\{\mu_A(a), \mu_A(y)\}.$$

Since $\mu_A(a) = 1$ and 1 is the neutral element of any t-norm, we conclude that

$$\mu_A(\lambda x + (1-\lambda)y) \geq \min\{\mu_A(x), \mu_A(y)\}.$$

□

While T-convexity may reflect the use of a particular t-norm T playing the role of the conjunction in a certain fuzzy logic model, notice that none of the special properties of a triangular norm has been mentioned in

Definition 5. Indeed, only one of such classical properties has been used in the proof of Proposition 1. This suggests the study of convexity in an even more general form, in which we will introduce a further generalization of the concept of T-convexity.

The following results are for fuzzy subsets of the real line. They could be easily generalized for $\mathbb{R}^n$.

Definition 6. Let $F : [0,1]^2 \to [0,1]$ be an arbitrary mapping. A fuzzy subset A of the real line is said to be **convex with respect to the map** F (or F-**convex**), if for each $x, y, z \in \mathbb{R}$ such that $x \le y \le z$ it holds that $\mu_A(y) \ge F(\mu_A(x), \mu_A(z))$, where μ_A is the membership function of the fuzzy set A.

Similarly to Definition 3 we can also consider here a linear combination of the points x, z in place of y.

For an infinite collection of fuzzy sets we will have to assume the upper semicontinuity of F, i.e. fulfilling

$$F(\inf_{i\in A} \alpha_i, \inf_{i\in A} \beta_i) = \inf_{i\in A} F(\alpha_i, \beta_i)$$

for an arbitrary index set A. However, this requirement on F will be explicitly mentioned when necessary.

Clearly for $F(\alpha, \beta) = \min\{\alpha, \beta\}$ we have the usual quasi-convexity, while by replacing F by a triangular norm T we obtain the notion of T-convexity introduced above.

In the following proposition we show that the only reasonable results we can obtain appear for maps F accomplishing that $F(\alpha, \beta) \le \min\{\alpha, \beta\}$ for every $0 \le \alpha, \beta \le 1$ (we write $F \le \min$, for short). Indeed, this happens even in the case in which we require the fulfillment of the inequality that arises in Definition 6 only for inner points of the interval $[x, z]$.

Proposition 2. *Let $F : [0,1]^2 \to [0,1]$ be an arbitrary mapping. Assume that $F(\alpha, \beta) > \min\{\alpha, \beta\}$ holds true for every $0 \le \alpha, \beta \le 1$. Then, there is no fuzzy subset A of $\mathbb{R}$ with the property $\mu_A(y) \ge F(\mu_A(x), \mu_A(z))$ for all $x, y, z \in [a, b], x < y < z$, where $[a, b]$ is an interval of real numbers.*

Proof. Let μ_A be the membership function of a fuzzy set with the required property for some F, where $F >$ min. Let $a, b \in \mathbb{R}, a < b$. Consider the value

$$S = \sup\{\mu_A(x), x \in [a, b]\}.$$

Suppose first that the supremum S is attained, i.e. there is a point $z \in [a, b]$ such that $\mu(z) = S$. Clearly such z is unique, because of the condition for μ_A: notice that if there would exist two such points $z_1 \ne z_2$ (suppose $z_1 < z_2$), then for any $y \in [z_1, z_2]$ we would have

$$\mu_A(y) \ge F(\mu_A(z_1), \mu_A(z_2)) >$$

$$> \min\{\mu_A(z_1), \mu_A(z_2)\} = S$$

which is a contradiction. Hence, either $z \ne a$ or $z \ne b$. Suppose that $z > a$. Take arbitrary points x, y such that $a \le x < y < z$. Then it follows that

$$\mu_A(y) \ge F(\mu_A(x), \mu_A(z)) > \min\{\mu_A(x), S\} = \mu_A(x).$$

We can see that μ_A is increasing on $[a, z)$. From the condition on the membership function μ_A we also see that it is discontinuous (from the right) at each $x \in [a, z)$. This is a contradiction to Froda's theorem (see e.g. [3, 8] or [11], Problem 7 in Ch. 20), claiming that the set of discontinuity points for a monotone function is at most countable.

Now let us suppose that S is not attained. Then there is a sequence $\{x_n\}_{n=1}^{\infty}$ in $[a, b]$ such that $\lim_{n\to\infty} \mu_A(x_n) = S$. Taking into account the topological properties of sequences in closed and bounded (i.e compact) intervals (see e.g. [2]) we may assume without loss of generality that this sequence is increasing and $\lim_{n\to\infty} x_n = y$. By the same argumentation as above we see that the sequence $\{\mu_A(x_n)\}_{n=1}^{\infty}$ is also increasing. Using Froda's theorem again we see that for each $n \in N$ the set

$$\{x \in [a, y]; \mu_A(x) < \mu_A(x_n)\}$$

is countable. Hence also the set

$$\cup_{n=1}^{\infty} \{x \in [a, y]; \mu_A(x) < \mu_A(x_n)\} =$$

$$\{x \in [a, y]; \mu_A(x) < S\}$$

is countable. So, there should be at least one (in fact uncountably many ones) point $z \in [a, y)$ such that $\mu_A(z) = S$ which is a contradiction. Hence no fuzzy set with required property can exist. □

Example 1. Hence, as a consequence of Proposition 2, reasonable results can be achieved only for functions F that satisfy the additional condition $F \le$ min. As an example we can consider the fuzzy set A whose membership function μ_A is given as follows:

$$\mu_A(x) = \begin{cases} \frac{1}{2}(x-1)^2 + \frac{1}{2}, & x \in [0, 2], \\ 0, & x \in \mathbb{R} \setminus [0, 2]. \end{cases}$$

This is an F-convex fuzzy subset of the real line, where $F(\alpha, \beta) = \frac{1}{2} \min\{\alpha, \beta\}$. Clearly the fuzzy set A is neither quasi-convex nor T-convex for any t-norm T, since $\mu_A(1) = \frac{1}{2} \not\ge T(\mu_A(0), \mu_A(2)) = T(1, 1) = 1$ holds for any t-norm T.

In the sequel, for a fixed mapping $F : [0,1]^2 \to [0,1]$ let us denote by $\mathcal{C}_F$ the system of all F-convex fuzzy sets. As a consequence of the Proposition 2 we get the following result.

Proposition 3. *Let $F : [0,1]^2 \to [0,1]$ be an arbitrary map. If the condition $F(\alpha,\beta) > \min\{\alpha,\beta\}$ holds for every $0 \le \alpha,\beta \le 1$, then the set $\mathcal{C}_F$ is empty, otherwise $\mathcal{C}_F \neq \emptyset$.*

Proof. The emptiness of $\mathcal{C}_F$ for $F > \min$ follows from the Proposition 2, while for $F \le \min$ at least the fuzzy set $\mu_A = 0$ belongs to $\mathcal{C}_F$, because, in such case $F(0,0) = 0$. □

In order to study generalized convexities from now on we will work only with mappings F with the property $F(\alpha,\beta) \le \min\{\alpha,\beta\}$ $(0 \le \alpha,\beta \le 1)$. Considering F as an aggregation function, we restrict ourselves to those maps that are conjunctive (see [4]).

From Definition 6 it is also clear that given two maps F, G such that or all $\alpha,\beta \in [0,1]$ it holds that $F(\alpha,\beta) \le G(\alpha,\beta)$, then $\mathcal{C}_G \subseteq \mathcal{C}_F$. Furthermore, we can see that the extreme cases are obtained whenever $F(\alpha,\beta) = 1$ for all $\alpha,\beta \in [0,1]$, so that $\mathcal{C}_F = \{0_{\mathbb{R}}, 1_{\mathbb{R}}\}$, as well as for $F(\alpha,\beta) = 0$ for all $\alpha,\beta \in [0,1]$, so that $\mathcal{C}_F = \mathcal{F}(X)$ (the system of all fuzzy subsets of X). Observe also that, for $F(\alpha,\beta) = \min\{\alpha,\beta\}$ the set $\mathcal{C}_F$ is exactly the system of all quasi-convex fuzzy subsets of $\mathbb{R}$.

Perhaps the most important property of classical convex sets is that they create a generalized convexity system, or, in other words, convexity is preserved under intersections. Keeping this in mind, we are also interested in conditions under which an intersection of F-convex fuzzy sets is again an F-convex fuzzy set. We will analyze this problem not only for intersections (represented by triangular norms), but also (and more generally) for arbitrary aggregations of fuzzy sets.

As usually, by a (binary) aggregation function on $[0,1]$ we will understand a mapping $A : [0,1]^2 \to [0,1]$ such that $A(0,0) = 0, A(1,1) = 1$ and, in addition, A s monotone in both variables. By an aggregation of fuzzy sets A_1, A_2 whose membership functions are, respectively, μ_{A_1} and μ_{A_2} we understand the fuzzy set B whose membership function is $\mu_B(x) = A(\mu_{A_1}(x), \mu_{A_2}(x))$. We usually denote it as follows: $\mu_B = A(\mu_{A_1}, \mu_{A_2})$

To formulate our following result we recall the notion of domination for real valued mappings of two variables. For more details on aggregation functions see [4], more information on domination can be found e.g. in [10].

Definition 7. Let $F, G : [0,1]^2 \to [0,1]$ denote two arbitrary mappings. Then we say that F **dominates** G $(F \gg G)$ if for any $\alpha_1,\alpha_2,\beta_1,\beta_2 \in [0,1]$ it holds that

$$F(G(\alpha_1,\beta_1), G(\alpha_2,\beta_2)) \ge G(F(\alpha_1,\alpha_2), F(\beta_1,\beta_2)).$$

The next proposition provides a sufficient and necessary condition for the preservation of convexity with respect to a mapping F.

Proposition 4. *Let $F : [0,1]^2 \to [0,1]$ be an arbitrary mapping, let A be a binary aggregation function on $[0,1]$. Then the following are equivalent:*

1. *$A(\mu_{A_1}, \mu_{A_2})$ is F-convex for any F-convex fuzzy subsets A_1, A_2 of the real line,*
2. *A dominates F.*

Proof. Let $F : [0,1]^2 \to [0,1]$ be a mapping, let A be an arbitrary binary aggregation function on the unit interval. Since throughout this proof we will work with F-convex subsets of the real line, due to the result already stated in Proposition 2 we will assume that $F(\alpha,\beta) \le \min\{\alpha,\beta\}$ $(0 \le \alpha,\beta \le 1)$.

First we suppose that for any F-convex fuzzy subsets A_1, A_2 of $\mathbb{R}$ their aggregation function of membership, namely $A(\mu_{A_1}, \mu_{A_2})$ is F-convex. Thus, let $\alpha_1,\alpha_2,\beta_1,\beta_2 \in [0,1]$. Consider a fixed interval $[a,b] \subseteq \mathbb{R}$, and the following fuzzy sets A_1, A_2 such that

$$\mu_{A_1}(x) = \begin{cases} \alpha_1, & x = a, \\ F(\alpha_1,\alpha_2), & x \in (a,b), \\ \alpha_2, & x = b, \end{cases}$$

$$\mu_{A_2}(x) = \begin{cases} \beta_1, & x = a, \\ F(\beta_1,\beta_2), & x \in (a,b), \\ \beta_2 & x = b \end{cases}$$

and $\mu_{A_1}(x) = \mu_{A_2}(x) = 0$ for $x \in \mathbb{R} \setminus [a,b]$.

We observe that both A_1 and A_2 are F-convex (here we make use of the assumption $F \le min$). Hence the fuzzy set $A(\mu_{A_1}, \mu_{A_2})$ is also F-convex. In other words, for any $z \in (a,b)$ it follows that

$$A(\mu_{A_1}, \mu_{A_2})(z) \ge F(A(\mu_{A_1}, \mu_{A_2})(a), A(\mu_{A_1}, \mu_{A_2})(b))$$

which is equivalent to

$$A(\mu_{A_1}(z), \mu_{A_2}(z)) \ge$$
$$F(A(\mu_{A_1}(a), \mu_{A_2}(a)), A(\mu_{A_1}(b), \mu_{A_2}(b)))$$

or

$$A(F(\alpha_1,\alpha_2), F(\beta_1,\beta_2)) \ge F(A(\alpha_1,\beta_1), A(\alpha_2,\beta_2))$$

and thus $A \gg F$.

To prove the converse assume that $A \gg F$. Let A_1, A_2 be arbitrary F-convex fuzzy subsets of the real line. Take $x, y, z \in \mathbb{R}$ such that $x < y < z$. Then we have that

$$\mu_{A_1}(y) \geq F(\mu_{A_1}(x), \mu_{A_1}(z)),$$
$$\mu_{A_2}(y) \geq F(\mu_{A_2}(x), \mu_{A_2}(z))$$

and from the monotonicity of the map A we obtain

$$A(\mu_{A_1}, \mu_{A_2})(y) = A(\mu_{A_1}(y), \mu_{A_2}(y)) \geq$$
$$A(F(\mu_{A_1}(x), \mu_{A_1}(z)), F(\mu_{A_2}(x), \mu_{A_2}(z))) \geq$$
$$F(A(\mu_{A_1}(x), \mu_{A_2}(x)), A(\mu_{A_1}(z), \mu_{A_2}(z)) =$$
$$F(A(\mu_{A_1}, \mu_{A_2})(x), A(\mu_{A_1}, \mu_{A_2})(z)).$$

Therefore the aggregation of μ_{A_1} and μ_{A_2} is also F-convex. □

This proposition shows in fact that the intersection of a finite collection of T-convex fuzzy sets based on the t-norm T is again T-convex, as the minimum t-norm dominates any other t-norm (see [9]).

Proposition 5. *If $F \ll \min$ and F is upper semicontinuous, the system $\mathcal{C}_F$ (assuming intersections based on the minimum t-norm) is a generalized convexity.*

Proof. Clearly the fuzzy sets with membership functions $\mu_A = 0, \mu_B = 1$ belong to $\mathcal{C}_F$. Let Γ be an arbitrary index set, let for each $\gamma \in \Gamma$ the set A_γ be F-convex. Then for arbitrary $x \leq y \leq z$ there is

$$\mu_{A_\gamma}(y) \geq F(\mu_{A_\gamma}(x), \mu_{A_\gamma}(z))$$

for all $\gamma \in \Gamma$. From the upper semicontinuity of F we get

$$\inf_{\gamma\in\Gamma} \mu_{A_\gamma}(y) \geq F(\inf_{\gamma\in\Gamma} \mu_{A_\gamma}(x), \inf_{\gamma\in\Gamma} \mu_{A_\gamma}(z)),$$

hence the standard intersection of the collection $\{A_\gamma\}_{\gamma\in\Gamma}$ is F-convex. □

Another similar condition to achieve the same conclusion is the monotonicity of F, as shown in the next proposition.

Proposition 6. *If $F \leq \min$ and F is upper semicontinuous, then $\mathcal{C}_F$ is a generalized convexity if and only if F is increasing.*

Proof. If F is increasing then by the results from [9] $F \ll \min$ and by Proposition 5 $\mathcal{C}_F$ is a generalized convexity.

Assume, by contradiction, that F fails to be increasing. Then there are $\alpha_1, \alpha_2, \beta_1, \beta_2 \in [0, 1]$ such that $\alpha_1 \leq \beta_1, \alpha_2 \leq \beta_2$, but $F(\alpha_1, \alpha_2) > F(\beta_1, \beta_2)$. Take an arbitrary interval $[a, b] \in \mathbb{R}$ and the fuzzy sets A_1 and A_2 whose membership functions are

$$\mu_{A_1}(x) = \begin{cases} \alpha_1, & x = a, \\ F(\alpha_1, \alpha_2), & x \in (a, b), \\ \alpha_2, & x = b, \end{cases}$$

$$\mu_{A_2}(x) = \begin{cases} \beta_1, & x = a, \\ F(\beta_1, \beta_2), & x \in (a, b), \\ \beta_2 & x = b \end{cases}$$

and $\mu_{A_1}(x) = \mu_{A_2}(x) = 0$ for $x \in \mathbb{R} \setminus [a, b]$. We may notice that both A_1 and A_2 are F-convex fuzzy sets. However, their intersection is the fuzzy set defined by means of the membership function

$$(\mu_{A_1} \cap \mu_{A_2})(x) = \begin{cases} \alpha_1, & x = a, \\ F(\beta_1, \beta_2), & x \in (a, b), \\ \alpha_2, & x = b, \end{cases}$$

which is not F-convex. The reason is that, for any $y \in (a, b)$, we have that

$$(\mu_{A_1} \cap \mu_{A_2})(y) = F(\beta_1, \beta_2) \ngeq F(\alpha_1, \alpha_2) =$$
$$F((\mu_{A_1} \cap \mu_{A_2})(a), (\mu_{A_1} \cap \mu_{A_2})(b)).$$

We conclude that $\mathcal{C}_F$ is not a generalized convexity. □

Finally, in the next proposition we explain the relationship between F-convex fuzzy sets and crisp convex sets (here we consider crisp sets as a special case of fuzzy sets).

Proposition 7. *Let $\mathcal{C}$ denote the system of all crisp convex subsets of the real line $\mathbb{R}$ and let $\mathcal{F} = \{F : [0, 1]^2 \to [0, 1]; F(1, 1) > 0\}$, F being a map. Then $\mathcal{C} = \cap_{F\in\mathcal{F}}(\mathcal{C}_F \cap \{0, 1\}^{\mathbb{R}})$.*

Here we consider crisp convex subsets of the real line (i.e. intervals) represented by their characteristic functions.

Proof. Suppose $C \in \mathcal{C}$. Let $F \in \mathcal{F}$. We will show that C is F-convex. To do so, take $x, y, z \in \mathbb{R}, x \leq y \leq z$.

If $C(y) = 0$, then from its convexity at least one of the values $C(x), C(z)$ should be zero. Hence $F(C(x), C(z)) = 0$ too. If $C(y) = 1$ then the inequality $C(x) \geq F(C(x), C(z))$ is fulfilled for any $F \in \mathcal{F}$. Thus C is F-convex for any F.

Now let C be an F-convex crisp set for any $F \in \mathcal{F}$. Suppose C is not convex. Then there are $x, y, z \in \mathbb{R}$ with $x < y < z$, and such that $C(x) = C(z) = 1, C(y) = 0$. Consider a mapping $F \in \mathcal{F}$. From the F-convexity of C we have

$$0 = C(y) \geq F(C(x), C(z)) = F(1, 1) > 0$$

which is a contradiction. This concludes the proof. □

4 SUGGESTIONS FOR FURTHER RESEARCH

A former suggestion for a further development of these ideas could be trying to avoid working with linear spaces, and defining generalized convexities (see [7]) on a nonempty set U, called universe, as suitable mappings $f : U \times U \times [0,1] \to U$ that accomplish certain conditions (e.g.: $f(x,y,\alpha) = f(y,x,1-\alpha)$ for every $x,y \in U$ and $\alpha \in [0,1]$) so that $f(x,y,\alpha)$ could play the role of the point "$\alpha \cdot x + (1-\alpha) \cdot y$" that is typical in the case in which U is a linear space.

Another suggestion could be trying to work with some more general kinds of fuzzy sets, as, for instance, those in which the membership function takes values in a lattice, instead of the unit interval $[0,1]$.

Acknowledgements

The first author acknowledges the partial support of the research Projects MTM2012-37894-C02-02 and TIN2013-40765-P (Spain) of the Spanish Ministry of Economy and Competitiveness.

References

[1] E. Ammar, J. Metz (1992). On fuzzy convexity and parametric fuzzy optimization, *Fuzzy Sets and Systems* **49**, pp 135–141.

[2] J. Dugundji (1996). Topology. Allyn and Bacon. Boston.

[3] A. Froda (1929). Sur la Distribution des Propriétés de Voisinage des Fonctions de Variables Réelles, Thèse, Hermann, Paris.

[4] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap (2009). Aggregation Functions, Encycopedia of Mathematics and its Applications, Cambridge University Press.

[5] T. Iglesias, I. Montes, V. Janiš, S. Montes (2012). T-convexity for lattice-valued fuzzy sets, Proceedings of the ESTYLF Conference, pp 26–30.

[6] V. Janiš, S. Montes, T. Iglesias (2012). Aggregation of Weakly Quasi-convex Fuzzy Sets, Communications in Computer and Information Science : 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU 2012, Catania, Italy, July 9-13, 2012, pp 353-359.

[7] J.V. Llinares, (2002). Abstract convexity, some relations and applications, *Optimization* **51**(6), pp 797–818.

[8] F. Riesz, B. Sz. Nagy (1995). Functional Analysis. Ungar. New York.

[9] S. Saminger-Platz, R. Mesiar, U. Bodenhofer (2002). Domination of aggregation operators and preservation of transitivity, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **10** (Suppl.) pp 11–35.

[10] S. Saminger-Platz, R. Mesiar, D. Dubois (2007). Aggregation operators and Commuting, *IEEE Transactions on Fuzzy Systems*, **15** (6), pp 1032–1045.

[11] M. Spivak (1967). Calculus. W. A. Benjamin Inc. New York.

[12] Y.R. Syau (2001). Some properties of weakly convex fuzzy mappings *Fuzzy Sets and Systems* **123**, pp 203–207.

[13] L. A. Zadeh (1965). Fuzzy sets, *Information and Control* **8**, pp 338–353.

FUNCTIONAL TRANSFORMATIONS OF SPECIAL CLASSES OF AGGREGATION FUNCTIONS

Anna Kolesárová
Fac. of Chem. and Food Technology,
Slovak University of Technology,
812 37 Bratislava, Slovakia
anna.kolesarova@stuba.sk

Radko Mesiar and Andrea Stupňanová
Fac. of Civil Engineering,
Slovak University of Technology,
810 05 Bratislava, Slovakia,
{radko.mesiar,andrea.stupnanova}@stuba.sk

Summary

Transformations of special classes of binary aggregation functions based on real functions of three variables are studied. Our attention is focused on functional transformations of averaging and conjunctive aggregation functions as well as on transformations of copulas based on quadratic polynomials and their stochastic interpretation.

Keywords: Aggregation function, copula, transformation.

1 INTRODUCTION

In this contribution we will deal with function-based transformations of aggregation functions. Recall that a function $A\colon [0,1]^n \to [0,1]$ is called an n-ary aggregation function ($n \in \mathbb{N}$, $n \geq 2$) whenever it is increasing in each variable and satisfies the boundary conditions $A(0,\ldots,0) = 0$ and $A(1,\ldots,1) = 1$. We will mostly deal with binary aggregation functions ($n = 2$). The set of all binary aggregation functions will be denoted by $\mathcal{A}$, and when no confusion can arise, we will call them simply aggregation functions.

Considering some additional properties, special classes of aggregation functions can be obtained, such as weighted arithmetic means, OWA operators, t-norms, copulas, uninorms, etc. For more details on these classes of aggregation functions we refer the reader, e.g., to [1, 2, 6, 8, 9, 14].

If we consider a fixed binary aggregation function A, at each point $(x,y) \in [0,1]^2$ we have three values as a basic information: the values x, y and $z = A(x,y)$, all from the unit interval $[0,1]$. If $f\colon [0,1]^3 \to [0,1]$ is a function of three variables, then a composite function $A_f\colon [0,1]^2 \to [0,1]$,

$$A_f(x,y) = f(x,y,A(x,y)) \tag{1}$$

is well defined. Clearly, if f is a ternary aggregation function then A_f is a binary aggregation function for any $A \in \mathcal{A}$. On the other hand, for some special aggregation functions A, A_f can belong to $\mathcal{A}$ though f is not a ternary aggregation function. For example, this is true in the case of the class $\mathcal{L}$ of all binary 1-Lipschitz aggregation functions and the function $f\colon [0,1]^3 \to [0,1]$,

$$f(x,y,z) = x + y - z. \tag{2}$$

Recall that an aggregation function $A\colon [0,1]^2 \to [0,1]$ is said to be *1-Lipschitz* if for all $x_1, x_2, y_1, y_2 \in [0,1]$,

$$|A(x_1,y_1) - A(x_2,y_2)| \leq |x_1 - x_2| + |y_1 - y_2|. \tag{3}$$

By [10], for each $A \in \mathcal{L}$ and f defined in (2), the function A_f is a 1-Lipschitz aggregation function. However, as f is decreasing in the third variable, f is not an aggregation function.

The aim of this contribution is to study under what constraints imposed on f, the function A_f defined by (1) is an aggregation function from some special subclass $\mathcal{B} \subset \mathcal{A}$ for each $A \in \mathcal{B}$. In other words, our aim is, for a chosen class $\mathcal{B} \subset \mathcal{A}$, to characterize functions f with the property that the assignment $A \mapsto A_f$ is a $\mathcal{B} \to \mathcal{B}$ mapping. The contribution is organized as follows. In the next section, some special classes of aggregation functions, which will be discussed later, are characterized. In Section 3, we introduce some general properties of functions f generating the studied transformations. In Section 4, we discuss function-based transformations of averaging aggregation functions, including transformations of some special subclasses of averaging aggregation functions. Similarly, in Section 5, transformations of conjunctive aggregation functions are studied, including transformations of some their special subclasses. Section 6 is devoted to the quadratic constructions of copulas and their stochastic interpretation. Finally, some concluding remarks are provided.

2 SOME SPECIAL CLASSES OF AGGREGATION FUNCTIONS

This section contains a classification of aggregation functions and a brief summary of special classes of aggregation functions which we will work with. We first recall that for each $A \in \mathcal{A}$, its dual $A^d \in \mathcal{A}$ is given by

$$A^d(x,y) = 1 - A(1-x, 1-y)$$

and clearly, for each $A \in \mathcal{A}$, $\left(A^d\right)^d = A$.

The basic classification of aggregation functions proposed by D. Dubois and H. Prade [4] distinguishes

- *conjunctive aggregation functions,* $\mathcal{C}on = \{A \in \mathcal{A} \mid A \leq Min\}$,
- *disjunctive aggregation functions,* $\mathcal{D}is = \{A \in \mathcal{A} \mid A \geq Max\}$,
- *averaging aggregation functions,* $\mathcal{A}v = \{A \in \mathcal{A} \mid Min \leq A \leq Max\}$,
- *mixed aggregation functions,* $\mathcal{R} = \mathcal{A} \setminus (\mathcal{C}on \cup \mathcal{D}is \cup \mathcal{A}v)$.

Note that the classes $\mathcal{A}v$ and $\mathcal{R}$ are closed under duality, while $A \in \mathcal{C}on$ if and only if $A^d \in \mathcal{D}is$. This fact allows us to rewrite straightforwardly all results valid for conjunctive aggregation functions into the corresponding results for disjunctive aggregation functions.

In this contribution we will work with special classes of aggregation functions, namely, with the class of all:

• semicopulas $\mathcal{S}$: aggregation functions with neutral element $e = 1$ (i.e., satisfying the property $A(x,1) = A(1,x) = x$ for each $x \in [0,1]$); $\mathcal{S} \subset \mathcal{C}on$.

• quasicopulas $\mathcal{Q}$: $\mathcal{Q} = \mathcal{L} \cap \mathcal{S}$, i.e., 1-Lipschitz aggregation functions with neutral element $e = 1$;

• copulas $\mathcal{C}$: quasicopulas characterized by supermodularity

$$A(\mathbf{x} \vee \mathbf{y}) + A(\mathbf{x} \wedge \mathbf{y}) \geq A(\mathbf{x}) + A(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in [0,1]^2$.

• weighted arithmetic means $\mathcal{W}$: aggregation functions characterized by additivity. Recall that $A \in \mathcal{W}$ if and only if $A(x,y) = cx + (1-c)y$ for some fixed $c \in [0,1]$ and all $x, y \in [0,1]$. $\mathcal{W} \subset \mathcal{A}v$.

• OWA operators $\mathcal{OWA}$: aggregation functions characterized by symmetry and comonotone additivity. Recall that $A \in \mathcal{OWA}$ if and only if

$$A(x,y) = cMin(x,y) + (1-c)Max(x,y)$$

for some fixed $c \in [0,1]$ and all $x, y \in [0,1]$. It holds that $\mathcal{OWA} \subset \mathcal{A}v$.

• aggregation functions having a neutral element $e \in]0,1[$. This class will be denoted by $\mathcal{A}_e$. Distinguished members of $\mathcal{A}_e$ are, for example, uninorms U_e. It holds that $\mathcal{A}_e \subset \mathcal{R}$.

More details on the mentioned classes of aggregation functions and their properties can be found, e.g., in [1, 2, 5, 6, 8].

3 FUNCTIONAL TRANSFORMATIONS OF AGGREGATION FUNCTIONS

For any subclass $\emptyset \neq \mathcal{B} \subseteq \mathcal{A}$, denote by $\mathcal{F}_{\mathcal{B}}$ the set of all functions $f\colon [0,1]^3 \to [0,1]$ such that $A_f \in \mathcal{B}$ for each $A \in \mathcal{B}$, i.e. the mapping $A \mapsto A_f$, where A_f is given by (1), defines a transformation on the class $\mathcal{B}$. Clearly, $\mathcal{F}_{\mathcal{B}}$ is closed under the composition $*$ given by

$$(f * g)(x,y,z) = f(x,y,g(x,y,z)).$$

Moreover, if the class $\mathcal{B}$ is convex, then $\mathcal{F}_{\mathcal{B}}$ is also a convex set.

Due to the boundary conditions of aggregation functions, for an arbitrary $\mathcal{B}$ and each $f \in \mathcal{F}_{\mathcal{B}}$ it holds that $f(0,0,0) = 0$ and $f(1,1,1) = 1$. Let $\mathcal{F}$ denote the set of all functions $f\colon [0,1]^3 \to [0,1]$ satisfying these two properties. Note that the function f given by (2) belongs to $\mathcal{F}$. Next, let $\mathcal{A}_3$ denote the class of all ternary aggregation functions. Clearly, $\mathcal{A}_3 \subset \mathcal{F}$, and as was already mentioned, $\mathcal{A}_3 \subset \mathcal{F}_{\mathcal{A}}$. With the previous notation, we can formulate the following stronger claim.

Proposition 1 *It holds that $\mathcal{F}_{\mathcal{A}} = \mathcal{A}_3$.*

The only function $f \in \mathcal{F}$ such that $f \in \mathcal{F}_{\mathcal{B}}$ for any subclass $\emptyset \neq \mathcal{B} \subseteq \mathcal{A}$, is the third projection given by $f(x,y,z) = z$. In that case, $A_f = A$ for any $A \in \mathcal{A}$. On the other hand, if for $A \in \mathcal{A}$, we write f_A for the function $f_A\colon [0,1]^3 \to [0,1]$ defined by $f_A(x,y,z) = A(x,y)$, then $f_A \in \mathcal{F}$. Moreover, for any subclass $\emptyset \neq \mathcal{B} \subseteq \mathcal{A}$ it holds that

$$\emptyset \neq \{f_A \mid A \in \mathcal{B}\} \subseteq \mathcal{F}_{\mathcal{B}}.$$

In some cases, the properties of functions f are important on some subdomain of $[0,1]^3$ only. For example, if $A \in \mathcal{L}$ then for each $(x,y) \in [0,1]^2$ it holds [10] that

$$\max\{0, x+y-1\} \leq A(x,y) \leq \min\{1, x+y\},$$

and hence the properties of A_f only depend on the properties of f on the subdomain D,

$$D = \{(x, y, z) \in [0, 1]^3 \mid \max\{0, x + y - 1\} \le z \le \min\{1, x + y\}\}. \quad (4)$$

Proposition 2 *Let $f \in \mathcal{F}$. Then $f \in \mathcal{F}_{\mathcal{L}}$ if and only if f is 1-Lipschitz on the set D (defined in (4)) and satisfies the inequalities*

$$\frac{\partial f}{\partial x} \ge 0,\ \frac{\partial f}{\partial y} \ge 0,\ 0 \le \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} \le 1,$$

$$\text{and}\ \ 0 \le \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z} \le 1$$

at each point from the interior of D where the partial derivatives exist.

Note that the function f given by (2) is in $\mathcal{F}_{\mathcal{L}}$ and satisfies the properties

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 1 \text{ and } \frac{\partial f}{\partial z} = -1$$

at any interior point of D. In general, if $f \in \mathcal{F}$ is a linear function, i.e., if $f(x, y, z) = ax + by + cz$ for some real constants constrained by the condition $a + b + c = 1$, then $f \in \mathcal{F}_{\mathcal{L}}$ if and only if $a, b \in [0, 1]$ and $c = 1 - a - b$. In what follows, we will denote the set of all such linear functions by $\mathcal{F}^*_{\mathcal{L}}$. Note that $\mathcal{F}_{\mathcal{L}} \setminus \mathcal{F}^*_{\mathcal{L}} \neq \emptyset$. As an example of a non-linear function $f \in \mathcal{F}_{\mathcal{L}}$ we can consider $f \in \mathcal{F}$, given by $f(x, y, z) = z(x + y - z)$.

4 FUNCTIONAL TRANSFORMATIONS OF AVERAGING FUNCTIONS

Let $\mathcal{B}$ be any class of averaging functions. Considering a transformation of an aggregation function $A \in \mathcal{B}$ based on a function $f \in \mathcal{F}$, the values of f at the points which are out of the subdomain

$$D_{\mathcal{A}v} = \left\{(x, y, z) \in [0, 1]^3 \mid \min\{x, y\} \le z \le \max\{x, y\}\right\}$$

are not entering into the construction (1). Let $\mathcal{A}v_3$, $\mathcal{W}_3$, $\mathcal{OWA}_3$ denote the set of all ternary averaging functions, weighted arithmetic means and OWA operators, respectively. For a subdomain $\emptyset \neq D \subset [0, 1]^3$ and a class $\mathcal{D} \subseteq \mathcal{A}_3$, let

$$\mathcal{F}_{D,\mathcal{D}} = \{f \in \mathcal{F} \mid f|D = A|D \text{ for some } A \in \mathcal{D}\}.$$

Then, keeping the previous notation, we can formulate the following results.

Proposition 3 *For averaging aggregation functions it holds that $\mathcal{F}_{\mathcal{A}v} = \mathcal{F}_{D_{\mathcal{A}v}, \mathcal{A}v_3}$.*

Proposition 4 *For weighted arithmetic means it holds that $\mathcal{F}_{\mathcal{W}} = \mathcal{F}_{D_{\mathcal{A}v}, \mathcal{F}^*_{\mathcal{L}}}$.*

Proposition 5 *For OWA operators it holds that*

$$\mathcal{F}_{\mathcal{OWA}} = \{f \in \mathcal{F} \mid \exists g \in \mathcal{F}_{\mathcal{W}},$$
$$f(x, y, z) = g\left(Min(x, y, z), Max(x, y, z), Med(x, y, z)\right)\}.$$

Example 1 *Consider the function $f \in \mathcal{F}$ given by (2). Then $f \in \mathcal{F}^*_{\mathcal{L}}$ and thus $f \in \mathcal{F}_{\mathcal{W}}$, too. The f-based transformation, which was defined in (1), applied to a weighted arithmetic mean $W_\lambda \in \mathcal{W}$ given by $W_\lambda(x, y) = \lambda x + (1 - \lambda)y$ yields*

$$(W_\lambda)_f\,(x, y) = f(x, y, \lambda x + (1 - \lambda)y)$$
$$= x + y - (\lambda x + (1 - \lambda)y) = (1 - \lambda)x + \lambda y = W_{1-\lambda}(x, y),$$

i.e., $W_\lambda = W_{1-\lambda}$.

Consider the function h induced by f and defined by

$$h(x, y, z) = \min\{x, y\} + \max\{x, y\} - med\{x, y, z\}.$$

Note that $h \in \mathcal{F}_{\mathcal{OWA}}$ and $h|D_{\mathcal{A}v} = f|D_{\mathcal{A}v}$. For an OWA operator OWA_λ given by

$$OWA_\lambda(x, y) = \lambda \min\{x, y\} + (1 - \lambda)\max\{x, y\}$$

it holds that

$$(OWA_\lambda)_h\,(x, y) = \min\{x, y\} + \max\{x, y\}$$
$$-(\lambda \min\{x, y\} + (1 - \lambda)\max\{x, y\}) = OWA_{1-\lambda}(x, y),$$

i.e., $(OWA_\lambda)_h = OWA_{1-\lambda}$.

5 FUNCTIONAL TRANSFORMATIONS OF CONJUNCTIVE AGGREGATION FUNCTIONS

In the case of conjunctive aggregation functions, an important role is played by the domain

$$D_{\mathcal{C}on} = \{(x, y, z) \in [0, 1]^3 \mid z \le \min\{x, y\}\}.$$

Let $\mathcal{K}$ denote the class of all ternary aggregation functions A such that $A \le Med$. Then we have:

Proposition 6 *$\mathcal{F}_{\mathcal{C}on} = \mathcal{F}_{D_{\mathcal{C}on}, \mathcal{K}}$.*

Proposition 7 *For the class $\mathcal{S}$ of all semicopulas it holds that*

$$\begin{aligned}\mathcal{F}_{\mathcal{S}} &= \{f \in \mathcal{F}_{\mathcal{C}on} \mid f(x, 1, x) = x \text{ and } f(1, y, y) \\ &= y \text{ for all } x, y \in [0, 1]\}.\end{aligned}$$

Proposition 8 *For the class $\mathcal{Q}$ of all quasicopulas it holds that*

$$\mathcal{F}_{\mathcal{Q}} \subset \mathcal{F}_{\mathcal{S}} \cap \mathcal{F}_{\mathcal{L}}.$$

Example 2 *Define* $f, g, h\colon [0,1]^3 \to [0,1]$ *by*

$$f(x,y,z) = z\sqrt{xy}, \quad g(x,y,z) = \sqrt{xyz}$$

and

$$h(x,y,z) = z(x+y-z).$$

Then

$f \in \mathcal{F}_{\mathcal{C}on} \setminus \mathcal{F}_{\mathcal{S}}$, $g \in \mathcal{F}_{\mathcal{S}} \setminus \mathcal{F}_{\mathcal{Q}}$, *and* $h \in \mathcal{F}_{\mathcal{Q}}$. *Consider the smallest copula* W, $W(x,y) = \max\{x+y-1, 0\}$. *Note that* W *is also the smallest quasicopula. Then the function* W_f *given by*

$$W_f(x,y) = \sqrt{xy}\max\{x+y-1,0\}$$

is conjunctive, but as $W_f(1,y) = y\sqrt{y}$, $W_f \notin \mathcal{S}$. *Next,* W_g *given by*

$$W_g(x,y) = \sqrt{xy\max\{x+y-1,0\}}$$

is a semicopula, but it is not 1-Lipschitz and thus $W_g \notin \mathcal{Q}$. *Finally,* $W_h = W$, *i.e.,* W_h *is a copula (and hence also a quasicopula).*

6 QUADRATIC TRANSFORMATIONS OF COPULAS

Consider quadratic polynomials $p_{a,b}\colon [0,1]^3 \to \mathbb{R}$ with $a, b \in [0,1]$ defined by

$$p_{a,b}(x,y,z) = axy + bz + (1-a-b)z(x+y-z).$$

Evidently, $p_{a,b}(0,0,0) = 0$ and $p_{a,b}(1,1,1) = 1$, however, the range of $p_{a,b}$ need not be contained in $[0,1]$. If we restrict the domain of $p_{a,b}$ to $D_{\mathcal{C}on}$, i.e., if $z \leq \min\{x,y\}$, then $p_{a,b}(x,y,z) \in [0,1]$ for all $a, b \in [0,1]$, and thus we can transform an arbitrary conjunctive aggregation function A into $A_{p_{a,b}}$, though $p_{a,b}$ need not be from $\mathcal{F}$. Obviously, defining $f_{a,b}\colon [0,1]^3 \to [0,1]$ by

$$f_{a,b} = \min(1, \max(0, p_{a,b})),$$

we can see that $f_{a,b} \in \mathcal{F}$, and $A_{f_{a,b}} = A_{p_{a,b}}$ for any $A \in \mathcal{C}on$.

Proposition 9 *Let* $\mathcal{H} = \{f_{a,b} \mid a, b \in [0,1]\}$. *Then*

$$\mathcal{H} \subset \mathcal{F}_{\mathcal{C}} \cap \mathcal{F}_{\mathcal{Q}} \cap \mathcal{F}_{\mathcal{S}},$$

i.e., for any copula (quasicopula, semicopula) A, *the function* $A_{f_{a,b}}$ *is a copula (quasicopula, semicopula) for any* $a, b \in [0,1]$.

Note that the polynomial class $\mathcal{P} = \{p_{a,b} \mid a, b \in [0,1]\}$ is convex, and it can be viewed as a convex closure of the extremal points $p_{0,0}$, $p_{0,1}$, $p_{1,0}$ and $p_{1,1}$. Note that, due to [11] and [12], for any copula (quasicopula, semicopula) A the corresponding transformations (written by means of the transforming functions $f_{a,b}$) are given by

$$\begin{aligned} A_{f_{0,0}}(x,y) &= A(x,y).(x+y-A(x,y)), \\ A_{f_{0,1}}(x,y) &= A(x,y), \\ A_{f_{1,0}}(x,y) &= xy, \end{aligned}$$

and

$$A_{f_{1,1}}(x,y) = xy - A(x,y).(x+y-1-A(x,y)).$$

Based on the results of [3], in the case of copulas we have the next stochastic interpretation of the above transformations.

Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of continuous random variables (all four being uniformly distributed over $]0,1[$) which are identically distributed, and let the stochastic dependence of their components be described by a copula C. Then:

• The copula $C_{f_{0,0}}$ describes the stochastic dependence structure of the random vector (Z_1, Z_2),

$$(Z_1, Z_2) = \begin{cases} (\min(X_1,X_2), \max(Y_1,Y_2)) \\ \qquad \text{with probability } 1/2, \\ (\max(X_1,X_2), \min(Y_1,Y_2)) \\ \qquad \text{with probability } 1/2. \end{cases}$$

• The copula $C_{f_{0,1}} = C$ describes the stochastic dependence structure of the random vector (X_1, Y_1).

• The copula $C_{f_{1,0}}$ describes the stochastic dependence structure of the random vector (X_1, X_2).

• The copula $C_{f_{1,1}}$ describes the stochastic dependence structure of the random vector (T_1, T_2), given by

$$(T_1, T_2) = \begin{cases} (\min(X_1,X_2), \min(Y_1,Y_2)) \\ \qquad \text{with probability } 1/2, \\ (\max(X_1,X_2), \max(Y_1,Y_2)) \\ \qquad \text{with probability } 1/2. \end{cases}$$

Note also that copulas invariant with respect to the $f_{a,b}$-based transformations are just the Plackett copulas [15], for more details see [11].

7 CONCLUDING REMARKS

We have introduced and discussed function-based transformations of binary aggregation functions. In some cases there is only a trivial identity transformation that preserves the considered classes of aggregation functions. This is, for example, the case of the class $\mathcal{A}_e$, $e \in]0,1[$. In some cases we have obtained interesting original results, for example for quadratic constructions of copulas. There are still several open problems. For example, a complete description of the classes $\mathcal{F}_{\mathcal{Q}}$ and $\mathcal{F}_{\mathcal{C}}$ is still missing.

Acknowledgements

The authors kindly acknowledge the support of the grants VEGA 1/0419/13, VEGA 1/0420/15 and the project APVV-14-0013.

References

[1] G. Beliakov, A. Pradera, and T. Calvo (2007). *Aggregation Functions: A Guide for Practitioners.* Springer, Heidelberg, Berlin, New York.

[2] T. Calvo, A. Kolesárová, M. Komorníková, R. Mesiar (2002). Aggregation Operators: Properties, Classes and Construction Methods. In *Aggregation Operators. New Trends and Applications*, T. Calvo, G. Mayor, R. Mesiar, eds. Physica-Verlag, Heidelberg, pp. 3-107.

[3] A. Dolati, M. Úbeda-Flores (2009). Constructing copulas by means of pairs of order statistics. *Kybernetika* **45**, pp. 992-1002.

[4] D. Dubois, H. Prade (1985). A review of fuzzy set aggregation connectives. *Inform. Sci.* **36**, pp. 85–121.

[5] F. Durante, C. Sempi (2005). Semicopulae. *Kybernetika* **41**, pp. 315–328.

[6] J.C. Fodor, R.R. Yager, and A. Rybalov (1997). Structure of uninorms. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems* **5**, pp. 411–427.

[7] C. Genest, J.J. Quesada Molina, J. A. Rodríguez-Lallena, and C. Sempi (1999). A characterization of quasi-copulas. *J. Multivariate Anal.* **69**, pp. 193-205.

[8] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap (2009). *Aggregation Functions.* Cambridge University Press, Cambridge.

[9] E. P. Klement, R. Mesiar, and E. Pap (2000). *Triangular norms.* Kluwer, Dordrecht.

[10] A. Kolesárová (2003). 1-Lipschitz aggregation operators and quasi-copulas. *Kybernetika* **39**, pp. 615-629.

[11] A. Kolesárová, G. Mayor, and R. Mesiar (2015). Quadratic constructions of copulas. *Inform. Sci.*, submitted.

[12] A. Kolesárová, R. Mesiar (2015). On linear and quadratic constructions of aggregation functions (2015). *Fuzzy Sets and Systems*, doi. 10.1016/fss.2014.04.015.

[13] A. Kolesárová, R. Mesiar, and J. Kalická (2013). On a new construction of 1-Lipschitz aggregation functions, quasi-copulas and copulas. *Fuzzy Sets and Systems* **224**, pp. 19-31.

[14] R. B. Nelsen (1999). *An Introduction to Copulas.* Lecture Notes in Statistics 139, Springer Verlag, New York.

[15] R.L. Plackett (1965). A class of bivariate distributions. *J. Amer. Statist. Assoc.* **60**, pp. 516-522.

[16] R.R. Yager, A. Rybalov (1996). Uninorm aggregation operators. *Fuzzy Sets and Systems* **80**, pp. 111–120.

PERTURBATIONS OF TAIL DEPENDENCIES OF COPULAS

Jozef Komorník
Faculty of Managment,
Comenius University,
Odbojárov 10, P.O.Box 95, 820 05 Bratislava,
Slovakia
jozef.komornik@fm.uniba.sk

Magda Komorníková and Jana Kalická
Faculty of Civil Engineering,
Slovak University of Technology,
Radlinského 11, 810 05 Bratislava,
Slovakia
{magdalena.komornikova,jana.kalicka}@stuba.sk

Summary

In this paper, we synthesize and substantially extend our recent investigations of specific class of perturbations of bivariate copulas and their effects on tail dependencies (along both diagonal sections). We show that those perturbations do not change the coefficients of tail dependencies along the main diagonal but linearly reduce their values along the second diagonal. An interesting possible application for analysing of dependencies along the second diagonal of copulas represent insurance data, where censoring introduces a negative dependence between the investigated components of the claims.

As a by–product, we present a new class of perturbations of copulas that linearly reduce the more popular coefficients of tail dependencies along the main diagonal, while preserving their values along the second diagonal.

Keywords: Copula; Perturbation of copula; Tail dependence; Survival copula; Reflections of copulas.

1 INTRODUCTION

Fitting of an appropriate copula to real data is one of major tasks in application of copulas. For this purpose, a large buffer of potential copulas has been designed (mainly parametric families of copulas). Once we know approximately a copula C appropriate to model the observed data, we look for a minor perturbation of C which fit them better then C itself.

The paper is organized as follows. The second section is presenting a brief overview of the theory of copulas. In the third section we discuss perturbations of bivariate copulas. The fourth section is devoted to an overview of the tail dependence coefficients and contains the main result of this paper.

2 COPULAS

Copula represents a multivariate distribution that captures the dependence structure among random variables. It is a great tool for building flexible multivariate stochastic models. Copula offers the choice of an appropriate model for the dependence between random variables independently from the selection of marginal distributions. This concept was introduced in the late 50's and became popular in several fields beyond statistics and probability theory, such as finance, actuarial science, fuzzy set theory, hydrology, civil engineering, etc.

Definition 1. A function $C : [0,1]^2 \to [0,1]$ is called a (bivariate) copula whenever it is

i) 2–increasing, i.e.,
$V_C([u_1,u_2]\times[v_1,v_2]) =$

$$= C(u_1,v_1)+C(u_2,v_2)-C(u_1,v_2)-C(u_2,v_1) \geq 0$$

for all $0 \leq u_1 \leq u_2 \leq 1$, $0 \leq v_1 \leq v_2 \leq 1$ (recall that $V_C([u_1,u_2]\times[v_1,v_2])$ is the C–volume of the rectangle $[u_1,u_2]\times[v_1,v_2]$);

ii) grounded, i.e., $C(u,0) = C(0,v) = 0$ for all $u,v \in [0,1]$;

iii) it has a neutral element $e = 1$, i.e., $C(u,1) = u$ and $C(1,v) = v$ for all $u,v \in [0,1]$.

Recall that for a 2–dimensional random vector (X,Y) with a joint distribution function F_{XY} and continuous marginal distribution functions F_X, F_Y a copula C satisfying the relations $F_{XY}(x,y) = C(F_X(x),F_Y(y))$ is the distribution function of the random vector (U,V),

ISBN: 978–83–8012–519–3

where $U = F_X(X)$ and $V = F_Y(Y)$ have uniform distributions on $[0, 1]$. For more details we recommend monographs Joe(1997) [5] and Nelsen(2006) [8].

For a better specifications of the tails of a bivariate distribution, Joe [5] introduced the *lower (left) and upper (right) tail dependence coefficients* λ_L and λ_U. The tail dependence coefficients can be calculated from the copula C of random vector (X, Y).

Definition 2. Let X and Y be continuous random variables with distribution functions F_X and F_Y and with copula C, then the *lower tail dependence coefficient* is defined by

$$\lambda_L(C) = \lim_{\delta\to 0^+} Pr\,(F_Y(y) \le \delta \,|\, F_X(x) \le \delta) = \tag{1}$$

$$= \lim_{\delta\to 0^+} \frac{C(\delta,\delta)}{\delta} = \lim_{\delta\to 0^+} Pr\,(F_X(x) \le \delta \,|\, F_Y(y) \le \delta)$$

(provided that the above indicated limits exist), and *upper tail dependence coefficient* is defined by

$$\begin{aligned}\lambda_U(C) &= \lim_{\delta\to 0^+} Pr\,(F_Y(y) \ge 1-\delta \mid F_X(x) \ge 1-\delta) = \\ &= \lim_{\delta\to 0^+} \frac{2\delta - 1 + C(1-\delta, 1-\delta)}{\delta} = \\ &= \lim_{\delta\to 0^+} Pr\,(F_X(x) \ge 1-\delta \mid F_Y(y) \ge 1-\delta)\,,\end{aligned} \tag{2}$$

(provided that the above indicated limits exist).

It is well known (see e.g. [4]) that the values of λ_U and λ_L for Normal and Frank copulas are equal to 0 and the Gumbel copula C_θ^G, $\theta \ge 1$ and Clayton copula C_θ^{Cl}, $\theta > 0$ satisfy the relation (see, e.g., [5, 8])

$$\lambda_U(C_\theta^G) = 2 - 2^{\frac{1}{\theta}}, \quad \lambda_L(C_\theta^G) = 0$$

and

$$\lambda_U(C_\theta^{Cl}) = 0, \quad \lambda_L(C_\theta^{Cl}) = 2^{-\frac{1}{\theta}}.$$

We follow the approach of Patton [9] and consider a so–called *survival copula* derived from a given copula C corresponding to the couple (X, Y) by

$$\widehat{C}(u, v) = u + v - 1 + C(1-u, 1-v) \tag{3}$$

which is the copula corresponding to the couple $(-X, -Y)$ with the marginal distribution functions

$$F_{-X}(x) = 1 - F_X(-x^+)$$

and

$$F_{-Y}(y) = 1 - F_Y(-y^+).$$

Obviously, the relations

$$\lambda_L(\widehat{C}) = \lambda_U(C) \quad \text{and} \quad \lambda_U(\widehat{C}) = \lambda_L(C)$$

hold.

Another natural transformations of the copula C are copulas LC and RC corresponding to the couples $(-X, Y)$ and $(X, -Y)$, respectively.

They have the form

$$LC(u, v) = v - C(1-u, v)$$

and

$$RC(u, v) = u - C(u, 1-v).$$

We will call the copulas LC and RC the *left* and the *right reflections* of the copula C, respectively (see, e.g. [1]). Since the survival copula $\widehat{C}$ can be obtained in the form of the right reflection of the copula LC as well as the left reflection of the copula RC, we included $\widehat{C}$ also in the family of the reflections of the copula C. Observe that if C is an absolutely continuous copula with density function $c_C(u, v)$, then also all its reflections are absolutely continuous with the respective density functions

$$c_{LC}(u, v) = c_C(1-u, v),$$

$$c_{RC}(u, v) = c_C(u, 1-v)$$

and

$$c_{\widehat{C}}(u, v) = c_C(1-u, 1-v).$$

We recall the definitions of upper–lower and lower–upper tail dependencies for the copula C (c.f. [6]) by

$$\lambda_{UL}(C) = \lambda_U(LC) = \lim_{u\to 0^+} \frac{u - C(u, 1-u)}{u} \tag{4}$$

and

$$\lambda_{LU}(C) = \lambda_L(LC) = \lim_{v\to 0^+} \frac{v - C(1-v, v)}{v}. \tag{5}$$

Obviously, for any copula C we have

$$\lambda_{UL}(LC) = \lambda_{LU}(RC) = \lambda_U(C) \tag{6}$$

and

$$\lambda_{LU}(RC) = \lambda_{LU}(LC) = \lambda_L(C). \tag{7}$$

Hence for a Gumbel copula $C = C_\theta^G$ the equalities

$$\lambda_{UL}(LC) = \lambda_{LU}(RC) = \lambda_U(C) = 2 - 2^{\frac{1}{\theta}}$$

hold. Moreover, $\lambda_{UL}(C) = \lambda_{LU}(C) = 0$ (see [6]).

Similarly, for the Clayton copula $C = C_\theta^{Cl}$ we have

$$\lambda_{UL}(RC) = \lambda_{LU}(LC) = \lambda_L(C) = 2^{-\frac{1}{\theta}}$$

as well as $\lambda_{UL}(C) = \lambda_{LU}(C) = 0$.

It is well known that for the convex sums of copulas, the corresponding density function is the convex sum (with the same weights) of incoming density functions. The same kind of mixing behaviour can be observed for the coefficients of tail dependencies λ_U, λ_L, λ_{LU} and λ_{UL}.

3 PERTUBATION OF BIVARIATE COPULAS

We will consider the bivariate copulas C_H that can be expressed in the form

$$C_H(u,v) = C(u,v) + H(u,v),$$

where C is a fixed copula and $H : [0,1]^2 \to \Re$ is a continuous function. The function H is called a *perturbation factor* and copula C_H is called a *perturbation* of C by means of H (see, e.g. [3]).

In [7] the next perturbation method (valid for any copula C) was introduced.

Theorem 1. Let $C : [0,1]^2 \to [0,1]$ be a copula and define $H_\alpha^C : [0,1]^2 \to \Re, \alpha \in [0,1]$ by

$$H_\alpha^C(u,v) = \alpha\,(u - C(u,v))\,(v - C(u,v)). \qquad (8)$$

Then $C_{H_\alpha^C} : [0,1]^2 \to [0,1]$ given by

$$C_{H_\alpha^C}(u,v) = C(u,v) + H_\alpha^C(u,v) \qquad (9)$$

is a copula for each $\alpha \in [0,1]$ and any copula C.

Observe that the perturbed copula $C_{H_1^C}$ has a following stochastic interpretation:

Consider two independent identically distributed continuous random vectors (X_1,Y_1) and (X_2,Y_2), $C_{(X_1,Y_1)} = C_{(X_2,Y_2)} = C$. Then $C_{H_1^C} = C_{(T_1,T_2)}$ is a copula characterizing the random vector (T_1,T_2) given by

$$(T_1,T_2) = \\ = \begin{cases} (\min(X_1,X_2), \min(Y_1,Y_2)) & \text{with probability } 0.5, \\ (\max(X_1,X_2), \max(Y_1,Y_2)) & \text{with probability } 0.5. \end{cases}$$

For more details see [2].

Concerning the perturbed copulas $C_{H_\alpha^C}$, $\alpha \in [0,1]$, they can be seen as a convex sum of the original copula C and the copula $C_{H_1^C}$,

$$C_{H_\alpha^C} = (1-\alpha)\,C + \alpha\,C_{H_1^C}.$$

Remark 1. It is straightforward to show that the following relation hold

$$LC_{H_\alpha^C}(u,v) = LC(u,v)(1 - \alpha\,RC(1-u,1-v)) \quad (10)$$

and

$$RC_{H_\alpha^C}(u,v) = RC(u,v)\,(1 - \alpha\,LC(1-u,1-v)). \quad (11)$$

Proof.

$$LC_{H_\alpha^C}(u,v) = v - C_{H_\alpha^C}(1-u,v) = v - [C(1-u,v) +$$

$$+\alpha\,(1-u-C(1-u,v))\,(v - C(1-u,v))] =$$

$$= -\alpha C(1-u,v)^2 - \alpha u C(1-u,v) + \alpha v C(1-u,v) +$$

$$+\alpha C(1-u,v) - C(1-u,v) + \alpha uv - \alpha v + v = v - C(1-u,v) -$$

$$-\alpha\,(1-u-C(1-u,v))\,(v - C(1-u,v)) =$$

$$= LC(u,v) - \alpha LC(u,v) RC(1-u,1-v) =$$

$$= LC(u,v)(1 - \alpha\,RC(1-u,1-v)).$$

Similarly

$$RC_{H_\alpha^C}(u,v) = u - C_{H_\alpha^C}(u,1-v) = u - [C(u,1-v) +$$

$$+\alpha\,(u - C(u,1-v))\,(v - C(u,1-v))] =$$

$$= -\alpha C(u,1-v)^2 + \alpha u C(u,1-v) - \alpha v C(u,1-v) +$$

$$+\alpha C(u,1-v) - C(u,1-v) - \alpha u + \alpha uv + u =$$

$$= u - C(u,1-v) - \alpha(u - C(u,1-v))(1 - v - C(u,1-v)) =$$

$$= RC(u,v) - \alpha RC(u,v) LC(1-u,1-v) =$$

$$= RC(u,v)\,(1 - \alpha\,LC(1-u,1-v)).$$

□

In the next section we will investigate tail dependencies for given perturbed copulas.

4 TAIL DEPENDENCE OF PERTURBED COPULAS

C is said to have lower (upper) tail dependence if and only if $\lambda_L \neq 0$ ($\lambda_U \neq 0$). As it can bee seen from Definition 2, the tail dependence coefficients are connected with the *diagonal section* of the bivariate copula C, which is defined by the function

$$\delta_C : [0,1] \to [0,1]\,, \delta_C(u) = C(u,u). \qquad (12)$$

Combining (12) with (3) we obtain

$$\begin{aligned} \delta_{\widehat{C}}(u) &= 2u - 1 + \delta_C(1-u,1-u) \\ \delta_C(u) &= 2u - 1 + \delta_{\widehat{C}}(1-u,1-u) \end{aligned} \qquad (13)$$

(because of $\widehat{\widehat{C}} = C$).

The coefficients of tail dependence (if they exist) can be expressed by means of formulas

$$\lambda_L(C) = \delta_C'(0^+) \qquad (14)$$

and (using (2))

$$\lambda_U(C) = 2 - \delta_C'(1^-). \qquad (15)$$

Example 1. We will consider Farlie–Gumbel–Morgenstern (FGM) $\left(C_\alpha^{FGM}\right)_{\alpha\in[-1,1]}$ family of copulas given by

$$C_\alpha^{FGM}(u,v) = u\,v + \alpha\,u\,(1-u)\,v\,(1-v), \quad \alpha\in[-1,1].$$

We have

$$\delta(u) = u^2\left[1+\alpha\,(1-u)^2\right] \text{ for } \alpha\in[-1,1].$$

Obviously, $\delta'(0^+) = 0$ and $\delta'(1^-) = 2$. Hence $\lambda_U(C_\alpha) = \lambda_L(C_\alpha) = 0$.

Remark 2. For $\alpha\in[0,1]$

$$C_\alpha^{FGM}(u,v) = u\,v + \alpha\,u\,(1-u)\,v\,(1-v) =$$

$$= u\,v + \alpha\,(u-u\,v)\,(v-u\,v) = \Pi_{H_\alpha^\Pi},$$

i.e., for $\alpha\in[0,1]$ is C_α^{FGM} the perturbation of Π.

However, we can not expect that (8) and (9) will yield copula H_α^C for any copula C and $\alpha<0$. For example, for the minimal copula $C_{min} = W$ given by $W(u,v) = \max(u+v-1,0)$ (which is the copula related to any couple $(X,-X)$ with a continuous distribution function F_X) we get for $0<u, 0<v$ and $v<1-u$, the value $W(u,v)=0$ and $H_\alpha^W(u,v) = \alpha\,u\,v < 0$. Hence H_α^W can not be extended using (8) and (9) for an arbitrary copula C and $\alpha<0$.

Theorem 2. Let C be a copula and $\lambda_L(C)$, $\lambda_U(C)$ exist. Let $C_{H_\alpha^C}$ be a copula given by (8) and (9). Then

$$\lambda_L(C_{H_\alpha^C}) = \lambda_L(C) \tag{16}$$

and

$$\lambda_U(C_{H_\alpha^C}) = \lambda_U(C). \tag{17}$$

Proof.

$$\lambda_L(C_{H_\alpha^C}) = \delta'_{C_{H_\alpha^C}}(0^+) = \delta'_C(0^+) + \alpha\frac{d\left[u-\delta_C(u)\right]^2}{du}|(0^+).$$

We have $0\le u-\delta_C(u)\le u$ for $u\in[0,1]$ and thus $\frac{d\,[u-\delta_C(u)]^2}{du}|(0^+) = 0$. Hence

$$\lambda_L(C_{H_\alpha^C}) = \delta'_C(0^+) = \lambda_L(C).$$

The upper tail dependence is given by

$$\lambda_U(C_{H_\alpha^C}) = 2 - \delta'_{C_{H_\alpha^C}}(1^-) =$$

$$= 2 - \left(\delta'_C(1^-) + \alpha\frac{d\left[u-\delta_C(u)\right]^2}{du}|(1^-)\right). \tag{18}$$

Combining (18) with (13) and substituting $u = 1-v$ we obtain

$$u - \delta_C(u) = v - \delta_{\widehat{C}}(v)$$

and thus

$$\frac{d\,[u-\delta_C(u)]^2}{du}|(1^-) = \frac{d\,[v-\delta_{\widehat{C}}(v)]^2}{dv}|(0^+) = 0.$$

Therefore,

$$\lambda_U(C_{H_\alpha^C}) = 2 - \delta'_C(1^-) = \lambda_U(C).$$

□

Thus, the perturbations H_α^C do not change the values of the coefficients of tail dependence along the first (main) diagonal.

The following result shows that the effect of those perturbations is much different along the second diagonal.

Theorem 3. Let for a copula C the coefficients λ_{LU} and λ_{UL} given by (5) and (4) exist. Let $C_{H_\alpha^C}$ be given by (8) and (9). Then the relations

$$\lambda_{UL}\left(C_{H_\alpha^C}\right) = \lambda_{UL}(C)\,(1-\alpha) \tag{19}$$

and

$$\lambda_{LU}\left(C_{H_\alpha^C}\right) = \lambda_{LU}(C)\,(1-\alpha). \tag{20}$$

hold.

Proof. For $D:[0,1]^2\to[0,1]$ we define

$$\beta_D : [0,1]\to[0,1]\,, \beta_D(u) = D(u,1-u). \tag{21}$$

Let C be a copula and $\lambda_{UL}(C)$ and $\lambda_{LU}(C)$ exist. Using (4), we get

$$\lambda_{UL}(C) = 1-\beta'_C(0^+),$$

similarly, from (5) we get

$$\lambda_{LU}(C) = 1+\beta'_C(1^-).$$

Since

$$H_\alpha^C(u,v) = \alpha\,\left(u-C(u,v)\right)\,\left(v-C(u,v)\right),$$

$$C_{H_\alpha^C}(u,v) = C(u,v)+H_\alpha^C(u,v) = C(u,v)+\alpha\,H_1^C(u,v).$$

$$\lambda_{UL}\left(C_{H_\alpha^C}\right) = \lambda_{UL}(C) - \alpha\,\beta'_{H_1^C}(0^+). \tag{22}$$

$$\beta_{H_1^C}(u) = \left(u-C(u,1-u)\right)\,\left((1-u)-C(u,1-u)\right),$$

$$\beta'_{H_1^C}(0^+) = \left(1-\beta'_C(0^+)\right)\,1+0\,\left(0-\beta'_C(0^+)\right) = \lambda_{UL}(C). \tag{23}$$

Combining (22) with (23) we get that (19) holds. Similarly,

$$\beta'_{H_1^C}(1^-) = 0\,\left(1-\beta'_C(1^-)\right)+1\,\left(-1-\beta'_C(1^-)\right) = -\lambda_{LU}(C),$$

and thus

$$\lambda_{LU}\left(C_{H_\alpha^C}\right) = \lambda_{LU}(C)+\alpha\,\beta'_{H_1^C}(1^-) = \lambda_{LU}(C)\,(1-\alpha),$$

hence (20) holds. □

Corollary 1. The above relations imply (together with (8) and (9)) the following equalities:

$$\lambda_{LU}\left(LC_{H_\alpha^C}\right) = \lambda_L\left(C_{H_\alpha^C}\right) = \lambda_L\left(C\right)$$

$$\lambda_{LU}\left(RC_{H_\alpha^C}\right) = \lambda_U\left(C_{H_\alpha^C}\right) = \lambda_U\left(C\right)$$

$$\lambda_{UL}\left(LC_{H_\alpha^C}\right) = \lambda_U\left(C_{H_\alpha^C}\right) = \lambda_U\left(C\right)$$

$$\lambda_{UL}\left(RC_{H_\alpha^C}\right) = \lambda_L\left(C_{H_\alpha^C}\right) = \lambda_L\left(C\right)$$

Moreover, from (6), (7), (19) and (20) we obtain

$$\lambda_U\left(LC_{H_\alpha^C}\right) = (1-\alpha)\lambda_U(LC)$$

$$\lambda_L\left(LC_{H_\alpha^C}\right) = (1-\alpha)\lambda_U(LC)$$

$$\lambda_U\left(RC_{H_\alpha^C}\right) = (1-\alpha)\lambda_U(RC)$$

$$\lambda_L\left(RC_{H_\alpha^C}\right) = (1-\alpha)\lambda_U(RC).$$

Remark 3. Note that our approach allows to introduce a parametric class of perturbed copulas with linearly varying lower and upper tail dependencies, too. Indeed, consider a copula C with lower and upper tail dependencies λ_L and λ_U, respectively. Then

$$\lambda_{UL}(C) = \lambda_U(LC)$$

and

$$\lambda_{LU}(C) = \lambda_L(LC).$$

Now, for the perturbed copulas $LC_{H_\alpha^{LC}}$ it holds

$$\lambda_U(LLC_{H_\alpha^{LC}}) = \lambda_{UL}(LC_{H_\alpha^{LC}}) = (1-\alpha)\lambda_{UL}(LC) =$$

$$= (1-\alpha)\lambda_U(LLC) = (1-\alpha)\lambda_U(C),$$

and similarly

$$\lambda_L(LLC_{H_\alpha^{LC}}) = (1-\alpha)\lambda_L(C).$$

Observe that if $\alpha = 0$ then $LC_{H_0^{LC}} = LC$ and thus $LLC_{H_0^{LC}} = LLC = C$.

If we denote the new class of perturbations as

$$D_\alpha = LLC_\alpha^{LC}, \tag{24}$$

we get

$$\lambda_{UL}(D_\alpha) = \lambda_U\left(LC_\alpha^{LC}\right) = \lambda_U(LC) = \lambda_{UL}(C).$$

Similarly,

$$\lambda_{LU}(D_\alpha) = \lambda_L\left(LC_\alpha^{LC}\right) = \lambda_L(LC) = \lambda_{LU}(C).$$

Thus the class D_α given by (24) does not change the values of the tail dependencies along the second diagonal, while it reduces them linearly along the main diagonal.

5 CONCLUSIONS

We have shown that the investigated class of perturbations that was introduced in [7] as a (partial) generalization of the FMG class of copulas does not change the values of the coefficients of tail dependencies along the main diagonal. However, it (quit surprisingly) yields their reduction along the second diagonal. Inspired by suggestions of Radko Mesiar, we introduced another class of perturbations that reduce linearly the coefficients of tail dependencies along the main diagonal while preserving their values along the second diagonal.

Acknowledgements

The support of the grants APVV-14-0013 and VEGA 1/0420/15 is kindly announced.

References

[1] B. De Baets, H. De Meyer, J. Kalická and R. Mesiar (2009). Flipping and cyclic shifting of binary aggregation functions. *Fuzzy Sets and Systems* **160(6)**, pp. 752–765.

[2] A. Dolati and M. Úbeda–Flores (2009). Constructing Copulas by Means of Pairs of Order Statistics. *Kybernetika* **45**, pp. 992–1002.

[3] F. Durante, J.F. Sánchez and M.U. Flores (2013). Bivariate copulas generated by perturbations. *Fuzzy Sets and System* **228**, pp. 137–144.

[4] P. Embrechts, F. Lindskog and A. McNeil (2001). Modeling dependence with copulas and applications to risk management. *Rachev, S. (Ed.) Handbook of Heavy Tailed Distributions in Finance.* Elsevier, Chapter 8, pp 329-384.

[5] H. Joe (1997). Multivariate model and dependence concept. *Monographs on Statistics and Applied Probability* **73**, Chapman and Hall.

[6] J. Komorník and M. Komorníková (2013) Modelling financial time series using reflections of copulas. *Kybernetika*, 49, **3** 487–497, 2013.

[7] R. Mesiar, M. Komorníková and J. Komorník (2015). Perturbation of bivariate copula. *Fuzzy Sets and System* **268**, pp. 127–140.

[8] R.B. Nelsen (2006). An introduction to copulas. Second Edition. *Springer Series in Statistics*, Springer-Verlag, New York.

[9] A.J. Patton (2006). Modelling Asymmetric Exchange Rate Dependence. *International Economic Review* **47**, No. 2, pp 527-556.

FUZZY (C, I)-EQUIVALENCES

Anna Król
Department of Mathematics and Natural Sciences, University of Rzeszów,
ul. Rejtana 16C, 35-959 Rzeszów, Poland
annakrol@ur.edu.pl

Summary

The article deals with the notion of fuzzy equivalence, whose definition depends on a fuzzy conjunction and implication. Moreover, the preservation of properties of such (C, I)-equivalences during aggregation process is considered and the way of generating new fuzzy equivalences from the given ones is indicated.

Keywords: Fuzzy (C, I)-equivalence, Fuzzy implication, Fuzzy conjunction, Aggregation function, Dominance.

1 INTRODUCTION

Aggregation functions can be useful in a variety of information fusion problems [3, 8]. The problem of aggregation of diverse mathematical objects is rather well known. We may aggregate for example fuzzy relations and consider the problem of preservation of fuzzy relation properties during aggregation process (e.g. [6, 11, 14, 15]) or examine fuzzy connectives and preservation of their axioms or properties by aggregation functions (e.g. [2, 5, 7, 10]).

Examination of preservation of axioms and properties of fuzzy connectives finds its applications in decision making, approximate reasoning and fuzzy control. Moreover, preservation in the aggregation process not only individual properties, but all axioms of given fuzzy connectives shows a way of generating a new fuzzy connective of the same kind. In this context, some of the fuzzy connectives (negation, conjunction, disjunction and implication) were examined for example in [7], a fuzzy implication was examined in [5], diverse kind of a fuzzy equivalences were considered in [2, 10].

In this paper, the definition of (C, I)-equivalence, as one of fuzzy connectives, generated by a fuzzy conjunction and implication is presented. Besides, the preservation of the various properties of (C, I)-equivalence during aggregation process is indicated and a way of generating new fuzzy equivalences from the given ones is presented.

In Section 2, basic notions useful in the paper are presented. In Section 3, fuzzy equivalences are discussed, and in Section 4, aggregation of fuzzy equivalences are examined.

2 PRELIMINARIES

Here we recall basic notions and their properties which will appear in the sequel. In particular, we consider fuzzy conjunctions, fuzzy implications and aggregation functions.

2.1 FUZZY CONJUNCTIONS

First, the definition and some properties of a fuzzy conjunction is presented.

Definition 1 ([7])**.** An operation $C : [0, 1]^2 \to [0, 1]$ is called a fuzzy conjunction if it is increasing with respect to each variable and

$$C(1, 1) = 1, \quad C(0, 0) = C(0, 1) = C(1, 0) = 0.$$

Directly from the definition we obtain a useful property of a fuzzy conjunction.

Corollary 1. *Any fuzzy conjunction has a zero element $z = 0$.*

Example 1. Consider the following family of fuzzy conjunctions for $\alpha \in [0, 1]$

$$C^\alpha(x, y) = \begin{cases} 1, & \text{if } x = y = 1 \\ 0, & \text{if } x = 0 \text{ or } y = 0 \\ \alpha & \text{otherwise} \end{cases}.$$

ISBN: 978–83–8012–519–3

Operations C^0 and C^1 are the least and the greatest fuzzy conjunction, respectively. Well-known t-norms T_M, T_P, T_L, T_D are other examples of fuzzy conjunction.

2.2 FUZZY IMPLICATIONS

Next, we recall the notion of a fuzzy implication.

Definition 2 ([1], pp. 2,9)**.** A binary operation $I\colon [0,1]^2 \to [0,1]$ is called a fuzzy implication if it is decreasing with respect to the first variable and increasing with respect to the second variable and

$$I(0,0) = I(0,1) = I(1,1) = 1, \quad I(1,0) = 0.$$

Corollary 2. *A fuzzy implication has the right absorbing element* 1 *and fulfils the condition*

$$I(0,y) = 1, \quad y \in [0,1].$$

We can also consider other properties of a fuzzy implications (for more information see e.g. [1]). For this contribution one of these is especially important.

Definition 3 ([1], pp. 9)**.** We say that a fuzzy implication I fulfils the identity principle (IP) if

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

Example 2 (cf. [1], pp. 4,5)**.** Let us present the following family of fuzzy implications for $\alpha \in [0,1]$

$$I^\alpha(x,y) = \begin{cases} 0, & \text{if } x = 1, y = 0 \\ 1, & \text{if } x = 0 \text{ or } y = 1 \\ \alpha & \text{otherwise} \end{cases}.$$

The operations I^0 and I^1 are the least and the greatest fuzzy implication, respectively, where

$$I^0(x,y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } \ y = 1 \\ 0, & \text{otherwise} \end{cases},$$

$$I^1(x,y) = \begin{cases} 0, & \text{if } x = 1, y = 0 \\ 1, & \text{otherwise} \end{cases}.$$

The following are the other examples of fuzzy implications.

$$I_{\mathrm{LK}}(x,y) = \min(1 - x + y, 1),$$

$$I_{\mathrm{GD}}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ y, & \text{if } x > y \end{cases},$$

$$I_{\mathrm{RC}}(x,y) = 1 - x + xy,$$

$$I_{\mathrm{DN}}(x,y) = \max(1 - x, y),$$

$$I_{\mathrm{GG}}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ \frac{y}{x}, & \text{if } x > y \end{cases},$$

$$I_{\mathrm{RS}}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ 0, & \text{if } x > y \end{cases},$$

$$I_{\mathrm{YG}}(x,y) = \begin{cases} 1, & \text{if } x, y = 0 \\ y^x, & \text{else} \end{cases},$$

$$I_{\mathrm{FD}}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ \max(1 - x, y), & \text{if } x > y \end{cases},$$

$$I_{\mathrm{WB}}(x,y) = \begin{cases} 1, & \text{if } x \le 1 \\ y, & \text{if } x = 1 \end{cases},$$

$$I_{\mathrm{DP}}(x,y) = \begin{cases} y, & \text{if } x = 1 \\ 1 - x, & \text{if } y = 0 \\ 1 & \text{otherwise} \end{cases}.$$

The implications fulfilling the property (IP) are: I^1, I_{LK}, I_{GD}, I_{GG}, I_{RS}, I_{WB}, I_{FD}, I_{DP}.

2.3 AGGREGATION FUNCTIONS

Now, we recall the notion of aggregation function.

Definition 4 (cf. [4], pp. 6-22, [12], pp. 216-218)**.** Let $n \in \mathbb{N}$. A function $A : [0,1]^n \to [0,1]$ which is increasing, i.e. for $x_i, y_i \in [0,1], x_i \leqslant y_i, i = 1, \ldots, n$

$$A(x_1, \ldots, x_n) \leqslant A(y_1, \ldots, y_n)$$

is called an aggregation function if

$$A(0, \ldots, 0) = 0, \ A(1, \ldots, 1) = 1.$$

Example 3 (cf. [4], pp. 44-56, [9])**.** A_0, A_1 are the least and the greatest aggregation functions, where

$$A_0(x_1, \ldots, x_n) = \begin{cases} 1, & (x_1, \ldots, x_n) = (1, \ldots, 1) \\ 0, & (x_1, \ldots, x_n) \neq (1, \ldots, 1) \end{cases},$$

$$A_1(x_1, \ldots, x_n) = \begin{cases} 0, & (x_1, \ldots, x_n) = (0, \ldots, 0) \\ 1, & (x_1, \ldots, x_n) \neq (0, \ldots, 0) \end{cases},$$

$x_1, \ldots, x_n \in [0,1]$. Other simple examples of aggregation function are projections

$$P_k(x_1, \ldots, x_n) = x_k, \text{ for } k = 1, 2, \ldots, n, \qquad (1)$$

and weighted means

$$A_w(x_1, \ldots, x_n) = \sum_{k=1}^{n} w_k x_k, \qquad (2)$$

for $w_k > 0, \sum_{k=1}^n w_k = 1$.

3 FUZZY EQUIVALENCES

In this section the definition of (C, I)-equivalence is presented. Moreover, the properties of such a fuzzy connective are examined according to axioms of other, well-known, notions of a fuzzy C-equivalence [13] as well as a fuzzy equivalence introduced by Fodor and Roubens [11].

In the literature one can meet various definitions of a fuzzy equivalence. A trivial case used in many contributions, for example those concerning generalized logical laws, is an equality, that is the function $E : [0,1]^2 \to [0,1]$ given by the formula (relation of identity)

$$E(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}. \tag{3}$$

Usually it is expected that such notion of a fuzzy equivalence is a generalization of the equivalence of classical propositional calculus, that is the function $E : [0,1]^2 \to [0,1]$ that fulfils conditions $E(0,1) = E(1,0) = 0$, $E(0,0) = E(1,1) = 1$. Here the approach will be applied in the proposed definition of a fuzzy (C, I)-equivalence as well as in other definitions of a fuzzy equivalence chosen from many other well-known notions.

Let us consider a fuzzy equivalence defined on the pattern of the following law of classical propositional calculus

$$(p \Leftrightarrow q) \Leftrightarrow [(p \Rightarrow q) \wedge (q \Rightarrow p)].$$

Definition 5. Let C, I be a fuzzy conjunction and implication, respectively. The function $E : [0,1]^2 \to [0,1]$ given by the formula

$$E_{C,I}(x,y) = C(I(x,y), I(y,x)), \; x,y \in [0,1] \tag{4}$$

will be called (C, I)-equivalence.

Remark 1. Let us notice that the function (4) fulfils zero-one table of crisp equivalence. Indeed,

$$\begin{aligned} E_{C,I}(0,0) &= C(I(0,0), I(0,0)) = C(1,1) = 1, \\ E_{C,I}(1,1) &= C(I(1,1), I(1,1)) = C(1,1) = 1, \\ E_{C,I}(0,1) &= C(I(0,1), I(1,0)) = C(1,0) = 0, \\ E_{C,I}(1,0) &= C(I(1,0), I(0,1)) = C(0,1) = 0. \end{aligned}$$

Additionally, if I fulfils (IP), then $E_{C,I}(x,x) = 1$, as in this case we have $E_{C,I}(x,x) = C(I(x,x), I(x,x)) = C(1,1) = 1$ for any $x \in [0,1]$. Besides, if C is a commutative fuzzy conjunction, then we obtain $E_{C,I}(x,y) = C(I(x,y), I(y,x)) = C(I(y,x), I(x,y)) = E_{C,I}(y,x)$ for any $x, y \in [0,1]$, which means that $E_{C,I}$ is also commutative.

Example 4. Let us consider the greatest fuzzy implication I^1 and an arbitrary fuzzy conjunction C. Then

$$E_{C,I^1}(x,y) = \begin{cases} 0, & \text{if } \{x,y\} = \{0,1\} \\ 1 & \text{otherwise} \end{cases}.$$

For other examples of (C, I)-equivalences see Example 5.

Now, let us focus on the notion of fuzzy equivalence considered in [11].

Definition 6 ([11], p. 33)**.** A fuzzy equivalence is a function $E : [0,1]^2 \to [0,1]$ which fulfils

$$E(0,1) = 0, \tag{5}$$

$$E(x,x) = 1, \; x \in [0,1], \tag{6}$$

$$E(x,y) = E(y,x), \; x,y \in [0,1], \tag{7}$$

$$E(x,y) \leqslant E(u,v), \; x \leqslant u \leqslant v \leqslant y, \; x,y,u,v \in [0,1]. \tag{8}$$

There exists a characterization of such defined fuzzy equivalence by the use of fuzzy implications fulfilling (IP).

Theorem 1 ([11], p. 33)**.** *A function $E : [0,1]^2 \to [0,1]$ is a fuzzy equivalence if and only if there exists such a fuzzy implication I fulfilling (IP) that*

$$E_I(x,y) = \min(I(x,y), I(y,x)), \; x,y \in [0,1]. \tag{9}$$

Corollary 3 ([11], p. 34)**.** *A function $E : [0,1]^2 \to [0,1]$ is a fuzzy equivalence if and only if there exists such a fuzzy implication I fulfilling (IP) that*

$$E_I(x,y) = I(\max(x,y), \min(x,y)), \; x,y \in [0,1]. \tag{10}$$

By Theorem 1 we obtain as follows.

Corollary 4. *Any fuzzy equivalence (Definition 6) is (C, I)-equivalence.*

Remark 2. A fuzzy equivalence is (C, I)-equivalence, however, it is not necessary $C = \min$, what shows Example 4. Thus, (C, I)-equivalence is a generalization of a fuzzy equivalence.

Example 5. The following table presents examples of fuzzy equivalences as well as $(\min, I)$-equivalences generated by the use of formula (10) and these of the fuzzy implications from Example 2 that fulfil (IP). For example, we will show how to obtain the fuzzy equivalence E_{GG}. Let us notice that for any $x, y \in [0,1]$ we have $\max(x,y) \leqslant \min(x,y) \Leftrightarrow x = y$. Thus, for any $x, y \in [0,1]$ we have

$$E_{GG}(x,y) = \begin{cases} 1, & \text{if } x = y \\ \frac{\min(x,y)}{\max(x,y)}, & \text{if } x \neq y \end{cases} = \begin{cases} 1, & \text{if } x = y \\ \frac{x}{y}, & \text{if } x < y \\ \frac{y}{x}, & \text{if } x > y \end{cases}.$$

Let us also see, that in the case of a fuzzy equivalence generated by the fuzzy implication I_{RS} we obtain the fuzzy equivalence E_{RS}, which is the fuzzy equality (3).

I	E_I
I_{LK}	$E_{LK}(x,y) = 1 - \|x-y\|$
I_{GD}	$E_{GD}(x,y) = \begin{cases} 1, & \text{if } x = y \\ x, & \text{if } x < y \\ y, & \text{if } x > y \end{cases}$
I_{GG}	$E_{GG}(x,y) = \begin{cases} 1, & \text{if } x = y \\ \frac{x}{y}, & \text{if } x < y \\ \frac{y}{x}, & \text{if } x > y \end{cases}$
I_{RS}	$E_{RS}(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}$
I_{WB}	$E_{WB}(x,y) = \begin{cases} 1, & \text{if } x \neq 1,\ y \neq 1 \\ x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \end{cases}$
I_{FD}	$E_{FD}(x,y) = \begin{cases} 1, & \text{if } x = y \\ \max(1-y, x), & \text{if } x < y \\ \max(1-x, y), & \text{if } x > y \end{cases}$
I_{DP}	$E_{DP}(x,y) = \begin{cases} x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \\ 1-x, & \text{if } y = 0 \\ 1-y, & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases}$
I^1	$E(x,y) = \begin{cases} 0, & \text{if } \{x,y\} = \{0,1\} \\ 1 & \text{otherwise} \end{cases}$

The next definition follows from the notion of a fuzzy equivalence relation, namely relation which is reflexive, symmetric and transitive.

Definition 7 (cf. [13], p. 33)**.** Let C be a fuzzy conjunction. A fuzzy C-equivalence is a function $E : [0,1]^2 \to [0,1]$ fulfilling the conditions (5)–(7) and C-transitivity, i.e.

$$C(E(x,y), E(y,z)) \leqslant E(x,z), \quad x,y,z \in [0,1]. \quad (11)$$

In the cited monograph [13] property (5) is omitted. However, in this case the constant function $E(x,y) = 1$, $x,y \in [0,1]$ fulfils the definition of a fuzzy equivalence although it is not a generalization of crisp equivalence. This is why this assumption is added to the definition.

Let us notice that one may weaken conditions put onto a fuzzy C-equivalence by replacing the C-transitivity property with the appropriate weaker transitivity condition. Such considerations have already been presented e.g. in [2, 10].

Theorem 2. *Let C, C_1 be arbitrary fuzzy conjunctions. (C, I^0)-fuzzy equivalence is C_1-transitive.*

Proof. Let $x,y,z \in [0,1]$. First, we consider the left and right sides of the transitivity equation (11).

$$L = C_1(E_{C,I^0}(x,y), E_{C,I^0}(y,z)) =$$
$$= C_1(C(I^0(x,y), I^0(y,x)), C(I^0(y,z), I^0(z,y))),$$
$$R = E_{C,I^0}(x,z) = C(I^0(x,z), I^0(z,x)).$$

Let us notice, that if $x = y = z = 0$ or $x = y = z = 1$, then all of the values $I^0(x,y)$, $I^0(y,x)$, $I^0(y,z)$, $I^0(z,y)$, $I^0(x,z)$, $I^0(z,x)$ are equal to 1. Thus, $L = R = 1$. Otherwise, by the formula of the fuzzy implication I^0, at least one of the values $I^0(x,y)$, $I^0(y,x)$, $I^0(y,z)$, $I^0(z,y)$ is equal to 0, and then, by Corollary 1, $L = 0 \leqslant R$. □

Example 6. Let C, C_1 be arbitrary fuzzy conjunctions. (C, I^1)-fuzzy equivalence is not C_1-transitive.

Proof. Let $x = 0$, $y = 0.5$, $z = 1$. On one hand we have

$$L = C_1(E_{C,I^1}(0, 0.5), E_{C,I^1}(0.5, 1)) =$$
$$C_1(C(I^1(0, 0.5), I^1(0.5, 0)), C(I^1(0.5, 1), I^1(1, 0.5))) =$$
$$C_1(C(1,1), C(1,1)) = C_1(1,1) = 1.$$

On the other hand $R = E_{C,I^1}(0,1) = 0$.
Thus $L = 1 > 0 = R$ and (C, I^1)-fuzzy equivalence is not C_1-transitive. □

4 PRESERVATION OF FUZZY EQUIVALENCES IN AGGREGATION PROCESS

Here we consider aggregation of fuzzy connectives defined in the previous section.

Definition 8 (cf. [11], p. 14)**.** Let $n \in \mathbb{N}$, A be an arbitrary aggregation function, $E_k : [0,1]^2 \to [0,1]$ for $k = 1, \ldots, n$. For given binary operations $E_1, \ldots, E_n$, we consider a binary operation $E : [0,1]^2 \to [0,1]$ such that for all $x, y \in [0,1]$

$$E(x,y) = A(E_1(x,y), \ldots, E_n(x,y)). \quad (12)$$

We say that an aggregation function A preserves a property of the given binary operations if the operation E defined by (12) has such a property for any $E_1, \ldots, E_n$ fulfilling this property.

At first we recall or prove some facts concerning preservation of given property of a binary operation.

Lemma 1 ([7])**.** *Any aggregation function preserves binary truth tables of aggregated fuzzy connectives of the same type.*

Lemma 2 ([7])**.** *Any aggregation function preserves symmetry of aggregated binary operations.*

Lemma 3 ([7]). *Any aggregation function preserves the property* (6) *of aggregated binary operations.*

Lemma 4. *Any aggregation function preserves the property* (8) *of the binary operations operations.*

Proof. Let $E_1, \ldots, E_n$ be binary operations fulfilling conditions (8). According to (12), for any aggregation function A one has for any $x, y, u, v \in [0,1]$ such that $x \leqslant u \leqslant v \leqslant y$

$$E(x,y) = A(E_1(x,y), \ldots, E_n(x,y)) \leqslant \\ \leqslant A(E_1(u,v), \ldots, E_n(u,v)) = E(u,v),$$

as $E_k(x,y) \leqslant E_k(u,v)$ for $k = 1, \ldots, n$ and by monotonicity of A. □

Next, let us notice that in the following theorem the notion of dominance is involved. Some characterizations and many examples of the dominance between aggregation functions and fuzzy conjunctions can be found in [2].

Theorem 3 (cf. [10]). *Let C be a fuzzy conjunction. An aggregation function $A : [0,1]^n \to [0,1]$ preserves condition* (11) *of the aggregated binary operations if and only if A dominates C, i.e.*

$$A(C(a_{1,1}, a_{1,2}), \ldots, C(a_{n,1}, a_{n,2})) \geqslant \\ C(A(a_{1,1}, \ldots, a_{n,1}), A(a_{1,2}, \ldots, a_{n,2})).$$

Remark 3. By Lemma 1 it follows that the operation E defined by (12) is a generalization of the equivalence of classical propositional calculus without any additional assumptions on an aggregation function A.

By Lemmas 1–4 it follows the following result.

Theorem 4. *Any aggregation function preserves axioms* (5)–(8) *of a fuzzy equivalence presented in Definition 6.*

Remark 4. By above theorem we see that formula (12) gives possibility of generating new fuzzy equivalences from given ones.

Example 7. Let us consider a binary weighted mean, that is an aggregation function given by (2) for n=2, where $w_1 = w_2 = 0.5$. By Theorem 4, from fuzzy equivalences $E_1 = E_{LK}$ and $E_2 = E_{GD}$, by the use of formula (12) we obtain a new fuzzy equivalence (Definition 6).

$$E(x,y) = A(E_{LK}(x,y), E_{GD}(x,y)) = \\ \begin{cases} 1, & \text{if } x = y \\ 0.5 + x - 0.5y, & \text{if } x < y \\ 0.5 - 0.5x + y, & \text{if } x > y \end{cases}.$$

By Lemmas 1–3 and Theorem 3 we obtain the following result.

Theorem 5 (cf. [2]). *Let C be a fuzzy conjunction, A an aggregation function such as A dominates C. A preserves axioms* (5)–(7), (11) *of a fuzzy C-equivalence.*

Corollary 5. *Any aggregation function preserves additional properties, such as* (5)–(8) *of* (C, I)*-equivalence. An aggregation function that dominates a fuzzy conjunction C preserve supplementary C-transitivity of* (C, I)*-equivalence.*

5 CONCLUSIONS

In the paper, the definition of a fuzzy equivalence that depends on a fuzzy conjunction and implication is proposed. Some of its properties according to axioms of other notions o fuzzy equivalence are examined. Moreover, the preservation of the properties in aggregation process are considered. As further work, other properties of (C, I)-equivalences may be examined taking into consideration additional properties of generators C and I.

Acknowledgements

This work was partially supported by the Center for Innovation and Transfer of Natural Sciences and Engineering Knowledge in Rzeszów, through Project Number RPPK.01.03.00-18-001/10.

References

[1] M. Baczyński, B. Jayaram (2008). Fuzzy implications, Springer, Berlin.

[2] U. Bentkowska, A. Król. Aggregation of fuzzy α-C-equivalences, Proc. IFSA-EUSFLAT 2015 (in press).

[3] B. Bouchon-Meunier, editor (1998). Studies in Fuzziness and Soft Computing vol. 12: Aggregation and Fusion of Imperfect Information, Physica-Verlag, Heidelberg.

[4] T. Calvo, A. Kolesárová, M. Komorniková, R. Mesiar (2002). Aggregation operators: Properties, classes and construction methods, in: *Aggregation Operators*, T. Calvo et al. (Eds.), Physica-Verlag, Heildelberg, pp 3-104.

[5] T. Calvo, J. Martín, G. Mayor (2013). Aggregation of implication functions, in: *Proceedings of the 8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-2013)*, G. Pasi et al. (Eds.), Atlantis Press, pp 569-574.

[6] B. De Baets, R. Mesiar (1998). T-partitions, *Fuzzy Sets and Systems* **97**, pp 211-223.

[7] J. Drewniak J., A. Król. (2010) A survey of weak connectives and the preservation of their properties by aggregations, *Fuzzy Sets and Systems* **161**, pp 202-215.

[8] D. Dubois, H. Prade (2004). On the use of aggregation operations in information fusion processes, *Fuzzy Sets and Systems* **142**, pp 143-161.

[9] U. Dudziak (2006). Development of the mean concept, *Journal of Mathematics and Applications* **28**, pp 25-39.

[10] U. Dudziak, A. Krl (2013). Aggregation of fuzzy equivalences, in: *New Trends in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics vol. 1: Foundations*, K.T. Atanassov et al. (Eds.), IBS PAN - SRI PAS, Warszawa, pp 85-103.

[11] J. C. Fodor, M. Roubens (1994). Fuzzy preference modelling and multicriteria decision support, Kluwer, Dordrecht.

[12] E. P. Klement, R. Mesiar, E. Pap (2000). Triangular norms, Kluwer Academic Publishers, Dordrecht.

[13] H.T. Nguyen, E. Walker (1996). A first course in fuzzy logic, CRC Press, Boca Raton.

[14] V. Peneva, I. Popchev (2003). Properties of the aggregation operators related with fuzzy relations, *Fuzzy Sets and Systems* **139**, pp 615-633.

[15] S. Saminger, R. Mesiar, U. Bodenhofer (2002). Domination of aggregation operators and preservation of transitivity, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **10** (Suppl.), pp 11-35.

MEASURES ON INTERVAL-VALUED FUZZY SETS AND IF-SETS AND PRINCIPLE OF INCLUSION AND EXCLUSION

Mária Kuková
Mensa Secondary School
Prague
Czech Republic
maja.kukova@gmail.com

Mirko Navara
Faculty of Electrical Engineering
Czech Technical University in Prague
Czech Republic
navara@cmp.felk.cvut.cz

Summary

In our previous work, we proved that the only continuous operations on fuzzy sets fulfilling the principle of inclusion and exclusion are Gödel and product operations and their ordinal sums. We have made a similar observation on t-representable measures on interval-valued fuzzy sets and IF-sets. Here we complete these results and discuss also operations which are not t-representable.

Keywords: Principle of inclusion and exclusion, fuzzy set, interval-valued fuzzy set, IF-set, (Atanassov's) intuitionistic fuzzy set.

1 THE PRINCIPLE OF INCLUSION AND EXCLUSION

We say that a function m and operations $\dot\cup, \underset{\cdot}{\cap}$ satisfy the principle of inclusion and exclusion if the following equation holds for any $A_1, \ldots, A_n$:

$$m\Big(\dot{\bigcup}_{i=1}^{n} A_i\Big) - \sum_{i=1}^{n} m(A_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} m(A_i \underset{\cdot}{\cap} A_j) - \cdots + (-1)^n\, m\Big(\underset{\cdot}{\bigcap}_{i=1}^{n} A_i\Big) = 0\,. \quad (1)$$

For $n = 2$, (1) reduces to the *valuation property* (see [2] or [17])

$$m(A \,\dot\cup\, B) - m(A) - m(B) + m(A \underset{\cdot}{\cap} B) = 0\,.$$

The classical result is that the principle of inclusion and exclusion holds if $\dot\cup, \underset{\cdot}{\cap}$ are the set-theoretical union and intersection and m is the cardinality of sets or a measure on an algebra of sets.

Example 1.1 *Formula (1) holds if $A_1, \ldots, A_n$ are numbers, m the identity mapping, and $\dot\cup, \underset{\cdot}{\cap}$ one of the following couples of operations:*

1. *the disjunction and conjunction on $\{0, 1\}$,*
2. *the maximum and minimum on any set of reals,*
3. *the product t-conorm (probabilistic sum) $x \overset{\mathrm{P}}{\vee} y = x + y - x \cdot y$ and the product on $[0, 1]$.*

We shall show that these results have the same principle, which can be expressed by their common generalization, the principle of inclusion and exclusion for fuzzy sets. Moreover, we further generalize it to IV-sets (interval-valued fuzzy sets) and IF-sets and show the limitations of this principle.

2 BASIC NOTIONS

Let us first introduce basic notions on fuzzy sets. See [11, 18, 19] for more details.

2.1 Fuzzy set operations

We fix a non-empty set (universe) X and a σ-algebra $\mathcal{S}$ of subsets of X. A fuzzy set A is described by an $\mathcal{S}$-measurable membership function $\mu_A \colon X \to [0, 1]$. The set of all $\mathcal{S}$-measurable fuzzy subsets of X will be denoted by $\mathcal{F}$. In particular $1_X, 0_X \in \mathcal{F}$ are the constant functions on X with values $1, 0$, respectively. An element $A \in \mathcal{F}$ is called *sharp* if $A \colon X \to \{0, 1\}$. They form a subalgebra of $\mathcal{F}$ isomorphic to the σ-algebra $\mathcal{S}$.

We shall restrict attention to *continuous* operations (fuzzy unions and intersections), thus the $\mathcal{S}$-measurability is preserved. The same applies to generalizations of fuzzy sets studied in the following sections. Most of the preceding papers studied only the collections of *all* fuzzy subsets of the given universe, i.e., the special case where $\mathcal{S}$ is the power set of X.

ISBN: 978–83–8012–519–3

We denote by $\dot\wedge$ and $\dot\vee$ a general t-norm and t-conorm. Particular types will be distinguished by indices in place of dots. We shall use mainly the following t-norms:

$$\begin{aligned} x \overset{\mathrm{G}}{\wedge} y &= \min(x,y)\,, && \text{(Gödel)} \\ x \overset{\mathrm{P}}{\wedge} y &= x \cdot y\,, && \text{(product)} \\ x \overset{\mathrm{L}}{\wedge} y &= \max(x+y-1,0)\,. && \text{(Łukasiewicz)} \end{aligned}$$

The respective t-conorms, $\overset{\mathrm{G}}{\vee}, \overset{\mathrm{P}}{\vee}, \overset{\mathrm{L}}{\vee}$, are obtained by duality with respect to the standard fuzzy negation $\neg\alpha = 1-\alpha$. Fuzzy unions and intersections are defined by the pointwise application of the respective t-conorms and t-norms:

$$\mu_{A \dot\cup B}(x) = \mu_A(x) \dot\vee \mu_B(x)\,, \tag{2}$$
$$\mu_{A \dot\cap B}(x) = \mu_A(x) \dot\wedge \mu_B(x)\,. \tag{3}$$

The notation $\dot\cup$ and $\dot\cap$ is used for a general union and intersection. Particular types will be distinguished by the same indices as the corresponding t-conorms and t-norms.

For the definition of a measure, the notion of convergence will be needed. The sign $\nearrow$ will appear in connection with objects of two different types, but in a similar sense: For fuzzy sets A, A_i $(i \in \mathbb{N})$,

$$A_i \nearrow A \iff \Big(A_i \subseteq A_{i+1}\ \forall i \in \mathbb{N} \text{ and } \overset{\mathrm{G}}{\bigcup_{i=1}^{\infty}} A_i = A\Big)$$

and for $\alpha, \alpha_i \in \mathbb{R}$ $(i \in \mathbb{N})$,

$$\alpha_i \nearrow \alpha \iff \Big(\alpha_i \le \alpha_{i+1}\ \forall i \in \mathbb{N} \text{ and } \bigvee_{i=1}^{\infty} \alpha_i = \alpha\Big)\,.$$

2.2 Interval-valued fuzzy sets and IF-sets

There are more generalizations of fuzzy sets. One of them is the concept of interval-valued fuzzy sets proposed by Zadeh (see [29]), which is based on the idea of ill-known membership degree. An *interval-valued fuzzy set* (*IV-set*) A on the universe X is described by two $\mathcal{S}$-measurable functions $\mu_A, \varrho_A \colon X \to [0,1]$, where $\mu_A(x) \le \varrho_A(x)$ for all $x \in X$. Another approach is the concept of *IF-sets* (derived from Atanassov's *intuitionistic fuzzy sets*[1]). An IF-set is given by two $\mathcal{S}$-measurable functions $\mu_A, \nu_A \colon X \to [0,1]$ satisfying the condition $\mu_A(x) + \nu_A(x) \le 1$ for all $x \in X$. Functions μ_A and ν_A are called the membership function and the non-membership function, respectively.

[1]The name "intuitionistic fuzzy sets" was introduced by K. Atanassov [1]. Due to criticism of this terminology in [9], we use the term "IF-sets".

Although the motivation is different, formally the IV-sets and IF-sets are isomorphic. The bijection is given by

$$\varrho_A = 1_X - \nu_A \tag{4}$$

and vice versa. Thus we consider these as two alternative descriptions of the same object. For an IV-set (or IF-set) A, functions μ_A, ν_A, ϱ_A are defined and linked by (4). The set of all such objects (IV-sets or IF-sets on a universe X) will be denoted by $\mathcal{IF}$. The following results are mostly formulated for IV-sets, while IF-sets were used, e.g. in [12]. Ordinary fuzzy sets can be canonically embedded into $\mathcal{IF}$ by the homomorphism which, for a fuzzy set $A \in \mathcal{F}$ with a membership function μ_A, defines additionally $\varrho_A = \mu_A$, $\nu_A = 1_X - \mu_A$. Thus we consider $\mathcal{F}$ as a subset of $\mathcal{IF}$ containing *ordinary fuzzy sets* (such an $A \in \mathcal{F}$ satisfies $\mu(A) = \varrho(A) = 1_X - \nu(A)$). The elements of $\mathcal{IF} \setminus \mathcal{F}$ are called *genuine IV-sets*.

When seen as interval-valued fuzzy sets, elements of $\mathcal{IF}$ can be alternatively represented by $\mathcal{S}$-measurable mappings from X into the set $\mathcal{J} = \{[\mu, \varrho] : 0 \le \mu \le \varrho \le 1\}$ of all closed subintervals of $[0,1]$. The partial ordering on $\mathcal{J}$ is componentwise, i.e.

$$[\mu_1, \varrho_1] \le [\mu_2, \varrho_2] \iff \mu_1 \le \mu_2 \text{ and } \varrho_1 \le \varrho_2\,.$$

The partial ordering is extended to $\mathcal{IF}$ so that $A \le B$ iff any of the following two equivalent conditions holds:

$$\mu_A \le \mu_B \text{ and } \varrho_A \le \varrho_B\,,$$
$$\mu_A \le \mu_B \text{ and } \nu_A \ge \nu_B\,.$$

We denote by Z the least element of $\mathcal{IF}$,

$$\mu_Z = 0_X\,, \qquad \varrho_Z = 0_X\,, \qquad \nu_Z = 1_X\,,$$

and by U the greatest element,

$$\mu_U = 1_X\,, \qquad \varrho_U = 1_X\,, \qquad \nu_U = 0_X\,.$$

An intersection $\dot\cap$ and union $\dot\cup$ on $\mathcal{IF}$ are defined by the pointwise application of the respective t-norm $\dot\sqcap$ and t-conorm $\dot\sqcup$ on $\mathcal{J}$:

$$\mu_{A \dot\cap B}(x) = \mu_A(x) \dot\sqcap \mu_B(x)\,, \tag{5}$$
$$\mu_{A \dot\cup B}(x) = \mu_A(x) \dot\sqcup \mu_B(x)\,, \tag{6}$$

where t-norms and t-conorms on $\mathcal{J}$ are defined by the usual requirements: commutativity, associativity, monotonicity, and neutral elements $[1,1]$ (of $\dot\sqcap$) and $[0,0]$ (of $\dot\sqcup$). We distinguish by notation t-norms and t-conorms on $[0,1]$ and on $\mathcal{J}$ because the former are used in definitions of the latter ones.

3 MEASURES

3.1 Measures on systems of fuzzy sets

Definition 3.1 *[2] A mapping $m\colon \mathcal{F} \to [0,1]$ is called a* state *if the following properties are satisfied:*

(M1) $m(1_X) = 1,\ m(0_X) = 0,$

(M2) $m(A \overset{Ł}{\cup} B) = m(A) + m(B) - m(A \underset{Ł}{\cap} B),$

(M3) $A_n \nearrow A \Rightarrow m(A_n) \nearrow m(A).$

States were characterized by D. Butnariu and E. P. Klement in [2]:

Theorem 3.2 *Every state m is of the form*

$$m(A) = \int \mu_A \,\mathrm{d}P\,, \tag{7}$$

where P is a state (probability measure) on the Boolean σ-algebra of sharp elements of $\mathcal{F}$ (P is the restriction of m).

A state of the form (7) is also called an *integral state*. It has been used in many previous studies, even in the pioneering work by Zadeh [28]. However, it was usually introduced without any deeper motivation. The axiomatic approach of Butnariu and Klement proves that the only *states* are integral states. We refer to [27] for an overview of different approaches to measures on systems of fuzzy sets and to [17] for the explanation of the particular role of Łukasiewicz operations in the definition of a state.

3.2 Measures on systems of IV-sets

The following definition comes from Riečan [20].

Definition 3.3 *A mapping $m\colon \mathcal{IF} \to [0,1]$ is called a* state *if the following properties are satisfied for all $A, B, A_n \in \mathcal{IF}$ ($n \in \mathbb{N}$):*

1. $m(U) = 1,\ m(Z) = 0,$
2. $A \underset{Ł}{\cap} B = Z \Rightarrow m(A \overset{Ł}{\cup} B) = m(A) + m(B),$
3. $A_n \nearrow A \Rightarrow m(A_n) \nearrow m(A),$

In [4], Ciungu and Riečan have proved the following representation theorem (see also [5, 22, 24, 26]):

Theorem 3.4 *[4] For any state $m\colon \mathcal{IF} \to [0,1]$ there exist probability measures $P, Q\colon \mathcal{S} \to [0,1]$ and $\alpha \in [0,1]$ such that $P \ge \alpha\, Q$ and, for all $A \in \mathcal{IF}$,*

$$m(A) = \int_X \mu_A \,\mathrm{d}P + \alpha \left(1 - \int_X (\mu_A + \nu_A)\,\mathrm{d}Q\right). \tag{8}$$

We need another formulation:

Theorem 3.5 *For any state $m\colon \mathcal{IF} \to [0,1]$ there exist probability measures $R, Q\colon \mathcal{S} \to [0,1]$ and $\alpha \in [0,1]$ such that, for all $A \in \mathcal{IF}$,*

$$m(A) = (1-\alpha) \int_X \mu_A \,\mathrm{d}R + \alpha \int_X \varrho_A \,\mathrm{d}Q\,.$$

Proof: We take the representation from Theorem 3.4. Then we find a state R satisfying

$$P = \alpha\, Q + (1-\alpha)\, R\,.$$

If $\alpha = 1$, R can be arbitrary (e.g., Q). Otherwise, we define

$$R = \frac{P - \alpha\, Q}{1-\alpha}\,.$$

Due to the assumption $P \ge \alpha\, Q$, R is non-negative. For each $B \in \mathcal{S}$, its complement, $B' = X \setminus B$, satisfies

$$\begin{aligned} 0 &\le (1-\alpha)\, R(B') = P(B') - \alpha\, Q(B') \\ &= 1 - P(B) - \alpha\,(1 - Q(B)) \\ &= 1 - \alpha - \big(P(B) - \alpha\, Q(B)\big) \le 1 - \alpha\,. \end{aligned}$$

Thus $R \le 1$. The σ-additivity and other conditions are obviously preserved, so R is a state. We reformulate (8):

$$\begin{aligned} m(A) &= \int_X \mu_A \,\mathrm{d}P + \alpha \left(1 - \int_X (\mu_A + \nu_A)\,\mathrm{d}Q\right) \\ &= \int_X \mu_A \,\mathrm{d}P - \alpha \int_X \mu_A \,\mathrm{d}Q + \alpha \left(1 - \int_X \nu_A \,\mathrm{d}Q\right) \\ &= \int_X \mu_A \,\mathrm{d}(P - \alpha\, Q) + \alpha \int_X (1 - \nu_A)\,\mathrm{d}Q \\ &= (1-\alpha) \int_X \mu_A \,\mathrm{d}R + \alpha \int_X \varrho_A \,\mathrm{d}Q\,. \end{aligned}$$

□

Theorem 3.5 says that each state is a convex combination of two states, one depending only on μ_A, the other only on ϱ_A for each A. We acknowledge that this observation was made, under different formulation, also in [22].

In particular, all constant functions $X \to \mathcal{J}$ form a sublattice of $\mathcal{S}$ isomorphic to $\mathcal{J}$. For a state m on $\mathcal{J}$, there is an $\alpha \in [0,1]$ such that

$$m([\mu, \varrho]) = (1-\alpha)\,\mu + \alpha\,\varrho\,.$$

This means that a state on $\mathcal{J}$ is just a convex combination of the entries.

As an alternative, Grzegorzewski and Mrówka [10] defined a *probability* $\mathcal{P}$ of each $A \in \mathcal{IF}$ by the interval

$$\mathcal{P}(A) = \left[\int_X \mu_A \,\mathrm{d}P, 1 - \int_X \nu_A \,\mathrm{d}P\right],$$

where P is a probability measure on the σ-algebra $\mathcal{S}$. A more general axiomatic approach to probability on $\mathcal{IF}$ was proposed by Riečan [21]. The main conclusion is that $\mathcal{P}(A) = \left[\mathcal{P}^\flat(A), \mathcal{P}^\sharp(A)\right]$, where $\mathcal{P}^\flat, \mathcal{P}^\sharp \colon \mathcal{F} \to [0,1]$ are states and $\mathcal{P}^\flat \le \mathcal{P}^\sharp$. The principal conclusion is that both bounds depend linearly on $\mu(A), \varrho(A)$. Here we restrict attention to states and refer to [12, 13] for details on the approach based on probabilities.

4 THE PRINCIPLE OF INCLUSION AND EXCLUSION FOR FUZZY SETS

It is natural to ask whether the principle of inclusion and exclusion holds for *fuzzy sets*. This question has many aspects because we can consider different fuzzy unions and intersections. We ask which operations $\dot\cup, \underset{\cdot}{\cap}$ satisfy (1) for *all* $A_1, \ldots, A_n$ *and all states* m. We require this for *all underlying* σ*-algebras* $\mathcal{S}$, thus avoiding trivial cases satisfying the principle of inclusion and exclusion because of degenerate $\mathcal{S}$.

Using the integral representation of states, the principle of inclusion and exclusion requires that the following expression must be zero:

$$\begin{aligned} 0 &= m\Big(\dot{\bigcup}_{i=1}^{n} A_i\Big) - \sum_{i=1}^{n} m(A_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} m(A_i \underset{\cdot}{\cap} A_j) \\ &\quad - \cdots + (-1)^n\, m\Big(\underset{\cdot}{\bigcap}_{i=1}^{n} A_i\Big) \\ &= \int \Big(\mu_{\dot\bigcup_{i=1}^{n} A_i} - \sum_{i=1}^{n} \mu_{A_i} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mu_{A_i \underset{\cdot}{\cap} A_j} \\ &\quad - \cdots + (-1)^n\, \mu_{\underset{\cdot}{\bigcap}_{i=1}^{n} A_i}\Big)\, \mathrm{d}P \\ &= \int \Big(\dot{\bigvee}_{i=1}^{n} \mu_{A_i} - \sum_{i=1}^{n} \mu_{A_i} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mu_{A_i} \underset{\cdot}{\wedge} \mu_{A_j} \\ &\quad - \cdots + (-1)^n\, \underset{\cdot}{\bigwedge}_{i=1}^{n} \mu_{A_i}\Big)\, \mathrm{d}P\,. \end{aligned}$$

This integral has to be zero for all fuzzy sets $A_1, \ldots, A_n$. This means that the integrand has to be zero P-almost everywhere:

$$\begin{aligned} 0 &= \dot{\bigvee}_{i=1}^{n} \mu_{A_i} - \sum_{i=1}^{n} \mu_{A_i} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mu_{A_i} \underset{\cdot}{\wedge} \mu_{A_j} \\ &\quad - \cdots + (-1)^n\, \underset{\cdot}{\bigwedge}_{i=1}^{n} \mu_{A_i}\,. \end{aligned}$$

The membership degrees can be replaced by any constants, $\mu_{A_i} := a_i$:

$$\begin{aligned} 0 &= \dot{\bigvee}_{i=1}^{n} a_i - \sum_{i=1}^{n} a_i + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_i \underset{\cdot}{\wedge} a_j \\ &\quad - \cdots + (-1)^n\, \underset{\cdot}{\bigwedge}_{i=1}^{n} a_i\,. \end{aligned} \tag{9}$$

Validity of equation (9) for any $a_1, \ldots, a_n \in [0,1]$ is equivalent to the validity of the principle of inclusion and exclusion in the form (1) for the respective fuzzy set operations. The complete answer for continuous fuzzy operations was given in [15] (we refer to [11, 25] for the definition and notation of ordinal sums of t-norms and t-conorms):

Theorem 4.1 *Let* $\dot\cup, \underset{\cdot}{\cap}$ *be a fuzzy union and intersection on* $\mathcal{F}$ *corresponding (by (2),(3)) to a continuous t-conorm* $\dot\vee$ *and a continuous t-norm* $\underset{\cdot}{\wedge}$*. Then* $\dot\cup, \underset{\cdot}{\cap}$ *satisfy the principle of inclusion and exclusion iff there is a (possibly empty) collection of disjoint intervals* $((a_\alpha, b_\alpha))_{\alpha\in I}$ *in* $[0,1]$ *such that* $\underset{\cdot}{\wedge}$ *and* $\dot\vee$ *are ordinal sums*

$$\underset{\cdot}{\wedge} \;=\; (\langle a_\alpha, b_\alpha, \underset{\mathrm{P}}{\wedge}\rangle)_{\alpha\in I}\,, \tag{10}$$

$$\dot\vee \;=\; (\langle a_\alpha, b_\alpha, \overset{\mathrm{P}}{\vee}\rangle)_{\alpha\in I}\,. \tag{11}$$

5 THE PRINCIPLE OF INCLUSION AND EXCLUSION FOR IV- AND IF-SETS

Further, the principle of inclusion and exclusion was generalized to IV- and IF-sets. In previous papers, e.g., [3, 4, 5, 10, 12, 23], only the Gödel, product, or Łukasiewicz operations were considered (with the positive answer for the first two and the negative answer for Łukasiewicz operations).

5.1 The principle of inclusion and exclusion for t-representable operations on $\mathcal{IF}$

For $A, B \in \mathcal{IF}$, t-representable intersections are computed by the following formulas:

$$\mu_{A\underset{\cdot}{\cap}B}(x) \;=\; \mu_A(x) \underset{1}{\wedge} \mu_B(x)\,, \tag{12}$$

$$\varrho_{A\underset{\cdot}{\cap}B}(x) \;=\; \varrho_A(x) \underset{2}{\wedge} \varrho_B(x)\,, \tag{13}$$

$$\nu_{A\underset{\cdot}{\cap}B}(x) \;=\; \nu_A(x) \overset{2}{\vee} \nu_B(x)\,,$$

where $\underset{1}{\wedge}, \underset{2}{\wedge}$ are t-norms on $[0,1]$ satisfying $\underset{1}{\wedge} \le \underset{2}{\wedge}$ and $\overset{2}{\vee}$ is the t-conorm dual to $\underset{2}{\wedge}$. Dually, t-representable unions are computed by

$$\mu_{A\dot\cup B}(x) \;=\; \mu_A(x) \overset{3}{\vee} \mu_B(x)\,, \tag{14}$$

$$\varrho_{A\dot\cup B}(x) \;=\; \varrho_A(x) \overset{4}{\vee} \varrho_B(x)\,, \tag{15}$$

$$\nu_{A\dot\cup B}(x) \;=\; \nu_A(x) \underset{4}{\wedge} \nu_B(x)\,,$$

where $\overset{3}{\vee}, \overset{4}{\vee}$ are t-conorms on $[0,1]$ satisfying $\overset{3}{\vee} \le \overset{4}{\vee}$ and $\underset{4}{\wedge}$ is the t-norm dual to $\overset{4}{\vee}$. These formulas are special cases of (5), (6) for a t-representable t-norm $\dot{\sqcap}$ and a t-representable t-conorm $\dot{\sqcup}$. Notice that t-representable operations may result in genuine IV-sets even if the arguments are ordinary fuzzy sets.

Only operations which satisfy the conditions of Theorem 4.1 give a chance to satisfy the principle of inclusion and exclusion on $\mathcal{IF}$. In [12], the principle of inclusion and exclusion on $\mathcal{IF}$ was proved for the Gödel and product operations, i.e. for the pair of operations $(\overset{G}{\cup}, \underset{G}{\cap})$ or $(\overset{P}{\cup}, \underset{P}{\cap})$ chosen for $(\dot{\cup}, \dot{\cap})$ in (1). We obtained the following generalization:

Theorem 5.1 *[16] Let $\dot{\cup}, \dot{\cap}$ be a fuzzy union and intersection on $\mathcal{IF}$ corresponding (by (12),(13),(14),(15)) to continuous t-norms $\underset{1}{\wedge}, \underset{2}{\wedge}$ and continuous t-conorms $\overset{3}{\vee}, \overset{4}{\vee}$. Then $\dot{\cup}, \dot{\cap}$ satisfy the principle of inclusion and exclusion iff there is a (possibly empty) collection of disjoint intervals $((a_\alpha, b_\alpha))_{\alpha \in I}$ in $[0,1]$ such that $\underset{1}{\wedge} = \underset{2}{\wedge}$ and $\overset{3}{\vee} = \overset{4}{\vee}$ are given by (10), (11), respectively.*

5.2 The principle of inclusion and exclusion for operations on $\mathcal{IF}$ which are not t-representable

There are t-norms and t-conorms on IF-sets which are not t-representable. The first example was published in [6], we generalize it:

Example 5.2 *Let $\dot{\vee}\colon [0,1]^2 \to [0,1]$ be a continuous t-conorm. Then the operation $\dot{\sqcup}\colon \mathcal{J}^2 \to \mathcal{J}$, defined by*

$$[\mu_1, \varrho_1] \,\dot{\sqcup}\, [\mu_2, \varrho_2] = \begin{cases} [\mu_1, \varrho_1] & \text{if } [\mu_2, \varrho_2] = [0,0], \\ [\mu_2, \varrho_2] & \text{if } [\mu_1, \varrho_1] = [0,0], \\ [\varrho_1 \dot{\vee} \varrho_2, \varrho_1 \dot{\vee} \varrho_2] & \text{otherwise,} \end{cases} \quad (16)$$

is a t-conorm on $\mathcal{J}$.

Notice that the t-conorm $\dot{\sqcup}$ from (16) results often in ordinary fuzzy sets even if the arguments are genuine IV-sets. The first published example of a t-conorm which is not t-representable [6] is its special case for $\dot{\vee} = \overset{G}{\vee}$ (Gödel t-conorm).

The remaining examples in [7, 8], when restricted to ordinary fuzzy sets, coincide with the Łukasiewicz operations. As the principle of inclusion and exclusion does not hold for Łukasiewicz operations on ordinary fuzzy sets, there is no chance to satisfy it for these operations on $\mathcal{IF}$. However, we generalized one of them:

Example 5.3 *Let $\dot{\vee}\colon [0,1]^2 \to [0,1]$ be a continuous t-conorm. Then the operation $\dot{\sqcup}\colon \mathcal{J}^2 \to \mathcal{J}$ defined by*

$$[\mu_1, \varrho_1] \,\dot{\sqcup}\, [\mu_2, \varrho_2] = \left[\min(\mu_1 \dot{\vee} \varrho_2, \varrho_1 \dot{\vee} \mu_2), \varrho_1 \dot{\vee} \varrho_2\right]$$

is a t-conorm on $\mathcal{J}$.

Nevertheless, none of the operations from Examples 5.2 and 5.3 satisfies the principle of inclusion and exclusion due to the following result:

Theorem 5.4 *Let a union $\dot{\cup}\colon \mathcal{IF}^2 \to \mathcal{IF}$ be based on a t-conorm $\dot{\sqcup}\colon \mathcal{J}^2 \to \mathcal{J}$ (via (6)). Suppose that there exist $\mu, \varrho, \varrho_1, \varrho_2$ such that*

$$[\mu, \varrho] = [0, \varrho_1] \,\dot{\sqcup}\, [0, \varrho_2]$$

and $\mu > 0$. Then there is no intersection $\dot{\cap}\colon \mathcal{IF}^2 \to \mathcal{IF}$ such that $(\dot{\cup}, \dot{\cap})$ satisfy the principle of inclusion and exclusion.

6 CONCLUSIONS

We have proved that the only continuous operations on fuzzy sets which satisfy the principle of inclusion and exclusion are the Gödel ones (minimum and maximum), the product operations, and some of their ordinal sums. The same holds for t-representable operations on IV-sets. We restricted the possibility of satisfying the principle of inclusion and exclusion by operations on IV-sets which are not t-representable.

Acknowledgements

The second author was supported by the Ministry of Education of the Czech Republic under Project RVO13000.

References

[1] K. Atanassov. *Intuitionistic Fuzzy Sets: Theory and Applications.* Springer-Verlag, Heidelberg, 1999.

[2] D. Butnariu and E. P. Klement. *Triangular Norm-Based Measures and Games with Fuzzy Coalitions.* Kluwer Academic Publishers, Dordrecht, 1993.

[3] L. Ciungu, J. Kelemenová, and B. Riečan. A New Point of View to the Inclusion - Exclusion Principle. In 6^{th} *IEEE International Conference on Intelligent Systems IS'12*, Varna, Bulgaria, 142–144, 2012.

[4] L. Ciungu and B. Riečan. General form of probabilities on IF-sets. In *Fuzzy Logic and Applications*, Proc. 8th Int. Workshop WILF, Lecture Notes in Artificial Intelligence, 101–107, 2009.

[5] L. Ciungu and B. Riečan. Representation theorem for probabilities on IFS-events. *Inform. Sci.*, 180:793–798, 2010.

[6] G. Deschrijver, C. Cornelis, and E.E. Kerre. Intuitionistic fuzzy connectives revisited. In: *Proc. 9th Int. Conf. Information Processing and Management of Uncertainty*, ESIA – Université de Savoie, Annecy, France, 1839–1844, 2002.

[7] G. Deschrijver, C. Cornelis, and E.E. Kerre. Classes of intuitionistic fuzzy t-norms satisfying the residuation principle. *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems* 11:691–709, 2003.

[8] G. Deschrijver, C. Cornelis, and E.E. Kerre. On the representation of intuitionistic fuzzy t-norms and t-conorms. *IEEE Trans Fuzzy Syst.*, 12:45–61, 2004.

[9] D. Dubois, S. Gottwald, P. Hájek, J. Kacprzyk, and H. Prade. Terminological difficulties in fuzzy set theory—The case of "Intuitionistic Fuzzy Sets", *Fuzzy Sets and Systems*, 156:485–491, 2005.

[10] P. Grzegorzewski and E. Mrówka. Probability of intuitionistic fuzzy events. In P. Grzegorzewski et al., editors, *Soft Methods in Probability, Statistics and Data Analysis*, Springer, New York, 105–115, 2002.

[11] E. P. Klement, R. Mesiar, and E. Pap. *Triangular Norms.* Kluwer Academic Publishers, Dordrecht, 2000.

[12] M. Kuková. The inclusion-exclusion principle for L-states and IF-events. *Inform. Sci.*, 224:165–169, 2013.

[13] M. Kuková. Probability on Algebraic Structures (in Slovak). PhD. thesis, Matej Bel University, Banská Bystrica, Slovakia, 2012.

[14] M. Kuková and M. Navara. Continuous t-norms and t-conorms satisfying the principle of inclusion and exclusion. In: H. Bustince, J. Fernandez, R. Mesiar, and T. Calvo (eds.), *Aggregation Functions in Theory and in Practise*, Springer, Heidelberg/New York/Dordrecht/London, 179–186, 2013.

[15] M. Kuková and M. Navara. Principles of inclusion and exclusion for fuzzy sets. *Fuzzy Sets Syst.* 232:98–109, 2013. DOI 10.1016/j.fss.2013.02.014

[16] M. Kuková and M. Navara. Principles of inclusion and exclusion for interval-valued fuzzy sets and IF-sets. To appear.

[17] M. Navara. Triangular norms and measures of fuzzy sets. In E.P. Klement and R. Mesiar, editors, *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier, 345–390, 2005.

[18] M. Navara and P. Olšák. *Basics of Fuzzy Sets* (in Czech). Czech Technical University in Prague, 2007.

[19] H.T. Nguyen and E. Walker. *A First Course in Fuzzy Logic.* Chapman & Hall/CRC, Boca Raton, 2nd edition, 2000.

[20] B. Riečan. On some contributions to quantum structures by fuzzy sets. *Kybernetika*, 43:481–490, 2007.

[21] B. Riečan. Probability theory on IF events. In S. Aguzzolii et al. (eds.), *Algebraic and Proof-Theoretic Aspects on Non-classical Logics, Lecture Notes on Comput. Sci.*, Springer, 290–308, 2007.

[22] B. Riečan. Analysis of fuzzy logic models. In V.M. Koleshko (ed.), *Intelligent Systems*, InTech, 219–244, 2012.

[23] B. Riečan and D. Mundici. Probability on MV-algebras. In E. Pap, editor, *Handbook of Measure Theory.* Elsevier Science, Amsterdam, Chapter 21, 869–910, 2002.

[24] B. Riečan and J. Petrovičová. On the Łukasiewicz Probability Theory on IF-sets. *Tatra Mt. Math. Publ.*, 46:125–146, 2010.

[25] B. Schweizer and A. Sklar. *Probabilistic Metric Spaces.* North-Holland, New York, 1983.

[26] V. Skřivánek. States on IF-events. Lecture at *Twelfth Int. Conf. Fuzzy Sets Theory Appl.*, Liptovský Ján, Slovak University of Technology in Bratislava, Slovakia, 2014.

[27] S. Weber and E. P. Klement. Fundamentals of a generalized measure theory. In U. Höhle et al., editors, *Mathematics of Fuzzy Sets. Logic, Topology, and Measure Theory*, Kluwer Academic Publishers, Dordrecht, 633–651, 1999.

[28] L. A. Zadeh. Probability measures of fuzzy sets, *J. Math. Anal. Appl.*, 23:421–427, 1968.

[29] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I, *Inform. Sci.*, 8:199–249, 1975.

A NOTE ON THE KRUSE INTEGRAL BASED ON λ-MEASURES

Jun Li*, **Suqin Min and Xiaoli Hu**
School of Science,
Communication University of China,
Beijing 100024, China
{lijun,msq,huxiaoli}@cuc.edu.cn

Summary

In this note, we discuss a pair of special pseudo-addition and pseudo-multiplication associated with the parameter λ ($\lambda > 0$), which constitutes a commutative isotonic semiring. We show that on λ-measure spaces (special types of monotone measure spaces) the pan-integral based on the commutative isotonic semiring coincides with the fuzzy integral introduced by Kruse [R. Kruse, Fuzzy Sets and Systems 10(1983)].

Keywords: λ-measure, Pan-integral, Kruse integral.

1 INTRODUCTION

In [20] the pan-integral was established based on a special type of commutative isotonic semiring $(\overline{R}_+, \oplus, \odot)$ (see also [19]). Sugeno and Murofushi [17] and Ichihashi *et al.* [3] defined the operations of pseudo-addition and pseudo-multiplication, respectively, and established integrals based on their own operations, respectively. The several kinds of common integrals, as the Riemann integral, the Lebesgue integral, the Sugeno integral [16] and Shilkret integral [15], were all special instance of pan-integral. A related concept of generalizing Lebesgue integral based on a generalized ring $(\overline{R}_+, \oplus, \otimes)$ (the commutativity of $\otimes$ is not required) , which is called *generalized Lebesgue integral*, was proposed and discussed in [21]. The abovementioned several types of integrals are all covered by the generalized Lebesgue integral [21].

In 1982 Kruse [5] showed that there exists a relationship between probability measures and λ-additive measures. This relationship was used to give the definition of a so called "fuzzy integral" of a fuzzy event with respect to λ-additive fuzzy measure, which generalizes the Lebesgue integral canonically ([6]). We call so kinds of fuzzy integrals introduced by Kruse [6] as *Kruse integral* (for short, K-integral).

In this paper, we introduce a pair of special pseudo-addition and pseudo-multiplication $(\oplus_\lambda, \otimes_\lambda)$ associated with the parameter λ ($\lambda > 0$). They form a commutative isotonic semiring $(\overline{R}_+, \oplus_\lambda, \otimes_\lambda)$. We can use the pseudo-addition $\oplus_\lambda$ to reformulate λ-additive measures. We shall show that on λ-measure spaces (see [19], special typse of monotone measure spaces) the Kruse-integral coincides with the pan-integral based on $(\overline{R}_+, \oplus_\lambda, \otimes_\lambda)$, i.e., the Kruse-integral is a special type of pan-integral.

*Corresponding author. Tel./fax: +86-10-65783583.

2 PAN-INTEGRALS

Let $X \neq \emptyset$ be a universe of discourse, Σ a σ-algebra of subsets of X, $\mu : \Sigma \to [0, \infty]$ be a monotone set function with $\mu(\emptyset) = 0$ (it is also known as *monotone measure* [19]), $\mathcal{F}_+$ denote the set of all finite nonnegative Σ-measurable functions on X, and $\hat{\mathcal{P}}$ be the set of all finite partitions of (X, Σ). Unless stated otherwise, all the subsets mentioned are supposed to belong to Σ.

The concept of a pan-integral [19, 20] involves two binary operations, the pseudo-addition $\oplus$ and pseudo-multiplication $\otimes$ of real numbers. We recall a commutative isotonic semiring [19] (see also [1, 2, 3, 8, 9, 13, 17, 18]). Without mentioning explicitly, in this paper, $R_+ = [0, +\infty)$, $\overline{R}_+ = [0, +\infty]$, $a, b, c, d, a_i, b_i (i = 1, 2, \cdots)$ and $a_t (t \in T$, where T is any given index set) are all elements in $\overline{R}_+$.

Definition 2.1 *A binary operation $\oplus$ on $\overline{R}_+$ is called a pseudo-addition on $\overline{R}_+$ if and only if it satisfies the following requirements:*
(PA1) $a \oplus b = b \oplus a$;

ISBN: 978–83–8012–519–3

(PA2) $(a \oplus b) \oplus c = a \oplus (b \oplus c)$;
(PA3) $a \le b \Longrightarrow a \oplus c \le b \oplus c$ *for any* c;
(PA4) $a \oplus 0 = a$;
(PA5) $a_n \to a$ *and* $b_n \to b \Rightarrow a_n \oplus b_n \to a \oplus b$.

From associativity (PA2), we may write $a_1 \oplus a_2 \oplus \cdots \oplus a_n$ as $\bigoplus_{i=1}^{n} a_i$, and denote $\bigoplus_{i=1}^{\infty} a_i = \lim_{n\to\infty} \bigoplus_{i=1}^{n} a_i$.

Definition 2.2 *A binary operation* $\otimes$ *on* $\overline{R}_+$ *is called a pseudo-multiplication (with respect to pseudo-addition* $\oplus$*) on* $\overline{R}_+$ *if and only if it fulfills the following conditions:*
(PM1) $a \otimes b = b \otimes a$;
(PM2) $(a \otimes b) \otimes c = a \otimes (b \otimes c)$;
(PM3) $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$;
(PM4) $a \le b \Longrightarrow a \otimes c \le b \otimes c$ *for any* c;
(PM5) $a \otimes b = 0 \Longleftrightarrow a = 0$ *or* $b = 0$;
(PM6) there exists $e \in \overline{R}_+$*, such that* $e \otimes a = a$ *for any* $a \in \overline{R}_+$;
(PM7) $a, b \in (0, +\infty)$, $a_n \to a$ *and* $b_n \to b \Longrightarrow a_n \otimes b_n \to a \otimes b$.

When $\oplus$ is a pseudo-addition on $\overline{R}_+$ and $\otimes$ is a pseudo-multiplication (with respect to $\oplus$) on $\overline{R}_+$, the triple $(\overline{R}_+, \oplus, \otimes)$ is called a *a commutative isotonic semiring* (on $\overline{R}_+$), and $(X, \Sigma, \mu, \overline{R}_+, \oplus, \otimes)$ is called a *pan-space* [19].

Definition 2.3 *([19, 20]) Let* $(X, \Sigma, \mu, \overline{R}_+, \oplus, \otimes)$ *be a pan-space. For any* $f \in \mathcal{F}_+$, $A \in \Sigma$*, the pan-integral of* f *over* A *with respect to* μ*, is defined by*

$$(Pan)\int_A^{(\oplus,\otimes)} f d\mu = \sup_{\mathcal{E}\in\hat{\mathcal{P}}} \big\{ s_f(\mathcal{E} \mid_A, \oplus, \otimes) \big\}, \qquad (2.1)$$

where

$$s_f(\mathcal{E} \mid_A, \oplus, \otimes) = \bigoplus_{E\in\mathcal{E}} \big[\big(\inf_{x\in A\cap E} f(x) \big) \otimes \mu(A \cap E) \big].$$

Note that in the case commutative isotonic semiring $(\overline{R}_+, \vee, \wedge)$, Sugeno integral [16] is recovered, while for $(\overline{R}_+, \vee, \cdot)$, Shilkret integral [15] is covered by formula (2.1).

3 λ-MEASURES AND THE KRUSE INTEGRALS

In this section, we shall discuss a pair of special pseudo-addition and pseudo-multiplication $(\oplus_\lambda, \otimes_\lambda)$ (associated with the parameter λ, $\lambda > 0$), and use the pseudo-addition $\oplus_\lambda$ to reformulate λ-additive measures. We recall a kinds of fuzzy integrals with respect to λ-additive measure introduced by Kruse [6]. Theorem 3.1 presents our main result. It is shown that on λ-measure spaces ($\lambda > 0$) the Kruse-integral and the pan-integral based on $(\overline{R}_+, \oplus_\lambda, \otimes_\lambda)$ coincide.

Let (X, Σ) be a measurable space, $\lambda \ge 0$. A set function $g_\lambda : \Sigma \to \overline{R}_+$ is called

(i) λ-*additive* [16, 19], if for any $A, B \in \Sigma, A \cap B = \emptyset$, we have

$$g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B);$$

(ii) σ-λ *additive* [6, 16, 19], if for any disjoint sequence of sets $\{A_n\}$ in Σ,

$$g_\lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \begin{cases} \frac{1}{\lambda}\{\prod_{n=1}^{\infty}[1 + \lambda g_\lambda(A_n)] - 1\} & \lambda \ne 0, \\ \sum_{n=1}^{\infty} g_\lambda(A_n) & \lambda = 0. \end{cases}$$

(iii) λ-*measure* [19], if g_λ is σ-λ additive with $g_\lambda(\emptyset) = 0$.

In the following we introduce a pair of special pseudo-addition and pseudo-multiplication on $\overline{R}_+$. For any given $\lambda > 0$, we define two binary operators $\oplus_\lambda$ and $\otimes_\lambda$ on $\overline{R}_+$, as follows:

$$a \oplus_\lambda b = \begin{cases} a + b + \lambda ab, & a, b \in [0, +\infty), \\ +\infty, & \text{otherwise}, \end{cases}$$

and

$$a \otimes_\lambda b = \begin{cases} \frac{1}{\lambda}\big[(1 + \lambda a)^{\log_{1+\lambda}(1+\lambda b)} - 1\big], & \text{if } a, b \in [0, +\infty), \\ 0, & \text{if } a = 0,\ b = +\infty \text{ or } a = +\infty,\ b = 0, \\ +\infty, & \text{otherwise}. \end{cases}$$

It is easy to verify that $\oplus_\lambda$ is a pseudo-addition in the sense of Definition 2.1, while $\otimes_\lambda$ is a pseudo-multiplication (with respect to $\oplus_\lambda$) in the sense of Definition 2.2. Thus $(\overline{R}_+, \oplus_\lambda, \otimes_\lambda)$ is a commutative isotonic semiring on $\overline{R}_+$.

In particular, we have

$$\bigoplus_{n=1}^{\infty}{}_\lambda\, a_i = \frac{1}{\lambda}\Big[\prod_{i=1}^{\infty}(1 + \lambda a_i) - 1\Big] \qquad (3.1)$$

and, for any $a_n \ge 0, b_n < +\infty, n = 1, 2, \cdots$

$$\bigoplus_{n=1}^{\infty}{}_\lambda\, \big(a_i \otimes_\lambda b_i\big) = \frac{1}{\lambda}\Big[\prod_{n=1}^{\infty}(1 + \lambda a_n)^{\log_{1+\lambda}(1+\lambda b_n)} - 1\Big]. \qquad (3.2)$$

We can reformulate λ-additivity and λ-measure by using the pseudo-addition $\oplus_\lambda$.

Proposition 3.1 *Let $\lambda > 0$.*
(i) g_λ is λ-additive if and only if g_λ is $\oplus_\lambda$-additive, i.e.,

$$g_\lambda(A \cup B) = g_\lambda(A) \oplus_\lambda g_\lambda(B).$$

(ii) g_λ is λ-measure if and only if g_λ is σ-$\oplus_\lambda$-decomposable measure, i.e., $\mu(\emptyset) = 0$ and for any disjoint sequence of sets $\{A_n\}$,

$$g_\lambda(\bigcup_{n=1}^{\infty} A_n) = \bigoplus_{n=1}^{\infty}{}_\lambda\, g_\lambda(A_n).$$

Proposition 3.2 *Let $h(x)$ be an extended real valued function on $\overline{R}_+$ and be defined by*

$$h(x) = \begin{cases} \log_{1+\lambda}(1+\lambda x) & x \in R_+, \\ \infty & x = +\infty, \end{cases}$$

then the inverse of h is

$$h^{-1}(x) = \begin{cases} \frac{1}{\lambda}[(1+\lambda)^x - 1] & x \in R_+, \\ \infty & x = +\infty. \end{cases}$$

Furthermore, for a_i and $b_i \in R_+, i = 1, 2, \cdots, n$, we have

(i) $h\big[\bigoplus_{i=1}^{n}{}_\lambda\, (a_i \bigotimes_\lambda b_i)\big] = \sum_{i=1}^{n} h(a_i) \cdot h(b_i)$;

(ii) $h^{-1}\big[\sum_{i=1}^{n}(a_i b_i)\big] = \bigoplus_{i=1}^{n}{}_\lambda\, \big[h^{-1}(a_i) \bigotimes_\lambda h^{-1}(b_i)\big]$;

(iii) h and h^{-1} are both self-isomorphic on $(\overline{R}_+, \vee)$ and $(\overline{R}_+, \wedge)$, i.e., for $\{a_t\}_{t\in T} \subset \overline{R}_+$, where T is an arbitrary index set, we have

$$\begin{aligned} h(\bigvee_{t\in T})a_t &= \bigvee_{t\in T} h(a_t), \\ h(\bigwedge_{t\in T})a_t &= \bigwedge_{t\in T} h(a_t), \\ h^{-1}(\bigvee_{t\in T})a_t &= \bigvee_{t\in T} h^{-1}(a_t), \\ h^{-1}(\bigwedge_{t\in T})a_t &= \bigwedge_{t\in T} h^{-1}(a_t). \end{aligned}$$

Proof. (i) follows directly from (3.2). So we have merely to show (ii) and (iii) holds.

(ii) From (i), we can immediately obtain that

$$h\Big\{\bigoplus_{i=1}^{n}{}_\lambda\, [h^{-1}(a_i) \otimes_\lambda h^{-1}(b_i)]\Big\} = \sum_{i=1}^{n} a_i \cdot b_i.$$

So we have

$$\bigoplus_{i=1}^{n}{}_\lambda \big[h^{-1}(a_i) \otimes_\lambda h^{-1}(b_i)\big] = h^{-1}\big(\sum_{i=1}^{n} a_i \cdot b_i\big).$$

(iii) We only prove the first equality, and the others can be proved similarly. On the one hand, suppose that $\{a_n\}$ is a subsequence of $\{a_t\}_{t\in T}$ satisfying $a_n \uparrow \bigvee_{t\in T} a_t$. Thus, from the continuity of h we have

$$h(\bigvee_{t\in T} a_t) = \lim_{n\to\infty} h(a_n) \le \bigvee_{t\in T} h(a_t).$$

On the other hand, since $\forall\ t \in T$, $\bigvee_{t\in T} a_t \ge a_t$, it follows from the monotone increasing of h that $\forall\ t \in T$, $h(\bigvee_{t\in T} a_t) \ge h(a_t)$. Therefore we have

$$h(\bigvee_{t\in T} a_t) \ge \bigvee_{t\in T} h(a_t).$$

The proof is completed.

The following result is due to Kruse [5].

Proposition 3.3 *Let g_λ be a λ-measure on Σ. Denote*

$$g^*_\lambda(A) = \log_{1+\lambda}\Big(1 + \lambda g_\lambda(A)\Big), \quad \forall A \in \Sigma, \qquad (3.3)$$

then g^*_λ is a classical measure on Σ. On the contrary, if m is a classical measure on Σ, then m_λ defined by

$$m_\lambda(A) = \frac{1}{\lambda}\Big[\big(1+\lambda\big)^{m(A)} - 1\Big], \quad \forall A \in \Sigma, \qquad (3.4)$$

is a λ-measure on Σ.

In [6] Kruse introduced a special type of fuzzy integral by using the relationship between classical measures and λ-measures described in Proposition 3.3. We call so kinds of fuzzy integrals as *Kruse integral* (for short, K-integral), as follows:

Definition 3.1 *Let (X, Σ) be a measurable space, g_λ a λ-measure on Σ, $f \in F_+$, and $A \in \Sigma$, then the Kruse integral of f over A, is defined by*

$$(K)\int_A f dg_\lambda = \frac{1}{\lambda}\left[\big(1+\lambda\big)^{\int_A \log_{1+\lambda}(1+\lambda f) dg^*_\lambda} - 1\right],$$

where g^*_λ is described by (3.3), and $\int_A \log_{1+\lambda}(1 + \lambda f) dg^*_\lambda$ is the Lebesgue integral of $\log_{1+\lambda}(1+\lambda f)$ over A with respect to g^*_λ.

The following is our main result.

Theorem 3.1 *Let $(X, \Sigma, g_\lambda, \overline{R}_+, \oplus_\lambda, \otimes_\lambda)$ be a pan-space, where g_λ is a λ-measure on Σ and $\lambda > 0$, and $f \in F_+$, $A \in \Sigma$, then*

$$(Pan)\int_A^{(\oplus_\lambda, \otimes_\lambda)} f dg_\lambda = (K)\int_A f dg_\lambda.$$

Proof. Denote

$$\begin{aligned} & s_f(\mathcal{E}\mid_A, \oplus_\lambda, \otimes_\lambda) \\ = & \bigoplus_{E\in\mathcal{E}}{}_\lambda\Big[\big(\inf_{x\in A\cap E} f(x)\big)\otimes_\lambda g_\lambda(A\cap E)\Big], \end{aligned}$$

from Proposition 3.2 (i), we have

$$\begin{aligned} & h\big[(s_f(\mathcal{E}\mid_A, \oplus_\lambda, \otimes_\lambda)\big] \\ = & h\left\{\bigoplus_{E\in\mathcal{E}}{}_\lambda\Big[\big(\inf_{x\in A\cap E} f(x)\big)\otimes_\lambda g_\lambda(A\cap E)\Big]\right\} \\ = & \sum_{E\in\mathcal{E}}\big(\inf_{x\in A\cap E} h(f(x))\big)\cdot h(g_\lambda(A\cap E)) \\ = & \sum_{E\in\mathcal{E}}\big(\inf_{x\in A\cap E}(h(f(x))\big)\cdot g_\lambda^*(A\cap E). \end{aligned}$$

Noting that g_λ^* is a classical measure on Σ and $h(f(x)) = \log_{1+\lambda}(1+\lambda f(x))$, we have

$$\begin{aligned} & \sup_{\mathcal{E}\in\hat{\mathcal{P}}}\Big\{h\big[s_f(\mathcal{E}\mid_A, \oplus_\lambda, \otimes_\lambda)\big]\Big\} \\ = & \sup_{\mathcal{E}\in\hat{\mathcal{P}}}\Big\{\sum_{E\in\mathcal{E}}\big(\inf_{x\in A\cap E}(h(f(x))\big)\cdot g_\lambda^*(A\cap E)\Big\} \\ = & (L)\int_A (h(f(x))dg_\lambda^* \\ = & (L)\int_A \log_{1+\lambda}(1+\lambda f)dg_\lambda^*. \end{aligned}$$

Therefore

$$\begin{aligned} & (Pan)\int_A^{(\oplus_\lambda,\otimes_\lambda)} f dg_\lambda \\ = & \sup\Big\{s_f(\mathcal{E}\mid_A, \oplus_\lambda, \otimes_\lambda) \mid \mathcal{E}\in\hat{\mathcal{P}}\Big\} \\ = & h^{-1}\Big(\sup\Big\{h\big[s_f(\mathcal{E}\mid_A, \oplus_\lambda, \otimes_\lambda)\big] \mid \mathcal{E}\in\hat{\mathcal{P}}\Big\}\Big) \\ = & h^{-1}\left((L)\int_A \log_{1+\lambda}(1+\lambda f)dg_\lambda^*\right) \\ = & \frac{1}{\lambda}\left[(1+\lambda)^{(L)\int_A \log_{1+\lambda}(1+\lambda f)dg_\lambda^*} - 1\right] \\ = & (K)\int_A f dg_\lambda. \end{aligned}$$

The proof of the theorem is completed.

4 CONCLUSIONS

We have shown that on λ-measure spaces ($\lambda > 0$) the Kruse integral coincides with the pan-integral based on $(\overline{R}_+, \oplus_\lambda, \otimes_\lambda)$.

Observe that for a general automorphism $g : [0,\infty] \to [0,\infty]$, Pap has introduced a g-integral [7, 12, 13], which is based on a generated ring $(R, \oplus_g, \otimes_g)$, where

$$x \oplus_g y = g^{-1}(g(x) + g(y))$$

and

$$x \otimes_g y = g^{-1}(g(x)\cdot g(y)).$$

Kruse integral is a particular case of g-integral, and all our results can be extended to g-integrals. Summarizing, it can be shown that, considering $\oplus_g - measures$, g-integral coincide with PAN-integral related to the generated ring $(R, \oplus_g, \otimes_g)$.

We also point out that Ichihashi et. al. [3] defined pseudo-addition and pseudo-multiplication on interval $[a, b]$, and then put forward a kind of fuzzy integral based on these two operators. Let $[a, b] = [0, +\infty] = \overline{R}_+$, then if we define pseudo-addition and pseudo-multiplication on $[a, b]$ in the similar way as we have done in Section 3, we shall obtain a pair of operators–pseudo-addition and pseudo-multiplication (with respect to Ichihashi et. al.'s [3]). In such case, the integral in the sense of Ichihashi et. al. [3] coincides with K-integral.

Mesiar et al. introduced *pseudo-concave integrals* [10] (see also [11]) and *pseudo-concave Benvenuti integrals* [4] by means of the pseudo-addition $\oplus$ and pseudo-multiplication $\otimes$ of reals based on a generalized ring $(\overline{R}_+, \oplus, \otimes)$ (see also [1, 2]). In next deeper study, we shall investigate the relationships among these two integrals and K-integral based on the pan-space $(X, \Sigma, g_\lambda, \overline{R}_+, \oplus_\lambda, \otimes_\lambda)$, where g_λ is a λ-measure.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grant No. 11371332).

References

[1] P. Benvenuti, R. Mesiar, D. Vivona (2002). Monotone set functions-based integrals. *Handbook of Measure Theory*, E. Pap, ed., Elsevier, Amsterdam, pp 1329-1379.

[2] P. Benvenuti, R. Mesiar (2004). Pseudo-arithmetical operations as a basis for the general measure and integration theory. *Information Sciences* **160**, pp 1-11.

[3] H. Ichihashi, M. Tanaka, K. Asai (1988). Fuzzy integrals based on pseudo-additions and multiplications. *J. Math. Anal. Appl.* **130**, pp 354–364.

[4] A. Kolesárová, J. Li, R. Mesiar (2012). Pseudo-concave Benvenuti integral. in: *Commun. Com-*

put. and Inf. Sci. **300**, Springer-Verlag Berlin Heidelberg, pp 565–570.

[5] R. Kruse (1982). A note on λ-additive fuzzy measures. *Fuzzy Sets and Systems* **8**, pp 219–222.

[6] R. Kruse (1983). Fuzzy integrals and conditional fuzzy measures. *Fuzzy Sets and Systems* **10**, pp 309–313.

[7] R. Mesiar, E. Pap (1999). Idempotent integral as limit of g-integrals. *Fuzzy Sets and Systems* **102**, pp 385-392.

[8] R. Mesiar, J. Rybárik (1995). Pan-operations structure. *Fuzzy Sets and Systems* **74**, pp 365-369.

[9] R. Mesiar (1995). Choquet-like integrals. *J. Math. Anl. Appl.* **194**, pp 477-488.

[10] R. Mesiar, J. Li, E. Pap (2011). Pseudo-concave integrals. in: *NLMUA2011*, in: *Adv. Intell. Syst. Comput.* **100**, Springer-Verlag, Berlin Heidelberg, pp. 43–49.

[11] R. Mesiar, J. Li, E. Pap (2013). Discrete pseudo–integrals. *International Journal of Approximative Reasoning* **54**, pp 357–364.

[12] E. Pap (1993). g-calculus, Univ. u Novom Sadu, Zb. rad Prirod.–Mat. Fak. Ser. Mat. **23(1)** pp 93–101.

[13] E. Pap (1995). *Null-additive Set Functions.* Kluwer, Dordrecht.

[14] E. Pap (2002). Pseudo-additive measure and their application. *Handbook of Measure Theory*, E. Pap, ed., Elsevier, Amsterdam, pp 1403-1468.

[15] N. Shilkret (1971). Maxitive measure and integration. *Indag. Math.* **33**, pp 109-116.

[16] M. Sugeno (1974). Theory of fuzzy integrals and its application. Ph. D. Thesis, Tokyo Institute of Technology, 1974.

[17] M. Sugeno, T. Murofushi (1987). Pseudo-additive measures and integrals. *J. Math. Anal. Appl.* **122**, pp 197-222.

[18] X. Tong, M. Chen, HX. Li (2004). Pan-operations structure with non-idempotent pan-addition. *Fuzzy Sets and Systems* **145**, pp 463-470.

[19] Z. Wang, G. J. Klir (2009). *Generalized Measure Theory*, Springer, 2009.

[20] Q. Yang (1985). The pan-integral on fuzzy measure space. *Fuzzy Mathematics* **3**, pp 107-114 (in Chinese).

[21] Q. Zhang, R. Mesiar, J. Li, P. Struk (2011). Generalized Lebesgue integral. *International Journal of Approximate Reasoning* **52**, pp 427-443.

SUBJECTIVE LINGUISTIC PREFERENCE RELATIONS AND THEIR APPLICATION IN GROUP DECISION MAKING

Sebastia Massanet, Juan Vicente Riera and Joan Torrens
Dept. Mathematics and Computer Science
University of the Balearic Islands
Palma de Mallorca, Spain
{s.massanet,jvicente.riera,jts224}@uib.es

Enrique Herrera-Viedma
Dept. Computer Science and Artificial Intelligence
University of Granada
Granada, Spain
viedma@decsai.ugr.es

Summary

In a medical system, experts have to take many critical decisions to choose which can be the best treatment for a concrete patient with a particular disease. Several group decision making problems have been applied with a relative success to this problem, usually based on preference relations. The aim of this paper is to introduce a new class of fuzzy preference relation based on subjective evaluations in the discrete fuzzy numbers framework and to present a group decision making problem based on these relations. Finally, an illustrative example of the model is given.

Keywords: Preference relation, discrete fuzzy number, subjective linguistic evaluation, aggregation operator.

1 INTRODUCTION

A general decision making (GDM) problem [14] may be defined as a decision problem with several alternatives and experts that try to achieve a common solution taking into account all their opinions. These opinions are often based on the use of preferences expressed by the experts, usually through the so-called preference relations. Depending on the nature or complexity of the problem, these preference relations can be expressed in different ways, many times embedded in a fuzzy environment (see for instance [9]). Interval-valued fuzzy preference relations, linguistic interval fuzzy preference relations or incomplete fuzzy linguistic preference relations are interesting examples of preference relations among many others [7, 9, 14, 21]. Thereby, for each problem, the most suitable class of fuzzy preference relations should be chosen. This fact is one of the main reasons why a great number of experts have investigated new families of fuzzy preference relations in recent years. In this sense, with the aim of better describing the experts' opinions allowing them to be more flexible, the concept of subjective evaluations is considered in [2, 3, 10, 13, 16, 17] as a kind of discrete fuzzy numbers whose support is a subinterval of the finite chain $L_n = \{0, 1, \ldots, n\}$. This class of fuzzy subsets can be interpreted as a flexibilization of linguistic expressions such as *better than Good, between Fair and Very Good* or even more complex expressions. Indeed, they have already been used successfully in decision making problems [3, 13, 16, 17].

On the other hand, the modelling of medical decision making has been among the leading research objectives for decades [11, 20]. From the early work in 1979 by E. Sanchez [19] who introduced the concept of *medical knowledge* studying relationships between symptoms and diseases by means of fuzzy relations until nowadays, many authors have proposed different approaches [1, 5, 11] (fuzzy cognitive maps, fuzzy soft sets, intuitionistic fuzzy sets...) in medical diagnoses. Thus, in this paper the authors propose a new class of fuzzy preference relations based on subjective evaluations. Some properties of them are studied and then a GDM model based on subjective linguistic preference relations is presented. Finally, a concrete application to medical decision is given.

2 PRELIMINARIES

In this section we will present the main concepts related to discrete fuzzy numbers that will be used later.

By a fuzzy subset of $\mathbb{R}$, we mean a function $A : \mathbb{R} \to [0, 1]$. For each fuzzy subset A, let $A^\alpha = \{x \in \mathbb{R} : A(x) \geq \alpha\}$ for any $\alpha \in (0, 1]$ be its α-level set (or α-cut). By *supp*(A), we mean the support of A, i.e. the set $\{x \in \mathbb{R} : A(x) > 0\}$.

Definition 1 *[23] A fuzzy subset A of $\mathbb{R}$ with membership mapping $A : \mathbb{R} \to [0, 1]$ is called a* discrete fuzzy number *if its support is finite, i.e., there exist $x_1, ..., x_n \in \mathbb{R}$ with $x_1 < x_2 < ... < x_n$ such that*

$supp(A) = \{x_1, ..., x_n\}$, and there are natural numbers s, t with $1 \leq s \leq t \leq n$ such that:

1. *$A(x_i)=1$ for all i with $s \leq i \leq t$.* (core)
2. *$A(x_i) \leq A(x_j)$ for all i, j with $1 \leq i \leq j \leq s$.*
3. *$A(x_i) \geq A(x_j)$ for all i, j with $t \leq i \leq j \leq n$.*

From now on, we will denote by $\mathcal{A}_1^{L_n}$ the set of discrete fuzzy numbers whose support is a subinterval of the finite chain $L_n = \{0, 1, \ldots, n\}$.

Let $A, B \in \mathcal{A}_1^{L_n}$ be two discrete fuzzy numbers. Note that the supports of A and B and their α-cuts are subintervals of L_n. Let $A^\alpha = [x_1^\alpha, x_p^\alpha]$, $B^\alpha = [y_1^\alpha, y_k^\alpha]$ be the α-level cuts for A and B, respectively.

The authors showed in [2] that $\mathcal{A}_1^{L_n}$ is a bounded distributive lattice while the set of discrete fuzzy numbers in general is not. Aggregation functions defined on L_n have been extended to $\mathcal{A}_1^{L_n}$ (see for instance [3, 16]) according to the next result.

Theorem 1 *[3, 16] Let consider a binary aggregation function F on the finite chain L_n. The binary operation on $\mathcal{A}_1^{L_n}$ defined as follows*

$$\begin{array}{c} \mathcal{F} : \mathcal{A}_1^{L_n} \times \mathcal{A}_1^{L_n} \longrightarrow \mathcal{A}_1^{L_n} \\ (A, B) \longmapsto \mathcal{F}(A, B) \end{array}$$

being $\mathcal{F}(A, B)$ the discrete fuzzy number whose α-cuts are the sets

$$\{z \in L_n \mid \min F(A^\alpha, B^\alpha) \leq z \leq \max F(A^\alpha, B^\alpha)\}$$

for each $\alpha \in [0, 1]$ is an aggregation function on $\mathcal{A}_1^{L_n}$. This function will be called the extension of the discrete aggregation function F to $\mathcal{A}_1^{L_n}$*. In particular, if F is a t-norm, a t-conorm, a uninorm, a nullnorm or a compensatory aggregation function, so is its extension $\mathcal{F}$.*

Remark 1 *Note that the previous result allows us, in this framework, to easily handle different classes of aggregation functions such as t-norms, t-conorms, nullnorms, uninorms or compensatory functions [12].*

Next proposition proposes a method to obtain a negation function on the bounded distributive lattice $\mathcal{A}_1^{L_n}$ from the unique strong negation $N_C(x) = n - x$ on the finite chain L_n.

Proposition 2 *[3] Let us consider the strong negation N_C on the finite chain $L_n = \{0, 1, \ldots, n\}$. The mapping*

$$\begin{array}{cccc} \mathcal{N}: & \mathcal{A}_1^L & \longrightarrow & \mathcal{A}_1^L \\ & A & \mapsto & \mathcal{N}(A) \end{array}$$

is a strong negation on $\mathcal{A}_1^{L_n}$ where $\mathcal{N}(A)$ is the discrete fuzzy number such that has as α-level cuts the sets $\mathcal{N}(A)^\alpha = [N_C(x_p^\alpha), N_C(x_1^\alpha)]$ for each $\alpha \in [0, 1]$ (being $A^\alpha = [x_1^\alpha, x_p^\alpha]$ the α-cuts of A).

3 LINGUISTIC MODEL BASED ON DISCRETE FUZZY NUMBERS

In this section we recall the fuzzy linguistic model based on discrete fuzzy numbers whose support is an interval of the finite chain $L_n = \{0, 1, \ldots, n\}$.

First of all, note that we can consider a bijective mapping between the ordinal scale $\mathfrak{L} = \{s_0, \ldots, s_n\}$ and the finite chain L_n which keeps the original order. Furthermore, each normal discrete convex fuzzy subset defined on the ordinal scale $\mathfrak{L}$ can be considered like a discrete fuzzy number belonging to $\mathcal{A}_1^{L_n}$, and vice-versa.

For example, consider the linguistic hedge

$$\mathfrak{L} = \{EB, VB, B, F, G, VG, EG\} \qquad (1)$$

where the letters refer to the linguistic terms Extremely Bad, Very Bad, Bad, Fair, Good, Very Good and Extremely Good and they are listed in an increasing order:

$$EB \prec VB \prec B \prec F \prec G \prec VG \prec EG$$

and the finite chain L_6. Thus, the discrete fuzzy number $A = \{0.6/2, 0.7/3, 1/4, 0.8/5\} \in \mathcal{A}_1^{L_6}$ can be also expressed as $A = \{0.6/B, 0.7/F, 1/G, 0.8/VG\}$ (see Figure 1). Note that this discrete fuzzy number, A, can be interpreted as a possible flexibilization of the linguistic label G (Good). Furthermore, in [10] and [15] it was shown that discrete fuzzy numbers can play also the role of a possible generalization of a Hesitant Fuzzy Linguistic Term Set (HFLTS) (see [18] and [22] for details). For instance, the discrete fuzzy number $B = \{0.5/1, 1/2, 1/3, 1/4, 0.2/5\}$ is a possible flexibilization of the HFLTS "between Bad and Good" (see Figure 2).

From the above discussion, we can introduce the definition of a subjective evaluation.

Definition 2 *Let $L_n = \{0, \ldots, n\}$ be a finite chain. We call a subjective evaluation to each discrete fuzzy number belonging to the partially ordered set $\mathcal{A}_1^{L_n}$.*

According to the previous comments, a subjective evaluation can be interpreted equivalently like a normal convex fuzzy subset defined on the ordinal scale $\mathfrak{L}$.

Moreover, our approach presents some interesting properties [10, 13]. Thus, a first aspect about the linguistic interpretation based on subjective evaluations

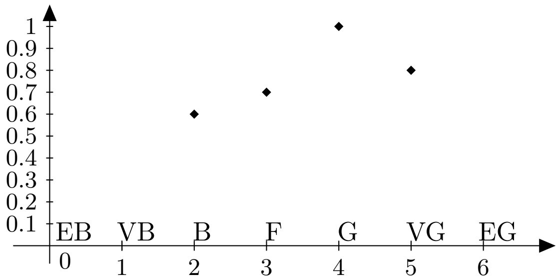


Figure 1: Graphical representation of a subjective evaluation that can interpret the linguistic label *"Good"*.

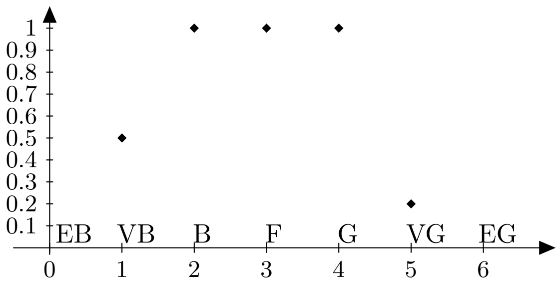


Figure 2: Graphical representation of a subjective evaluation that can interpret the linguistic expression *"Between Bad and Good"*.

is that it does not need making previous transformations when we wish to aggregate the information. Each subjective evaluation considered by an expert is directly interpreted as a discrete fuzzy number of the bounded set $\mathcal{A}_1^{L_n}$ and by this reason we can handle this information directly according to Theorem 1. Also, our approach allows more flexibilization of the linguistic term sets. In this way, it is possible to define different *flexibilizations* of a linguistic expression (see Figures 1 and 2). Finally, it is worth to mention that this linguistic model [13] permits that experts can use different representation formats to express their preferences. In this sense, experts can use linguistic scales with different granularity (see Figure 3) in order to make their corresponding assessment. The model is capable of expressing a final decision encompassing all the assessments expressed in these different linguistic scales.

4 THE GDM PROBLEM BASED SUBJECTIVE LINGUISTIC PREFERENCE RELATIONS

In a classical Group Decision Making situation [9], there is a set of possible alternatives, $X = \{x_1, \ldots, x_n\}(n \geq 2)$ and a group of experts, $E = \{e_1, \ldots, e_m\}(m \geq 2)$, characterized by their back-

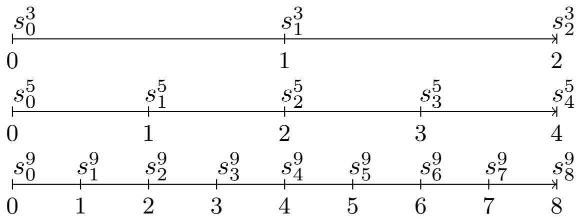


Figure 3: Linguistic hierarchy of three, five and nine linguistic terms.

ground and knowledge, who express their opinions about X to achieve a common solution. Preference relations are a common method to model experts' preferences in group decision making problems. It is well known that there exist different types of preference relations that can be used according to the considered domain to evaluate the intensity of the preference [9]. Now, in this paper we introduce a new linguistic preference relation where the degree of the linguistic preference of the alternative x_i over x_j is expressed as a subjective evaluation.

Definition 3 *A subjective linguistic preference relation P on a finite set of alternatives X is characterized by a membership function $\mu_{P_{e_i}} : X \times X \to \mathcal{A}_1^{L_n}$ where $\mu_{P_{e_i}}(x_l, x_k) = p_i^{lk} \in \mathcal{A}_1^{L_n}$ denotes the linguistic preference for the alternative x_l over the alternative x_k expressed by the expert e_i.*

When the cardinality of X is small, these preference relations may be represented by the matrix $P_{e_i} = (p_i^{lk})$. Furthermore, from now on we will assume that n is even, $m = n/2$ and that 1_a denotes the discrete fuzzy number whose support is the singleton $\{a\}$ for any $a \in L_n$. With these assumptions, subjective linguistic preferences will be interpreted as follows:

- The linguistic preference $p_i^{lk} = 1_m$ indicates indifference between x_l and x_k.
- If the subjective evaluation $p_i^{lk} \neq 1_m$ and its support is a subinterval of $[m, n]$ $((m, n])$ then x_l is preferred (strictly preferred) to x_k. In particular, the linguistic preference $p_i^{lk} = 1_n$ indicates that x_l is absolutely preferred to x_k.
- If the subjective evaluation $p_i^{lk} \neq 1_m$ and its support is a subinterval of $[0, m]$ $([0, m))$ then x_k is preferred (strictly preferred) to x_l. In particular, the linguistic preference $p_i^{lk} = 1_0$ indicates that x_k is absolutely preferred to x_l.
- Finally, if the support of the subjective evaluation p_i^{lk} contains values less than m and greater than m at the same time, then the expert has some hesitation about his preference.

In this work we will use subjective linguistic preference relations that satisfy the reciprocity property given in the next definition.

Definition 4 *A subjective linguistic preference relation P satisfies the* reciprocity property *if for all $p_i^{lk}, p_i^{kl} \in P$ the following identities hold:*

$$a_1^{lk,\alpha} + b_2^{kl,\alpha} = n$$
$$a_2^{lk,\alpha} + b_1^{kl,\alpha} = n \text{ for all } \alpha \in [0,1]$$

being $[a_1^{lk,\alpha}, a_2^{lk,\alpha}]$ and $[b_1^{lk,\alpha}, b_2^{lk,\alpha}]$ the α-cuts of the subjective evaluations p_i^{lk}, p_i^{kl} respectively.

Remark 2 *It is worth to point out that this condition is similar to the classical one considered in the interval fuzzy preference relations. Also, note that if P satisfies the reciprocity property then $p_i^{kl} = \mathcal{N}(p_i^{lk})$, where $\mathcal{N}$ is the extension in $\mathcal{A}_1^{L_n}$ of the unique strong negation N in L_n (see [3]).*

Next example shows a reciprocal linguistic preference relation given by an expert e_1 on the preference set $X = \{x_1, x_2, x_3\}$ where each subjective evaluation is expressed in $\mathfrak{L} = \{EB, B, F, G, EG\}$, where the letters refer to the same linguistic terms of the chain (1).

Example 1

$$P_{e_1} = \begin{pmatrix} - & p_1^{12} & p_1^{13} \\ p_1^{21} & - & p_1^{23} \\ p_1^{31} & p_1^{32} & - \end{pmatrix}$$

where

$$\begin{aligned}
p_1^{12} &= \{0.5/0, 1/1\} = \{0.5/EB, 1/B\}, \\
p_1^{13} &= \{0.6/3, 1/4\} = \{0.6/G, 1/EG\}, \\
p_1^{23} &= \{0.3/1, 0.6/2, 1/3\} = \{0.3/B, 0.6/F, 1/G\}, \\
p_1^{21} &= \mathcal{N}(p_1^{12}) = \{1/3, 0.5/4\} = \{1/G, 0.5/EG\}, \\
p_1^{31} &= \mathcal{N}(p_1^{13}) = \{1/0, 0.6/1\} = \{1/EB, 0.6/B\}, \\
p_1^{32} &= \mathcal{N}(p_1^{23}) = \{1/1, 0.6/2, 0.3/3\} \\
&= \{1/B, 0.6/F, 0.3/G\}.
\end{aligned}$$

Remark 3 *Note that when the support of each one of the values of the matrix $P_{e_i} = (p_i^{lk})$ are discrete fuzzy numbers whose support coincides with the core, the subjective linguistic preference relation is a linguistic interval fuzzy preference relation in the sense of [21]. That is, our model can be interpreted as a possible generalization of the linguistic interval fuzzy preference relation.*

As the experts' opinions can be expressed in different linguistic scales, a first step is to unify each one of these preferences into a single linguistic domain.

4.1 MAKING THE LINGUISTIC INFORMATION UNIFORM

In this phase, all experts' multi-granular linguistic preferences are unified into a single linguistic domain. For this reason, it is necessary to construct a transformation function among the levels of a linguistic hierarchy. These transformation functions will be based on the concept of completion of a discrete fuzzy number [13]. In this work, we will use the α-completion (see [13] for details). Let us show an example of this step.

Example 2 *Consider the subjective evaluation of Example 1. The expression of this assessment in the linguistic scale L_8, identified with the linguistic scale $\mathfrak{L} = \{D, EB, VB, B, F, G, VG, EG, P\}$, where the labels have the same meaning as in (1) and D denotes* Dreadful *and P denotes* Perfect *corresponding to the third level of the linguistic hierarchy given in Figure 3, is as follows:*

$$\tilde{P}_{e_1} = \begin{pmatrix} - & \tilde{p}_1^{12} & \tilde{p}_1^{13} \\ \tilde{p}_1^{21} & - & \tilde{p}_1^{23} \\ \tilde{p}_1^{31} & \tilde{p}_1^{32} & - \end{pmatrix} \tag{2}$$

where

$$\begin{aligned}
\tilde{p}_1^{12} &= \{0.5/0, 0.5/1, 1/2\}, \\
\tilde{p}_1^{13} &= \{0.6/6, 0.6/7, 1/8\}, \\
\tilde{p}_1^{23} &= \{0.3/2, 0.3/3, 0.6/4, 0.6/5, 1/6\}, \\
\tilde{p}_1^{lk} &= \mathcal{N}(\tilde{p}_1^{kl}) \text{ with } k = 1, 2,\ l = 2, 3 \text{ and } k < l.
\end{aligned}$$

4.2 RESOLUTION PROCESS OF THE GDM PROBLEM

Once the linguistic information is uniform, the resolution process of the GDM problem usually relies on obtaining a set of solution alternatives from the preferences given by the experts. This selection process is composed by two procedures [6, 8]: aggregation and exploitation.

- *Aggregation phase*

  In this phase, a collective subjective linguistic preference relation $P_c = (P^{ij})$ is obtained by means of the aggregation of all the individual subjective linguistic preference relations $\{P_{e_1}, \ldots, P_{e_m}\}$ at the level of pairs of alternatives. This aggregation is carried out using the method proposed in Theorem 1.

- *Exploitation phase*

  In this phase, the global and collective information about the alternatives is transformed into a global ranking among them and after that, we can choose the set of solution alternatives. To do so, firstly we compute for each preference x_i the choice function, according to [7], given by the expression

$$px_i = \mathcal{F}(P^{i1}, \ldots, P^{in}) \quad (3)$$

where P^{ii} is not considered. Finally, if we use a ranking method, we obtain a classification of the alternatives through

$$\text{if } px_i > px_j \text{ then } x_i \text{ is preferable to } x_j.$$

There exist a great number of ranking methods. For instance, we can highlight the one proposed by L. Chen and H. Lu in [4] or the centroid method.

5 APPLICATION TO MEDICAL DECISION: AN ILLUSTRATIVE EXAMPLE

Suppose that a group of three oncologists $E = \{e_1, e_2, e_3\}$ hesitate about the best treatment, $X = \{x_1$=chemotherapy, x_2 = radiation therapy, x_3=hormonal therapy$\}$, to deliver the highest quality possible patient care. For this reason, each oncologist provides his subjective linguistic preference relations.

The first expert provides his preference P_{e_1}, which has the same expression of Example 1. The other two oncologists e_2 and e_3 use the linguistic scale L_8 and their preferences are the next ones:

$$P_{e_2} = \begin{pmatrix} - & p_2^{12} & p_2^{13} \\ p_2^{21} & - & p_2^{23} \\ p_2^{31} & p_2^{32} & - \end{pmatrix}, P_{e_3} = \begin{pmatrix} - & p_3^{12} & p_3^{13} \\ p_3^{21} & - & p_3^{23} \\ p_3^{31} & p_3^{32} & - \end{pmatrix}$$

$$\begin{array}{ll} p_2^{12} & = \{1/0, 0.8/1, 0.7/2\}, \\ p_2^{13} & = \{0.5/0, 0.8/1, 1/2, 0.9/3\}, \\ p_2^{23} & = \{0.6/5, 0.8/6, 1/7, 0.7/8\}, \\ p_3^{12} & = \{0.4/0, 0.5/1, 1/2, 0.6/3\}, \\ p_3^{13} & = \{0.6/3, 0.7/4, 1/5, 0.9/6\}, \\ p_3^{23} & = \{0.7/5, 0.8/6, 0.9/7, 1/8\}, \\ p_i^{lk} & = \mathcal{N}(p_i^{kl}) \text{ with } k = 1,2,\ l,i = 2,3, \text{ and } k < l. \end{array}$$

According to Section 4.1, the first step is making uniform the valuations of each expert. In our case, if we consider L_8 as the common linguistic scale then, we must only transform the valuations of the first expert. Example 2 shows this transformation. Then, the following two steps are:

Aggregation: For instance, we consider the extension $\mathcal{F}$ in $\mathcal{A}_1^{L_n}$ of the kernel function [12], a compensatory class of aggregation functions on the finite chain L_n given by $F(x_1, \ldots, x_m) =$

$$\max\{\min\{x_1, \ldots, x_m\}, \max\{x_1, \ldots, x_m\} - k\}.$$

Thus, taking $k = 4$, the collective relation P_c is expressed as

$$P_c = \begin{pmatrix} - & P^{12} & P^{13} \\ P^{21} & - & P^{23} \\ P^{31} & P^{32} & - \end{pmatrix}$$

where $P^{ij} = \mathcal{F}(\tilde{p}_1^{ij}, p_2^{ij}, p_3^{ij})$ with $i < j$, $i,j = 1,2,3$ and

$$\begin{array}{ll} P^{12} & = \{1/0, 0.8/1, 0.7/2\}, \\ P^{13} & = \{0.6/2, 0.6/3, 1/4\}, \\ P^{23} & = \{0.3/2, 0.3/3, 0.6/4, 0.7/5, 1/6\}, \\ P^{21} & = \mathcal{N}(P^{12}) = \{0.7/6, 0.8/7, 1/8\}, \\ P^{31} & = \mathcal{N}(P^{13}) = \{1/4, 0.6/5, 0.6/6\}, \\ P^{32} & = \mathcal{N}(P^{23}) = \{1/2, 0.7/3, 0.6/4, 0.3/5, 0.3/6\}. \end{array}$$

Exploitation: We compute for each preference the values of the choice functions:

$$px_1 = \mathcal{F}(P^{12}, P^{13}) = \{1/0, 0.8/1, 0.7/2\},$$
$$px_2 = \mathcal{F}(P^{21}, P^{23}) = \{0.3/2, 0.3/3, 0.6/4, 0.7/5, 1/6\},$$
$$px_3 = \mathcal{F}(P^{31}, P^{32}) = \{1/2, 0.7/3, 0.6/4, 0.3/5, 0.3/6\}.$$

Finally, if we use the ranking method proposed by L. Chen and H. Lu in [4] we obtain for any $\beta \in [0, 1]$

$$p_{x_2} > p_{x_3} > p_{x_1}.$$

According to this ranking, based on the experts' opinions, the system recommends *the radiation therapy* as the highest quality patient care of the considered ones.

6 CONCLUSIONS AND FUTURE WORK

In this paper we have presented a new class of fuzzy preference relations based on subjective evaluations (interpreted as discrete fuzzy numbers whose support is an interval of the finite chain $L_n = \{0, \ldots, n\}$) called subjective linguistic preference relations. In addition, we have proposed a resolution process for group decision making problems based on these new fuzzy preferences relations. Finally, an example of the application of the model to a medical decision making problem is given. As a future work, we want to propose a consensus procedure for our model with this kind of new preference relations to apply to group decision making problems.

Acknowledgements

This paper has been partially supported by the Spanish Grants TIN2013-42795-P and TIN2013-40658-P.

References

[1] J. H. T. Bates and M. P. Young. Applying fuzzy logic to medical decision making in the intensive care unit. *Am. J. Respir. Crit. Care Med.*, 167:948–952, 2003.

[2] J. Casasnovas and J. Riera. Lattice properties of discrete fuzzy numbers under extended min and max. In *Proceedings of IFSA-EUSFLAT2009, Lisbon*, pages 647–652, 2009.

[3] J. Casasnovas and J. Riera. Extension of discrete t-norms and t-conorms to discrete fuzzy numbers. *Fuzzy Sets and Systems*, 167(1):65–81, 2011.

[4] L.-H. Chen and H.-W. Lu. An approximate approach for ranking fuzzy numbers based on left and right dominance. *Computers & Mathematics with Applications*, 41(12):1589 – 1602, 2001.

[5] S. Dasa and S. Karb. Group decision making in medical system: An intuitionistic fuzzy softset approach. *Applied Soft Computing*, 24:196–201, 2014.

[6] F. Herrera and E. Herrera-Viedma. Aggregation operators for linguistic weighted information. *IEEE Transactions on Systems, Man and Cybernetics*, 27(5):646 –656, 1997.

[7] F. Herrera and E. Herrera-Viedma. Choice functions and mechanisms for linguistic preference relations. *European Journal of Operational Research*, 120:144–161, 2000.

[8] F. Herrera, E. Herrera-Viedma, and J. L. Verdegay. A sequential selection process in group decision making with a linguistic assessment approach. *Information Sciences*, 85(4):223 – 239, 1995.

[9] E. Herrera-Viedma, F. J. Cabrerizo, J. Kacprzyk, and W. Pedrycz. A review of soft consensus models in a fuzzy environnment. *Information Fusion*, 17:4–13, 2014.

[10] E. Herrera-Viedma, J. V. Riera, S. Massanet, and J. Torrens. Some remarks on the fuzzy linguistic model based on discrete fuzzy numbers. In P. Angelov et al., editor, *Intelligent Systems'2014*, volume 322 of *Advances in Intelligent Systems and Computing*, pages 319–330. Springer International Publishing, 2015.

[11] D. K. Iakovidis and E. Papageorgiou. Intuitionistic fuzzy cognitive maps for medical decision making. *IEEE Transactions on Information Technology in Biomedicine*, 15(1):100–107, 2011.

[12] M. Mas, M. Monserrat, and J. Torrens. Kernel aggregation functions on finite scales. constructions from their marginals. *Fuzzy Sets and Systems*, 241:27–40, 2014.

[13] S. Massanet, J. Riera, J. Torrens, and E. Herrera-Viedma. A new linguistic computational model based on discrete fuzzy numbers for computing with words. *Information Sciences*, 258:277–290, 2014.

[14] F. Mata, L. Martínez, and E. Herrera-Viedma. An adaptative consensus support model for group decision-making problems in a multigranular fuzzy linguistic contexts. *IEEE Transactions of Fuzzy Systems*, 17(2):279–290, 2009.

[15] J. V. Riera, S. Massanet, E. Herrera-Viedma, and J. Torrens. Some interesting properties of the fuzzy linguistic model based on discrete fuzzy numbers to manage hesitant fuzzy linguistic information. 2015. Submitted to Applied Soft Computing.

[16] J. V. Riera and J. Torrens. Aggregation of subjective evaluations based on discrete fuzzy numbers. *Fuzzy Sets and Systems*, 191:21 – 40, 2012.

[17] J. V. Riera and J. Torrens. Using discrete fuzzy numbers in the aggregation of incomplete qualitative information. *Fuzzy Sets and Systems*, 264:121–137, 2015.

[18] R. Rodríguez, L. Martínez, and F. Herrera. Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 20(1):109 –119, 2012.

[19] E. Sanchez. *Medical Diagnosis and Composite Fuzzy Relations*, pages 437–444. Amsterdam: North-Holland, 1979.

[20] R. Seising and M. E. Tabacchi, editors. *Fuzziness and Medicine: Philosophical Reflections and Application Systems in Health Care.* Springer, 2013.

[21] J. Tapia, M. del Moral, M. Martínez, and E. Herrera-Viedma. A consensus model for group decision making problems with linguistic interval fuzzy preference relations. *Experts Systems with Applications*, 39:10022–10030, 2012.

[22] V. Torra. Hesitant fuzzy sets. *Int. J. Intell. Syst.*, 25(6):529–539, 2010.

[23] W. Voxman. Canonical representations of discrete fuzzy numbers. *Fuzzy Sets and Systems*, 118(3):457 – 466, 2001.

DEFINING COPULAS FROM FUZZY IMPLICATION FUNCTIONS

Sebastia Massanet, Daniel Ruiz-Aguilera and Joan Torrens
Department of Mathematics and Computer Science
University of the Balearic Islands
E-07122 Palma, Spain
{s.massanet,daniel.ruiz,jts224}@uib.es

Summary

Copulas have been deeply investigated because of their applications in many fields. From the theoretical point of view, a key point in this research lies in the search of new construction methods of parametrized families of copulas. This paper presents some of these construction methods based on fuzzy implication functions. The ideas come from recent papers by P. Grzegorzewski, where some construction methods of implication functions from copulas were introduced.

Keywords: Fuzzy implication function, copula, t-conorm, survival copula.

1 INTRODUCTION

Aggregation functions and fuzzy implication functions are two types of operations extremely related between them and in fact, there are many published works dealing with this relationship (see [3, 15] and references therein). In particular, many authors have investigated methods to construct fuzzy implication functions from aggregation functions and vice versa. A well known example of these methods comes from the family of *residual implications* or R-implications in short. Indeed, one of the earliest methods for obtaining implications was from conjunctions as their residuals, and many different classes of conjunctions have been used in this line. Mainly, left-continuous t-norms [15], left-continuous conjunctive uninorms (leading to the so-called RU-implications) [1, 7], but also semi-copulas and copulas [9] and many other types of conjunctive aggregation functions (see [17] and references therein). In the major part of these cases, it is also possible to construct the initial conjunction from the corresponding residual implication.

Another construction method comes from the *material implications*. In this case, given a negation N it is possible to construct an implication function from a t-conorm S, obtaining the well known (S, N)-implications, and vice versa [3, 4]. Again, one can use a disjunctive uninorm instead of a t-conorm (obtaining the so-called (U, N)-implications) [5], and also many other kinds of disjunctive aggregation functions (see [17] and references therein), and again the process can be done in both directions, from disjunctions to implications and vice versa.

Recently, some different ways to obtain implication functions from copulas have appeared in [8, 11, 12, 13, 14] and some properties have been studied in [2]. The common thread in these articles is the search for a type of implication functions that "*takes into consideration both imprecision modelled by fuzzy concepts and randomness described by tools originated by probability theory*" [11]. Trying to give adequate answers to this problem, the so-called *probabilistic implications*, *probabilistic S-implications*, *survival implications*, and *survival S-implications* appeared in the successive papers [11, 12, 13, 14]. However, in all these works only one direction was investigated, that is, the way of obtaining implication functions from copulas.

In this paper we want to deal with the reverse direction, that is, we want to study how to construct copulas from implication functions, just by reversing the methods given in the papers mentioned before. It is well known through the Sklar Theorem that copulas link the joint distribution function of two random variables X, Y, to their one-dimensional marginal distributions. In this sense, we know that a copula C expresses the dependence among these two random variables. This is the main interest of copulas and for this reason they have also been studied from a theoretical point of view. In particular, the research of different methods to construct copulas with appropriate properties is a constant in this theoretical investigation. Thus, this paper is also located in this line of research, since we will present some new methods to construct copulas, this time from fuzzy implication functions.

The paper is organized as follows. First some preliminaries on copulas and implication functions are recalled in Section 2. Section 3 and 4 are devoted to the construction of copulas from fuzzy implication functions with examples and properties, by reversing the methods used in probabilistic S-implications and survival probabilistic S-implications, respectively. We end the paper with a section devoted to conclusions and future work.

2 PRELIMINARIES

We will suppose the reader to be familiar with the basic results on t-norms, t-conorms and fuzzy negation functions (for details see [15]). We recall here only some concepts on copulas and fuzzy implication functions in order to make the work as self-contained as possible. To see more details on copulas see [18] and for more details on implications see [3, 17].

Definition 1 ([18]) *A function* $C : [0,1] \times [0,1] \to [0,1]$ *is a* copula *if it satisfies:*

- $C(0,x) = C(x,0) = 0$ *for all* $x \in [0,1]$.
- $C(1,x) = C(x,1) = x$ *for all* $x \in [0,1]$.
- *C is 2-increasing, that is, for all* $x' \leq x$ *and* $y' \leq y$ *it satisfies* $C(x,y') + C(x',y) \leq C(x',y') + C(x,y)$.

Definition 2 ([18]) *Given a copula C, the corresponding* survival copula *is another copula C^* that is given by*

$$C^*(x,y) = x+y-1+C(1-x,1-y), \text{ for all } x,y \in [0,1].$$

Definition 3 ([16]) *A copula C is said to be* radially symmetric *or* invariant *with respect to the construction of the survival copula (invariant for short), whenever* $C^* = C$.

Definition 4 ([3, 10]) *A binary operation* $I : [0,1]^2 \to [0,1]$ *is said to be a* fuzzy implication function*, or a* fuzzy implication*, if it satisfies:*

(I1) $I(x,z) \geq I(y,z)$ *when* $x \leq y$, *for all* $z \in [0,1]$.

(I2) $I(x,y) \leq I(x,z)$ *when* $y \leq z$, *for all* $x \in [0,1]$.

(I3) $I(0,0) = I(1,1) = 1$ *and* $I(1,0) = 0$.

Note that, from the definition, it follows that $I(0,x) = 1$ and $I(x,1) = 1$ for all $x \in [0,1]$ whereas the symmetrical values $I(x,0)$ and $I(1,x)$ are not derived from the definition. Let us recall here only two properties that will be used along the paper.

Definition 5 ([3, 10]) *Let I be a fuzzy implication.*

- *The function N_I defined by* $N_I(x) = I(x,0)$ *for all* $x \in [0,1]$*, is called the* natural negation *of I and it is always a fuzzy negation.*
- *It is said that I satisfies the* left neutrality principle *whenever* $I(1,y) = y$ *for all* $y \in [0,1]$.

The following four construction methods of fuzzy implication functions from copulas were presented in [11, 12, 13, 14].

Proposition 1 ([11, 14]) *Let C be a copula. The function* $I_C : [0,1]^2 \to [0,1]$ *given by*

$$I_C(x,y) = \begin{cases} 1 & \text{if } x = 0, \\ \frac{C(x,y)}{x} & \text{if } x > 0, \end{cases}$$

is an implication function if and only if $C(x_1,y)x_2 \geq C(x_2,y)x_1$ *for all* $x_1 \leq x_2$*. In this case, I_C is called a* probabilistic implication *(based on copula C).*

Proposition 2 ([14]) *Let C be a copula. The function* $\tilde{I}_C : [0,1]^2 \to [0,1]$ *given by*

$$\tilde{I}_C(x,y) = C(x,y) - x + 1 \quad \text{for all } x,y \in [0,1]$$

is always an implication function, which is called a probabilistic S-implication *(based on copula C).*

Proposition 3 ([13]) *Let C be a copula. The function* $I_C^* : [0,1]^2 \to [0,1]$ *given by*

$$I_C^*(x,y) = \begin{cases} 1 & \text{if } x = 0, \\ \frac{x+y-1+C(1-x,1-y)}{x} & \text{if } x > 0, \end{cases}$$

is an implication function if and only if

$$C(1-x_1,1-y)x_2 - C(1-x_2,1-y)x_1 \geq (1-y)(x_2-x_1)$$

for all $x_1 \leq x_2$*. In this case, I_C is called a* survival implication *(based on copula C).*

Proposition 4 ([13]) *Let C be a copula. The function* $\tilde{I}_C^* : [0,1]^2 \to [0,1]$ *given by*

$$\tilde{I}_C^*(x,y) = y + C(1-x,1-y) \quad \text{for all } x,y \in [0,1]$$

is always an implication function, which is called a survival S-implication *(based on copula C).*

3 PROBABILISTIC S-IMPLICATION COPULAS

In this work we want to deal with the construction of copulas from fuzzy implication functions. We want to study the possibility of reversing the methods used in the construction of probabilistic S-implications and survival S-implications. We think that the cases of probabilistic implications and survival implications are also worth to be studied, but we left them for a future work because of lack of space.

Let us begin with the case of probabilistic S-implications defined by $\tilde{I}_C(x,y) = C(x,y) - x + 1$. We can obtain the following definition by reversing this method.

Definition 6 *Let I be a fuzzy implication function. We define the* probabilistic S-implication function *(PSI-function for short) derived from I, C_I^{PSI}, as the function defined by*

$$C_I^{PSI}(x,y) = I(x,y) + x - 1 \quad \text{for all} \quad x,y \in [0,1].$$

When C_I^{PSI} is in fact a copula, it will be called a PSI-copula.

Note that we always have $\tilde{I}_{C_I^{PSI}} = I$ and $C_{\tilde{I}_C}^{PSI} = C$ for any copula C and fuzzy implication function I, by construction.

However, it is clear that the function C_I^{PSI} is not always a copula. For instance, it is obvious that the implication function I must be continuous in order to derive a copula. In fact, the following theorem gives necessary and sufficient conditions on a fuzzy implication function I, to obtain a copula via the PSI-function defined before.

Theorem 5 *Let I be a fuzzy implication function and C_I^{PSI} its derived PSI-function. Then C_I^{PSI} is a copula if and only if the following items hold:*

i) The natural negation of I is $N_c(x) = 1 - x$.

ii) I satisfies the left-neutrality principle.

iii) I is 2-increasing.

Among all classes of fuzzy implication functions we want to deal with (S,N)-implication functions given by $I(x,y) = S(N(x),y)$, whenever S is a t-conorm and N is a fuzzy negation. In this case, it is clear that condition $ii)$ in the previous theorem is always satisfied and, to satisfy condition $i)$ we must take N the classical negation N_c. Thus, the only condition to be checked is concerning the 2-increasingness.

Remark 1 *Since any copula C is always continuous, it is clear that the implication function I must be also continuous in order to C_I^{PSI} be a copula (note that continuity of I follows from the three conditions stated in Theorem 5). In the case of (S,N)-implications, since N must be N_c, the t-conorm used to construct the implication must be continuous.*

Theorem 6 *Let S be a t-conorm, I_S the (S,N)-implication given by $I_S(x,y) = S(1-x,y)$ for all $x,y \in [0,1]$, and $C_{I_S}^{PSI}$ the PSI-function derived from I_S. Then the following conditions are equivalent:*

i) $C_{I_S}^{PSI}$ is a copula.

ii) S is 2-decreasing, that is, $S(x_1,y_1) - S(x_1,y_2) - S(x_2,y_1) + S(x_2,y_2) \leq 0$ for all $x_1 \leq x_2$ and $y_1 \leq y_2$.

iii) S satisfies the Lipschitz property with constant 1, that is,

$$S(x_2,y) - S(x_1,y) \leq x_2 - x_1$$

for all $x_1, x_2, y \in [0,1]$ such that $x_1 \leq x_2$.

From the theorem above we can derive the following corollary that characterizes all (S,N)-implications such that their derived PSI-function is a copula.

Corollary 7 *Let S be a t-conorm, N a fuzzy negation, I the corresponding (S,N)-implication and C_I^{PSI} the PSI-function derived from I. Then C_I^{PSI} is a copula if and only if $N = N_c$ and S is one of the following t-conorms: the maximum, an Archimedean t-conorm with convex additive generator or an ordinal sum of Archimedean t-conorms with convex additive generators.*

Let us now show some illustrative examples. First we recall the examples given in [14] using the three basic copulas: the minimum (upper Fréchet-Hoeffding bound) copula M, the product copula Π and the Łukasiewicz (lower Fréchet-Hoeffding bound) copula W.

Example 1 i) *It was proved in [14] that the Łukasiewicz implication (which is the (S,N)-implication obtained from N_c and the Łukasiewicz t-conorm) is obtained from the copula M. Conversely, it is easy to see that C_I^{PSI} is the copula M when we start from the Łukasiewicz implication.*

ii) *Similarly, the Reichenbach implication (which is the (S,N)-implication obtained from N_c and the probabilistic sum t-conorm) can be obtained from the copula Π. Conversely, it is easy to see that C_I^{PSI} is the copula Π when we start from the Reichenbach implication.*

iii) *Finally, the Kleene-Dienes implication (which is the (S,N)-implication obtained from N_c and the maximum t-conorm) can be obtained from the copula W. Conversely, it is easy to see that C_I^{PSI} is the copula W when we start from the Kleene-Dienes implication.*

Of course, from Corollary 7 we can obtain copulas from many different (S,N)-implications taking for instance

Archimedean or ordinal sums t-conorms from some well known families, like Schweizer-Sklar, Hamacher, Frank, Yager, Dombi, Sugeno-Weber, Alsina-Sklar, or Mayor-Torrens t-conorms (see [15]). Let us detail here the copulas from three of these families.

Example 2 **i)** *It is well known that the* Hamacher-family *of t-norms, $T_\lambda^{\mathbf{H}}$, with $\lambda \in [0,2]$ are also copulas known as the family of* Ali-Mikhail-Haq *copulas (see [15]). All of them are strict with convex additive generator and so, their dual t-conorms, given by*

$$S_\lambda^{\mathbf{H}}(x,y) = \begin{cases} 1 & \text{if } \lambda = 0 \text{ and } x = y = 1, \\ \frac{x+y-xy-(1-\lambda)xy}{1-(1-\lambda)xy} & \text{otherwise,} \end{cases}$$

are also strict with convex additive generator. Consequently, it is possible to derive PSI-copulas from the corresponding (S,N)-implications with $N = N_c$. An easy computation shows that this family of PSI-copulas, denoted by $C_{\lambda,\mathbf{H}}^{PSI}$, is given by

$$C_{\lambda,\mathbf{H}}^{PSI}(x,y) = \begin{cases} 0 & \text{if } \lambda = x = 0 \text{ and } y = 1, \\ \frac{\lambda xy+(1-\lambda)x^2y}{1-(1-\lambda)(1-x)y} & \text{otherwise,} \end{cases}$$

and it will be called the family of Hamacher PSI-copulas. *Note that for $\lambda = 1$ the Hamacher t-conorm is the probabilistic sum and then the corresponding Hamacher PSI-copula is the copula Π.*

ii) *Another well known family of t-norms that are copulas is the family of* Sugeno-Weber *t-norms, $T_\lambda^{\mathbf{SW}}$ with $\lambda \in [0,+\infty]$. In this case, all of them are nilpotent (except for $\lambda = +\infty$ when the product t-norm is reached) and so their dual t-conorms $S_\lambda^{\mathbf{SW}}$ with $\lambda \in [-1,0]$, given by*

$$S_\lambda^{\mathbf{SW}}(x,y) = \min(x+y+\lambda xy, 1) \quad \text{for all } x,y \in [0,1],$$

are also nilpotent (except for $\lambda = -1$ leading to the probabilistic sum) with additive convex generator. So, we can derive again suitable (S,N)-implications with $N = N_c$ to generate a family of PSI-copulas. This family, denoted by $C_{\lambda,\mathbf{SW}}^{PSI}$, is given by

$$C_{\lambda,\mathbf{SW}}^{PSI}(x,y) = \min(x, y(1+\lambda-\lambda x)) \quad \text{for all } x,y \in [0,1],$$

and it will be called the family of Sugeno-Weber PSI-copulas. *In this case we obtain a parametrized family of copulas from the copula Π $(\lambda = -1)$ to the copula M $(\lambda = 0)$.*

iii) *Finally, a well known family of t-norms given by ordinal sums that are also copulas is the family of* Mayor-Torrens *t-norms, $T_\lambda^{\mathbf{MT}}$ with $\lambda \in [0,1]$. They are given by ordinal sums of only one Łukasiewicz summand. Their dual t-conorms, which are given by*

$$S_\lambda^{\mathbf{MT}}(x,y) = \begin{cases} \min(x+y-\lambda, 1) & \text{if } x,y \in [\lambda,1], \\ \max(x,y) & \text{otherwise,} \end{cases}$$

where $\lambda \in [0,1]$, are consequently suitable to generate PSI-copulas from the corresponding (S,N)-implications with $N = N_c$. This family of PSI-copulas, denoted by $C_{\lambda,\mathbf{MT}}^{PSI}$, is given by

$$C_{\lambda,\mathbf{MT}}^{PSI}(x,y) = \begin{cases} \min(x, y-\lambda) & \text{if } 1-x, y \in [\lambda,1], \\ \max(0, x+y-1) & \text{otherwise,} \end{cases}$$

and it will be called the family of Mayor-Torrens PSI-copulas. *In this case we obtain a parametrized family of copulas from the copula M $(\lambda = 0)$ to the copula W $(\lambda = 1)$.*

The structure of Mayor-Torrens PSI-copulas can be viewed in Figure 1. Note that these copulas are in fact W-ordinal sums as defined in [6].

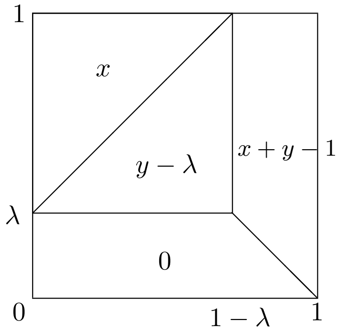


Figure 1: Structure of Mayor-Torrens PSI-copulas given in Example 2, taking $\lambda < \frac{1}{2}$.

Another well known family of implication functions is the family of residual implications derived from left-continuous t-norms, also called R-implications. However, from this family we can derive only one PSI-copula as the following result shows.

Theorem 8 *Let T be a left-continuous t-norm, I_T its R-implication and $C_{I_T}^{PSI}$ the PSI-function derived from I_T. Then C_{I_T} is a copula if and only if T is the Łukasiewicz t-norm. In this case, I_T is the Łukasiewicz implication and $C_{I_T}^{PSI}$ is the copula M.*

4 SURVIVAL S-IMPLICATION COPULAS

In this section we will deal with the case of survival S-implications defined, from a copula C, by $I_C^*(x,y) = y + C(1-x, 1-y)$ for all $x,y \in [0,1]$. It is clear that we can obtain the following definition by reversing this method.

Definition 7 *Let I be a fuzzy implication function. We define the* survival S-implication function *(SSI-function for short) derived from I, C^{SSI} as the function defined by*

$$C_I^{SSI}(x,y) = I(1-x, 1-y)+y-1 \quad \text{for all} \quad x,y \in [0,1].$$

When C_I^{SSI} is in fact a copula, it will be called a SSI-copula.

Example 3 *The three classical copulas can be obtained as SSI-copulas from the same implication functions as in the case of PSI-copulas. That is, we obtain the copula W from the Kleene-Dienes implication, the copula Π from the Reichenbach implication, and the copula M from the Łukasiewicz implication.*

As in the section above, the function C_I^{SSI} is not always a copula. In fact, the following theorem gives necessary and sufficient conditions on a fuzzy implication function I, to obtain a copula via the SSI-function defined before.

Theorem 9 *Let I be a fuzzy implication function and C_I^{SSI} its derived SSI-function. Then the following conditions are equivalent:*

i) C_I^{SSI} is a copula.

ii) The natural negation of I is N_c, I satisfies the left-neutrality principle and I is 2-increasing.

iii) C_I^{PSI} is a copula.

Moreover, the relation between C_I^{PSI} and C_I^{SSI} is extremely closed as the following theorem shows.

Theorem 10 *Let I be a fuzzy implication function such that C_I^{PSI} is a copula. Then C_I^{SSI} is also a copula and the following conditions are equivalent:*

i) $C_I^{SSI}(x,y) = C_I^{PSI}(y,x)$ for all $x, y \in [0,1]$.

ii) I satisfies contraposition with respect to N_c, that is, $I(x,y) = I(1-y, 1-x)$ for all $x, y \in [0,1]$.

From Theorem 9, the fuzzy implication functions from which we can obtain PSI-copulas and those from we can obtain SSI-copulas are exactly the same. Thus, if we deal with R-implications only the copula M is available from the Łukasiewicz implication (see Theorem 8). However, if we deal with (S,N)-implications we derive the following obvious corollary.

Corollary 11 *Let S be a t-conorm, N a fuzzy negation, I the corresponding (S,N)-implication and C_I^{SSI} the SSI-function derived from I. Then C_I^{SSI} is a copula if and only if $N = N_c$ and S is one of the following t-conorms: the maximum, an Archimedean t-conorm with convex additive generator or an ordinal sum of Archimedean t-conorms with convex additive generators.*

Now, we can derive also examples from (S,N)-implications with S a t-conorm in one of the families in Example 2, obtaining new parametrized families of copulas. Note that, since (S,N)-implications obtained from any t-conorm S and N_c always satisfy contraposition with respect to N_c, these new copulas will coincide with the *symmetric* copulas of those obtained in Example 2 (see Theorem 10).

Example 4 **i)** *If we consider the Hamacher family of t-conorms $S_\lambda^{\mathbf{H}}$ and their (S,N)-implications with $N = N_c$, the corresponding SSI-copulas are given by $C_{\lambda,\mathbf{H}}^{SSI}(x,y) = C_{\lambda,\mathbf{H}}^{PSI}(y,x)$. That is,*

$$C_{\lambda,\mathbf{H}}^{SSI}(x,y) = \begin{cases} 0 & \text{if } \lambda = y = 0 \text{ and } x = 1, \\ \frac{\lambda xy + (1-\lambda)xy^2}{1-(1-\lambda)x(1-y)} & \text{otherwise.} \end{cases}$$

ii) *When we take the Sugeno-Weber t-conorms $S_\lambda^{\mathbf{SW}}$ and their (S,N)-implications with $N = N_c$, the Sugeno-Weber SSI-copulas again satisfy $C_{\lambda,\mathbf{SW}}^{SSI}(x,y) = C_{\lambda,\mathbf{SW}}^{PSI}(y,x)$ and so they are given by*

$$C_{\lambda,\mathbf{SW}}^{SSI}(x,y) = \min(x(1+\lambda-\lambda y), y) \quad \text{for all } x, y \in [0,1].$$

iii) *Finally, if we consider the Mayor-Torrens t-conorms $S_\lambda^{\mathbf{MT}}$ and their (S,N)-implications with $N = N_c$, the Mayor-Torrens SSI-copulas satisfy $C_{\lambda,\mathbf{MT}}^{SSI}(x,y) = C_{\lambda,\mathbf{MT}}^{PSI}(y,x)$ and so they are given by*

$$C_{\lambda,\mathbf{MT}}^{SSI}(x,y) = \begin{cases} \min(x-\lambda, y) & \text{if } x, 1-y \in [\lambda, 1], \\ \max(0, x+y-1) & \text{otherwise.} \end{cases}$$

From Theorem 9, we also obtain the following interesting characterization of invariant copulas.

Theorem 12 *Let I be a fuzzy implication function such that C_I^{PSI} is a copula (and then also C_I^{SSI} is a copula). Then the following conditions are equivalent:*

i) C_I^{PSI} is an invariant copula.

ii) $C_I^{PSI} = C_I^{SSI}$.

iii) C_I^{PSI} is commutative.

Applying this last result to the case of (S,N)-implications we have the following proposition.

Proposition 13 *Let S be a t-conorm with the Lipschitz property, $N = N_c$ and I the corresponding (S,N)-implication. Let C_I^{PSI} and C_I^{SSI} the corresponding PSI-copula and the SSI-copula derived from I. Then C_I^{PSI} (or C_I^{SSI}) is an invariant copula if and only if*

$$S(1-x,y) + x = S(x, 1-y) + y \quad \text{for all } x, y \in [0,1].$$

It is worth to study the functional equation given in the previous proposition for continuous t-conorms. As a first step note that the set of solutions includes at least the three basic t-conorms: maximum, probabilistic sum and the Łukasiewicz t-conorm. This can be easily viewed directly or as a direct consequence of Examples 1 and 3.

5 CONCLUSIONS AND FUTURE WORK

Based on ideas on probabilistic S-implications and survival S-implications introduced in [13, 14] we have presented in this work two new methods of constructing copulas from fuzzy implication functions. We have dealt with (S, N) and R-implications leading to some new families of parametrized copulas. As a future work, other kinds of implications can be considered and we are also working in a similar study for probabilistic implications and survival implications.

Acknowledgements

This paper has been partially supported by the Spanish Grant TIN2013-42795-P.

References

[1] I. Aguiló, J. Suñer, and J. Torrens. A characterization of residual implications derived from left-continuous uninorms. *Information Sciences*, 180(20):3992–4005, 2010.

[2] M. Baczyński, P. Grzegorzewski, and W. Niemyska. Laws of contraposition and law of importation for probabilistic implications and probabilistic S-implications. In A. Laurent et al., editor, *Proceedings of IPMU-2014*, volume 442 of *Communications in Computer and Information Science*, pages 158–167. Springer Berlin Heidelberg, 2014.

[3] M. Baczyński and B. Jayaram. *Fuzzy Implications*, volume 231 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin Heidelberg, 2008.

[4] M. Baczyński and B. Jayaram. (S,N)- and R-implications: A state-of-the-art survey. *Fuzzy Sets and Systems*, 159:1836–1859, 2008.

[5] M. Baczyński and B. Jayaram. (U,N)-implications and their characterizations. *Fuzzy Sets and Systems*, 160:2049–2062, 2009.

[6] B. De Baets and H. De Meyer. Orthogonal grid constructions of copulas. *IEEE Transactions on Fuzzy Systems*, 15(6):1053–1062, 2007.

[7] B. De Baets and J. C. Fodor. Residual operators of uninorms. *Soft Computing*, 3:89–100, 1999.

[8] A. Dolati, J. F. Sánchez, and M. Úbeda-Flores. A copula-based family of fuzzy implication operators. *Fuzzy Sets and Systems*, 211:55 – 61, 2013.

[9] F. Durante, E. Klement, R. Mesiar, and C. Sempi. Conjunctors and their residual implicators: Characterizations and construction methods. *Mediterranean Journal of Mathematics*, 4:343–356, 2007.

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

[11] P. Grzegorzewski. Probabilistic implications. In S. Galichet et al., editor, *Proc. of the 7th conference of the European Society for Fuzzy Logic and Technology, EUSFLAT 2011, Aix-Les-Bains, France*, pages 254–258. Atlantis Press, 2011.

[12] P. Grzegorzewski. On the properties of probabilistic implications. In P. Melo-Pinto et al., editor, *Eurofuse 2011*, volume 107 of *Advances in Intelligent and Soft Computing*, pages 67–78. Springer Berlin Heidelberg, 2012.

[13] P. Grzegorzewski. Survival implications. In S. Greco et al., editor, *Advances on Computational Intelligence - 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2012. Proceedings, Part II*, volume 298 of *Communications in Computer and Information Science*, pages 335–344. Springer, 2012.

[14] P. Grzegorzewski. Probabilistic implications. *Fuzzy Sets and Systems*, 226:53–66, 2013.

[15] E. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[16] E. Klement, R. Mesiar, and E. Pap. Invariant copulas. *Kybernetika*, 38:275–286, 2002.

[17] S. Massanet and J. Torrens. An overview of construction methods of fuzzy implications. In M. Baczynński et al., editor, *Advances in Fuzzy Implication Functions*, volume 300 of *Studies in Fuzziness and Soft Computing*, pages 1–30. Springer, 2013.

[18] R. Nelsen. *An introduction to copulas*. Springer, New York, EUA, 2006.

IMPLICATIONS SATISFYING THE LAW OF IMPORTATION WITH A GIVEN UNINORM REVISITED

Sebastia Massanet, Daniel Ruiz-Aguilera and Joan Torrens
Department of Mathematics and Computer Science
University of the Balearic Islands
E-07122 Palma, Spain
{s.massanet,daniel.ruiz,jts224}@uib.es

Summary

In this paper the characterization of all fuzzy implications with continuous α-natural negation that satisfy the law of importation with a given uninorm U is resumed. The cases when the considered uninorm U is continuous in the open unit square or when U is an idempotent uninorm are studied separately and characterizations of those implications with continuous α-natural negation that satisfy the law of importation with a uninorm in these classes can be derived.

Keywords: Fuzzy implication, law of importation, uninorm, fuzzy negation.

1 INTRODUCTION

Fuzzy implication functions are used in fuzzy control and approximate reasoning to perform fuzzy conditionals [8, 12, 17] and also to perform forward and backward inferences in any fuzzy rules based system through the inference rules of modus ponens and modus tollens or through the similarity based reasoning [11, 17, 22].

Moreover, fuzzy implication functions have proved to be useful not only in fuzzy control and approximate reasoning, but also in many other fields like fuzzy relational equations [17], fuzzy DI-subsethood measures and image processing [2, 3], fuzzy morphological operators [6, 7, 13] and data mining [24], among others. In each one of these fields, there are some additional properties that the fuzzy implication functions to be used should have to ensure good results in the mentioned applications. The analysis of these additional properties of fuzzy implication functions is one of the most important topics in this field, and usually reduces to the solution of specific functional equations involving implication functions.

One of the most studied properties in this area is the so-called *Law of Importation*, which is extremely related to the exchange principle (see [18]) and it has proved to be useful in simplifying the process of applying the compositional rule of inference in many cases, see [1] and [10]. The law of importation can be written as

$$I(T(x,y),z) = I(x,I(y,z)) \quad \text{for all } x,y,z \in [0,1], \tag{LI}$$

where T is a t-norm, or a more general conjunction (for instance a conjunctive uninorm) and I is a fuzzy implication function. The law of importation has been studied by many authors in the last years [1, 10, 15, 16, 18], dealing with different aspects of this property and with different classes of implication functions. In particular, in [18] the law of importation has been used to give new characterizations of some classes of implications like (S,N)-implications and R-implications. Finally, it has been a crucial property to characterize Yager's implications in [19].

Despite of all these works devoted to the law of importation, there are still some open problems involving this property. In particular, given any t-norm T (conjunctive uninorm U), it is an open problem to find all fuzzy implications I such that they satisfy the law of importation with respect to this fixed t-norm T (conjunctive uninorm U). Recently, the authors have studied this problem, for implications with continuous natural negation, in the cases of the minimum t-norm and any continuous Archimedean t-norm (see [20]). Similarly, the characterization of implications with continuous α-natural negation for some $\alpha \in]0,1[$ satisfying the law of importation with a uninorm in $\mathcal{U}_{\min}$ and with a representable uninorm was done in [21].

As a second part of this latest work [21], in this paper we want to deal with this problem in the cases when U is a conjunctive uninorm lying in the class of uninorms continuous in the open unit square, and when U lies in the class of idempotent uninorms. We will give again partial solutions in the sense that we will find all solutions involving fuzzy implications with the additional property of having continuous α-natural negation for some $\alpha \in]0,1[$ related to the neutral element of the uninorm U.

2 PRELIMINARIES

We will suppose the reader to be familiar with the theory of t-norms and t-conorms (all necessary results and notations can be found in [14]) and uninorms (see [5] and Chapter 5 in [1]). To make this work self-contained, we recall here some of the concepts and results used in the rest of the paper.

As we have already said, with respect to uninorms we will only focus on idempotent uninorms and conjunctive uninorms continuous in the open unit square.

Definition 1 *A binary operator $U : [0,1]^2 \to [0,1]$ is said to be* idempotent *whenever $U(x,x) = x$ for all $x \in [0,1]$.*

The idempotent uninorms have been recently characterized in [23]. We recall here only the conjunctive case.

Theorem 2 ([23, Theorem 4]) *Consider $e \in (0,1)$. The following items are equivalent:*

(i) U is an idempotent uninorm with neutral element e.

(ii) There exists a decreasing function $g : [0,1] \to [0,1]$, symmetric with respect to the identity, with fixed point e such that U is given by $U(x,y) =$

$$\begin{cases} \min\{x,y\} & \text{if } y < g(x) \text{ or } (y = g(x) \\ & \text{and } x < g(g(x))), \\ \max\{x,y\} & \text{if } y > g(x) \text{ or } (y = g(x) \\ & \text{and } x > g(g(x))), \\ x \text{ or } y & \text{if } y = g(x) \text{ and } x = g(g(x)), \end{cases}$$

being commutative on the set of points $(x, g(x))$ such that $x = g(g(x))$.

In this case, such a uninorm is conjunctive if and only if $g(0) = 1$.

Any idempotent uninorm U with neutral element e and associated function g, will be denoted by $U \equiv \langle g, e\rangle_{\text{ide}}$ and the class of idempotent uninorms will be denoted by $\mathcal{U}_{\text{ide}}$.

Definition 3 ([5]) *Let e be in $]0,1[$. A binary operation $U : [0,1]^2 \to [0,1]$ is a representable uninorm if and only if there exists a strictly increasing function $h : [0,1] \to [-\infty,+\infty]$ with $h(0) = -\infty$, $h(e) = 0$ and $h(1) = +\infty$ such that*

$$U(x,y) = h^{-1}(h(x) + h(y))$$

for all $(x,y) \in [0,1]^2 \setminus \{(0,1),(1,0)\}$ and $U(0,1) = U(1,0) \in \{0,1\}$. The function h is usually called an additive generator *of U.*

A more general class containing representable uninorms are those continuous in the open unit square $]0,1[^2$, that were characterized in [9] as follows.

Theorem 4 ([9]) *Suppose U is a uninorm continuous in $]0,1[^2$ with neutral element $e \in]0,1[$. Then either one of the following cases is satisfied:*

(a) There exist $u \in [0,e[$, $\lambda \in [0,u]$, two continuous t-norms T_1 and T_2 and a representable uninorm R such that U can be represented as $U(x,y) =$

$$\begin{cases} \lambda T_1\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) & \text{if } x,y \in [0,\lambda], \\ \lambda + (u-\lambda)T_2\left(\frac{x-\lambda}{u-\lambda}, \frac{y-\lambda}{u-\lambda}\right) & \text{if } x,y \in [\lambda,u], \\ u + (1-u)R\left(\frac{x-u}{1-u}, \frac{y-u}{1-u}\right) & \text{if } x,y \in]u,1[, \\ 1 & \text{if } \min(x,y) \in]\lambda,1] \\ & \text{and } \max(x,y) = 1, \\ \min(x,y) \text{ or } 1 & \text{if } (x,y) \in \{(\lambda,1),(1,\lambda)\}, \\ \min(x,y) & \text{elsewhere.} \end{cases} \tag{1}$$

(b) There exist $v \in]e,1]$, $\omega \in [v,1]$, two continuous t-conorms S_1 and S_2 and a representable uninorm R such that U can be represented as $U(x,y) =$

$$\begin{cases} v + (\omega-v)S_1\left(\frac{x-v}{\omega-v}, \frac{y-v}{\omega-v}\right) & \text{if } x,y \in [v,\omega], \\ \omega + (1-\omega)S_2\left(\frac{x-\omega}{1-\omega}, \frac{y-\omega}{1-\omega}\right) & \text{if } x,y \in [\omega,1], \\ vR\left(\frac{x}{v}, \frac{y}{v}\right) & x,y \in]0,v[, \\ 0 & \text{if } \max(x,y) \in [0,\omega[\\ & \text{and } \min(x,y) = 0, \\ \max(x,y) \text{ or } 0 & \text{if } (x,y) \in \{(0,\omega),(\omega,0)\}, \\ \max(x,y) & \text{elsewhere.} \end{cases} \tag{2}$$

The class of all uninorms continuous in $]0,1[^2$ will be denoted by $\mathcal{U}_{\text{cos}}$. A uninorm as in (1) will be denoted by $U \equiv \langle T_1, \lambda, T_2, u, (R,e)\rangle_{\text{cos,min}}$ and the class of all uninorms continuous in the open unit square of this form will be denoted by $\mathcal{U}_{\text{cos,min}}$. Analogously, a uninorm as in (2) will be denoted by $U \equiv \langle (R,e), v, S_1, \omega, S_2\rangle_{\text{cos,max}}$ and the class of all uninorms continuous in the open unit square of this form will be denoted by $\mathcal{U}_{\text{cos,max}}$. Now, we give some definitions and results concerning fuzzy negations.

Definition 5 ([4, Definition 1.1]) *A decreasing function $N : [0,1] \to [0,1]$ is called a fuzzy negation, if $N(0) = 1$, $N(1) = 0$. A fuzzy negation N is called* *(i) strict, if it is strictly decreasing and continuous,*

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

Next lemma plays an important role in the results presented in this paper. Essentially, given a fuzzy negation, it defines a new fuzzy negation which in some sense can perform the role of the inverse of the original negation.

Lemma 6 ([1, Lemma 1.4.10]) *If N is a continuous fuzzy negation, then the function $\mathfrak{R}_N : [0,1] \to [0,1]$ defined by*

$$\mathfrak{R}_N(x) = \begin{cases} N^{(-1)}(x) & \text{if } x \in (0,1], \\ 1 & \text{if } x = 0, \end{cases}$$

where $N^{(-1)}$ stands for the pseudo-inverse of N given by $N^{(-1)}(x) = \sup\{z \in [0,1] \mid N(z) > x\}$ for all $x \in [0,1]$, is a strictly decreasing fuzzy negation. Moreover, $\mathfrak{R}_N^{(-1)} = N$, $N \circ \mathfrak{R}_N = id_{[0,1]}$ and $\mathfrak{R}_N \circ N|_{Ran(\mathfrak{R}_N)} = id|_{Ran(\mathfrak{R}_N)}$, where $Ran(\mathfrak{R}_N)$ stands for the range of function $\mathfrak{R}_N$.

Now, we recall the definition of fuzzy implications.

Definition 7 ([4, Definition 1.15]) *A binary operator $I : [0,1]^2 \to [0,1]$ is said to be a fuzzy implication if it is non-increasing in the first variable, non-decreasing in the second variable and it satisfies $I(0,0) = I(1,1) = 1$ and $I(1,0) = 0$.*

Note that, from the definition, it follows that $I(0,x) = 1$ and $I(x,1) = 1$ for all $x \in [0,1]$ whereas the symmetrical values $I(x,0)$ and $I(1,x)$ are not derived from it. Fuzzy implications can satisfy additional properties coming from tautologies in crisp logic. In this paper, we are going to deal with the law of importation, already presented in the introduction.

The natural negation with respect to α of a fuzzy implication will be also useful in our study.

Definition 8 ([1, Definition 5.2.1]) *Let I be a fuzzy implication. If $I(1,\alpha) = 0$ for some $\alpha \in [0,1)$, then the function $N_I^\alpha : [0,1] \to [0,1]$ given by $N_I^\alpha(x) = I(x,\alpha)$ for all $x \in [0,1]$, is called the* natural negation *of I with respect to α.*

Remark 9

1. *If I is a fuzzy implication, N_I^0 is always a fuzzy negation.*
2. *Given a binary function $F : [0,1]^2 \to [0,1]$, we will denote by $N_F^\alpha(x) = F(x,\alpha)$ for all $x \in [0,1]$ its α-horizontal section. In general, N_F^α is not a fuzzy negation. In fact, it is trivial to check that N_F^α is a fuzzy negation if, and only if, $F(x,\alpha)$ is a non-increasing function satisfying $F(0,\alpha) = 1$ and $F(1,\alpha) = 0$.*

3 ON THE SATISFACTION OF (LI) WITH A GIVEN UNINORM U

In this section, the main goal is the characterization of all fuzzy implications with a continuous natural negation with respect to $e \in [0,1)$ which satisfy the Law of Importation (LI) with a fixed conjunctive uninorm U. It was proved in [21] that the property:

$$\begin{array}{ll} \text{if} & N(y) = N(y') \text{ for some } y, y' \in [0,1], \\ \text{then} & N(U(x,y)) = N(U(x,y')) \text{ for all } x \in [0,1], \end{array} \quad (3)$$

plays a key role in this problem as the following proposition shows.

Proposition 10 ([21]) *Let $I : [0,1]^2 \to [0,1]$ be a binary function such that N_I^α is a fuzzy negation for some $\alpha \in [0,1)$. If I satisfies* (LI) *with a conjunctive uninorm U, then N_I^α and U satisfy Property (3).*

Moreover, the following characterization was done for the general case of a conjunctive uninorm U.

Theorem 11 ([21]) *Let $I : [0,1]^2 \to [0,1]$ be a binary function with N_I^α a continuous fuzzy negation for some $\alpha \in [0,1)$ and U a conjunctive uninorm with neutral element e such that $N_I^\alpha(e) = \alpha$. Then I satisfies* (LI) *with U if and only if N_I^α and U satisfy Property (3) and I is given by*

$$I(x,y) = N_I^\alpha(U(x, \mathfrak{R}_{N_I^\alpha}(y))).$$

Note that, from this characterization, it remains to know when N_I^α and U satisfy Property (3), for each concrete conjunctive uninorm U. We will study it for some classes of uninorms in the following section.

4 ON THE SATISFACTION OF PROPERTY (3) FOR SOME UNINORMS

First of all, we want to stress again that the goal of this paper is to characterize all fuzzy implications with a continuous natural negation with respect to some $\alpha \in [0,1)$ satisfying (LI) with a concrete conjunctive uninorm U. Therefore, there are other implications satisfying (LI) with a conjunctive uninorm U than those given in the results of this section. Of course, these implications must have non-continuous natural negations with respect to any $\alpha \in [0,1)$ such that $I(1,\alpha) = 0$. An example of an implication with this property is the least fuzzy implication.

Proposition 12 *Let I_{Lt} be the least fuzzy implication given by*

$$I_{Lt}(x,y) = \begin{cases} 1 & \text{if } x = 0 \text{ or } y = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then I_{Lt} satisfies (LI) *with any conjunctive uninorm U.*

Consequently, although I_{Lt} satisfies (LI) with any conjunctive uninorm U, we will not obtain this implication in the next results since it has no continuous natural negation at any level $\alpha \in [0, 1)$.

4.1 IDEMPOTENT UNINORMS

Let us deal in this section with conjunctive idempotent uninorms $U \equiv \langle g, e\rangle_{\text{ide}}$ ([23]). First, we will set a negation and determine some necessary conditions on the uninorm.

Proposition 13 *Let N be a continuous fuzzy negation that is constant, $N(x) = k$, in an interval $[a, b]$, where $a = \min\{x \in [0, 1] \mid N(x) = k\}$ and $b = \max\{x \in [0, 1] \mid N(x) = k\}$ and let $U \equiv \langle g, e\rangle_{\text{ide}}$ be a conjunctive uninorm in $\mathcal{U}_{\text{ide}}$ such that U and N satisfy Equation (3). Then one of the following items holds:*

i) $\{g(x) \mid x \in [0, 1]\} \cap]a, b[= \emptyset$ *whenever* $a \geq e$ *or* $b \leq e$.

ii) $\{g(x) \mid x \in [0, 1] \setminus [a, b]\} \cap]a, b[= \emptyset$ *whenever* $a < e < b$.

Moreover, if $g(x) = a$ for some $x \in [0, 1]$ then it must be $U(x, a) = \max(x, a)$, whereas if $g(x) = b$ for some $x \in [0, 1]$ then it must be $U(x, b) = \min(x, b)$.

Now, let us set the uninorm and determine necessary conditions on the negation. In the case when g has a discontinuity point s, we will denote the corresponding lateral limits by s^- and s^+.

Proposition 14 *Let N be a continuous fuzzy negation and $U \equiv \langle g, e\rangle_{\text{ide}}$ be a conjunctive uninorm in $\mathcal{U}_{\text{ide}}$ such that U and N satisfy Equation (3).*

- *If g is continuous and strictly decreasing on an interval $]a, b[$ then N is also strictly decreasing on that interval.*
- *If g is constant on an interval $]a, b[$, or s is a discontinuity point of g with $s^+ = a$ and $s^- = b$, then N can be any continuous decreasing function on the interval $[a, b]$ with the only two restrictions:*

  - If N is constant in a subinterval $[a, c]$ with $c \leq b$ and $g(x) = a$ for some x, then $U(x, a) = \max(x, a)$.

  - If N is constant in a subinterval $[c, b]$ with $c \geq a$ and $g(x) = b$ for some x, then $U(x, b) = \min(x, b)$.

Now, if we fix a conjunctive idempotent uninorm $U \equiv \langle g, e\rangle_{\text{ide}}$, those continuous negations N that satisfy equation (3) with U can be easily characterized using the previous propositions. Finally, using Theorem 11, we can derive all fuzzy implications satisfying (LI) with U.

4.2 UNINORMS CONTINUOUS IN THE OPEN UNIT SQUARE

In this section we deal with uninorms continuous in the open unit square and we divide our study into two cases, one for uninorms in $\mathcal{U}_{cos,\min}$ and the other for uninorms in $\mathcal{U}_{cos,\max}$.

4.2.1 UNINORMS IN $\mathcal{U}_{cos,\min}$

In this case all uninorms are conjunctive and we have the following results which follow a similar pattern to the previous subsection.

Proposition 15 *Let N be a continuous fuzzy negation and $U \equiv \langle T_1, \lambda, T_2, u, (R, e)\rangle_{\text{cos,min}}$ be a conjunctive uninorm in $\mathcal{U}_{cos,\min}$. Suppose that $N(x) = 1$ in an interval $[0, a]$ where $0 < a = \max\{x \in [0, 1] \mid N(x) = 1\}$. If U and N satisfy Equation (3) then $\lambda \geq a$. Moreover, if $\lambda = a$, then $U(1, \lambda) = \lambda$.*

Proposition 16 *Let N be a continuous fuzzy negation and $U \equiv \langle T_1, \lambda, T_2, u, (R, e)\rangle_{\text{cos,min}}$ be a conjunctive uninorm in $\mathcal{U}_{cos,\min}$. Suppose that N is constant, $N(x) = k$ with $0 < k < 1$, in an interval $[a, b]$, where $a = \min\{x \in [0, 1] \mid N(x) = k\}$ and $b = \max\{x \in [0, 1] \mid N(x) = k\}$. If U and N satisfy Equation (3) then one of the following cases hold:*

1. *$\lambda \geq b$ and $T_1 = T_M$. Moreover, if $\lambda = b$, then $U(1, \lambda) = \lambda$.*
2. *$\lambda \leq a < b \leq u$ and $T_2 = T_M$. Moreover, if $\lambda = a$, then $U(1, \lambda) = 1$.*

Proposition 17 *Let N be a continuous fuzzy negation and $U \equiv \langle T_1, \lambda, T_2, u, (R, e)\rangle_{\text{cos,min}}$ be a conjunctive uninorm in $\mathcal{U}_{cos,\min}$. Suppose that $N(x) = 0$ in a interval $[a, 1]$ where $1 > a = \min\{x \in [0, 1] \mid N(x) = 0\}$. Then there is no such uninorm U satisfying Equation (3) with N.*

Proposition 18 *Let N be a continuous fuzzy negation with constant regions $\{([a_i, b_i], k_i)\}_{i \in I}$ where $k_i > 0$, $a_i = \min\{x \in [0, 1] \mid N(x) = k_i\}$ and $b_i = \max\{x \in [0, 1] \mid N(x) = k_i\}$ and let $U \equiv \langle T_1, \lambda, T_2, u, (R, e)\rangle_{\text{cos,min}}$ be a conjunctive uninorm in $\mathcal{U}_{cos,\min}$ such that U and N satisfy Equation (3). Then for each $i \in I$, either $[a_i, b_i] \subseteq [0, \lambda]$ or $[a_i, b_i] \subseteq [\lambda, u]$ and the following restrictions hold:*

- *If* $U(1,\lambda) = 1$*, then for all* $a_i < \lambda$*, it must be* $b_i < \lambda$.
- *If* $U(1,\lambda) = \lambda$*, then for all* $b_i > \lambda$*, it must be* $a_i > \lambda$.
- *If* T_1 *is an Archimedean t-norm, there is only one possible constant region with* $b_i \leq \lambda$ *and in this case,* $a_i = 0$ *and* $k_i = 1$.
- *If* T_2 *is an Archimedean t-norm, then* $b_i \leq \lambda$ *for all* $i \in I$.

Now, if we fix a conjunctive uninorm $U \equiv \langle T_1, \lambda, T_2, u, (R, e)\rangle_{\cos,\min}$ in $\mathcal{U}_{cos,\min}$ with Archimedean or minimum t-norms, those continuous negations N that satisfy Equation (3) with U can be easily characterized using the previous propositions. Finally, again using Theorem 11, we can derive all fuzzy implications satisfying (LI) with U.

4.2.2 UNINORMS IN $\mathcal{U}_{cos,\max}$

There is a family of uninorms that are conjunctive although they lie in $\mathcal{U}_{cos,\max}$. They are those uninorms of the form (2) with parameter $\omega = 1$ and $U(\omega, 0) = U(1, 0) = 0$. A uninorm in this class will be denoted by $U \equiv \langle (R, e), v, S\rangle_{\cos,\max}$

Proposition 19 *Let* N *be a continuous fuzzy negation and* $U \equiv \langle (R, e), v, S\rangle_{\cos,\max}$ *be a conjunctive uninorm in* $\mathcal{U}_{cos,\max}$*. Suppose that* $N(x) = 1$ *in a interval* $[0, a]$ *where* $0 < a = \max\{x \in [0, 1] \mid N(x) = 1\}$*. Then necessarily* $v \leq a$*. Moreover, if* S *is Archimedean then it must be* $v = a$.

Proposition 20 *Let* N *be a continuous fuzzy negation and* $U \equiv \langle (R, e), v, S\rangle_{\cos,\max}$ *be a conjunctive uninorm in* $\mathcal{U}_{cos,\max}$*. Suppose that* N *is constant,* $N(x) = k$ *with* $0 < k < 1$*, in an interval* $[a, b]$*, where* $a = \min\{x \in [0, 1] \mid N(x) = k\}$ *and* $b = \max\{x \in [0, 1] \mid N(x) = k\}$*. Then necessarily* $v \leq a$*. Moreover, if* S *is Archimedean this case can not be done.*

Proposition 21 *Let* N *be a continuous fuzzy negation and* $U \equiv \langle (R, e), v, S\rangle_{\cos,\max}$ *be a conjunctive uninorm in* $\mathcal{U}_{cos,\max}$*. Suppose that* $N(x) = 0$ *in a interval* $[a, 1]$ *where* $1 > a = \min\{x \in [0, 1] \mid N(x) = 0\}$*. Then necessarily* $v \leq a$.

With all the previous results it can be easily characterized when N and U satisfy Equation (3) in the case of conjunctive uninorms in $\mathcal{U}_{cos,\max}$ with S an Archimedean or maximum t-conorm as follows.

Proposition 22 *Let* N *be a continuous fuzzy negation and* $U \equiv \langle (R, e), v, S\rangle_{\cos,\max}$ *be a conjunctive uninorm in* $\mathcal{U}_{cos,\max}$.

- *If* $S = Max$ *then* U *and* N *satisfy Equation (3) if and only if* N *is constant or strictly decreasing on the interval* $[0, v]$.
- *If* S *is Archimedean, let us denote* $a = \min\{x \in [0, 1] \mid N(x) = 0\}$*. Then* U *and* N *satisfy Equation (3) if and only if one of the following cases holds:*

  - $N(x) = 1$ *for all* $x \in [0, v]$ *and* N *is strictly decreasing on the interval* $[v, a]$.

  - N *is strictly decreasing on the interval* $[0, a]$.

Finally, again using Theorem 11, we can derive all fuzzy implications with a continuous α-natural negation satisfying (LI) with U.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we have resumed the study of the characterization of all fuzzy implication functions satisfying (LI) with a conjunctive uninorm U when the natural negation of the implication with respect to some $\alpha \in [0, 1)$ is continuous. In particular, we have analysed the problem when the conjunctive uninorm U is continuous in the open unit square or when U is an idempotent uninorm. From the results obtained in this paper, the expression of the fuzzy implication functions satisfying (LI) with a conjunctive uninorm U of one of these two classes can be derived.

As a future work, we want to study the more general case when U is a uninorm with continuous underlying t-norm and t-conorm. In addition, we want to establish the relation between the new class of implications introduced in this paper in Theorem 11 and (U, N)-implications.

Acknowledgements

This paper has been partially supported by the Spanish Grant TIN2013-42795-P.

References

[1] M. Baczyński and B. Jayaram. *Fuzzy Implications*, volume 231 of *Studies in Fuzziness and Soft Computing*. Springer, Berlin Heidelberg, 2008.

[2] H. Bustince, V. Mohedano, E. Barrenechea, and M. Pagola. Definition and construction of fuzzy DI-subsethood measures. *Information Sciences*, 176:3190–3231, 2006.

[3] H. Bustince, M. Pagola, and E. Barrenechea. Construction of fuzzy indices from fuzzy DI-subsethood measures: application to the global

comparison of images. *Information Sciences*, 177:906–929, 2007.

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

[5] J. C. Fodor, R. R. Yager, and A. Rybalov. Structure of uninorms. *Int. J. of Uncertainty, Fuzziness and Knowledge-based Systems*, 5:411–127, 1997.

[6] M. González-Hidalgo and S. Massanet. A fuzzy mathematical morphology based on discrete t-norms: fundamentals and applications to image processing. *Soft Computing*, 18(11):2297–2311, 2014.

[7] M. González-Hidalgo, S. Massanet, A. Mir, and D. Ruiz-Aguilera. On the choice of the pair conjunction-implication into the fuzzy morphological edge detector. *IEEE Transactions on Fuzzy Systems*, 2014. doi:10.1109/TFUZZ.2014.2333060.

[8] S. Gottwald. *A Treatise on Many-Valued Logic*. Research Studies Press, Baldock, 2001.

[9] S. Hu and Z. Li. The structure of continuous uninorms. *Fuzzy Sets and Systems*, 124:43–52, 2001.

[10] B. Jayaram. On the law of importation $(x \wedge y) \to z \equiv (x \to (y \to z))$ in fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 16:130–144, 2008.

[11] B. Jayaram. Rule reduction for efficient inferencing in similarity based reasoning. *International Journal of Approximate Reasoning*, 48(1):156 – 173, 2008.

[12] E. Kerre, C. Huang, and D. Ruan. *Fuzzy Set Theory and Approximate Reasoning*. Wu Han University Press, Wu Chang, 2004.

[13] E. Kerre and M. Nachtegael. *Fuzzy techniques in image processing*, volume 52 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag, New York, 2000.

[14] E. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[15] M. Mas, M. Monserrat, and J. Torrens. The law of importation for discrete implications. *Information Sciences*, 179:4208–4218, 2009.

[16] M. Mas, M. Monserrat, and J. Torrens. A characterization of (U,N), RU, QL and D-implications derived from uninorms satisfying the law of importation. *Fuzzy Sets and Systems*, 161:1369–1387, 2010.

[17] M. Mas, M. Monserrat, J. Torrens, and E. Trillas. A survey on fuzzy implication functions. *IEEE Transactions on Fuzzy Systems*, 15(6):1107–1121, 2007.

[18] S. Massanet and J. Torrens. The law of importation versus the exchange principle on fuzzy implications. *Fuzzy Sets and Systems*, 168(1):47 – 69, 2011.

[19] S. Massanet and J. Torrens. On the characterization of Yager's implications. *Information Sciences*, 201:1–18, 2012.

[20] S. Massanet and J. Torrens. Implications satisfying the law of importation with a given t-norm. In H. Bustince, J. Fernandez, R. Mesiar, and T. Calvo, editors, *Aggregation Functions in Theory and in Practise*, volume 228 of *Advances in Intelligent Systems and Computing*, pages 417–428. Springer Berlin Heidelberg, 2013.

[21] S. Massanet and J. Torrens. Implications satisfying the law of importation with a given uninorm. In A. Laurent, O. Strauss, B. Bouchon-Meunier, and R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, volume 442 of *Communications in Computer and Information Science*, pages 148–157. Springer International Publishing, 2014.

[22] D. Ruan and E. E. Kerre, editors. *Fuzzy IF-THEN Rules in Computational Intelligence: Theory and Applications*. Kluwer Academic Publishers, Norwell, MA, USA, 2000.

[23] D. Ruiz-Aguilera, J. Torrens, B. De Baets, and J. Fodor. Some remarks on the characterization of idempotent uninorms. In E. Hllermeier, R. Kruse, and F. Hoffmann, editors, *Computational Intelligence for Knowledge-Based Systems Design*, volume 6178 of *Lecture Notes in Computer Science*, pages 425–434. Springer Berlin / Heidelberg, 2010.

[24] P. Yan and G. Chen. Discovering a cover set of ARsi with hierarchy from quantitative databases. *Information Sciences*, 173:319–336, 2005.

THE ORDERED WEIGHTED AVERAGE SUM

José M. Merigó
Department of Management Control
and Information Systems,
University of Chile, Santiago, Chile
jmerigo@fen.uchile.cl

Ronald R. Yager
Machine Intelligence Institute,
Iona College, New Rochelle,
New York, USA
yager@panix.com

Summary

The ordered weighted average sum is an aggregation operator that provides a parameterized family of sums between the minimum sum and the maximum one. It is a particular case of the OWA norm when the function used is a sum. Some key properties and particular cases are studied including the average sum, the olympic sum and the centered sum. A simple numerical example is also presented. Further extensions are presented including the generalized OWAS operator and the Choquet integral sum.

Keywords: OWA operators, Norm aggregations, Sums.

1 INTRODUCTION

The ordered weighted average (OWA) [20] is an aggregation operator that analyzes the information providing a parameterized family of aggregation operators between the minimum and the maximum. Since its introduction, many authors have developed many extensions and applications [1,4,25-26].

The OWA operator can be extended and generalized in different ways. Recently, Yager, suggested the use of norms in the OWA operator [23], obtaining a more general representation that included a wide range of particular cases including distance measures [7,12,18]. This work was extended by Merigó and Yager [14] by using induced aggregation operators. They also presented a generalization that considered a unified framework between the OWA operator and the weighted average. Some other studies have also considered the use of norms in other frameworks including situations with heavy aggregation operators [10]. The work of Merigó and Yager [14] found a wide range of particular cases of great interest including the OWA sum, the OWA subtraction and the OWA multiplication. However, this was only a result in the analysis of the OWA operators with norms.

The aim of this paper is to analyze the use of sums in the OWA operator. For doing so, the aggregation with sums is seen from the perspective of the OWA norm (OWAN). However, it is also possible to build it following the methodology used with OWA distances and other related operators [9,13]. The article presents the OWA sum (OWAS). It is an aggregation operator that provides a parameterized family of aggregation operators from the minimum sum to the maximum one. It is very useful to aggregate a set of sums. Some key properties are studied. Several particular types of OWAS operators are presented including the average sum, the minimum sum, the maximum sum, the olympic sum and the centered sum. The main advantage is the flexibility to adapt to the specific needs of the problem considered. Moreover, it does not loose information in the analysis because it considers any scenario from the minimum to the maximum one.

The applicability of the OWAS operator is studied. The OWAS is useful in aggregating a set of sums in order to provide a final result. This is very common in real life because the aggregation of data many times depends on sums. Some key examples are mentioned including the sum of costs, sales and incomes. A simple numerical example is also presented in order to understand the usefulness of the new approach. Further generalizations are also developed by using generalized means. The result is the generalized OWAS (GOWAS) operator. The main advantage of this operator is that it is much more general than the OWAS being able to consider geometric and quadratic aggregations. Another generalization when using Choquet integrals is also presented obtaining the Choquet integral sum (CIS).

The paper is organized as follows. Section 2 briefly

ISBN: 978–83–8012–519–3

reviews some basic preliminaries. Section 3 introduces the OWAS operator. Section 4 studies the GOWAS operator and Section 5 the CIS operator. Section 6 summarizes the main findings and conclusions of the paper.

2 PRELIMINARIES

This Section briefly presents the OWA operator and the OWA norm.

2.1 OWA OPERATOR

The OWA operator [20] is a well-known aggregation operator that aggregates the information providing a parameterized system of aggregation operators that fluctuate between the minimum and the maximum. The OWA operator is defined as follows.

Definition 1. An OWA operator of dimension n is a mapping OWA: $R^n \rightarrow R$ that has an associated weighting W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$OWA(a_1, ..., a_n) = \sum_{j=1}^{n} w_j b_j, \tag{1}$$

where b_j is the jth largest of the a_i.

The OWA operator is commutative, monotonic, idempotent and bounded by the minimum and the maximum. In order to characterize the weighting vector of the OWA operator, several measures have been suggested including the orness measure [20] and the entropy of dispersion [20] which follows a similar methodology than the Shannon entropy [15].

2.2 NORMS WITH OWA OPERATORS

A norm is a representation that allows the use of functions and other related techniques in order to deal with a set of data. A norm is a function $R^n \rightarrow [0, \infty)$ that has the following properties [23]:

1) $f(x_1, x_2, ..., x_n) = 0$ if and only if all $x_i = 0$ and $x_i > 0$.

2) $f(aX) = |a| f(X)$.

Moreover, note that an important property to consider is the triangle inequality: $f(X) + f(Y) f(X+Y)$. However, it is not a requirement for a norm to accomplish this property since some norms may violate this axiom.

Norms can be used in the aggregation of the information by using averaging operators. When using the weighted average in the aggregation of norms, the formulation is as follows:

$$f(a_1, \ldots, a_n) = G(|a_1|, \ldots, |a_n|) = \sum_{i=1}^{n} |a_i|, \tag{2}$$

where w_i is the ith weight of the weighted average and ai is the norm used in the aggregation function G for two sets of elements $X = x_1, ..., x_n$ and $Y = y_1, ..., y_n$. Note that if$w_i = 1/n$ for all i, the weighted average norm becomes the simple average norm.

Recently, Yager [23] has suggested the use of OWA operators in norm aggregations. This aggregation can be expressed in the following way:

$$f(a_1, \ldots, a_n) = G(|a_1|, \ldots, |a_n|) = \sum_{j=1}^{n} w_j N_j, \tag{3}$$

where N_j is the jth largest of the $|a_i|$ arguments of the aggregation function G.

Further extensions can be developed by using induced aggregations [14], heavy aggregations [10] and a unified framework between the weighted average and the OWA operator [14].

3 THE OWA SUM

The ordered weighted average sum (OWAS) is an aggregation operator that aggregates a set of sums from the minimum sum to the maximum one. It is very useful for dealing with complex scenarios taking into account the attitude of the decision maker in the analysis and considering sums in the aggregation process. This is very common in business and economics when dealing with the sum of the costs, sales, benefits and assets. The OWAS operator can be defined as follows for two sets $X = x_1, , x_n$ and $Y = y_1, , y_n$.

Definition 2. An OWAS operator of dimension n is a mapping $OWAS$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$OWAS([x_1 + y_1], \ldots, [x_n + y_n]) = \sum_{j=1}^{n} w_j b_j, \tag{4}$$

where b_j is the jth largest of the $[x_i + y_i]$.

The OWAS operator is commutative, monotonic, idempotent and bounded. It is commutative because $f([x_1 + y_1], ..., [x_n + y_n]) = f([c_1 + d_1], ..., [c_n + d_n])$ where $([x_1 + y_1], ..., [x_n + y_n])$ is any permutation of the arguments $([c_1 + d_1], ..., [c_n + d_n])$. It also accomplishes the commutativity because $f([x_1+y_1], ..., [x_n+y_n]) = f([y_1 + x_1], ..., [y_n + x_n])$. It is monotonic because if $[x_i + y_i] \geq [c_i + d_i]$, for all i, then, $f([x_1+y_1], ..., [x_n+y_n]) \geq f([c_1+d_1], ..., [c_n+d_n])$. It is bounded by the minimum and the maximum sum. Therefore, $\min[x_i + y_i] \leq f([x_1 + y_1], ..., [x_n + y_n]) \leq \max[x_i + y_i]$. It is idempotent because if $[x_i + y_i] = a$, for all i, then $f([x_1 + y_1], ..., [x_n + y_n]) = a$. Note that the neutral (or identity) element is 0 because $f([x_1 + 0], ..., [x_n + 0]) = f(x_1, ..., x_n)$.

The OWAS operator is ordered in a descending way although it is possible to consider an ascending version by using $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the DOWAS operator and w^*_{n-j+1} the jth weight of the AOWAS operator.

Observe that if the weighting vector does not sum up to one, i.e., $W = \sum_{j=1}^{n} w_j \neq 1$, then, the OWAS is formulated as follows:

$$OWAS([x_1 + y_1], ..., [x_n + y_n]) = \frac{1}{W} \sum_{j=1}^{n} w_j b_j, \quad (5)$$

The weights used in the OWAS aggregation represent the attitudinal character of the decision maker rather than weighting some specific variables. In order to characterize the weights of the OWAS operator, let us present the measures commonly used in the OWA operator [20]. The degree of orness (degree of optimism) is used in the OWAS as follows:

$$\alpha(W) = \sum_{j=1}^{n} w_j \left(\frac{n-j}{n-1} \right), \quad (6)$$

The entropy of dispersion of the OWAS weighting vector is defined as:

$$H(W) = -\sum_{j=1}^{n} w_j \ln(w_j), \quad (7)$$

Finally, the balance operator of the OWAS operator is:

$$BAL(W) = \sum_{j=1}^{n} \left(\frac{n+1-2j}{n-1} \right) w_j, \quad (8)$$

In order to understand numerically the OWAS operator, let us look into a numerical example. Assume the sales of a company are classified in two regions The enterprise wants to forecast the sales for the next period considering that five scenarios or states of nature may occur. The expected sales in the two regions depending on the state of nature that occur are $X = (30, 80, 90, 80, 20)$ and $Y = (70, 30, 40, 60, 40)$. The decision maker does not know the probability that the states of nature occur. Therefore, he uses a weighting vector that represents his attitudinal character: $W = (0.3, 0.3, 0.2, 0.1, 0.1)$. He aggregates the information using the OWAS operator as follows.

$$OWAS = 0.3 \times (80 + 60) + 0.3 \times (90 + 40) + 0.2 \times (80 + 30) + 0.1 \times (30 + 70) + 0.1 \times (20 + 40) = 119.$$

Note that the OWAS develops a descending order of the individual sums. Thus, $140 \geq 130 \geq 110 \geq 100 \geq 60$.

Note that the aggregation of sums is very common in real life problems. For example, in business and economics there are many operations that require the sum of variables including the sum of sales, costs, assets, products and benefits. Therefore, this approach may have a huge potential in real world applications.

Next, let us look into some of the main families of OWAS operators [11,21]. The maximum sum is found if $w_1 = 1$ and $w_j = 0$ for all $j \neq 1$. The minimum sum is obtained if $w_n = 1$ and$w_j = 0$ for all $j \neq n$. From a general point of view, if $w_k = 1$ and $w_j = 0$ for all $j \neq k$, we get the step-OWAS operator. If $wj = 1/n$ for all i, the OWAS operator becomes the simple average sum (AS) which is formulated as:

$$AS([x_1 + y_1], ..., [x_n + y_n]) = \frac{1}{n} \sum_{i=1}^{n} (x_i + y_i), \quad (9)$$

Although it is not strictly a particular case of the OWAS operator, it is also interesting to mention the weighted average sum which is defined as follows:

$$WAS([x_1 + y_1], ..., [x_n + y_n]) = \sum_{i=1}^{n} w_i (x_i + y_i), \quad (10)$$

The olympic-OWAS operator is obtained if $w_1 = w_n = 0$, and for all others, $w_j* = 1/(n-2)$. More generally [11], $w_j = 0 for j = 1, 2, ..., k, n, n-1, ..., n-k+1$; and for all others, $w_j* = 1/(n-2k)$, where $k < n/2$.

Another type of family that could be considered is the S-OWAS operator. The generalized S-OWAS operator is obtained when $w_1 = (1/n)(1-(\alpha+\beta))+\alpha$, $w_n = (1/n)(1-(\alpha+\beta))+\beta$, and $w_j = (1/n)(1-(\alpha+\beta))$ for $j = 2$ to $n-1$, where $\alpha, \beta \in [0, 1]$ and $\alpha + \beta \leq 1$. If $\alpha = 0$, the generalized S-OWAS operator becomes the andlike S-OWAS operator, and if $\beta = 0$, the orlike S-OWAS operator.

Finally, the centered-OWAS operator [11] is obtained if the aggregation is symmetric, strongly decaying and inclusive. It is symmetric if $w_j = w_{j+n-1}$. It is strongly decaying if $i < j \leq (n+1)/2$, then $w_i < w_j$, and if $i > j \geq (n+1)/2$, 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 $wi \leq wj$ instead of $wi < wj$ (softly decaying centered-OWAS operator).

Note that many other particular types of OWAS operators could be studied following the OWA literature [8,21].

Finally, note that in Eq. (4), the OWAS operator has been defined for two sets X and Y. However, it is possible to use as many sets as needed in the aggregation by using the following definition for m sets.

Definition 3. An OWAS operator of dimension n is a mapping $OWAS$: $R^n \times R^n \times ... \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0,1]$, such that:

$$f([x_1^1 + ... + x_1^m], ..., [x_n^1 + ... + x_n^m]) = \sum_{j=1}^{n} w_j b_j, \quad (11)$$

where b_j is the jth largest of the $[x_i^1 + ... + x_i^m]$.

4 GENERALIZED OWAS OPERATOR

The OWAS operator can be generalized by using generalized means [22]. The result is the generalized OWAS (GOWAS) operator. Its main advantage is that it provides a more general framework that includes geometric, quadratic and harmonic aggregations. It can be defined as follows for two sets $X = x_1, , x_n$ and $Y = y_1, , y_n$.

Definition 4. A GOWAS operator of dimension n is a mapping $GOWAS$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0,1]$, such that:

$$GOWAS([x_1 + y_1], ..., [x_n + y_n]) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \quad (12)$$

where b_j is the jth largest of the $[x_i + y_i]$ and λ is a parameter such that $\lambda \in (-\infty, \infty) - \{0\}$.

Following Eq. (11), the GOWAS operator could be expressed as:

$$f([x_1^1 + ... + x_1^m], ..., [x_n^1 + ... + x_n^m]) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda}. \quad (13)$$

By using a different value in the parameter the GOWAS operator can form a wide range of particular cases. For example:

• If $\lambda = 1$, the GOWAS becomes the usual OWAS operator.

• If $\lambda = 2$, we get the quadratic OWAS (OWQAS) operator.

$$OWQAS([x_1 + y_1], ..., [x_n + y_n]) = \sqrt{\left(\sum_{j=1}^{n} w_j b_j^2 \right)}. \quad (14)$$

• If $\lambda \longrightarrow 0$, we get the geometric OWAS (OWGAS) operator.

$$OWGAS([x_1 + y_1], ..., [x_n + y_n]) = \prod_{j=1}^{n} b_j^{w_j}. \quad (15)$$

• If $\lambda = -1$, we get the harmonic OWAS (OWHAS) operator.

$$OWHAS([x_1 + y_1], ..., [x_n + y_n]) = \frac{1}{\sum_{j=1}^{n} \frac{w_j}{b_j}}. \quad (16)$$

• If $\lambda = 3$, we get the cubic OWAS (OWCAS) operator.

$$OWCAS([x_1 + y_1], ..., [x_n + y_n]) = \left(\sum_{j=1}^{n} w_j b_j^3 \right)^{1/3}. \quad (17)$$

Note that a lot of other particular cases could be studied [8,11]. Finally, let us present a further generalization by using quasi-arithmetic means forming the Quasi-OWAS operator.

Definition 5. A Quasi-OWAS operator of dimension n is a mapping $QOWAS$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0,1]$, such that:

$$f([x_1 + y_1], ..., [x_n + y_n]) = g^{-1} \left(\sum_{j=1}^{n} w_j g(b_j) \right), \quad (18)$$

where b_j is the jth largest of the $[x_i + y_i]$ and g is a strictly continuous monotonic function.

5 CHOQUET INTEGRALS WITH THE OWAS OPERATOR

The OWAS operator can also be extended with Choquet integrals following a similar methodology as other authors have followed in previous work [2,9]. In this case, we get the Choquet integral sum (CIS). It is very similar to the OWAS operator but not it follows the Choquet integral aggregation process. Before introducing the CIS operator, let us briefly recall the concept of a fuzzy measure. The fuzzy measure is also known as non-additive measure and can be defined as follows [6,16].

Definition 6. Let X be a universal set $X = x_1, x + 2, ..., x_n$ and $P(X)$ the power set of X. A fuzzy measure on X is a set function on m: $P(X) \longrightarrow [0,1]$ that accomplishes the following conditions:

1. $m(\emptyset) = 0, m(X) = 1$ (boundary conditions).

2. If $A, B \in P(X)$ and $A \subseteq B$, then $m(A) \leq m(B)$ (monotonicity).

Another basic definition important to remark is the Choquet integral [3] in its discrete form:

Definition 7. Let f be a positive real-valued function $f : X \to R^+$ and m be a fuzzy measure on X. The discrete Choquet integral of f with respect to m is:

$$C_m(f_1, f_2, \ldots, f_n) = \sum_{i=1}^{n} f_{(i)} \left[m(A_{(i)}) - m(A_{(i-1)})\right], \tag{19}$$

where $f_{(i)}$ is the ith largest value in the set $f_1, f_2, ..., f_n$; $A_{(i)} = x_{(1)}, ..., x_{(i)}$ $i \geq 1$; and $A_{(0)} = \emptyset$.

Next, let us look into some extensions of the Choquet integral when using the OWAS operator. First, it is possible to consider the CIS operator which is a Choquet aggregation where the set of arguments are formed by subsets of sums. The CIS operator is defined as follows.

Definition 8. Let m be a fuzzy measure on$X$. A Choquet integral sum (CIS) operator of dimension$n$ is a function CIS: $R^n \times R^n \to R$, such that:

$$\begin{aligned} CIS\left([x_1 + y_1], [x_2 + y_2], \ldots, [x_n + y_n]\right) = \\ = \sum_{j=1}^{n} b_j \left[m(A_{(j)}) - m(A_{(j-1)})\right] \end{aligned}, \tag{20}$$

where b_j is the jth largest of the $[x_i + y_i]$ value, the $[x_i + y_i]$ is the argument variable represented in the form of individual sums; $A_{(j)} = x_{(1)}, ..., x_{(j)}$ $j \geq 1$; and $A_{(0)} = \emptyset$.

This approach can be generalized by using generalized and quasi-arithmetic means [5,8]. For example, by using quasi-arithmetic means, we get the quasi-arithmetic Choquet integral sum (Quasi-CIS) operator, which is defined as follows.

Definition 9. Let m be a fuzzy measure on X. A quasi-arithmetic Choquet integral sum (Quasi-CIS) of dimension n is a function $QCIS$: $R^n \times R^n \to R$, such that:

$$\begin{aligned} Quasi - CIS\left([x_1 + y_1], \ldots, [x_n + y_n]\right) = \\ = g^{-1}\left(\sum_{j=1}^{n} g(b_j) \left[m(A_{(j)}) - m(A_{(j-1)})\right]\right) \end{aligned}, \tag{21}$$

where g is a strictly continuous monotonic function and the rest is equivalent to Eq. (20).

Note that if $g = b^{\lambda}$, the Quasi-CIS operator is equivalent to the generalized CIS (GCIS) operator. Form here, we could consider arithmetic aggregations when $\lambda = 1$, quadratic ones when $\lambda = 2$ and geometric ones when $\lambda \to 0$.

6 CONCLUSIONS

This article has presented the OWAS operator. It is an aggregation operator that aggregates a set of data assessed with sums providing a complete representation of the information from the minimum sum to the maximum one. Some general properties are studied including those coming from the OWA operator and those from the norm aggregation. The OWAS operator includes many particular cases including the simple average sum, the minimum sum and the maximum sum. These operators are useful for some specific situations according to the needs of the problem considered.

Some extensions have been considered by using generalized aggregation operators and Choquet integrals. The GOWAS operator is more general than the OWAS operator because it includes many other particular cases including the quadratic OWAS and the geometric OWAS. The CIS operator gives a more formal representation to the OWAS operator. The CIS has also been extended with quasi-arithmetic operators forming the Quasi-CIS operator.

This approach is very useful in a wide range of real world problems that deals with sums. Some interesting examples are the forecast of business and economic variables such as the sales and costs that are formed from a set of scenarios that sums different costs and sales. In future research, other additional developments will be considered by using other type of frameworks including induced [24] and weighted aggregation operators [17,19]. Some real world applications will be considered in decision making, economics and business.

Acknowledgements

Support from the University of Chile and valuable comments given by two anonymous reviewers are gratefully acknowledged.

References

[1] G. Beliakov, A. Pradera, T. Calvo (2007). *Aggregation functions: A guide for practitioners.* Springer-Verlag, Berlin.

[2] J. Belles-Sampera, J.M. Merigó, M. Guillén, M. Santolino (2014). Indicators for the characterization of discrete Choquet integrals. *Information Sciences* **267**, pp 201-216.

[3] G. Choquet (1953). Theory of capacities. *Annals Institute Fourier* **5**, pp 131-295.

[4] A. Emrouznejad, M. Marra (2014). Ordered weighted averaging operators 1988-2014: A citation based literature survey. *International Journal of Intelligent Systems* **29**, pp 994-1014.

[5] J. Fodor, J.L. Marichal, M. Roubens (1995). Characterization of the ordered weighted averaging operators. *IEEE Transactions on Fuzzy Systems* **3**, pp 236-240.

[6] M. Grabisch, J.L. Marichal, R. Mesiar, E. Pap (2009). *Aggregation Functions*. Cambridge: Cambridge University Press.

[7] N. Karayiannis (2000). Soft learning vector quantization and clustering algorithms based on ordered weighted aggregation operators. *IEEE Transactions on Neural Networks* **11**, pp 1093-1105.

[8] J.M. Merigó, M. Casanovas (2011). The uncertain induced quasi-arithmetic OWA operator. *International Journal of Intelligent Systems* **26**, pp 1-24.

[9] J.M. Merigó, M. Casanovas (2011). Decision making with distance measures and induced aggregation operators. *Computers & Industrial Engineering* **60**, pp 66-76.

[10] J.M. Merigó, M. Casanovas, S.Z. Zeng (2014). Distance measures with heavy aggregation operators. *Applied Mathematical Modelling* **38**, pp 3142-3153.

[11] J.M. Merigó, A.M. Gil-Lafuente (2009). The induced generalized OWA operator. *Information Sciences* **279**, pp 729-741.

[12] J.M. Merigó, A.M. Gil-Lafuente (2010). New decision-making techniques and their application in the selection of financial products. *Information Sciences* **180**, pp 2085-2094.

[13] J.M. Merigó, R.R. Yager (2013). Generalized moving averages, distance measures and OWA operators. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **21**, pp 533-559.

[14] J.M. Merigó, R.R. Yager (2013). Norms aggregations and OWA operators. *AGOP 2013*, Pamplona, Spain, pp 141-151.

[15] C.E. Shannon (1948). A mathematical theory of communication. *Bell Systems Technical Journal* **27**, pp 379-423.

[16] M. Sugeno (1974). *Theory of fuzzy integrals and its applications*. PhD Thesis, Tokyo Institute of Technology, Japan.

[17] V. Torra (1997). The weighted OWA operator. *International Journal of Intelligent Systems* **12**, pp 153-166.

[18] Z.S. Xu, J. Chen (2008). Ordered weighted distance measure. *Journal of Systems Science and Systems Engineering* **17**, pp 432-445.

[19] Z.S. Xu, Q.L. Da (2003). An overview of operators for aggregating information. *International Journal of Intelligent Systems* **18**, pp 953-969.

[20] R.R. Yager (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics B* **18**, pp 183-190.

[21] R.R. Yager (1993). Families of OWA operators. *Fuzzy Sets and Systems* **59**, pp 125-148.

[22] R.R Yager (2004). Generalized OWA aggregation operators. *Fuzzy Optimization and Decision Making* **3**, pp 93-107.

[23] R.R. Yager (2010). Norms induced from OWA operators. *IEEE Transactions on Fuzzy Systems* **18**, pp 57-66.

[24] R.R. Yager, D.P. Filev (1999). Induced ordered weighted averaging operators. *IEEE Transactions on Systems, Man and Cybernetics B* **29**, pp 141-150.

[25] R.R.Yager, J.Kacprzyk (1997). *The ordered weighted averaging operators: Theory and applications*. Kluwer Academic Publishers, Norwell, MA.

[26] R.R. Yager, J. Kacprzyk and G. Beliakov (2011). *Recent developments on the ordered weighted averaging operators: Theory and practice*. Springer-Verlag, Berlin.

ON THE FUNCTIONAL EQUATION $f(m_1(x+y)) = m_2(f(x)+f(y))$ FOR INJECTIVE FUNCTION m_2

Wanda Niemyska and Michał Baczyński
Institute of Mathematics, University of Silesia
40-007 Katowice, ul. Bankowa 14, Poland
{wniemyska,michal.baczynski}@us.edu.pl

Summary

Recently, in some considerations connected with the distributivity laws of fuzzy implications over triangular norms and conorms, the following functional equation

$$f(\min(x+y,a)) = \min(f(x)+f(y),b),$$

connected with the Cauchy equation, appeared, where a, b are finite or infinite nonnegative constants (see [4]). In [7] we have considered a generalized version of this equation in the case when both a and b are finite, namely the equation

$$f(m_1(x+y)) = m_2(f(x)+f(y)),$$

where m_1, m_2 are functions defined on some finite intervals of $\mathbb{R}$ satisfying additional assumptions. In this article we consider the above equation when m_1, m_2 are defined on some finite or infinite sets and satisfy only one additional assumption: m_2 is injective.

Keywords: Fuzzy connectives, fuzzy implication, distributivity, functional equation, Jensen equation.

1 PRELIMINARIES

Distributivity of fuzzy implication functions over different fuzzy logic connectives has been thoroughly investigated in recent past by many authors (see [1, 19, 8, 17, 18, 4, 2, 3, 16, 15, 5]). In general we can consider four such distributivity equations:

$$I(x, C_1(y,z)) = C_2(I(x,y), I(x,z)), \tag{D1}$$

$$I(x, D_1(y,z)) = D_2(I(x,y), I(x,z)), \tag{D2}$$

$$I(C(x,y), z) = D(I(x,z), I(y,z)), \tag{D3}$$

$$I(D(x,y), z) = C(I(x,z), I(y,z)), \tag{D4}$$

satisfied for all $x, y, z \in [0,1]$, where I is some generalization of classical implication, C, C_1, C_2 are some generalizations of classical conjunction and D, D_1, D_2 are some generalizations of classical disjunction.

The importance of such equations in Fuzzy Control and Fuzzy Systems has been first emphasized by Combs and Andrews [11], wherein they exploit the following classical tautology

$$(p \wedge q) \to r \equiv (p \to r) \vee (q \to r)$$

in their inference mechanism towards reduction in the complexity of fuzzy "IF-THEN" rules. Subsequently, there were many discussions [9, 10, 12, 14], most of them pointing out the need for a theoretical investigation required for employing such equations.

If we use continuous Archimedean t-norms and t-conorms in above distributivity laws (D1) – (D4), then from their representation theorems (see [13]) we obtain the following four equations

$$f_x(\min(t_1(y)+t_1(z), t_1(0))) = \min(f_x(t_1(y)) + f_x(t_1(z)), t_2(0)),$$

$$g_x(\min(s_1(y)+s_1(z), s_1(1))) = \min(g_x(s_1(y)) + g_x(s_1(z)), s_2(1)),$$

$$h^z(\min(t(x)+t(y), t(0))) = \min(h^z(s(x)) + h^z(s(y)), s(1)),$$

$$k^z(\min(s(x)+s(y), s(1))) = \min(k^z(t(x)) + k^z(t(y)), t(0)),$$

where

- t_1, t_2, t are functions occurring in the representations of T_1, T_2, T, respectively,

ISBN: 978–83–8012–519–3

- s_1, s_2, s are functions occurring in the representations of S_1, S_2, S, respectively,
- $f_x(\cdot) = t_2 \circ I(x, t_1^{-1}(\cdot))$, for a fixed $x \in [0,1]$,
- $g_x(\cdot) = s_2 \circ I(x, s_1^{-1}(\cdot))$, for a fixed $x \in [0,1]$,
- $h^z(\cdot) = s \circ I(x, t^{-1}(\cdot))$, for a fixed $z \in [0,1]$,
- $k^z(\cdot) = t \circ I(x, s^{-1}(\cdot))$, for a fixed $z \in [0,1]$.

Observe that the first equation may be written in the following form

$$f_x(\min(u+v, t_1(0))) = \min(f_x(u) + f_x(v), t_2(0)),$$

where $u, v \in [0, t_1(0)]$, and f_x is an unknown function. Similarly, the second equation may be written in the form

$$g_x(\min(u+v, s_1(1))) = \min(g_x(u) + g_x(v), s_2(1)),$$

where $u, v \in [0, s_1(1)]$, and g_x is an unknown function. The other two equations can be written in the similar way. Thus, in the papers [4, 2], authors have found the solutions of the following functional equations

$$\begin{aligned} f(\min(x+y, r_1)) &= \min(f(x) + f(y), r_2), \\ f(\min(x+y, r_1)) &= f(x) + f(y), \\ f(x+y) &= \min(f(x) + f(y), r_2), \\ f(x+y) &= f(x) + f(y), \end{aligned}$$

when in the first case $f\colon [0, r_1] \to [0, r_2]$, in the second case $f\colon [0,\infty] \to [0, r_2]$, in the third case $f\colon [0, r_1] \to [0,\infty]$ and finally in the last case $f\colon [0,\infty] \to [0,\infty]$. Observe that the last equation is in fact the Cauchy equation, while the other equations can be seen as some modifications of classical Cauchy equation.

In the articles [6, 7] we considered the generalized version of the first equation i.e., we have replaced both functions $\min(\cdot, r_1)$, $\min(\cdot, r_2)$ occurring directly in this equation, by functions m_1, m_2 satisfying some assumptions. This means that we studied there the following equation

$$f(m_1(x+y)) = m_2(f(x) + f(y)). \tag{1}$$

In this article we continue these investigations and we show the solutions of the equation (1) when m_1, m_2 are defined on some finite or infinite sets and satisfy only one additional assumption: m_2 is injective.

2 SOME NEW RESULTS PERTAINING TO JENSEN EQUATION

In this section we would like to present new results connected to Jensen equation

$$f(x) + f(y) = 2f(\frac{x+y}{2}). \tag{J}$$

We consider here few cases where domain or codomain are extended to the infinity. Results from this section will be used in the main theorems of the next section.

Lemma 2.1. *Let $D \subset \mathbb{R}^N$ be convex set such that $intD \neq \emptyset$, and let $f\colon D \to \mathbb{R}$ be a solution of Jensen equation* (J)*. If f is bounded above or bounded below on D, then*

$$f(x) = cx + a, \quad x \in D, \tag{2}$$

for some constants $c \in \mathbb{R}^N$, $a \in \mathbb{R}$.

Proof. This lemma is a simple corollary from results stated in [13]. First, since f is a solution of the equation (J) and it is a bounded function (from above or from below), thus f is continuous [see [13, Theorem XIII.2.3]]. And then, by [13, Theorem XIII.2.2], we have that f takes the form (2). □

Proposition 2.2. *Let $f\colon [0,\infty] \to [0,b]$, for some finite $b \in [0,\infty)$. Then the following statements are equivalent:*

(i) Function f satisfies the Jensen equation (J) *for all $x, y \in [0,\infty]$.*

(ii) Function f is constant, i.e., there exists $d \in [0,b]$, such that $f = d$.

Proof. $(i) \Leftarrow (ii)$ It is obvious that every constant function satisfies (J).

$(i) \Rightarrow (ii)$ Let us define function $g\colon [0,\infty) \to [0,b]$ as a truncation f to real domain, i.e., $g = f\big|_{\mathbb{R}}$. Obviously function g satisfies Jensen equation (J) as well, and since g is bounded, thus from Lemma 2.1 there exist $c, d \in \mathbb{R}$ such that

$$g(x) = cx + d, \quad x \in [0,\infty).$$

Of course c and d are nonnegative, since $f \geq 0$. Therefore $f(x) = cx + d$, for $x \in [0,\infty)$, as well. Putting $x = 0$ and $y = \infty$ to the equation (J) for function f, we get

$$\begin{aligned} f(0) + f(\infty) = 2f\left(\frac{0+\infty}{2}\right) \\ \iff d + f(\infty) = 2f(\infty), \end{aligned}$$

thus $f(\infty) = d$, since f is finite. Now putting any $x \in (0,\infty)$ and $y = \infty$ to the equation (J) we get

$$\begin{aligned} f(x) + f(\infty) = 2f\left(\frac{\infty+x}{2}\right) \\ \iff (cx+d) + d = 2f(\infty) = 2d, \end{aligned}$$

thus $c = 0$. Therefore the only solution is $f = d$, for $d \in [0,\infty)$. □

Proposition 2.3. *For a function $f\colon [0,\infty) \to [0,\infty]$ the following statements are equivalent:*

(i) Function f satisfies the Jensen equation (J) *for all $x, y \in [0,\infty)$.*

(ii) Either $f = \infty$, or there exists $d \in [0,\infty)$ such that

$$f(x) = \begin{cases} d, & \text{if } x = 0, \\ \infty, & \text{if } x > 0, \end{cases} \tag{3}$$

or there exist $c, d \in [0,\infty)$ such that

$$f(x) = cx + d, \tag{4}$$

for all $x \in [0,\infty)$.

Proof. $(i) \Leftarrow (ii)$ It is a direct calculation that all the above functions satisfy the equation (J).

$(i) \Rightarrow (ii)$ First let us assume that f takes only real values, i.e., $f\colon [0,\infty) \to [0,\infty)$. Function f is nonnegative and satisfies (J), thus from Lemma 2.1 there exist $c, d \in [0,\infty)$ such that

$$f(x) = cx + d, \quad x \in [0,\infty).$$

This way we get the solution (4).
In the opposite case, there exists $x_\infty \in [0,\infty)$ such that $f(x_\infty) = \infty$. Putting, to the equation (J), $x = x_\infty$ and $y = 2z - x_\infty$, for any $z \geq \frac{1}{2}x_\infty$, we get

$$\begin{aligned} f(x_\infty) + f(2z - x_\infty) &= 2f\left(\frac{x_\infty + (2z - x_\infty)}{2}\right) \\ \iff \quad \infty + f(2z - x_\infty) &= 2f(z) \\ \iff \quad \infty &= f(z). \end{aligned}$$

Therefore $f(z) = \infty$ for all $z \geq \frac{1}{2}x_\infty$. Now let us assume that $f([\frac{1}{2^n}x_\infty, \infty)) = \{\infty\}$ for some $n \in \mathbb{N}$. Then for any $z \geq \frac{1}{2^{n+1}}x_\infty$ we have $(2z - \frac{1}{2^n}x_\infty) \geq 0$ and

$$\infty = f(\frac{1}{2^n}x_\infty) + f(2z - \frac{1}{2^n}x_\infty) = 2f(z).$$

Thus $f(z) = \infty$, which gives us $f([\frac{1}{2^{n+1}}x_\infty, \infty)) = \{\infty\}$. This way we have proved by an induction that $f([\frac{1}{2^n}x_\infty, \infty)) = \{\infty\}$ for any $n \in \mathbb{N}$. Therefore $f((0,\infty)) = \{\infty\}$, since for any $z \in (0,\infty)$ there exists $n \in \mathbb{N}$ such that $z > \frac{1}{2^n}x_\infty$. The remaining value for 0 is not specified, thus we can get $f = \infty$ or the solution (3). □

Proposition 2.4. *Let $f\colon [0,a] \to [0,\infty]$, for some finite $a \in [0,\infty)$. Then the following statements are equivalent:*

(i) Function f satisfies the Jensen equation (J) *for all $x, y \in [0,a]$.*

(ii) Either $f = \infty$, or there exists $d \in [0,\infty)$ such that

$$f(x) = \begin{cases} d, & \text{if } x = 0, \\ \infty, & \text{if } x > 0, \end{cases} \tag{5}$$

or

$$f(x) = \begin{cases} \infty, & \text{if } x < a, \\ d, & \text{if } x = a, \end{cases} \tag{6}$$

or there exist nonnegative $c, d \in \mathbb{R}$ such that

$$f(x) = cx + d, \tag{7}$$

for all $x \in [0,a]$.

Proof. $(i) \Leftarrow (ii)$ It is a direct calculation that all the above functions satisfy the equation (J).

$(i) \Rightarrow (ii)$ First, as in the proof of Proposition 2.3, let us assume that f takes only real values, i.e., $f\colon [0,a] \to [0,\infty)$. Function f is nonnegative and satisfies (J), thus from Lemma 2.1 there exist $c, d \in [0,\infty)$ such that $f(x) = cx + d$, for all $x \in [0,a]$. This way we get the solution (7).
Next let us assume that there exists $x_\infty \in [0,a]$ such that $f(x_\infty) = \infty$. Putting $x = x_\infty$ and $y = a - x_\infty$ to the Jensen equation (J) we get

$$\begin{aligned} \infty = f(x_\infty) &= f(x_\infty) + f(a - x_\infty) \\ &= 2f\left(\frac{x_\infty + a - x_\infty}{2}\right) = 2f\left(\frac{a}{2}\right). \end{aligned}$$

Then putting $x = 0$, $y = a$ to the Jensen equation (J) we get $f(0) + f(a) = 2f(\frac{a}{2}) = \infty$, thus $f(0) = \infty$ or $f(a) = \infty$.

1) $f(0) = \infty$. First observe that for any $x \in [0, \frac{a}{2}]$ we have

$$\infty = f(0) + f(2x) = 2f(x),$$

thus $f(x) = \infty$ and we have $f([0, \frac{a}{2}]) = \{\infty\}$. Next we can easily prove by an induction that $f([0, \frac{2^n-1}{2^n}a]) = \{\infty\}$, for any $n \in \mathbb{N}$. Therefore $f([0,a)) = \{\infty\}$, since for any $x \in [0,a)$ there exists $n \in \mathbb{N}$ such that $x < \frac{2^n-1}{2^n}a$. The remaining value for a is not specified, thus we can get $f = \infty$ or the solution (6).

2) $f(a) = \infty$. In a similar way as in 1) we get $f = \infty$ or the solution (5).

□

Proposition 2.5. *For a function $f\colon [0,\infty] \to [0,\infty]$ the following statements are equivalent:*

(i) Function f satisfies the Jensen equation (J) *for all $x, y \in [0,\infty]$.*

(ii) Either $f = \infty$, or $f = d$ for some $d \in [0, \infty)$, or there exists $d \in [0, \infty)$ such that

$$f(x) = \begin{cases} d, & \text{if } x = 0, \\ \infty, & \text{if } x > 0, \end{cases} \tag{8}$$

or there exist $c, d \in [0, \infty)$ such that

$$f(x) = \begin{cases} cx + d, & \text{if } x < \infty, \\ \infty, & \text{if } x = \infty, \end{cases} \tag{9}$$

for all $x \in [0, \infty]$.

Proof. $(i) \Leftarrow (ii)$ It is a direct calculation that all the above functions satisfy the equation (J).

$(i) \Rightarrow (ii)$ Here we apply Proposition 2.3 for $f\big|_{\mathbb{R}}$ and then analyze the remaining values f of ∞. We omit the details of that proof, since they are quite similar to previous cases. □

3 SOLUTIONS OF EQUATION $f(m_1(x + y)) = m_2(f(x) + f(y))$ WHEN m_2 IS INJECTIVE

In this section we characterize solutions to the equation (1) for different domains and codomains of functions f, m_1 and m_2, namely when $m_1 \colon [0, 2r_1] \to [0, r_1]$, $m_2 \colon [0, 2r_2] \to [0, r_2]$ and $f \colon [0, r_1] \to [0, r_2]$, where constants $r_1, r_2 \in [0, \infty]$ may be finite or infinite. All the time we have only one additional assumption that m_2 is injective.

First we want to recall the lemma that we obtained in [7]. The only difference here is that the constants r_1 and r_2 may also be infinite, but the proofs in these additional cases are very similar.

Lemma 3.1 (cf [7, Lemma 2.1])**.** *Let $r_1, r_2 \in [0, \infty]$ be some numbers, finite or infinite, and let $m_1 \colon [0, 2r_1] \to [0, r_1]$, $m_2 \colon [0, 2r_2] \to [0, r_2]$ be given functions. If m_2 is injective and a function $f \colon [0, r_1] \to [0, r_2]$ satisfies the functional equation (1), then f satisfies the Jensen equation (J).*

Next, we consider four cases corresponding to all possible combinations of finite and infinite values of r_1, r_2. First one, when both r_1 and r_2 are finite, was solved in [7].

Theorem 3.2 (cf. [7, Theorem 2.2])**.** *Let $r_1, r_2 \in [0, \infty)$ be some numbers and let $m_1 \colon [0, 2r_1] \to [0, r_1]$, $m_2 \colon [0, 2r_2] \to [0, r_2]$, $f \colon [0, r_1] \to [0, r_2]$ be given functions. Further, let m_2 be injective. Then the following statements are equivalent:*

(i) The triple of functions m_1, m_2, f satisfies the equation (1) for all $x, y \in [0, r_1]$.

(ii) Either $f = d$ for some $d \in [0, r_2]$ and $m_2(2d) = d$, or $f(x) = cx + d$ for some $c, d \in \mathbb{R}$, $c \neq 0$ such that $cx + d \in [0, r_2]$ for all $x \in [0, r_1]$ and

$$m_1(x) = \frac{m_2(cx + 2d) - d}{c}.$$

Obviously all the solutions in the last theorem are continuous as they will be in the next case, when $r_1 = \infty$, but r_2 stays finite.

Theorem 3.3. *Let $r_2 \in [0, \infty)$ be some number and let $m_1 \colon [0, \infty] \to [0, \infty]$, $m_2 \colon [0, 2r_2] \to [0, r_2]$ and $f \colon [0, \infty] \to [0, r_2]$ be given functions. Further, let m_2 be injective. Then the following statements are equivalent:*

(i) The triple of functions m_1, m_2, f satisfies the equation (1) for all $x, y \in [0, \infty]$.

(ii) $f = d$, for some $d \in [0, r_2]$, and $m_2(2d) = d$.

Proof. $(i) \Leftarrow (ii)$ It is a direct calculation.

$(i) \Rightarrow (ii)$ From Lemma 3.1 we have that f satisfies the Jensen equation (J). Then from Proposition 2.2 we get that function f is constant, i.e., there exists $d \in [0, r_2]$ such that $f = d$. Finally, from the equation (1) we have $m_2(2d) = d$. □

In the following two cases, when we extend a codomain of a function f to the infinity, there appear also discontinuous solutions.

Theorem 3.4. *Let $r_1 \in [0, \infty)$ be some number and let $m_1 \colon [0, 2r_1] \to [0, r_1]$, $m_2 \colon [0, \infty] \to [0, \infty]$ and $f \colon [0, r_1] \to [0, \infty]$ be given functions. Further, let m_2 be injective. Then the following statements are equivalent:*

(i) The triple of functions m_1, m_2, f satisfies the equation (1) for all $x, y \in [0, r_1]$.

(ii) Either $f = \infty$ and $m_2(\infty) = \infty$, or there exists $d \in [0, \infty)$ such that

$$f(x) = \begin{cases} d, & \text{if } x = 0, \\ \infty, & \text{if } x > 0, \end{cases} \quad x \in [0, r_1], \tag{10}$$

and ($m_1(0) = 0, m_1(x) > 0$ for all $x > 0$, $m_2(2d) = d$ and $m_2(\infty) = \infty$), or ($m_1(0) > 0, m_1(x) = 0$ for all $x > 0$, $m_2(2d) = \infty$ and $m_2(\infty) = d$), or

$$f(x) = \begin{cases} \infty, & \text{if } x < r_1, \\ d, & \text{if } x = r_1, \end{cases} \quad x \in [0, r_1], \tag{11}$$

and ($m_1(2r_1) = r_1, m_1(x) < r_1$ for all $x < 2r_1$, $m_2(2d) = d$ and $m_2(\infty) = \infty$), or ($m_1(2r_1) <$

$r_1, m_1(x) = r_1$ *for all* $x < 2r_1$, $m_2(2d) = \infty$ *and* $m_2(\infty) = d$*), or there exist* $c, d \in [0, \infty)$ *such that*

$$f(x) = cx + d, \tag{12}$$

and $m_2(2d) = d$*, if* $c = 0$*, while*

$$m_1(x) = \frac{m_2(cx + 2d) - d}{c},$$

for all $x \in [0, r_1]$*, if* $c > 0$.

Proof. $(i) \Leftarrow (ii)$ It is quite easy to check that these functions satisfy (1). However, this is kind of a tedious work and we omit it here.

$(i) \Rightarrow (ii)$ From Lemma 3.1 we have that f satisfies the Jensen equation (J). Then from Proposition 2.4 we get that $f = \infty$ or f takes the form (10), (11) or (12). Let us consider all those four possibilities.

1) $f = \infty$. Then the equation (1) takes the form $\infty = m_2(\infty)$.

2) f takes the form (10). Putting $x = y = 0$ to the equation (1) we get

$$f(m_1(0)) = m_2(f(0) + f(0)) = m_2(2d).$$

Two cases are possible:

(A1) $m_1(0) = 0$, and then $m_2(2d) = f(0) = d$;

(A2) $m_1(0) > 0$, and then $m_2(2d) = \infty$.

Next, let us put to the equation (1) $x, y \in [0, r_1]$ such that $x + y > 0$. Assume $x > 0$, without losing generality. We have then $f(m_1(x + y)) = m_2(\infty + f(y)) = m_2(\infty)$, which is equivalent to

$$f(m_1(x)) = m_2(\infty),$$

for all $x \in (0, 2r_1]$. If there exists $x_1 \in (0, 2r_1]$ such that $m_1(x_1) = 0$, then $m_2(\infty) = f(m_1(x_1)) = f(0) = d$. And if there exists $x_2 \in (0, 2r_1]$ such that $m_1(x_2) > 0$ then $m_2(\infty) = f(m_1(x_2)) = \infty$. Obviously, existence of x_1 and x_2 at the same time leads to the contradiction, thus we have two possibilities:

(B1) $m_1(x) = 0$ for all $x \in (0, 2r_1]$, and then $m_2(\infty) = d$;

(B2) $m_1(x) > 0$ for all $x \in (0, 2a]$, and then $m_2(\infty) = \infty$.

Since m_2 is injective and $d \neq \infty$, thus (A1) cannot occur together with (B1), and (A2) cannot occur together with (B2). The remaining combinations (A1) with (B2) and (A2) with (B1) complete the thesis for f in the form (10).

3) f takes the form (11). We can analyze this case in a similar way as the last one. We omit here the details.

4) f takes the form (12), i.e., there exist $c, d \in [0, \infty)$ such that $f(x) = cx + d$ for all $x \in [0, r_1]$. If $c = 0$ then $f = d$ and the equation (1) takes the form $d = m_2(2d)$. On the other hand, if $c > 0$ then the equation (1) takes the form $cm_1(x + y) + d = m_2(cx + d + cy + d) = m_2(c(x + y) + 2d)$, for all $x, y \in [0, r_1]$, which is equivalent to

$$m_1(x) = \frac{m_2(cx + 2d) - d}{c},$$

for all $x \in [0, 2r_1]$. That completes the proof.

□

Finally we consider the last case when both r_1 and r_2 are infinite.

Theorem 3.5. *Let* $m_1, m_2, f\colon [0, \infty] \to [0, \infty]$ *be given functions. Further, let* m_2 *be injective. Then the following statements are equivalent:*

(i) The triple of functions m_1, m_2, f *satisfies the equation (1) for all* $x, y \in [0, \infty]$.

(ii) Either $f = \infty$ *and* $m_2(\infty) = \infty$*, or there exists* $d \in [0, \infty)$ *such that* $f = d$ *and* $m_2(2d) = d$*, or*

$$f(x) = \begin{cases} d, & \text{if } x = 0, \\ \infty, & \text{if } x > 0, \end{cases} \quad x \in [0, \infty], \tag{13}$$

and $(m_1(0) = 0, m_1(x) > 0$ *for all* $x > 0$, $m_2(2d) = d$ *and* $m_2(\infty) = \infty$*), or* $(m_1(0) > 0, m_1(x) = 0$ *for all* $x > 0$, $m_2(2d) = \infty$ *and* $m_2(\infty) = d$*), or*

$$f(x) = \begin{cases} d, & \text{if } x < \infty, \\ \infty, & \text{if } x = \infty, \end{cases} \quad x \in [0, \infty], \tag{14}$$

and $(m_1(\infty) = \infty, m_1(x) < \infty$ *for all* $x < \infty$, $m_2(2d) = d$ *and* $m_2(\infty) = \infty$*), or* $(m_1(\infty) < \infty, m_1(x) = \infty$ *for all* $x < \infty$, $m_2(2d) = \infty$ *and* $m_2(\infty) = d$*), or there exist* $c, d \in [0, \infty), c \neq 0$*, such that*

$$f(x) = \begin{cases} cx + d, & \text{if } x < \infty, \\ \infty, & \text{if } x = \infty, \end{cases} \quad x \in [0, \infty], \tag{15}$$

and

$$m_1(x) = \frac{m_2(cx + 2d) - d}{c},$$

for all $x \in [0, \infty]$.

4 CONCLUSIONS

In this article we have discussed some solutions of the functional equation $f(m_1(x+y)) = m_2(f(x)+f(y))$, which generalizes some equations connected with the solutions of the distributivity laws for fuzzy implications over triangular norms and triangular conorms. It would be interesting to find applications of our solutions in fuzzy control and/or fuzzy logic. As a byproduct we obtained solutions of the Jensen equation in few cases when domain or codomain of f is extended to the infinity. Note that Baczyński, Jayaram in [4] and Baczyński in [2] did similar work with the Cauchy equation. It is of our interest to analyze the connections between those results.

References

[1] M. Baczyński. On a class of distributive fuzzy implications. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 9:229–238, 2001.

[2] M. Baczyński. On the distributivity of fuzzy implications over continuous and Archimedean triangular conorms. *Fuzzy Sets and Systems*, 161(10):1406–1419, 2010.

[3] M. Baczyński. On the distributivity of fuzzy implications over representable uninorms. *Fuzzy Sets and Systems*, 161(17):2256–2275, 2010.

[4] M. Baczyński and B. Jayaram. On the distributivity of fuzzy implications over nilpotent or strict triangular conorms. *IEEE Trans. Fuzzy Syst.*, 17(3):590–603, June 2009.

[5] M. Baczyński and F. Qin. Some remarks on the distributive equation of fuzzy implication and the contrapositive symmetry for continuous, Archimedean t-norms. *Internat. J. Approx. Reason.*, 54(2):290–296, 2013.

[6] M. Baczyński, T. Szostok, and W. Niemyska. On a functional equation related to distributivity of fuzzy implications. In *Proc. FUZZ IEEE 2013*, pages 1–5, Hyderabad, India, 2013.

[7] M. Baczyński, T. Szostok, and W. Niemyska. On a functional equation related to the distributivity of fuzzy implications over triangular norms and conorms. *Kybernetika*, 50(4):679–695, 2014.

[8] J. Balasubramaniam and C. J. M. Rao. On the distributivity of implication operators over T and S norms. *IEEE Trans. Fuzzy Syst.*, 12(2):194–198, Apr. 2004.

[9] W. E. Combs. Author's reply. *IEEE Trans. Fuzzy Syst.*, 7(3):371–373, June 1999.

[10] W. E. Combs. Author's reply. *IEEE Trans. Fuzzy Syst.*, 7(4):477–478, Aug. 1999.

[11] W. E. Combs and J. E. Andrews. Combinatorial rule explosion eliminated by a fuzzy rule configuration. *IEEE Trans. Fuzzy Syst.*, 6(1):1–11, Feb. 1998.

[12] S. Dick and A. Kandel. Comments on "Combinatorial rule explosion eliminated by a fuzzy rule configuration". *IEEE Trans. Fuzzy Syst.*, 7(4):475–477, Aug. 1999.

[13] M. Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities. Cauchy's Equation and Jensen's Inequality.* Państwowe Wydawnictwo Naukowe (Polish Scientific Publishers) and Uniwersytet Śląski, Warszawa–Kraków–Katowice, 1985.

[14] J. M. Mendel and Q. Liang. Comments on "Combinatorial rule explosion eliminated by a fuzzy rule configuration". *IEEE Trans. Fuzzy Syst.*, 7(3):369–371, June 1999.

[15] F. Qin, M. Baczyński, and A. Xie. Distributive equations of implications based on continuous triangular norms (I). *IEEE Trans. Fuzzy Syst.*, 20(1):153–167, Feb. 2012.

[16] F. Qin and L. Yang. Distributive equations of implications based on nilpotent triangular norms. *Internat. J. Approx. Reason.*, 51:984–992, 2010.

[17] D. Ruiz-Aguilera and J. Torrens. Distributivity of strong implications over conjunctive and disjunctive uninorms. *Kybernetika*, 42(3):319–336, 2006.

[18] D. Ruiz-Aguilera and J. Torrens. Distributivity of residual implications over conjunctive and disjunctive uninorms. *Fuzzy Sets and Systems*, 158(1):23–37, 2007.

[19] E. Trillas and C. Alsina. On the law $[(p \wedge q) \to r] = [(p \to r) \vee (q \to r)]$ in fuzzy logic. *IEEE Trans. Fuzzy Syst.*, 10(1):84–88, Feb. 2002.

SIMILARITY MEASURE DEFINED FROM OVERLAP FUNCTION

Barbara Pękala
Faculty of Mathematics and Natural Sciences University of Rzeszów,
Pigonia 1, 35-310 Rzeszów, Poland
bpekala@ur.edu.pl

Summary

The problem of measuring the degree of inclusion and similarity measure for two intuitionistic interval-valued fuzzy sets is considered. Moreover, some properties of inclusion and similarity measure and some correlation, between them are examined.

Keywords: Degree of inclusion, Fuzzy set, Similarity, Interval-valued intuitionistic fuzzy set.

1 INTRODUCTION

In this paper interval-valued intuitionistic fuzzy sets are studied. Many new approaches and theories treating imprecision and uncertainty have been proposed since fuzzy set were introduced by Zadeh. As extensions of classical fuzzy set theory, intuitionistic fuzzy sets [1], interval-valued fuzzy sets are very useful in dealing with imprecision and uncertainty. Especially, interval-valued intuitionistic fuzzy set introduced by Atanassov [2], as a combining concept of intuitionistic fuzzy set and interval-valued fuzzy set, greatly furnishes the additional capability to deal with vague information and model non-statistical uncertainty by providing a membership interval and a nonmembership interval. Therefore, interval-valued intuitionistic fuzzy sets have played a significant role in the uncertain system and received much attention from researchers. Recently many scholars have investigated interval-valued intuitionistic fuzzy sets and obtained some meaningful results in the fields of multicriteria decision making [18] and group decision making with interval-valued intuitionistic fuzzy sets [22]. Many researchers have examined different types of transitivity and have proposed some distance measures, similarity measures ([4], [5], [11], [12], [13], [14], [15], [16], [21] or [23]) and correlation measures of interval-valued intuitionistic fuzzy sets ([22], [25]) moreover there were presented applications of such considerations to real-life problems involving pattern recognition, medical diagnosis and decision-making. Additionally, some inclusion measures of intuitionistic fuzzy sets and interval-valued intuitionistic fuzzy sets have been proposed in [19], which also play important roles in the application areas such as approximate reasoning, statistical inference and decision making. Thus the motivation of the present paper is to propose a more natural tools for estimating the degree of inclusion between interval-valued intuitionistic fuzzy sets and explore their properties. The paper is organized as follows. In Section 2 we recall basic information on interval-valued intuitionistic fuzzy sets. We also show there crisp definition of inclusion. In Section 3 we present some inclusion measure for interval-valued intuitionistic fuzzy sets. Then some properties of inclusion measure for interval-valued intuitionistic fuzzy sets are examined. Section 4 is concerning similarity measure and properties connected with transitivity and bisymmetry properties.

2 INTERVAL-VALUED INTUITIONISTIC FUZZY SETS

Throughout this paper the discourse set is denoted as $X = \{x_1, x_2, \ldots, x_n\}$, $P(X)$ stands for the set of all crisp subsets in X, respectively. A fuzzy set ρ in X is defined as the set of ordered pairs

$$\rho = \{< x_i, R(x_i) >: x_i \in X\},$$

where $\rho : X \to [0, 1]$ is the membership function of ρ and $R(x_i)$ is the grade of belongingness of x_i into ρ. A family of all fuzzy sets in X will be denoted by $FS(X)$. According to Zadeh seminal paper [24] introducing fuzzy sets we define inclusion for two fuzzy sets ρ and σ in X as follows

$$\rho \subset \sigma \Leftrightarrow R(x_i) \le S(x_i) \quad (1)$$

ISBN: 978–83–8012–519–3

for each $i \in \{1, ..., n\}, n \in N$, where we denote membership functions of the sets $\rho = \{< x_i, R(x_i) >: x_i \in X\}$ and $\sigma = \{< x_i, S(x_i) >: x_i \in X\}$, respectively. For each fuzzy set ρ the grade of nonbelongingness of x into ρ is automatically equal to $1 - R(x_i)$. However, in real life the linguistic negation does not always identify with logical negation. This situation is very common in natural language processing, computing with words and their applications in many areas. Thus, although fuzzy set theory provides useful tools for dealing with uncertain information, Atanassov [1] suggested a generalization of classical fuzzy set, called an interval-valued intuitionistic fuzzy set.

Definition 1 ([3])**.** An interval-valued intuitionistic fuzzy set ρ in the finite universe X can be expressed in the form

$$\rho = \{< x_i, R(x_i), \mathfrak{R}(x_i) >: x_i \in X\},$$

where $R(x_i) = [\underline{R}(x_i), \overline{R}(x_i)]$ is called the interval membership degree of an element x_i to interval-valued intuitionistic fuzzy set ρ, while $\mathfrak{R}(x_i) = [\underline{\mathfrak{R}}(x_i), \overline{\mathfrak{R}}(x_i)]$ is the interval non-membership degree of this element to the set, and the condition $0 \leq \overline{R}(x_i) + \overline{\mathfrak{R}}(x_i) \leq 1$ must hold for any $x_i \in X$.

For shorter notation of an interval-valued intuitionistic fuzzy set we use a pair $\rho = (R, \mathfrak{R})$. The family of all interval-valued intuitionistic fuzzy sets described in the given sets X is denoted by $IVIFS(X)$.

The boundary elements in $IVIFS(X)$ are $1^{\cdot} = (\mathbf{1}, \mathbf{0})$ and $0^{\cdot} = (\mathbf{0}, \mathbf{1})$, where $\mathbf{0} = [0, 0], \mathbf{1} = [1, 1]$. Basic operations for $\rho = (R, \mathfrak{R})$, $\sigma = (S, \mathfrak{S}) \in IVIRS(X)$ are the union, the intersection and the complement, respectively

$$\rho \vee \sigma = ([\underline{R} \vee \underline{S}, \overline{R} \vee \overline{S}], [\underline{\mathfrak{R}} \wedge \underline{\mathfrak{S}}, \overline{\mathfrak{R}} \wedge \overline{\mathfrak{S}}]),$$
$$\rho \wedge \sigma = ([\underline{R} \wedge \underline{S}, \overline{R} \wedge \overline{S}], [\underline{\mathfrak{R}} \vee \underline{\mathfrak{S}}, \overline{\mathfrak{R}} \vee \overline{\mathfrak{S}}]),$$
$$\rho' = (\mathfrak{R}, R).$$

Moreover, the order is defined by

$$\rho \leq \sigma \Leftrightarrow (\underline{R} \leq \underline{S}, \overline{R} \leq \overline{S}, \ \underline{\mathfrak{S}} \leq \underline{\mathfrak{R}}, \overline{\mathfrak{S}} \leq \overline{\mathfrak{R}}). \quad (2)$$

The pair $(IVIFS(X), \leq)$ is a partially ordered set. Operations $\vee, \wedge$ are the binary supremum and infimum in the family $IVIFS(X)$, respectively. The family $(IVIFS(X), \vee, \wedge)$ is a complete, distributive lattice.

Similarly to [5], [14], [16] or [23] we will examine the transitivity property and equivalence relation, i.e. a relation $\psi \subset X \times X$ is an F-equivalence relation if it fulfils

- reflexivity, $\psi(x, x) = 1$;
- symmetry, $\psi(x, y) = \psi(y, x)$;
- F-transitivity, $F(\psi(x, z), \psi(z, y)) \leq \psi(x, y)$ for $x, y, z \in X$.

3 INCLUSION MEASURE

Firstly, we recall some definition of aggregation operators.

Definition 2 ([9])**.** An operation $\mathcal{A} : L^n \to L$ is called an aggregation function on a bounded lattice L if it is increasing and

$$\mathcal{A}(\underbrace{0_L, ..., 0_L}_{n\times}) = 0_L, \ \ \mathcal{A}(\underbrace{1_L, ..., 1_L}_{n\times}) = 1_L.$$

Especially we have

Definition 3 ([10])**.** A triangular norm T on a bounded lattice L is an increasing, commutative, associative operation $T : L^2 \to L$ with a neutral element 1_L.
A triangular conorm S on L is an increasing, commutative, associative operation $S : L^2 \to L$ with a neutral element 0_L.

Now, we recall the definition of an overlap function which generalizes intersection operators such as the minimum. Overlap functions are special kinds of aggregation operators that have been recently proposed for applications involving the overlap problem and/or when the associativity property is not strongly required, as in imaging processing and decision making based on fuzzy preference relations, respectively. Therefore, in those cases, it is not necessary the use of t-norms or t-conorms as the combination/separation operators. For example, overlap functions allowed the development of some construction methods for the concepts of indifference and incomparability, as introduced by Bustince et al. in [7]. The notions of overlapping arise from a common problem in many fields: how to assign a given element or object to exactly one class among several available. The notion of overlap function was presented in [6] and [20] to address the former difficulty in the context of image processing. So it will be interested to examine similarity measure create from the overlap functions.

Definition 4. $G_0 : [0, 1]^2 \to [0, 1]$ is an overlap function if
(GO1) $G_O(x, y) = G_O(y, x)$ for all $x; y \in [0, 1]$;
(GO2) $G_O(x, y) = 0$ if and only if $x = 0$ or $y = 0$;
(GO3) $G_O(x, y) = 1$ if and only if $x = y = 1$;
(GO4) G_O is increasing;
(GO5) G_O is continuous.

For our further considerations the most important will be the condition (GO3) which justifies the choice of the overlap function to create a similarity measure. Many researches tried to relax the rigidity of Zadeh definition (1) of inclusion to get a soft approach which is more compatible with the spirit of fuzzy logic. Instead of binary discrimination: being or not being a subset, they proposed several indicators giving the degree to which an interval-valued intuitionistic fuzzy set is a subset of another interval-valued intuitionistic fuzzy set. More precisely, let us consider a mapping $Inc : IVIFS(X) \times IVIFS(X) \to [0,1]$, called an inclusion measure (or subsethood measure) such that the value $Inc(\rho,\sigma)$ quantifies the degree of inclusion of ρ into σ.

Definition 5 (cf. [27]). Let $\rho,\sigma \in IVIFS(X)$. We denote $Inc : IVIFS(X) \times IVIFS(X) \to [0,1]$ as inclusion measure if it satisfies the following conditions:
(IM1) If $\rho = 1^{\cdot}$, $\sigma = 0^{\cdot}$, then $Inc(\rho,\sigma) = 0$;
(IM2) $Inc(\rho,\sigma) = 1 \Leftrightarrow \rho \le \sigma$;
(IM3) If $\rho \le \sigma \le \gamma$, then $Inc(\gamma,\rho) \le Inc(\sigma,\rho)$ and $Inc(\gamma,\rho) \le Inc(\gamma,\sigma)$.

Definition 6 (cf. [26]). Let $\rho,\sigma \in IVIFS(X)$. We denote $Inc^H : IVIFS(X) \times IVIFS(X) \to [0,1]$ as hybrid monotonic inclusion measure if it satisfies the following conditions:
(HIM1) If $\rho = 1^{\cdot}$, $\sigma = 0^{\cdot}$, then $Inc(\rho,\sigma) = 0$;
(HIM2) $Inc(\rho,\sigma) = 1 \Leftrightarrow \rho \le \sigma$;
(HIM3) If $\rho \le \sigma$, then $Inc(\sigma,\gamma) \le Inc(\rho,\gamma)$ and $Inc(\gamma,\rho) \le Inc(\gamma,\sigma)$.

Here are some examples of inclusion measures:

(i)

$$Inc_{\wedge}(\rho,\sigma) = \begin{cases} 1, & \text{if } \rho(x) = \sigma(x) = 0^{\cdot}, \\ & \forall_x \in X \\ \frac{|\rho\wedge\sigma|}{|\rho|}, & \text{otherwise} \end{cases},$$

where $|\rho| = \sum_{x_i \in X} \frac{\underline{R}+\overline{R}+2-\underline{\mathfrak{R}}-\overline{\mathfrak{R}}}{4}$.

(ii) The inclusion description via inclusion indicator comprises formally Zadehs definition (1), because we may define

$$Inc_z(\rho,\sigma) = \begin{cases} 1, & \text{if } \rho(x) \le \sigma(x), \forall_x \in X \\ 0, & \text{otherwise} \end{cases}.$$

Now, we examine the transitivity property using the overlap functions.

Proposition 1 ([17]). *Inc_z is a partially ordered relation with $\wedge$-transitivity.*

Proposition 2. *Inc_z is a partially ordered operation with G_0-transitivity, where G_0 is an overlap function.*

Proof. Reflexivity is obvious. We consider antisymmetry. We examine the following implication:

$$Inc_z(\rho,\sigma) = Inc_z(\sigma,\rho) \Rightarrow \rho = \sigma.$$

1. If $\rho \le \sigma$ and $Inc_z(\rho,\sigma) = Inc_z(\sigma,\rho)$, then $Inc_z(\rho,\sigma) = 1 = Inc_z(\sigma,\rho)$, so $\sigma \le \rho$. Thus $\rho = \sigma$.
2. If $\rho \ge \sigma$ and $Inc_z(\rho,\sigma) = Inc_z(\sigma,\rho)$, then $Inc_z(\rho,\sigma) = 0 = Inc_z(\sigma,\rho)$, so $\sigma \ge \rho$ and we obtain a contradiction. Thus the considered implication is true.

Now, we consider G_0-transitivity, i.e.

$$G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) \le Inc_z(\rho,\gamma).$$

1. If $\rho \le \sigma \le \gamma$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(1,1) = 1 \le 1 = Inc_z(\rho,\gamma)$.
2. If $\rho \le \gamma \le \sigma$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(1,0) = 0 \le 1 = Inc_z(\rho,\gamma)$.
3. If $\sigma \le \rho \le \gamma$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(0,1) = 0 \le 1 = Inc_z(\rho,\gamma)$.
4. If $\sigma \le \gamma \le \rho$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(0,1) = 0 \le 0 = Inc_z(\rho,\gamma)$.
5. If $\gamma \le \rho \le \sigma$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(1,0) = 0 \le 0 = Inc_z(\rho,\gamma)$.
6.If $\gamma \le \sigma \le \rho$, then $G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\gamma)) = G_0(0,1) = 0 \le 0 = Inc_z(\rho,\gamma)$.

So Inc_z is a partially ordered operation. □

Remark 1. If $G_0 \le \wedge$, then Inc^H has G_0-transitive property, where G_0 is an overlap function.

4 SIMILARITY MEASURE

The similarity measure is employed to indicate the similarity degrees of two models or two rules in a system. In this section, we first recall definition of the similarity measure.

Definition 7 ([27]). A real function $Sim : IVIFS(X) \times IVIFS(X) \to [0,1]$ is named a similarity measure of IVIFSs on universe X, if it satisfies the following properties:
(SM1) $Sim(\rho,\rho') = 0$, if $\rho \in P(X)$;
(SM2) $Sim(\rho,\sigma) = 1 \Leftrightarrow \rho = \sigma$;
(SM3) $Sim(\rho,\sigma) = Sim(\sigma,\rho)$;
(SM4) If $\rho \le \sigma \le \gamma$, then $Sim(\rho,\gamma) \le Sim(\rho,\sigma)$ and $Sim(\rho,\gamma) \le Sim(\sigma,\gamma)$.

Here are some examples of inclusion measures:

(i) $Sim_d(\rho,\sigma) =$
$1 - \frac{1}{n}\sum_{i=1}^{n} \max(|\underline{R}(x_i) - \underline{S}(x_i)|,$
$|\overline{R}(x_i) - \overline{S}(x_i)|, |\underline{\mathfrak{R}}(x_i) - \underline{\mathfrak{S}}(x_i)|, |\overline{\mathfrak{R}}(x_i) - \overline{\mathfrak{S}}(x_i)|),$

(ii) $Sim_{\wedge}(\rho,\sigma) = Inc(\rho,\sigma) \wedge Inc(\sigma,\rho)$.

We observe an interesting connection between the similarity and the inclusion measure.

Proposition 3 ([27]). *$Inc(\rho,\sigma) = Sim(\sigma, \rho\vee\sigma)$ is an inclusion measure.*

Moreover, we present some similarity measures

Proposition 4. *Let G_0 be an overlap function.*

$$Sim_{G_0}(\rho,\sigma) = G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\rho))$$

is a similarity measure.

Proof. The proof of (SM1) is obvious. By (GO3) we have

$$\begin{aligned} Sim_{G_0}(\rho,\sigma) = 1 &\Leftrightarrow G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\rho)) = 1 \\ &\Leftrightarrow Inc_z(\rho,\sigma) = 1 \end{aligned}$$

and

$$Inc_z(\sigma,\rho) = 1 \Leftrightarrow \rho \le \sigma,\ \sigma \le \rho.$$

Thus $\rho = \sigma$, which proves (SM2).
Now, we consider (SM3). Let $\rho \le \sigma \le \gamma$. Then by properties of Inc_z and (GO2) we obtain

$$\begin{aligned} Sim_{G_0}(\rho,\gamma) &= G_0(Inc_z(\rho,\gamma), Inc_z(\gamma,\rho)) \\ &= G_0(1,0) = 0. \end{aligned}$$

In a similar way we obtain $Sim_{G_0}(\rho,\sigma) = 0$ and $Sim_{G_0}(\sigma,\gamma) = 0$. Thus (SM3) holds. □

Proposition 5. *If G_0 is an overlap function with the neutral element 1, then*

$$Sim_{G_0}(\rho,\sigma) = G_0(Inc^H(\rho,\sigma), Inc^H(\sigma,\rho))$$

is a similarity measure.

Proof. (SM1) and (SM2) hold. We consider (SM3). Let $\rho \le \sigma \le \gamma$ and 1 be the neutral element of G_0. Then

$$\begin{aligned} Sim_{G_0}(\rho,\gamma) &= G_0(Inc^H(\rho,\gamma), Inc^H(\gamma,\rho)) \\ &= G_0(1, Inc^H(\gamma,\rho)) = Inc^H(\gamma,\rho). \end{aligned}$$

In a similar way we obtain

$$\begin{aligned} Sim_{G_0}(\rho,\sigma) &= Inc^H(\sigma,\rho), \\ Sim_{G_0}(\sigma,\gamma) &= Inc^H(\gamma,\sigma). \end{aligned}$$

Now, by (HIM3) we have $Inc^H(\gamma,\rho) \le Inc^H(\sigma,\rho)$ and $Inc^H(\gamma,\rho) \le Inc^H(\gamma,\sigma)$. So $Sim_{G_0}(\rho,\gamma) \le Sim_{G_0}(\rho,\sigma)$ and $Sim_{G_0}(\rho,\gamma) \le Sim_{G_0}(\gamma,\sigma)$. Thus we finished the proof. □

Moreover, we consider, when the similarity measure is a G_0-equivalence relation.

Proposition 6. *Let G_0 be a bisymmetric overlap function. Then*

$$Sim_{G_0}(\rho,\sigma) = G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\rho))$$

is the G_0-equivalence relation.

Proof. Let G_0 be a bisymmetric overlap function. If $\rho = \sigma$, then

$$\begin{aligned} Sim_{G_0}(\rho,\sigma) &= G_0(Inc_z(\rho,\rho), Inc_z(\rho,\rho)) \\ &= G_0(1,1) = 1. \end{aligned}$$

So Sim_{G_0} has the reflexivity property.
Now, we consider the symmetry property. By symmetry of G_0 we obtain

$$\begin{aligned} Sim_{G_0}(\rho,\sigma) &= G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\rho)) \\ &= G_0(Inc_z(\sigma,\rho), Inc_z(\rho,\sigma)) \\ &= Sim_{G_0}(\sigma,\rho). \end{aligned}$$

If G_0 has bisymmetry property, then we can prove G_0-transitivity:

$$G_0(Sim_{G_0}(\rho,\sigma), Sim_{G_0}(\sigma,\delta)) \le Sim_{G_0}(\rho,\delta).$$

By G_0-transitivity of Inc_z we have

$$G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\delta)) \le Inc_z(\rho,\delta)$$

and

$$G_0(Inc_z(\delta,\sigma), Inc_z(\sigma,\rho)) \le Inc_z(\delta,\rho)$$

Then by non-decreasingness, symmetry and bisymmetry property of G_0 we calculate
$G_0(G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\delta)), G_0(Inc_z(\delta,\sigma), Inc_z(\sigma,\rho)))$
$\le G_0(Inc_z(\rho,\delta), Inc_z(\delta,\rho))$

$$\Leftrightarrow$$

$G_0(G_0(Inc_z(\rho,\sigma), Inc_z(\sigma,\rho)), G_0(Inc_z(\sigma,\delta), Inc_z(\delta,\sigma)))$
$\le G_0(Inc_z(\rho,\delta), Inc_z(\delta,\rho))$,
which finishes the proof. □

Remark 2. Let G_0 be a bisymmetric overlap function, $G_0 \le \wedge$. Then

$$Sim_{G_0}(\rho,\sigma) = G_0(Inc^H(\rho,\sigma), Inc^H(\sigma,\rho))$$

is the G_0-equivalence relation.

5 CONCLUSIONS

In this paper we comment on the existing axiomatical definitions of similarity measures for IVIFSs. In Definitions 4-6 we follow the literature but in future work we would like to consider other order relations ($\le$ - more compatible with the semantic of the relation of inclusion and similarity) and other interval structures. Some general formula calculating the similarity between IVIFSs have been proposed. The relationships among the similarity measures and the inclusion measures of IVIFSs and G_0-transitivity have been investigated. In future, we would like to examine some modifies of Inc_Z by using an admissible order defined by Bustince [8], where the order is constructed by aggregation functions, and finally we would like to propose definitions of similarity with the above inclusion measure. Furthermore, another transitivity properties and their connection with similarity will be examined. Moreover, dependence between similarity and preference property will be investigated.

Acknowledgements

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge in Rzeszów, through Project Number RPPK.01.03.00-18-001/10.

References

[1] K.T. Atanassov (1986). Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **20**, pp. 87 - 96.

[2] K.T. Atanassov (1994). Operators over interval valued intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **64**, pp. 159 - 174.

[3] K. Atanassov (1999). Intuitionistic Fuzzy Sets: Theory and Applications. *Physica-Verlag.*

[4] U. Bodenhofer, B. De Baets, J. Fodor (2007). A compendium of fuzzy weak orders:Representations and constructions. *Fuzzy Sets Syst.* **158**, pp. 811 - 829.

[5] K. Bosteels, E.E. Kerre (2009). On a reflexivity-preserving family of cardinality-based fuzzy comparison measures. *Inform. Sci.* **179**, pp. 2342 - 2352.

[6] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna (2010). Overlap functions. *Nonlinear Anal.: Theory Methods Appl.* **72**, pp. 1488 - 1499.

[7] H. Bustince, M. Pagola, R. Mesiar, E. Hüllermeier, F. Herrera, Grouping, overlaps, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. *IEEE Trans. Fuzzy Syst.* **20 (3)** (2012) 405 - 415.

[8] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar (2013). Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets Syst.* **220**, pp. 69 - 77.

[9] T. Calvo, A. Kolesárová, M. Komornková, and R. Mesiar (2002). Aggregation operators: properties, classes and construction methods. In T. Calvo, G. Mayor, and R. Mesiar, editors, Aggregation Operators **97** of Studies in Fuzziness and Soft Computing, pp. 3 - 104. *Physica-Verlag HD.*

[10] B. De Baets, R. Mesiar (1999). Triangular norms on product lattices. *Fuzzy Sets Syst.* **104**, pp. 61–76.

[11] B. De Baets, H. De Meyer, H. Naessens (2001). A class of rational cardinality-based similarity measures. *J. Comp. Appl. Math.* **132**, pp. 51 - 69.

[12] B. De Baets, H. De Meyer (2005). Transitivity frameworks for reciprocal relations:cycle-transitivity versus FG-transitivity. *Fuzzy Sets Syst.* **152**, pp. 249 - 270.

[13] B. De Baets, S. Janssens, H. De Meyer (2009). On the transitivity of a parametric family of cardinality-based similarity measures. *Int. J. Appr. Reason.* **50**, pp. 104 - 116.

[14] M. De Cock, E.E. Kerre (2003). Why fuzzy T-equivalence relations do not resolve the Poincar'e paradox, and related issues. *Fuzzy Sets Syst.* **133**, pp. 181 - 192.

[15] S. Freson, B. De Baets, H. De Meyer (2014). Closing reciprocal relations w.r.t. stochastic transitivity. *Fuzzy Sets Syst.* **241**, pp. 2 - 26.

[16] B. Jayaram, R. Mesiar (2009). I-Fuzzy equivalence relations and I-fuzzy partitions. *Information Sci.* **179**, pp. 1278 - 1297.

[17] A. Kehagias, M. Konstantinidou (2003). L-fuzzy valued inclusion measure, L-fuzzy similarity and L-fuzzy distance. *Fuzzy Sets Syst.* **136**, pp. 313 - 332.

[18] D.F. Li (2010). Toposis-based nonlinear-programming methodology for multiattribute decision making with interval-valued intuitionistic fuzzy sets. *IEEE Trans. Fuzzy Syst.* **18**, pp. 299 - 311.

[19] X.D. Liu, S.H. Zheng, F.L. Xiong (2005). Entropy and subsethood for general interval-valued intuitionistic fuzzy sets. *Lecture Notes Artif. Intell.* **3613**, pp. 42 - 52.

[20] N. Madrid, A. Burusco, H. Bustince, J. Fernandez, I. Perfilieva (2015). Upper bounding overlaps by groupings. *Fuzzy Sets Syst.* **264**, pp. 76 - 99.

[21] S. Ovchinnikov (2002). Numerical representation of transitive fuzzy relations. *Fuzzy Sets Syst.* 126, pp. 225 - 232.

[22] D.G. Park, Y.C. Kwun, J.H. Park, I.Y. Park (2009). Correlation coefficient of interval-valued intuitionistic fuzzy sets and its application to multiple attribute group decision making problems. *Math. Comput. Modell.* **50**, pp. 1279 - 1293.

[23] Z. Świtalski (2003). General transitivity conditions for fuzzy reciprocal preference matrices. *Fuzzy Sets Syst.* 137, pp. 85 - 100.

[24] L.A. Zadeh (1965). Fuzzy sets. *Information Contr.* **8**, pp. 338 - 353.

[25] W.Y. Zeng, P. Guo (2008). Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship. *Inform. Sci.* **178**, pp. 1334 - 1342.

[26] H.-Y. Zhang, W.-X. Zhang (2009). Hybrid monotonic inclusion measure and its use in measuring similarity and distance between fuzzy sets. *Fuzzy Sets Syst.* **160**, pp. 107 - 118.

[27] Q. Zhang, H. Xing, F. Liu, J. Ye, P. Tang (2014). Some new entropy measures for interval-valued intuitionistic fuzzy sets based on distances and their relationships with similarity and inclusion measures. *Inform. Sci.* **283**, pp. 55 - 69.

ANALYZING SUBDISTRIBUTIVITY AND SUPERDISTRIBUTIVITY ON OVERLAP AND GROUPING FUNCTIONS

H. Santos[a], **L. Lima**[b], **B. Bedregal**[a]
Universidade Federal do Rio Grande do Norte, Brazil
[a]{helida,bedregal}@dimap.ufrn.br
[b]lucelialimastm@gmail.com

G. P. Dimuro
Universidade Federal do Rio Grande, Brazil
gracaliz@furg.br

M. Rocha
Universidade Federal do Pará, Brazil
marcus.rocha60@gmail.com

H. Bustince
Universidad Pública de Navarra, Spain
bustince@unavarra.es

Summary

Overlap and grouping functions have been used in different fields and for various purposes, namely for instance in image processing, classification problems and also in decision making. Thus, motivated by practical and fundamental reasons, we propose in this paper to apply the distributive property on overlap and grouping functions, which are dual to each other, and to analyze their behavior on the unit interval $[0, 1]$.

Keywords: Fuzzy connectives, Distributivity, Aggregations, Overlaps, Groupings.

1 INTRODUCTION

Constructing the membership degree is one of the main challenges in the context of Fuzzy Set Theory since its introduction by Zadeh in 1965 [24]. Moreover, many generalizations of the logical connectives have been proposed and their properties have also been discussed and analyzed. The distributive property, for instance, has been studied for many years and has a lot of aspects considered diversely [1, 3, 4]. Some of the approaches which examine the aforementioned property can be found within triangular norms and conorms [2, 10], aggregation functions and quasi-arithmetic means [14], fuzzy implications [5], etc.

Overlap functions were presented in [12] in order to deal with the difficulty of assigning a membership degree to an element belonging to more than one class simultaneously in the context of image processing. Thereafter, grouping functions were introduced in [13] as the dual notion of overlaps. In this sense, our main goal in this work is to provide an initial analysis of the behavior of these functions concerning the distributivity law as well as the subdistributivity and superdistributivity inequalities. Thus we take some examples of overlap and grouping functions, which are dual to each other, and then we analyze their behavior on the unit interval $[0, 1]$.

The structure of this paper is given as follows. In the subsequent section we recall some preliminary concepts which will be used throughout the paper, then in the third section we present the formal definitions of overlap and grouping functions providing some examples. Sections 4 and 5 contain our main contributions, that is, the results of the analysis proposed and a study on the class of distributive overlap and grouping functions. Finally, we present our conclusions, future works and references.

2 PRELIMINARIES

In this section we review some concepts which will be used throughout this paper.

2.1 AGGREGATION FUNCTIONS

Definition 2.1 *Let $m \in \mathbb{N}$ such that $m \geq 2$. A function $\mathcal{A} : [0, 1]^m \to [0, 1]$ is a m-ary aggregation operator,*

i. If $x_i \leq y_i$ for each $i = 1, \ldots, m$, then $\mathcal{A}(x_1, \ldots, x_m) \leq \mathcal{A}(y_1, \ldots, y_m)$, for each $x_1, \ldots, x_m, y_1, \ldots, y_m \in [0, 1]$;

ii. $\mathcal{A}(0, \ldots, 0) = 0$;

iii. $\mathcal{A}(1, \ldots, 1) = 1$.

2.2 DISTRIBUTIVITY

Definition 2.2 *Let $\mathcal{A}, \mathcal{B}$ be binary aggregation operators. We say $\mathcal{A}$ distributes over $\mathcal{B}$ if both of the following laws hold:*

ISBN: 978–83–8012–519–3

(LD) $\mathcal{A}$ *is left distributive over* $\mathcal{B}$, *that is* $\mathcal{A}(x, \mathcal{B}(y, z)) = \mathcal{B}(\mathcal{A}(x, y), \mathcal{A}(x, z))$ *for all* $x, y, z \in [0, 1]$.

(RD) $\mathcal{A}$ *is right distributive over* $\mathcal{B}$, *that is* $\mathcal{A}(\mathcal{B}(y, z), x) = \mathcal{B}(\mathcal{A}(y, x), \mathcal{A}(z, x))$ *for all* $x, y, z \in [0, 1]$.

Definition 2.3 *Let* $\mathcal{A}, \mathcal{B}$ *be binary aggregation operators:*

- $\mathcal{A}$ *left subdistributes over* $\mathcal{B}$ *if* $\mathcal{A}(x, \mathcal{B}(y, z)) \leq \mathcal{B}(\mathcal{A}(x, y), \mathcal{A}(x, z))$ *for all* $x, y, z \in [0, 1]$.
- $\mathcal{A}$ *left superdistributes over* $\mathcal{B}$ *if* $\mathcal{A}(x, \mathcal{B}(y, z)) \geq \mathcal{B}(\mathcal{A}(x, y), \mathcal{A}(x, z))$ *for all* $x, y, z \in [0, 1]$.

Observe that the right subdistributivity and right superdistributivity can be defined analogously. Notice that whenever an aggregation function is commutative, then the concept of left and right (sub,super)distributivity coincide.

2.3 AUTOMORPHISMS AND NEGATIONS

Automorphisms play an important role in fuzzy logic and its extensions as they allow a simple characterization of some classes of operators (for instance, the ones proposed in [6, 8, 9, 23]).

Definition 2.4 *[20, 22] A mapping* $\varphi : [0, 1] \to [0, 1]$, *is an automorphism if it is bijective and monotonic, i.e.* $x \leq y \Rightarrow \varphi(x) \leq \varphi(y)$.

An equivalent definition was given in [11]:

Definition 2.5 *Automorphisms are continuous and strictly increasing functions* $\varphi : [0, 1] \to [0, 1]$ *such that* $\varphi(0) = 0$ *and* $\varphi(1) = 1$.

Automorphisms are closed under composition, i.e., taking the set $Aut([0, 1])$ as the set of all automorphisms on $[0, 1]$, if $\varphi_1, \varphi_2 \in Aut([0, 1])$ then $\varphi_1 \circ \varphi_2(x) = \varphi_1(\varphi_2(x)) \in Aut([0, 1])$. Besides, the inverse φ^{-1} of an automorphism φ is also an automorphism.

Definition 2.6 *[7] The action of an automorphism* φ *on a function* $f : [0, 1]^n \to [0, 1]$ *and denoted by* f^φ *is called the conjugate of* f *and is defined as:*

$$f^\varphi(x_1, \ldots, x_n) = \varphi^{-1}(f(\varphi(x_1), \ldots, \varphi(x_n))).$$

A strict negation ([23]) is a continuous and strictly decreasing mapping, i.e., $N : [0, 1] \to [0, 1]$, such that $N(0) = 1$ and $N(1) = 0$. A strict negation N is strong if $N(N(x)) = x$, for all $x \in [0, 1]$.

3 OVERLAP AND GROUPING FUNCTIONS

The idea of overlap functions introduced in [12] emerged in the context of image processing and the inherent difficult of classifying pixels of an image containing an object on a background. In some situations, it is a hard task to decide whether a pixel belongs to the class of the object (background) either because the separation between the classes is unclear or due to an overlap between classes. Thus, overlap functions provide a way to measure to what extent a pixel belongs to both (object and background) classes. In the same way, grouping functions were introduced in [13] as the dual form of overlap functions and they were used to measure to what extent a pixel belongs to at least one of the classes.

The formal definitions of overlap and grouping are given as follows.

Definition 3.1 *An operator* $\mathcal{O} : [0, 1]^2 \to [0, 1]$ *is an overlap if it is: commutative, increasing, continuous, and satisfies the boundary conditions:*

- $\mathcal{O}(x, y) = 0$ *if and only if* $x = 0$ *or* $y = 0$, *and*
- $\mathcal{O}(x, y) = 1$ *if and only if* $x = y = 1$.

Definition 3.2 *An operator* $\mathcal{G} : [0, 1]^2 \to [0, 1]$ *is a grouping if it is: commutative, increasing, continuous, and satisfies the boundary conditions:*

- $\mathcal{G}(x, y) = 0$ *if and only if* $x = y = 0$, *and*
- $\mathcal{G}(x, y) = 1$ *if and only if* $x = 1$ *or* $y = 1$.

Remark 1 *[21] Note that by the boundary conditions given in Definitions 3.1 and 3.2, a grouping can never be less than an overlap and not always for all* $x, y \in [0, 1]$ *the following inequality holds:* $\mathcal{O}(x, y) \leq \mathcal{G}(x, y)$.

Bustince et al. (in [13]) demonstrated the relation between grouping and overlap functions through the following theorem.

Given a negation N, and a function $\mathcal{F} : [0, 1]^2 \to [0, 1]$. The N-dual of $\mathcal{F}$ is the function $\mathcal{F}_N : [0, 1]^2 \to [0, 1]$ defined by

$$\mathcal{F}_N(x, y) = N(\mathcal{F}(N(x), N(y))).$$

Theorem 3.3 *[13] Let* $\mathcal{O}$ *be an overlap function,* $\mathcal{G}$ *be a grouping function and* N *be a strict negation. Then it holds:*

1. $\mathcal{O}_N(x, y) = N(\mathcal{G}(N(x), N(y)))$ *is an overlap function;*

2. $\mathcal{G}_N(x,y) = N(\mathcal{O}(N(x),N(y)))$ *is a grouping function.*

Now, we provide some examples of overlap and grouping functions (dual to each other) which were given by Dimuro et al. in [15, 16, 17, 18].

Example 3.4 *Some examples of overlap functions:*

i.
$$\mathcal{O}_{mM}(x,y) = \min\{x,y\}\max\{x^2,y^2\}.$$

ii.
$$\mathcal{O}_{DB}(x,y) = \begin{cases} \frac{2xy}{x+y} & \text{if } x+y \neq 0, \\ 0 & \text{if } x+y = 0. \end{cases}$$

iii.
$$\mathcal{O}_P(x,y) = x^p y^p, \text{ with } p > 0.$$

iv.
$$\mathcal{O}_V(x,y) = \begin{cases} \frac{1+(2x-1)^2(2y-1)^2}{2} & \text{if } x,y \in [0.5,1], \\ \min\{x,y\} & \text{otherwise.} \end{cases}$$

v.
$$\mathcal{O}_{min}(x,y) = \min\{x,y\}.$$

Note that $\mathcal{O}_{mM}, \mathcal{O}_{DB}$ and $\mathcal{O}_V$ are examples of overlaps that are not t-norms. $\mathcal{O}_P$ is not a t-norm if $p > 1$ and if $p = 1$, then we have the product t-norm which is an overlap. The last example ($\mathcal{O}_{min}$) is actually the minimum t-norm.

Example 3.5 *Some examples of grouping functions:*

vi.
$$\mathcal{G}_{mM}(x,y) = 1-\min\{1-x,1-y\}\max\{(1-x)^2,(1-y)^2\}.$$

vii.
$$\mathcal{G}_{DB}(x,y) = \begin{cases} \frac{x+y-2xy}{2-(x+y)} & \text{if } x \neq 1 \text{ or } y \neq 1, \\ 1 & \text{if } x = y = 1. \end{cases}$$

viii.
$$\mathcal{G}_P(x,y) = 1-(1-x)^p(1-y)^p, \text{ with } p > 0.$$

ix.
$$\mathcal{G}^V_{mM}(x,y) = \begin{cases} \frac{\mathcal{G}_{mM}(2x,2y)}{2} & \text{if } x,y \in [0,0.5[, \\ \max\{x,y\} & \text{otherwise.} \end{cases}$$

x.
$$\mathcal{G}_{max}(x,y) = \max\{x,y\}.$$

We can also note that $\mathcal{G}_{mM}, \mathcal{G}_{DB}, \mathcal{G}_P$ and $\mathcal{G}^V_{mM}$ are examples of groupings that are not t-conorms, and the last one ($\mathcal{G}_{max}$) is the maximum t-conorm.

4 ANALYSIS OF THE DISTRIBUTIVITY OF OVERLAP/GROUPING FUNCTIONS

As previously mentioned, our main goal is to analyze the behavior of overlap and grouping functions regarding the distributivity law. However, we should firstly recall two important results, items (a) and (b) given below, presented by Fodor and Roubens in [19] concerning the distributivity law.

Let T be a t-norm, T_{min} be the minimum t-norm, S be a t-conorm and S_{max} be the maximum t-conorm, then for all $x,y,z \in [0,1]$, we have:

(a) $S(x,T(y,z)) = T(S(x,y),S(x,z))$ holds if and only if $T = T_{min}$;

(b) $T(x,S(y,z)) = S(T(x,y),T(x,z))$ holds if and only if $S = S_{max}$.

So, in our analysis, we take the overlaps and groupings presented on examples 3.4 and 3.5, respectively, and study the distributivity property on each possible combination among the ten examples. Or, in other words, we compared $\mathcal{A}(x,\mathcal{B}(y,z))$ and $\mathcal{B}(\mathcal{A}(x,y),\mathcal{A}(x,z))$ and organized the results in Table 1, where $=$ means both functions distribute over one another, such that for all $x,y,z \in [0,1]$, (1) $\mathcal{A}(x,\mathcal{B}(y,z)) = \mathcal{B}(\mathcal{A}(x,y),\mathcal{A}(x,z))$. Analogously, $\leq$ means that $\mathcal{A}$ subdistributes over $\mathcal{B}$, such that (2) $\mathcal{A}(x,\mathcal{B}(y,z)) \leq \mathcal{B}(\mathcal{A}(x,y),\mathcal{A}(x,z))$, for all $x,y,z \in [0,1]$; $\geq$ means that $\mathcal{A}$ superdistributes over $\mathcal{B}$, such that (3) $\mathcal{A}(x,\mathcal{B}(y,z)) \geq \mathcal{B}(\mathcal{A}(x,y),\mathcal{A}(x,z))$ for all $x,y,z \in [0,1]$; and, finally, $\|$ means that functions $\mathcal{A}$ and $\mathcal{B}$ are not comparable, i.e. (4) $\mathcal{A}(x,\mathcal{B}(y,z)) \parallel \mathcal{B}(\mathcal{A}(x,y),\mathcal{A}(x,z))$ for all $x,y,z \in [0,1]$. We state that, in some cases, functions $\mathcal{A}$ and $\mathcal{B}$ cannot be compared as they have different behavior concerning the distributivity, in the sense that either $\mathcal{A}$ may subdistribute over $\mathcal{B}$ or $\mathcal{A}$ may superdistribute over $\mathcal{B}$ depending on the values chosen for $x,y,z \in [0,1]$ (see Example 4.4).

In order to understand more clearly the results shown in Table 1, we will discuss some results and provide examples as follows. On the first row, for instance, it is clear that $\mathcal{G}_{mM}$ subdistributes over $\mathcal{G}_{mM}, \mathcal{G}_P, \mathcal{G}^V_{mM}$ and $\mathcal{O}_{DB}$; distributes over $\mathcal{G}_{max}$ and $\mathcal{O}_{min}$; and cannot be compared with $\mathcal{G}_{DB}, \mathcal{O}_{mM}, \mathcal{O}_P$ and $\mathcal{O}_V$. In the same way, on the second row we see that $\mathcal{G}_{DB}$ just subdistributes over $\mathcal{O}_{DB}$; superdistributes over $\mathcal{O}_{mM}$ and $\mathcal{O}_P$; distributes over $\mathcal{G}_{max}$ and $\mathcal{O}_{min}$; and cannot be compared with $\mathcal{G}_{mM}, \mathcal{G}_{DB}, \mathcal{G}_P, \mathcal{G}^V_{mM}$ and $\mathcal{O}_V$. Finally, some numerical examples (4.1, 4.2, 4.3 and 4.4) are given below. Notice that despite showing the examples for specific points in $[0,1]$, the computation of

Table 1: Distributivity on overlap and grouping functions

$\mathcal{A}/\mathcal{B}$	$\mathcal{G}_{mM}$	$\mathcal{G}_{DB}$	$\mathcal{G}_P$	$\mathcal{G}^V_{mM}$	$\mathcal{G}_{max}$	$\mathcal{O}_{mM}$	$\mathcal{O}_{DB}$	$\mathcal{O}_P$	$\mathcal{O}_V$	$\mathcal{O}_{min}$
$\mathcal{G}_{mM}$	$\leq$	$\|$	$\leq$	$\leq$	$=$	$\|$	$\leq$	$\|$	$\|$	$=$
$\mathcal{G}_{DB}$	$\|$	$\|$	$\|$	$\|$	$=$	$\geq$	$\leq$	$\geq$	$\|$	$=$
$\mathcal{G}_P$	$\leq$	$\leq$	$\leq$	$\|$	$=$	$\|$	$\|$	$\|$	$\|$	$=$
$\mathcal{G}^V_{mM}$	$\|$	$\|$	$\|$	$\|$	$=$	$\geq$	$\|$	$\|$	$\|$	$=$
$\mathcal{G}_{max}$	$\leq$	$\leq$	$\leq$	$\|$	$=$	$\geq$	$\leq$	$\geq$	$\geq$	$=$
$\mathcal{O}_{mM}$	$\|$	$\|$	$\|$	$\|$	$=$	$\geq$	$\|$	$\geq$	$\leq$	$=$
$\mathcal{O}_{DB}$	$\leq$	$\geq$	$\leq$	$\|$	$=$	$\|$	$=$	$\|$	$\|$	$=$
$\mathcal{O}_P$	$\|$	$\|$	$\|$	$\|$	$=$	$\geq$	$\geq$	$\geq$	$\leq$	$=$
$\mathcal{O}_V$	$\leq$	$\geq$	$\|$	$\leq$	$=$	$\leq$	$\|$	$\geq$	$\geq$	$=$
$\mathcal{O}_{min}$	$\leq$	$\geq$	$\leq$	$\leq$	$=$	$\geq$	$\geq$	$\geq$	$\geq$	$=$

all values in the range between 0 and 1 was done to check the distributivity among the functions.

Example 4.1 *We see in Table 1 that $\mathcal{O}_P$ distributes over $\mathcal{G}_{max}$, i.e.* (1). *Let us take, for instance, $x = 0.1$, $y = 0.5$, $z = 0.6$ and $p = 2$. We have:*

$$\mathcal{O}_P(0.1, \mathcal{G}_{max}(0.5, 0.6)) = 0.0036 = \mathcal{G}_{max}(\mathcal{O}_P(0.1, 0.5), \mathcal{O}_P(0.1, 0.6)).$$

Example 4.2 *We can also observe in Table 1 that $\mathcal{G}_{mM}$ subdistributes over $\mathcal{O}_{DB}$, i.e.* (2). *Let us take $x = 0.7$, $y = 0.2$ and $z = 1$. We have:*

$$\mathcal{G}_{mM}(x, \mathcal{O}_{DB}(y, z)) \cong 0.867 \leq \mathcal{O}_{DB}(\mathcal{G}_{mM}(x, y), \mathcal{G}_{mM}(x, z)) \cong 0.894.$$

Example 4.3 *In Table 1, note that $\mathcal{O}_{mM}$ superdistributes over $\mathcal{O}_{mM}$, i.e.* (3). *Let us take $x = 0.2$, $y = 0.5$ and $z = 1$. We have:*

$$\mathcal{O}_{mM}(x, \mathcal{O}_{mM}(y, z)) = 0.05 \geq \mathcal{O}_{mM}(\mathcal{O}_{mM}(x, y), \mathcal{O}_{mM}(x, z)) = 0.002.$$

As previously mentioned, some of the functions cannot be compared regarding the distributivity. This occurs because some of them can both subdistribute and superdistribute for different values of x, y, z taken on the interval $[0, 1]$, as pointed out in the following example.

Example 4.4 *We can say $\mathcal{G}_{mM}$ cannot be compared with $\mathcal{O}_P$, i.e.* (4). *Take, for instance $x_1 = 0.5$, $y_1 = 0.2$, $z_1 = 0.9$ and $p = 2$, we have:*

$$\mathcal{G}_{mM}(x, \mathcal{O}_P(y, z)) \cong 0.5318 \geq \mathcal{O}_P(\mathcal{G}_{mM}(x, y), \mathcal{G}_{mM}(x, z)) \cong 0.4395.$$

However, if we take $x_2 = 0.8$, $y_2 = 0.6$, $z_2 = 0.8$ and $p = 2$, we have:

$$\mathcal{G}_{mM}(x, \mathcal{O}_P(y, z)) \cong 0.8815 \leq \mathcal{O}_P(\mathcal{G}_{mM}(x, y), \mathcal{G}_{mM}(x, z)) \cong 0.9221.$$

5 A STUDY ON THE CLASS OF DISTRIBUTIVE OVERLAP/GROUPING FUNCTIONS

In this section we can draw some conclusions from the distributive property applied on overlap and grouping functions.

As shown in Table 1, any considered aggregation function is distributive over $\mathcal{O}_{min}$ and $\mathcal{G}_{max}$. We have the following results:

Theorem 5.1 *Let $\mathcal{A} : [0, 1]^2 \to [0, 1]$ be an aggregation function. Then:*

(A1) $\mathcal{A}(x, \mathcal{O}(y, z)) = \mathcal{O}(\mathcal{A}(x, y), \mathcal{A}(x, z))$ *holds if and only if $\mathcal{O} = \mathcal{O}_{min}$;*

(A2) $\mathcal{A}(x, \mathcal{G}(y, z)) = \mathcal{G}(\mathcal{A}(x, y), \mathcal{A}(x, z))$ *holds if and only if $\mathcal{G} = \mathcal{G}_{max}$.*

Proof. It is immediate, since any aggregation function is monotonic. □

Proposition 5.2 *Let $\mathcal{G} : [0, 1]^2 \to [0, 1]$ be a grouping function with 0 as neutral element and $\mathcal{O} : [0, 1]^2 \to [0, 1]$ an overlap function. If $\mathcal{G}$ is distributive over $\mathcal{O}$, then $\mathcal{O}$ is idempotent.*

Proof. If $\mathcal{G}$ is distributive over $\mathcal{O}$, then, for all $x \in [0, 1]$ it holds that:

$$x = \mathcal{G}(x, \mathcal{O}(0, 0)) = O(\mathcal{G}(x, 0), \mathcal{G}(x, 0)) = \mathcal{O}(x, x),$$

since 0 is the neutral element of $\mathcal{G}$. □

Analogously, one can prove that:

Proposition 5.3 *Let $\mathcal{O} : [0, 1]^2 \to [0, 1]$ be an overlap function with 1 as neutral element and $\mathcal{G} : [0, 1]^2 \to [0, 1]$ a grouping function. If $\mathcal{O}$ is distributive over $\mathcal{G}$, then $\mathcal{G}$ is idempotent.*

Observe that the converse of the two previous propositions does not hold. For example, according to Table 1, $\mathcal{G}_{mM}$ is a grouping function with 0 as neutral element, which is not distributive over the idempotent overlap function $\mathcal{O}_{DB}$. Also, $\mathcal{O}_{mM}$ is an overlap function with 1 as neutral element, which is not distributive over the idempotent grouping function $\mathcal{G}_{DB}$.

Proposition 5.4 *Let $\mathcal{O} : [0,1]^2 \to [0,1]$ be an overlap function, $\mathcal{G} : [0,1]^2 \to [0,1]$ be a grouping function and $\varphi : [0,1] \to [0,1]$ be an automorphism. Then it holds:*

1. *$\mathcal{O}$ is (sub,super)distributive over $\mathcal{G}$ if and only if $\mathcal{O}^\varphi$ is (sub,super)distributive over $\mathcal{G}^\varphi$;*

2. *$\mathcal{G}$ is (sub,super)distributive over $\mathcal{O}$ if and only if $\mathcal{G}^\varphi$ is (sub,super)distributive over $\mathcal{O}^\varphi$.*

Proof. One has that:

1. If $\mathcal{O}$ is distributive over $\mathcal{G}$, then it follows that:

$$\begin{aligned}
&\mathcal{O}^\varphi(x, \mathcal{G}^\varphi(y,z)) \\
&\quad = \varphi^{-1}(\mathcal{O}(\varphi(x), \varphi(\varphi^{-1}(\mathcal{G}(\varphi(y), \varphi(z)))))) \\
&\quad = \varphi^{-1}(\mathcal{G}(\mathcal{O}(\varphi(x), \varphi(y)), \mathcal{O}(\varphi(x), \varphi(z)))) \\
&\quad = \mathcal{G}^\varphi(\mathcal{O}^\varphi(x,y), \mathcal{O}^\varphi(x,z)).
\end{aligned}$$

   On the other hand, if $\mathcal{O}^\varphi$ is distributive over $\mathcal{G}^\varphi$, then by the previous paragraph, it follows that $(\mathcal{O}^\varphi)^{\varphi^{-1}} = \mathcal{O}$ is distributive over $(\mathcal{G}^\varphi)^{\varphi^{-1}} = \mathcal{G}$.

   For the subdistributivity and superdistributivity the proofs are analogous.

2. It is analogous to item 1. □

Proposition 5.5 *Let $N : [0,1] \to [0,1]$ be a strong negation, and let $\mathcal{O} : [0,1]^2 \to [0,1]$ and $\mathcal{G} : [0,1]^2 \to [0,1]$ be an overlap function and a grouping function, respectively, defined according to Theorem 3.3. Then it holds:*

1. *$\mathcal{O}$ is (sub,super)distributive over $\mathcal{G}$ if and only if $\mathcal{O}_N$ is (sub,super)distributive over $\mathcal{G}_N$;*

2. *$\mathcal{G}$ is (sub,super)distributive over $\mathcal{O}$ if and only if $\mathcal{G}_N$ is (sub,super)distributive over $\mathcal{O}_N$.*

Proof. ($\Rightarrow$)

1. If $\mathcal{O}$ is distributive over $\mathcal{G}$, then by Theorem 3.3 we have that:

$$\begin{aligned}
&\mathcal{O}_N(x, \mathcal{G}_N(y,z)) \\
&\quad = N(\mathcal{G}(N(x), N(\mathcal{G}_N(y,z)))) \\
&\quad = N(\mathcal{G}(N(x), N(N(\mathcal{O}(N(y), N(z)))))) \\
&\quad = N(\mathcal{G}(N(x), \mathcal{O}(N(y), N(z)))) \\
&\quad = N(\mathcal{O}(\mathcal{G}(N(x), N(y)), \mathcal{G}(N(x), N(z)))) \\
&\quad = N(\mathcal{O}(N(N(\mathcal{G}(N(x), N(y)))), \\
&\qquad N(N(\mathcal{G}(N(x), N(z)))))) \\
&\quad = \mathcal{G}_N(\mathcal{O}_N(x,y), \mathcal{O}_N(x,z)).
\end{aligned}$$

   ($\Leftarrow$) Conversely, if the grouping $\mathcal{O}_N$ is distributive over the overlap $\mathcal{G}_N$, then by the previous item it follows that the overlap $(\mathcal{O}_N)_N = \mathcal{O}$ is distributive over the grouping $(\mathcal{G}_N)_N = \mathcal{G}$. Analogously, when the overlap $\mathcal{G}_N$ is distributive over the grouping $\mathcal{G}_N$, then by the previous item it follows that the grouping $(\mathcal{G}_N)_N = \mathcal{G}$ is distributive over the overlap $(\mathcal{O}_N)_N = \mathcal{O}$.

   For the subdistributivity and superdistributivity the proofs are analogous.

2. It is analogous to item 1. □

6 CONCLUSIONS

Our main goal in this work was to present an initial study about the distributivity property on overlap and grouping functions.

The open question that remain is: "Is there any idempotent overlap function $\mathcal{O}$, different from $\mathcal{O}_{\min}$, such that any grouping function $\mathcal{G}$ having 0 as neutral element distributes over $\mathcal{O}$? If not, is there an idempotent overlap function $\mathcal{O}$, different from $\mathcal{O}_{\min}$, and a grouping function $\mathcal{G}$ having 0 as neutral element, such that $\mathcal{G}$ distributes over $\mathcal{O}$?"

Similar questions may be posed for idempotent grouping functions different from $\mathcal{G}_{\max}$ and overlap functions having 1 as neutral element.

Observe that, if the above open questions have a positive answer, then it would be important to develop a characterization of distributive overlap/grouping functions, which is let for further work.

Acknowledgements

This work is supported by the Brazilian funding agencies CNPq (under the process numbers 307681/2012-2, 406503/2013-3, 481283/2013-7, 306970/2013-9, 202606/2014-7) and CAPES (under the process number 5778/2014-00), and NSF (ref. TIN-2013-40765-P, TIN2011-29520).

References

[1] J. Aczél (1966). Lectures on Functional Equations and their Applications. *Academic Press*, New York.

[2] C. Alsina (1985). On a family of connectives for fuzzy sets. *Fuzzy Sets and Systems*, **16**, pp. 231–235.

[3] C. Alsina, G. Mayor, M.S. Tomas, J. Torrens (1993). Subdistributive De Morgan triplets. *Serdica*, **19**, pp. 258–266.

[4] C. Alsina, E. Trillas (1992). On almost distributive Łukasiewicz triplets. *Fuzzy Sets and Systems*, **50**, pp. 175–178.

[5] M. Baczyński, B. Jayaram (2008). Fuzzy Implications. *Studies in Fuzziness and Soft Computing*, **231**, Springer, Berlin.

[6] B.C. Bedregal (2010). On interval fuzzy negations. *Fuzzy Sets and Systems*, **161**, pp. 2290–2313.

[7] B. Bedregal, G. Dimuro, H. Bustince, E. Barrenechea (2013). New results on overlap and grouping functions. *Information Sciences*, **249**, pp. 148–170.

[8] B.C. Bedregal, G.P. Dimuro, R.H.N. Santiago, R.H.S. Reiser (2010). On interval fuzzy S-implications. *Information Sciences*, **180**, pp. 1373–1389.

[9] B. Bedregal, R.H.N. Santiago (2013). Interval representations, Łukasiewicz implicators and Smets-Magrez axioms. *Information Sciences*, **221**, pp. 192–200.

[10] C. Bertoluzza, V. Doldi (2004). On the distributivity between t-norms and t-conorms. *Fuzzy Sets and Systems*, **142**, pp. 85–104.

[11] H. Bustince, P. Burillo, F. Soria (2003). Automorphism, negations and implication operators. *Fuzzy Sets and Systems*, **134**, pp. 209–229.

[12] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna (2010). Overlap functions. *Nonlinear Analysis: Theory, Methods & Applications*, **72(3–4A)**, pp. 1488–1499.

[13] H. Bustince, M. Pagola, R. Mesiar, E. Hüllermeier, F. Herrera (2012). Grouping, overlaps, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. *IEEE Transactions on Fuzzy Systems*, **20(3)**, pp. 405–415.

[14] T. Calvo (1999). On some solutions of the distributivity equation. *Fuzzy Sets and Systems*, **104**, pp. 85–96.

[15] G. Dimuro, B. Bedregal (2014). Archimedean overlap functions: The ordinal sum and the cancellation, idempotency and limiting properties. *Fuzzy Sets and Systems*, **252**, pp. 39–54.

[16] G. Dimuro, B. Bedregal (2015). On residual implications derived from overlap functions. *Information Sciences*, **312**, pp. 78–88.

[17] G.P. Dimuro, B. Bedregal, H. Bustince, M.J. Asiáin, R. Mesiar (2015). On additive generators of overlap functions. *Fuzzy Sets and Systems*. DOI: 10.1016/j.fss.2015.02.008.

[18] G. Dimuro, B. Bedregal, R. Santiago (2014). On (G, N)-implications derived from grouping functions. *Information Sciences*, **279**, pp. 1–17.

[19] J. Fodor, M. Roubens (1994). Fuzzy Preference Modelling and Multicriteria Decision Support. *Theory and Decision Library*, Kluwer Academic Publishers.

[20] E. Klement, M. Navara (1999). A survey on different triangular norm-based fuzzy logics. *Fuzzy Sets and Systems*, **101**, pp. 241–251.

[21] N. Madrid, A. Burusco, H. Bustince, J. Fernández, I. Perfilieva (2015). Upper bounding overlaps by groupings. *Fuzzy Sets and Systems*, **264**, pp. 76–99.

[22] M. Navara (1999). Characterization of measures based on strict triangular norms. *Journal of Mathematical Analysis and Applications*, **236(2)**, pp. 370–383.

[23] E. Trillas (1979). Sobre funciones de negación en la teoria de los conjuntos difusos. *Stochastica*, **3(1)**, pp. 47–59.

[24] L.A. Zadeh (1965). Fuzzy sets. *Information and Control*, **8**, pp. 338–353.

SUPER- AND SUBADDITIVE CONSTRUCTIONS OF AGGREGATION FUNCTIONS

Alexandra Šipošová and Ladislav Šipeky
Department of Mathematics
Faculty of Civil Engineering
Slovak Univ. of Technology
Radlinského 11
810 05 Bratislava, Slovakia
{alexandra.siposova,ladislav.sipeky}@stuba.sk

Fabio Rindone
Dept. of Economics and Business
University of Catania
95029 Catania
Italy
frindone@unict.it

Summary

Two construction methods for aggregation functions based on a restricted a priori known decomposition set and decomposition weighting function are introduced and studied. The outgoing aggregation functions are either superadditive or subadditive. Several examples, including illustrative figures, show the potential of the introduced construction methods.

Keywords: aggregation function, subadditive transformation, superadditive transformation.

1 INTRODUCTION

Aggregation functions play an important rôle in many domains where an n-dimensional input representation is represented by a single value. For more information and details we recommend monographs [1], [2]. Recall that for $n \in N$ a monotone function $A : [0,1]^n \to [0,1]$ is called an aggregation function whenever it satisfies two boundary conditions $A(0,...,0) = A(\mathbf{0}) = 0$ and $A(1,...,1) = A(\mathbf{1}) = 1$. Observe that we will not consider the usual convention $A(x) = x$ for 1-dimensional aggregation functions. Note also that, in general, some other interval I can be considered instead of the unit interval $[0,1]$. However, our results related to $[0,1]$ domain can be easily generalized to the domain I.

In several practical situations, the aggregation function A is not known on its full domain $[0,1]^n$, but only on a subdomain $\mathcal{H} \subseteq [0,1]^n$. More often the boundary condition $A(\mathbf{1}) = 1$ is not important, i.e., A and λA gives the same information for the user, independently of $\lambda \in]0,\infty[$. This is, e.g., the case when A is considered as a utility function. The above facts have inspired us to introduce two construction methods for aggregation functions when only a partial information is known. Our approach was motivated by the ideas from [3] dealing with superadditive and subadditive transformations of aggregation functions on $[0,\infty[$.

Our contribution is organized as follows. In Section 2, we introduce superadditive and subadditive functions B^* and B_*, and the related aggregation functions $A^{\mathcal{H},B}$ and $A_{\mathcal{H},B}$, including some preliminary results. In Section 3, we exemplify the functions B^* and B_* for several decomposition pairs $(\mathcal{H}, B)$. Finally, some concluding remarks are added.

2 SUPER- AND SUBADDITIVE CONSTRUCTIONS OF AGGREGATION FUNCTIONS

In what follows, an arbitrary subset $\mathcal{H}$ of $[0,1]^n$ such that $\mathbf{0} \in \mathcal{H}$ will be called a *decomposition set.* A function $B : \mathcal{H} \to [0,1]$, not identically equal to zero, with $B(\mathbf{0}) = 0$ and such that $B(\mathbf{x}) \le B(\mathbf{y})$ whenever $\mathbf{x} \le \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathcal{H}$, will be called a *decomposition weighing function.*

Although a decomposition weighing function is defined only on $\mathcal{H}$ which, in the extreme case, may consist besides $\mathbf{0}$ just of a single point, one may introduce its transformation to the entire unit n-cube $[0,1]^n$ by letting (with the convention that $\inf \emptyset = \infty$ and $\sup \emptyset = 0$)

$$B_*(\mathbf{x}) = \inf \left\{ \sum_{i=1}^{k} B(\mathbf{y}^{(i)}) \mid (\mathbf{y}^{(i)})_{i=1}^{k} \in \mathcal{H}^k;\ \sum_{i=1}^{k} \mathbf{y}^{(i)} \ge \mathbf{x} \right\} \quad (1)$$

ISBN: 978–83–8012–519–3

and

$$B^*(\mathbf{x}) = \sup\left\{\sum_{i=1}^{k} B(\mathbf{y}^{(i)}) \mid (\mathbf{y}^{(i)})_{i=1}^{k} \in \mathcal{H}^k;\ \sum_{i=1}^{k} \mathbf{y}^{(i)} \le \mathbf{x}\right\}. \quad (2)$$

Observe that, in general, B_* and B^* are mappings from $[0,1]^n \to [0,\infty]$. The pair $(\mathcal{H}, B)$ will be called *subadmissible* if $B_*(1) \in]0,\infty[$, and *superadmissible* if $B^*(1) \in]0,\infty[$. The set of all subadmissible and superadmissible pairs will be denoted simply by Sub_n and $Super_n$, respectively.

For any subadmissible (superadmissible) pair $(\mathcal{H}, B)$ we may introduce normalized versions of the transformation of B introduced above by letting

$$A_{\mathcal{H},B} :\ [0,1]^n \to [0,1];\ \mathbf{x} \mapsto B_*(\mathbf{x})/B_*(\mathbf{1}) \quad (3)$$

and

$$A^{\mathcal{H},B} :\ [0,1]^n \to [0,1];\ \mathbf{x} \mapsto B^*(\mathbf{x})/B^*(\mathbf{1}), \quad (4)$$

where in both cases $\mathbf{1} \in [0,1]^n$ denotes the all-one vector.

Quite expectedly, these normalized versions are subadditive and superadditive, respectively:

Proposition 1. *If $(\mathcal{H}, B)$ is a subadmissible pair, then $A_{\mathcal{H},B}$ is a subadditive aggregation function. Analogously, if $(\mathcal{H}, B)$ is a superadmissible pair, then $A^{\mathcal{H},B}$ is a superadditive aggregation function.*

We illustrate our proporsals in the next simple example. Let $n = 1$ and consider a trivial decomposition system $\mathcal{H} = \{0, 1/t\}$ for some fixed positive integer t. Further, let B be a decomposition weighing function defined by $B(0) = 0$ and $B(1/t) = b > 0$. Obviously, $B_*(0) = 0$. For any $x \in]0,1]$, letting $k = \lceil tx \rceil$ (the ceiling of tx) we have $x \in](k-1)/t, k/t]$, so that $B_*(x) = kb$ and hence $B_*(1) = tb$; it follows that $A_{\mathcal{H},B}(x) = B_*(x)/B_*(1) = \lceil tx \rceil/t$, which is a subadditive aggregation function. By the same token, letting $\ell = \lfloor tx \rfloor$ (the floor of tx) we have $x \in [\ell t, (\ell+1)t[$, so that $B^*(x) = \ell b$, $B^*(x) = tb$, and $A^{\mathcal{H},B}(x) = B^*(x)/B^*(1) = \lfloor tx \rfloor/t$, which is a superadditive aggregation function.

Proposition 2. *If $(\mathcal{H}, B)$ is a subadmissible pair, then $A_{\mathcal{H},B} = B$ if and only if $\mathcal{H} = [0,1]^n$ and B is subadditive, with $B(\mathbf{1}) = 1$. Analogously, if $(\mathcal{H}, B)$ is a superadmissible pair, then $A^{\mathcal{H},B} = B$ if and only if $\mathcal{H} = [0,1]^n$ and B is superadditive, with $B(\mathbf{1}) = 1$.*

Observe that on the space of subadmissible pairs Sub_n we have a natural partial order $\preceq_{Sub}$ defined by

$$(\mathcal{H}_1, B_1) \preceq_{Sub} (\mathcal{H}_2, B_2) \text{ if and only if } \mathcal{H}_1 \supseteq \mathcal{H}_2 \text{ and } B_1|\mathcal{H}_2 \le B_2\ . \quad (5)$$

Similarly, on the space of superadmissible pairs $Super_n$ we have a natural partial order $\preceq_{Super}$ defined by

$$(\mathcal{H}_1, B_1) \preceq_{Super} (\mathcal{H}_2, B_2) \text{ if and only if } \mathcal{H}_1 \subseteq \mathcal{H}_2 \text{ and } B_1 \le B_2|\mathcal{H}_1\ . \quad (6)$$

This allows us to compare the values of the corresponding aggregation functions as follows.

Proposition 3. *Let $(\mathcal{H}_1, B_1), (\mathcal{H}_2, B_2) \in Sub_n$ and $(\mathcal{H}_1, B_1) \preceq_{Sub} (\mathcal{H}_2, B_2)$. If $(B_1)_*(\mathbf{1}) = (B_2)_*(\mathbf{1})$, then $A_{\mathcal{H}_1,B_1} \le A_{\mathcal{H}_2,B_2}$. Analogously, if $(\mathcal{H}_1, B_1), (\mathcal{H}_2, B_2) \in Super_n$ are such that $(\mathcal{H}_1, B_1) \preceq_{Super} (\mathcal{H}_2, B_2)$. If $B_1^*(\mathbf{1}) = B_2^*(\mathbf{1})$, then $A^{\mathcal{H}_1,B_1} \le A^{\mathcal{H}_2,B_2}$.*

Remark. The above result will not be valid in general if, say, in the $Super_n$ case, the assumption $B_1^*(\mathbf{1}) = B_2^*(\mathbf{1})$ is dropped. To see this, for $n = 1$, $\mathcal{H}_1 = \{0, 1/2\}$, $\mathcal{H}_2 = \{0, 1/2, 1\}$, $B_1(1/2) = 1$ and $B_2(1/2) = 1$, $B_2(1) = 4$, so that $B_1^*(1) = 2$ and $B_2^*(1) = 4$. It is then easy to see that, for example, $A^{\mathcal{H}_1,B_1}(1/2) = 1/2$, while $A^{\mathcal{H}_2,B_2}(1/2) = 1/4$, violating the inequality $A^{\mathcal{H}_1,B_1} \le A^{\mathcal{H}_2,B_2}$.

We continue with an auxiliary result in dimension 1.

Proposition 4. *Let $\mathcal{H} \subseteq [0,1]$ be a decomposition set and let $B :\ \mathcal{H} \to [0,1]$ be a decomposition weighing function of dimension 1. Then,*

(a) $(\mathcal{H}, B) \in Sub_1$ *if and only if* $\inf\{B(x)/x \mid x \in \mathcal{H} \setminus \{0\}\} > 0$, *and*

(b) $(\mathcal{H}, B) \in Super_1$ *if and only if* $\sup\{B(x)/x \mid x \in \mathcal{H} \setminus \{0\}\} < \infty$.

Proposition 5. *Let $\mathcal{H} \subset [0,1]^n$ be a decomposition set and let $B :\ \mathcal{H} \to [0,1]$ be a decomposition weighing function of dimension $n \ge 1$. Then,*

(a) $(\mathcal{H}, B) \in Sub_n$ *if and only if for each* $i \in \{1, ..., n\}$ *there is an* $\mathbf{x} \in \mathcal{H}$ *such that* $(\mathbf{x})_i > 0$ *and* $\inf\left\{\frac{B(\mathbf{x})}{(\mathbf{x})_i} \mid \mathbf{x} \in \mathcal{H},\ (\mathbf{x})_i > 0\right\} > 0$ *for some* $i \in \{1, \ldots, n\}$;

(b) $(\mathcal{H}, B) \in Super_n$ *if and only if* $\sup\left\{\frac{B(\mathbf{x})}{\max(\mathbf{x})} \mid \mathbf{x} \in \mathcal{H} \setminus \{0\}\right\} < \infty$.

3 EXAMPLES

In this section we will present examples of functions B_* and B^* for specific decomposition sets and related decomposition weighting functions.

Example 1. *Let* $\mathcal{H}$ $=\{(0,0),(0.1,0.1)\}$ *and* $B(\mathbf{y})= y_1$*, where* y_1 *is the first coordinate of* $\mathbf{y}$*. Values of* B_* *and* B^* *are depicted in Figure 1 and Figure 2, respectively. Observe that* $\sum_{i=1}^{k}\mathbf{y}^{(i)}$ *appearing in expression* (1) *and* (2) *always have the form* $k(0.1,0.1)$ *for* $k \in \{0,1,...,10\}$*, because the only vector that can be used for summation is* $(0.1,0.1)$*. This explains the shape of the graphs in Figure 1 and Figure 2.*

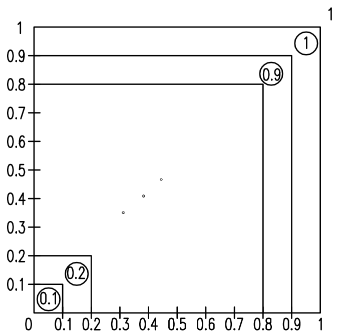


Figure 1: B_* from Example 1

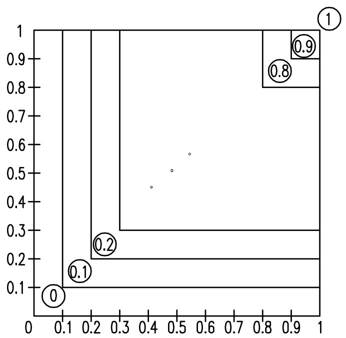


Figure 2: B^* from Example 1

Example 2. *Let* $\mathcal{H}=\{(0,0),(0.8,0.3),(0.2,0.7)\}$ *and let* $B(0.8,0.3) = 0.8$*,* $B(0.2,0.7) = 0.6$*. It can be shown that in this case we have* $B^*(1,1) = B_*(1,1) = 1.4$*. The corresponding values of* B_* *and* B^* *are depicted in Figure 3 and Figure 4, respectively.*

Example 3. *Let* $\mathcal{H} = \{(0,0),(0.2,0.3),(0.5,0.7)\}$ *and let* $B = \prod$*. A schematic description of* B^* *is in Figure 5.*

Example 4. *In this example we will use a segment for* $\mathcal{H}$ *by letting* $\mathcal{H} = \{(x,y)|\ x \in [0.1,1], y = 0.1 - x\} \cup$

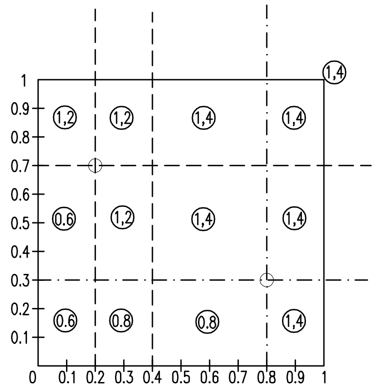


Figure 3: B_* from Example 2

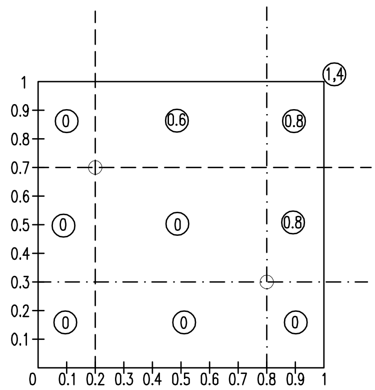


Figure 4: B^* from Example 2

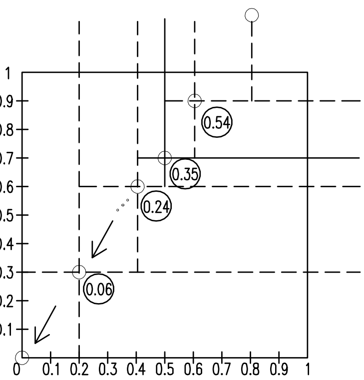


Figure 5: B^* from Example 3

$\{(0,0)\}$*. The aggregation function is defined as follows* $B : \mathcal{H} \setminus \{0,0\} \to 0.05$*. The function* B^* *is depicted in Figure 6 (2D) and in Figure 7 (3D).*

4 CONCLUDING REMARKS

We have introduced two methods of constructing aggregation functions on $[0,1]$ in situation when only a partial information is available. We have exemplified the superadditive functions B^* and the subadditive

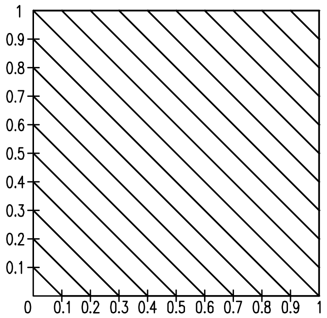


Figure 6: Contour lines of B^* from Example 4

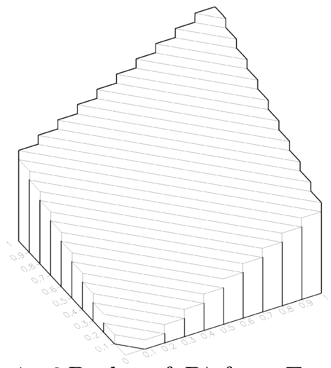


Figure 7: $3D$ plot of B^* from Example 4

functions B_*, having in mind that the related aggregation functions $A^{\mathcal{H},B}$ and $A_{\mathcal{H},B}$ are easily obtained by normalization of B^* and B_*, respectively. We expect applications of our approach in economics, social sciences, etc., and especially in multicriteria decision support.

Acknowledgements

The first and the second author acknowledge support from the project VEGA 1/0420/15 and APVV-14-0013.

References

[1] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation Functions: A Guide for Practitioners. Springer, Berlin, ISBN: 978-3-540-73720-9 (Print) 978-3-540-73721-6 (Online).

[2] M. Grabisch, J-L. Marichal, R. Mesiar, E. Pap (2009). Aggregation Functions (Encyklopedia of Mathematics and its Applications Vol. 127). Cambridge University Press.

[3] S. Greco, R. Mesiar, F. Rindone, L. Šipeky (2014). The superadditive and the subadditive transformations of integrals and aggregation functions. *Fuzzy Sets and Systems*, submitted.

STANDARD, WEAK AND DIRECTIONAL MONOTONICITY OF MIXTURE FUNCTIONS

Jana Špirková
Faculty of Economics, Matej Bel University,
Tajovského 10, 975 90 Banská Bystrica, Slovakia
jana.spirkova@umb.sk

Summary

Mixture functions represent a special class of weighted averaging functions whose weights are determined by continuous weighting functions which depend on the input values. If they are non-decreasing (standard) monotone, they also belong to the important class of aggregation functions. This paper presents properties of selected mixture functions with special stress on their standard, weak and directional monotonicity.

Keywords: Mixture function, Standard monotonicity, Weak monotonicity, Directional monotonicity.

1 INTRODUCTION

Standard monotonicity of mixture functions and their generalizations was studied in several papers, for example, [1], [5], and [8], [9]. We also studied this problem in papers [6], [7] and [10], where we provided several sufficient conditions for standard non-decreasing monotonicity of mixture functions.
However, standard monotonicity is violated in the case of several fusion techniques frequently applied in real data processing, in the case of implications, which are characterized by hybrid monotonicity or the mode function or various types of means [1]. Nevertheless, in many applications it is sufficient if processing functions are *weakly monotone*, see [14] and [15].
The property of weak monotonicity is very useful for calculating representative values of clusters of data in the presence of outliers. According to [13], cluster structure may change when only some inputs are increased (or decreased), but it does not change when all inputs are changed by the same value.
Moreover, regarding the generalization of the concept of aggregation functions, in [3], authors introduced and discussed so-called fusion functions and their directional monotonicity. Their results generalize the results of [15] concerning so-called weak monotonicity. In our recent research, we have focused our attention not only on the study of the standard non-decreasing monotonicity of mixture functions but also on their weak and directional monotonicity. The latest results of our investigation in this area can be found in [2], [11] and [12].
The paper consists of four sections. Section 1 presents an overview of the latest results concerning on monotonicity of mixture functions. Section 2 contains the basic definitions related to standard, weak and directional monotonicity. Section 3 provides sufficient conditions of standard and weak monotonicity of mixture functions and presents the latest sufficient conditions of weak and directional monotonicity and relevant results concerning this topic. The attention is mainly focused on three types of monotonicity of the mixture functions generated by selected weighting functions. Conclusion summarizes the presented results and brings some ideas for the future research.

2 PRELIMINARIES

Throughout the paper, the following notations will be used. Consider any closed non-empty interval $\mathbb{I} = [a, b] \subset \overline{\mathbb{R}} = [-\infty, \infty]$. Then $\mathbb{I}^n = \{\mathbf{x} = (x_1, \ldots, x_n) \mid x_i \in \mathbb{I}, i = 1, \ldots, n\}$ is the set of all input vectors $\mathbf{x}$.
In this section, we provide definitions of functions whose properties will be studied. Firstly, we recall the definitions of aggregation and mixture functions.

Aggregation functions

Definition 2.1 *A function $A : \mathbb{I}^n \to \mathbb{I}$ is called an aggregation function if it is monotone non-decreasing in each variable and satisfies the boundary conditions $A(\mathbf{a}) = a$, $A(\mathbf{b}) = b$, where $\mathbf{a} = (a, a, \ldots, a)$, $\mathbf{b} =$*

ISBN: 978–83–8012–519–3

$(b, b, \ldots, b)$.

Mixture functions

Definition 2.2 *A mapping $M_g : \mathbb{I}^n \to \mathbb{I}$ given by*

$$M_g(x_1, \ldots, x_n) = \frac{\sum_{i=1}^{n} g(x_i) \cdot x_i}{\sum_{i=1}^{n} g(x_i)}, \quad (1)$$

where $g : \mathbb{I} \to [0, \infty[$ is a continuous weighting function, is called a mixture function.

In the following part, we discuss three types of the monotonicity of mixture functions, namely, standard monotonicity, weak monotonicity and directional monotonicity.

2.1 MONOTONICITY

In this part, we start by recalling the basic definitions.

Standard monotonicity

Definition 2.3 *A function $A : \mathbb{I}^n \to \mathbb{I}$ is monotone non–decreasing if and only if, for all $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$, such that $\mathbf{x} \leq \mathbf{y}$, it holds that $A(\mathbf{x}) \leq A(\mathbf{y})$.*

Weak monotonicity

Definition 2.4 ([15]) *A function $A : \mathbb{I}^n \to \mathbb{I}$ is weakly monotone non-decreasing if $A(\mathbf{x} + k\mathbf{1}) \geq A(\mathbf{x})$ for all $\mathbf{x}$ and for any $k > 0$, $\mathbf{1} = (\underbrace{1, 1, \cdots, 1}_{n \text{ times}})$, such that $\mathbf{x}, \mathbf{x} + k\mathbf{1} \in \mathbb{I}^n$.*

It is clear that each monotone non-decreasing function is also weakly monotone non-decreasing.

Here, we also recall the definition of shift-invariantness that is related to weak monotonicity.

Shift-invariance

Definition 2.5 *A function $A : \mathbb{I}^n \to \mathbb{I}$ is shift–invariant if $A(\mathbf{x}+k\mathbf{1}) = A(\mathbf{x})+k$ whenever $\mathbf{x}, \mathbf{x}+k\mathbf{1} \in \mathbb{I}^n$ and $A(\mathbf{x}) + k \in \mathbb{I}$.*

It is easy to see that each shift-invariant function A is also weakly monotone non-decreasing, [15].

Inspired by the notion of weak monotonicity the researchers have recently opened investigation of the so-called directional monotonicity of $\mathbb{I}^n \to \mathbb{I}$ functions which is defined as follows:

Directional monotonicity

Definition 2.6 ([3]) *Let $\mathbf{r}$ be a real n-dimensional vector, $\mathbf{r} \neq \mathbf{0}$. A function $A : \mathbb{I}^n \to \mathbb{I}$ is $\mathbf{r}$-non-decreasing if for all $\mathbf{x} \in \mathbb{I}^n$ and all $k > 0$ such that $\mathbf{x} + k\mathbf{r} \in \mathbb{I}^n$, it holds that $A(\mathbf{x} + k\mathbf{r}) \geq A(\mathbf{x})$.*

Vectors $\mathbf{r} \neq \mathbf{0}$ are called *directions*. It is clear, that weakly monotone functions are $\mathbf{r}$-non-decreasing in the direction of vector $\mathbf{r} = (1, 1, \ldots, 1)$.
The monotone non-decreasingness of a function $A : \mathbb{I}^n \to \mathbb{I}$ is equivalent to the $\mathbf{e}_i$-directional non-decreasingness of A for each $i = 1, 2, \ldots, n$, where $\mathbf{e}_i$ is the vector, whose ith coordinate is equal to 1, and other coordinates are equal to 0.

2.2 SUFFICIENT CONDITIONS OF MONOTONICITY

In this part, we mention sufficient conditions of standard and weak monotonicity of mixture functions.

2.2.1 Standard monotonicity of mixture functions

In case $\mathbb{I} = [0, 1]$, Ribeiro and Marques Pereira in [8] showed that any non-decreasing differentiable weighting function $g : [0, 1] \to [0, \infty[$ such that

$$g \geq g' \quad (2)$$

yields a non-decreasing mixture function (1).

The non-decreasing monotonicity of mixture functions we studied deeply in [10]. We provided there more general sufficient conditions ensuring the non-decreasing monotonicity than that one in (2). For the convenience of the reader we repeat the relevant results from [10].

Theorem 2.7 *Mixture function $M_g : \mathbb{I}^n \to \mathbb{I}$, $\mathbb{I} = [0, 1]$, given by (1), with non-decreasing piecewise differentiable weighting function g, is monotone non-decreasing if the weighting function g satisfies at least one from the following conditions:*

1. $$g(x) \geq g'(x), \quad (3)$$

2. $$g(x) \geq g'(x) \cdot (1 - x), \quad (4)$$

3. *for a fixed n, $n > 1$,*
$$\frac{g^2(x)}{(n-1)g(1)} + g(x) \geq g'(x) \cdot (1 - x). \quad (5)$$

Other sufficient conditions of the standard monotonicity of mixture functions and their generalizations can be found in [10].

2.2.2 Weak monotonicity of mixture functions

Weak monotonicity of mixture functions and their generalizations we studied in [2], [11] and [12]. Here, we give a sufficient condition of weak monotonicity of mixture functions.

Theorem 2.8 ([12]) *Let $M_g : \mathbb{I}^n \to \mathbb{I}$, be a mixture function* (1) *with differentiable weighting function $g : \mathbb{I} \to [0, \infty[$. Then M_g is weakly monotone non-decreasing if*

$$\left(\sum_{i=1}^{n} g(x_i)\right)^2 + \left(\sum_{i=1}^{n} g(x_i)\right) \cdot \left(\sum_{i=1}^{n} x_i \cdot g'(x_i)\right) - \left(\sum_{i=1}^{n} x_i \cdot g(x_i)\right) \cdot \sum_{i=1}^{n} g'(x_i) \geq 0 \qquad (6)$$

for all $\mathbf{x} \in \mathbb{I}^n$.

Proof is based on non-negativity of directional derivative of M_g, i.e., $D_{\mathbf{1}}(M_g)(\mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial M_g}{\partial x_i} \geq 0$. □

3 MONOTONICITY OF MIXTURE FUNCTIONS

In this section, we give our latest results and introduce sufficient conditions of weak and also directional monotonicity of mixture functions with selected weighting functions.
We concentrate our investigation on mixture functions generated by shift-invariant linear weighting function $g(x) = x + l$, $l \geq 0$; and special linear function in the shape $g(x) = cx + 1 - c$, $c \geq 0$.

3.1 MIXTURE FUNCTION WITH SHIFT-INVARIANT LINEAR WEIGHTING FUNCTION

Using the following sufficient conditions of standard and weak monotonicity, we can show that the weak monotonicity gives us a wider choice of coefficients l, i.e., it gives us a wider choice of aggregations by mixture functions.

Proposition 3.1 *Let $M_g : [0,1]^n \to [0,1]$ be a mixture function defined by* (1)*, and $g : [0,1] \to [0, \infty[$ be a weighting function given by $g(x) = x + l$, $l \geq 0$.*
Then M_g is:

1. *standard monotone non-decreasing for $l \geq 1$;*
2. *for a fixed n:*
   a) *[11] standard monotone non-decreasing for*
   $$l \geq \sqrt{\frac{n-1}{n}};$$
   b) *[11], [14] weakly monotone non-decreasing for*
   $$l \geq \frac{n-2}{n}.$$

Proof

1. With respect to (3) and (4).
2. a) With respect to (5) and also Definition 2.3.
   b) With respect to (6) and also Definition 2.4.

□

The next theorem expresses the global bound of coefficient l in weighting function $g(x) = x + l$ for weak monotonicity of mixture function (1). In fact, based on the number of input values, we can obtain a lower bound for l smaller than $\frac{\sqrt{2}-1}{2}$, which can be seen from conditions in Proposition 3.1.

Theorem 3.2 ([2]) *Let $M_g : [0,1]^n \to [0,1]$ be a mixture function defined by* (1) *with shift-invariant weighting function $g(x) = x + l$, $l \geq 0$. Then M_g is weakly monotone non-decreasing for*

$$l \geq \frac{\sqrt{2}-1}{2}. \qquad (7)$$

In the next theorem, we introduce our initial results related to directional monotonicity of mixture functions.

Theorem 3.3 *Let $M_g : [0,1]^2 \to [0,1]$ be a mixture function defined by* (1) *with shift-invariant weighting function $g(x) = x + l$, $l \geq 0$. Then M_g is directionally monotone non-decreasing only for vectors $\mathbf{r} = (r_1, r_2)$ which satisfy the condition*

$$r_1 = r_2 > 0.$$

Proof Let $\mathbf{r} = (r_1, r_2) \neq \mathbf{0}$. Let $\mathbf{x} = (x, y) \in \mathbb{I}^2$ and $k > 0$ such that $\mathbf{x} + k\mathbf{r} \in \mathbb{I}^2$.
From Definition 2.6 we get

$$\frac{(x + kr_1)(x + kr_1 + l) + (y + kr_2)(y + kr_2 + l)}{x + y + 2l + k(r_1 + r_2)} \geq \geq \frac{x(x + l) + y(y + l)}{x + y + 2l},$$

whence

$$(x + y + 2l)\left(2(xr_1 + yr_2) + k(r_1^2 + r_2^2) + l(r_1 + r_2)\right) \geq \\ \geq (r_1 + r_2)(x^2 + y^2 + l(x + y)).$$

Without loss of generality, for $k \to 0$ and after some modification, we obtain inequality

$$(x^2 - y^2)(r_1 - r_2) + 2xy(r_1 + r_2) + 4l(xr_1 + yr_2) + \\ +2l^2(r_1 + r_2) \geq 0.$$

Without loss of generality, for $r_1 > 0$, $r_2 > 0$, $x \to 0$, $y \to 1$ or $x \to 1$, $y \to 0$, we obtain consequently the conditions

$$l^2 + 2\frac{r_2}{r_1 + r_2}l + \frac{r_2 - r_1}{r_1 + r_2} \geq 0$$
$$l^2 + 2\frac{r_1}{r_1 + r_2}l + \frac{r_1 - r_2}{r_1 + r_2} \geq 0$$

For $l \to 0$, the conditions $r_2 \geq r_1$ and $r_2 \leq r_1$ must be satisfied. This means that the considered mixture function is directionally monotone increasing only in the direction (r_1, r_2) where $r_1 = r_2 > 0$. □

Corollary 1 *Let $M_g : [0,1]^2 \to [0,1]$ be a mixture function defined by (1) with shift-invariant weighting function $g(x) = x + l$, $l \geq 0$. Then M_g is directionally monotone non-decreasing only in the direction $(1,1)$, and in no other directions, i.e., it is only weakly monotone non-decreasing.*

Remark 1 *From Corollary 1 we can state that the Lehmer mean $LeM(x,y) = \dfrac{x^2 + y^2}{x + y}$ is only weakly monotone non-decreasing.*

3.2 MIXTURE FUNCTIONS WITH SPECIAL LINEAR WEIGHTING FUNCTION

This part focusses on all mentioned types of monotonicity of mixture function with $g(x) = cx + 1 - c$, $c \in [0,1]$.
On the basis of the mentioned sufficient conditions of standard and weak monotonicity, we can show that a weak monotonicity gives us a wider choice of coefficients c on aggregation.

Proposition 3.4 *Let $M_g : [0,1]^n \to [0,1]$ be a mixture function defined by (1) with the weighting function $g(x) = cx + 1 - c$, $c \in [0,1]$. Then M_g is monotone non-decreasing:*

1. *for $c \in [0, 0.5]$,*
2. *for a fixed n:*
   - $n = 2$ $\quad c \in [0, 0.585786]$,
   - $n = 3$ $\quad c \in [0, 0.550510]$,
3. *weakly monotone non-decreasing for:*
   - $n = 2$ $\quad c \in [0, 1]$,
   - $n = 3$ $\quad c \in [0, 0.878679]$.

Proof

1. With respect to (3) and (4).
2. With respect to (5).
3. With respect to (6) and the Mathematica 8 system.

□

Similarly, as in Proposition 3.2, we can determine minimal interval of c, to be mixture function (1) weakly monotone non-decreasing. This interval of c can be written as follows.

Proposition 3.5 *Let $M_g : [0,1]^n \to [0,1]$ be a mixture function defined by (1) with the weighting function $g(x) = cx + 1 - c$, $c \in [0,1]$. Then M_g is weakly monotone non-decreasing for*

$$c \in [0, 2\sqrt{2} - 2]. \tag{8}$$

Proof Using sufficient condition (6), we get

$$(c\sum_{i=1}^{n} x_i + n - nc)^2 + (c\sum_{i=1}^{n} x_i + n - nc)(c\sum_{i=1}^{n} x_i) - \\ -(\sum_{i=1}^{n} x_i(cx_i + 1 - c))na \geq 0,$$

which can be written as

$$c^2\left[2(\sum_{i=1}^{n} x_i)^2 - n\sum_{i=1}^{n} x_i^2 - 2n\sum_{i=1}^{n} x_i + n^2\right] + \\ +c(2n\sum_{i=1}^{n} x_i - 2n^2) + n^2 \geq 0. \tag{9}$$

Since we consider mixture function on the unit cube, we can formulate minimization of the left-hand side of the previous inequality at the vertices of the unit cube, see [2], [4].
Assume, let $(\underbrace{1, 1, \ldots, 1}_{k\text{-times}}, \underbrace{0, 0, \ldots, 0}_{(n-k)\text{-times}})$ be the input vector, at which the expression on left-hand side attains

the smallest value. Using substitution $z = \bar{x} = \frac{k}{n}$ we can rewrite condition (9) as

$$c^2(2z^2 - 3z + 1) + 2c(z - 1) + 1 \geq 0. \qquad (10)$$

The expression $2z^2 - 3z + 1$ is positive for $z \in [0, \frac{1}{2}[$. We need to find the smallest value of

$$c^2(2z^2 - 3z + 1) + 2c(z - 1) + 1$$

such that $z \in [0, \frac{1}{2}[$. Therefore, we assume a root of the quadratic expression in the shape

$$c = \frac{-z + 1 - \sqrt{z - z^2}}{2z^2 - 3z + 1},$$

and it has minimal value for $z = \frac{1}{4}(2 - \sqrt{2})$. It follows that $c \leq 2\sqrt{2} - 2$, and hence $c \in [0, 2\sqrt{2} - 2[$ ensures weak monotonicity of the mentioned mixture function M_g.

We recall that $c = 2\sqrt{2} - 2$ represents global minimum, what means that this number can be exceeded. For instance, if $z = \frac{1}{2}$, we get the condition $c \leq 1$, which is consistent with our result in Proposition 3.1 and Theorem 3.3. □

The next proposition gives us the sufficient condition of non-decreasing directional monotonicity of mixture function (1) generated by $g(x) = cx + 1 - c$, $c \in [0, 1]$.

Proposition 3.6 *Let $M_g : [0, 1]^2 \to [0, 1]$ be a mixture function defined by (1) with the weighting function $g(x) = cx + 1 - c$, $c \in [0, 1]$. Then M_g is directionally non-decreasing for vectors $\mathbf{r} = (r_1, r_2)$ which satisfy conditions*

- *for $c \in [0, 2 - \sqrt{2}[$,*

$$r_2 \geq \frac{c^2 - 4c + 2}{c^2 - 2} r_1 \text{ and } r_2 \geq \frac{c^2 - 2}{c^2 - 4c + 2} r_1; \qquad (11)$$

- *for $c \in]2 - \sqrt{2}, 1]$,*

$$r_2 \geq \frac{c^2 - 4c + 2}{c^2 - 2} r_1 \text{ and } r_2 \leq \frac{c^2 - 2}{c^2 - 4c + 2} r_1; \qquad (12)$$

- *for $c = 2 - \sqrt{2}$, $r_1 \geq 0$ and $r_2 \geq 0$.*

Proof On the basis of Definition 2.6, let

$$\frac{(x + kr_1)(c(x + kr_1) + 1 - c)(y + kr_2)(c(y + kr_2) + 1 - c)}{c(x + y) + 2 - 2c + ck(r_1 + r_2)} \geq \frac{x(cx + 1 - c) + y(cy + 1 - c)}{c(x + y) + 2 - 2c}.$$

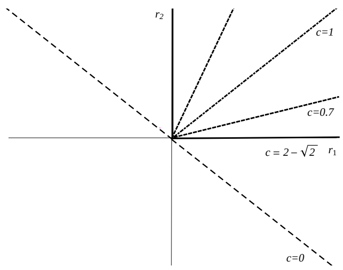


Figure 1: Space of directional non-decreasing monotonicity of M_g with $g(x) = cx + 1 - c$.

Without loss of generality, for $k \to 0$, we get

$$\begin{aligned} c^2((x^2 - y^2)(r_1 - r_2) + 2xy(r_1 + r_2) \\ -4(xr_1 + yr_2) + 2(r_1 + r_2)) \\ +4c(r_1(x - 1) + r_2(y - 1)) + 2(r_1 + r_2) \geq 0, \end{aligned} \qquad (13)$$

and so we write (13), for $x \to 0$, $y \to 1$, and subsequently for $x \to 1$, $y \to 0$, we obtain the conditions

$$(r_1 - r_2)c^2 - 4r_1 c + 2(r_1 + r_2) \geq 0$$

and

$$(r_2 - r_1)c^2 - 4r_2 c + 2(r_1 + r_2) \geq 0.$$

From the previous inequalities it follows that

$$r_2 \geq -\frac{c^2 - 4c + 2}{2 - c^2} r_1$$

and

$$r_2 \geq -\frac{c^2 - 2}{c^2 - 4c + 2} r_1, \quad \text{for } c \in [0, 2 - \sqrt{2}[;$$

or

$$r_2 \leq -\frac{c^2 - 2}{c^2 - 4c + 2} r_1, \quad \text{for } c \in]2 - \sqrt{2}, 1].$$

□

This means that the considered mixture function is directionally monotone non-decreasing, for example:

1. for $c = 0$ in all directions $\mathbf{r} \neq \mathbf{0}$ from the half-plane highlighted in Figure 1;
2. for $c = 2 - \sqrt{2}$ in all directions in the first quadrant;
3. for $c = 0.7$ in all directions in the highlighted the acute angle.

From our proof, as well as from Figure 1, it is clear that the space of directionally non-decreasing monotonicity decreases gradually from upper half-plane bounded by line $r_2 = -r_1$ (for $c = 0$) to half line $r_2 = r_1$ (for $c = 1$).

4 CONCLUSIONS

In the paper, we have provided sufficient conditions for the standard and weak non-decreasing monotonicity of mixture functions generated by selected weighting functions. Moreover, we introduced new sufficient conditions for their weak and also directional monotonicity.

In our future research, we intend to continue in a deep study of weak and directional monotonicity of mixture functions generated by special types of weighting functions, e.g., by Gaussian weighting functions $g(x) = \exp\left\{-\left(\frac{x-c}{\sigma}\right)^2\right\}$, $c, \sigma \in \mathbb{R}$, shift-invariant exponential functions of the form $g(x) = q^{cx+d}$, $q > 0, c, d \in \mathbb{R}$, as well as power functions $g(x) = x^m + l, m, l \in \mathbb{R}$. We expect the obtained results to be used in image processing applications.

Acknowledgements

This paper has been supported by the Project VEGA no. 1/0542/13 Risk management and the actuarial function in life insurance.

References

[1] G. Beliakov, A. Pradera, T. Calvo (2007). Aggregation functions: A Guide for Practitioners, Springer, Berlin.

[2] G. Beliakov, J. Špirková (2014). Weak monotonicity of mixture functions with special weighting functions, submitted to the journal Fuzzy Sets and Systems.

[3] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar (2014). Fusion functions and their directional monotonicity. *Information Processing and Management of Uncertainty in Knowledge-Based Systems Communications in Computer and Information Science*, Volume 444, pp 262-268.

[4] P. L. De Angelis, P. M. Pardalos, G. Toraldo (1997). Developments in Global Optimization, Quadratic programming with box constraints, editor: I. M. Bomze et al., Kluwer Academic Publishers, Dordrecht.

[5] R. A. Marques Pereira, G. Pasi (1999). On non-monotonic aggregation: Mixture operators. *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, pp 513-517.

[6] R. Mesiar, J. Špirková (2006). Weighted means and weighting functions, *Kybernetika* 42, 2, pp 151-160.

[7] R. Mesiar, J. Špirková, L. Vavríková (2008). Weighted aggregation operators based on minimization, *Information Sciences*, Volume 17, Issue 4, pp 1133-1140.

[8] R. A. Ribeiro, R. A. Marques Pereira (2003). Generalized mixture operators using weighting functions: A comparative study with WA and OWA, *European Journal of Operational Research*, 145, pp 329-342.

[9] R. A. Ribeiro, R. A. Marques Pereira (2003) Aggregation with generalized mixture operators using weighting functions, *Fuzzy sets and systems*, 137, pp 43-58.

[10] J. Špirková (2008). Dissertation thesis, Weighted aggregation operators and their applications, Bratislava.

[11] J. Špirková (2014). Weighted operators based on dissimilarity function. *Information Sciences*, Volume 281, pp 172-181.

[12] J. Špirková (2015). Induced weighted operators based on dissimilarity functions. *Information Sciences*, Volume 294, pp 530-539.

[13] T. Wilkin (2013). Image reduction operators based on no-monotonic averaging functions. In: Proceedings of the 10th IEEE Int. Conf. on Fuzzy Systems, pp 1-8.

[14] T. Wilkin, G. Beliakov, T. Calvo (2014). Weakly monotone averaging functions. *Information Processing and Management of Uncertainty in Knowledge-Based Systems Communications in Computer and Information Science*, Volume 444, pp 364-373.

[15] T. Wilkin, G. Beliakov (2014). Weakly monotonic averaging functions. *International Journal of Intelligent Systems*, Vol. 30, Issue 2, pp 144-169.

PRESERVATION OF THE EXCHANGE PRINCIPLE UNDER LATTICE OPERATIONS ON FUZZY IMPLICATIONS

Nageswara Rao Vemuri and Balasubramaniam Jayaram
Department of Mathematics
Indian Instititute of Technology Hyderabad
Yeddumailaram, India 502205.
{ma10p001,jbala}@iith.ac.in

Summary

In this work, we solve an open problem related to the exchange principle of fuzzy implications [Problem 3.1, Fuzzy Sets and Systems 261(2015) 112-123]. We show that two important generalizations of the exchange principle, namely, the generalized exchange principle(GEP) and the mutual exchangeability(ME) are sufficient conditions for the solution of the problem. We also show that, under some conditions, these are necessary too. Finally, we investigate the pairs (I, J) from different families of fuzzy implications such that the exchange principle is preserved under the join and meet operations.

Keywords: Fuzzy implication, the exchange principle, the generalized exchange principle, the mutual exchangeability, lattice operations.

1 INTRODUCTION

Fuzzy implications are one of the important logical connectives in fuzzy logic. These operators generalize the classical implication from $\{0, 1\}$-setting to the $[0, 1]$- setting. They are defined as follows:

Definition 1.1 ([1], Definition 1.1.1). *A function $I : [0, 1]^2 \longrightarrow [0, 1]$ is called a* fuzzy implication *if it satisfies, for all $x, x_1, x_2, y, y_1, y_2 \in [0, 1]$, the following conditions:*

$$\text{if } x_1 \leq x_2, \text{ then } I(x_1, y) \geq I(x_2, y), \tag{I1}$$

$$\text{if } y_1 \leq y_2, \text{ then } I(x, y_1) \leq I(x, y_2), \tag{I2}$$

$$I(0, 0) = 1,\ I(1, 1) = 1,\ I(1, 0) = 0. \tag{I3}$$

Let $\mathbb{I}$ denote the set of all fuzzy implications defined on $[0, 1]$. Fuzzy implications have many applications in fuzzy logic, approximate reasoning, decision making, fuzzy image processing, fuzzy control etc. Due to their applicational value it is always essential to generate fuzzy implications that satisfy various properties and functional equations.

However, it is always not straight-forward to generate fuzzy implications that preserve the desirable basic properties. For example, the lattice operations proposed by Bandler and Kohout as follows

$$(I \vee J)(x, y) := \max(I(x, y), J(x, y)),$$
$$(I \wedge J)(x, y) := \min(I(x, y), J(x, y)),$$

do not always preserve the exchange principle, which is defined as follows:

Definition 1.2 ([1], Definition 1.3.1). *A fuzzy implication I is said to satisfy the exchange principle* (EP)*, if for all $x, y, z \in [0, 1]$*

$$I(x, I(y, z)) = I(y, I(x, z)). \tag{EP}$$

For more about the lattice operations of fuzzy implications and the preservation of the basic properties, please see Chapter 6 of [1].

Thus this fact has become the main motivation to propose the following open problem.

Problem 1.3 ([3], Problem 3.1). *Characterize the subfamily of all fuzzy implications ((S,N)-implications, R-implications, etc.) which preserve* (EP) *for lattice operations.*

In this paper, we investigate the solutions of Problem 1.3 in a more general context, i.e., we attempt to characterize all fuzzy implications which preserve (EP) under the lattice operations.

Towards this end, in Section 2, we present some examples of $I, J \in \mathbb{I}$ such that $I \vee J$ and $I \wedge J$ preserve (EP) and investigate some basic conditions for a pair (I, J) to satisfy the same. We also recall some important generalizations of the exchange principle,

ISBN: 978–83–8012–519–3

viz., the generalized exchange principle(GEP) and the mutual exchangeability(ME). Following this, in Section 3, we show that either of (GEP) or (ME) is a sufficient condition for $I \vee J$ and $I \wedge J$ to preserve (EP). Later on, in Section 4, we show that the properties (GEP) and (ME) are also necessary under some conditions, namely, the Lattice Exchangeability Inequlaities (LEI). Finally, we present some results pertaining to the solutions of fuzzy implications I, J that satisfy (ME) and (GEP) separately in Sections 5 and 6, respectively.

2 PRELIMINARIES

In this section, we first show that there exist solutions of Problem 1.3. Following this, we investigate the basic characterizations of pairs (I, J) of fuzzy implications that become the solutions of Problem 1.3. Finally, we recall two important generalizations of (EP), namely, (GEP) and (ME), that will be helpful in obtaining the solutions of Problem 1.3.

Example 2.1. *(i) Let $I, J \in \mathbb{I}$ satisfy* (EP) *and $I \leq J$ under the usual point-wise ordering of functions. Clearly, $I \vee J = J$ and $I \wedge J = I$, which satisfy* (EP)*. Thus when I, J are comparable, $I \vee J$ and $I \wedge J$ always preserve* (EP)*.*

(ii) Let $I, J \in \mathbb{I}$ be defined as follows:

$$I(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ \sin(\frac{\pi y}{2}), & \text{if } x > 0, \end{cases}$$

$$\text{and } J(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ y^2, & \text{if } x > 0. \end{cases}$$

Then it is easy to see that the implications $I, J, I \vee J$, and $I \wedge J$ satisfy (EP)*. Note also that I, J are not comparable.*

In fact, one can generalize Example 2.1(ii) to obtain further solutions of Problem 1.3 as in the following.

Remark 2.2. *Let $I, J \in \mathbb{I}$ be defined as follows:*

$$I(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ \varphi(y), & \text{if } x > 0, \end{cases}$$

$$\text{and } J(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ \psi(y), & \text{if } x > 0, \end{cases}$$

where $\varphi, \psi : [0,1] \longrightarrow [0,1]$ are increasing bijections such that $\varphi(0) = 0 = \psi(0)$ and $\varphi(1) = 1 = \psi(1)$. Then it is easy to see that the implications $I, J, I \vee J, I \wedge J$ satisfy the exchange principle. In the case, if φ, ψ are incomparable then I, J are also incomparable.

From Example 2.1, it follows that lattice operations of comparable fuzzy implications always preserve (EP) and there exist some incomparable fuzzy implications whose lattice operations also preserve the same.

Now, in the following we present some important results that will be useful in the investigations of pairs (I, J) of fuzzy implications such that $I \vee J$ and $I \wedge J$ preserve (EP).

Proposition 2.3 ([1], Propositions 7.2.15 and 7.2.26)**.** *For a function $I\colon [0,1]^2 \to [0,1]$ the following statements are equivalent:*

(i) I is increasing in the second variable, i.e., I satisfies (I2)*.*

(ii) I satisfies $I(x, \min(y,z)) = \min(I(x,y), I(x,z))$ for all $x, y, z \in [0,1]$.

(iii) I satisfies $I(x, \max(y,z)) = \max(I(x,y), I(x,z))$ for all $x, y, z \in [0,1]$.

From the above result the following Lemma follows directly.

Lemma 2.4. *Let $I, J \in \mathbb{I}$ satisfy* (EP)*. Then the following statements are equivalent:*

(i) $A_i(I,J)$ satisfies (EP)*, where $A_1(I,J) = I \wedge J$ and $A_2(I,J) = I \vee J$.*

(ii) $A_i(I(x,I(y,z)), I(x,J(y,z)), J(x,I(y,z)), J(x,J(y,z))) = A_i(I(y,I(x,z)), I(y,J(x,z)), J(y,I(x,z)), J(y,J(x,z)))$, for $i = 1, 2$, where $A_1 = \min$ and $A_2 = \max$.

In the following, we recall two important generalizations of (EP) proposed in different contexts, which play an important role in the sequel.

Definition 2.5 (cf. [4], Proposition 5.5)**.** *A pair (I, J) of fuzzy implications is said to satisfy the* generalized exchange principle (GEP)*, if for all $x, y, z \in [0,1]$,*

$$\left.\begin{aligned} I(x, J(y,z)) = I(y, J(x,z)), \\ J(x, I(y,z)) = J(y, I(x,z)). \end{aligned}\right\} \qquad \text{(GEP)}$$

Remark 2.6. *Note that, in the original definition of* (GEP) *in [4], the pair (I, J) satisfies* (GEP) *if only the first of the above two conditions, viz., $I(x, J(y,z)) = I(y, J(x,z))$, is true. In that sense, given $I, J \in \mathbb{I}$, our definition requires both the pairs (I, J) and (J, I) to satisfy* (GEP)*. However, to avoid cumbersome repetitions, we continue to consider the definition given in Definition 2.5 in this work.*

Example 2.7. *Let $I, J \in \mathbb{I}$ be defined as follows:*

$$I(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ y^3, & \text{if } x > 0, \end{cases}$$

$$\text{and } J(x,y) = \begin{cases} 1, & \text{if } x = 0, \\ y^4, & \text{if } x > 0. \end{cases}$$

Then it is easy to see that the pair (I, J) *satisfies* (GEP).

Definition 2.8 ([8], Definition 3.9)**.** *A pair (I, J) of fuzzy implications is said to be* mutually exchangeable*, if for all* $x, y, z \in [0,1]$*,*

$$I(x, J(y,z)) = J(y, I(x,z)). \tag{ME}$$

From Remark 3.10 in [7], it follows that (GEP) is different from (ME).

3 SUFFICIENT CONDITIONS ON I, J SUCH that $I \vee J, I \wedge J$ PRESERVE (EP)

In this section, we show that either of (GEP) and (ME) is a sufficient condition for a pair (I, J) to be a solution of Problem 1.3.

Theorem 3.1. *Let* $I, J \in \mathbb{I}$ *satisfy* (EP)*. If the pair* (I, J) *satisfies either* (GEP) *or* (ME)*, then both* $I \vee J$ *and* $I \wedge J$ *satisfy* (EP)*.*

Proof. Let $I, J \in \mathbb{I}$ satisfy (EP).

(i) Let the pair (I, J) satisfy (GEP). Let $K_1 = I \vee J$ amd $x, y, z \in [0,1]$. Now, from (EP), (GEP) of I, J and Lemma 2.4, it follows that

$$\begin{aligned} K_1(x, K_1(y,z)) &= \max(I(x, I(y,z)), I(x, J(y,z)), \\ &\quad J(x, I(y,z)), J(x, J(y,z))) \\ &= \max(I(y, I(x,z)), I(y, J(x,z)), \\ &\quad J(y, I(x,z)), J(y, J(x,z))) \\ &= K_1(y, K_1(x,z)), \end{aligned}$$

or equivalently, $K_1 = I \vee J$ satisfies (EP). Similarly, one can show that $I \wedge J$ also satisfies (EP).

(ii) Let the pair (I, J) satisfy (ME). Now, from (ME) it follows that

$$I(y, J(x,z)) = J(x, I(y,z)), \qquad x, y, z \in [0,1].$$

Now let $x, y, z \in [0,1]$. Then, once again, by using (EP) and (ME) of I, J and Lemma 2.4, we get

$$\begin{aligned} K_1(x, K_1(y,z)) &= \max(I(x, I(y,z)), I(x, J(y,z)), \\ &\quad J(x, I(y,z)), J(x, J(y,z))) \\ &= \max(I(y, I(x,z)), J(y, I(x,z)), \\ &\quad I(y, J(x,z)), J(y, J(x,z))) \\ &= K_1(y, K_1(x,z)). \end{aligned}$$

Thus $K_1 = I \vee J$ satisfies (EP). Similarly, one can show that $I \wedge J$ also satisfies (EP).

□

4 NECESSARY CONDITIONS ON I, J SUCH THAT $I \vee J, I \wedge J$ PRESERVE (EP)

In Theorem 3.1, we have shown that either (GEP) and (ME) of I, J is a sufficient condition for $I \vee J$ and $I \wedge J$ to preserve (EP). In this section, we show that these properties also become necessary under some conditions.

Towards this end, we define the following:

Definition 4.1. *Let* $I, J \in \mathbb{I}$ *satisfy* (EP)*. Then we say that the pair* (I, J) *satisfies* ***Lattice Exchangeable Inequalities*** *(LEI) if it satisfies the following inequalities: For all* $x, y, z \in [0,1]$*,*

$$\max(I(x, I(y,z)), J(x, J(y,z))) \leq \max(I(x, J(y,z)), J(x, I(y,z))), \tag{LEI-1}$$

$$\min(I(x, I(y,z)), J(x, J(y,z))) \geq \min(I(x, J(y,z)), J(x, I(y,z))). \tag{LEI-2}$$

Example 4.2. *It can be easily verified that the pair of fuzzy implications* (I_1, J_1) *does satisfy the (LEI) inequalities, while the pair* (I_2, J_2) *does not:*

$$I_1(x,y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } y = 1, \\ 0, & \text{if } x = 1 \text{ and } y = 0, \\ 0.4, & \text{otherwise}, \end{cases}$$

$$J_1(x,y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } y = 1, \\ 0, & \text{if } x = 1 \text{ and } y = 0, \\ 0.6, & \text{otherwise}, \end{cases}$$

$$I_2(x,y) = \begin{cases} 1, & \text{if } x \leq 0.4, \\ y^2, & \text{if } x > 0.4, \end{cases}$$

$$J_2(x,y) = \begin{cases} 1, & \text{if } x \leq 0.6, \\ y^4, & \text{if } x > 0.6. \end{cases}$$

Lemma 4.3. *Let* $I, J \in \mathbb{I}$ *satisfy* (EP)*.* (LEI-1) *is equivalent to* (LEI-1')*:*

$$\max(I(x, I(y,z)), J(x, J(y,z))) \leq \max(I(y, J(x,z)), J(y, I(x,z))), \tag{LEI-1'}$$

and (LEI-2) *is equivalent to* (LEI-2')*:*

$$\min(I(x, I(y,z)), J(x, J(y,z))) \geq \min(I(y, J(x,z)), J(y, I(x,z))). \tag{LEI-2'}$$

Proof. Let $I, J \in \mathbb{I}$ satisfy (EP). In the following we show that (LEI-1) is equivalent to (LEI-1'), since the proof for the other can be similarly obtained.

(LEI-1) $\implies$ (LEI-1'): Let the pair (I, J) satisfy (LEI-1). From (LEI-1), one can always write

$$\max(I(y, I(x, z)), J(y, J(x, z))) \leq \max(I(y, J(x, z)), J(y, I(x, z))). \quad (1)$$

Since I, J satisfy (EP), the inequality (1) becomes

$$\max(I(x, I(y, z)), J(x, J(y, z))) \leq \max(I(y, J(x, z)), J(y, I(x, z))),$$

which is equal to (LEI-1').

(LEI-1') $\implies$ (LEI-1): Follows similarly. □

Theorem 4.4. *Let $I, J, I \vee J, I \wedge J \in \mathbb{I}$ satisfy* (EP). *If the pair (I, J) satisfies (LEI) then it also satisfies the following equations:*

$$\max(I(x, J(y, z)), J(x, I(y, z))) = \max(I(y, J(x, z)), J(y, I(x, z))), \quad (2)$$

$$\min(I(x, J(y, z)), J(x, I(y, z))) = \min(I(y, J(x, z)), J(y, I(x, z))). \quad (3)$$

Proof. Let $I, J, I \vee J, I \wedge J \in \mathbb{I}$ satisfy (EP). Let the pair (I, J) also satisfy (LEI). In the following, we prove only the equation (2), since the proof for (3) can be obtained similarly. Let $x, y, z \in [0, 1]$. Since $I \vee J$ satisfies (EP), from Lemma 2.4, the pair (I, J) satisfies the equation in Lemma 2.4(ii), with $i = 2$. Thus we have, for all $x, y, z \in [0, 1]$,

$$\max\Big\{I(x, I(y, z)), I(x, J(y, z)), J(x, I(y, z)), J(x, J(y, z))\Big\} = \max\Big\{I(y, I(x, z)), I(y, J(x, z)), J(y, I(x, z)), J(y, J(x, z))\Big\}. \quad (*)$$

Since the pair (I, J) also satisfies (LEI), from (LEI-1), we get

L.H.S. of $(*) = \max(I(x, J(y, z)), J(x, I(y, z)))$ $=$ **L.H.S.** of (2).

Since I, J satisfy (EP) from Lemma 4.3, it follows that (LEI-1) is equivalent to (LEI-1'), from whence we obtain that

R.H.S. of $(*) = \max(I(y, J(x, z)), J(y, I(x, z)))$ $=$ **R.H.S.** of (2).

□

Let $I \vee J, I \wedge J$ preserve (EP) and I, J satisfy (LEI). Then from Theorem 4.4, it follows that the pair (I, J) satisfies (2) and (3). In other words, this fact implies that the solutions of the equations (2), (3) also become the solutions of Problem 1.3. In the following, we investigate the solutions of (2) and (3). Before doing so, we recall the following important result which is useful in the sequel.

Lemma 4.5 ([2], page. 366)**.** *Let L be any distributive lattice. Let $a, b, c \in L$ satisfy*

$$\max(a, b) = \max(a, c), \quad (4)$$
$$\min(a, b) = \min(a, c). \quad (5)$$

Then $b = c$.

Remark 4.6. *Since $([0, 1], \leq, \vee, \wedge)$ is also a distributive lattice, Lemma 4.5 is also true for all $a, b, c \in [0, 1]$.*

Remark 4.7. *Let $a, b, c, d \in [0, 1]$ satisfy*

$$\max(a, b) = \max(c, d), \quad (6)$$
$$\min(a, b) = \min(c, d). \quad (7)$$

Then either $a = c$ or $a = d$. Further,

(i) if $a = c$ then $b = d$.

(ii) if $a = d$ then $b = c$.

Lemma 4.8. *Let the pair $(I, J) \in \mathbb{I}$ satisfy the equations* (2) *and* (3)*. Then it satisfies either* (GEP) *or* (ME).

Proof. Follows from Remark 4.7. □

Theorem 4.9. *Let $I, J, I \vee J, I \wedge J \in \mathbb{I}$ satisfy* (EP) *and let the pair (I, J) satisfy (LEI). Then the pair (I, J) satisfies either* (GEP) *or* (ME).

Proof. Follows from Theorem 4.4 and Lemma 4.8. □

Theorem 4.10. *Let $I, J \in \mathbb{I}$ satisfy* (EP) *and (LEI). Then the following statements are equivalent:*

(i) $I \vee J, I \wedge J$ satisfy (EP).

(ii) The pair (I, J) satisfies either (GEP) *or* (ME).

Proof. Follows from Theorems 4.9 and 3.1. □

Remark 4.11. *Let $I, J \in \mathbb{I}$ satisfy* (EP). *If the pair (I, J) satisfies (LEI) then from Theorem 4.10, it follows that $I \vee J, I \wedge J$ satisfy* (EP). *However, the converse need not be true. For example, take $I = I_2$ and $J = J_2$ of Example 4.2.*

Since (ME) and (GEP) play an important role in the characterizations of solutions of Problem 1.3, it is of interest to know the pairs (I, J) of fuzzy implications that do satisfy (ME) or (GEP). We take up this investigation in the following sections.

5 PAIRS OF FUZZY IMPLICATIONS SATISFYING (ME)

Due to the variety of fuzzy implications and the complexity of the functional equation, it is not an easy task to investigate the pairs of fuzzy implications that do satisfy (ME). However, Vemuri [5] has investigated the solutions of (ME), but only for the families of fuzzy implications whose characterizations are well established. In the following, we recall some of the most important results that give the solutions of (ME) and thus the solutions of Problem 1.3. For details about definitions, properties, characterizations and representations of different families of fuzzy implications, please see [1].

5.1 (S, N)-IMPLICATIONS SATISFYING (ME)

Proposition 5.1 ([5], Proposition 4.1)**.** *Let I be an (S, N)-implication whose negation N has trivial range, i.e., $N(x) \in \{0, 1\}$ for all $x \in [0, 1]$. Then I satisfies* (ME) *with every $J \in \mathbb{I}$.*

From Proposition 5.1, it follows that if at least one of I, J is an (S, N)-implication with trivial range negation N then the pair (I, J) satisfies (ME) and hence becomes the solution of Problem 1.3.

In the case if I, J are two (S, N)-implications with continuous negations and satisfy (ME) then as the following result suggests the two t-conorms involved in the definition must be the same.

Theorem 5.2 ([7], Theorem 6.7)**.** *Let $I(x, y) = S_1(N_1(x), y), J(x, y) = S_2(N_2(x), y)$ be two (S, N)-implications such that N_1, N_2 are continuous negations. Then the following statements are equivalent:*

(i) The pair (I, J) satisfies (ME).

(ii) $S_1 = S_2$.

5.2 R-IMPLICATIONS SATISFYING (ME)

Theorem 5.3 ([5], Theorem 5.1)**.** *Let $I = I_{T_1}$ and $J = I_{T_2}$ be two R-implications generated from left-continuous t-norms T_1, T_2 respectively. Then the following statements are equivalent:*

(i) The pair (I, J) satisfies (ME).

(ii) $I = J$.

Before presenting the solutions of f and g-implications that satisfy (ME), we recall two important definitions that will be useful in the sequel.

Definition 5.4 ([6, 7])**.** *For any $I, J \in \mathbb{I}$, we define $I \circledast J \colon [0, 1]^2 \to [0, 1]$ as*

$$(I \circledast J)(x, y) = I(x, J(x, y)), \qquad x, y \in [0, 1].$$

Definition 5.5 ([8], Definition 5.1)**.** *Let $I \in \mathbb{I}$. For any $n \in \mathbb{N}$, we define the n-th power of I w.r.t. the binary operation $\circledast$ as follows: For $n = 1$,*

$$I_{\circledast}^{[n]} = I,$$

and for $n \geq 2$,

$$I_{\circledast}^{[n]}(x, y) = I\left(x, I_{\circledast}^{[n-1]}(x, y)\right) = I_{\circledast}^{[n-1]}(x, I(x, y)),$$

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

5.3 f-IMPLICATIONS SATISFYING (ME)

Theorem 5.6 ([5], Theorem 6.6)**.** *Let I, J be two f-implications. Then the following statements are equivalent:*

(i) The pair (I, J) satisfies (ME).

(ii) $J = I_{\circledast}^{[n]}$ for some $n \in \mathbb{N}$.

5.4 g-IMPLICATIONS SATISFYING (ME)

Theorem 5.7 ([5], Theorem 7.4)**.** *Let I, J be two g-implications. Then the following statements are equivalent:*

(i) The pair (I, J) satisfies (ME).

(ii) $J = I_{\circledast}^{[n]}$ for some $n \in \mathbb{N}$.

6 PAIRS OF FUZZY IMPLICATIONS SATISFYING (GEP)

In this section, we attempt to find the pairs (I, J) of fuzzy implications that do satisfy (GEP). Once again keeping the complexity of the functional equation (GEP) in mind, we restrict ourselves to do so for the families (S, N)-, R-, f- and g- of fuzzy implications.

Note that all of these families of fuzzy implications satisfy the following left neutrality property (NP):

Definition 6.1 (cf. [1], Definition 1.3.1). *An $I \in \mathbb{I}$ is said to satisfy the left neutrality property* (NP) *if*

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

Lemma 6.2. *Let $I, J \in \mathbb{I}$ satisfy* (NP). *If the pair (I, J) satisfies* (GEP) *then $I = J$.*

Proof. The substitution of $x = 1$ in (GEP) and (NP) of $I, J \in \mathbb{I}$ will yield $I = J$. □

From the above results, it is clear that if $I, J \in \mathbb{I}$ belong to one of the following families of fuzzy implications, viz.,(S, N)-, R-, f-, g- implications, and satisfy (GEP), then $I = J$ and hence it trivially follows that both $I \vee J$ and $I \wedge J$ preserve (EP).

7 CONCLUSIONS

In this paper, we have investigated the solutions of an open problem [Problem 3.1, Fuzzy Sets and Systems 261(2015) 112-123] related to the preservation of the exchange principle (EP) of fuzzy implications under lattice operations. Our study has shown the importance of two of the generalizations of (EP), viz., (GEP) and (ME) in obtaining the solutions of the problem.

While (GEP), (ME) are independently sufficient for the lattice operations of fuzzy implications to preserve (EP), these conditions are not necessary. However, the newly proposed pair of inequalities, namely the Lattice Exchangeable Inequalities (LEI-1) and (LEI-2) make (GEP) and (ME) also a necessity for a pair of fuzzy implications to be a solution of Problem 1.3.

Since the pairs (I, J) of fuzzy implications satisfying either (GEP) or (ME) are the most general solutions of the problem, we have investigated them but for some well known families of fuzzy implications. However, this problem has to be investigated in the most general setting. Further, the solutions of (LEI) are worthy of study. We intend to explore these in detail in the near future.

References

[1] Baczyński, M., Jayaram, B.: Fuzzy implications (Studies in Fuzziness and Soft Computing, vol. 231), Springer, Berlin 2008.

[2] Birkhoff, G., Mac Lane, S.: A Survey of Modern Algebra. Universities Press 1997.

[3] Mesiar, R., Stupňanovà, A.: Open problems posed at International Conference on Fuzzy Sets : Theory and Its Applications (FSTA-2014). Fuzzy Sets and Systems 261(2015) 112-123.

[4] Reiser, R.H., Bedregal, B. and Baczyński, M.: Aggregating fuzzy implications. Information Sciences 253(2013) 126 – 146.

[5] Vemuri, N. R.: Mutually exchangeable fuzzy implications. Accepted for publication in Information Sciences. doi:10.1016/j.ins.2015.04.038

[6] Vemuri, N.R., Jayaram, B.: Fuzzy implications: Novel generation process and the consequent algebras, in: Advances on Computational Intelligence II (IPMU 2012), 365 – 374.

[7] Vemuri, N. R., Jayaram, B.: Representations through a monoid on the set of fuzzy implications. Fuzzy Sets and Systems 247(2014) 51–67.

[8] Vemuri, N. R., Jayaram, B.: The ⊛-composition of fuzzy implications: Closures with respect to properties, powers and families. Accepted for publication in Fuzzy Sets and Systems. doi:10.1016/j.fss.2014.10.004

ON INFORMATION RECEPTION AND SOME OF ITS AGGREGATION OPERATORS

Doretta Vivona and Maria Divari
"Sapienza" Universitá di Roma,
Dip.SBAI, via A.Scarpa n.16
00161 Roma (Italy)
{doretta.vivona,maria.divari}@sbai.uniroma1.it

Summary

The aim of this paper is to present a class of measures for information reception on crisp and fuzzy setting. Then, some aggregation operators for this reception are proposed.

Keywords: Crisp sets, Fuzzy sets, Information measure, Aggregation operators.

1 INTRODUCTION

The goal of this paper is to introduce a class of measures of *information reception* on crisp and fuzzy setting. The structure of the paper is the following: in Section 2 we recall some preliminaries about general information given by Kampé De Fériet and Forte for crisp sets and by Benvenuti and ourselves for fuzzy sets. This information is called general because it is defined without probability or fuzzy measure. In Section 3 we introduce the definition of information reception. In Section 4 we present an hypothesis for measure of information reception with its properties expressed in a system of functional equations and inequalities. This system is solved in Section 5. Finally, in Section 6, we give some forms of aggregation operators for information reception. Section 7 is devoted to the conclusions.

2 PRELIMINARIES

In this paragraph, we present some preliminaires on crisp and fuzzy setting.

2.1 CRISP SETTING

Let Ω be an abstract space and $\mathcal{A}$ a $\sigma-$algebra of all subsets of Ω, such that $(\Omega, \mathcal{A})$ is measurable. We refer to [6] for all knowledge about crisp sets.

Following Kampé de Fériet, Forte and Benvenuti [10, 7, 9, 5, 8], we recall that measure $J(\cdot)$ of general information (i.e. without probability) is a mapping

$$J(\cdot) : \mathcal{A} \to [0, +\infty]$$

such that $\forall\ A, A' \in \mathcal{A}$:

(i) $\quad A^{'} \supset A \Rightarrow J(A^{'}) \leq J(A)$,

(ii) $\quad J(\emptyset) = +\infty$,

(iii) $\quad J(\Omega) = 0$.

2.2 FUZZY SETTING

Let X be an abstract space and $\mathcal{F}$ a $\sigma-$algebra of all fuzzy sets of X, such that $(X, \mathcal{F})$ is measurable. We refer to [19, 12] for all knowledge about fuzzy sets.

In 1990, in [1], we introduced measure of general information $\hat{J}(\cdot)$ (i.e. without probability or fuzzy measure) on fuzzy setting as a mapping

$$\hat{J}(\cdot) : \mathcal{F} \to [0, +\infty]$$

such that $\forall\ F, F' \in \mathcal{F}$:

$(\hat{i})$ $\quad F^{'} \supset F \Rightarrow \hat{J}(F^{'}) \leq \hat{J}(F)$,

$(\hat{ii})$ $\quad \hat{J}(\emptyset) = +\infty$,

$(\hat{iii})$ $\quad \hat{J}(X) = 0$.

3 STATEMENT OF THE PROBLEM

3.1 CRISP AND FUZZY SETTING

Given an information J on $\mathcal{A}$ or $\hat{J}$ on $\mathcal{F}$, we can ask to us if it has been received. In fact, it is possible that information has not been received well, for this reason we introduce the *reception of information* J (or *information reception*) of a crisp event $A : R(J(A))$, or *information reception* $\hat{J}$ of a fuzzy set $F : R(\hat{J}(F))$. Shortly, from now on, we put $R(J(A)) = R_J(A)$, $R(\hat{J}(F)) = R_{\hat{J}}(F)$, respectively .

Moreover, if information has not been received or it has been received completely, we say, in the first case, that information reception of an event is null, in the second one we assume that the information reception of an event gives us the whole information of the event itself. Finally, we think that information reception is an increasing function with respect to information of the given set.

Let $\mathcal{J}$ be the family of information measures on crisp setting as in Sect.2.1.

So, we introduce the following:

Definition 3.1. Given a set $A \in \mathcal{A}$ and an information measure $J \in \mathcal{J}$, *information reception* R_J is a mapping

$$R_J : \mathcal{J} \to [0, \text{J(A)}]$$

such that $\forall J, J' \in \mathcal{J}$:

$$J \leq J' \Rightarrow R_J(A) \leq R_{J'}(A).$$

It is clear that:

$$R_J(\Omega) = J(\Omega) = 0, \; R_J(\emptyset) = J(\emptyset) = +\infty.$$

On fuzzy setting the meaning of information reception is the same.

Let $\hat{\mathcal{J}}$ be the family of information measure on fuzzy setting as in Sect.2.2.

Definition 3.2. Given a set $F \in \mathcal{F}$ and an information measure $\hat{J} \in \hat{\mathcal{J}}$, *information reception* $R_{\hat{J}}$ is a mapping

$$R_{\hat{J}} : \hat{\mathcal{J}} \to [0, \hat{\text{J}}(\text{F})]$$

such that $\forall \hat{J}, \hat{J}' \in \hat{\mathcal{J}}$:

$$\hat{J} \leq \hat{J}' \Rightarrow R_{\hat{J}}(F) \leq R_{\hat{J}'}(F).$$

It is clear that:

$$R_{\hat{J}}(X) = \hat{J}(X) = 0, \; R_{\hat{J}}(\emptyset) = \hat{J}(\emptyset) = +\infty.$$

4 THE FUNCTIONS Ψ and Φ

In this paragraph we should like to find some particular forms of information reception. Given an information, we think that information reception is linked not only to information but also to the degree of accepting of reception which we express through a parameter $k, k \in [0, 1]$. Maximum value of k is $k = 1$ and minimum $k = 0$. For the meaning of information and of k we must exclude the cases:

$$k = 0 \;\text{ when } J(A) = +\infty, \; \forall \; A \in \mathcal{A}, \tag{1}$$

$$k = 0 \;\text{ when } \hat{J}(F) = +\infty, \; \forall \; F \in \mathcal{F}.$$

4.1 CRISP SETTING

For reasons expressed above, given $A \in \mathcal{A}$, we are going to look for information reception as a function Ψ depending on information $J(A)$ and the parameter k :

$$R_J(A) = \Psi\Big(k, J(A)\Big), \tag{2}$$

where $\Psi : [0, 1] \times [0, +\infty] \to [0, +\infty]$, and continuous.

From the previous properties, $\forall \; J, J' \in \mathcal{J}$ and $\forall \; k, k' \in [0, 1]$ the continuous function Ψ is defined axiomatically in the following way:

p_1. $\Psi(0, J(A)) = 0$,

p_2.$\Psi(1, J(A)) = J(A)$,

p_3. $k \leq k' \Longrightarrow \Psi(k, J(A)) \leq \Psi(k', J(A))$,

p_4. $J(A) = 0 \Longrightarrow \Psi(k, 0) = 0$,

p_5. $J(A) = +\infty$ and $k \neq 0 \Longrightarrow \Psi(k, +\infty) = +\infty$,

p_6. $J \leq J' \Longrightarrow \Psi(k, J(A)) \leq \Psi(k, J'(A))$.

It is clear that p_5. translates the condition (1).

Now, we are going to introduce the properties above $[p_1. - p_6.]$ in a system of functional equations and inequalities: setting $J(A) = x, J'(A) = x', \; x, x' \in [0, +\infty], \; k, k' \in [0, 1]$ we have

$$\begin{cases} P_1. \; \Psi(0, x) = 0, \\ P_2. \; \Psi(1, x) = x, \\ P_3. \; \Psi(k, x) \leq \Psi(k', x) \quad\; k \leq k', \\ P_4. \; \Psi(k, 0) = 0, \\ P_5. \; \Psi(k, +\infty) = +\infty \quad\; k \neq 0, \\ P_6. \; \Psi(k, x) \leq \Psi(k, x') \quad\; x \leq x'. \end{cases}$$

We are looking for an universal continuous solution of the system above in ever proper space.

4.2 FUZZY SETTING

In analogous way, given a fuzzy set F, we propose that information reception $R_{\hat{J}}$ of is a function Φ of $\hat{J}(F)$ and k, where the parameter k has the same meaning as in the crisp case:

$$R_{\hat{J}}(F) = \Phi\Big(k, \hat{J}(F)\Big), \tag{3}$$

where $\Phi : [0, 1] \times [0, +\infty] \to [0, +\infty]$, and continuous and it satisfies the conditions $[p_1. - p_6.]$.

5 THE SOLUTIONS

5.1 CRISP SETTING

On a crisp setting, we are going to consider the solution Ψ for separation of variables, as a product of two function $\alpha(k)$ and $f(x)$:

$$\Psi(k,x) = \alpha(k) \cdot f(x), \qquad (4)$$

with $\alpha : [0,1] \to [0,1]$ and $f : [0,+\infty] \to [0,+\infty]$. By substituting (4) in the system, the conditions P_1. and P_3. give $\alpha(0) = 0$ and α increasing, respectively. For $P_2., \alpha(1)f(x) = x \Longleftrightarrow \alpha(1) = 1$ and $f(x) = x$. If $\alpha(k)$ is continuous, a function as been found satisfies also $[P_4. - P_6.]$.

So, we have the following

Proposition 5.1. *The function $\Psi(k,x) = \alpha(k) \cdot f(x)$ is a solution of the system $[P_1. - P_6.]$ if and only if*

a)$f(x) = x,\ x \in [0,+\infty]$,

b)$\alpha : [0,1] \to [0,1]$, *increasing, continuous,* $\alpha(0) = 0,\ \alpha(1) = 1$,

i.e.

$$\Psi(k,x) = \alpha(k)x, \quad x \in [0,+\infty], \qquad (5)$$

with the condition b).

Proof. It is easy to see that the function (5) with the condition *b*) is solution of the system $[P_1. - P_6.]$. The viceversa has been shown above. □

Moreover, it is possible to generalize the previous result in

Proposition 5.2. *A class of solutions of the system $[P_1. - P_6.]$ are the functions*

$$\Psi_h(k,x) = h^{-1}\Big(\alpha(k)h(x)\Big), \qquad (6)$$

if and only if

a) $h : [0,+\infty] \to [0,+\infty]$, *strictly increasing, continuous with* $h(0) = 0,\ h(+\infty) = +\infty$,

b) $\alpha : [0,1] \to [0,1]$, *increasing, continuous, with* $\alpha(0) = 0,\ \alpha(1) = 1$.

For $h(x) = x$, we have Proposition 5.1.

5.2 FUZZY SETTING

In a fuzzy setting, we are going to find the function Φ solution of the system $[P_1. - P_6.]$. The previous Propositions 5.1 and 5.2 are valid also in a fuzzy setting:

Proposition 5.3. *The function $\Phi(k,x) = \alpha(k) \cdot f(x)$ is a solution of the system $[P_1. - P_6.]$ if and only if*

a)$f(x) = x, \quad x \in [0,+\infty]$,

b) $\alpha : [0,1] \to [0,1]$, *increasing, continuous,* $\alpha(0) = 0,\ \alpha(1) = 1$,

i.e.

$$\Phi(k,x) = \alpha(k)x, \quad x \in [0,+\infty], \qquad (7)$$

with the condition b).

Proposition 5.4. *A class of solutions of the system $[P_1. - P_6.]$ are the functions*

$$\Phi_g(k,x) = g^{-1}\Big(\alpha(k)g(x)\Big), \qquad (8)$$

if and only if

a) $g : [0,+\infty] \to [0,+\infty]$, *strictly increasing, continuous, with* $g(0) = 0, g(+\infty) = +\infty$,

b) $\alpha : [0,1] \to [0,1]$ *increasing, continuous,* $\alpha(0) = 0,\ \alpha(1) = 1$.

6 SOME AGGREGATION OPERATORS

Many authors have introduced the aggregation operators L, with the properties of idempotence, monotonicity and continuity from below [3, 4, 11]. We have introduced aggregation operators in particular cases [13, 14, 15, 16, 18].

6.1 CRISP SETTING

Given an information J and n crisp events $A_i, i = 1, ..., n$ with their information receptions $R_J(A_i), i = 1, ..., n$ following the procedure presented in [2], it is possible to obtain the following result:

Proposition 6.1. *Some aggregation operators of n crisp events $A_i, i = 1, ..., n$ with information reception $a_i = R_J(A_i), i = 1, ..., n$ are*

$$L\Big(a_1, ..., a_i, ..., a_n\Big) = \bigvee_{i=1}^{n} a_i,$$

$$L\Big(a_1, ..., a_i, ..., a_n\Big) = \bigwedge_{i=1}^{n} a_i,$$

$$L\Big(a_1, ..., a_i, ..., a_n\Big) = \frac{h^{-1}\Big(\sum_{i=1}^{n} h(a_i)\Big)}{n},$$

with $h : [0,1] \to [0,M]$, $(0 \le M \le +\infty)$ *continuous, strictly increasing,* $h(0) = 0, h(1) = M$.

6.2 FUZZY SETTING

In analogous way, given an information $\hat{J}$, m fuzzy events $F_l, l = 1, ...m$, with information receptions $R_{\hat{J}}(F_l), l = 1, ..., m$, we obtain the same result:

Proposition 6.2. *Some aggregation operators of m fuzzy events $F_l, l = 1, ..., m$ with information reception $b_l = R_{\hat{J}}(F_l), l = 1, ..., m$ are*

$$L\Big(b_1, ..., b_l, ..., b_m\Big) = \bigvee_{l=1}^{m} b_l,$$

$$L\Big(b_1, ..., b_l, ..., b_m\Big) = \bigwedge_{l=1}^{m} b_l,$$

$$L\Big(b_1, ..., b_l, ..., b_m\Big) = \frac{h^{-1}\Big(\sum_{l=1}^{m} h(b_l)\Big)}{m},$$

with $h : [0,1] \to [0,M]$ $(0 \leq M \leq +\infty)$ continuous, strictly increasing, $h(0) = 0, h(1) = M$.

7 CONCLUSIONS

In this paper, for crisp and fuzzy sets, we have introduced information reception depending on a given information measure and a parameter: we have found some classes of reception, from Propositions 5.1 and 5.2:

$$R_J(A) = \alpha(k)J(A), \forall\ A \in \mathcal{A},$$

with $\alpha : [0,1] \to [0,1]$, increasing, continuous, $\alpha(0) = 0,\ \alpha(1) = 1$, and

$$R_{J,h}(k, J(A)) = h^{-1}\Big(\alpha(k)h(J(A))\Big),$$

with $h : [0,+\infty] \to [0,+\infty]$, strictly increasing, continuous, $h(0) = 0, h(+\infty) = +\infty$, and $\alpha : [0,1] \to [0,1]$ increasing, continuous, $\alpha(0) = 0,\ \alpha(1) = 1$.

From Propositions 5.1 and 5.2:

$$R_{\hat{J}}(F) = \alpha(k)\hat{J}(F), \forall\ F \in \mathcal{F}$$

$\alpha : [0,1] \to [0,1],\ \alpha(0) = 0,\ \alpha(1) = 1$, increasing, continuous, and

$$\hat{R}_{\hat{J},g}(k, \hat{J}(F)) = g^{-1}\Big(\alpha(k)g(\hat{J}(F))\Big),$$

with $g : [0,+\infty] \to [0,+\infty]$, strictly increasing, continuous, $g(0) = 0, g(+\infty) = +\infty,\ \alpha : [0,1] \to [0,1]$, increasing, continuous, $\alpha(0) = 0,\ \alpha(1) = 1$.

Then, we have presented some aggregation operators of these information receptions.

Acknowledgements

The first author wishes to thank the research centre CRITEVAT of "Sapienza"-University of Rome and GNFM of MIUR (Italy).

References

[1] P. Benvenuti, D. Vivona, M. Divari (1990). A general information for fuzzy sets. *Uncertainty in Knowledge Bases, Lectures Notes in Computer Sciences* **521**, pp. 307-316.

[2] P. Benvenuti, D. Vivona, M. Divari (2001). Aggregation operators and associated fuzzy mesures. *International Journal of Uncertainty, Fuzziness and Knownledge-Based Systems* **2**, pp. 197-204.

[3] T. Calvo, G. Mayor, R. Mesiar (2002). *Aggregation Operators. New Trend and Applications.* Physisc-Verlag, Heidelberg.

[4] M. Grabish, G. Mayor, R. Mesiar, E. Pap (2009). *Aggregation Functions.* Cambridge University Press. Cambridge.

[5] B. Forte (1969). Measures of information: the general axiomatic theory. *Revue de Information et Recherche Operative (R.A.I.R.O.)* **R3**, pp. 63-90.

[6] P.R. Halmos (1969). *Measure Theory.* Van Nostrand Company. New Jork.

[7] J. Kampé De Fériet (1969). Mesures de l'information par un ensamble d'observateures. *Comptes Rendus de l'Académie des Sciences Paris* **269 A**, pp. 1081-1086.

[8] J. Kampé De Fériet (1970). Mesures de l'information fornie par un evénement. *Coll.Int.Centre Nat.de la Rech. Scien.* **186**, pp. 191-221.

[9] J. Kampé De Fériet, P. Benvenuti (1969). Sur une classe d'informations. *Comptes Rendus de l'Académie des Sciences Paris* **269A**, pp. 97-101.

[10] J. Kampé De Fériet, B. Forte (1967). Information et Probabilité. *Comptes Rendus de l'Académie des Sciences Paris* **265**, pp. 110-114, 142-146, 350-353.

[11] E.P. Klement (2005). Binary aggregation operators which are bounded by the minimum (a mathematical overview). *Proceedings of the 3rd International Summer School on Aggrgation Operators and their applications (AGOP'05)* pp. 13-16.

[12] G.J. Klir, T.A. Folger (1988). *Fuzzy sets, Uncertainty, and Information.* Prentice-Hall International editions. New Jork.

[13] D. Vivona, M. Divari (2008). Aggregation operators for conditional information without probability. *Proceedings of Information Processing and*

management of Uncertainty in Knowledge-Based Systems (IPMU08), pp. 258-260.

[14] D. Vivona, M. Divari (2009). Aggregation operators for conditional crispness. *Proceeding of International Fuzzy Systems and Applications-European Society for Fuzzy Logic and Tecnology (IFSA-EUSFLAT2009)*, pp. 88-91.

[15] D. Vivona, M. Divari (2011). Aggregation operators of general aimed informations. *Proceeding of European Society for Fuzzy Logic and Tecnology - Logic Floue et ses Applications (EUSFLAT-LFA2011)*, pp. 75-78.

[16] D. Vivona, M. Divari (2012). Aggregation operators for the Awareness of General Information. *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU2012)* Part III, CCSI 299, pp. 271-275.

[18] D. Vivona, M. Divari (2014). Aggregation operators of General Misinformation. *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU2014)* Part III, CCIS 444, pp. 559-564.

[19] L.A. Zadeh (1965). Fuzzy sets. *Information and Control* **8**, pp. 338-353.

INTERVAL-VALUED AGGREGATION AS A TOOL TO IMPROVE MEDICAL DIAGNOSIS

Andrzej Wójtowicz, Patryk Żywica, Anna Stachowiak and Krzysztof Dyczkowski
Faculty of Mathematics and Computer Science
Adam Mickiewicz University in Poznań
Umultowska 87, 61-614 Poznań, Poland
min@amu.edu.pl

Summary

In the paper we present experimental results on the problem of an effective decision making on incomplete data. In order to investigate this problem we examined a variety of interval aggregation methods. Exemplary results are based on a medical diagnosis support system. Our research shows that an application of the aggregation in this problem leads to promising results.

Keywords: interval-valued aggregation, missing data, incomplete information, decision making under uncertainty, supporting medical diagnosis.

1 INTRODUCTION

According to recent statistics, the annual numbers of deaths due to ovarian cancer in some countries are alarmingly high and still grow. The correct diagnosis became a serious problem that the medicine is trying to face. The correct classification of a tumor as malignant or benign is particularly important since a given type of the tumor determines whether the patient must undergo a surgery. Moreover, incorrect indication of malignant tumor as benign, in the longer term causes deterioration of the patient's health and results in a high risk of failure of the surgery.

For this reason, a wide range of preoperative diagnostic models have been developed, where the goal is to predict the type of malignancy. Both the sensitivity and specificity of the models rarely exceeds 90% in external evaluation [14, 8]. The Table 1 presents six most common ones (two based on scoring systems [1, 11] and four based on logistic regressions [13, 12, 7]) and a list of used attributes. The attributes are divided into two groups: objective medical history and the rest (which are subjective medical history, ultrasound and blood markers).

Our previous research indicated possible problems with collecting all the data by a physician during examinations [15, 10]. It is common that some examinations might be omitted by a gynaecologist, either due to the their unavailability or because of medical reasons. The possible lack of data can be due to e.g. the technical limitations of the health care unit, high costs of medical examination and high risk of patient's health deterioration after potential examination. Obviously, lack of data hinders making a final decision.

The main issue we investigate in this paper is how to overcome the problem of low-quality diagnosis in the presence of missing data. The approach presented in the following sections focuses on aggregating knowledge that comes from many diagnostic scales, in order to minimise the impact of incomplete data. In Section 2 we introduce notions of an interval-based model of a patient, a diagnostic scale and aggregation operator. In Section 3 we describe an aggregation strategy that allows to improve a diagnosis and we give a methodology of calculating, analysing and compering different aggregation methods. Section 4 concludes our results.

2 DESCRIPTION OF THE PROBLEM

2.1 INTERVAL-VALUED PATIENT MODEL

In a classical approach, a patient is modelled as a vector $\mathbf{p} = (p_1, p_2, ..., p_n)$ in a space $P = D_1 \times D_2 \times ... \times D_n$, where $D_1, D_2, ..., D_n$ are real closed intervals denoting domains of attributes that describe patients.

We extend this representation by introducing a possibility to model incomplete data. Each attribute D_i is substituted by its interval version $\hat{D}_i = \mathcal{I}_{D_i}$. Analo-

ISBN: 978–83–8012–519–3

Table 1: Attributes used by most common preoperative diagnostic models.

group	attribute	SM [11] m_1	Alc. [1] m_2	LR1 [13] m_3	LR2 [13] m_4	Timm. [12] m_5	RMI1 [7] m_6
objective medical history	age	-	-	✓	✓	-	✓
	menopausal status	✓	-	-	-	✓	✓
	ovarian cancer in family	-	-	✓	-	-	-
	hormonal therapy	-	-	✓	-	-	-
	hysterectomy	-	-	-	-	-	✓
other	pain during examination	-	-	✓	-	-	-
	lesion volume	✓	-	✓	-	-	-
	internal cyst walls	✓	-	✓	✓	-	-
	septum thickness	✓	-	-	-	-	-
	echogenicity	✓	✓	-	-	-	-
	localisation	✓	-	-	-	-	✓
	ascites	✓	-	✓	✓	-	✓
	papillary projections	-	✓	-	-	✓	-
	solid element size	-	✓	✓	✓	-	✓
	blood flow location	-	✓	✓	✓	-	-
	resistance index	-	✓	-	-	-	-
	acoustic shadow	-	-	✓	✓	-	-
	amount of blood flow	-	-	✓	-	✓	-
	CA-125 blood marker	-	-	-	-	✓	✓
	lesion quality class	-	-	-	-	-	✓

gously as before we define $\hat{P} = \hat{D}_1 \times \hat{D}_2 \times ... \times \hat{D}_n$. Consequently, for each vector $\mathbf{p}^* \in P^*$ we can define its interval equivalent $\hat{\mathbf{p}} \in \hat{P}$ that has a form $\hat{\mathbf{p}} = \left([\underline{p}_1, \overline{p}_1], ..., [\underline{p}_n, \overline{p}_n]\right)$, where

$$\underline{p}_i = \begin{cases} p_i & \text{if } p_i \neq \emptyset \\ \min\limits_{d \in D_i} d & \text{if } p_i = \emptyset \end{cases}, \overline{p}_i = \begin{cases} p_i & \text{if } p_i \neq \emptyset \\ \max\limits_{d \in D_i} d & \text{if } p_i = \emptyset \end{cases}. \quad (1)$$

2.2 INTERVAL-VALUED DIAGNOSTIC SCALES

Diagnostic scale can be formalised as a function $m : P \to [0, 1]$. Values returned by a function indicate malignancy of a tumor and are interpreted in the following way:

- $m(\mathbf{p}) > 0.5$ – diagnosis towards malignant;
- $m(\mathbf{p}) < 0.5$ – diagnosis towards benign;
- $m(\mathbf{p}) = 0.5$ – indicates the impossibility of determining the nature of malignancy.

We construct an extended diagnostic scale $\hat{m} : \hat{P} \to \mathcal{I}_{[0,1]}$ defined as:

$$\hat{m}(\hat{\mathbf{p}}) = \left\{ m(\mathbf{p}) : \forall_{1 \leq i \leq n} \quad \underline{p}_i \leq p_i \leq \overline{p}_i \right\} = \left[\min_{\mathbf{p} \in \hat{\mathbf{p}}} m(\mathbf{p}), \max_{\mathbf{p} \in \hat{\mathbf{p}}} m(\mathbf{p}) \right] \quad (2)$$

where by $\mathbf{p} \in \hat{\mathbf{p}}$ we denote that $\mathbf{p}$ is an embedded vector of $\hat{\mathbf{p}}$.

Such extended diagnostic scale is able to operate on interval-valued representation of a patient. The resultant interval represents all the possible diagnoses that can be made basing on a patient description, in which every missing value was substituted with all possible values for that attribute. The more incomplete description, the more uncertain the diagnosis. However, worth noticing is that in many cases it is still possible to make a proper diagnosis, since some amount of missing values is acceptable and would not affect the final result significantly.

2.3 INTERVAL-VALUED AGGREGATION

A diagnosis in a form of an interval (2) has its advantages and drawbacks. An advantage is that such model gives a diagnosis even in the presence of missing data. A drawback is that the diagnosis is often uncertain and not so easy to apply by a physician. A main problem is thus how to efficiently support a physician in making a final diagnosis under incomplete information.

In order to solve this problem we make a following observation. As presented in Table 1 different scales denoted by $m_1, \ldots m_n$ use different attributes describing the patient, and therefore are subjected to different levels of uncertainty. The main idea is thus to improve the final diagnosis by taking advantage of the diversity of diagnostic scales. Given n scales $\hat{m}_1, \ldots \hat{m}_n$ we construct a function $\mathrm{Agg} : {\mathcal{I}_{[0,1]}}^n \to \mathcal{I}_{[0,1]}$. Its result $\mathrm{Agg}(\hat{m}_1, \ldots \hat{m}_n)$ is an interval that gathers and integrates information from the input sets. Thanks to this interpretation we immediately see the relationship with the issue of group decision making and information aggregation [3].

An n-argument interval–valued aggregation operator is a mapping $\mathrm{Agg} : {\mathcal{I}_{[0,1]}}^n \to \mathcal{I}_{[0,1]}$ with the following properties [6]:

1. if $\hat{x}_i \preceq \hat{y}_i$ for all $i \in 1, ..., n$, then $\mathrm{Agg}(\hat{x}_1, ..., \hat{x}_n) \preceq \mathrm{Agg}(\hat{y}_1, ..., \hat{y}_n)$,
2. $\mathrm{Agg}([1, 1], ..., [1, 1]) = [1, 1]$,
3. $\mathrm{Agg}([0, 0], ..., [0, 0]) = [0, 0]$,

where relation $\preceq$ is defined as follows:

$$[x_1, x_2] \preceq [y_1, y_2] \iff x_1 \leq y_1 \text{ and } x_2 \geq y_2 .$$

Recent research has led to the construction of many interval–valued aggregation methods [16, 5, 6, 2]. The most commonly used aggregation methods in group decision making are based on the weighted arithmetic average [3]. We propose to use various aggregation methods to improve the quality of diagnosis as well as minimise the impact of the lack of data and uncertainty in decision making.

In medical decision making problem, final diagnosis obtained from aggregation must indicate whether tumor is malignant or not. However, supporting decision in the case when there is not enough information may led to wrong diagnosis. Thus we accept situation when no diagnosis recommendation is made. The conversion of interval diagnosis into final diagnosis (binarization) is very important and may influence overall efficacy.

3 AGGREGATION PROCESS

3.1 METHODOLOGY DESCRIPTION

The proposed methodology is aimed at evaluating aggregation operators for coping with the lack of data. For this purpose, an essential element of the methodology is to simulate different levels of missing data. In order to better reflect the reality we have divided the attributes that describe the patient into two separate groups: those that were subjected to obscuration and those that were not. This separation naturally exists in many problems, including the problem of medical diagnosis because some data about the patient, such as age and other objective data from the medical history, are always available to the physician.

The results of all diagnostic scales are represented as intervals. The list of those intervals forms an input to the aggregation operators. Each operator synthesise input diagnoses in accordance with its principle of operation. The result of the aggregation is an interval representing the synthesised diagnosis. In order to make the final diagnosis it is required to perform the binarization process. Resulting diagnoses are compared to reference values in order to calculate the necessary statistics. The final statistics for a given level of missing data are calculated by averaging the results of all iterations.

Our study group was 268 women diagnosed and treated due to ovarian tumor in the Division of Gynaecological Surgery, Poznań University of Medical Sciences between 2005 and 2012. Among them, 62% was diagnosed with a benign tumor and 38% with a malignant one. In each iteration we chose 50 patients for positive and negative groups. All patients had no missing values in attributes required by diagnostic scales. Whole dataset is described in details in [8].

In evaluation we used six different diagnostic scales $\hat{m}_1, ..., \hat{m}_6$ obtained from basic scales listed in Table 1.

3.2 AGGREGATION STRATEGY

For the experiment we chose the simplest methods of aggregation, which base on weighted average, sum and intersection of sets, and majority vote. Such methods are most often used in the problem of group decision making [3]. However, the authors are aware that these methods do not cover recent research in that field (e.g. [4]).

A construction of a certain method of aggregation consists in choices of aggregation strategy and binarization strategy. Moreover, methods which base on weighted average require definition of weight of intervals. We chose 10 methods of aggregation presented in Table 2.

First group of aggregation operators (A–C) is based on arithmetic mean with use of interval arithmetic:

$$\mathrm{Agg}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_n) = \frac{\sum_{i=1}^{n} \omega(\hat{x}_i) \times \hat{x}_i}{\sum_{i=1}^{n} \omega(\hat{x}_i)} .$$

Second group of the operators (D–G) is based on weighted mean which is calculated with reference to a representative (rep) of the interval. Selected strategies in choosing representatives are minimum, maximum and centre of a interval:

$$\mathrm{Agg}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_n) = \frac{\sum_{i=1}^{n} \omega(\hat{x}_i) \cdot \mathrm{rep}(\hat{x}_i)}{\sum_{i=1}^{n} \omega(\hat{x}_i)} .$$

Next two operators (H–I) are based on sum and in-

Table 2: Selected aggregation methods.

ID	aggregation	strategy of weight calc.	binarization
A	Interval avg.	width	margin
B	Interval avg.	entropy	no margin
C	Interval avg.	constant	no margin
D	Lower bound avg.	width	margin
E	Upper bound avg.	width	margin
F	Center avg.	width	margin
G	Center avg.	entropy	margin
H	Intersection	-	no margin
I	Sum	-	no margin
J	Majority vote	-	margin

tersection from the set theory:

$$\mathrm{Agg}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_n) = \bigcup_{i=1}^{n} \hat{x}_i$$

and

$$\mathrm{Agg}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_n) = \bigcap_{i=1}^{n} \hat{x}_i \,.$$

Last method (J) differs from former in a way that it firstly binarizes input intervals, and after that it chooses a diagnose which appeared more frequently. In case of draw, the diagnosis is not taken.

3.3 WEIGHT CALCULATION STRATEGY

In our evaluation we selected three strategies for choosing weights:

- constant value: $\omega([a,b]) = 1$,
- interval length: $\omega([a,b]) = b - a$,
- normalised interval distance from 0.5 (entropy):

$$\omega([a,b]) = \begin{cases} 0 & \text{if } a \leq 0.5 \leq b \\ 2(a-0.5) & \text{if } a \geq 0.5 \\ 2(0.5-b) & \text{otherwise}\,. \end{cases}$$

3.4 BINARIZATION STRATEGY

In our research we chose the simplest variant of interval binarization:

$$\tau_\epsilon([a,b]) = \begin{cases} 0 & \text{if } b < 0.5 + \epsilon \\ 1 & \text{if } a \geq 0.5 - \epsilon \\ \mathrm{NA} & \text{otherwise}\,. \end{cases} \quad (3)$$

In this approach, an instance is classified as a positive when whole interval is greater than 0.5 with respect to ϵ margin. The negative case is defined similarly. In case when first or second conditions are not met, it is not possible to make a decision. For example, diagnosis $[0.1, 0.3]$ will be classified as benign, but for interval $[0.1, 0.6]$ it is not possible to make a decision, when margin is set to $\epsilon = 0.025$.

In our evaluation we arbitrarily chose two values for ϵ: 0 (no margin) and 0.025.

3.5 EVALUATION

Statistical evaluation as well as implementation of proposed methodology were performed using R software, version 3.1.1 [9]. We set levels of missing data to vary from 0% to 50% with 5% step. For each level we made 1000 repeats of random data obscuration with other calculations. With such number of repeats, the averaged results are stable, so that it is possible to reliably analyse them. We set a baseline to 69% accuracy that is achieved by a classifier based only on menopausal status – all methods of aggregation should be better than a baseline classifier and single diagnostic scales.

The most significant results are presented on Fig. 1. The figure presents how the aggregators and single diagnostic scales perform with increasing level of missing data. Sub-figure (a) presents diagnostic accuracy (ACC) and sub-figure (b) presents percentage of patients for whom the decision could be made. Upper and lower bounds of the shaded regions correspond to the biggest and the smallest values achieved among diagnostic scales.

The diagrams show that preserving high diagnosability frequently prevent models from achieving high accuracy – and *vice versa*.

4 RESULTS AND CONCLUSIONS

The developed methodology led us to conclusion that in the case of our medical diagnosis problem, aggregation is useful as a tool to solve the problem of missing data. Even the simplest methods presented in this paper received efficacy which exceed individual diagnostic scales, both in terms of accuracy and the number of diagnosed patients, despite missing data. There are three interesting cases:

1. result of an aggregation is an achievement of very high and stable accuracy (over 95%, regardless to level of missing data) at the expense of small number of patients, in which it was possible to make a diagnose (below 50%, less than individual diagnostic scales) – see e.g. I aggregation operator,
2. result of an aggregation is an achievement of very high diagnosability (over 90%) regardless of the

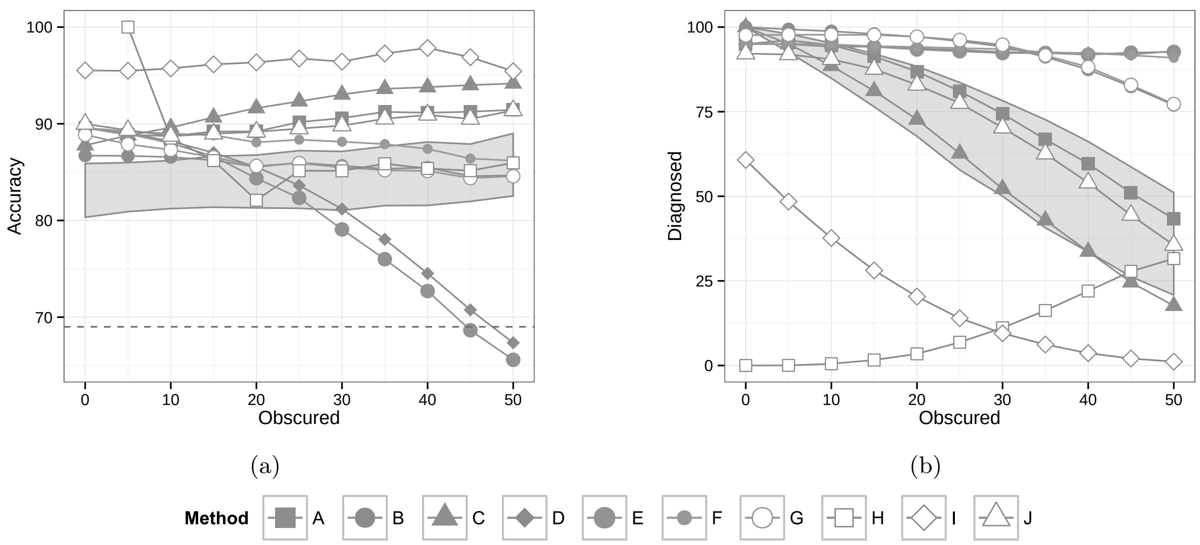


Figure 1: Results of evaluation of selected aggregation methods. Shaded region indicates bounds of single diagnostic methods. A dashed horizontal line in sub-figure (a) indicates accuracy of the baseline classifier.

level of missing data, at the expense of decreasing accuracy with respect to increasing level of missing data (accuracy might be even lower than these achieved by individual diagnostic scales) – see e.g. *E* and *F* aggregation operators,

3. result of an aggregation is an achievement of persistent high accuracy which is comparable to those achieved by individual diagnostic scales, with simultaneous high level of diagnosability (significantly higher than those achieved by individual diagnostic scales) – see e.g. *F* i *G*.

In the problem of ovarian tumor diagnosis, it seems that the most promising results were obtained by aggregation operator *F*. It is capable of maintaining high accuracy and diagnosability. Its little sensitivity to the lack of data makes it a promising candidate in the search for robust aggregation operators.

An interesting result is obtained in the case of the aggregation operator *H*, which is based on a set intersection. In the case of complete data it is not able to make any decision. This is due to the fact that in such situation the intervals are degenerated to a single points, and the intersection of such intervals is mostly an empty set.

The authors are aware that since the evaluation was performed on the whole dataset with arbitrarily chosen binarization margins, general conclusions on the performance of the presented aggregation operators should not be drawn. To make the results more reliable, the performance should be validated on a separate dataset with optimised aggregation parameters.

In the future work, we are going to study broader range of aggregation methods and determine guidelines on their applicability to various problems.

The presented results are promising and show that the competent selection and use of aggregation methods can significantly improve the quality of decisions taken by a diagnostic system. The problem is particularly significant when the knowledge is based on incomplete information. Proper selection of the method of aggregation is essential for weakening the negative impact of the incomplete data on the quality of decisions. Because the design of aggregation method depends on the particular problem, each time its extensive evaluation is needed. It can be done using our proposed method.

References

[1] J. L. Alcázar, L. T. Mercé, et al. A new scoring system to differentiate benign from malignant adnexal masses. *Obstetrical & Gynecological Survey*, 58(7):462–463, 2003.

[2] G. Beliakov, H. Bustince, et al. Aggregation for atanassov's intuitionistic and interval valued fuzzy sets: The median operator. *Fuzzy Systems, IEEE Transactions on*, 20(3):487–498, 2012.

[3] G. Beliakov, A. Pradera, et al. *Aggregation functions: A guide for practitioners.* Springer, 2008.

[4] L. De Miguel, H. Bustince, J. Fernandez, E. Induráin, A. Kolesárová, and R. Mesiar. Construction of admissible linear orders for interval-valued atanassov intuitionistic fuzzy sets with an application to decision making. *Information Fusion*, 2015, doi:10.1016/j.inffus.2015.03.004.

[5] G. Deschrijver and E. Kerre. Implicators based on binary aggregation operators in interval-valued fuzzy set theory. *Fuzzy Sets and Systems*, 153(2):229–248, 2005.

[6] G. Deschrijver and E. Kerre. Aggregation operators in interval-valued fuzzy and Atanassov's intuitionistic fuzzy set theory. In *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, pages 183–203. Springer, 2008.

[7] I. Jacobs, D. Oram, et al. A risk of malignancy index incorporating ca 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *BJOG: An International Journal of Obstetrics & Gynaecology*, 97(10):922–929, 1990.

[8] R. Moszyński, P. Żywica, et al. Menopausal status strongly influences the utility of predictive models in differential diagnosis of ovarian tumors: An external validation of selected diagnostic tools. *Ginekologia Polska*, 85(12):892–899, 2014.

[9] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2014.

[10] A. Stachowiak, P. Żywica, et al. An interval-valued fuzzy classifier based on an uncertainty-aware similarity measure. In P. Angelov et al., editors, *Intelligent Systems'2014*, volume 322 of *Advances in Intelligent Systems and Computing*, pages 741–751. Springer International Publishing, 2015.

[11] D. Szpurek, R. Moszyński, et al. An ultrasonographic morphological index for prediction of ovarian tumor malignancy. *European Journal of Gynaecological Oncology*, 26(1):51–54, 2005.

[12] D. Timmerman, T. H. Bourne, et al. A comparison of methods for preoperative discrimination between malignant and benign adnexal masses: the development of a new logistic regression model. *American Journal of Obstetrics and Gynecology*, 181(1):57–65, 1999.

[13] D. Timmerman, A. C. Testa, et al. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. *Journal of Clinical Oncology*, 23(34):8794–8801, 2005.

[14] C. Van Holsbeke, B. Van Calster, et al. External validation of mathematical models to distinguish between benign and malignant adnexal tumors: a multicenter study by the International Ovarian Tumor Analysis Group. *Clinical Cancer Research*, 13(15):4440–4447, 2007.

[15] A. Wójtowicz, P. Żywica, et al. Dealing with Uncertainty in Ovarian Tumor Diagnosis. In K. Atanassov et al., editors, *Modern Approaches in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics*, pages 151–158. IBS PAN – SRI PAS, Warsaw, 2014.

[16] R. Yager. Owa aggregation of intuitionistic fuzzy sets. *International Journal of General Systems*, 38(6):617–641, 2009.

Michał Baczyński, Bernard De Baets, Radko Mesiar (Eds.)
Proceedings of 8th International Summer School on Aggregation Operators (AGOP 2015)

ISBN 978 - 83 - 8012 - 519 - 3

9 788380 125193 >