35 Years of Fuzzy Set Theory
35 Years of Fuzzy Set Theory

Celebratory Volume Dedicated to the Retirement of Etienne E. Kerre
Foreword

The volume “35 Years of Fuzzy Set Theory: Celebratory Volume Dedicated to the Retirement of Etienne E. Kerre”, edited by Chris Cornelis, Glad Deschrijver, Mike Nachtegaele, Steven Schockaert and Yun Shi, is a fitting tribute to Professor Etienne Kerre as a scientist, educator and a member of the profession. For me, to write a preface to this volume is an opportunity to offer a personal tribute to Professor Kerre – my long time friend. I should like to precede my tribute with an expression of regret. First, because Etienne is retiring from his full professorship at Ghent University and his position as Director of the Fuzziness and Uncertainty Modelling Research Unit of Ghent University; and second, because in the world of science and academia the likes of Etienne are so few and far between.

My acquaintance with Etienne goes back to the seventies when Etienne entered the academic world and shortly after he became interested in fuzzy set theory and its applications. Since those days, Etienne has become a very prominent figure within the fuzzy logic community. His many seminal contributions have brought him both national and international recognition. Particularly worthy of note are his contributions to fuzzy topology, fuzzy algebraic structures, fuzzy relational calculus, fuzzy reasoning and if-then rules, possibility theory, reliability theory and fuzzy image processing. He has authored or co-authored close to ten books and over 390 papers on fuzzy set theory and its applications. It is a remarkable record. Equally impressive are Etienne’s contributions as an educator. He has guided the research of over twenty Ph.D.s and many graduate students. In addition, Etienne has played an active role in organizing and participating in international conferences. Etienne has a very warm personality and works very closely with his students. He is a role model in all respects.

I should like to take this opportunity to comment on the profound changes in the world of fuzzy logic since Etienne began his career in the academic world. By 1980 there were 521 papers in the INSPEC database which contained “fuzzy” in the title. Today, there are 68073. In addition, there are 18302 papers with “fuzzy” in the title in the MathSciNet database. Fuzzy logic is
employed in a wide variety of products, among them cameras, appliances, hybrid engines and transmissions, TV sets, copiers, elevators, blood pressure meters, etc. In recent years, applications of fuzzy logic within human sciences, notably sociology and political science, began to grow in visibility and importance.

These numbers paint a rosy picture. But beyond the rosy picture one can see some problems which call for attention. A problem which has existed since the birth of fuzzy set theory and fuzzy logic is related to the meaning of “fuzzy”. In English, the word “fuzzy” has a pejorative connotation. In the United States, in particular, pejorative connotation of “fuzzy” has seriously impeded acceptance of fuzzy logic. Despite the profusion of products which employ fuzzy logic, what is still visible is a residue of skepticism, particularly in the academic world.

Pejorative connotation of “fuzzy” is rooted in a deep-seated attitude in science – an attitude of respect for precision and lack of respect for imprecision. What this means is that “fuzzy” is pejorative because it is related to imprecision. This is a problem that cuts across most languages and is hard to resolve.

A problem which is more visible in the United States than in other countries is that at many universities courses on fuzzy logic are not taught. As a consequence, most engineering graduates are unfamiliar with fuzzy logic and hence are not in a position to employ it even when fuzzy logic can result in a superior solution to a problem. However, in coming years the problem may be solved, at least in part. More specifically, what is happening is that computational intelligence, which is just another name for soft computing, is rapidly growing in popularity. Computational intelligence (soft computing) is an amalgam of fuzzy logic, neurocomputing and evolutionary computing. Thus, in coming years fuzzy logic will be taught in many universities in courses on computational intelligence.

What we observe today is that fuzzy logic, along with many other fields of science and technology, is developed and applied more actively in Asian countries – China, Japan, South Korea, Taiwan and Singapore – than in Western countries and the United States. This is a manifestation of a global shift in power and influence from the West to the East. That is why Western science needs so badly the likes of Etienne Kerre.

In my view, in years to come fuzzy logic – its name notwithstanding – is certain to grow in visibility, importance and acceptance. There is a basic reason. In essence, fuzzy logic is the logic of classes with unsharp boundaries. In the world of human cognition, classes with unsharp boundaries are the rule rather than exception. What this implies is that it is only a matter of time before the essentiality of fuzzy logic becomes apparent not only in science but in most domains of human thought and culture. In the history of fuzzy logic and its applications, Etienne’s contributions will occupy a prominent place.

Berkeley (California), June 1, 2010
Lotfi A. Zadeh
Preface

About Etienne E. Kerre

Etienne E. Kerre: husband, father, grandfather, and ... pioneering researcher in fuzzy set theory and fuzzy logic. This book is dedicated to the long and exceptionally fruitful career of Etienne, and is published on the occasion of his retirement. Let us use this preface to provide a bird’s eye view on the life of Etienne, both as a person and as a researcher, as a way to say “thank you” for the wonderful years that we worked together and enjoyed his guidance.

Etienne E. Kerre – ‘E’ stands for Emiel – was born in Zele, Belgium, on May 8, 1945, the very day World War II ended in Europe. He is the oldest son in a large family of 8 children. His father, who passed away earlier this year, worked for the Belgian railway company. His mother, who is now 90 years old, stayed at home to take care of the children. In the social context of the decades right after World War II, Etienne should be proud of having succeeded in starting what would prove to be a successful academic career. At that time, it was not obvious to complete a full cycle of secondary school. It were Etienne’s primary school teachers who could convince his parents to allow him to go to secondary school, given the fact that Etienne was an excellent student. We are all very lucky that Etienne got and took this opportunity!

Etienne is married to Andrea De Kegel and they have one daughter Tessa. Andrea is director of the cultural department of the province Oost-Vlaanderen (together they share a passion for culture), Tessa is a hematologist and has a Ph.D. in medical sciences, and his son-in-law Bruno is a jazz photographer and reviewer. Whenever Etienne speaks about his family, it is always with love. We particularly remember his happiness when his grandson Henri was born on October 21, 1999. On many occasions Etienne told us entertaining anecdotes about this boy of whom he is so proud. Without any doubt, his retirement will allow him to spend even more time with his grandson.
After graduating from Sint-Lodewijkscollege Secondary School (Lokeren, Belgium) in 1963, Etienne studied mathematics at Ghent University where he obtained a Master’s degree in 1967, and the certificate of teacher education one year later. After an unusually short period of less than three years, in 1970 he was able to complete his Ph.D. in Mathematics concerning “Low Energy Electron Diffraction”, a topic on which he has published five papers in international scientific journals. It was the first year of what Etienne calls his “3 Fat Years”: in 1970 he got his Ph.D., in 1971 he got married, and in 1972 he became a father. Etienne worked as an assistant of Ghent University from 1969 to 1978. In 1978 he became a work leader at Ghent University, and a lecturer in 1984. In 1991 he obtained a full professor position at Ghent University. In 2004 he furthermore obtained honorary professorships from Southwest Jiaotong University and Xihua University in Chengdu, China.

Etienne started his research on fuzzy set theory in 1976. In this year, he also founded the Fuzziness and Uncertainty Modelling Research Unit, which he guided until his retirement. Since then he has published, together with many colleagues from around the world (DBLP lists 91 different co-authors, ISI Web of Science more than 130), more than 390 research papers on fuzzy set theory (both theoretical and applied) in international scientific journals and international conference proceedings. His research interests cover Fuzzy Topology, Fuzzy Algebraic Structures, Fuzzy Relational Calculus, Fuzzy Reasoning and If-Then Rules, Possibility Theory and Reliability Theory, Fuzzy Image Processing and so on. He has also written 8 books on basic principles of Fuzzy Set Theory and its applications, and is editor of 17 other books. Besides these publications, Etienne has given 155 lectures at international conferences and institutes, and acted as organizer or chairman of sessions in 82 international conferences and workshops. He has been a member of the editorial boards of 24 international journals and has been a referee for more than 65 international journals. He has also been a member of the program committee of 127 international conferences and has been a referee for more than 83 international conferences. Already 29 students from Belgium, Bulgaria, China and Egypt obtained a Ph.D. degree under his guidance, and Etienne has been a member of the jury for another 15 Ph.D. theses.

These are all impressive numbers, and if in any way possible they quantify an impressive academic career. All editors of this book belong to the group of Etienne’s 29 former Ph.D. students, and all enjoyed his guidance. We can rightly say that Etienne has always made great efforts to attract young people to academic research in fuzzy set theory. This seduction already started during his classes on fuzzy set theory, which he taught with great enthusiasm and dynamism. Quite logically, a lot of students made their master thesis in the field of fuzzy set theory, and this was a good recruiting area for young researchers. Etienne has obtained several large research projects that enabled him to employ these young people, and to train them into successful academics.
Etienne was also very active in teaching. He has given 7 courses about Mathematical Analysis and 11 courses about Fuzzy Set Theory and Fuzzy Logic to both undergraduates and postgraduate levels in Computer Science, Mathematics and Physics from 1969 until now. In fact, thanks to Etienne, Ghent University was one of the first universities worldwide where fuzzy set theory was taught in the basic curriculum. Students from Ghent University will certainly remember his white duster, which he wore whenever he was teaching. Etienne is also known as a firm but fair professor to his students (including his Ph.D. students): he demanded a lot of effort from his students and insisted that they formulated everything as precisely as possible. This was of course justified, from an academic point of view in general and from a mathematical point of view in particular. However, especially in the first year of an academic education many students lack the required level of formal rigour, and on occasion Etienne would come in our offices to complain about it. Teaching students this skill has always been one of his main missions. It is also interesting to discover how well Etienne is known by many current secondary school teachers in mathematics. If they took their university study in Ghent, chances are high that they had at least one course of Etienne, as an assistant in his beginning period and as a professor later on, and the teachers we encountered do remember him vividly. Teaching was also done outside of Ghent. Etienne has given 3 courses to undergraduates at the University Nebraska (Lincoln, U.S.A.) from 1987 to 1990, and 2 courses to postgraduates at Alcatel Bell Company (Antwerp, Belgium) in 1993 and 1994.

Beside his strong involvement in the academic world, Etienne always stayed a family man. He would not hesitate to offer his help if someone needed it. We often wondered how he found the time to do all these things. Etienne just has an extraordinary amount of energy, which is prefectly illustrated by the fact that he only needs a couple of hours of sleep every night. Chances are high that he is reading this book in the middle of the night!

When we are writing this preface, Etienne still has three months before his retirement. We know that he will retire with mixed feelings. On the one hand Etienne certainly deserves the rest of a retirement, which will allow him to spend more time with his family and to engage in other activities. On the other hand, Etienne is filled with passion for research and is not able to stop completely. Luckily, the university offers the possibility to stay active, of course at a slower pace, such that Etienne will be able to get a good mix between retirement and a continued involvement in the academic world. We wish him all the best with his retirement and his new or more intensified activities, and hope to keep meeting him, either in the context of research or just as friends to have a drink on a sunny day. Good luck Etienne, and thank you for being the wonderful person and researcher that you are!
About This Volume

This volume contains contributions from researchers that have been close to Etienne in one way or another: former Ph.D. students, co-authors, and friends in general. The idea of compiling a volume dedicated to the retirement of Etienne originated in January 2010, and was carried out by the postdoctoral researchers in his group at that time. Several people have kindly offered their help in assisting us with the reviewing process: Kim Bauters, Martine De Cock, Jeroen Janssen, Bart Van Gasse, and Patricia Victor. We also thank the authors for the time and efforts they spent in preparing their contribution, within a tight time frame. We are grateful to Janusz Kacprzyk and the people from Springer-Verlag for their help, and for giving us the opportunity to publish this book in the series of Studies in Fuzziness and Soft Computing.

The chapters of this volume are divided in three parts: i) logics and connectives, ii) data analysis, and iii) media applications. The first part gathers contributions that are primarily of a theoretical nature, covering fuzzy logic, (interval-valued and intuitionistic) fuzzy logic connectives, and aggregation operators. The chapter by R. Mesiar and M. Komorníková introduces and discusses a general concept of aggregation on bounded partially ordered sets and on bounded lattices. The weakest and strongest aggregation functions on a poset are identified, and several construction methods are discussed. The properties of t-norms and related operators on a poset are discussed and several characterizations are given. The chapter ends with a discussion of aggregation functions on the underlying lattice of Atanassov’s intuitionistic fuzzy set theory. Next, the chapter by K.T. Atanassov contains an overview of and new results on 23 of the 138 different intuitionistic fuzzy implications and 5 of the different intuitionistic fuzzy negations already introduced by the author. Some properties of these operators are discussed, and a partial solution to a problem by Baczynski and Jayaram relating the contrapositive property is given using intuitionistic fuzzy implications and negations. The axioms of intuitionistic logic are checked for the intuitionistic fuzzy implications and negations. The chapter by V. Novák and I.Perfilieva deals with the practicality of fuzzy logic, both in the narrow and in the broad sense. In the field of mathematical fuzzy logic in the narrow sense, traditional and evaluated syntax are reviewed, as well as fuzzy type theory. In the field of fuzzy logic in the broader sense, the formal theories of evaluative linguistic expressions, fuzzy if–then rules, perception based logical deduction and intermediate and generalized quantifiers are treated. Finally an example in commonsense human reasoning is given, and two theories related to fuzzy logic are discussed, namely fuzzy approximation and the fuzzy transform. The subsequent chapter by B. Van Gasse, C. Cornelis and G. Deschrijver studies propositional calculi for interval-valued fuzzy logics. Starting from the notion of the triangularization of a lattice, the authors focus on a specific case called interval-valued residuated lattices. Then the notion of a triangle algebra is introduced, and shown to be isomorphic to a particular class
of interval-valued residuated lattices. Finally, triangle algebras are used to define several interval-valued fuzzy logics, for which soundness and completeness is analyzed. The last chapter of the first part by Y. Shi, B. Van Gasse and D. Ruan deals with fuzzy implications. The complete dependencies and independencies between eight commonly considered potential properties of implications are studied. Each independency is shown by a counterexample. Two of these counterexamples lead to a new class of implications which are only determined by a negation, and this new class is extensively investigated.

The second part of the book pertains to applications of fuzzy set theory to data analysis, covering classification, decision making, rough sets and formal concept analysis. The chapter by A.M. Radzikowska deals with \( L \)-fuzzy rough sets as a further generalization of rough sets. Rough sets were originally proposed as a formal tool for analyzing and processing incomplete information represented in data tables. Later on, fuzzy generalizations of rough sets were introduced and investigated to be able to deal with imprecision. Anja and Etienne cooperated on this topic for more than a decade. The next chapter is authored by Y. Djouadi, D. Dubois and H. Prade. One of Etienne Kerre’s favorite topics is undoubtedly fuzzy relational calculus. Formal concept analysis (FCA) is known to be a successful and elegant application of relational calculus, revolving around a set of objects, a set of properties, and a binary relation between them. This chapter covers different graded extensions of formal concept analysis that allow to account for gradual properties, to handle uncertainty, and to acknowledge typicality of properties and importance of objects. The chapter offers insights in existing work on fuzzy FCA as well as interesting new results. The chapter by X. Wang, C. Wu and X. Wu gives a comprehensive overview of the research on fuzzy choice functions. First, it discusses Banerjee’s framework, and then moves on to study various preferences derived from choice functions, where the focus is on rationality conditions and their interrelationships. At the end, Georgescu’s framework, in which also the set of alternatives to choose from is fuzzified, is investigated. Finally, the chapter by G. Chen, Y. Xiong and Q. Wei discusses the use of fuzzy partitions to generate classification rules in domains with numerical attributes. In particular, they develop an extension of GARC, a classifier which is based on association rules. Experimental results show that the accuracy of GARC is maintained, while using a fewer number of rules.

The final part of the book comprises chapters in which fuzzy sets are used to represent or to manipulate different types of media objects, such as text documents and images. The first chapter is a contribution of the 5 former PhD students of Etienne that worked on fuzzy techniques in image processing (M. Nachtegael, T. Mélange, S. Schulte, V. De Witte and D. Van der Weken). In a prozaic style, the chapter gives an overview of the 12 years of research in this field, focussing on fuzzy mathematical morphology, similarity measures for images and noise reduction filters for both grayscale and colour images and video sequences. Next, the chapter by H. Bustince, M. Pagola, E. Barrenechea and J. Fernández deals with adapting a fuzzy thresholding
algorithm to some extensions of fuzzy set theory, namely interval-valued fuzzy sets, type-2 fuzzy sets, interval type-2 fuzzy sets and intuitionistic fuzzy sets. The resulting fuzzy thresholding algorithms are used for image segmentation to solve the problem of accurate elicitation of membership functions. Experimental results suggest that in some cases, the adapted thresholding algorithm indeed outperforms the existing algorithm. The chapter by S. Schockaert, N. Makarytska, and M. De Cock provides an overview of fuzzy approaches and related techniques in the web intelligence field. In particular, the authors explain the need for, and discuss the potential of fuzzy concepts and fuzzy set theory for three classes of applications that are becoming increasingly important on the web: information retrieval methods, recommender systems and the semantic web. Finally, the chapter by S. Zadrozny, J. Kacprzyk and K. Nowacka is centered around the problem of text categorization. The authors propose a weighting scheme to represent text documents in which the weight of each term is a fuzzy set. An existing classification algorithm for binary attributes is generalized to cope with fuzzy set valued attributes, after which some preliminary experimental results are provided.

Ghent, July 2010
Chris Cornelis
Glad Deschrijver
Mike Nachtegaele
Steven Schockaert
Yun Shi
Contents

Part I: Logics and Connectives

Aggregation Functions on Bounded Posets ....................... 3
Radko Mesiar, Magda Komorníková

On the Intuitionistic Fuzzy Implications and Negations ...... 19
Krassimir T. Atanassov

Mathematical Fuzzy Logic: A Good Theory for Practice ...... 39
Vilém Novák, Irina Perfilieva

Interval-Valued Algebras and Fuzzy Logics ...................... 57
Bart Van Gasse, Chris Cornelis, Glad Deschrijver

Implications in Fuzzy Logic: Properties and a New Class .... 83
Yun Shi, Bart Van Gasse, Da Ruan

Part II: Data Analysis

On Lattice–Based Fuzzy Rough Sets ............................... 107
Anna Maria Radzikowska

Graduality, Uncertainty and Typicality in Formal Concept
Analysis ............................................................... 127
Yassine Djouadi, Didier Dubois, Henri Prade

Choice Functions in Fuzzy Environment: An Overview ....... 149
Xuzhu Wang, Caiping Wu, Xia Wu
A Fuzzy Extension to Compact and Accurate Associative Classification ................................................... 171
Guoqing Chen, Yixi Xiong, Qiang Wei

Part III: Media Applications

Fuzzy Techniques in Image Processing at Ghent University: Summary of a 12-Year Journey ................................................ 197
Mike Nachtegael, Tom Mélange, Stefan Schulte, Valérie De Witte, Dietrich Van der Weken

The Need to Use Fuzzy Extensions in Fuzzy Thresholding Algorithms .................................................. 219
Humberto Bustince, Miguel Pagola, Edurne Barrenechea, Javier Fernández

Fuzzy Methods on the Web: A Critical Discussion .............. 237
Steven Schockaert, Nataliya Makarytska, Martine De Cock

Using Fuzzy and Interval-Valued Fuzzy Sets in Automatic Text Categorization Based on a Fuzzy Information Retrieval Model ............................................. 267
Sławomir Zadrozny, Janusz Kacprzyk, Katarzyna Nowacka

Author Index ................................................ 293
Aggregation Functions on Bounded Posets

Radko Mesiar and Magda Komorníková

Abstract. A general concept of aggregation on bounded posets is introduced and discussed. Up to general results, several particular results due to E.E. Kerre and his research group are recalled in the light of our approach. A special attention is paid to triangular norms on bounded posets.

1 Introduction

Two basic attributes of aggregation functions introduced and discussed in [6, 28, 33, 34] are the monotonicity and the boundary conditions. Recall that an aggregation function $A : [0, 1]^n \to [0, 1]$ (an extended aggregation function $A : \bigcup_{n \in \mathbb{N}} [0, 1]^n \to [0, 1]$) is supposed to be increasing, i.e., $A(x) \leq A(y)$ whenever $x \leq y$, and $A(0, \ldots, 0) = 0$, $A(1, \ldots, 1) = 1$. The framework of aggregation functions on $[0, 1]$ can be modified into any closed interval $[a, b] \subseteq [-\infty, \infty]$, or even to any interval $I \subseteq [-\infty, \infty]$ (with modified boundary conditions). Obviously, aggregation functions can be introduced to act on any (partially) ordered structure with bounds.

Definition 1. Let $(P, \leq, 0, 1)$ be a bounded poset (partially ordered set). Let $n \in \mathbb{N}$ be fixed. A mapping $A : P^n \to P$ is called an $(n$-ary) aggregation function on $P$ whenever it is increasing, i.e.,

$$A(x) \leq A(y) \text{ whenever } x \leq y, \text{ (i.e., } x_1 \leq y_1, \ldots, x_n \leq y_n)$$

and it satisfies boundary conditions

Radko Mesiar
Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68
Bratislava, Slovakia
e-mail: Radko.Mesiar@stuba.sk

Magda Komorníková
Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68
Bratislava, Slovakia
e-mail: Magdalena.Komornikova@stuba.sk

C. Cornelis et al. (Eds.): 35 Years of Fuzzy Set Theory, STUDFUZZ 261, pp. 3–17.
springerlink.com © Springer-Verlag Berlin Heidelberg 2010
A mapping $B : \bigcup_{n \in \mathbb{N}} P^n \to P$ is called an extended aggregation function on $P$ whenever $B / P^n$ is an $n$–ary aggregation function on $P$ for any $n \in \mathbb{N}$.

Note that in the case $n = 1$, often a convention $A(x) = x$ is used for unary aggregation function on $[0,1]$, see [6, 28, 34]. We will not insist on this convention in our approach.

Remark 1. i) Note that for any poset $(P, \leq, 0, 1)$ a dual poset $(P^*, \leq^*, 0^*, 1^*)$ can be introduced, where $P^* = P$, $x \leq^* y$ if and only if $y \leq x$, and $0^* = 1$, $1^* = 0$. Evidently, any aggregation function $A : P^n \to P$ on $P$ can be considered also as an aggregation function on $P^*$. Several properties of $A$ on $P$ are the same as those of $A$ on $P^*$ (namely, all algebraic properties not linked to the orderings $\leq$ and $\leq^*$). However, properties based on the ordering should be modified by the above $^*$–duality (for example, conjunctivity on $P$ is equivalent to the disjunctivity on $P^*$, see Definition 2).

ii) Extended aggregation functions acting on a poset $P$ with convention of identity for the unary aggregation were introduced in [10], where their categorical foundations were studied. Particularly, aggregation processes in probabilistic metric spaces were considered. As a typical example, convolution of distribution functions can be seen as a special aggregation function.

In this contribution, we discuss first aggregation functions which are defined on any poset, see Section 2. In Section 3, aggregation functions on general bounded lattices are studied. Special class of triangular norms and related aggregation functions are covered by Section 4. Section 5 brings an overview of some aggregation functions acting on special lattice $L^*$, with several results of E.E. Kerre’s group. Finally, some concluding remarks are added.

## 2 Aggregation Functions on General Bounded Posets

Algebraic properties introduced for the classical aggregation functions acting on $[0,1]$ can be straightforwardly introduced for aggregation functions on general bounded posets. We recall some of these properties (for more details and discussion in the case $P = [0,1]$ we recommend [28]):

- symmetry
- associativity
- neutral element
- annihilator
- strict monotonicity
- decomposability (for extended aggregation functions)
- idempotency
- internality (i.e., $A(x_1, \ldots, x_n) \in \{x_1, \ldots, x_n\}$)
- bisymmetry.
Several other properties, like additivity, continuity, etc., can be introduced only in the case of particular posets equipped with additional properties, and thus we will not consider them. Note also that we can introduce conjunctive and disjunctive aggregation functions (even the concept of $k$–tolerantness and $k$–intolerantness due to Marichal \[35\]), while the concept of averaging aggregation functions is restricted to the class of lattices.

**Definition 2.** Let $A : P^n \rightarrow P$ be an aggregation function on $P$. Then $A$ is called:

i) conjunctive whenever $A(x_1, \ldots, x_n) \leq x_i$ for all $i \in \{1, \ldots, n\}$,

ii) disjunctive whenever $A(x_1, \ldots, x_n) \geq x_i$ for all $i \in \{1, \ldots, n\}$,

iii) $k$–tolerant with $k \in \{1, \ldots, n\}$ whenever $\text{card}(\{i|A(x_1, \ldots, x_n) \geq x_i\}) \geq n - k + 1$,

iv) $k$–intolerant with $k \in \{1, \ldots, n\}$ whenever $\text{card}(\{i|A(x_1, \ldots, x_n) \leq x_i\}) \geq n - k + 1$.

Evidently, conjunctive aggregation functions coincide with $1$-intolerant, and disjunctive aggregation functions coincide with $1$-tolerant. The concepts of tolerant and intolerant aggregation functions allows to introduce a stronger form of averaging property.

**Definition 3.** Let $A : P^n \rightarrow P$ be an aggregation function on $P$. Then $A$ is called strongly averaging whenever $A$ is simultaneously $n$-tolerant and $n$-intolerant, i.e., for each $(x_1, \ldots, x_n) \in P^n$, there are $i, j \in \{1, \ldots, n\}$ so that

$$x_i \leq A(x_1, \ldots, x_n) \leq x_j.$$

Evidently, each internal aggregation function $A$ on $P$ is strongly averaging (but not vice-versa).

**Example 1.** Let $P = [0, 1]^2$ be equipped with the standard partial order $\leq$. Observe that then $(P, \leq)$ is a lattice. Then $\text{Min}, \text{Max} : P^n \rightarrow P$ defined by

$$\text{Min}((x_1, y_1), \ldots, (x_n, y_n)) = (\min(x_1, \ldots, x_n), \min(y_1, \ldots, y_n))$$

and

$$\text{Max}((x_1, y_1), \ldots, (x_n, y_n)) = (\max(x_1, \ldots, x_n), \max(y_1, \ldots, y_n))$$

are idempotent $n$-ary aggregation functions on $P$. Neither $\text{Min}$ nor $\text{Max}$ are strongly averaging. However, $\text{Min}$ is conjunctive and $\text{Max}$ is disjunctive.

On the other side, projection $P_{ri} : P^n \rightarrow P$ defined by

$$P_{ri} ((x_1, y_1), \ldots, (x_n, y_n)) = (x_i, y_i)$$

are simultaneously $n$–tolerant and $n$–intolerant and thus strongly averaging (indeed, projections are internal aggregation functions).
We introduce some distinguished aggregation functions on \( P \) (for any arity thus also as extended aggregation functions):

- the weakest aggregation function \( A_w : \bigcup_{n \in \mathbb{N}} P^n \to P \),

\[
A_w(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } x_1 = \ldots = x_n = 1, \\ 0 & \text{else}, \end{cases}
\]

- the strongest aggregation function \( A_s : \bigcup_{n \in \mathbb{N}} P^n \to P \),

\[
A_s(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } x_1 = \ldots = x_n = 0, \\ 1 & \text{else}, \end{cases}
\]

- the weakest t-norm (drastic product) \( T_D : \bigcup_{n \in \mathbb{N}} P^n \to P \),

\[
T_D(x_1, \ldots, x_n) = \begin{cases} x_i & \text{if } x_j = 1 \text{ for each } j \neq i, \\ 0 & \text{else}, \end{cases}
\]

- the strongest t-conorm (drastic sum) \( S_D : \bigcup_{n \in \mathbb{N}} P^n \to P \),

\[
S_D(x_1, \ldots, x_n) = \begin{cases} x_i & \text{if } x_j = 0 \text{ for each } j \neq i, \\ 1 & \text{else}. \end{cases}
\]

All these four aggregation functions are associative (i.e., their binary form is associative, and n-ary form is the genuine extension of that binary form).

Aggregation functions \( A_w, T_D, A_s, S_D \) are symmetric. For any fixed \( n \in \mathbb{N}, n \geq 2, \) projection functions \( Pr_i : P^n \to P, Pr_i(x_1, \ldots, x_n) = x_i \), are non–symmetric \( n \)-ary aggregation functions on \( P \).

Observe that on the Boolean chain \( P = \{0, 1\} \), there are exactly four binary aggregation functions, namely \( A_w = T_D, A_s = S_D, Pr_1 \) and \( Pr_2, Pr_1(x, y) = x \) and \( Pr_2(x, y) = y \). Each of them is associative. This fact excludes the possibility to define a non-associative aggregation function \( A : P^2 \to P \) by a unique formula independent of the underlying poset \( P \). Obviously, in particular cases of posets \( P \) with \( card P > 2 \), there are several non-associative aggregation functions. For example, the arithmetic mean \( M : [0, 1]^2 \to [0, 1] \) given by \( M(x, y) = \frac{x+y}{2} \) is not associative.

There are several construction methods appropriate for constructing aggregation functions on posets. First of all, for any fixed bounded poset \( (P, \leq, 0, 1) \), the composition of aggregation functions can be applied. It is not difficult to check that for any fixed \( m \in \mathbb{N}, n_1, \ldots, n_m \in \mathbb{N} \), and aggregation functions \( A : P^m \to P, B_i : P^{n_i} \to P, i = 1, \ldots, m \), the mapping \( D : P^n \to P, n = \sum_{i=1}^m n_i, D(x_1, \ldots, x_n) = A(B_1(x_1, \ldots, x_{n_1}), \ldots, B_m(x_{n_m+1}, \ldots, x_n)) \), is an \( n \)-ary aggregation function on \( P \).

As a special case we can consider an \( n \)-ary aggregation function \( A : P^n \to P \) and unary aggregation functions \( f, g_1, \ldots, g_n : P \to P \), yielding a transformed \( n \)-ary aggregation function \( D : P^n \to P \) given by

\[
D(x_1, \ldots, x_n) = f(A(g_1(x_1), \ldots, g_n(x_n))).
\]
Observe that [3] generalizes the standard isomorphic transformation of $A$, which corresponds to the case when $g_1 = \ldots = g_n$ and $f \circ g = id_P, f \left(g(x)\right) = x$.

Another construction method is linked to the ordinal sums of posets introduced by Birkhoff [3].

**Proposition 1.** Let $(K, \preceq, 0_K, 1_K)$ be a bounded chain, and let $\left((P_k, \preceq_k, 0_k, 1_k)\right)_{k \in K}$ be a family of posets such that for $k_1, k_2 \in K, k_1 \preceq k_2$, either $P_k \cap P_{k_2}$ is empty, or $1_{k_1} = 0_{k_2}$ and $P_k \cap P_{k_2} = \{1_{k_1}\}$. For each $k \in K$, let $A_k : P^n_k \rightarrow P_k$ be an aggregation function on $P_k$. Then:

(i) $\left(P, \preceq, 0, 1\right)$ is a poset, where $P = \bigcup_{k \in K} P_k, 0 = 0_{0_K}, 1 = 1_{1_K}$, and for any $x, y \in P$, $x \preceq y$ if and only if $x \in P_{k_1}, y \in P_{k_2}$ and $k_1 \preceq k_2$ or $k_1 = k_2 = k$ and $x \preceq_k y$.

(ii) A mapping $A : P^n \rightarrow P(A : \bigcup_{n \in N} P^n \rightarrow P)$ given by $A(x_1, \ldots, x_n) = A_k(x_1^k, \ldots, x_n^k)$, where $k \in K$ satisfies $x_i \in P_k$ for some $i \in \{1, \ldots, n\}$ and $\{x_1, \ldots, x_n\} \cap (\bigcup_{r \neq k} P_r) = \emptyset$, and

$$x_j^k = \begin{cases} x_j & \text{if } x_j \in P_k, \\ 1_k & \text{if } x_j > 1_k, \\ 0_k & \text{if } x_j < 0_k, \end{cases}$$

is an aggregation function on $P$. $A$ is called conjunctive ordinal sum, and $A_k, k \in K$, are called summands with the notation $A = C - (\langle P_k, A_k \rangle | k \in K)$.

(iii) A mapping $B : P^n \rightarrow P(B : \bigcup_{n \in N} P^n \rightarrow P)$ given by $B(x_1, \ldots, x_n) = A_k(x_1^k, \ldots, x_n^k)$, where $k \in K$ satisfies $x_i \in P_k$ for some $i \in \{1, \ldots, n\}$ and $\{x_1, \ldots, x_n\} \cap (\bigcup_{r \neq k} P_r) = \emptyset$, is an aggregation function on $P$. $B$ is called a disjunctive ordinal sum with summands $A_r$, with the notation $B = D - (\langle A_r, P_r \rangle | r \in K)$.

**Proof.** Item (i) is due to Birkhoff [3]. Items (ii) and (iii) can be checked directly. □

The next properties can be either checked directly, or they follow from Clifford’s ordinal sum of semigroups construction [7].

**Corollary 1.** Let $A = C - (\langle P_k, A_k \rangle | k \in K)$, $(B = D - (\langle A_k, P_k \rangle | k \in K))$ be a conjunctive (disjunctive) ordinal sum of aggregation functions. Then

(i) $A$ is a conjunctive aggregation function ($B$ is a disjunctive aggregation function) on $P$ if and only if all summands $A_k, k \in K$, are conjunctive (all summands $A_k, k \in K$, are disjunctive);

(ii) $A$ ($B$) is symmetric if and only if all $A_k, k \in K$, are symmetric;

(iii) $A$ is associative ($B$ is associative) if and only if all $A_k, k \in K$, are associative, and if for $k_1 < k_2$, $P_k \cap P_{k_2} = \emptyset$, then $1_{k_1}$ is a neutral element of $A_k$ and $0_{k_2} = 1_{k_1}$ is the annihilator of $A_{k_2}$ ($1_{k_1}$ is an annihilator of $A_k$ and $0_{k_2} = 1_{k_1}$ is a neutral element of $A_{k_2}$).
3 Aggregation Functions on Bounded Lattices

Two basic binary aggregation functions on a bounded lattice \((L, \leq, 0, 1)\) are the lattice operations, i.e., the join \(\text{Max}(x, y) = x \lor y\) and the meet \(\text{Min}(x, y) = x \land y\). Due to their associativity, both \(\text{Max}\) and \(\text{Min}\) can be extended to \(n\)-ary aggregation functions, as well as to extended aggregation functions (keeping the original notation). Observe that \(\text{Min}\) is the strongest conjunctive aggregation function on \(L\), while \(\text{Max}\) is the weakest disjunctive aggregation function on \(L\).

For aggregation on lattices, the concept of averaging can be introduced as follows.

**Definition 4.** Let \((L, \leq, 0, 1)\) be a lattice and let \(A : L^n \to L\) be an aggregation function on \(L\). \(A\) is called averaging whenever it satisfies \(\text{Min} \leq A \leq \text{Max}\).

In general, each strongly averaging aggregation function on a lattice \(L\) is averaging.

Note that if \(L\) is a chain then the strong averaging and averaging properties coincide. In any other case, \(\text{Min}\) and \(\text{Max}\) are averaging aggregation functions which are not strongly averaging, compare also Example 1. Moreover, an aggregation function \(A\) on a lattice \(L\) is averaging if and only if it is idempotent.

Lattice operations of the join \(\lor\) and of the meet \(\land\) allows to introduce lattice polynomials \(p : L^n \to L\) in the following way (see [27], section I.4):

(i) For any \(i \in \{1, \ldots, n\}\), \(P_r : L^n \to L\) is a lattice polynomial.
(ii) If \(A, B : L^n \to L\) are lattice polynomial functions, then also \(A \lor B\) and \(A \land B\) are lattice polynomial functions from \(L^n\) to \(L\).
(iii) Every lattice polynomial function \(A : L^n \to L\) is formed by finitely many applications of rules (i) and (ii).

**Example 2.** Define \(A, B : L^3 \to L\) by

\[
A(x, y, z) = (x \lor y) \land z
\]

and

\[
B(x, y, z) = (x \land z) \lor (y \land z).
\]

Then both \(A\) and \(B\) are lattice polynomials on \(L\), and in general \(A \neq B\). Obviously, if \(L\) is a distributive lattice, then \(A = B\).

Note that each lattice polynomial \(A\) is an averaging aggregation function, which is strongly averaging if and only if \(A\) is a projection function, \(A = Pr_i\) for some \(i\), or \(L\) is a chain.

An interesting method how to construct \(n\)-ary aggregation functions on a lattice \(L\) is linked to sections of aggregation functions on \(L\) with higher dimensions (this method can be applied on posets, too):

consider an aggregation function \(B : L^{k+n} \to L\), and choose elements \(a_1, \ldots, a_k \in L\) so that
Aggregation Functions on Bounded Posets

\[ B(a_1, \ldots, a_k, 0, \ldots, 0) = 0 \text{ and } B(a_1, \ldots, a_k, 1, \ldots, 1) = 1 \]

(observe that such elements need not exist, in general); then the mapping \( A : L^n \to L \) given by

\[ A(x_1, \ldots, x_n) = B(a_1, \ldots, a_k, x_1, \ldots, x_n) \]

is an \( n \)-ary aggregation function on \( L \).

**Example 3.**

(i) Let \( \text{Med} : L^3 \to L \) be given by \( \text{Med}(x, y, z) = (x \wedge y) \vee (x \wedge z) \vee (y \wedge z) \).

Then \( \text{Med} \) is a ternary lattice polynomial on \( L \) (\( \text{Med} \) is called median, following the standard terminology from the case when \( L = [\mathbb{R}] \) is the extended real line). Obviously, for any \( a \in L \), \( \text{Med}_a : L^2 \to L \) given by \( \text{Med}_a(x, y) = \text{Med}(a, x, y) \) is a binary aggregation function (so called \( a \)-median, compare [23, 26]). Note that \( \text{Med}_0 = \text{Min} \) and \( \text{Med}_1 = \text{Max} \).

(ii) Define \( B : L^4 \to L \) by \( B(x_1, x_2, x_3, x_4) = (x_1 \wedge x_3) \vee (x_2 \wedge x_4) \vee (x_3 \wedge x_4) \), i.e., \( B \) is a quaternary lattice polynomial on \( L \). For any fixed \( a, b \in L \), \( A : L^2 \to L \) given by \( A(x, y) = B(a, b, x, y) \) is a binary aggregation function on \( L \). Evidently, if \( a = b \) then \( A = \text{Med}_a \). Moreover, \( A \) is always associative, and if \( L = [0, 1] \) then it coincides with the binary Sugeno integral, see [40].

On any lattice \( L \), we can introduce also the following conjunctive aggregation function \( Z : L^2 \to L \) and its dual \( Z^* : L^2 \to L \) given by

\[ Z(x, y) = \begin{cases} x \wedge y & \text{if } x \vee y = 1, \\ 0 & \text{else}, \end{cases} \]

and

\[ Z^*(x, y) = \begin{cases} x \vee y & \text{if } x \wedge y = 0, \\ 1 & \text{else}. \end{cases} \]

Observe that the duality follows from Remark 1. Indeed, the formula for \( Z \) on \( (L, \leq, 0, 1) \) is exactly the same as the formula for \( Z^* \) on \( (L^*, \leq^*, 0^*, 1^*) \) taking into account that \( \wedge^* = \vee \) and \( \vee^* = \wedge \).

Note that \( Z = T_D \) \( (Z^* = S_D) \) if and only if \( L \) has no unit multipliers, i.e., \( x < 1 \) and \( y < 1 \) imply \( x \vee y < 1 \) \( (L \) has no zero divisors, i.e., \( x > 0 \) and \( y > 0 \) imply \( x \wedge y > 0 \)). Both these constraints are satisfied if \( L \) is a bounded chain. In remaining cases, \( Z \geq T_D \) \( (Z \neq T_D) \), and \( Z^* \leq S_D \) \( (Z^* \neq S_D) \).

The functions \( T_D, S_D, Z, Z^*, \text{Min} \) and \( \text{Max} \) are examples of triangular norms and related aggregation functions which we will discuss in the next section.

## 4 Triangular Norms and Related Aggregation Functions

E. E. Kerre with G. De Cooman [9] have introduced order norms on bounded posets, covering triangular norms, their weakening triangular seminorms, and related dual aggregation functions.

**Definition 5.** Let \( A : P^2 \to P \) be a binary aggregation function on a fixed bounded poset \( (P, \leq, 0, 1) \). Then \( A \) is called:
(i) a triangular seminorm (t–seminorm in short) whenever \(1\) is a neutral element of \(A\), i. e., \(A(x, 1) = A(1, x) = x\) for all \(x \in P\);

(ii) a triangular semiconorm (t–seminorm in short) whenever \(0\) is a neutral element of \(A\).

Moreover:

(iii) a symmetric associative \(t\)–seminorm \(A\) is called a triangular norm (t–norm) on \(P\);

(iv) a symmetric associative \(t\)–semiconorm \(A\) is called a triangular conorm (t–conorm) on \(P\).

Finally, if \(A\) is a \(t\)-norm, or a \(t\)-seminorm, or a \(t\)-conorm, or a \(t\)-semiconorm on \(P\), then it is called an order norm on \(P\).

Due to the associativity, triangular norms and triangular conorms can be extended to \(n\)–ary aggregation functions as well as to extended aggregation functions on \(P\) (for \(n = 1\), identity on \(P\) is considered in both cases). Note that triangular norms are conjunctive, while triangular conorms are disjunctive aggregation functions. Moreover, it is possible to introduce \(n\)–ary \(t\)–seminorms (t–seminorms) as aggregation functions \(A : P^n \rightarrow P\) with a neutral element \(1\) (0), i. e., satisfying \(A(x_1, \ldots, x_n) = x_i\) whenever \(x_j = 1\) for all \(j \neq i\) \((A(x_1, \ldots, x_n) = x_j\) whenever \(x_j = 0\) for all \(j \neq i\)).

As already mentioned in Section 2, \(T_0\) is the weakest \(t\)–norm on \(P\). Evidently, it is also the weakest \(t\)–seminorm on \(P\). Similarly, \(S_D\) is both the strongest \(t\)–conorm and the strongest \(t\)–semiconorm on \(P\). In general, the strongest \(t\)–norm either \(t\)–seminorm need not exist. By duality, the weakest \(t\)–conorm (t–semiconorm) need not exist.

**Example 4.** Define a bounded poset \(P\) by the next Hasse diagram:

![Hasse diagram of a poset P](image)

and two mappings \(A, B : P^2 \rightarrow P\) by the next tables 1 and 2.
Then both $A$ and $B$ are maximal $t$-seminorms on $P$. Moreover, both $A$ and $B$ are triangular norms on $P$.

On the other hand, if $P$ is a lattice, then $\text{Min}$ is the strongest $t$–norm as well as the strongest $t$–seminorm on $P$, and $\text{Max}$ is the weakest $t$–conorm as well as the weakest $t$–semiconorm on $P$.

Example 4 shows that there can be several idempotent $t$–norms ($t$–seminorms, $t$–conorms, $t$–semiconorms) on a general bounded poset $P$. However, if $P$ is a lattice, then the requirement of idempotency leads to the uniqueness of discussed order norms.

**Proposition 2.** Let $(L, \leq, 0, 1)$ be a bounded lattice and let $A : L^2 \to L$ be an idempotent (i. e., averaging) aggregation function on $L$. Then:

(i) $A$ is a $t$–norm if and only if $A$ is a $t$–seminorm if and only if $A = \text{Min}$;

(ii) $A$ is a $t$–conorm if and only if $A$ is a $t$–semiconorm if and only if $A = \text{Max}$.

Observe that due to the boundary conditions required for order norms, when considering a pair $A, B : P^2 \to P$ of a $t$–norm ($t$–seminorm) $A$ and a $t$–conorm ($t$–semiconorm) $B$ on $P$, the distributivity

$$A(x, B(y, z)) = B(A(x, y), A(x, z))$$

implies the absorption

$$B(A(x, y), x) = x$$
(see [9], Proposition 3.5), and this absorption implies the idempotency of $B$ (see [9], Proposition 3.6).

Similarly, the distributivity

$$A(B(x, y), z) = B(A(x, z), A(y, z))$$

implies the absorption

$$B(x, A(y, x)) = x$$

which implies the idempotency of $B$.

Note that the role of $A$ and $B$ can be reversed in the above claims.

Several other properties (and their relationships) of order norms are introduced and discussed in [9], having in mind in particular the role of $t$-norms and $t$-conorms as conjunctions and disjunctions in many valued logics with truth values domain $P$.

Order norms (in particular triangular norms) on special kinds of posets and/or possessing some special properties were discussed in several papers. For example, triangular norms on product lattices were studied, among others, in [8]. Evidently, for any product poset

$$(P, \leq, 0, 1) = \prod_{i \in I} (P_i, \leq_i, 0_i, 1_i),$$

and any system $(T_i)_{i \in I}$ of $t$–norms $T_i : P_i^2 \to P_i$, $i \in I$, the mapping $T : P^2 \to P$ given by

$$T \left( (x_i)_{i \in I}, (y_i)_{i \in I} \right) = (T_i(x_i, y_i))_{i \in I}$$

is a triangular norm on $P$. A similar result is valid also for the remaining types of order norms, i. e., for $t$–seminorms, for $t$–conorms, and for $t$–semiconorms. Interesting is the opposite claim, which is valid only under some special constraints.

Proposition 3. [8] For a fixed chain $(L, \leq, 0, 1)$ and $n > 1$, let $T : (L^n)^2 \to L^n$ be a $t$–norm on the product lattice $L^n$. Then $T$ is the product of $t$–norms $T_1, \ldots, T_n$ on $L$, i. e., $T \left( (x_1, \ldots, x_n), (y_1, \ldots, y_n) \right) = (T_1(x_1, y_1), \ldots, T_n(x_n, y_n))$ if and only if $T(x \land y, z) = T(x, z) \land T(y, z)$ for all $x, y, z \in L^n$ (or $T(x \lor y, z) = T(x, z) \lor T(y, z)$ for all $x, y, z \in L^n$).

Not all $t$–norms are direct products of $t$–norms on $L$. For example, in [30] a method for constructing $t$–norms on product lattices which are not direct products is presented.

Many particular results on $t$–norms on lattices (thus by duality also on $t$–conorms) can be found in the domain of integral monoids, see, for example, [29].

Problems concerning ordinal sums of $t$–norms on bounded lattices were investigated in [36], and [38]. We recall only the next result from [38].

Proposition 4. [38] Let $(L, \leq, 0, 1)$ be a bounded lattice, $(K, \leq, a, b)$ its complete bounded sublattice, and $H : K^2 \to K$ a $t$–norm on $K$. Then $T : L^2 \to L$ given by
Aggregation Functions on Bounded Posets

\[ T(x, y) = \begin{cases} 
  x \land y & \text{if } 1 \in \{x, y\}, \\
  0 & \text{if } 0 \in \{x, y\}, \\
  H(x^*, y^*) & \text{else,}
\end{cases} \]

where \( x^* = \sup \{z \in K \cup \{0, 1\} | z \leq x\} \), is a \( t \)-norm on \( L \) extending \( H \), i.e., \( T|K^2 = H \).

For interested readers we recall some works on order norms on posets (lattices) which may be an inspiration for a deeper study of this topic. Some particular methods to construct triangular norms on special types of posets can be found for example in [42]. \( \lor \)-distributivity and infinitely \( \lor \)-distributivity of triangular norms on complete bounded lattices is studied in [31]. An interesting relation between triangular norms and triangular seminorms on the standard lattice \( \mathcal{L} = [0, 1] \) (i.e., classical triangular norms are considered, as they were introduced in [39] and studied in [1, 32]) was shown in [24] (observe that triangular seminorms on \([0, 1]\) are often called semicopulas [25]).

**Proposition 5.** The class of all symmetric \( t \)-seminorms on \([0, 1]\) is the \( \lor \)-closure (\( \land \)-closure) of the class of all \( t \)-norms on \([0, 1]\).

**Open problem:** Let \((L, \leq, 0, 1)\) be a complete lattice. Is it possible to rewrite Proposition 5 for this case, i.e., is the class of all symmetric \( t \)-seminorms on \( L \) the \( \lor \)-closure (\( \land \)-closure) of the class of all \( t \)-norms on \( L \)?

### 5 Aggregation Functions on Lattice \( L^* \)

Consider the set \( L^* = \{ (x_1, x_2) | (x_1, x_2) \in [0, 1]^2, x_1 + x_2 \leq 1 \} \) and define a relation \( \leq \) on \( L^* \) by

\( (x_1, x_2) \leq (y_1, y_2) \) whenever \( x_1 \leq x_2 \) and \( x_2 \geq y_2 \).

Then \((L^*, \leq, 0, 1)\) is a bounded complete lattice, where \( 0 = (0, 1) \) and \( 1 = (1, 0) \).

Note that then the lattice operations on \( L^* \) are defined as follows:

\( (x_1, x_2) \land (y_1, y_2) = (x_1 \land y_1, x_2 \lor y_2) \),

\( (x_1, x_2) \lor (y_1, y_2) = (x_1 \lor y_1, x_2 \land y_2) \).

Recall that the lattice \( L^* \) is a background of intuitionistic fuzzy set theory introduced by [2] and it is isomorphic to the standard lattice of all closed subintervals of \([0, 1]\).

The major contribution to the theory of aggregation functions on \( L^* \) is due to the school of E. E. Kerre [13, 14, 15, 16, 17, 18, 19, 20, 21] and [4, 5, 22, 41]. In this section, we recall some of results presented in these manuscripts.

Note first of all that due to [12], a mapping \( N : L^* \rightarrow L^* \) is decreasing involution (i.e., a negator), if and only if there is decreasing involution \( n : [0, 1] \rightarrow [0, 1] \) such that \( N((x_1, x_2)) = (n(1 - x_2), 1 - n(x_1)) \). It is not difficult to see that
for any aggregation function \( A : (L^*)^n \to L^* \) (\( A : \bigcup_{n \in \mathbb{N}} (L^*)^n \to L^* \)), also the mapping \( A^N : (L^*)^n \to L^* \) \((A^N : \bigcup_{n \in \mathbb{N}} (L^*)^n \to L^*)\) given by \( A^N ((x_1,y_1), \ldots, (x_n,y_n)) = N (A (N(x_1,y_1), \ldots, N(x_n,y_n))) \) is an aggregation function on \( L^* \) \((A^N \) is then called \( N\)–dual of \( A \)). Recall that \( N\)–dual to a \( t\)–norm, \( t\)–seminorm, \( t\)–conorm, \( t\)–seminorm on \( L^* \) is a a \( t\)–conorm, \( t\)–seminorm, \( t\)–norm, \( t\)–seminorm on \( L^* \), respectively. A simple method how to construct aggregation functions on \( L^* \) is based on aggregation functions on \([0,1] \).

**Lemma 1.** Let \( A : [0,1]^n \to [0,1] \) be an aggregation function on \([0,1] \). Then \( A^*: (L^*)^n \to L^* \) given by

\[
A^* ((x_1,y_1), \ldots, (x_n,y_n)) = (A(x_1, \ldots, x_n), 1 - A(1 - y_1, \ldots, 1 - y_n))
\]

is an aggregation function on \( L^* \).

This method can be further generalized to so called representable aggregation functions on \( L^* \) \([11][12] \).

**Proposition 6.** Let \( A,B : [0,1]^n \to [0,1] \) \((A,B : \bigcup_{n \in \mathbb{N}} [0,1]^n \to [0,1] \)) be two aggregation functions of \([0,1] \) such that for all \( x_1, \ldots, x_n \in [0,1] \), the inequality

\[
A(x_1, \ldots, x_n) + B(1 - x_1, \ldots, 1 - x_n) \leq 1
\]

holds. Then the mapping \( C : (L^*)^n \to L^* \) \((C : \bigcup_{n \in \mathbb{N}} (L^*)^n \to L^*)\) given by \( C((x_1,y_1), \ldots, (x_n,y_n)) = (A(x_1, \ldots, x_n), B(y_1, \ldots, y_n)) \)

is an aggregation function on \( L^* \). \( C \) is called a representable aggregation function.

For arbitrary aggregation functions \( A,B \) on \([0,1] \) such that \( A \leq B^d \) \((B^d \) is the dual aggregation function to \( B \), based on duality w.r.t. \( n, n(x) = 1 - x \)) \( A \) and \( B \) satisfy requirements of Proposition \([6] \). Thus Lemma \([1] \) reflects the fact \( A \leq A \) (observe that \((A^d)^d = A \)). Note that not all aggregation functions on \( L^* \) are representable.

**Example 5.**

(i) Let \( T : [0,1]^2 \to [0,1] \) be a \( t\)–norm on \([0,1] \). Then \( T^* : (L^*)^2 \to L^* \) given by

\[
T^* ((x_1,y_1), (x_2,y_2)) = (T(x_1,x_2), S(y_1,y_2)),
\]

where \( S : [0,1]^2 \to [0,1] \) is a \( t\)–conorm on \([0,1] \) dual to \( T \) (based on duality w.r.t. \( n, n(x) = 1 - x \)). Obviously, \( T^* \) is a representable \( t\)–norm on \( L^* \).

Also \( F : (L^*)^2 \to L^* \) given by

\[
F((x_1,y_1), (x_2,y_2)) = (x_1x_2, y_1 \lor y_2)
\]

is a representable \( t\)–norm on \( L^* \) (linked to \( A = T_P \), product \( t\)–norm on \([0,1] \), and \( B = S_M \), maximum \( t\)–conorm on \([0,1] \)).

(ii) For a fixed \( t\)–norm \( T : [0,1]^2 \to [0,1] \), define a mapping \( D : (L^*)^2 \to L^* \) by

\[
D((x_1,y_1), (x_2,y_2)) = (T(x_1,x_2), S(1 - x_1,y_2) \land S(x_1,1 - x_2)).
\]
Evidently, $D$ is a symmetric $t$–seminorm on $L^*$ which is not representable. Observe that $D$ is associative and thus a $t$–norm on $L^*$. For more details see [12].

(iii) Define $C: (L^*)^3 \rightarrow L^*$ by

$$C((x_1,y_1),(x_2,y_2),(x_3,y_3)) = \left( \frac{x_1 + 2x_2 + x_3}{4}, \frac{y_1y_2^2y_3}{4} \right).$$

Then $C$ is an idempotent representable aggregation function on $L^*$ which is bisymmetric but not symmetric ($C$ is linked to a weighted arithmetic mean and to a weighted geometric mean).

Several interesting details, including deep results on Archimedean $t$–norms and $t$–norms making $L^*$ a residuated monoid can be found in [11,17].

6 Concluding Remarks

We have introduced and discussed aggregation functions on bounded partially ordered sets and on bounded lattices, stressing among others the pioneering work of E. E. Kerre and his research group in this domain. Our work opens new doors to the investigation of aggregation functions. For example, the class of strongly idempotent aggregation functions deserves a deeper look not only in the case of general bounded posets, but especially in the case of particular bounded lattices, such as $L^*$, $[0,1]^n$, etc. Our approach covers also the aggregation on cardinal and ordinal scales, on linguistic scales, as well as on special systems such as distributive functions.

Acknowledgement. The research summarized in this paper was supported by the Grants APVV–0012–07 and VEGA 1/0080/10.

References


On the Intuitionistic Fuzzy Implications and Negations. Part 1

Krassimir T. Atanassov

To my friend Prof. Etienne Kerre

Abstract. The paper contains review and new results related to the first 23 of the 138 already introduced by the author different intuitionistic fuzzy implications and the first 5 of the 36 different intuitionistic fuzzy negations. Some of their basic properties are studied. The validity and non-validity of the Law of Excluded Middle and its modifications, and of De Morgan’s Laws and their modifications are discussed. The axioms of Intuitionistic Logic are checked for the intuitionistic fuzzy implications and negations.

1 Introduction

During the last years, a discussion started related to the name “intuitionistic” given to the extension of the fuzzy set, introduced by the author in 1983 in [1]. He hopes that now, after [16], it is clear that the concept of “intuitionistic fuzzy set” introduced by Takeuti and Titani a year and a half later in [28] is a particular case of the ordinary fuzzy sets, while the author’s concept of an Intuitionistic Fuzzy Set (IFS) essentially extends the fuzzy sets. Also, the author hopes that it is already clear that the construction of the IFSs contains Brouwer’s idea for intuitionism [20]. In every way, the discussion was very useful for IFS theory.

First, and it is possible, that the most important author’s mistake was that more than 20 years the IFS theory used only the classical negation (denoted below by \(\neg_1\)), that generates a classical implication (denoted below by \(\rightarrow_4\)). Really, another implication (denoted below by \(\rightarrow_{11}\)) that generates

Krassimir T. Atanassov
Department of Bioinformatics and Mathematical Modelling,
Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 105,
Sofia-1113, Bulgaria
e-mail: krat@bas.bg
an intuitionistic fuzzy negation, was introduced in IFS theory in 1988 [2], but for a long period of time the author had not estimated their merits as an intuitionistic implication and negation and he had not used them because they are more complex than the standard ones. Now, the author realizes his mistake (better late than never) and hence the properties of the different versions of both these operations will be discussed below.

Second, searching arguments to defend the name of IFS, the author found a lot of new implications and negations. One part of them exhibits classical properties, another – intuitionistic properties and a third group exhibits more nonstandard properties. These new operations will be an object of discussion of the present and next author’s research.

We must mention that another approach for introducing intuitionistic fuzzy implications and negations is demonstrated by a series of papers by Chris Cornelis, Glad Deschrijver and Etienne Kerre [21, 22, 23, 24].

2 Definitions and Properties of 23 Intuitionistic Fuzzy Implications and 5 Intuitionistic Fuzzy Negations

2.1 Definitions of 23 Intuitionistic Fuzzy Implications

The first ten variants of intuitionistic fuzzy implications are discussed in [4], using the book [26] by George Klir and Bo Yuan as a basis, where the conventional fuzzy implications are given. Other five implications, defined by the author and his colleagues Boyan Kolev and Trifon Trifonov, are introduced in [2, 3, 7, 8, 15, 17]. These fifteen implications generated five negations by the formulae

\[ \neg x = x \rightarrow 0. \]  

On the other hand, these negations are the basis for new eight implications introduced in [9] through the formulae

\[ x \rightarrow y = \neg x \lor y \]  

and

\[ x \rightarrow y = \neg x \lor \neg \neg y. \]

The new (eight) implications generate negations that coincide with the respective negations generated by the corresponding implications, i.e., the generating process finishes. Therefore, using this scheme, we finally obtain 23 implications.

Below, we shall describe some properties of these implications and negations.

Let us denote below each of these implications by \( I(x, y) \) and the negations by \( N(x) \).
In intuitionistic fuzzy logic, if $x$ is a variable, then its truth-value is represented by the ordered pair 

$$V(x) = \langle a, b \rangle,$$

such that $a, b, a + b \in [0, 1]$, where $a$ and $b$ are degrees of validity and of non-validity of $x$.

Assume that for the three variables $x, y$ and $z$ there hold the equalities: 

$$V(x) = \langle a, b \rangle, V(y) = \langle c, d \rangle, V(z) = \langle e, f \rangle$$ 

$(a, b, c, d, e, f, a + b, c + d, e + f \in [0, 1])$.

For the needs of the discussion below, we define the notion of Intuitionistic Fuzzy Tautology (IFT, see [2, 3]) by,

$x$ is an IFT if and only if $a \geq b$,

while $x$ will be a tautology iff $a = 1$ and $b = 0$.

We define the following relation:

$$V(x) \leq V(y) \text{ iff } a \leq c \text{ and } b \geq d.$$ 

In some definitions, we use functions $sg$ and $\overline{sg}$, defined by

$$sg(x) = \begin{cases} 
1 & \text{if } x > 0 \\
0 & \text{if } x \leq 0 
\end{cases}, \quad \overline{sg}(x) = \begin{cases} 
0 & \text{if } x > 0 \\
1 & \text{if } x \leq 0 
\end{cases}$$

For two variables $x$ and $y$ operations “conjunction” ($\&$) and “disjunction” ($\lor$) are defined (see [2, 3]) by:

$$V(x \& y) = \langle \min(a, c), \max(b, d) \rangle, \quad V(x \lor y) = \langle \max(a, c), \min(b, d) \rangle.$$ 

In Table 1, we include the implications from [4], but also the implications, introduced by the author in [2, 7, 8, 15, 17] with coauthors Boyan Kolev [15] and Trifon Trifonov [17].

The correctness of the above definitions of implications is directly checked. For example, to check the validity of the definition of $\rightarrow_{15}$, we check that for every $a, b, c, d \in [0, 1]$ such that $a + b \leq 1$ and $c + d \leq 1$ for the expression,

$$X \equiv 1 - (1 - \min(b, c)).sg(\min(a, c) + \max(d - b)) - \min(b, c).sg(a - c).sg(d - b) + 1 - (1 - \max(a, d)).sg(\min(a, c) + \min(d - b)) - \max(a, d).\overline{sg}(a - c).\overline{sg}(d - b)$$

we obtain,

If $a \leq c$ and $b \geq d$, then,

$$X = 1 - (1 - \min(b, c)).sg(0 + 0) - \min(b, c).0 + 1$$

$$- (1 - \max(a, d)).sg(1 + 1) - \max(a, d).1$$

$$= 1 + 1 - (1 - \max(a, d)) - \max(a, d) = 1.$$
<table>
<thead>
<tr>
<th>Notation</th>
<th>Name</th>
<th>Form of implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rightarrow_1)</td>
<td>Zadeh</td>
<td>((\max(b, \min(a, c)), \min(a, d)))</td>
</tr>
<tr>
<td>(\rightarrow_2)</td>
<td>Gaines-Rescher</td>
<td>((1 - \text{sg}(a - c), \text{d.sg}(a - c)))</td>
</tr>
<tr>
<td>(\rightarrow_3)</td>
<td>G&quot;ödel</td>
<td>((1 - (1 - c)).\text{sg}(a - c), \text{d.sg}(a - c)))</td>
</tr>
<tr>
<td>(\rightarrow_4)</td>
<td>Kleene-Dienes</td>
<td>((\max(b, c), \min(a, d)))</td>
</tr>
<tr>
<td>(\rightarrow_5)</td>
<td>Lukasiewicz</td>
<td>((\min(1, b + c), \max(0, a + d - 1)))</td>
</tr>
<tr>
<td>(\rightarrow_6)</td>
<td>Reichenbach</td>
<td>((b + ac, ad))</td>
</tr>
<tr>
<td>(\rightarrow_7)</td>
<td>Willmott</td>
<td>((\min(\max(b, c), \max(a, b), \max(c, d)), \max(\min(a, d), \min(a, b), \min(c, d))))</td>
</tr>
<tr>
<td>(\rightarrow_8)</td>
<td>Wu</td>
<td>((1 - (1 - \min(b, c)).\text{sg}(a - c), \max(a, d).\text{sg}(a - c).\text{sg}(d - b)))</td>
</tr>
<tr>
<td>(\rightarrow_9)</td>
<td>K\l i\r e-\d\i\n\e\e-\D\i\n\e\n\s\s\s 1</td>
<td>((b + a''c, ab + a''d))</td>
</tr>
<tr>
<td>(\rightarrow_{10})</td>
<td>K\l i\r e\ and\ Y\u\a\n\ 2</td>
<td>((c.\text{sg}(1 - a) + \text{sg}(1 - a)))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>((\text{sg}(1 - c) + b.\text{sg}(1 - c)), \text{d.sg}(1 - a) + a.\text{sg}(1 - a).\text{sg}(1 - c)))</td>
</tr>
<tr>
<td>(\rightarrow_{11})</td>
<td>Atanassov 1</td>
<td>((1 - (1 - c)).\text{sg}(a - c), \text{d.sg}(a - c).\text{sg}(d - b)))</td>
</tr>
<tr>
<td>(\rightarrow_{12})</td>
<td>Atanassov 2</td>
<td>((\max(b, c), 1 - \max(b, c)))</td>
</tr>
<tr>
<td>(\rightarrow_{13})</td>
<td>Atanassov and Kolev</td>
<td>((b + c - b, c, a, d))</td>
</tr>
<tr>
<td>(\rightarrow_{14})</td>
<td>Atanassov and Trifonov</td>
<td>((1 - (1 - c)).\text{sg}(a - c) - d.\text{sg}(a - c).\text{sg}(d - b), \text{d.sg}(d - b)))</td>
</tr>
<tr>
<td>(\rightarrow_{15})</td>
<td>Atanassov 3</td>
<td>((1 - (1 - \min(b, c)).\text{sg}(\text{sg}(a - c) + \text{sg}(d - b)), \min(\max(b, c), \min(1 - b, d))))</td>
</tr>
<tr>
<td>(\rightarrow_{16})</td>
<td>(\text{max}(1 - \text{sg}(a), c), \min(\text{sg}(a), d)))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{17})</td>
<td>((\max(b, c), \min(a.b + a''c, d)))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{18})</td>
<td>((\max(b, c).\text{min}(1 - b, d)))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{19})</td>
<td>((\max(1 - \text{sg}(\text{sg}(a) + \text{sg}(1 - b)), c), \min(\text{sg}(1 - b), d)))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{20})</td>
<td>((\max(\max(1 - \text{sg}(a), 1 - \text{sg}(1 - \text{sg}(c))), \text{min}(\text{sg}(a), \text{sg}(1 - \text{sg}(c))))))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{21})</td>
<td>((\max(b, c(1 + d)), \min(a(a + b), d(c^2 + d + cd))))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{22})</td>
<td>((\max(b, 1 - d), \min(1 - b, d)))</td>
<td></td>
</tr>
<tr>
<td>(\rightarrow_{23})</td>
<td>((1 - \text{min}(\text{sg}(1 - b), \text{sg}(1 - \text{sg}(1 - d))), \text{min}(\text{sg}(1 - b), \text{sg}(1 - \text{sg}(1 - d)))))</td>
<td></td>
</tr>
</tbody>
</table>
If \( a \leq c \) and \( b < d \), then,
\[
X = 1 - (1 - \min(b, c)) \cdot \sg(0 + 1) - \min(b, c) \cdot 0.1 \\
+ 1 - (1 - \max(a, d)) \cdot \sg(1 + 0) - \max(a, d) \cdot 1.0 \\
= 1 - 1 + \min(b, c) + 1 - 1 + \max(a, d) \\
= \min(b, c) + \max(a, d) \leq 1.
\]
If \( a > c \) and \( b \geq d \), then,
\[
X = 1 - (1 - \min(b, c)) \cdot \sg(1 + 0) - \min(b, c) \cdot 1.0 \\
+ 1 - (1 - \max(a, d)) \cdot \sg(0 + 1) - \max(a, d) \cdot 0.1 \\
= 1 - 1 + \min(b, c) + 1 - 1 + \max(a, d) \\
= \min(b, c) + \max(a, d) \leq 1.
\]
If \( a > c \) and \( b < d \) then,
\[
X = 1 - (1 - \min(b, c)) \cdot \sg(1 + 1) - \min(b, c) \cdot 1.1 \\
+ 1 - (1 - \max(a, d)) \cdot \sg(0 + 0) - \max(a, d) \cdot 0.0 \\
= 1 - (1 - \min(b, c)) \cdot 1 - \min(b, c) + 1 - (1 - \max(a, d)) \cdot 0 \\
= 1 - 1 + \min(b, c) - \min(b, c) + 1 = 1.
\]
Therefore, implication \( \rightarrow_{15} \) is valid.

**Theorem 1.** If \( I \) is any of the 23 implications, then
\[
I((0, 1), (0, 1)) = (1, 0), \\
I((0, 1), (1, 0)) = (1, 0), \\
I((1, 0), (1, 0)) = (1, 0), \\
I((1, 0), (0, 1)) = (0, 1).
\]
Therefore, the restriction of each of these implications coincides over the constants false and true with the implication from the ordinary propositional calculus.

Now, we introduce an implication, inspired by G. Takeuti and S. Titani’s paper [28] and its answer by T. Trifonov and the author [16].

In [28] G. Takeuti and S. Titani introduced the following implication for \( p, q \in [0, 1] \)
\[
p \rightarrow q = \bigvee \{ r \in [0, 1] \mid p \land r \leq q \} = \begin{cases} 1, & \text{if } p \leq q \\ q, & \text{if } p > q \end{cases}
\]
In [16], its intuitionistic fuzzy extension is given in the form
\[ \langle a, b \rangle \rightarrow \langle c, d \rangle = \langle \max(c, \overline{\text{sg}}(a - c)), \min(d, \text{sg}(a - c)) \rangle \]
\[ = \begin{cases} 
(1, 0), & \text{if } a \leq c \text{ and } b \geq d \\
(1, 0), & \text{if } a \leq c \text{ and } b < d \\
\langle c, d \rangle, & \text{if } a > c \text{ and } b \geq d \\
\langle c, d \rangle, & \text{if } a > c \text{ and } b < d 
\end{cases} \]

There, it is proved that the latter implication coincides with Gödel’s implication \((\rightarrow_3)\), because if
\[ X \equiv 1 - (1 - c) \cdot \text{sg}(a - c) - \max(c, \overline{\text{sg}}(a - c)), \]
\[ Y \equiv d \cdot \text{sg}(a - c) - \min(d, \text{sg}(a - c)). \]
then we obtain the following. If \( a > c \)
\[ X = 1 - (1 - c) - \max(c, 0) = c - c = 0, \]
\[ Y = d \cdot 1 - \min(d, 1) = d - d = 0. \]
If \( a \leq c \)
\[ X = 1 - (1 - c) \cdot 0 - \max(c, 1) = 1 - 1 = 0, \]
\[ Y = d \cdot 0 - \min(d, 0) = 0 - 0 = 0, \]
i.e.
\[ \langle 1 - (1 - c) \cdot \text{sg}(a - c), d \cdot \text{sg}(a - c) \rangle = \langle \max(c, \overline{\text{sg}}(a - c)), \min(d, \text{sg}(a - c)) \rangle. \]

All constructions in [28] are re-written in [16] for intuitionistic fuzzy case with the aim to show that Takeuti and Titani’s sets are a particular case of intuitionistic fuzzy sets in the sense of [3]. Practically, the set constructed by them is an ordinary fuzzy set with elements, satisfying intuitionistic logic axioms. In [16], we show that there are intuitionistic fuzzy sets with elements, satisfying intuitionistic logic axioms.

Extending the research from [3], we shall study some properties of the 23 implications.

Let us introduce the expression
\[ I_i = x \rightarrow_i y, \]
where \( 1 \leq i \leq 23 \). We say that \( I_i \) is more powerful than \( I_j \) for \( 1 \leq i, j \leq 23 \), if
\[ V(I_i) \geq V(I_j) \text{ for all } x \text{ and } y. \]
We can construct Table 2, in which the lack of relation between two implications is denoted by “\(*\)".
Table 2 Relations between elements of set \( \{ I_i | 1 \leq i \leq 23 \} \)

| \( I_1 \) | \( I_2 \) | \( I_3 \) | \( I_4 \) | \( I_5 \) | \( I_6 \) | \( I_7 \) | \( I_8 \) | \( I_9 \) | \( I_{10} \) | \( I_{11} \) | \( I_{12} \) | \( I_{13} \) | \( I_{14} \) | \( I_{15} \) | \( I_{16} \) | \( I_{17} \) | \( I_{18} \) | \( I_{19} \) | \( I_{20} \) | \( I_{21} \) | \( I_{22} \) | \( I_{23} \) |
| \( I_1 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_2 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_3 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_4 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_5 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_6 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_7 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_8 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_9 \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{10} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{11} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{12} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{13} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{14} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{15} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{16} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{17} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{18} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{19} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{20} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{21} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{22} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| \( I_{23} \) | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |

There are many other implications, defined by Lilija Atanassova, Dimiter Dimitrov and the author, that will be discussed in a next author’s research.

2.2 Definitions of 5 Intuitionistic Fuzzy Negations

Now, following [5] and using, as a basis the equality (1) that has the intuitionistic fuzzy form

\[ \neg \langle a, b \rangle = \langle a, b \rangle \rightarrow \langle 0, 1 \rangle, \]

we shall construct negations, corresponding to the above implications. These negations are introduced in Table 3, but we must note that some of them have better (more compact) forms than those published already.

The following three properties are checked in [6] [7] for the separate negations:

- **Property P1:** \( A \rightarrow \neg \neg A \),
- **Property P2:** \( \neg \neg A \rightarrow A \),
- **Property P3:** \( \neg \neg \neg A = \neg A \).
Each of the negations from Table 3 satisfies Property 1.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Form of negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬₁</td>
<td>⟨b, a⟩</td>
</tr>
<tr>
<td>¬₂</td>
<td>⟨1 − sg(a), sg(a)⟩</td>
</tr>
<tr>
<td>¬₃</td>
<td>⟨b, a.b + a²⟩</td>
</tr>
<tr>
<td>¬₄</td>
<td>⟨b, 1−b⟩</td>
</tr>
<tr>
<td>¬₅</td>
<td>⟨sg(1−b), sg(1−b)⟩</td>
</tr>
</tbody>
</table>

Only negation ¬₁ satisfies Property 2, while the other negations do not satisfy it.

Each of the negations from Table 3 satisfies Property 3.

As an illustration of how the checks are performed, we shall prove the validity of the latest assertion for the case of negation ¬₅.

\[
\neg\neg\neg⟨a, b⟩ = \neg(\neg⟨\neg(1−b), \neg(1−b)⟩) \\
= \neg(\neg(1−\neg(1−b)), \neg(1−\neg(1−b))) \\
= (\neg(1−\neg(1−\neg(1−b))), \neg(1−\neg(1−\neg(1−b)))).
\]

Let

\[
X \equiv \neg(1−\neg(1−\neg(1−b))) − \neg(1−b).
\]

If \(b = 1\), then

\[
X = \neg(1−\neg(1)) − \neg(0) = \neg(0) − 1 = 0.
\]

If \(b < 1\), then

\[
X = \neg(1−\neg(1−1)) − 0 = \neg(1) = 0.
\]

Therefore Property 3 is valid for ¬₅.

Following [6], we show the relations between the different negations. By direct checks we can see the validity of these relations in Table 4.

The lack of relation between two implications is denoted in Table 4 by “∗”.

<table>
<thead>
<tr>
<th></th>
<th>¬₁</th>
<th>¬₂</th>
<th>¬₃</th>
<th>¬₄</th>
<th>¬₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬₁</td>
<td>≡</td>
<td>*</td>
<td>≤</td>
<td>≥</td>
<td>≥</td>
</tr>
<tr>
<td>¬₂</td>
<td>*</td>
<td>≡</td>
<td>*</td>
<td>*</td>
<td>≥</td>
</tr>
<tr>
<td>¬₃</td>
<td>≥</td>
<td>*</td>
<td>≡</td>
<td>≥</td>
<td>≥</td>
</tr>
<tr>
<td>¬₄</td>
<td>≤</td>
<td>*</td>
<td>≤</td>
<td>≡</td>
<td>≥</td>
</tr>
<tr>
<td>¬₅</td>
<td>≤</td>
<td>≤</td>
<td>*</td>
<td>≤</td>
<td>≡</td>
</tr>
</tbody>
</table>
Therefore, we can construct the ordered graph (see Fig. 1) with vertices, corresponding to negations $\neg_1, \ldots, \neg_5$.

George Klir and Bo Yuan discuss the following 9 axioms related to fuzzy implications in their book [26].

Axiom 1: $\forall x, y(x \leq y \rightarrow \forall z(I(x, z) \geq I(y, z)))$.
Axiom 2: $\forall x, y(x \leq y \rightarrow \forall z(I(z, x) \leq I(z, y)))$.
Axiom 3: $\forall y(I(0, y) = 1)$.
Axiom 4: $\forall y(I(1, y) = y)$.
Axiom 5: $\forall x(I(x, x) = 1)$.
Axiom 6: $\forall x, y, z(I(x, I(y, z)) = I(y, I(x, z)))$.
Axiom 7: $\forall x, y(I(x, y) = 1 \text{ iff } x \leq y)$.
Axiom 8: $\forall x, y(I(x, y) = I(N(y), N(x)))$, where $N$ is an operation for a negation.
Axiom 9: $I$ is a continuous function.

Following and extending [29] we shall note that Table 5 summarizes the list of axioms of Klir and Yuan which are satisfied by the 23 implications.

If some axiom is valid as an IFT, its number in Table 5 is marked by an asterisk (*). We should note that Axiom 8 is checked using the classical intuitionistic fuzzy negation ($\neg_1$); if it is valid using the respective generated by implication $I$ negation $N(x)$, then the axiom is marked as $8^N$. We should also note that the validity of Axiom 7 does not imply the validity of Axiom $7^*$. The forms of the axioms with asterisks are

Axiom 3*: $\forall y(I(0, y) \text{ is an IFT})$.
Axiom 5*: $\forall x(I(x, x) \text{ is an IFT})$.
Axiom 7*: $\forall x, y(I(x, y) \text{ is an IFT } \text{ iff } x \leq y)$.

The validity of each of these assertions can be checked directly. In [14] all 36 IF negations are given.
Table 5 List of axioms of Klir and Yuan that are satisfied by the 23 intuitionistic fuzzy implications

<table>
<thead>
<tr>
<th>Notation</th>
<th>Axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>→₁</td>
<td>2,3,4,5⁺,9</td>
</tr>
<tr>
<td>→₂</td>
<td>1,2,3,5</td>
</tr>
<tr>
<td>→₃</td>
<td>1,2,3,4,5,6</td>
</tr>
<tr>
<td>→₄</td>
<td>1,2,3,4,5⁺,6,8,9</td>
</tr>
<tr>
<td>→₅</td>
<td>1,2,3,4,5⁺,6,8,9</td>
</tr>
<tr>
<td>→₆</td>
<td>2,3,4,5⁺,9</td>
</tr>
<tr>
<td>→₇</td>
<td>3⁺,4,5⁺,8,9</td>
</tr>
<tr>
<td>→₈</td>
<td>1,2,3,5</td>
</tr>
<tr>
<td>→₉</td>
<td>2,3,4,5⁺</td>
</tr>
<tr>
<td>→₁₀</td>
<td>2,3,4</td>
</tr>
<tr>
<td>→₁₁</td>
<td>1,2,3,4,5,6</td>
</tr>
<tr>
<td>→₁₂</td>
<td>1,2,3,6,8,9</td>
</tr>
<tr>
<td>→₁₃</td>
<td>1,2,3,4,5⁺,6,8,9</td>
</tr>
<tr>
<td>→₁₄</td>
<td>1,2,3,4,5,6,7</td>
</tr>
<tr>
<td>→₁₅</td>
<td>1,2,3,5,7⁺,7⁺,8</td>
</tr>
<tr>
<td>→₁₆</td>
<td>1,2,3,4,6</td>
</tr>
<tr>
<td>→₁₇</td>
<td>2,3,4,5⁺,6,9</td>
</tr>
<tr>
<td>→₁₈</td>
<td>1,2,3,4,5⁺,6,9</td>
</tr>
<tr>
<td>→₁₉</td>
<td>1,2,3,4,6</td>
</tr>
<tr>
<td>→₂₀</td>
<td>1,2,3,5,6,8⁺N</td>
</tr>
<tr>
<td>→₂₁</td>
<td>1,2,3,5⁺,9</td>
</tr>
<tr>
<td>→₂₂</td>
<td>1,2,3,5⁺,6,8⁺N⁺,9</td>
</tr>
<tr>
<td>→₂₃</td>
<td>1,2,3,5,6,8⁺N</td>
</tr>
</tbody>
</table>

2.3 On the Law of Excluded Middle and Its Modifications

Following [6], we shall study the validity of the Law of Excluded Middle (LEM) in the forms:

\[ \langle a, b \rangle \lor \neg_i \langle a, b \rangle = \langle 1, 0 \rangle \]

(tautology-form) and

\[ \langle a, b \rangle \lor \neg_i \langle a, b \rangle = \langle p, q \rangle \]

(IFT-form), where \(1 \geq p \geq q \geq 0\) and \(i = 1, 2, ..., 5\).

We will also study a Modified LEM in the forms:

\[ \neg_i \neg_i \langle a, b \rangle \lor \neg_i \langle a, b \rangle = \langle 1, 0 \rangle \]

(tautology-form) and

\[ \neg_i \neg_i \langle a, b \rangle \lor \neg_i \langle a, b \rangle = \langle p, q \rangle , \]
On the Intuitionistic Fuzzy Implications and Negations

(IFT-form), where $1 \geq p \geq q \geq 0$ and $i = 1, 2, ..., 5$.
None of our negations satisfies the LEM in the tautological form.
Negations $\neg_1, \neg_3$ and $\neg_4$ satisfy the LEM in the IFT-form.
Only $\neg_2$ and $\neg_5$ satisfy the Modified LEM in the tautological form.
All negations satisfy the Modified LEM in the IFT-form.

As illustration, we shall calculate

$$\neg_5 \neg_5 \langle a, b \rangle \lor \neg_5 \langle a, b \rangle$$

$$= \langle \overline{sg}(1 - sg(1 - b)), sg(1 - sg(1 - b)) \lor \overline{sg}(1 - b), sg(1 - b) \rangle$$

$$= \langle \max(\overline{sg}(1 - sg(1 - b)), \overline{sg}(1 - b)), \min(sg(1 - sg(1 - b)), sg(1 - b)) \rangle.$$

If $b = 1$, then

$$\max(\overline{sg}(1 - sg(1 - b)), \overline{sg}(1 - b)) = \max(\overline{sg}(1), 1) = \max(0, 1) = 1.$$ and

$$\min(sg(1 - sg(1 - b)), sg(1 - b)) = \min(sg(1), 0) = \min(1, 0) = 0.$$

If $b < 1$, then

$$\max(\overline{sg}(1 - sg(1 - b)), \overline{sg}(1 - b)) = \max(\overline{sg}(0), 0) = \max(1, 0) = 1.$$ and

$$\min(sg(1 - sg(1 - b)), sg(1 - b)) = \min(sg(0), 1) = \min(0, 1) = 0.$$

Therefore, negation $\neg_5$ satisfies the Modified LEM in the IFT-form. On the other hand, in all cases the evaluation of the expression is equal to $\langle 1, 0 \rangle$, i.e., this negation satisfies the Modified LEM in the tautological form.

2.4 On De Morgan’s Laws and Their Modifications

Below, following [11], we shall discuss some forms of De Morgan’s Laws.
Usually, De Morgan’s Laws have the forms:

$\neg x \land \neg y = \neg(x \lor y),$

$\neg x \lor \neg y = \neg(x \land y).$

In [11] it is proved that for every two propositional forms $x$ and $y$:

$\neg_i x \land \neg_i y = \neg_i(x \lor y),$

$\neg_i x \lor \neg_i y = \neg_i(x \land y)$

for $i = 1, 2, 4, 5$, while negation $\neg_3$ does not satisfy these equalities.
We shall illustrate only the fact that the De Morgan’s Laws are not valid for \( i = 3 \). For example, if \( a = b = 0.5, c = 0.1, d = 0 \), then
\[
V(\neg_3 x \land \neg_3 y) = (0, 0.5),
\]
\[
V(\neg_3 (x \lor y)) = (0, 0.25).
\]

The above mentioned change of the LEM inspired the idea to study the validity of De Morgan’s Laws, which the classical negation \( \neg \) (here it is negation \( \neg_1 \)) satisfies. Really, it can be easily proved that the expressions
\[
\neg_1(\neg_1 x \vee \neg_1 y) = x \land y
\]
and
\[
\neg_1(\neg_1 x \land \neg_1 y) = x \lor y
\]
are IFTs, but the other negations do not satisfy these equalities. For them the following assertion is proved in [11].

For every two propositional forms \( x \) and \( y \) it holds that
\[
\neg_i(\neg_i x \lor \neg_i y) = \neg_i \neg_i x \land \neg_i \neg_i y
\]
and
\[
\neg_i(\neg_i x \land \neg_i y) = \neg_i \neg_i x \lor \neg_i \neg_i y
\]
for \( i = 2, 4, 5 \), while negation \( \neg_3 \) does not satisfy these equalities.

Now, we shall discuss another form of the LEM, also known as Law of Contradiction, that in propositional calculus is equivalent with the standard LEM, if De Morgan’s Laws were valid. It is the following:
\[
\neg(x \land \neg x).
\]

Negations \( \neg_2 \) and \( \neg_5 \) satisfy it as tautologies, negations \( \neg_1, \neg_3 \) and \( \neg_4 \) satisfy it as IFTs.

### 2.5 On One Baczynski and Jayaram Problem and Its Extension

In [18] Michal Baczynski and Balasubramaniam Jayaram formulated some problems related to fuzzy implications \( I \) and negations \( N \). Here we give a solution to one of them:

**Problem 1.7.1.** Give examples of fuzzy implications \( I \) such that
(i) \( I \) satisfies only property
\[
(CP) \quad I(x, y) = I(N(y), N(x))
\]
(ii) $I$ satisfies only property

\[(L - CP) \quad I(N(x), y) = I(N(y), x)\]

(iii) $I$ satisfies both (CP) and (L-CP), but not

\[(R - CP) \quad I(x, N(y)) = I(y, N(x))\]

with some fuzzy negation $N$, where $x, y \in [0, 1]$.

We must note that in [18] no example is given.

Here we shall give examples of pairs of implications and negations that satisfy Problem 1.7.1 (ii) and other problems. Let the pair $(m, n)$ denote the expression with $m$-th implication and $n$-th negation.

Following [12], first, we shall formulate the following

**Theorem 2.** The pairs $(4, 1), (5, 1), (7, 1), (12, 1), (13, 1), (15, 1), (20, 2), (22, 4), (23, 5)$, satisfy the three axioms.

We must note that no one pair for $1 \leq m \leq 23$ and $1 \leq n \leq 5$ satisfy exactly two axioms and more precisely, axioms $(L - CP)$ and $(R - CP)$. In [12], solutions are given for $1 \leq m \leq 138$ and $1 \leq n \leq 34$.

We had not found any pair of implication and negation that are solution of Problem 1.7.1 (iii).

Also, we had not found any pair of implication and negation that satisfy only the first axiom, i.e., we cannot give examples for the case of Problem 1.7.1 (i).

Another result of our search [12] is

**Theorem 3.** The pairs $(2, 2), (3, 2), (8, 2), (11, 2), (16, 2), (12, 3), (17, 3), (12, 4), (18, 4), (14, 5), (15, 5), (19, 5)$ satisfy only the axiom $(R - CP)$.

The most interesting is the following

**Theorem 4.** Only the pair $(21, 3)$ satisfies only Axiom $(L - CP)$.

We must note that this assertion has 135 solutions, described in [12] in the case when $23 < m \leq 138$ or $5 < n \leq 34$.

Now, we shall extend the Baczynski and Jayaram problem and shall give solutions to its new form.

Using the idea for Modified LEM, we can replace the equalities $(CP)$, $(L - CP)$ and $(R - CP)$ with the equalities:

\[
I(N(N(x)), N(N(y))) = I(N(y), N(x)), \quad (CP')
\]
\[
I(N(x), N(N(y))) = I(N(y), N(N(x))), \quad (L - CP')
\]
\[
I(N(N(x)), N(y)) = I(N(N(y)), N(x)). \quad (R - CP')
\]
Here, we give examples of pairs of implications and negations that satisfy an extended form of Problem 1.7.1 (ii), where the assertions are related to $(CP')$, $(L-CP')$ and $(R-CP')$.

In [13], it is shown that there are 1322 pairs $(m, n)$ satisfying the three equalities, when $1 \leq m \leq 138$ and $1 \leq n \leq 34$. But, no one pair $(m, n)$ for $1 \leq m \leq 23$ and $1 \leq n \leq 5$ satisfies exactly two equalities, namely $- (L-CP')$ and $(R-CP')$.

We had not found any pair of implication and negation for $1 < m \leq 138$ and $1 < n \leq 34$, that solve the extended Problem 1.7.1 (iii).

Also, we had not found any pair of implication and negation that satisfy the first equality only, i.e., we cannot give examples for the case of Problem 1.7.1 (i).

Another result of our search is

**Theorem 5.** Only the pair $(21, 3)$ satisfies only equality $(R-CP')$.

The most interesting is the following

**Theorem 6.** The pairs $(12, 3), (17, 3)$ satisfy only equality $(L-CP')$.

We must note that this assertion has 375 solutions, described in [13] in the case when $23 < m \leq 138$ or $5 < n \leq 34$.

### 2.6 On the Axioms of Propositional Intuitionistic Logic

The next and more important question is which of the introduced implications satisfy all the axioms of Propositional Intuitionistic Logic (IL) (see for example [27]).

The validity of the IL axioms was already checked for some implications in [15] [25] [29]. Here, we give a full list of valid axioms for each of the 23 implications. We will again verify the validity axioms in two variants - tautological validity (Table 6) and IFT validity (Table 7).

We use the following list of axioms for propositional intuitionistic logic:

1. $A \rightarrow A$,
2. $A \rightarrow (B \rightarrow A)$,
3. $A \rightarrow (B \rightarrow (A \land B))$,
4. $(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$,
5. $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$,
6. $A \rightarrow \neg \neg A$,
7. $\neg(A \land \neg A)$,
8. $(\neg A \lor B) \rightarrow (A \rightarrow B)$,
9. $\neg(A \lor B) \rightarrow (\neg A \land \neg B)$,
10. $(\neg A \land \neg B) \rightarrow (\neg A \lor B)$,
11. $(\neg A \lor B) \rightarrow (\neg A \land B)$,
12. $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$,
13. $(A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$,
(n) \( \neg \neg \neg A \rightarrow \neg A \),  
(o) \( \neg A \rightarrow \neg \neg A \),  
(p) \( \neg (A \rightarrow B) \rightarrow (A \rightarrow \neg B) \),  
(q) \( (C \rightarrow A) \rightarrow ((C \rightarrow (A \rightarrow B)) \rightarrow (C \rightarrow B)) \).

Table 6 List of axioms of the intuitionistic logic that are satisfied by intuitionistic fuzzy implications as tautologies

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
<th>n</th>
<th>o</th>
<th>p</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The most important of the results collected in Table 6 can be formulated as the following

**Theorem 7.** Implications \( \rightarrow_3, \rightarrow_{11}, \rightarrow_{14}, \rightarrow_{20}, \rightarrow_{23} \) satisfy all intuitionistic logic axioms as tautologies.

The validity of axioms for cells marked by a question mark (?) in Tables 6 and 7 is not yet clear and it is an open problem.

The most important of the results collected in Table 7 can be formulated as follows:

**Theorem 8.** Implications \( \rightarrow_1, \rightarrow_3, \rightarrow_4, \rightarrow_5, \rightarrow_{11}, \rightarrow_{14}, \rightarrow_{18}, \rightarrow_{20}, \rightarrow_{22}, \rightarrow_{23} \) satisfy all intuitionistic logic axioms as IFTs.
Table 7 List of axioms of the intuitionistic logic that are satisfied by intuitionistic fuzzy implications as IFTs

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
<th>n</th>
<th>o</th>
<th>p</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>→1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>→23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 A New Argument that the Intuitionistic Fuzzy Sets Are ‘Intuitionistic’ in Nature

The above assertions show that all negations, except the first one, satisfy conditions of the intuitionistic logic, but not of the classical logic. A part of these negations were generated by implications, that were generated by fuzzy implications. Now, let us return from the intuitionistic fuzzy negations to ordinary fuzzy negations. The result is shown on Table 8, where \( b = 1 - a \).

Another approach to obtaining fuzzy negations from IF negations is to use the ideas from [19], where a procedure for de-intuitionistic-fuzzification is discussed. It will give the negations from Table 9.

Now, if we put \( b = 1 - a \), we will again obtain the values from Table 8.

Therefore, from the intuitionistic fuzzy negations we can generate fuzzy negations, so that two of them (\( \neg_3 \) and \( \neg_4 \)) coincide with the standard fuzzy negation (\( \neg_1 \)). Therefore, there are intuitionistic fuzzy negations that lose
Table 8 List of the fuzzy negations, generated by intuitionistic fuzzy negations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Form of the intuitionistic fuzzy negation</th>
<th>Form of the fuzzy negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬₁</td>
<td>⟨b, a⟩</td>
<td>1 − a</td>
</tr>
<tr>
<td>¬₂</td>
<td>⟨1 − sg(a), sg(a)⟩</td>
<td>1 − sg(a)</td>
</tr>
<tr>
<td>¬₃</td>
<td>⟨b, a.b + a²⟩</td>
<td>1 − a</td>
</tr>
<tr>
<td>¬₄</td>
<td>⟨b, 1 − b⟩</td>
<td>1 − a</td>
</tr>
<tr>
<td>¬₅</td>
<td>⟨sg(1 − b), sg(1 − b)⟩</td>
<td>1 − sg(a)</td>
</tr>
</tbody>
</table>

Table 9 List of the fuzzy negations, generated by intuitionistic fuzzy negations using the procedure for de-intuitionistic-fuzzification

<table>
<thead>
<tr>
<th>Notation</th>
<th>Form of the intuitionistic fuzzy negation</th>
<th>Form of the fuzzy negation</th>
</tr>
</thead>
<tbody>
<tr>
<td>¬₁</td>
<td>⟨b, a⟩</td>
<td>1 − a + b</td>
</tr>
<tr>
<td>¬₂</td>
<td>⟨1 − sg(a), sg(a)⟩</td>
<td>1 − sg(a)</td>
</tr>
<tr>
<td>¬₃</td>
<td>⟨b, a.b + a²⟩</td>
<td>1 + b − a.b − a²</td>
</tr>
<tr>
<td>¬₄</td>
<td>⟨b, 1 − b⟩</td>
<td>b</td>
</tr>
<tr>
<td>¬₅</td>
<td>⟨sg(1 − b), sg(1 − b)⟩</td>
<td>sg(1 − b)</td>
</tr>
</tbody>
</table>

their properties when they are restricted to ordinary fuzzy case. In other words, the construction of the intuitionistic fuzzy estimation

⟨degree of membership/validity, degree of non-membership/non-validity⟩

which is specific for the intuitionistic fuzzy sets, is the reason for the intuitionistic behaviour of these sets. Over them we can define both intuitionistic fuzzy, and classical negations.

In the fuzzy case, negations ¬₂ and ¬₅ coincide, generating a fuzzy negation that satisfies properties P1 and P3 and does not satisfy Property P2 from Section 2.2, i.e., it is ‘intuitionistic’ in nature. In these cases, the implication from properties P1 and P2 may be any that generates negation ¬₂.

4 Etienne Kerre’s Group and Intuitionistic Fuzzy Implications

In the beginning of the century, Prof. Etienne Kerre established in Ghent University one of the most active research groups in the world in the area of IFS theory. The author regrets that he does not dispose of all their publications, but among these, that he has got, there are very important ones. Papers [21, 22, 23, 24] are some of them. Their authors are Etienne Kerre and two of his then PhD-students – Chris Cornelis and Glad Deschijver, who nowadays are active and prospective scientists.
In a relation to our research, the results of Kerre’s group on intuitionistic fuzzy implications generate some interesting problems. Below, we will mention two of them:

1. By analogy with our research on axioms of Klir and Yuan that are satisfied by the 23 intuitionistic fuzzy implications, to check which of these implications satisfy the axioms discussed in [21, 24].
2. To check the properties of the implications, discussed in [21, 22, 23, 24] for the introduced here 23 implications and for the rest implications, that already are defined in IFS theory.

5 Conclusion

The new operations open a very large field for future research. We hope that soon many other properties will be clarified and new results will be obtained. There are some possible directions that will be an object of an author’s future research:

1. Study the relationships between the different implications and order them as vertices of an oriented graph with respect to the ordering “≤”. Similarly to the relationships between the different negations, discussed above.
2. Study all pairs (→i, ¬j) and determine which of them have nice properties, e.g. satisfy all axioms of intuitionistic logic.
3. Construct new sets of the existing implications and study their properties.

The most interesting problem is related to the theory of T- and S-norms. The results, related to De Morgan Laws show that the present theory of T- and S-norms must be revised totally, because it is based on De Morgan Laws and therefore, on the classical negation. In future, it must be modified, based on the different types of non-classical negations.

Acknowledgements. This paper is partially supported by grant BIn-2/09 “Design and development of intuitionistic fuzzy logic tools in information technologies” of the Bulgarian National Science Fund.

The author wishes to express his gratitude to Glad Deschrijver and Peter Vassilev for the valuable advice and corrections.

References

Mathematical Fuzzy Logic: A Good Theory for Practice*

Vilém Novák and IrinaPerfilieva

Abstract. We discuss the present state of mathematical fuzzy logic in narrow sense, its extension — fuzzy logic in broader sense (FLb) as a logic of natural human reasoning and also some related theories, e.g., the fuzzy transform. We argue that these are good theories with potential to be very practical.

Keywords: Soft computing, Fuzzy logic in broader sense, Fuzzy decision-making, Evaluative linguistic expressions, Fuzzy transform.

1 Introduction

“There is nothing more practical than a good theory” — this sentence, which has probably been first proclaimed by the psychologist K. Lewin in [27], characterizes in a very succinct form essentially important balance between basic and applied research. We are convinced that only such balance can lead to really new innovative solutions which can drive our society forward and bring also tangible economical benefit. This can be especially visible in physics where only good theory brought us to the atomic age (with all its positive and also negative impacts), but also many other mathematical theories such as the theory of differential equations, elasticity theory, control theory, statistics, etc.

One can ask whether also fuzzy logic is such a practical theory. We argue for the positive: the mathematical fuzzy logic (MFL), its extension — fuzzy logic in broader...
sense (FLb) and some related theories such as fuzzy approximation of functions or fuzzy transform are very practical.

Recall that MFL is a special mathematical theory the goal of which is to provide a working mathematical model of the vagueness phenomenon and to become a well established sound formal system that makes its applications well justified. The mathematization of vagueness is based on introduction of degrees of truth taken from some scale. Many people have significantly contributed to its development, for example P. Hájek, F. Esteva, L. Godo, S. Gottwald, F. Montagna, P. Cintula, D. Mundici, G. Gerla, V. Novák, and others. There are several books [3, 16, 18, 19, 46, 48] and many papers published on this topic, which are hard to be listed. Perhaps the most comprehensive list can be found on the WEB page www.mathfuzzlog.org/. In this paper, we will provide a brief overview of the present state of mathematical fuzzy logic and the mentioned related topics.

2 Mathematical Fuzzy Logic

2.1 Fuzzy Logic in Narrow Sense

Let us remember that the seminal paper [67] on fuzzy set theory written by L. A. Zadeh has been followed by the paper on fuzzy logic [17] written by J. A. Goguen. The latter can be taken as the seminal paper on MFL because the residuated lattice has been proposed there as a convenient structure of truth values. The paper contains a lot of solid material accompanied by the detailed philosophical discussion but it is not sufficiently formal. The first, mathematically highly sophisticated paper is that of J. Pavelka [50]. Since then, MFL has been significantly developed.

Fuzzy logic with traditional syntax

This branch of MFL has been elaborated in a deep and seminal way by P. Hájek in [19]. It is specific for it that formulas are dealt with classically, i.e. the language of such fuzzy logic differs from that of classical logic especially by having more connectives. On the other hand, the concepts of inference rule, theory, proof and many other ones remain classical.

The semantics is many valued and based on the concept of a commutative, integral, bounded residuated lattice \( L \) (shortly, residuated lattice) which is an algebra of type \((2, 2, 2, 0, 0)\)

\[
L = \langle L, \lor, \land, \otimes, \rightarrow, 0, 1 \rangle,
\]

where \( \langle L, \lor, \land, 0, 1 \rangle \) is a lattice with the bottom element \( 0 \) and the top element \( 1 \).

The operation \( \otimes \) is called multiplication and \( \langle L, \otimes, 1 \rangle \) is a commutative monoid. The operation \( \rightarrow \) is the residuation operation tied with the multiplication by adjointness: for all \( a, b, c \in L \) it holds that

\[
a \otimes b \leq c \quad \text{iff} \quad a \leq b \rightarrow c.
\]
A BL-algebra is a residuated lattice fulfilling, moreover, prelinearity \(((a \rightarrow b) \lor (b \rightarrow a) = 1)\) and divisibility \((a \otimes (a \rightarrow b) = a \land b)\).

The most prominent is the standard Łukasiewicz MV-algebra (taken as the residuated lattice)

\[
\mathcal{L}_\mathbb{L} = \langle [0,1], \lor, \land, \otimes, \rightarrow, 0, 1 \rangle
\]

where

\[
a \otimes b = \max(0, a + b - 1), \quad \text{(Łukasiewicz conjunction)} \\
a \rightarrow b = \min(1, 1 - a + b) \quad \text{(Łukasiewicz implication)}
\]

Quite recently, a new algebra of truth values called EQ-algebra has been introduced in \[44\]. It is not identical with residuated lattice. It is specific by taking fuzzy equality \(\sim\) as the basic operation so that the implication operation is derived:

\[a \rightarrow b = (a \land b) \sim a.\]

The study of residuated lattices is closely connected with the study of t-norms (cf. \[26\] and also \[22\], \[24\]) and it is important for the study of fuzzy logics (with traditional syntax). One of the first logics in which conjunction is interpreted by a t-norm has been proposed in \[6\]. At present, there are many formal systems of MFL, both propositional as well as predicate. The prominent fuzzy logic calculi are MTL, IMTL, BL, product, or Łukasiewicz (cf. \[40\]). All of them are embraced in a wide class of the, so called, core fuzzy logics \[4, 20\].

All core fuzzy logics enjoy completeness in the form

\[
\vdash A \iff \models A
\]

for every formula \(A\) where \(\models A\) means that \(A\) is true in the degree 1 \((M(A) = 1)\) in every interpretation \(M\) of the language of the given core fuzzy logic.

MFL has been studied in detail especially from the algebraic point of view. One can hardly estimate, how far this work can still continue. It seems that MFL is in the position of a theory promising a deeply justified technique for modeling of various manifestations of the vagueness phenomenon and for various applications. There are not so many results in the latter, however. One of the exceptions is the program of formalization of fuzzy set theory — fuzzy mathematics — using fuzzy logic as a metatheory that was announced by P. Cintula and L. Běhounek in \[1\], \[2\]. Several papers have already been published and the program seems to be very promising because it encompasses and quite often even generalizes most (if not all) of the results in fuzzy set theory obtained so far. Another promising program is development of fuzzy logic in broader sense which we will discuss below.

**Fuzzy type theory**

It turns out that a fully-fledged treatment of vagueness cannot be made using first-order fuzzy logic only, especially when the meaning of natural language is at play. Thus, the need for higher-order fuzzy logic raised up. This has been developed in
And called the fuzzy type theory (FTT). It generalizes classical type theory initiated by B. Russel, A. Church and L. Henkin. Let us mention that another version of higher-order fuzzy logic has been developed in [1] which focuses especially on construction of special classes and which thus aims at becoming the general theory of “fuzzy mathematics” mentioned above.

The structure of truth values of FTT is generally supposed to form one of the following: a complete IMTL$_{\Delta}$-algebra (see [13]), standard Łukasiewicz algebra, ŁΠ-algebra, or BL-algebra. The most important for applications in linguistics is the standard Łukasiewicz algebra (3) extended by the operation

$$\Delta(a) = \begin{cases} 1 & \text{if } a = 1, \\ 0 & \text{otherwise.} \end{cases}$$

which raises boolean structure from the corresponding algebra.

Important concept in FTT is that of a fuzzy equality. This is a fuzzy relation $\equiv : M \times M \rightarrow L$ which has the following properties:

(i) reflexivity $[m \equiv m] = 1$,
(ii) symmetry $[m \equiv m'] = [m' \equiv m]$,
(iii) $\otimes$-transitivity $[m \equiv m'] \otimes [m' \equiv m''] \leq [m \equiv m'']$

for all $m, m', m'' \in M$ where $[m \equiv m']$ denotes a truth value of $m \equiv m'$.

A special case of fuzzy equality on the algebra of truth values is biresiduation. Example of a fuzzy equality on $M = \mathbb{R}$ with respect to the standard Łukasiewicz algebra is

$$[m \equiv n] = 1 - (1 \wedge |m - n|), \quad m, n \in \mathbb{R}.$$ 

Syntax of FTT is a generalization of the lambda-calculus that is constructed in a classical way. The main difference from classical type theory is thus in definition of additional special connectives. Note that all essential syntactical elements of FTT are formulas (alternatively, they can be called lambda-terms as is usual in classical type theory).

As usual, each formula $A$ has a certain type. The basic types are $\omega$ (truth values) and $\varepsilon$ (elements). These can be then iterated to more complex types.

Formulas of type $\omega$ (truth value) can be joined by the following connectives: $\equiv$ (equivalence), $\lor$ (disjunction), $\land$ (conjunction), $\&$ (strong conjunction), $\nabla$ (strong disjunction), $\Rightarrow$ (implication). General $(\forall)$ and existential $(\exists)$ quantifiers are defined as special formulas. For the details about their definition and semantics — see [36].

If $A \in \text{Form}_{\omega\alpha}$, then $A$ represents a fuzzy set of elements. It can also be understood as a first-order property of elements of the type $\alpha$. Similarly, $A_{(\omega\alpha)\alpha}$ is a fuzzy relation (between elements of type $\alpha$).

There are over 15 logical axioms in FTT depending on the chosen structure of truth values. FTT has two inference rules and classical concept of provability. The rules of modus ponens and generalization are derived rules.

A theory $T$ is a set of formulas of type $\omega$. A specific formula of type $\omega$ is a formula $\dagger$ representing the most indefinite truth value for which $\vdash \neg \neg \dagger \equiv \dagger$ holds. Note that
such a truth value, in general, needs not be present in the given algebra of truth values. In the standard Łukasiewicz algebra is \( \dagger \) interpreted by 0.5.

A formula \( \Delta A_o \) is crisp, i.e., its interpretation is either 0 or 1. There are formulas which are not crisp.

Quite useful are the following special crisp formulas:

\[
\Upsilon_{oo} \equiv \lambda z_o \cdot \neg (\neg z_o),
\]

\[
\hat{\Upsilon}_{oo} \equiv \lambda z_o \cdot \neg (z_o \lor \neg z_o).
\]

Thus, the formula \( \Upsilon_{oo} z_o \) says that \( z_o \) is non-zero truth value and \( \hat{\Upsilon}_{oo} z_o \) says that \( z_o \) is a general truth value between 0 and 1.

Note that \( \Delta \) corresponds to \( D \)-operator of supervaluation theory (e.g., \( A \vdash C \) implies \( \vdash \Delta A \Rightarrow C \)) and \( \hat{\Upsilon} \) corresponds to \( I \)-operator indefinitely.

The semantics of FTT is defined using generalization of the concept of a frame that is a system \( \langle (M_\alpha, =_\alpha)_{\alpha \in \text{Types}}, L \rangle \) where \( =_\alpha \) is a special fuzzy equality in each set \( M_\alpha \). Interpretation of each formula \( A_\beta \alpha \) is a function assigning to every \( m \in M_\alpha \) an element from \( M_\beta \).

**Theorem 1 (completeness, [36])**

(a) A theory \( T \) of FTT is consistent iff \( T \) has a general model.

(b) \( T \vdash A_o \) iff \( T \models A_o \) holds for every theory \( T \) and a formula \( A_o \).

We claim that all essential properties of vague predicates are formally expressible in FTT and so, they have a many-valued model.

**Fuzzy logic with evaluated syntax (EvŁ)**

This logic is specific by considering formulas that are evaluated also on syntactic level. Namely, we deal with evaluated formulas of the form \( a / A \) where \( A \) is a formula and \( a \in L \) is its syntactic evaluation. This has nice interpretation since it allows to consider axioms which need not be fully satisfactory and so, their initial truth value can be lower than 1.

This logic has been in propositional version established by J. Pavelka in [50] and extended to predicate version by V. Novák in [33] and especially in the book [48]. EvŁ has been established as a generalization of classical mathematical logic with clearly distinguished syntax and semantics. The syntax consists of precise definitions of evaluated formula, evaluated proof, fuzzy theory, model, evaluated provability, etc. and the semantics is formed by the standard Łukasiewicz algebra only. The reason for choosing the latter follows from Theorem 1.7 in [50], Part III, which states that completeness in the sense of [7] below requires the following four equations to hold in the given algebra of truth values:

\[
\bigvee_{i \in I} (a \rightarrow b_i) = a \rightarrow \bigvee_{i \in I} b_i, \quad \bigwedge_{i \in I} (a \rightarrow b_i) = a \rightarrow \bigwedge_{i \in I} (b_i) \tag{5}
\]

\[
\bigvee_{i \in I} (a_i \rightarrow b) = \bigwedge_{i \in I} a_i \rightarrow b, \quad \bigwedge_{i \in I} (a_i \rightarrow b) = \bigvee_{i \in I} a_i \rightarrow b \tag{6}
\]
These equations are in $[0, 1]$ equivalent to continuity of $\rightarrow$, which is fulfilled only by the Łukasiewicz implication and its isomorphs. The resulting logic is quite strong since it is complete with respect to the generalized syntax, in which all formulas are evaluated by elements from the underlying algebra. A fuzzy theory $T$ is determined by a fuzzy set of axioms. Hence, besides the concept of truth of a formula $A$ in a theory $T$, $T \models_a A$, which is infimum of truth values of $A$ in all models of $T$, also the provability degree of $A$ in $T$ is introduced: $T \vdash_a A$, which is supremum of values of all proofs of $A$ in $T$. Then we obtain generalization of the Gödel-Henkin completeness theorem:

$$T \vdash_a A \iff T \models_a A, \quad a \in L,$$

for all all formulas $A \in F_{J(T)}$ and all fuzzy theories $T$. The system of $Ev_L$ is open to extension by new connectives.

### 2.2 Fuzzy Logic in Broader Sense

The paradigm of fuzzy logic in broader sense\footnote{It should be noted that FLb should be distinguished from the very general concept of fuzzy logic in wide sense announced informally by L. A. Zadeh in the nineties.} was proposed by V. Novák in 1995 in [34] as a program for extension of $FL_n$, which aims at developing a formal theory of natural human reasoning that would include mathematical models of special expressions in natural language with regard to their vagueness. This program overlaps with two other paradigms proposed in the literature, namely commonsense reasoning and precisiated natural language (PNL).

The idea of commonsense reasoning has been proposed by J. McCarthy in [31] as a part of the program of logic-based artificial intelligence. Its paradigm is to develop formal commonsense theories and systems using mathematical logic that exhibit commonsense behavior. The reason is that commonsense reasoning is a central part of human behavior and no real intelligence is possible without it.

The main drawback of the up-to-date formalizations of commonsense reasoning, in our opinion, is that it neglects vagueness present in the meaning of natural language expressions (cf. [5] and the citations therein).

The concept of PNL, which has been proposed by L. A. Zadeh in [70], is based on two main premises:

(a) Much of the world knowledge is perception based,
(b) perception based information is intrinsically fuzzy.

It is important to stress that the term precisiated natural language means “a reasonable working formalization of semantics of natural language without pretension to capture it in detail and fineness”. Its goal is to provide an acceptable and applicable technical solution. It should also be noted that the term perception is not considered here as a psychological term but rather as a result of human, intrinsically imprecise measurement.
PNL methodology requires the presence of World Knowledge Database and Multiagent, Modular Deduction Database. The former contains all the necessary information including perception based propositions describing the knowledge acquired by direct human experience, which can be used in the deduction process. No exact formalization of PNL, however, has been developed until recently, and so it should be taken mainly as a reasonable methodology.

The concept of FLb is a glue between both paradigms that should consider the best of each. During the years, it has been slowly developed and so far, it consists of the following theories:

(a) Formal theory of evaluative linguistic expressions, which is in detail explained in [42].
(b) Formal theory of fuzzy IF-THEN rules which is presented in [11, 45].
(c) Formal theory of perception-based logical deduction, which is presented in [9, 37, 38, 47].
(d) Formal theory of intermediate and generalized quantifiers presented in [10, 21, 39, 43].

The most convenient formal system as the basis for further development of FLb seems to be the fuzzy type theory. One of the reasons supporting the latter is argumentation of many logicians and linguists (cf. [29, 30, 66]) that the first order logic is not sufficient for capturing semantics of natural language.

Evaluative linguistic expressions

This theory is one of the most essential constituents of FLb. Recall that evaluative expressions are expressions of natural language, for example, small, medium, big, about twenty five, roughly one hundred, very short, more or less deep, not very tall, roughly warm or medium hot, quite roughly strong, roughly medium size, and many others. They form a small but very important part of natural language and they are present in its everyday use any time. The reason is that people very often need to evaluate phenomena around them. Moreover, they often make important decisions based on them, learn how to control, and many other activities. Note that from the linguistic point of view, they also include the, so called, gradable adjectives (cf. [23]). We argue that the meaning of evaluative expressions is a fundamental bearer of the vagueness phenomenon and namely, that its vagueness is a consequence of the indiscernibility between objects.

All the details about formal theory of evaluative linguistic expressions can be found in [42]. We distinguish intension (a property), and extension in a given context of use (i.e., a possible world; see [14]). Mathematical representation of intension is a function defined on a set of contexts which assigns to each context a fuzzy set of elements. Intension is invariant with respect to change of context. Extension is a class of elements (i.e., a fuzzy set) determined by intension when setting a specific context. It depends on the particular context of use and changes whenever the context is changed. For example, the expression “large town” is a name of an intension being a property of some feature of objects, i.e. number of people in a town. Its meaning
can be, e.g., 100 thousand people in the Czech Republic, 1 mil. people in France, or 10 mil. people in Asia.

The pure evaluative expressions are expressions of the form

\[ \langle \text{linguistic hedge} \rangle \langle \text{atomic evaluative expression} \rangle \]

where linguistic hedges are e.g. very, more or less and atomic evaluative expressions are gradable adjectives small, medium and big, and many others. We must also distinguish the meaning of evaluative predications that are linguistic expressions of the form

\[ X \text{ is } \langle \text{linguistic hedge} \rangle \langle \text{atomic expression} \rangle \]

where \( X \) is a variable for values of some specific feature of objects (e.g., temperature, pressure, height, depth, etc.).

There are good means in FTT using which notions of context, horizon, and many other general characteristics of the meaning of evaluative expressions can be modeled. The result is a formal logical theory of evaluative linguistic expressions \( T_{Ev} \) which is determined by 11 special axioms.

Interpretation of extensions of evaluative expressions in a model is depicted in Fig. 1. The context is determined by the triple \( \langle v_L, v_S, v_R \rangle \) of points (left bound, middle point, and right bound). The \( LH, MH \) and \( RH \) are fuzzy sets interpreting left, medium and right horizon, respectively which are determined by a special fuzzy equality \( \approx_w \) constructed for each context \( w \) from one universal one. The \( v_{a,b,c} \) is a function specific for each linguistic hedge \( \nu \) which represents the corresponding horizon shift. The composition of horizon and its shift provides extension of the evaluative predication (8) being a fuzzy set of elements. Note that there are various mathematical models of linguistic hedges which, however, in most cases do not meet all the necessary requirements. A nice overview of them can be found in [25].

It is possible to construct a model of the theory \( T_{Ev} \) of evaluative expression. Hence, using the completeness theorem, we can prove the following:

**Theorem 2.** The theory of evaluative linguistic expressions is consistent.

This theorem is a strong theoretical support for the functionality of our theory.
Fuzzy IF-THEN rules and intermediate quantifiers

The theory of evaluative expressions is the point of departure for more advanced theories of FLb. The first of them is the theory of fuzzy IF-THEN rules that are conditional expressions of natural language having the form

\[
\text{IF } X \text{ is } A \text{ THEN } Y \text{ is } B. \tag{9}
\]

The \( A, B \) in (9) are just evaluative linguistic expressions, or more precisely, the fuzzy IF-THEN rule consists of two (or more) evaluative linguistic predications joined by an implication. Sets (finite) of fuzzy IF-THEN rules are called linguistic descriptions and they are the fundamental tool, on the basis of which are can describe various decision, control or other situations which can give us more information about the reality.

The principal method for derivation of a conclusion on the basis of a linguistic description in perception-based logical deduction (see [37, 47]). We will explain the main idea on the following example.

**Example 1.** Given a linguistic description \( LD \)

\[
\mathcal{R}_1 := \text{IF } X \text{ is small THEN } Y \text{ is very big}, \\
\mathcal{R}_2 := \text{IF } X \text{ is very big THEN } Y \text{ is small}.
\]

Each rule provides us with a certain knowledge (related to the concrete application). We are able to distinguish between the rules despite the fact that their meaning is vague.

Let us now consider specific linguistic contexts: \( w \in W \) for values of \( X \), say, \( w = \langle 150, 330, 600 \rangle \) (for example, temperature in some oven) and \( w' \in W \) for values of \( Y \), \( w' = \langle 0, 36, 90 \rangle \) (for example, a turncock position in degrees). Then “small \( X \)” are values of \( X \) around 180–210 (and smaller) and “very big \( X \)” are values around, at least, 550 or higher. Similarly, “small \( Y \)” are values of \( Y \) around 8 (and smaller) and “very big \( Y \)” values around, at least, 80–85 (and higher).

Let the value \( X = 180 \) be given as an observation of \( X \) (this may be, e.g., a result of some measurement). To derive a conclusion on the basis of the given linguistic description, we must first test whether this value falls in its topic. Since the value of 180 is in the context \( w = \langle 150, 330, 600 \rangle \) apparently small, we may apply knowledge determined by \( LD \) and because of rule \( \mathcal{R}_1 \), we expect a very big value of \( Y \) (a value around 85). Similarly, \( X = 560 \) which is very big leads to a value of \( Y \) around 8 due to the rule \( \mathcal{R}_2 \).

More details, formal explanation and description of some applications can be found also in [41].

An interesting formal theory elaborated in the frame of FLb is the theory of intermediate quantifiers. These are expressions such as most, a lot of, many, a few, a great deal of, large part of, small part of which have been in detail informally elaborated in [63]. The main idea is to introduce intermediate quantifiers as special formulas of fuzzy type theory which express that they are, in fact, classical general or existential
quantifiers but limited to smaller (fuzzy) sets (for example, fuzzy sets with smaller support) whose size is measured using some measure. Formulas representing intermediate quantifiers are constructed in a certain extension of the formal theory $T^{Ev}$ of evaluative linguistic expressions. The theory is sufficiently general to encompass a wide class of the generalized quantifiers and provides a unique definition for all of them.

We have proved in [32] validity of all 105 generalizations of classical Aristotelian syllogisms presented in [63]. Their form is, for example, as follows:

\begin{align*}
\text{APK-I:} & \quad \text{All } M \text{ are } Y \\
\text{ATT-I:} & \quad \text{Most } X \text{ are } M \\
\text{Many } X \text{ are } Y & \quad \text{Most } X \text{ are } Y
\end{align*}

There is a great application potential of the theory presented above. One of interesting possibilities is the commonsense human reasoning.

### 2.3 Model of Commonsense Human Reasoning

We will very briefly mention an FLb-model of a detective story inspired by one episode of the famous TV series about Lt. Columbo. Note that detective stories are typical examples of human reasoning where crucial role is played also by natural language, knowledge of common things and situations, and standard logical deduction. FLb can provide a formal logical mechanism that can mimic Lt. Columbo’s reasoning.

**The plot:** Mr. John Smith has been shot dead in his house. He was found by his friend, Mr. Robert Brown. Lt. Columbo suspects Mr. Brown to be the murderer.

**Mr. Brown’s testimony:** I have started from my home at about 6:30, arrived to John’s house at about 7, found John dead and went immediately to the phone box to call police. They came immediately.

**Lt. Columbo’s evidence:** Mr. Smith had high quality suit, broken wristwatch stopped at 5:45. No evidence of strong strike on his body. Lt. Columbo touched engine of Mr. Brown’s car and found it to be more or less cold.

Lt. Columbo concluded that Mr. Brown lies because:

(i) Mr. Brown’s car engine is *more or less cold*, however he went *long* (more than about 30 minutes). He could not arrive and continue to call the police (which came *immediately*).

(ii) *High quality wristwatch* does not break after *not too strong strike*. A man having *high quality dress* and a *luxurious house* is supposed to have also *high quality wristwatch*. Wristwatch of John Smith is of *low quality* and so, it does not belong to him. It does not display time of Mr. Smith’s death.

**The method:** combination of logical rules, world knowledge and evidence with the help of *non-monotonic reasoning* realized on syntactic level in FTT.
Mathematical Fuzzy Logic: A Good Theory for Practice

• **Context**
  
  (a) *Drive duration to heat the engine (minutes):* \( w_D = (0, 5, 30) \)
  
  (b) *Temperature of engine (degrees Celsius):* \( w_T = (0, 45, 100) \)
  
  (c) *Abstract degrees: quality, state, strike strength:* \( (0, 0.5, 1) \)

• **Logical rules** – logical theorems of FTT and theorems given by some considered theory, for example

  \[
  \begin{align*}
  &\text{IF } X \text{ is } Sm \text{ THEN } X \text{ is not } Bi \\
  &\text{IF } X \text{ is } Bi \text{ THEN } X \text{ is not } Sm
  \end{align*}
  \]

**World knowledge**

• **Knowledge from physics**

  \[
  \begin{align*}
  &\text{IF } \text{drive duration is } Bi \text{ THEN engine temperature is } Bi \\
  &\text{IF } \text{drive duration is } Sm \text{ THEN engine temperature is } ML Sm
  \end{align*}
  \]

• **Customs of people**

  \[
  \begin{align*}
  &\text{IF quality of } x_\pi \text{’s suit is } Bi \text{ AND quality of } x_\pi \text{’s house is } Ve Bi \\
  &\text{THEN wealth of } x_\pi \text{ is } Bi
  \end{align*}
  \]

• etc.

  *(the following shorts for evaluative expressions have been used: *Sm*-small, *Bi*-big, *Ve*-very, *ML*-more or less)*

  We construct a specific model \( M \) given by the evidence of the contexts of wealthy people and perceptions of Lt. Columbo, for example:

  • Touching Mr. Brown’s car engine by hand does not burn; its temperature is *ML Sm*,
  
  • Quality of Mr. Smith’s house is *Ve Bi*.

**Lt. Columbo’s conclusion:** *The two special constructed theories are contradictory. Since my evidence is correct, Mr. Brown lies and had an opportunity to kill Mr. Smith.*

The vagueness is manifested in this reasoning by the fact that a slight change in the evidence may lead to disappearance of the contradiction, despite the fact that the basic knowledge has not changed. For example, the truth of “temperature of engine is low” can be higher, or vice versa, the truth of “temperature of engine is more or less big” can be nonzero, etc. so that the conclusion can be made more, or less, convincing. Consequently, further evidence might be necessary, or vice-versa. We conclude that vagueness plays an important role in reasoning and drawing conclusions. Classical boolean logic can provide solutions only at the price of imposing,
quite often, improper precision and thus, making the model of reasoning unrealistic. Introduction of truth degrees may help to overcome such restrictions since we can balance them: the higher the truth values, the more convincing the conclusion — but the limit (i.e., full truth 1 or full falsity 0) is unnecessary.

We should see the above example of commonsense reasoning as a general methodology which can have various kinds of applications. For example, in [12] it has also been applied to analysis of economic texts.

3 Related Theories with Great Application Potential

One could conclude from the above discussion that fuzzy logic is indeed a good well established theory which can be very practical. But we can go farther. Besides the linguistic interpretation of the rules (9), they are also interpreted in the frame of predicate fuzzy logic as fuzzy relations (cf. [19,48]).

3.1 Fuzzy Approximation

The linguistic description is assigned certain predicate formulas in BL- or Łukasiewicz fuzzy logic in one of two possible forms ([19,48,51,52,55]):

1. **Disjunctive normal form** (DNF) where each rule is interpreted by a conjunction of the antecedent and consequent and the fuzzy relations of all rules are joined by disjunction. This approach is largely known under the name *Mamdani-Assilian interpretation* [28], although it has its roots in [68,69]. Interpretation of DNF provides the following fuzzy relation

   \[ R_{\text{DNF}}(x,y) = \bigvee_{i=1}^{n} (A_i(x) \ast B_i(y)), x \in X, y \in Y. \]  

2. **Conjunctive normal form** (CNF) where the individual rules are interpreted as implications and joined by conjunction (see also [8]). Interpretation of CNF provides the following fuzzy relation

   \[ R_{\text{CNF}}(x,y) = \bigwedge_{i=1}^{n} (A_i(x) \rightarrow B_i(y)), x \in X, y \in Y. \]

The linguistic description characterizes a specific dependence between values of \( X \) and \( Y \), where the values of \( X \) are the possible inputs and the values of \( Y \) are the possible outputs of the given system described by \( R \). The constructed fuzzy relation \( R_{\text{DNF}} \) or \( R_{\text{CNF}} \) can then be used to compute an output for a given input. This input is either a crisp value \( x \in X \) or a fuzzy set \( A \in \mathcal{F}(X) \). The process how the output is computed using fuzzy relations is usually called the *Computational Rule of Inference* (CRI) first considered in [63]. The core of such an inference mechanism is the computation of an image of the input fuzzy set via a given fuzzy relation. The
result is an output fuzzy set \( B \in \mathcal{F}(Y) \). Logical justification of this process has been provided in [19, 48].

One of the practical problems is how the fuzzy sets \( A_i, B_i \) in (10), (11) should be specified. This problem turned out to be closely related to solvability of fuzzy relation equations which has been initiated by E. Sanchez in [65]. The result says that the fuzzy relations \( R_{DNF} \) or \( R_{CNF} \) and, consequently, CRI rule are incorrectly defined if the corresponding system of fuzzy relation equations is not solvable (cf. [54, 59, 60, 61]). This also implies that evaluative linguistic expressions cannot be, in general, used with CRI (see [35]) and so, the latter is not convenient tool for logical reasoning inside natural language (i.e., to fit the paradigm of FLb).

Relational interpretation of linguistic descriptions has thousands of various kinds of applications. It can be considered as one of most practical theory based on fuzzy predicate logic.

### 3.2 Fuzzy (F)-Transform

The fuzzy transform (F-transform for short) introduced in detail by I. Perilieva in [56] is a special technique that can be applied to a continuous function, defined on a fixed real interval \([a, b] \subset \mathbb{R}\). The essential idea is to transform a given function defined in one space into another, usually simpler space, and then to transform it back. The simpler space consists of a finite vector of numbers obtained on the basis of the well established fuzzy partition of the domain of the given function. The reverse transform then leads to a function approximately reconstructing the original one. Thus, the first step, sometimes called the direct F-transform, results in a vector of averaged functional values. The second step, called the inverse transform, converts this vector into another continuous function, which approximates the original one.

The essential idea of F-transform is defined with respect to a fuzzy partition which is a special system of fuzzy sets defined on a set of nodes in \([a, b]\) which fulfils several conditions. Some examples of fuzzy partitions see Fig. 2.

![Fig. 2 Graphical presentation of several fuzzy partitions.](image)

Let a fuzzy partition of \([a, b]\) be given by \(A_1, \ldots, A_n, n \geq 2\), and let \(f : [a, b] \rightarrow \mathbb{R}\) be an arbitrary continuous function. The \(n\)-tuple of real numbers \([F_1, \ldots, F_n]\) given by

\[
F_i = \frac{\int_a^b f(x)A_i(x)dx}{\int_a^b A_i(x)dx}, \quad i = 1, \ldots, n,
\]

is a direct fuzzy transform (F-transform) of \(f\) with respect to the given fuzzy partition. The numbers \(F_1, \ldots, F_n\) are called the components of the F-transform of \(f\).
In practice, the function $f$ is usually not given analytically, but we are at least provided some data, obtained, for example, by some measurements. In this case, the previous definition can be modified in such a way that the definite integrals in Formula (12) are replaced by finite summations. It has been proven [56] that the components of the F-transform are weighted mean values of the original function, where the weights are determined by the basic functions.

The original function $f$ can be approximately reconstructed from the direct F-transform of $f$ with respect to $A_1, \ldots, A_n \in \mathcal{F}([a, b])$ using the following inversion formula:

$$f_{F,n}(x) = \sum_{i=1}^{n} F_i A_i(x).$$

(13)

The function $f_{F,n}$ is called the inverse F-transform of $f$. It is a continuous function on $[a, b]$.

**Theorem 3.** Let $f$ be a continuous function on $[a, b]$. Then for any $\varepsilon > 0$ there exist $n_{\varepsilon}$ and a fuzzy partition $A_1, \ldots, A_{n_{\varepsilon}}$ of $[a, b]$ such that for all $x \in [a, b]$

$$|f(x) - f_{F,n_{\varepsilon}}(x)| \leq \varepsilon$$

(14)

where $f_{F,n_{\varepsilon}}$ is the inverse F-transform of $f$ with respect to the fuzzy partition $A_1, \ldots, A_{n_{\varepsilon}}$.

It should be noted that this theorem holds true independently on shapes of the fuzzy sets $A_1, \ldots, A_{n_{\varepsilon}}$ (of course, provided that they fulfil the general required conditions). For various properties of the F-transform and detailed proofs — see [56]. There are numerous applications of the F-transform in image processing, data mining, time series analysis, solution of differential equations, some numerical methods and elsewhere (see, e.g. [7, 49, 53, 57, 58, 62, 64]). One can see that this theory is very practical.

4 Conclusion

As can be seen from the previous brief presentation, mathematical fuzzy logic is a well developed formal theory which gave rise to several more special theories. Its structure is now already known quite well. For example, in the paper [4], almost 60 various fuzzy logics are analyzed together. We still do not know exactly which logic is the most convenient to solve problems related to models of vagueness and their applications but one can guess from the discussion above that the leading position is played by MFL based on standard Łukasiewicz MV-algebra of truth values.

We think that there are good reasons to continue the development of FLb and move it further towards human (i.e., commonsense) reasoning, following the methodology of PNL and results obtained in the AI theory of commonsense reasoning. Another promising direction is to include also uncertainty in FLb in the sense that has been nicely established by T. Flaminio and F. Montagna in [15]. In their paper, fuzzy logic is joined with probability theory.
To summarize, we have presented several theories and mentioned their applications. We are convinced that this demonstrates that MFL and the related theories are very practical.

References

Interval-Valued Algebras and Fuzzy Logics

Bart Van Gasse, Chris Cornelis, and Glad Deschrijver

Abstract. In this chapter, we present a propositional calculus for several interval-valued fuzzy logics, i.e., logics having intervals as truth values. More precisely, the truth values are preferably subintervals of the unit interval. The idea behind it is that such an interval can model imprecise information. To compute the truth values of ‘p implies q’ and ‘p and q’, given the truth values of p and q, we use operations from residuated lattices. This truth-functional approach is similar to the methods developed for the well-studied fuzzy logics. Although the interpretation of the intervals as truth values expressing some kind of imprecision is a bit problematic, the purely mathematical study of the properties of interval-valued fuzzy logics and their algebraic semantics can be done without any problem. This study is the focus of this chapter.

1 Introduction

Classical logic is a two-valued logic: propositions in this logic are either true or false. In the first case, the truth value 1 is attributed to the proposition, while in the second case the attributed truth value is 0. Given the truth values of two propositions
and $q$, it is possible to derive the truth values of the negation ‘not $p$’ (and ‘not $q$’),
the conjunction ‘$p$ and $q$’, the disjunction ‘$p$ or $q$’ and the implication ‘$p$ implies $q$’. These formulas are denoted as $\neg p$, $p \& q$, $p \lor q$ and $p \rightarrow q$. The truth values are calculated using the operations $\neg$, $\ast$, $\sqcup$ and $\Rightarrow$. The truth tables of these operations are given in Table 1.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>$\neg x$</th>
<th>$x \ast y$</th>
<th>$x \sqcup y$</th>
<th>$x \Rightarrow y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1 Truth tables of the operations in classical logic

For more complicated formulas the truth values can be computed in the same way. For example, if $p$ is true and $q$ is false, then the truth value of $(p \lor q) \rightarrow ((p \rightarrow q) \rightarrow q)$ is calculated as follows: the truth value of $p \rightarrow q$ is $1 \Rightarrow 0 = 0$, so the truth value of $(p \rightarrow q) \rightarrow q$ is $0 \Rightarrow 0 = 1$. The truth value of $p \lor q$ is $1 \sqcup 0 = 1$. So we conclude that the truth value of $(p \lor q) \rightarrow ((p \rightarrow q) \rightarrow q)$ is $1 \Rightarrow 1 = 1$. Interestingly, the truth value of this formula is always 1, even if other truth values are attributed to $p$ and $q$. Such formulas are called tautologies. If a formula $\phi$ is a tautology, this is denoted as $\models \phi$. More generally, for a set of formulas $\Gamma$, $\Gamma \models \phi$ means “no matter what truth values are attributed to the propositions, if the truth values of the formulas in $\Gamma$ are 1, then the truth value of $\phi$ is 1”.

The two values 0 and 1, together with the defined operations, form a Boolean algebra. Therefore we say that this Boolean algebra is the semantics of classical logic. Saying that $(p \lor q) \rightarrow ((p \rightarrow q) \rightarrow q)$ is a tautology in classical logic, is the same as saying that $(x \sqcup y) \Rightarrow ((x \Rightarrow y) \Rightarrow y) = 1$ is an identity in this Boolean algebra (meaning “whatever value of the Boolean algebra we give to $x$ and $y$, the calculation of $(x \sqcup y) \Rightarrow ((x \Rightarrow y) \Rightarrow y)$ yields 1”). Now, identities in this Boolean algebra are also identities in every other Boolean algebra (we say that this Boolean algebra generates all Boolean algebras). Therefore classical logic does not only have the Boolean algebra with two elements as semantics, but also the whole variety of Boolean algebras: the general semantics of classical logic consist of all Boolean algebras.

Interestingly it is also possible to describe classical logic without using semantics. This is done with axioms and deduction rules, which allow to prove a formula from a set of formulas. When a formula $\phi$ is provable from a theory $\Gamma$, this is denoted as $\Gamma \vdash \phi$. Two important results in classical logic are soundness (if $\Gamma \vdash \phi$, then $\Gamma \models \phi$) and completeness (if $\Gamma \models \phi$, then $\Gamma \vdash \phi$). We write this shortly as $\Gamma \vdash \phi$ iff $\Gamma \models \phi$.

Note that we use different symbols, to distinguish the logical connectives from the corresponding operations. Only for the negation we employ the same symbol.
Now, for the truth values of several propositions one might prefer more than the two options 0 (false) and 1 (true). Indeed, for vague propositions like ‘it is raining hard’, it would be useful if one could attribute an intermediate truth value, somewhere between ‘false’ and ‘true’. This can be done using fuzzy set theory, in which every element of the unit interval $[0,1]$ serves as a truth value, instead of only 0 and 1. The operations for the negation, conjunction, disjunction and implication were generalized to this setting. Later the structure of the unit interval was generalized to an arbitrary bounded lattice to allow for incomparabilities among elements, and triangular norms and conorms are quite common nowadays as generalized representations of logical conjunction and disjunction, respectively. An interesting class of these generalizations, especially from the logical point of view, are MTL-algebras [16]. In these structures, the operations modelling (strong) conjunction and implication are connected by the residuation principle. These MTL-algebras form the general semantics of monoidal t-norm based logic (MTL, [16]), in the same way Boolean algebras form the general semantics of classical logic. Similarly as for classical logic, these general semantics can be restricted. Indeed, MTL is also sound and complete w.r.t. standard MTL-algebras, i.e., MTL-algebras on the unit interval. Therefore MTL is called a (formal) fuzzy logic. But it is definitely not the only fuzzy logic. Indeed, by adding more axioms and/or deduction rules to the axioms and deduction rule of MTL, we obtain other fuzzy logics. It is even possible to retrieve classical logic in this way. Semantically speaking, this means that Boolean algebras are special cases of MTL-algebras. Some well-known fuzzy logics, situated between MTL and classical logic, are Hájek’s Basic Logic (BL) [24], Łukasiewicz Logic (Ł) [30] and Gödel Logic (GL) [14, 22]. Also Intuitionistic Logic (IL) [25] can be seen as a fuzzy logic. These logics are sound and complete w.r.t. BL-algebras, MV-algebras (or, equivalently, Wajsberg algebras [20]), G-algebras and Heyting algebras, respectively. We refer to [17] for a comprehensive overview of these and other logics. Other general references on fuzzy logics are [6, 23, 24].

In [43], Zadeh introduced type-2 fuzzy sets, a generalization of fuzzy sets. The idea behind these structures is that they provide a way to express incomplete as well as graded knowledge; as opposed to fuzzy sets, which only express gradedness, not incompleteness. Unfortunately, type-2 fuzzy sets are quite complicated to work with. Therefore often interval-valued fuzzy sets are used. These special cases of type-2 fuzzy sets are easier to handle. Indeed, truth values in this setting are closed subintervals of the unit interval, and such an interval is determined by just two values: its lower and upper bound. The aim of this chapter is to develop a logic that has intervals as truth values. The intended semantics are residuated lattices on the set of closed subintervals of the unit interval. We call this set the triangularization of the unit interval. A particular subset of this triangularization is its so-called diagonal, consisting of those intervals for which the lower and upper bound coincide. These intervals are called exact intervals and represent truth values of propositions about which the knowledge is complete. Intuitively, the truth values of formulas constructed with these propositions should be exact intervals as well (because in these cases, the situation is similar to working with formulas in fuzzy logics). The semantics of so-called interval-valued fuzzy logics have already been
examined by different authors. Especially interval-valued triangular norms, triangular conorms and implicators have received ample attention. Most of these authors [1, 2, 15, 21, 26] only consider interval-valued operations that map the diagonal on the diagonal, although the most general definitions of triangular norms, triangular conorms and implicators allow other operations as well [7, 10, 11, 28]. Generally speaking, interval-valued operations do not satisfy as many properties as operations on the unit interval. For example, standard interval-valued residuated lattices can never satisfy prelinearity [8]. A lot of other properties can hold though. There are even interval-valued implicators that satisfy all the Smets-Magrez axioms [9].

The three main sections of this chapter are conceived as follows:

• In Section 2 we elaborate the theory of interval-valued residuated lattices, which include the intended semantics of interval-valued fuzzy logics.
• In Section 3 we give the definition of triangle algebras, which are algebraic structures describing interval-valued residuated lattices.
• In Section 4 we then introduce several interval-valued fuzzy logics and examine their properties, in particular the soundness and completeness w.r.t. the intended and the general semantics.

Before we continue, we recall some algebraic concepts that will be used in this chapter.

• An algebra of type $(n_1, n_2, \ldots, n_m)$, with $n_1, n_2, \ldots, n_m$ non-negative integers, is a structure $(A, f_1, \ldots, f_m)$ in which $A$ is a set, $f_1$ an $n_1$-ary operation on $A$, . . . and $f_m$ an $n_m$-ary operation on $A$. If $n_i$ is 0, then $f_i$ is a constant.
• A reduct of an algebra is an algebra on the same set, but in which some of the operations are left out. An algebra $A$ is an expansion of an algebra $B$ if $B$ is a reduct of $A$.
• A subalgebra of an algebra $A = (A, f_1, \ldots, f_m)$ is an algebra on a subset $A'$ of $A$ in which all operations of $A$ are restricted to $A'$. Of course, this is only possible if $A'$ is closed under all these operations, i.e., if for every operation $f_i$ of $A$, $f_i(a_1, \ldots, a_n) \in A'$ whenever the arguments $a_1, \ldots, a_n$ are in $A'$ (with $n$ the arity of $f_i$).
• A morphism from an algebra $A = (A, f_1, \ldots, f_m)$ to an algebra $B = (B, g_1, \ldots, g_m)$ of the same type, is a mapping $h$ from $A$ to $B$ such that $h(f_i(a_1, \ldots, a_n)) = g_i(h(a_1), \ldots, h(a_n))$ for all operations $f_i$ of $A$ and all $a_1, \ldots, a_n$ in $A$ (with $n$ the arity of $f_i$).
• An embedding of an algebra $A = (A, f_1, \ldots, f_m)$ in an algebra $B = (B, g_1, \ldots, g_m)$ of the same type, is a morphism $h$ from $A$ to $B$ such that $h(a_1) \neq h(a_2)$ whenever $a_1 \neq a_2$.
• An isomorphism from an algebra $A = (A, f_1, \ldots, f_m)$ to an algebra $B = (B, g_1, \ldots, g_m)$ of the same type, is an embedding of $A$ in $B$ such that for every element $b$ of $B$, $b = h(a)$ for some $a$ in $A$. 
2 Interval-Valued Structures

The most general semantics of fuzzy logics do not only contain algebraic structures on the unit interval, it consists of all residuated lattices. For interval-valued fuzzy logics, the situation is comparable: the most general semantics are interval-valued residuated lattices. In this section, we propose a definition of these structures.

2.1 Triangularizations of Partially Ordered Sets

**Definition 1.** [39] Given any partially ordered set (shortly: poset) $\mathcal{P} = (P, \leq)$, we can define its triangularization $\mathbb{T}(\mathcal{P}) = (\text{Int}(\mathcal{P}), \preceq)$ in the following way:

- $\text{Int}(\mathcal{P}) = \{ [p_1, p_2] \mid (p_1, p_2) \in P^2 \text{ and } p_1 \leq p_2 \}$,
- $[p_1, p_2] \preceq [q_1, q_2]$ iff $p_1 \leq q_1$ and $p_2 \leq q_2$, for all $[p_1, p_2]$ and $[q_1, q_2]$ in $\text{Int}(\mathcal{P})$.

The elements of $\text{Int}(\mathcal{P})$ are called the intervals of $\mathcal{P}$. The first and the second projection $\text{pr}_1$ and $\text{pr}_2$ are the mappings from $\text{Int}(\mathcal{P})$ to $P$, defined by $\text{pr}_1([x_1, x_2]) = x_1$ and $\text{pr}_2([x_1, x_2]) = x_2$, for all $[x_1, x_2]$ in $\text{Int}(\mathcal{P})$.

The vertical and the horizontal projection $\text{pr}_v$ and $\text{pr}_h$ are the mappings from $\text{Int}(\mathcal{P})$ to $\text{Int}(\mathcal{P})$, defined by $\text{pr}_v([x_1, x_2]) = [x_1, x_1]$ and $\text{pr}_h([x_1, x_2]) = [x_2, x_2]$, for all $[x_1, x_2]$ in $\text{Int}(\mathcal{P})$.

It is straightforward to verify that for any poset $\mathcal{P}$, $\mathbb{T}(\mathcal{P})$ is also a poset. Moreover, the original poset $(P, \leq)$ is contained in $\mathbb{T}(\mathcal{P})$ in some way: indeed, the mapping $i: P \to \text{Int}(\mathcal{P})$ defined by $i(p) = [p, p]$ for all $p$ in $P$, is injective and preserves the ordering (if $p \leq q$, then $i(p) = [p, p] \preceq [q, q] = i(q)$). The image $i(P)$ consists of the intervals $[p_1, p_2]$ in $\text{Int}(\mathcal{P})$ for which $p_1 = p_2$. The elements of $i(P)$ are called exact intervals. The subset $i(P)$ of $\text{Int}(\mathcal{P})$ is often referred to as the diagonal of $\mathbb{T}(\mathcal{P})$. Note that $\text{pr}_v = i \circ \text{pr}_1$ and $\text{pr}_h = i \circ \text{pr}_2$, and that $i(P) = \text{pr}_v(\text{Int}(\mathcal{P})) = \text{pr}_h(\text{Int}(\mathcal{P}))$.

**Example 1.** The poset that will be of central interest in this chapter is $\mathbb{T}([0, 1], \leq)$: the closed subintervals of the unit interval. This poset is complete and its order is not linear. Its graphical representation as a triangle is shown in Figure 1. The diagonal is the hypothenuse of this triangle. Note that the shape of this representation is triangular. This holds for all triangularizations of bounded linear posets, hence the name ‘triangularization’.

2.2 Triangular Lattices

Recall that a lattice is a poset in which the supremum and infimum of every two elements exist; on the other hand, often the following equivalent definition is also used.

**Definition 2.** A lattice is an algebra $(L, \sqcap, \sqcup)$ of type $(2, 2)$ such that $\sqcap$ (‘meet’) and $\sqcup$ (‘join’) are idempotent, commutative and associative operations satisfying the following absorption laws: for all $x$ and $y$ in $L$, $x \sqcup (x \sqcap y) = x$ and $x \sqcap (x \sqcup y) = x$.

The lattice order $\leq$ is defined by $x \leq y$ iff $x \sqcap y = x$ (or, equivalently, iff $x \sqcup y = y$), for all $x$ and $y$ in $L$.
Because lattices can be seen as posets, we can consider their triangularizations. One easily observes that the infimum (resp. supremum) on the triangularization of a lattice is obtained by taking the infimum (resp. supremum) of the first and second projections. More precisely: for any lattice \( \mathcal{L} = (L, \sqcap, \sqcup) \), the infimum \( \sqcap \) and supremum \( \sqcup \) on its triangularization \( \mathbb{T}(\mathcal{L}) \) are given by

\[
\begin{align*}
[x_1, x_2] \sqcap [y_1, y_2] &= [x_1 \sqcap y_1, x_2 \sqcap y_2], \\
[x_1, x_2] \sqcup [y_1, y_2] &= [x_1 \sqcup y_1, x_2 \sqcup y_2],
\end{align*}
\]

for all \([x_1, x_2]\) and \([y_1, y_2]\) in \( \text{Int}(\mathcal{L}) \). Note that we use big \( \sqcap \)- and \( \sqcup \)-symbols for the intervals of the triangularization, and small \( \sqcap \)- and \( \sqcup \)-symbols for the elements of the original lattice.

It can be verified that \( \mathbb{T}(\mathcal{L}) \) is a lattice iff \( \mathcal{L} \) is a lattice, that \( i \) (as defined in Section 2.1) is a morphism from \( (L, \sqcap, \sqcup) \) to \( (\text{Int}(\mathcal{L}), \sqcap, \sqcup) \), and that the set \( i(L) \) of exact intervals is therefore closed under \( \sqcap \) and \( \sqcup \) and forms a sublattice \( \mathcal{E}(\mathcal{L}) = (i(L), \sqcap, \sqcup) \).

On the other hand, \( \mathbb{T}(\mathcal{L}) \) is bounded iff \( \mathcal{L} \) is bounded. In this case, the smallest (resp. greatest) element of \( \mathcal{L} \) is usually denoted by 0 (resp. 1), and the smallest (resp. greatest) element of \( \mathbb{T}(\mathcal{L}) \) by \( [0, 0] \) (resp. \( [1, 1] \)). As we will see later on, the element \([0, 1]\) will also play an important role, along with the projections \( \text{pr}_v \) and \( \text{pr}_h \).

For any triangularization \( (\text{Int}(\mathcal{L}), \sqcap, \sqcup) \) of a bounded lattice \( \mathcal{L} = (L, \sqcap, \sqcup) \) (with smallest element 0 and greatest element 1), we call \( (\text{Int}(\mathcal{L}), \sqcap, \sqcup, \text{pr}_v, \text{pr}_h, [0, 0], [0, 1], [1, 1]) \) the extended triangularization of \( \mathcal{L} \). Below, we show that extended triangularizations can be captured by a class of algebraic structures defined only with identities: the variety of triangular lattices.

**Definition 3.** [36] A triangular lattice\(^2\) is an algebra \( (L, \sqcap, \sqcup, v, \mu, 0, u, 1) \) of type \( (2, 2, 1, 1, 0, 0, 0) \) such that \( (L, \sqcap, \sqcup) \) is a bounded lattice with smallest element 0 and greatest element 1 such that

\(^2\) The reason we call the last two conditions (T.10) and (T.10’) instead of (T.7) and (T.7’) is that we would like to keep the same notations as in the papers [36, 37, 38, 40, 42].
In the remainder, we will use the set of exact elements lattice. Conversely, every extended triangularization of a bounded lattice is a triangular use in this work – is that of residuated lattices.

For most formal fuzzy logics the semantics require not only a partial order on the set of truth values, but also some extra operations that model ‘AND’ (the strong conjunction - the infimum being the weak conjunction) and ‘IMPLIES’ (the implication). A very commonly used structure – and also the basic structure that we will use in this work – is that of residuated lattices.
**Definition 4.** A residuated lattice\(^3\) is a structure \(\mathcal{L} = (L, \sqcap, \sqcup, *, \Rightarrow, 0, 1)\) in which

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

The binary operations \(*\) and \(\Rightarrow\) are called product and implication, respectively. We will use the notations \(\neg x\) for \(x \Rightarrow 0\) (negation), \(x \Leftrightarrow y\) for \((x \Rightarrow y) \sqcap (y \Rightarrow x)\).

The following kinds of residuated lattices are used in this chapter.

**Definition 5**

- An MTL-algebra\(^4\) is a prelinear residuated lattice, i.e., a residuated lattice in which \((x \Rightarrow y) \sqcup (y \Rightarrow x) = 1\) for all \(x, y\) in \(L\).
- A BL-algebra\(^5\) is a divisible MTL-algebra, i.e., an MTL-algebra in which \(x \sqcap y = x \ast (x \Rightarrow y)\) for all \(x, y\) in \(L\). The weaker property \(x \sqcap y = (x \ast (x \Rightarrow y)) \sqcup (y \ast (y \Rightarrow x))\) is called weak divisibility.
- An MV-algebra\(^6\) is a BL-algebra in which the negation is an involution, i.e., \((x \Rightarrow 0) \Rightarrow 0 = x\) for all \(x\) in \(L\).
- A Boolean algebra\(^7\) is an MV-algebra that is also a Heyting-algebra, i.e., in which \(x \ast x = x\) for all \(x\) in \(L\), or, equivalently, in which \(* = \sqcap\).

In a residuated lattice, the operator \(*\) is always a residuated t-norm, with \(\Rightarrow\) as its residual implicator\(^8\). Conversely, if \(T\) is a residuated t-norm on a bounded lattice \((L, \sqcap, \sqcup)\), then \((L, \sqcap, \sqcup, T, I_T, 0, 1)\) is a residuated lattice. Note however that not all t-norms are residuated. In complete lattices \((L, \sqcap, \sqcup)\), a t-norm \(T\) is residuated (and therefore induces a residuated lattice) iff it satisfies \(T(x, \text{sup}Y) = \text{sup}\{T(x, y) \mid y \in Y\}\) for all \(x\) in \(L\) and \(Y \subseteq L\)\(^9\).

**Example 2.** Let \(T\) be a t-norm on \([0, 1], \text{min}, \text{max}\). It is well-known (see, e.g.,\(^{17, 24}\)) that

- \(T\) is residuated iff \(T\) is left-continuous,
- \([0, 1], \text{min}, \text{max}, T, I_T, 0, 1\) is an MTL-algebra\(^8\) iff \(T\) is left-continuous,
- \([0, 1], \text{min}, \text{max}, T, I_T, 0, 1\) is a BL-algebra iff \(T\) is continuous.

---

3 In the literature (e.g. in\(^{26}\)), the name residuated lattice is sometimes used for structures more general than what we call residuated lattices. In the most general terminology, our structures would be called bounded integral commutative residuated lattices.

4 Recall that a triangular norm (t-norm, for short) on a poset \((P, \leq)\) with largest element 1, is a binary, increasing, commutative and associative operator \(T: P^2 \rightarrow P\) that satisfies \(T(x, 1) = 1\), for all \(x\) in \(P\). If for every pair \((x, y)\) in \(P^2\), \(\sup\{z \in P \mid T(x, z) \leq y\}\) exists, then the map \(I_T\) defined by \(I_T(x, y) = \sup\{z \in P \mid T(x, z) \leq y\}\) is called the residual implicator of \(T\). A t-norm \(T\) is called residuated if it has a residual implicator satisfying \(I_T(x, y) = \max\{z \in P \mid T(x, z) \leq y\}\), in other words if for any pair \((x, y)\) in \(P^2\) the set \(\{z \in P \mid T(x, z) \leq y\}\) has a maximum.

5 Because \([0, 1], \text{min}, \text{max}\) is linear, every residuated lattice on this lattice is automatically an MTL-algebra.
• $([0,1], \min, \max, T, I_T, 0, 1)$ is an MV-algebra iff $T$ is conjugated to the Łukasiewicz t-norm $T_W$, i.e., iff there exists a strictly increasing bijection $\phi : [0,1] \rightarrow [0,1]$ such that $T(x,y) = \phi^{-1}(T_W(\phi(x),\phi(y)))$, where $T_W(x,y) = \max(0,x+y-1)$.

Several t-norms can be defined on triangularizations of bounded lattices. A class that will be of great importance later on, consists of t-norms constructed from a t-norm on and an element of the original lattice: if $T$ is a t-norm on a lattice $\mathcal{L} = (L, \sqcap, \sqcup)$ which has a greatest element 1, and $t$ is an element of $L$, then $\mathcal{T}_{T,t}$, defined by

$$\mathcal{T}_{T,t}([x_1,x_2],[y_1,y_2]) = [T(x_1,y_1),T(T(x_2,y_2),t) \sqcup T(x_1,y_2) \sqcup T(x_2,y_1)],$$

(2)

for all $[x_1,x_2]$ and $[y_1,y_2]$ in $Int(\mathcal{L})$, is a t-norm on $\mathcal{T}(\mathcal{L})$. These t-norms were introduced by Deschrijver and Kerre in [11].

Recall that we are working towards a variety of algebraic structures suitable as semantics for a logic with intervals as truth values. At this point it might seem a good idea to choose residuated lattices on triangularizations (or, equivalently, residuated lattices on triangular lattices). However, in these structures the set of exact intervals is not necessarily closed under the product and implication. This counters the intuition that the truth values of the propositions $p \& q$ and $p \rightarrow q$ should be exact if the truth values of $p$ and $q$ are exact. Therefore, residuated lattices on triangularizations are too general to serve as the desired semantics. This leads us to the definition of interval-valued residuated lattices.

**Definition 6.** [41]

• An interval-valued residuated lattice (IVRL) is a residuated lattice $(Int(\mathcal{L}), \sqcap, \sqcup, *, \Rightarrow, 0, 1)$ on the triangularization $\mathcal{T}(\mathcal{L})$ of a bounded lattice $\mathcal{L} = (L, \sqcap, \sqcup)$, in which the diagonal $i(L)$ is closed under $*$ and $\Rightarrow$, i.e., $[x_1,x_1] * [y_1,y_1] \in i(L)$ (with $i$ the injection defined in Section 2.1) and $[x_1,x_1] \Rightarrow [y_1,y_1] \in i(L)$ for all $x_1$ and $y_1$ in $L$.

• When we add $[0,1]$ as a constant, and $pr_v$ and $pr_h$ (as defined in Section 2.1) as unary operators, the structure $(Int(\mathcal{L}), \sqcap, \sqcup, *, \Rightarrow, pr_v, pr_h, 0, 0, [0,1], [1,1])$ is called an extended IVRL.

An IVRL in which $\mathcal{L} = ([0,1], \min, \max)$ is called a standard IVRL. An extended IVRL in which $\mathcal{L} = ([0,1], \min, \max)$ is called a standard extended IVRL.

By Proposition 1 if $(Int(\mathcal{L}), \sqcap, \sqcup, *, \Rightarrow, pr_v, pr_h, 0, 0, [0,1], [1,1])$ is an extended IVRL, then it is also a triangular lattice. Now we show that the extra operations $*$ and $\Rightarrow$ satisfy the following two properties, for all $x$ and $y$ in $Int(\mathcal{L})$:

• $pr_v(x) * pr_v(y) \leq pr_v(pr_v(x) * pr_v(y))$,

• $pr_v(x) \Rightarrow pr_v(y) \leq pr_v(pr_v(x) \Rightarrow pr_v(y))$.

Indeed, the first property is equivalent to $pr_v(x) * pr_v(y) = pr_v(pr_v(x) * pr_v(y))$, which means that for any $a$ and $b$ in $i(L)$, $a * b = pr_v(a * b)$, in other words $a * b \in i(L)$. So
it tells us exactly that the diagonal \( i(L) \) is closed under \(*\). And similarly, the second property means the diagonal \( i(L) \) is closed under \( \Rightarrow \).

These two properties suggest a way to describe these IVRLs, which seem suitable as semantics for interval-valued fuzzy logic, using only identities. This leads us to the next section, where we will introduce this variety (called triangle algebras) and study its properties in detail.

### 3 Triangle Algebras

In this section, we introduce a variety of structures called triangle algebras, and show that they are isomorphic to extended IVRLs. Then, we investigate the product and implication of triangle algebras and show that they are determined by their action on the exact elements and by one specific product: \( u \ast u \). This characterization is used to uncover the connections between properties on triangle algebras and properties on their subalgebras of exact elements.

#### 3.1 Definition and Connection with IVRLs

In the definition of a triangle algebra we want to combine the structure of a residuated lattice and the structure of intervals (equipped with the order in Definition 1), plus the desired property that the subset of exact intervals is closed under all defined operations. This leads us to the following definition.

**Definition 7.** [41][36] A triangle algebra is a structure \((A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)\) of type \((2, 2, 2, 1, 1, 0, 0, 0)\) such that \((A, \sqcap, \sqcup, \nu, \mu, 0, u, 1)\) is a triangular lattice, \((A, \sqcap, \sqcup, *, \Rightarrow, 0, 1)\) is a residuated lattice, and satisfying for all \(x\) and \(y\) in \(A\),

\[
\begin{align*}
(T.7') & \quad \nu x \ast \nu y \leq \nu (\nu x \ast \nu y), \\
(T.9) & \quad \nu x \Rightarrow \nu y \leq \nu (\nu x \Rightarrow \nu y).
\end{align*}
\]

In other words, a structure \((A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)\) of type \((2, 2, 2, 1, 1, 0, 0, 0)\) such that \((A, \sqcap, \sqcup, *, \Rightarrow, 0, 1)\) is a residuated lattice, and satisfying for all \(x\) and \(y\) in \(A\),

\[
\begin{align*}
(T.1) & \quad \nu x \leq x, \\
(T.2) & \quad \nu x \leq \nu \nu x, \\
(T.3) & \quad \nu (x \sqcap y) = \nu x \sqcap \nu y, \\
(T.4) & \quad \nu (x \sqcup y) = \nu x \sqcup \nu y, \\
(T.5) & \quad \nu u = 0, \\
(T.6) & \quad \nu \mu x = \mu x, \\
(T.7') & \quad \nu x \ast \nu y \leq \nu (\nu x \ast \nu y), \\
(T.9) & \quad \nu x \Rightarrow \nu y \leq \nu (\nu x \Rightarrow \nu y), \\
(T.10) & \quad x = \nu x \sqcup (\mu x \sqcap u), \\
(T.10') & \quad x = \mu x \sqcap (\nu x \sqcup u).
\end{align*}
\]
A triangle algebra \((A, \cap, \sqcup, *, \Rightarrow, \nu, \mu, 0_A, u_A, 1_A)\) is called a standard triangle algebra iff \((A, \cap, \sqcup) = T([0, 1], \min, \max)\).

In a standard triangle algebra \((A, \cap, \sqcup, *, \Rightarrow, \nu, \mu, 0_A, u_A, 1_A)\) it holds that \(0_A = [0, 0], 1_A = [1, 1], u = [0, 1], \nu [x_1, x_2] = [x_1, x_1] \text{ and } \mu [x_1, x_2] = [x_2, x_2] \text{ for all } [x_1, x_2] \text{ in } \text{Int}([0, 1], \min, \max)\) \[41\]. Because a triangle algebra is an expansion of both a triangular lattice and a residuated lattice, the properties of these kinds of structures remain valid in triangle algebras. The connections between triangle algebras and several related algebraic structures from the literature are studied extensively in \[39, 41, 36\]. In this chapter, we focus on the relationship with IVRLs.

**Proposition 2.** \[41\] Let \((\text{Int}(\mathcal{L}), \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, [0, 0], [0, 1], [1, 1])\) be an extended IVRL. Then \((\text{Int}(\mathcal{L}), \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, [0, 0], [0, 1], [1, 1])\) is a triangle algebra. Conversely, let \(\mathcal{A} = (A, \cap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)\) be a triangle algebra. Then \(\mathcal{A}\) is isomorphic to an extended IVRL.

The isomorphism \(\chi\) that is used in the proof of Proposition 2 is depicted graphically in Figure 2.

![Figure 2](image.png)

Fig. 2 The isomorphism \(\chi\) from a triangle algebra to an extended IVRL

### 3.2 Characterization of Product and Implication. Decomposition Theorem.

The following important proposition reveals that the implication \(\Rightarrow\) and the product \(*\) are completely determined by their action on the diagonal and the value of \(u * u\):
Proposition 3. \([\[40\]]\) In a triangle algebra \(\mathcal{A} = (A, \sqcap, \sqcup, \ast, \Rightarrow, \nu, \mu, 0, u, 1)\), it holds that

- \(\nu(x \Rightarrow y) = (\nu x \Rightarrow \nu y) \sqcap (\mu x \Rightarrow \mu y)\),
- \(\mu(x \Rightarrow y) = (\mu x \Rightarrow (\mu(u \ast u) \Rightarrow \mu y)) \sqcap (\nu x \Rightarrow \mu y)\),
- \(\nu(x \ast y) = \nu x \ast \nu y\),
- \(\mu(x \ast y) = (\nu x \ast \mu y) \sqcup (\mu x \ast \nu y) \sqcup (\mu x \ast \mu y \ast \mu(u \ast u))\)

and therefore (by \((T.10)\) and \((T.10')\))

\[
x \Rightarrow y = \left( (\mu x \Rightarrow (\mu(u \ast u) \Rightarrow \mu y)) \sqcap (\nu x \Rightarrow \nu y) \right) \sqcup \left( (\mu x \Rightarrow \mu y) \sqcap (\nu x \Rightarrow \nu y) \right)
\]

and

\[
x \ast y = \left( (\nu x \ast \mu y) \sqcup (\mu x \ast \nu y) \sqcup (\mu x \ast \mu y \ast \mu(u \ast u)) \right) \sqcap (\nu x \ast \nu y) \sqcup \left( (\nu x \ast \nu y) \sqcup (\mu x \ast \mu y \ast \mu(u \ast u)) \right).
\]

Because of Proposition\([5]\) the product and implication in triangle algebras are always of a specific form, which implies that a triangle algebra is completely determined by its subalgebra of exact elements and the value \(u \ast u\) (in triangle algebras, \(u\) is a constant, playing the role of the interval \([0, 1]\) in IVRLs). Conversely, for a fixed residuated lattice \(\mathcal{L}\) and element \(\alpha\) in that lattice, we can construct a triangle algebra with \(\mathcal{L}\) as subalgebra of exact elements and \(u \ast u\) determined by \(\mu(u \ast u) = \alpha\). So we can conclude that there is a one-to-one correspondence between triangle algebras and couples \((\mathcal{L}, \alpha)\), in which \(\alpha\) is an element of the residuated lattice \(\mathcal{L}\). This characterization implies that every property that can be imposed on triangle algebras, can be formulated in terms of such couples.

In particular, any property defined in Definition\([5]\) can be weakened, by imposing it on \(E(\mathcal{A})\) (instead of \(\mathcal{A}\)) only. We will denote this with the prefix ‘pseudo’. For example, a triangle algebra is said to be pseudo-linear if its set of exact elements is linearly ordered (by the original ordering, restricted to the diagonal). Another example: a triangle algebra is pseudo-divisible if \(\nu x \sqcap \nu y = \nu x \ast (\nu x \Rightarrow \nu y)\) for all \(x\) and \(y\) in \(A\) (\(E(\mathcal{A})\) consists exactly of the elements of the form \(\nu x\)).

It is well-known that MTL-algebras are isomorphic to subdirect products of linear residuated lattices \([16, 24]\). This is a very useful result, as it implies that identities valid in all linear residuated lattices are also valid in all MTL-algebras, which significantly simplifies several proofs, and which is also needed for the chain-completeness of the corresponding logic MTL (see Section\([4]\)). The ‘interval-valued counterpart’ of this result is given below:
Theorem 1. Every pseudo-prelinear triangle algebra $\mathcal{A}$ is isomorphic to a subalgebra of the direct product of a system of pseudo-linear triangle algebras.

4 Interval-Valued Fuzzy Logics

In Section 2 we have given the definition of interval-valued residuated lattices (IVRLs), and in Section 3 we have introduced triangle algebras to capture their structure by means of identities and/or inequalities. Using this characterization, we give the definition of several (propositional) interval-valued fuzzy logics in Section 4.2. But first, in Section 4.1 we give an overview of the well-studied common fuzzy logics, on which our interval-valued fuzzy logics are based, and mention their most important properties. In Section 4.3 we then investigate which of these properties hold for interval-valued fuzzy logics as well. In particular, we prove the soundness and completeness with respect to the algebraic semantics and the deduction theorem.

4.1 Formal Fuzzy Logics

Because in interval-valued fuzzy logics there will be more formulae, we make a distinction between formulae for fuzzy logics (FL-formulae) and formulae for interval-valued fuzzy logics (IVFL-formulae).

Definition 8. FL-formulae are built up from a countable set of propositional variables (denoted by $p$, $q$, $r$, $p_1$, $p_2$, ...) and the constant 0. These symbols are FL-formulae by definition. The other FL-formulae are defined recursively: if $\phi$ and $\psi$ are FL-formulae, then so are $(\phi \land \psi)$, $(\phi \lor \psi)$, $(\phi \& \psi)$ and $(\phi \rightarrow \psi)$.

The set of FL-formulae is denoted by $F_{FL}$.

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

- among the connectives, $\&$ has the highest priority; furthermore $\land$ and $\lor$ take precedence over $\rightarrow$,
- the outermost brackets are not written.

The following notations are used: $\top$ for $0 \rightarrow 0$, $\neg \phi$ for $\phi \rightarrow 0$, $\phi^2$ for $\phi \& \phi$, $\phi^n$ (with $n \in \{3, 4, 5, \ldots\}$) for $(\phi^{n-1}) \& \phi$ (moreover, $\phi^0$ is $\top$ and $\phi^1$ is $\phi$), and $\phi \leftrightarrow \psi$ for $(\phi \rightarrow \psi) \land (\psi \rightarrow \phi)$, for FL-formulae $\phi$ and $\psi$.

The FL-formulae $\phi \& \psi$, $\phi \rightarrow \psi$ and $\neg \phi$ stand for what we understand intuitively by ‘$\phi$ and $\psi$’ (strong conjunction), ‘$\phi$ implies $\psi$’ (implication) and ‘not $\phi$’ (negation).

It is impossible to list all true FL-formulae of a specific fuzzy logic, because their number is not finite. Therefore axioms and deduction rules are used. In the logics we deal with, an FL-formula is true if it is provable from the axioms using the deduction rules. We will explain this in more detail. This method also allows to prove FL-formulae from a given set of FL-formulae (usually called a theory). This means that in the proof of an FL-formula not only axioms of the logic can be used.
but also formulae of the theory. If an FL-formula $\phi$ is provable from a theory $\Gamma$ in a fuzzy logic $L$, this is denoted as $\Gamma \vdash_L \phi$. The relation $\vdash_L$ is called provability relation or syntactic consequence.

It is often not very easy to find out if an FL-formula is true in a specific fuzzy logic. A proof might be difficult to find and such a proof can become very long. This is why soundness and completeness of (fuzzy) logics is so important. It provides a way to determine if a formula is true or provable from a theory in a purely algebraic way. Indeed, soundness and completeness of a fuzzy logic are two properties relative to a class of algebraic structures. We call such a class a semantics of the fuzzy logic. To explain this connection between formal logic and algebra in more detail, we need some terminology first.

**Definition 9.** Let $L = (L, \sqcap, \sqcup, *, \Rightarrow, 0, 1)$ be a residuated lattice, $\Gamma$ a theory (i.e., a set of FL-formulae). An $L$-evaluation is a mapping $e$ from the set of FL-formulae to $L$ that satisfies, for each two formulae $\phi$ and $\psi$:

- $e(\phi \land \psi) = e(\phi) \sqcap e(\psi)$,
- $e(\phi \lor \psi) = e(\phi) \sqcup e(\psi)$,
- $e(\phi & \psi) = e(\phi) * e(\psi)$,
- $e(\phi \rightarrow \psi) = e(\phi) \Rightarrow e(\psi)$ and
- $e(0) = 0$.

If an $L$-evaluation $e$ satisfies $e(\chi) = 1$ for every $\chi$ in $\Gamma$, it is called an $L$-model for $\Gamma$. We write $\Gamma \models_L \phi$ if $e(\phi) = 1$ for all $L$-models $e$ for $\Gamma$. If $\Gamma$ is empty, we simply write $\models_L \phi$. FL-formulae $\phi$ for which $\models_L \phi$ are called $L$-tautologies. The relation $\models_L$ is called semantic consequence.

Evaluations form a connection between the connectives of the logic and the algebraic operators in residuated lattices. Note that $e(1) = e(0 \Rightarrow 0) = e(0) \Rightarrow e(0) = 0 \Rightarrow 0 = 1$ and $e(\neg \phi) = e(\phi \rightarrow 0) = e(\phi) \Rightarrow e(0) = \neg e(\phi)$.

Now let $\mathcal{C}$ be a class of residuated lattices and $L$ a fuzzy logic.

- We say $L$ is sound w.r.t. $\mathcal{C}$ if for all $\Gamma \subseteq \mathcal{F}_FL$ and $\phi \in \mathcal{F}_FL$, $\Gamma \vdash_L \phi$ implies $\Gamma \models_L \phi$ for all $L$ in $\mathcal{C}$.
- We say $L$ is complete w.r.t. $\mathcal{C}$ if for all $\phi \in \mathcal{F}_FL$, ($\models_L \phi$ for all $L$ in $\mathcal{C}$) implies $\Gamma \models_L \phi$.
- We say $L$ is strong complete w.r.t. $\mathcal{C}$ if for all $\Gamma \subseteq \mathcal{F}_FL$ and $\phi \in \mathcal{F}_FL$, ($\Gamma \models_L \phi$ for all $L$ in $\mathcal{C}$) implies $\Gamma \models_L \phi$.

We will illustrate these definitions in the following subsections.

---

6 Note that an $\mathcal{L}$-evaluation is completely determined by its action on the propositional variables.

7 Note that $\mathcal{L}$-models for the empty set are just $\mathcal{L}$-evaluations.

8 For every logic $L$ appearing in this chapter, completeness w.r.t. a class $\mathcal{C}$ of residuated lattices (or, in the next sections, triangle algebras) implies that, for every finite theory $\Gamma \subseteq \mathcal{F}_FL$ and $\phi \in \mathcal{F}_FL$, $\Gamma \vdash_L \phi$ if $\Gamma \models_L \phi$ for all $L$ in $\mathcal{C}$. In other words, completeness implies ‘strong completeness for finite theories’.
4.1.1 Monoidal Logic

Monoidal logic (ML) was introduced by Höhle in [26]. Its axioms are:

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

This means that for all possible choices of FL-formulae for \(\phi\), \(\psi\) and \(\chi\), the above FL-formulae are provable in ML. For example, \((\overline{0} \rightarrow q) \rightarrow ((p \land q) \lor (\overline{0} \rightarrow q))\) is provable in ML, because it is an instance of the third axiom, with \(\psi = (\overline{0} \rightarrow q)\) and \(\phi = p \land q\).

To show that other FL-formulae are provable in ML, ML has one deduction rule. This deduction rule is called modus ponens (MP) and states that if \(\phi\) and \(\phi \rightarrow \psi\) are provable then so is \(\psi\). Now we can formally define what a proof in ML of an FL-formula \(\phi\) from a theory \(\Gamma\) is: it is a finite sequence of FL-formulae in which every FL-formula is either an instance of an axiom of ML, an element of \(\Gamma\) or the result of an application of the modus ponens to two FL-formulae appearing earlier in the sequence. If such a proof exists, this is denoted as \(\Gamma \vdash_{ML} \phi\). If \(\Gamma\) is empty, we simply write \(\vdash_{ML} \phi\) instead of \(\overline{0} \vdash_{ML} \phi\). An important result about ML is its soundness and completeness w.r.t. residuated lattices.

**Theorem 2.** ([26]) Monoidal logic is sound and strongly complete w.r.t. residuated lattices. In other words: for all \(\Gamma \subseteq \mathcal{F}_{FL}\) and \(\phi \in \mathcal{F}_{FL}\), we have \(\Gamma \vdash_{ML} \phi\) iff \(\Gamma \models_{\mathcal{L}} \phi\) for all residuated lattices \(\mathcal{L}\).

---

9 In [26] there are two axioms instead of the last one. In these two axioms the negation \(\neg\) appears (as a unary connective), but not the constant \(\overline{0}\). In this work we have chosen another way, namely to define the negation based on the constant \(\overline{0}\) (instead of the other way around).

Moreover, also \((\phi \land (\psi \land \chi)) \rightarrow ((\phi \land \psi) \land \chi)\) was listed as an axiom. But FL-formulae of this form can be proven from the other axioms, so it can be left out.

10 An instance of an axiom of ML is any FL-formula obtained by replacing \(\phi\), \(\psi\) and \(\chi\) with FL-formulae. For example, \(((p_1 \rightarrow (p_2 \lor \overline{0})) \land (p_2 \land q)) \rightarrow (p_1 \rightarrow (p_2 \lor \overline{0}))\) is an instance of \((\phi \land \psi) \rightarrow \phi\).
Each identity or inequality that is valid in residuated lattices can easily be transformed into a scheme of FL-formulae that are provable in ML. Therefore we just need to change such an identity or inequality to an equivalent ‘equal to 1 identity’.

- An example with an identity: \( x \Rightarrow (y \sqcap z) = (x \Rightarrow y) \sqcap (x \Rightarrow z) \) holds in all residuated lattices. This is equivalent with \( (x \Rightarrow (y \sqcap z)) \iff ((x \Rightarrow y) \sqcap (x \Rightarrow z)) = 1 \), which can be immediately transformed into the scheme \( (\phi \rightarrow (\psi \wedge \chi)) \leftrightarrow ((\phi \rightarrow \psi) \wedge (\phi \rightarrow \chi)) \) of FL-formulae that are provable in ML.

- An example with an inequality: \( x \ast y \leq x \sqcap y \) holds in all residuated lattices. This is equivalent with \( (x \ast y) \Rightarrow (x \sqcap y) = 1 \), which can be transformed into the scheme \( (\phi \& \psi) \rightarrow (\phi \wedge \psi) \) of FL-formulae that are provable in ML.

So there is a close connection between identities (and inequalities) in residuated lattices and (schemes of) FL-formulae that are provable in ML.

ML enjoys a so-called local deduction theorem:

**Theorem 3.** [24] Let \( \Gamma \cup \{ \phi, \psi \} \) be a set of FL-formulae. Then the following are equivalent:

- \( \Gamma \cup \{ \phi \} \vdash_{ML} \psi \),
- there is an integer \( n \) such that \( \Gamma \vdash_{ML} \phi^n \rightarrow \psi \).

This local deduction theorem, as well as the soundness and completeness of ML, remain valid in axiomatic extensions of ML. An axiomatic extension of ML is a logic having the same axioms and deduction rule as ML, plus one or more other axioms. Axiomatic extensions of ML are sound and complete w.r.t. residuated lattices that satisfy the identities corresponding to the extra axioms.

### 4.1.2 Monoidal t-Norm Based Logic

Monoidal t-norm based logic (MTL) [16] is an axiomatic extension of ML. The extra axiom is \( (\phi \rightarrow \psi) \lor (\psi \rightarrow \phi) \). This axiom corresponds with the identity \( (x \Rightarrow y) \sqcup (y \Rightarrow x) = 1 \) (prelinearity) in residuated lattices. Therefore we have the following soundness and completeness theorem.

**Theorem 4.** [16] Monoidal t-norm based logic is sound and strong complete w.r.t. prelinear residuated lattices (MTL-algebras). In other words: for all \( \Gamma \subseteq \mathcal{F}_{FL} \) and \( \phi \in \mathcal{F}_{FL} \), we have \( \Gamma \vdash_{MTL} \phi \) iff \( \Gamma \models_{L} \phi \) for all MTL-algebras \( L \).

The provability relation \( \vdash_{MTL} \) is defined in the same way as \( \vdash_{ML} \), only now also instances of the axiom \( (\phi \rightarrow \psi) \lor (\psi \rightarrow \phi) \) may appear in proofs of FL-formulae.

Because MTL-algebras are isomorphic to subdirect products of MTL-chains, the strong completeness can be strengthened to so-called (strong) chain completeness, i.e., (strong) completeness w.r.t. MTL-chains.

---

11 A scheme of FL-formulae consists of all FL-formulae of a particular form (the instances of the scheme). Note that the axioms of ML are schemes.
Theorem 5. [16] Monoidal t-norm based logic is sound and strong chain complete. In other words: for all $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{MTL} \phi$ iff $\Gamma \models_{L} \phi$ for all MTL-chains $L$.

Chain completeness and strong chain completeness are properties that remain valid for axiomatic extensions of MTL. Important theorems about MTL that do not necessarily remain valid for axiomatic extensions, are standard completeness and strong standard completeness. Recall that standard MTL-algebras are MTL-algebras on the unit interval.

Theorem 6. [29] Monoidal t-norm based logic is sound and strong standard complete. In other words: for all $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{MTL} \phi$ iff $\Gamma \models_{L} \phi$ for all standard MTL-chains $L$.

Because of this theorem, and because standard MTL-chains are induced by left-continuous t-norms, we can say that MTL is ‘the logic of left-continuous t-norms’.

MTL enjoys the same local deduction theorem as ML:

Theorem 7. [24] Let $\Gamma \cup \{\phi, \psi\}$ be a set of FL-formulae. Then the following are equivalent:

1. $\Gamma \cup \{\phi\} \vdash_{MTL} \psi$,
2. there is an integer $n$ such that $\Gamma \vdash_{MTL} \phi^n \rightarrow \psi$.

4.1.3 Basic Logic

Basic logic (BL) [24] is an axiomatic extension of MTL (and therefore also of ML). The extra axiom is $(\phi \land \psi) \rightarrow (\phi \& (\phi \rightarrow \psi))$. This axiom corresponds with the identity $(x \sqcap y) \Rightarrow (x \ast (x \Rightarrow y)) = 1$ in residuated lattices, in other words $x \sqcap y \leq x \ast (x \Rightarrow y)$. This is equivalent with $x \sqcap y = x \ast (x \Rightarrow y)$ (divisibility) because $x \sqcap y \geq x \ast (x \Rightarrow y)$ is valid in all residuated lattices. Because BL is an axiomatic extension of MTL and because divisible MTL-algebras are BL-algebras, we immediately have the following soundness and completeness result.

Theorem 8. [24] Basic logic is sound and strong chain complete. In other words: for all $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{BL} \phi$ iff $\Gamma \models_{L} \phi$ for all BL-chains $L$.

Note that BL is also (strong) complete w.r.t. all BL-algebras, which is a weaker property than (strong) chain completeness. BL also satisfies standard completeness (which is stronger than chain completeness), but not strong standard completeness. This implies that the following theorem in general does not hold for infinite theories $\Gamma$.

Theorem 9. [13] For all finite $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{BL} \phi$ iff $\Gamma \models_{L} \phi$ for all standard BL-chains $L$.

Because standard BL-chains are induced by continuous t-norms, we can say BL is ‘the logic of continuous t-norms’.

The local deduction theorem is also valid for BL.
Theorem 10. Let $\Gamma \cup \{\phi, \psi\}$ be a set of FL-formulae. Then the following are equivalent:

- $\Gamma \cup \{\phi\} \vdash_{BL} \psi$,
- there is an integer $n$ such that $\Gamma \vdash_{BL} \phi^n \rightarrow \psi$.

4.1.4 Łukasiewicz Logic

Łukasiewicz logic ($\mathcal{L}$) is an axiomatic extension of BL (and therefore also of MTL and ML). The extra axiom is $\neg\neg \phi \rightarrow \phi$, in which $\neg \phi$ is a short notation for $\phi \rightarrow \overline{0}$. This axiom corresponds with the identity $\neg\neg x \Rightarrow x = 1$ in residuated lattices, in other words $\neg\neg x \leq x$. This is equivalent with $\neg\neg x = x$ (involutive negation) because $x \leq \neg\neg x$ is valid in all residuated lattices. Because $\mathcal{L}$ is an axiomatic extension of MTL, we immediately have the following soundness and completeness result.

Theorem 11. Łukasiewicz logic is sound and strong chain complete. In other words: for all $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{\mathcal{L}} \phi$ iff $\Gamma \models_{\mathcal{L}} \phi$ for all MV-chains $\mathcal{L}$.

Note that $\mathcal{L}$ is also (strong) complete w.r.t. all MV-algebras, which is a weaker property than (strong) chain completeness. $\mathcal{L}$ also satisfies standard completeness (which is stronger than chain completeness), but not strong standard completeness. This implies that the following theorem in general does not hold for infinite theories.

Theorem 12. For all finite $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{\mathcal{L}} \phi$ iff $\Gamma \models_{\mathcal{L}} \phi$ for all standard MV-chains $\mathcal{L}$.

Because standard MV-chains are induced by t-norms that are conjugated to the Łukasiewicz t-norm, we can say that $\mathcal{L}$ is ‘the logic of the Łukasiewicz t-norm’.

The local deduction theorem is also valid for $\mathcal{L}$.

Theorem 13. Let $\Gamma \cup \{\phi, \psi\}$ be a set of FL-formulae. Then the following are equivalent:

- $\Gamma \cup \{\phi\} \vdash_{\mathcal{L}} \psi$,
- there is an integer $n$ such that $\Gamma \vdash_{\mathcal{L}} \phi^n \rightarrow \psi$.

4.1.5 Classical Logic

Classical logic (CPC\(^{12}\)) is an axiomatic extension of $\mathcal{L}$ (and therefore also of BL, MTL and ML). The extra axiom is $\phi \rightarrow (\phi \& \phi)$. This axiom corresponds with the identity $x \Rightarrow (x \ast x) = 1$ in residuated lattices, in other words $x \leq x \ast x$. This is equivalent with $x = x \ast x$ (contraction) because $x \ast x \leq x$ is valid in all residuated lattices. Because CPC is an axiomatic extension of MTL, we immediately have the following soundness and completeness result.

Theorem 14. Classical logic is sound and strong chain complete. In other words: for all $\Gamma \subseteq \mathcal{F}_{FL}$ and $\phi \in \mathcal{F}_{FL}$, we have $\Gamma \vdash_{\text{CPC}} \phi$ iff $\Gamma \models_{\mathcal{L}} \phi$ for all linear Boolean algebras $\mathcal{L}$.

\(^{12}\) The abbreviation, taken from [6], stands for ‘classical propositional calculus’.
Note that there is only one linear Boolean algebra (apart from the trivial one with one element), namely the Boolean algebra with two elements, 0 and 1. Therefore CPC cannot satisfy standard completeness, in the sense that CPC is not complete w.r.t. Boolean algebras on the unit interval (because there are no such Boolean algebras).

The local deduction theorem is also valid for CPC. But because in CPC the FL-formulae $\phi$ and $\phi^n$ (with $n$ a strictly positive integer) are equivalent (meaning $\vdash_{CPC} \phi \leftrightarrow \phi^n$ holds\footnote{In the logics we are concerned with in this work, we have the following property. If a subformula of a formula is replaced by an equivalent subformula, the resulting formula is equivalent with the original one. This can be proven using soundness and completeness. For example, in ML and its axiomatic extensions $(p_1 \& p_2) \to q$ is equivalent with $(p_2 \& p_1) \to q$ because the subformulae $p_1 \& p_2$ and $p_2 \& p_1$ are equivalent.}), the theorem can be strengthened.

**Theorem 15.** \cite{[32]} Let $\Gamma \cup \{\phi, \psi\}$ be a set of FL-formulae. Then the following are equivalent:

- $\Gamma \cup \{\phi\} \vdash_{CPC} \psi$,
- $\Gamma \vdash_{CPC} \phi \to \psi$.

### 4.1.6 Other Fuzzy Logics

Apart from the examples in the previous sections, many other fuzzy logics can be defined by adding axioms to ML, MTL, . . . Examples can be found in e.g. \cite{[6]}. In \cite{[6]} it was proven that an axiomatic extension of MTL is strong standard complete iff it has the real-chain embedding property, i.e., iff each countable MTL-chain in its semantics is embeddable in a standard MTL-chain in its semantics.

## 4.2 Interval-Valued Monoidal Logic: Definition

As semantics of interval-valued fuzzy logics, we choose triangle algebras. Because triangle algebras have more operators than residuated lattices, IVFL-formulae can contain more connectives than FL-formulae.

**Definition 10.** \cite{[41]} IVFL-formulae are built up from a countable set of propositional variables (denoted by $p, q, r, p_1, p_2, \ldots$) and the constants $\underline{0}$ and $\underline{1}$. These symbols are IVFL-formulae by definition. The other IVFL-formulae are defined recursively: if $\phi$ and $\psi$ are IVFL-formulae, then so are $(\phi \land \psi), (\phi \lor \psi), (\phi \& \psi), (\phi \to \psi), \Box \phi$ and $\Diamond \phi$. The set of IVFL-formulae is denoted by $\mathcal{F}_{IVFL}$. Note that $\mathcal{F}_{FL} \subseteq \mathcal{F}_{IVFL}$.

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

- unary operators always take precedence over binary ones, while
- among the connectives, $\&$ has the highest priority; furthermore $\land$ and $\lor$ take precedence over $\to$,
- the outermost brackets are not written.

The same notations (T is $\underline{0} \to \underline{1}, \ldots$) as for FL-formulae are used. Now we are ready to introduce interval-valued monoidal logic (IVML) \cite{[41]}. Its axioms are those of
ML, i.e., (ML.1)–(ML.12), complemented with axioms corresponding to 13 properties of triangle algebras:

\[
\begin{align*}
(IVML.2) & \quad \phi \rightarrow \Box \Box \phi, \\
(IVML.3) & \quad \Box (\phi \land \psi) \rightarrow \Box (\phi \land \psi), \\
(IVML.4) & \quad \Box (\phi \lor \psi) \rightarrow (\Box \phi \lor \Box \psi), \\
(IVML.5) & \quad \neg \Box \psi \rightarrow \neg \Box \psi, \\
(IVML.6) & \quad \Box \phi \rightarrow \Box \Box \phi, \\
(IVML.7) & \quad \Box (\phi \rightarrow \psi) \rightarrow (\Box \phi \rightarrow \Box \psi), \\
(IVML.8) & \quad (\Box \phi \rightarrow \Box \psi) \rightarrow (\Box \phi \rightarrow \Box \psi), \\
(IVML.9) & \quad \phi \leftrightarrow (\Box \phi \lor (\Box \phi \land \Box \psi)).
\end{align*}
\]

All instances of these axioms are by definition provable in IVML. To determine which other IVFL-formulae are provable, there are three deduction rules: modus ponens (MP, if \( \phi \) and \( \phi \rightarrow \psi \) are provable in IVML, then so is \( \psi \)), generalization (G, if \( \phi \) is provable in IVML, then so is \( \Box \phi \)) and monotonicity of \( \Box \) (M\( \Box \), if \( \phi \rightarrow \psi \) is provable, then so is \( \Box \phi \rightarrow \Box \psi \)). Proofs in IVML and the provability relation \( \vdash_{IVML} \) are defined in the usual way, similarly as for ML (and the other fuzzy logics from Section 4.1). If \( \Gamma \) is a theory, i.e., a set of IVFL-formulae, then a (formal) proof of an IVFL-formula \( \phi \) in \( \Gamma \) is a finite sequence of IVFL-formulae with \( \phi \) at its end, such that every IVFL-formula in the sequence is either an instance of an axiom of IVML, an IVFL-formula of \( \Gamma \), or the result of an application of a deduction rule to previous IVFL-formulae in the sequence. If a proof for \( \phi \) exists in \( \Gamma \), we denote this by \( \Gamma \vdash_{IVML} \phi \).

**Definition 11.** \([41] \) Let \( \mathcal{A} = (A, \cap, \cup, *, \Rightarrow, v, \mu, 0, 1) \) be a triangle algebra, \( \Gamma \) a theory (i.e., a set of IVFL-formulae). An \( \mathcal{A} \)-evaluation is a mapping \( e \) from the set of IVFL-formulae to \( A \) that satisfies, for each two IVFL-formulae \( \phi \) and \( \psi \):

- \( e(\phi \land \psi) = e(\phi) \cap e(\psi) \),
- \( e(\phi \lor \psi) = e(\phi) \cup e(\psi) \),
- \( e(\phi \& \psi) = e(\phi) * e(\psi) \),
- \( e(\phi \rightarrow \psi) = e(\phi) \Rightarrow e(\psi) \),
- \( e(\Box \phi) = v e(\phi) \),
- \( e(\Box \phi) = \mu e(\phi) \),
- \( e(\Box \phi) = 0 \) and
- \( e(\Box \phi) = 1 \).

If an \( \mathcal{A} \)-evaluation \( e \) satisfies \( e(\chi) = 1 \) for every \( \chi \) in \( \Gamma \), it is called an \( \mathcal{A} \)-model for \( \Gamma \). We write \( \Gamma \models_{\mathcal{A}} \phi \) if \( e(\phi) = 1 \) for all \( \mathcal{A} \)-models \( e \) for \( \Gamma \).

Soundness, completeness and strong completeness are defined similarly as for formal fuzzy logics. We just have to replace ‘residuated lattice’ by ‘triangle algebra’ and ‘FL-formula’ by ‘IVFL-formula’.

Now we introduce some axiomatic extensions of IVML, by adding well-known axioms. Note that these axioms are applied to IVFL-formulae of the form \( \Box \phi \).
instead of to all IVFL-formulae. As the image of a triangle algebra \((A, \sqcap, \sqcup, *, \Rightarrow, \nu, \mu, 0, u, 1)\) under \(\nu\) is the set \(E(\mathcal{A})\) of exact elements\(^{14}\), this means that the axioms schemes do not hold for all truth values, but only for exact truth values. This is not a drawback. On the contrary, it is precisely what we want because the exact truth values are easier to interpret and handle. Moreover, using Proposition\(^3\) for all axioms equivalent axioms can be found that only involve IVFL-formulae of the form \(\Box \phi\) and \(\Diamond \phi\), and \(\overline{\pi}\).

**Definition 12.**\(^{38}\)

- Interval-valued monoidal t-norm based logic (IVMTL) is IVML extended with the axiom scheme pseudo-prelinearity
  \[(\Box \phi \rightarrow \Box \psi) \lor (\Box \psi \rightarrow \Box \phi).\]
- Interval-valued basic logic (IVBL) is IVMTL extended with the axiom scheme pseudo-divisibility
  \[(\Box \phi \land \Box \psi) \rightarrow (\Box \phi \& (\Box \phi \rightarrow \Box \psi)).\]
- Interval-valued Łukasiewicz logic (IVŁ) is IVBL extended with the axiom scheme pseudo-involution
  \[\neg \neg \Box \phi \rightarrow \Box \phi.\]
- Interval-valued classical propositional calculus (IVCPC) is IVŁ extended with the axiom scheme pseudo-contraction
  \[\Box \phi \rightarrow (\Box \phi \& \Box \phi).\]

### 4.3 Soundness and Completeness

It is easy to check that IVML is sound w.r.t. the variety of triangle algebras, i.e., that if an IVFL-formula \(\phi\) can be proven from a theory \(\Gamma\) in IVML \((\Gamma \vdash_{IVML} \phi)\), then for every triangle algebra \(\mathcal{A}\) and for every \(\mathcal{A}\)-model \(e\) of \(\Gamma\), \(e(\phi) = 1\) (in other words: for every triangle algebra \(\mathcal{A}\), \(\Gamma \models_{\mathcal{A}} \phi\)). To show that IVML is also strong complete (w.r.t. triangle algebras), i.e., that the converse of soundness also holds, a general result from abstract algebraic logic (shortly AAL, see e.g.\(^{19}\) for a survey) can be applied. It proceeds by showing that IVML is an implicative logic (in the sense of Rasiowa\(^{33}\)). From this we can deduce (according to e.g.\(^{18}\)) that IVML is strong complete w.r.t. the variety of triangle algebras if it is sound w.r.t. it and if in triangle algebras \(x = y\) if \(x \Rightarrow y = 1\) and \(y \Rightarrow x = 1\). Triangle algebras indeed satisfy these conditions, so we can conclude that IVML is sound and strong complete w.r.t. triangle algebras.

---

\(^{14}\) Note that the image under \(\mu\) is also \(E(\mathcal{A})\). All axiom schemes in Definition\(^{12}\) can also be given in an equivalent way by changing \(\Box \phi\) to \(\Diamond \phi\) and/or \(\Box \psi\) to \(\Diamond \psi\).
Theorem 16. (Soundness and strong completeness of IVML) \[41\] An IVFL-formula $\phi$ can be deduced from a theory $\Gamma$ in IVML iff for every triangle algebra $\mathcal{A}$ and for every $\mathcal{A}$-model $e$ of $\Gamma$, $e(\phi) = 1$.

A more basic proof for the completeness of IVML, using the Lindenbaum algebra of IVML and inspired by a similar, commonly used procedure for formal fuzzy logics, was given in [36].

Theorem 16 implies similar results for axiomatic extensions (e.g. the interval-valued fuzzy logics in Definition 12), in the same way as the soundness and completeness of ML remains valid for axiomatic extensions. This can be seen by taking the set of all instances of the extra axioms as $\Gamma$ in Theorem 16. In particular, all extensions of IVML introduced in Section 4.2 are sound and (strong) complete w.r.t. their corresponding subvariety of the variety of triangle algebras. For example, IVBL is sound and complete w.r.t. the variety of triangle algebras satisfying $(vx \Rightarrow vy) \sqcup (vy \Rightarrow vx) = 1$ and $vx \sqcap vy \leq vx \ast (vx \Rightarrow vy)$.

For IVMTL and its axiomatic extensions we can prove a stronger version of completeness, namely strong pseudo-chain completeness. This is similar to axiomatic extensions of MTL being strong chain-complete. Together with Theorem 16, Theorem 1 implies the following result:

**Theorem 17.** [42] For each set of IVFL-formulae $\Gamma \cup \{ \phi \}$, the following three statements are equivalent:

- $\phi$ can be deduced from a theory $\Gamma$ in IVMTL ($\Gamma \vdash_{IVMTL} \phi$),
- for every pseudo-prelinear triangle algebra $\mathcal{A}$ and for every $\mathcal{A}$-model $e$ of $\Gamma$, $e(\phi) = 1$,
- for every pseudo-linear triangle algebra $\mathcal{A}$ and for every $\mathcal{A}$-model $e$ of $\Gamma$, $e(\phi) = 1$.

This completeness result remains valid for axiomatic extensions of IVMTL. The reason is that Theorem 1 also holds for subvarieties of pseudo-prelinear triangle algebras.

Using Theorem 6 and the real-chain embedding property, also the strong standard completeness of IVMTL can be proven.

**Theorem 18.** [38] (Strong standard completeness) For each set of IVFL-formulae $\Gamma \cup \{ \phi \}$, the following four statements are equivalent:

1. $\phi$ can be deduced from $\Gamma$ in IVMTL ($\Gamma \vdash_{IVMTL} \phi$),
2. for every pseudo-prelinear triangle algebra $\mathcal{A}$, $\Gamma \models_{\mathcal{A}} \phi$ (i.e., for every $\mathcal{A}$-model $e$ of $\Gamma$, $e(\phi) = 1$),
3. for every pseudo-linear triangle algebra $\mathcal{A}$, $\Gamma \models_{\mathcal{A}} \phi$,
4. for every standard triangle algebra $\mathcal{A}$, $\Gamma \models_{\mathcal{A}} \phi$.

Because of Proposition 2, every standard triangle algebra is isomorphic to a standard extended IVRL, and every standard extended IVRL is a standard triangle algebra. This result leads to the following corollary of Theorem 18.

**Corollary 1.** For each set of IVFL-formulae $\Gamma \cup \{ \phi \}$, the following statements are equivalent:
1. $\phi$ can be deduced from $\Gamma$ in IVMTL ($\Gamma \vdash_{IVMTL} \phi$).
2. for every standard extended IVRL $\mathcal{A}$, $\Gamma \models_{\mathcal{A}} \phi$.

So we can truly state that IVMTL is an interval-valued fuzzy logic. It is the logic of the t-norms $\mathcal{T}_{T,t}$ in (2), with $T$ a left-continuous t-norm on the unit interval.

Finally, it is also possible to prove the local deduction theorem for IVML (and its axiomatic extensions), which gives a connection between $\vdash_L$ and $\to$.

**Theorem 19.** [38] Let $\Gamma \cup \{ \phi, \psi \}$ be a set of IVFL-formulae, and $L$ be an axiomatic extension of IVML. Then the following are equivalent:

- $\Gamma \vdash_L \psi$.
- There is an integer $n$ such that $\Gamma \vdash_L (\Box \phi)^n \to \psi$.

### 4.4 Interpretation

Interval-valued fuzzy logics as we introduced them, are truth-functional logics: the truth degree of a compound proposition is determined by the truth degree of its constituent parts. This causes some counterintuitive results, if we want to interpret the element $[0, 1]$ of an IVRL as uncertainty regarding the actual truth value of a proposition. For example: suppose we don’t know anything about the truth value of propositions $p$ and $q$, i.e., $e(p) = e(q) = [0, 1]$. Then yet the implication $p \to q$ is definitely valid: $e(p \to q) = e(p) \Rightarrow e(q) = [1, 1]$. However, if $\neg [0, 1] = [0, 1]^{15}$ (which is intuitively preferable, since the negation of an uncertain proposition is still uncertain), then we can take $q = \neg p$, and obtain that $p \to \neg p$ is true. Or, equivalently (using the residuation principle), that $p \& p$ is false. This does not seem intuitive, as one would rather expect $p \& p$ to be uncertain if $p$ is uncertain.

Another consequence of $[0, 1] \Rightarrow [0, 1] = [1, 1]$ is that it is impossible to interpret the intervals as a set in which the ‘real’ (unknown) truth value is contained, and $X \Rightarrow Y$ as the smallest closed interval containing every $x \Rightarrow y$, with $x$ in $X$ and $y$ in $Y$ (as in [15]). Indeed: $1 \in [0, 1]$ and $0 \in [0, 1]$, but $1 \Rightarrow 0 = 0 \notin [1, 1]$.

On the other hand, for t-norms it is possible that $X \ast Y$ is the smallest closed interval containing every $x \ast y$, with $x$ in $X$ and $y$ in $Y$, but only if they are t-representable (described by the axiom $\mu(x \ast y) = \mu x \ast \mu y$). However, in this case $\neg[0, 1] = [0, 0]$, which does not seem intuitive (‘the negation of an uncertain proposition is absolutely false’).

These considerations seem to suggest that IVML and its axiomatic extensions are not suitable to reason with uncertainty. This does not mean that intervals are not a good way for representing degrees of uncertainty, only that they are not suitable as truth values in a truth-functional logical calculus when we interpret them as expressing uncertainty. It might even be impossible to model uncertainty as a truth value in any truth-functional logic. This question is discussed in [12, 13]. However, nothing prevents the intervals in interval-valued fuzzy logics from having more adequate interpretations.

---

15 This is for example the case if $\neg$ is involutive.
Acknowledgements. Bart Van Gasse and Chris Cornelis would like to thank the Research Foundation–Flanders for funding their research.

References

34. Shoenfield, J.: Mathematical Logic. Addison-Wesley, Reading (1967)


44. Zalta, E.: Basic concepts in modal logic (1995),

http://mally.stanford.edu/notes.pdf
Implications in Fuzzy Logic: Properties and a New Class

Yun Shi, Bart Van Gasse, and Da Ruan

Abstract. An implication in fuzzy logic, commonly defined as a two-place operation on the unit interval, is an extension of the classical binary implication. It plays important roles in both mathematical and applied sides of fuzzy set theory. Besides the basic properties, there are many potential properties for implications, among which eight are widely used in the literature. Different implications satisfying different subgroups of these eight properties can be found. However, certain interrelationships exist between these eight properties. This chapter aims to lay bare the interrelationships between these eight properties. When searching counterexamples to prove the independencies we discover a new class of implications determined only by a negation. We then examine under which conditions the eight properties are satisfied. Finally, we obtain the intersection of the new class of implications with the S- and R- implications.

1 Introduction to Implication Operators in Fuzzy Logic

Since Lotfi A. Zadeh introduced the concept of fuzzy sets in his pioneering work in 1965 [28], a huge amount of research work about fuzzy sets and fuzzy logic has appeared in the literature of mathematics and computer science. Fuzzy logic is an extension of fuzzy sets. Fuzzy logic in the narrow sense refers to a kind of many-valued logic, i.e., logic that maintains more than two truth values. Fuzzy logic in the broad sense refers to the theory of approximate reasoning and the theory of linguistic logic whose truth values are linguistic terms represented by fuzzy sets (i.e., membership functions) [18, 29].

One of the most important research areas in fuzzy logic is to extend the connective operators NOT(¬), AND(∧), OR(∨) and IMPLY(→) from binary logic to fuzzy logic. Our research work focuses mainly on implications in fuzzy logic.

1.1 Negations, Conjunctions and Disjunctions

A negation in fuzzy logic is an extension of the negation operator ¬ in binary logic: A [0, 1] → [0, 1] mapping N is a negation if it satisfies N(0) = 1, N(1) = 0,
and \( x \leq y \Rightarrow N(x) \geq N(y) \), for all \( x, y \in [0, 1] \). Moreover, a strong negation \( N \) is a negation that satisfies: \( N(N(x)) = x \), for all \( x \in [0, 1] \).

Strong negations are always continuous. But the converse is not true. One of the famous classes of strong negations are the Sugeno negations \( N_a \): there exists an \( a \in ]-1, +\infty[ \) such that for all \( x \in [0, 1] \), \( N_a(x) = \frac{1-x}{1+ax} \). Notice that if \( a = 0 \), then \( N_a \) is the standard negation \( N_0 \), \( N_0(x) = 1 - x \). An example of a class of non-continuous negations is:

\[
N_A(x) = \begin{cases} 
1, & \text{if } x \in A, \\
0, & \text{if } x \notin A \end{cases}, \quad \text{for all } x \in [0, 1],
\]

where \( A = [0, \alpha[ \), with \( \alpha \in ]0, 1] \), or \( A = [0, \alpha] \), with \( \alpha \in [0, 1] \). Notice that \( N_A \) is the class of negations that take values only in \( \{0, 1\} \). Another class of negations that will be useful later is:

\[
N_{A,\beta}(x) = \begin{cases} 
1, & \text{if } x \in A, \\
\frac{1-x}{1+\beta x}, & \text{if } x \notin A \end{cases}, \quad \text{for all } x \in [0, 1],
\]

where \( A = [0, \alpha[ \), with \( \alpha \in ]0, 1] \), or \( A = [0, \alpha] \), with \( \alpha \in [0, 1] \), and \( \beta \in ]-1, +\infty[ \). Notice that \( N_{(0),\beta} \) is the class of Sugeno negations.

A conjunction in fuzzy logic is an extension of the conjunction operator \( \land \) in binary logic. Widely used are triangular norms (\( t \)-norms) for short. A \( [0,1]^2 \to [0,1] \) mapping \( T \) is a \( t \)-norm if for all \( x, y, z \in [0, 1] \) it satisfies: \( T(x, 1) = x, y \leq z \Rightarrow T(x,y) \leq T(x,z) \), \( T(x,y) = T(y,x) \) and \( T(x, T(y,z)) = T(T(x,y),z) \).

A disjunction in fuzzy logic is an extension of the disjunction operator \( \lor \) in binary logic. Widely used are triangular conorms (\( t \)-conorms) for short. A \( [0,1]^2 \to [0,1] \) mapping \( S \) is a \( t \)-conorm if for all \( x, y, z \in [0, 1] \) it satisfies: \( S(x,0) = x, y \leq z \Rightarrow S(x,y) \leq S(x,z) \), \( S(x,y) = S(y,x) \) and \( S(x, S(y,z)) = S(S(x,y),z) \).

### 1.2 Implications

Define a statement as a sentence that can be attached with a truth value in \([0, 1]\). In binary logic, a statement is either attached with 0, which means that it is false, or attached with 1, which means that it is true. The implication operator in binary logic, denoted by \( \rightarrow \), has the following truth table:

<table>
<thead>
<tr>
<th>( p ) ( q )</th>
<th>( p \rightarrow q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 \ 0</td>
<td>1</td>
</tr>
<tr>
<td>0 \ 1</td>
<td>1</td>
</tr>
<tr>
<td>1 \ 0</td>
<td>0</td>
</tr>
<tr>
<td>1 \ 1</td>
<td>1</td>
</tr>
</tbody>
</table>

The implication operator \( \rightarrow \) represents the meaning of ‘if...then...’ in a conditional rule. An implication \( I \) is extended from the implication operator \( \rightarrow \) in binary logic. In order to coincide with \( \rightarrow \) on the set \( \{0, 1\} \), it should be a \( [0,1]^2 \to [0,1] \) mapping that at least satisfies: \( I(0,0) = I(0,1) = I(1,1) = 1 \) and \( I(1,0) = 0 \). However, there is no standard definition for an implication as for a conjunction, a disjunction and a
Implications in Fuzzy Logic: Properties and a New Class

negation in fuzzy logic. As Kerre [11] states ‘One of the main difficulties I have met during the preparation of lecture notes on some basic material concerning fuzzy set theory, consisted of a lack of standard definitions for basic elementary notions’. By taking into account the extensive literature about implications [3, 4, 5, 8, 10, 26], we have proposed in our work the following definition:

An implication $I$ is a $[0, 1]^2 \rightarrow [0, 1]$ mapping that satisfies:

FI1. the first place antitonicity FA:
  $$(\forall (x_1, x_2, y) \in [0, 1]^3)(x_1 < x_2 \Rightarrow I(x_1, y) \geq I(x_2, y));$$

FI2. the second place isotonicity SI:
  $$(\forall (x, y_1, y_2) \in [0, 1]^3)(y_1 < y_2 \Rightarrow I(x, y_1) \leq I(x, y_2));$$

FI3. dominance of falsity of antecedent DF:
  $$(\forall x \in [0, 1])(I(0, x) = 1);$$

FI4. dominance of truth of consequent DT:
  $$(\forall x \in [0, 1])(I(x, 1) = 1);$$

FI5. boundary condition BC:
  $$I(1, 0) = 0.$$ 

There are three important ways to construct implications from the other connective operators in fuzzy logic:

1. An strong implication (S-implication for short) $I_{S,N}$:
   $$(\forall (x, y) \in [0, 1]^2)(I_{S,N}(x, y) = S(N(x), y)).$$

2. An residuated implication (R-implication for short) $I_T$:
   $$(\forall (x, y) \in [0, 1]^2)(I_T(x, y) = \sup\{t \in [0, 1] \land T(x, t) \leq y\}).$$

3. A quantum logic implication operator (QL-implication operator for short) $I_{S,N,T}$:
   $$(\forall (x, y) \in [0, 1]^2)(I_{S,N,T}(x, y) = S(N(x), T(x, y))).$$

Notice that a QL-implication operator is not always an implication because it does not always satisfy FI1. For under which conditions does a QL-implication satisfy FI1, i.e., it is an implication, we refer to the articles [7,15,23], and Section 2.6 of book [3].

2 Dependencies and Independencies between Potential Properties FI6-FI13 of Implications

Besides FI1-FI5, sometimes an implication is required to have some additional properties to fulfill different requirements [3,4,5,9,8,17,20,27,30], among which the most important ones are, for all $x, y$ and $z \in [0, 1]$:

FI6. $I(1, x) = x$ (neutrality of truth, NT for short);

FI7. $I(x, I(y, z)) = I(y, I(x, z))$ (exchange principle, EP for short);

FI8. $I(x, y) = 1 \iff x \leq y$ (ordering principle, OP for short);

FI9. the mapping $N_I$ defined by
\( (\forall x \in [0,1])(N_I(x) = I(x,0)) \),

is a strong negation (strong negation principle, SN for short);

FI10. \( I(x,y) \geq y \) (consequent boundary, CB for short);

FI11. \( I(x,x) = 1 \) (identity, ID for short);

FI12. \( I(x,y) = I(N(y),N(x)) \), where \( N \) is a strong negation (contrapositive principle, CP for short);

FI13. \( I \) is a continuous mapping (continuity, CO for short).

Certain interrelationships exist between these eight properties. This section aims to lay bare the interrelationships between these eight properties. The result is instrumental to propose a classification of implications.

### 2.1 Getting Neutrality of Truth (NT) from the Other Properties

Theorem 1. ([3], Lemma 1.54(v), Corollary 1.57 (iii)) an implication \( I \) satisfying SN and CP w.r.t. a strong negation \( N \) satisfies NT iff \( N_I = N \).

In the rest of this section we consider the condition that \( N_I \neq N \).

Proposition 1. ([11], Lemma 6) An implication \( I \) satisfying EP and OP satisfies NT.

Proposition 2. ([3], Lemma 1.56(ii)) an implication \( I \) satisfying EP and SN satisfies NT.

Proposition 3. An implication \( I \) satisfying EP and CO satisfies NT.

Proof. Because \( I \) satisfies EP, we have for all \( x \in [0,1] \),

\[
I(1,N_I(x)) = I(1,I(x,0)) = I(x,I(1,0)) = I(x,0) = N_I(x). \tag{3}
\]

Because \( I \) is a continuous mapping, \( N_I \) is a continuous mapping. Thus expression (3) is equivalent to \( I(1,a) = a \), for all \( a \in [0,1] \). Hence \( I \) satisfies NT.

Proposition 4. There exists an implication \( I \) satisfying EP, CB, ID, CP and not NT.

Counterexample: The implication \( I_1 \) stated in [6] is defined by

\[
I_1(x,y) = \begin{cases} 0 & \text{if } x = 1 \text{ and } y = 0, \\ 1 & \text{else} \end{cases}, \quad \text{for all } x,y \in [0,1].
\]

\( I_1 \) satisfies CB, ID and CP w.r.t. any strong negation \( N \). However, in case that \( x \neq 1 \), \( I_1(1,x) = 1 \neq x \). Therefore \( I_1 \) does not satisfy NT.

Proposition 5. There exists an implication \( I \) satisfying OP, SN, CB, ID, CP, CO and not NT.
Counterexample: Let an implication $I_2$ be defined by

$$I_2(x, y) = \begin{cases} 
1 & \text{if } x \leq y \\
\frac{1}{\sqrt{1 - (x - y)^2}} & \text{if } x > y,
\end{cases} \quad \text{for all } x, y \in [0, 1].$$

$I_2$ satisfies OP, SN, CB, ID, CP w.r.t. the standard strong negation $N_0$, and CO. However, in case that $x \neq 1$ and $x \neq 0$, $I_2(1, x) = \sqrt{2x - x^2} \neq x$. Therefore $I_2$ does not satisfy NT.

Fig. 1 The implication $I_2$

So we considered all the possibilities that NT can be implied from the other seven properties. Moreover we stated for each independent case a counterexample.

2.2 Getting Exchange Principle (EP) from the Other Properties

**Proposition 6.** There exists an implication $I$ satisfying NT, OP, SN, CB, ID, CP, CO and not EP.

Counterexample: Let an implication $I_3$ be defined by

$$I_3(x, y) = \begin{cases} 
1 & \text{if } x \leq y \\
1 - (1 - y + xy)(x - y) & \text{if } x > y, \text{ for all } x, y \in [0, 1].
\end{cases}$$

$I_3$ satisfies NT, OP, SN, CB, ID, CP w.r.t. the standard strong negation $N_0$, and CO. However, take $x_0 = 0.3$, $y_0 = 0.9$ and $z_0 = 0.1$, we obtain $I(x_0, I(y_0, z_0)) \approx 0.9214$ and $I(y_0, I(x_0, z_0)) \approx 0.9210$. Therefore $I_3$ does not satisfy EP.
EP is thus independent of any of the other seven properties.

2.3 Getting FI8(OP) from the Other Properties

**Proposition 7.** There exists an implication \( I \) satisfying NT, EP, SN, CB, ID, CP, CO and not OP.

Counterexample: Given the strong negation \( N(x) = \sqrt{1 - x^2} \), for all \( x \in [0, 1] \). The S-implication \( I_4 \) generated by the t-conorm \( S_L(x,y) = \min(x+y,1) \) and the strong negation \( N \) is defined by

\[
I_4(x,y) = S_L(N(x), y) = \min(\sqrt{1 - x^2} + y, 1), \text{ for all } x, y \in [0, 1].
\]

Because \( I_4 \) is an S-implication generated from a continuous t-conorm and a strong negation, it satisfies NT, EP, SN, CB, CP w.r.t. the strong negation \( N \) and CO \[8\]. Moreover, for all \( x \in [0, 1] \), \( I_4(x,x) = 1 \). Therefore \( I_4 \) also satisfies ID. However, take \( x_0 = 0.5 \) and \( y_0 = 0.4 \), we obtain \( I(x_0, y_0) = 1 \) while \( x_0 > y_0 \). Therefore \( I_4 \) does not satisfy OP.

Therefore OP is thus independent of any of the other seven properties.
2.4 Getting Strong Negation Principle (SN) from the Other Properties

Proposition 8. ([3], Lemma 1.5.4(v)) An implication I satisfying NT and CP w.r.t. a strong negation N satisfies SN. Moreover, $N_I = N$.

Corollary 1. An implication I satisfying EP, OP and CP w.r.t. a strong negation N satisfies SN. Moreover, $N_I = N$.

Proof. Straightforward from Propositions 1 and 8.

Corollary 2. An implication I satisfying EP, CP w.r.t. a strong negation N and CO satisfies SN. Moreover, $N_I = N$.

Proof. Straightforward from Propositions 3 and 8.

Proposition 9. ([1], Lemma 14) ([8], Corollary 1.1) An implication I satisfying EP, OP and CO satisfies SN.

Proposition 10. ([8], Table 1.1) There exists an implication satisfying NT, EP, OP, CB, ID and not SN.

Proof. The Gödel implication

$$I_{GD}(x, y) = \begin{cases} 1, & \text{if } x \leq y, \\ y, & \text{if } x > y, \end{cases} \quad \text{for all } x, y \in [0, 1].$$
is an R-implication generated by the continuous t-norm \( T_M(x, y) = \min(x, y) \). Therefore \( I_{GD} \) satisfies NT, EP, CB and ID \([8]\). However we have for all \( x \in [0, 1] \):

\[
N_{I_{GD}}(x) = I_{GD}(x, 0) = \begin{cases} 
1, & \text{if } x = 0 \\
0, & \text{if } x > 0.
\end{cases}
\]

Therefore \( I_{GD} \) does not satisfy SN.

**Proposition 11.** There exists an implication \( I \) satisfying NT, EP, CB, ID, CO and not SN.

**Counterexample:** Given the negation \( N(x) = 1 - x^2 \), for all \( x \in [0, 1] \). The S-implication generated from the t-conorm \( S_L \) and the negation \( N \) is defined by

\[
I_5(x, y) = \min(1 - x^2 + y, 1), \text{ for all } x, y \in [0, 1].
\]

\( I_5 \) satisfies NT, CB, ID and CO. Moreover, because \( I_5 \) is an \((S, N)\)-implication generated from the Łukasiewicz t-conorm and the strict negation \( N(x) = 1 - x^2 \), it then also satisfies EP(\([3]\), Proposition 2.4.3(i)). However, we have for all \( x \in [0, 1] \)

\[
N_{I_5}(x) = I_5(x, 0) = 1 - x^2
\]

which is not a strong negation. Therefore \( I_5 \) does not satisfy SN.

![Fig. 4 The implication \( I_5 \)](image-url)
**Proposition 12.** There exists an implication $I$ satisfying $NT$, $OP$, $CB$, $ID$, $CO$ and not $SN$.

Counterexample: Let an implication $I_6$ be defined by

$$I_6(x, y) = \begin{cases} 
1, & \text{if } x \leq y \\
\frac{y}{1 + \sqrt{1 - x}} + \sqrt{1 - x}, & \text{if } x > y,
\end{cases} \text{ for all } x, y \in [0, 1].$$

$I_6$ satisfies $NT$, $OP$, $CB$, $ID$ and $CO$. However, we have for all $x \in [0, 1]$

$$N_{I_6}(x) = I_6(x, 0) = \sqrt{1 - x}$$

which is not a strong negation. Therefore $I_6$ does not satisfy $SN$.

**Proposition 13.** There exists an implication $I$ satisfying $EP$, $CB$, $ID$, $CP$ and not $SN$.

The implication $I_1$ stated in the proof of Proposition 4 satisfies $EP$, $CB$, $ID$ and $CP$ w.r.t. any strong negation $N$. However, we have

$$N_{I_1}(x) = I_1(x, 0) = \begin{cases} 
1, & \text{if } x < 1 \\
0, & \text{if } x = 1
\end{cases}, \text{ for all } x \in [0, 1],$$

which is not a strong negation. Therefore $I_1$ does not satisfy $SN$.

**Proposition 14.** There exists an implication $I$ satisfying $OP$, $CB$, $ID$, $CP$, $CO$ and not $SN$. 

![Fig. 5 The implication $I_6$](image-url)
Counterexample: Let an implication \( I_7 \) be defined by

\[
I_7(x, y) = \begin{cases} 
1, & \text{if } x \leq y \\
\sqrt{1 - (x - y)}, & \text{if } x > y 
\end{cases}
\text{ for all } x, y \in [0, 1].
\]

\( I_7 \) satisfies OP, CB, ID, CP w.r.t. the standard strong negation \( N_0 \), and CO. However, we have for all \( x \in [0, 1] \)

\[
N_{I_7}(x) = I_7(x, 0) = \sqrt{1 - x},
\]

which is not a strong negation. Therefore \( I_7 \) does not satisfy SN.

![Fig. 6 The implication \( I_7 \)]

So we considered all the possibilities that SN can be implied from the other seven properties. Moreover we stated for each independent case a counterexample.

### 2.5 Getting Consequent Boundary (CB) from the Other Properties

**Proposition 15.** ([7], Lemma 1 (viii)) An implication \( I \) satisfying NT satisfies CB.

**Corollary 3.** An implication \( I \) satisfying EP and SN satisfies CB.

**Proof.** Straightforward from Propositions 2 and 15  \( \square \)

**Corollary 4.** An implication \( I \) satisfying EP and CO satisfies CB.
Fig. 7 The implication $I_9$

**Proof.** Straightforward from Propositions 3 and 15. □

**Proposition 16. ([1], Lemma 6) An implication $I$ satisfying EP and OP satisfies CB.**

**Proposition 17. There exists an implication $I$ satisfying EP, ID, CP and not CB.**

Counterexample: Let an implication $I_8$ be defined by

$$I_8(x, y) = \begin{cases} 1, & \text{if } x \leq 0.5 \text{ or } y \geq 0.5 \\ 0, & \text{else} \end{cases}, \text{ for all } x, y \in [0, 1].$$

$I_8$ satisfies ID and CP w.r.t. the standard strong negation $N_0$. However, take $x_0 = 1$ and $y_0 = 0.1$, we obtain $I_8(x_0, y_0) = 0 < y_0$. Therefore $I_8$ does not satisfy CB.

**Proposition 18. There exists an implication $I$ satisfying OP, SN, ID, CP, CO and not CB.**

Counterexample: Let an implication $I_9$ be defined by

$$I_9(x, y) = \begin{cases} 1, & \text{if } x \leq y \\ (1 - \sqrt{x-y})^2, & \text{if } x > y \end{cases}, \text{ for all } x, y \in [0, 1].$$

$I_9$ satisfies OP, SN, ID, CP w.r.t. the standard strong negation $N_0$, and CO. However, take $x_0 = 1$ and $y_0 = 0.64$, we obtain $I_9(x_0, y_0) = 0.16 < y_0$. Therefore $I_9$ does not satisfy CB.

So we considered all the possibilities CB can be implied from the other seven properties. Moreover we stated for each independent case a counterexample.
2.6 Getting Identity (ID) from the Other Properties

**Proposition 19.** An implication $I$ satisfying OP satisfies ID.

*Proof.* Straightforward. □

**Proposition 20.** There exists an implication $I$ satisfying NT, EP, SN, CB, CP, CO and not ID.

*Proof.* The Kleene-Dienes implication $I_{KD}(x,y) = \max(1-x,y)$, for all $(x,y) \in [0,1]^2$ is an S-implication generated from the t-conorm $S_M(x,y) = \max(x,y)$ and the standard strong negation $N_0$. Therefore $I_{KD}$ satisfies NT, EP, SN, CB, CP w.r.t. the standard strong negation $N_0$, and CO. However, for $x_0 = 0.1$, we obtain $I_{KD}(x_0,x_0) = 0.9 \neq 1$. Therefore $I_{KD}$ does not satisfy ID. □

So we considered all the possibilities that ID can be implied from the other seven properties, and stated for the independent case a counterexample.

2.7 Getting Contrapositive Principle (CP) from the Other Properties

**Proposition 21.** (Lemma 1(ix)) An implication $I$ satisfying EP and SN satisfies CP w.r.t. the strong negation $N_I$.

**Proposition 22.** (Lemma 1) An implication $I$ satisfying EP, OP and CO satisfies CP w.r.t. the strong negation $N_I$.

**Proposition 23.** There exists an implication $I$ satisfying NT, EP, OP, CB, ID and not CP.

According to the proof of Proposition 10, the Gödel implication $I_{GD}$ satisfies NT, EP, CB, ID. However, for any strong negation $N$ we obtain

$$I_{GD}(N(y), N(x)) = \begin{cases} 1, & \text{if } x \leq y \\ N(x), & \text{if } x > y \end{cases}$$

for all $x, y \in [0,1]$. In case that $x > y$ and $N(x) \neq y$, $I_{GD}(N(y), N(x)) \neq I_{GD}(x, y)$. Therefore $I_{GD}$ does not satisfy CP w.r.t. any strong negation.

**Proposition 24.** There exists an implication $I$ satisfying NT, EP, CB, ID, CO and not CP.

The implication $I_5$ stated in the proof of Proposition 11 satisfies NT, EP, CB, ID and CO. However, because for all $x \in [0,1]$, $N_5(x) = 1 - x^2$, which is not a strong negation, according to Corollary 1.5.5 in [3], $I_5$ does not satisfy CP w.r.t. any strong negation.

**Proposition 25.** There exists an implication $I$ satisfying NT, OP, SN, CB, ID, CO and not CP.
Counterexample: Let an implication $I_{10}$ be defined by

$$I_{10}(x,y) = \begin{cases} 
1, & \text{if } x \leq y \\
\frac{y+(x-y)\sqrt{1-x^2}}{x}, & \text{if } x > y 
\end{cases} \quad \text{for all } x, y \in [0, 1].$$

$I_{10}$ satisfies NT, OP, SN, ID and CO. If $I_{10}$ satisfies CP w.r.t. a strong negation $N$, then for all $x \in [0, 1]$, we obtain

$$N(x) = I_{10}(1,N(x)) = I_{10}(x,0) = N_{I_{10}}(x) = \sqrt{1-x^2}.$$ 

However, take $x_0 = 0.8$ and $y_0 = 0.1$, we obtain $I_{10}(x_0,y_0) = 0.65$ and $I_{10}(N(y_0),N(x_0)) \approx 0.643$. Therefore $I_{10}$ does not satisfy CP w.r.t. any strong negation $N$.

![Fig. 8 The implication $I_{10}$](image)

So we considered all the possibilities that CP can be implied from the other seven properties. Moreover we stated for each independent case a counterexample.

### 2.8 Getting Continuity (CO) from the Other Properties

**Proposition 26.** There exists an implication $I$ satisfying NT, EP, OP, SN, CB, ID, CP and not CO.
Counterexample: Let $N$ be a strong negation. Recall the $R_0$-implication stated in [19] which is defined by

$$(I_{\min_0})_N(x, y) = \begin{cases} 1, & \text{if } x \leq y \\ \max(N(x), y), & \text{if } x > y \end{cases}, \quad \text{for all } x, y \in [0, 1].$$

$(I_{\min_0})_N$ is the R-implication generated by the left-continuous t-norm, nilpotent minimum [9]:

$$(T_{\min_0})_N(x, y) = \begin{cases} \min(x, y), & \text{if } y > N(x) \\ 0, & \text{if } y \leq N(x) \end{cases}, \quad \text{for all } x, y \in [0, 1].$$

$(I_{\min_0})_N$ satisfies NT, EP, OP, SN, CB, ID and CP w.r.t. $N$, and is right-continuous in the second place [19] but it is not continuous.

Therefore CO is independent of any of the other seven properties.

3 A New Class of Implications

In the previous section we have studied dependencies and independencies of eight potential properties for implications, and found different implications satisfying different subgroups of these eight properties, while in these implications $I_6$ and $I_{10}$ actually have the same form. Indeed they can be represented by $I^N$ where $N$ is a negation:

$$I^N(x, y) = \begin{cases} 1, & \text{if } x \leq y \\ \frac{1 - N(x)}{x} + N(x), & \text{if } x > y \end{cases}, \quad \text{for all } x, y \in [0, 1]. \quad (5)$$

If $I^N$ is always an implication, then (5) is an interesting new class of implications because it is only determined by a negation. In this section we check that $I^N$ is always an implication, and then study this new class of implications.

3.1 Is $I^N$ Defined by (5) Always an Implication?

We examine whether the mapping $I^N$ defined by (5) takes its values in $[0, 1]$ and it satisfies properties FI1-FI5. First we rewrite $I^N$ as

$$I^N(x, y) = S_p(N(x), I_{GG}(x, y)), \quad (6)$$

where $I_{GG}$ is the Goguen implication: $I_{GG}(x, y) = \begin{cases} 1, & \text{if } x \leq y \\ y/x, & \text{if } x > y \end{cases}, \quad \text{for all } x, y \in [0, 1], \text{ and } S_p \text{ is the probabilistic sum: } S_p(x, y) = x + y - xy, \quad \text{for all } x, y \in [0, 1]. \text{ It is then straightforward that } I^N(x, y) \in [0, 1], \text{ and } I^N \text{ satisfies F1-F5. Therefore, } I^N \text{ is an implication.}
3.2 The Potential Properties of the New Class of Implications

Now we work out whether \( I^N \) defined by (5) satisfies the potential properties FI6-FI13. If not always, then under which conditions \( I^N \) satisfies these properties.

1. NT: We see immediately from (6) that \( I^N \) always satisfies NT.

2. EP: We obtain the following theorem:

**Theorem 2.** The implication \( I^N \) defined by (5) satisfies EP iff \( N \) belongs to one of the following two classes of negations:

1. \( N_A \) defined by (7),
2. \( N_{A,\beta} \) defined by (2).

**Proof.** Necessity: Suppose \( I^N \) satisfies EP. We will show that if \( N \) is not of the form \( N_A \), \( N \) must be of the form \( N_{A,\beta} \). We will do this in three steps: first we will show that we can find a \( y_0 \) such that \( 0 < N(y_0) < y_0 < 1 \). Second we prove that for \( x \geq y_0 \), \( N(x) = 1 - \frac{x}{1 + \beta x} \) for some fixed \( \beta \). And finally we use this second step to prove that for \( x < y_0 \), \( N(x) = 1 \) or \( N(x) = 1 - \frac{1}{1 + \beta x} \).

Indeed, if \( I^N \) satisfies EP, then for all \( x, y, z \in [0, 1] \), \( I^N(x, I^N(y, z)) = I^N(y, I^N(x, z)) \).

Take \( z = 0 \), we obtain:

\[
(\forall (x, y) \in [0, 1]^2) (I^N(x, N(y)) = I^N(y, N(x))). \tag{7}
\]

Suppose \( N \neq N_A \). Then in particular \( N \neq N_{[0, 1]} \). So there exists a \( y_1 \in [0, 1] \) such that \( N(y_1) < 1 \). Now take \( y_0 \in \max(y_1, N(y_1)) \), then \( N(y_0) \leq N(y_1) < y_0 < 1 \). We first show that \( N(y_0) > 0 \). Indeed, if \( N(y_0) = 0 \), then for all \( x \in [0, 1] \), we obtain:

\[
N(x) = I^N(x, N(y_0)) = I^N(y_0, N(x)) = \begin{cases} 1, & \text{if } y_0 \leq N(x) \\ \frac{N(x)}{y_0}, & \text{if } y_0 > N(x) \end{cases}
\]

\[
\Rightarrow \ N(x) \in \{0, 1\}, \text{ for all } x \in [0, 1] \\
\Rightarrow \ N = N_A, \text{ for a certain } A,
\]

which we have already excluded. Therefore \( N(y_0) > 0 \). For all \( x \in [y_0, 1] \), \( x > N(y_0) \) and \( N(x) < y_0 \). We obtain:

\[
(7) \Rightarrow \frac{1 - N(x)}{x} N(y_0) + N(x) = \frac{1 - N(y_0)}{y_0} N(x) + N(y_0) \\
\Rightarrow \frac{1 - N(x) - x}{x} = \frac{1 - N(y_0) - y_0}{y_0 N(y_0)} N(x)
\]

If \( N(x) = 0 \), then \( \frac{1 - N(x) - x}{x} = 0 \Rightarrow x = 1 \), which we have already excluded. Therefore we obtain:
\[
\frac{1 - N(x) - x}{xN(x)} = \frac{1 - N(y_0) - y_0}{y_0 N(y_0)}
\]
\[
\Rightarrow N(x) = \frac{1 - x}{1 + \beta x} \text{ (with } \beta = \frac{1 - N(y_0) - y_0}{y_0 N(y_0)}, \beta \in ]-1, +\infty[). \]

Now we prove, for any \( x \in [0, y_0] \), that if \( N(x) \neq 1 \), then \( N(x) = \frac{1-x}{1+\beta x} \). In other words that, because \( N \) is decreasing, \( N = N_{A,\beta} \) defined by (2). Indeed, if \( N(x) \neq 1 \), then we can take \( y \) in \( \max(N(x), y_0), 1[ \) such that \( N(y) \leq x \) (this is possible because we have just proved that for \( y \in [y_0, 1[, N(y) = \frac{1-x}{1+\beta y} \)). We obtain:

\[
(7) \Rightarrow \frac{1 - N(x) - x}{xN(x)} = \frac{1 - N(y) - y}{yN(y)} = \beta.
\]

Thus \( N(x) = \frac{1-x}{1+\beta x} \).

Sufficiency of \( N_A \): We obtain: \( I_{N_A}(x,y) = \begin{cases} 1, & \text{if } x \in A \\ I_{GG}(x,y), & \text{if } x \notin A \end{cases} \) for all \( x, y \in [0, 1] \). Thus

\[
I_{N_A}(x, I_{N_A}(y,z)) = \begin{cases} 1, & \text{if } x \in A \text{ or } y \in A \\ I_{GG}(x, I_{GG}(y,z)), & \text{if } x \notin A \text{ and } y \notin A \end{cases} \text{ for all } x, y, z \in [0, 1].
\]

According to (1), \( I_{GG} \) satisfies EP. Therefore \( I_{N_A} \) satisfies EP.

Sufficiency of \( N_{A,\beta}, A = [0, \alpha[ \text{, with } \alpha \in [0, 1], \text{ or } A = [0, \alpha[ \text{, with } \alpha \in [0, 1[ \): We obtain: \( I_{N_A,\beta}(x, y) = \begin{cases} 1, & \text{if } x \leq y \text{ or } x \in A \\ \frac{1-x+(1+\beta y)}{1+\beta x}, & \text{if } x > y \text{ and } x \notin A \end{cases} \) for all \( x, y \in [0, 1] \). Thus for all \( x, y, z \in [0, 1] \):

\[
I_{N_A,\beta}(x, I_{N_A,\beta}(y,z)) = \begin{cases} 1, & \text{if } x \in A \text{ or } y \in A \\ \frac{2+\beta-x-y+\beta xy+(1+\beta)z}{(1+\beta x)(1+\beta y)}, & \text{else} \\ I_{N_A,\beta}(y, I_{N_A,\beta}(x,z)). \end{cases}
\]

(3) OP: We obtain the following theorem:

**Theorem 3.** The implication \( I^N \) defined by (5) satisfies OP iff \( x > 0 \Rightarrow N(x) < 1 \).

**Proof.** This follows from, for all \( 0 \leq y < x \leq 1 \),

\[
I^N(x,y) < 1 \Leftrightarrow \frac{1-N(x)}{x}y + N(x) < 1 \Leftrightarrow 1 - N(x) > 0 \Leftrightarrow N(x) < 1. \]

(4) SN: It is straightforward that \( N_{N}(x) = I^N(x,0) \) is a strong negation iff \( N \) is a strong negation, because \( N_{N} = N \).

(5) CB: Because \( I_{GG} \) satisfies CB, \( I^N \) satisfies CB according to (6).
(6) ID: We see immediately through the definition that \( I^N(x, x) = 1 \) for all \( x \in [0, 1] \).

(7) CP: We obtain the following theorem:

**Theorem 4.** The implication \( I^N \) defined by (5) satisfies CP w.r.t. a strong negation \( N' \) iff \( N \) is a Sugeno negation \( N_a \), \( a \in ]-1, +\infty[ \), and \( N' = N_a \).

**Proof.** Necessity: Recall that \( I^N \) always satisfies NT. If \( I^N \) satisfies CP w.r.t. \( N' \), then according to Proposition 8, \( I^N \) also satisfies SN, and for all \( x \in [0, 1] \), \( N'(x) = I^N(x, 0) = N(x) \). Therefore, \( N \) is strong and \( I^N \) satisfies CP w.r.t. \( N \). We obtain

\[
I^N(N(y), N(x)) = I^N(x, y) \Rightarrow (\forall x \in [0, 1])(\forall y \in [0,x])(\frac{1 - y - N(y)}{N(y)} N(x) = \frac{1 - N(x) - x}{x} y) \\
\Rightarrow (\forall x \in [0, 1])(\forall y \in [0,x])(\frac{1 - y - N(y)}{y N(y)} = \frac{1 - x - N(x)}{x N(x)}) \\
\Rightarrow (\exists a \in [-1, +\infty])(\forall x \in [0, 1])(\frac{1 - x - N(x)}{x N(x)} = a).
\]

If \( a = -1 \) or \( a = +\infty \), then \( N = N_A \) defined in (1) with \( A = [0, 1] \) or \( A = \{0\} \), which is not a strong negation. Thus \( N = N_a \), which is a Sugeno implication.

Sufficiency: If \( N = N_a \), then for all \( x, y \in [0, 1] \):

\[
I^N(x, y) = \begin{cases} 
1, & \text{if } x \leq y \\
\frac{1 + x y - 1 - x}{1 + ax}, & \text{if } x > y
\end{cases}
\]

and

\[
I^N(N(y), N(x)) = \begin{cases} 
1, & \text{if } x \leq y \\
\frac{1 - y}{N(y)} N(x) + y, & \text{if } x > y
\end{cases}
\]

Hence \( I^N(x, y) = I^N(N(y), N(x)) \).  

(8) CO: We obtain the following theorem:

**Theorem 5.** The implication \( I^N \) defined by (5) satisfies CO iff \( N \) is continuous.

**Proof.** It is easily verified that if \( N \) is continuous, \( I^N \) is continuous in each variable. Therefore by Corollary 1.2.2 in [3], \( I^N \) is continuous. The converse follows immediately from \( I^N(x, 0) = N(x) \).

Combining the four theorems in this section and ([1], Theorem 1), we obtain the following two corollaries:

**Corollary 5.** For the implication \( I^N \) defined in (5), the following four conditions are equivalent:

1. \( N \) is a Sugeno negation \( N_a \), \( a \in ]-1, +\infty[ \),
2. \( I^N \) satisfies EP and \( N \) is a continuous negation,
(3) $I^N$ satisfies CP (w.r.t. $N$),
(4) $I^N$ is conjugate with the Łukasiewicz implication $I_L$:

$$I_L(x,y) = \min(1-x+y,1).$$

Notice that if $a = 0$, then $N = N_0$. Then $I^N = I_L$.

**Corollary 6.** An implication $I^N$ defined by (5) satisfying EP and CO also satisfies OP.

The converse of Corollary 6 is not true. For example, the implication $I^{N(0)} = I_{GG}$: $I_{GG}$ satisfies OP but it is not continuous at the point $(0,0)$.

### 3.3 Intersection of the New Class of Implications with the S- and R-Implications

#### 3.3.1 Intersection of the New Class of Implications and S-Implications

In this section we find the intersection of the new class of implications defined in (5) and the class of all S-implications as well as the class of all S-implications generated by a t-conorm and a strong negation.

**Theorem 6.** The implication $I^N$ defined in (5) is an S-implication $S(N'(x),y)$ iff $N = N'$ and $N$ belongs to one of the following two classes of negations:

1. $N_A$ defined by (7) with $A = [0,1[$,
2. $N_{A,\beta}$ defined by (2).

**Proof.** Necessity: Because for all $x \in [0,1]$, $N(x) = I^N(x,0) = S(N'(x),0) = N'(x)$, $N = N'$.

According to (3), Proposition 2.4.6), any S-implication satisfies EP. Then according to Theorem 2, if $I^N$ is an S-implication, then $N = N_{A}, A = [0,\alpha[\ , \text{with } \alpha \in [0,1], \text{or } A = [0,\alpha[\ , \text{with } \alpha \in [0,1], \text{or } N = N_{A,\beta}$. Nevertheless,

$$I^{N_A}(x,y) = \begin{cases} 1, & \text{if } x \leq y \text{ or } x \in A, \\ \frac{y}{x}, & \text{if } x > y \text{ and } x \notin A, \end{cases} \quad \text{for all } x,y \in [0,1],$$

while

$$S(N_A(x),y) = \begin{cases} 1, & \text{if } x \in A, \\ y, & \text{if } x \notin A\ . \end{cases} \quad \text{for all } x,y \in [0,1].$$

If $A \neq [0,1[\ , \text{then we can take } x \text{ and } y \text{ such that } 0 < y < x < 1 \text{ and } x \notin A$. Then $S(N_A(x),y) = y \neq \frac{y}{x} = I^{N_A}(x,y)$. Thus (5) $\neq (9)$ provided $A \neq [0,1[\ . \text{Therefore } I^{N_{[0,\alpha[}}(\alpha < 1) \text{ and } I^{N_{[0,\alpha[}} \text{ are not S-implications.}$

Sufficiency of $N = N_{[0,1[}$: $I^{N_{[0,1[}}(x,y) = S(N_{[0,1[}(x),y)$ for any t-conorm $S$.

Sufficiency of $N = N_{A,\beta}$: Take $S(x,y) = \min(1,x+y+\beta xy)$. We can verify that $S$ is a t-conorm (for the associativity, for all $x,y,z \in [0,1]$):

$$S(x,S(y,z)) = \min(1,x+y+z+\beta xy+\beta yz+\beta xz+\beta^2 xyz) = S(S(x,y),z),$$

and that
Implications in Fuzzy Logic: Properties and a New Class

\[
S(N_A, \beta)(x, y) = \begin{cases} 
1, & \text{if } x \in A \text{ or } x \leq y \\
\frac{1 - x + y + \beta y}{1 + \beta x}, & \text{if } x \notin A \text{ and } x > y
\end{cases} = I^{N_A, \beta}(x, y).
\]

Consequently, \(I^{N_A, \beta}\) is an S-implication. \(\square\)

Combining Corollary 5 and Theorem 6 we obtain the following corollary.

**Corollary 7.** For the implication \(I^N\) defined by (5), the following three conditions are equivalent:

1. \(I^N\) is an S-implication generated by a t-conorm and a strong negation,
2. \(N\) is a Sugeno negation \(N_a\), \(a \in [-1, +\infty[\),
3. \(I^N\) is conjugate with the Łukasiewicz implication \(I_L\).

### 3.4 Intersection of the New Class of Implications and R-Implications

In this section we find the intersection of the new class of implications defined in (5) and the class of the R-implications generated by left-continuous t-norms.

**Theorem 7.** The implication \(I^N\) defined in (5) is an R-implication generated by a left-continuous t-norm if \(N\) belongs to one of the following two classes of negations:

1. a Sugeno negation \(N_a\), \(a \in [-1, +\infty[\),
2. \(N\) defined by (1) with \(A = \{0\}\).

**Proof.** Necessity: If \(I^N\) is an R-implication generated by a left-continuous t-norm, then according to ([8], Theorem 1.14), \(I^N\) satisfies EP and OP. According to Theorem 2, \(N = N_A\) defined by (1), or \(N = N_{A, \beta}\) defined by (2). According to Theorem 3, \(N(x) < 1\) provided \(x > 0\). Therefore \(N = N_a\), or \(N = N_{\{0\}}\).

Sufficiency of \(N = N_a\): According to Corollary 5 if \(N = N_a\), then \(I^N\) is conjugate with \(I_L(x, y) = \max(x + y - 1, 0)\). According to ([1], Theorem 1), \(I^N\) is an R-implication.

Sufficiency of \(N = N_{\{0\}}\): \(I_{N_{\{0\}}} = I_{GG}\), the R-implication generated by the continuous t-norm \(T_P(x, y) = xy\). \(\square\)

Notice that although \(I_{N_{\{0\}}, 1}\) is not an R-implication generated by a left-continuous t-norm, it is the R-implication \(I_{LR}\) generated by the non-left-continuous t-norm \(T_D(x, y) = \begin{cases} \min(x, y), & \text{if } x = 1 \text{ or } y = 1 \\
0, & \text{otherwise}
\end{cases}\) for all \(x, y \in [0, 1]\).

### 3.5 Conclusion Remarks

In this chapter we first studied the complete dependencies and independencies between the eight potential properties FI6-FI13 of implications. Two of the counterexamples we define to show the independencies lead us to a new class of implications determined by a negation \(N\), i.e.,
\[ I^N(x, y) = \begin{cases} 
1, & \text{if } x \leq y \\
\frac{1 - N(x)y}{x} + N(x), & \text{if } x > y,
\end{cases} \text{ for all } x, y \in [0, 1]. \]

We first checked that \( I^N \) is always an implication. For each of the properties FI6-FI13 we found sufficient and necessary conditions in terms of \( N \). We also obtained the intersection of \( I^N \) with any S-implications and R-implications generated by a left-continuous t-norm. An example of \( I^N \) being an R-implication generated by a non-left-continuous t-norm was also given.

It is worth mentioning that if we take \( N \) as a Sugeno negation, then \( I^N \) is an implication that is conjugate with the Łukasiewicz implication \( I_L \).

At the end of this paper we mention a possible generalization of the new class of implications \( I^N \). If we replace \( S_p \) and \( I_{GG} \) in (6) with any \( t \)-conorm \( S \) and any implication \( I \), then we obtain a class of implications defined by \( N, S \) and \( I \):

\[ I^{N,S,I}(x, y) = S(N(x), I(x, y)), \text{ for all } x, y \in [0, 1]. \]

Using the same proof for \( I^N \) we obtain that \( I^{N,S,I} \) is always an implication. This class of implications helps us to generate new implications from existing ones, which will be the topic of further research.

References

On Lattice–Based Fuzzy Rough Sets

Anna Maria Radzikowska

Abstract. Rough sets were originally proposed by Pawlak as a formal tool for analyzing and processing incomplete information represented in data tables. Later on, fuzzy generalizations of rough sets were introduced and investigated to be able to deal with imprecision. In this paper we present $L$–fuzzy rough sets as a further generalization of rough sets. As an underlying algebraic structure we take an extended residuated lattice, that is a residuated lattice endowed with a De Morgan negation. The signature of these structures gives algebraic counterparts of main fuzzy logical connectives. Properties of $L$–fuzzy rough sets are presented. We show that under some conditions families of all lower (resp. upper) $L$–fuzzy rough sets are complete (distributive) lattices. It is also pointed out that in some specific cases lower and upper fuzzy rough approximation operators are $L$–fuzzy topological operators of interior and closure, respectively.

1 Preface

I have enjoyed Etienne Kerre’s acquaintance since the late nineties. In June 1998 I visited Gent for a short–term scientific mission in the framework of the project COST Action 15 “Many–Value Logics for Computer Science Applications. I was impressed and moved by Etienne’s scientific open–mindness and his abilities to combine, in a mastery way, pure mathematics with modern approaches in mathematics of fuzziness and their wide applications in diverse areas. With great enthusiasm he
introduced me into various aspects of fuzzy set theory. We discussed about links between fuzzy set theory and rough set theory, which started our long–lasting research cooperation on fuzzy rough sets, fuzzy modal logics, and fuzzy information relations (see, for example, [41],[5],[22],[26]–[36]). Since then I came to Gent every year in June for 2–3 weeks. Each visit was a very fruitful research experience and, due to Etienne’s openness and kindness, it was also a fantastic, unforgettable stay.

I would like to take the opportunity to thank Etienne and Andrea for their great hospitality during my stays. I fondly remember our lunches, when I enjoyed their guidance on delicious Flemish cuisine. Thank you, Etienne, for being a reference, for your stimulating, challenging ideas, for your real, warm friendship – in summary, for being as you are.

2 Introduction

In real–life problems the available information is usually incomplete and/or imprecise. On one hand, not all relevant information is known or accessible, on the other hand, user’s data may be imprecise or vague (e.g., when expressed by means of linguistic terms like “quite good” or “rather cold”). Rough set theory was developed by Pawlak ([24],[25]) as a formal tool for analyzing and processing information in data tables. As a consequence, it acts on partial information. This formalism proved to be a natural instrument to inquire into many theoretical and practical problems related to data analysis and knowledge discovery (see, e.g., [21],[39],[40],[45],[47]).

Fuzzy set theory ([46]), on the other hand, offers a wide variety of techniques for analyzing imprecise data. Basically, both theories address the problem of information granulation: the theory of fuzzy sets is centered upon fuzzy information granulation, whereas rough set theory is focused on crisp information granulation. Originally, the basic notion in rough set theory was indistinguishability (i.e., indistinguishability between objects in information systems induced by different values of attributes characterizing these objects), yet in recent extensions ([21]) the focus moved to the notion of similarity, which is in fact a fuzzy concept. It is therefore apparent that these two theories have become much closer to each other, so it seems natural to combine methods developed within both theories in order to construct hybrid structures capable to deal with both incompleteness and imprecision. Such structures, called fuzzy rough sets and rough fuzzy sets, were proposed in the literature ([8],[9],[23],[41]). In [27] Radzikowska and Kerre investigated fuzzy rough sets taking the unit interval [0,1] as the basic structure.

In [15] Gougen pointed out that in many situations the linearly ordered set may be insufficient to adequately represent degrees of membership. From this reason, L–fuzzy sets were introduced, where degrees of membership are elements of a lattice L.

As mentioned above, the concept of rough set is recently focused on the notion of similarity. In many situations it is hard to specify whether similarity between objects x and y is stronger (resp. weaker) than similarity between x and z. Hence it is naturally justified to consider L–fuzzy rough sets ([29],[30],[35]). In the present paper
these structures are presented and investigated. Residuated lattices endowed with a
De Morgan negation, referred to as \textit{extended residuated lattices}, are taken as the
underlying algebraic structures. These structures allow us to obtain algebraic counter-
parts of main fuzzy logical connectives: triangular norms and triangular conorms,
\(R\)-implications and \(S\)-implications, and two types of fuzzy negations (including a
De Morgan negation). Using two classes of fuzzy implications the respective two
classes of \(L\)-fuzzy rough sets are investigated. We show that the families of lower
and upper \(L\)-fuzzy rough sets constitute a lattice structure and point out conditions
under which these lattices are distributive. Finally, we point out that under some
conditions \(L\)-fuzzy rough approximation operators are fuzzy interior and fuzzy
closure topological operators.

The present paper contains some important results on fuzzy rough sets that were
obtained during my long cooperation with Etienne, but there are also some studies
that were only a topic of our discussions, but not yet been published. This tribute
book is a nice occasion to supplement our joint research results.

3 Rough Sets

Let \(X\) be a non-empty universe and let \(R\) be a non-empty binary relation on \(X\).
The set \(X\) is viewed as a set of \textit{objects} and \(R\) represents relationships between these
objects determined by their features (attributes). The question addressed in rough
set theory is: \textit{how to approximate a concept} \(A \subseteq X\) \textit{using classes of the relation} \(R\)?

The following notation will be used. Given a subset \(A \subseteq X\), \(\neg A\) will stand for the
set complement of \(A\), i.e., \(\neg A = X \setminus A\). By \(2^X\) we will denote the powerset of \(X\). For
a binary relation \(R\) on \(X\) and for every \(x \in X\), we will write \(xR\) to denote the class of
\(R\) with the representant \(x\), that is the set \(xR = \{y \in X : (x, y) \in R\}\).

Let \(X\) be viewed as a set of objects (examples), and let \(R\) be a binary relation on \(X\)
representing relationships between these objects (indistinguishability or similarity).
The pair \(A_S = (X, R)\) is called an \textit{approximation space}.

Given an approximation space \(A_S = (X, R)\), the following two mappings \(\overline{A}_S, \overline{A}_S : 2^X \rightarrow 2^X\) are defined as: for every \(A \subseteq X\),
\[
\overline{A}_S(A) = \{x \in X : xR \subseteq A\} \quad (1)
\]
\[
\overline{A}_S(A) = \{x \in X : xR \cap A \neq \emptyset\}. \quad (2)
\]

It is easily noted that the operations (1) and (2) coincide with modal operators of \textit{nec-
essity} and \textit{possibility}, respectively (see, e.g., [3],[6]). For this reason, if \(x \in \overline{A}_S(A)\),
then we say that \(x\) \textit{certainly belongs to} \(A\), whereas if \(x \in \overline{A}_S(A)\), then it is said that \(x\)
\textit{possibly belongs to} \(A\).

For any \(A \subseteq X\), \(\overline{A}_S(A)\), and \(\overline{A}_S(A)\) are respectively called a \textit{lower} and an \textit{upper}
rough approximation of \(A\) in \(A_S\). A pair \((A_1, A_2) \in 2^X \times 2^X\) is a rough set in \(A_S\)
iff \(A_1 = \overline{A}_S(A)\) and \(A_2 = \overline{A}_S(A)\) for some \(A \subseteq X\). A subset \(A \subseteq X\) is called \textit{exact}
iff \(\overline{A}_S(A) = A = \overline{A}_S(A)\).
Originally, rough sets were defined with respect to approximation spaces with equivalence relations. Later generalizations of this notion lead to so-called generalized rough sets. In this paper we simply refer to these structures as rough sets.

Let us recall basic properties of rough sets (see, e.g., [6], [21], [24], [25], [45]).

**Theorem 1.** For every approximation space $\mathcal{A} = (X, R)$,

(i) $\mathcal{A}(X) = X$, $\overline{\mathcal{A}}(\emptyset) = \emptyset$

(ii) for all $A, B \subseteq X$, $A \subseteq B$ implies $\overline{\mathcal{A}}(A) \subseteq \overline{\mathcal{A}}(B)$

(iii) for every $A \subseteq X$, $\mathcal{A}(A) = \overline{\mathcal{A}}(\overline{A})$

(iv) for every indexed family $(A_i)_{i \in I}$ of subsets of $X$,

$$\mathcal{A}(\bigcap_{i \in I} A_i) = \bigcap_{i \in I} \mathcal{A}(A_i)$$

$$\overline{\mathcal{A}}(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} \overline{\mathcal{A}}(A_i)$$

$$\mathcal{A}(\bigcup_{i \in I} A_i) \supseteq \bigcup_{i \in I} \mathcal{A}(A_i)$$

$$\overline{\mathcal{A}}(\bigcap_{i \in I} A_i) \subseteq \bigcap_{i \in I} \overline{\mathcal{A}}(A_i).$$

It is well-known ([6]) that modal operators, and in consequence lower and upper rough approximation operators, are useful for characterizing properties of binary relations. The following theorem presents these results.

**Theorem 2. ([6])** For every approximation space $\mathcal{A} = (X, R)$ and for every $A \subseteq X$,

(i) $R$ is serial iff $\mathcal{A}(A) \subseteq \overline{\mathcal{A}}(A)$

(ii) $R$ is reflexive iff $\mathcal{A}(A) \subseteq A$ and $\overline{\mathcal{A}}(\overline{A}) \subseteq \mathcal{A}(A)$

(iii) $R$ is symmetric iff $A \subseteq \mathcal{A}(\overline{\mathcal{A}}(A))$ and $\overline{\mathcal{A}}(\mathcal{A}(A)) \subseteq A$

(iv) $R$ is Euclidean iff $\overline{\mathcal{A}}(A) \subseteq \mathcal{A}(\mathcal{A}(A))$ and $\mathcal{A}(\overline{\mathcal{A}}(A)) \subseteq \overline{\mathcal{A}}(A)$

(v) $R$ is transitive iff $\mathcal{A}(A) \subseteq \mathcal{A}(\mathcal{A}(A))$ and $\overline{\mathcal{A}}(\mathcal{A}(A)) \subseteq \overline{\mathcal{A}}(A)$.

Recall that a relation $R \subseteq X \times X$ is called serial iff for every $x \in X$ there is $y \in X$ such that $(x, y) \in R$; it is called Euclidean iff for all $x, y, z \in X$, $(x, y) \in R$ and $(x, z) \in R$ imply $(y, z) \in R$.

### 4 Algebraic Foundations

Let $L$ be a non-empty domain, let $1 \in L$ be its distinguished element, and let $\otimes$ be a binary operation in $L$. A structure $(L, \otimes, 1)$ is called a monoid iff $\otimes$ is associative operation satisfying $1 \otimes a = a \otimes 1 = 1$ for every $a \in L$. A monoid is called commutative iff $\otimes$ is commutative.
Typical examples of monoid operations are triangular norms and triangular conorms (see \cite{38}). A \textit{triangular norm} \( t \) (resp. a \textit{triangular conorm} \( s \)) is a commutative, associative and non-decreasing in both arguments \([0,1]^2 \rightarrow [0,1]\) mapping satisfying the following border condition \( t(1,a) = a \) (resp. \( s(a,0) = a \)) for every \( a \in [0,1] \). Three most popular triangular norms are:

\begin{itemize}
  \item the min operation (the Zadeh’s triangular norm) \( t_Z(a,b) = \min(a,b) \)
  \item the algebraic product \( t_P(a,b) = a \cdot b \)
  \item the \L ukasiewicz triangular norm \( t_L(a,b) = \max(0,a+b-1) \).
\end{itemize}

A triangular norm is called \textit{left–continuous} iff its partial mappings \( t(a,\cdot) \) and \( t(\cdot,a) \) are left–continuous for every \( a \in [0,1] \).

The well–known triangular conorms are:

\begin{itemize}
  \item the max operation (the Zadeh’s triangular conorm) \( s_Z(a,b) = \max(a,b) \)
  \item the bounded sum \( s_P(a,b) = a+b-a \cdot b \)
  \item the \L ukasiewicz triangular conorm \( s_L(a,b) = \min(1,a+b) \).
\end{itemize}

Triangular norms and triangular conorms are the basis for constructing main classes of \textit{fuzzy implications} \cite{19,20,44}, that is mappings \( i : [0,1]^2 \rightarrow [0,1] \), non–increasing in the first argument, non–decreasing in the second argument, which satisfy the boundary conditions: \( i(1,1) = i(0,0) = i(0,1) = 1 \) and \( i(1,0) = 0 \). There are two main classes of these connectives:

\begin{itemize}
  \item \textit{R–implications}, also called a residual implications based on a left–continuous triangular norm \( t \), or the residuum of \( t \), defined as:
    \[ i_t(x,y) = \sup\{z \in [0,1] : t(x,z) \leq y\} \]
  \item \textit{S–implications} based on a triangular conorm \( s \) and a fuzzy negation \footnote{A \textit{fuzzy negation} is a non–increasing mapping \( n : [0,1] \rightarrow [0,1] \) such that \( n(1) = 0 \) and \( n(0) = 1 \). The standard fuzzy negation \( \eta \) is the mapping \( \eta(x) = 1-x, x \in [0,1] \).} \( n \) defined by: \( i_{s,n}(x,y) = s(n(x),y) \).
\end{itemize}

The most popular R–implications, being the residua of \( t_Z \), \( t_P \) and \( t_L \), respectively, are:

\begin{itemize}
  \item the Gödel implication \( i_Z(x,y) = 1 \) iff \( x \leq y \) and \( i_Z(x,y) = y \) elsewhere
  \item the Gaines implication \( i_P(x,y) = 1 \) iff \( x \leq y \) and \( i_P(x,y) = \frac{x}{y} \) elsewhere
  \item the \L ukasiewicz implication \( i_L(x,y) = \min(1,1-x+y) \).
\end{itemize}

The well–known S–implications are:

\begin{itemize}
  \item the Kleene–Dienes implication \( i_{\kappa,\eta}(x,y) = \max(1-x,y) \) (based on \( s_Z \) and \( \eta \))
  \item the Reichenbach implication \( i_{\kappa,\eta}(x,y) = 1-x+x \cdot y \) (based on \( s_P \) and \( \eta \))
  \item the \L ukasiewicz implication (based on \( s_L \) and \( \eta \))
\end{itemize}

Let \((L,\leq)\) be a poset\footnote{Recall that a \textit{poset} is a structure \((X,\leq)\) such that \( X \neq \emptyset \) and \( \leq \) is a partial order, that is \( \leq \) is reflexive \((x \leq x)\), transitive \((x \leq y \& y \leq z \implies x \leq z)\), and antisymmetric \((x \leq y \& y \leq x \implies x = y)\).} and let \( \circ \) be a unary operation in \( L \). We say that \( \circ \) is \textit{isotone} (order–preserving) iff for all \( a,b \in L \), \( a \leq b \) implies \( \circ a \leq \circ b \); \( \circ \) is called \textit{antitone}
(order–reversing) iff for all \( a, b \in L \), \( a \leq b \) implies \( o b \leq o a \). We say that \( o \) is involution iff \( o o a = a \) for every \( a \in L \).

The following algebra is of main importance, in particular in fuzzy logics (e.g., [7], [10], [11], [14], [16], [43]).

**Definition 1.** A residuated lattice is an algebra \((L, \wedge, \vee, \otimes, \rightarrow, 0, 1)\) such that

1. \((L, \wedge, \vee, 0, 1)\) is a bounded lattice with the top element 1 and the bottom element 0;
2. \((L, \otimes, 1)\) is a commutative monoid;
3. \(\rightarrow\) is a binary operation in \(L\) defined as:
   \[ a \rightarrow b = \sup\{c \in L : a \otimes c \leq b\} \tag{3} \]
   
   The operation \(\otimes\) is called a product and \(\rightarrow\) is its residuum.

Given a residuated lattice \((L, \wedge, \vee, \otimes, \rightarrow, 0, 1)\), the following precomplement operation is defined as: for every \( a \in L \),

\[ \neg a = a \rightarrow 0. \]

The precomplement operation is a generalization of the pseudo–complement in a lattice ([37]). If \(\wedge = \otimes\), then \(\rightarrow\) is the relative pseudo–complement, \(\neg\) is the pseudo–complement, and \((L, \wedge, \vee, \rightarrow, \neg, 0, 1)\) is a Heyting algebra.

For recent results of residuated lattices we refer, for example, to [1], [17], and [18].

**Example 1.** Let \(\otimes\) be a left–continuous t–norm and let \(\rightarrow\) be the residual implication based on \(t\). The structure \(([0, 1], \min, \max, \otimes, \rightarrow, 0, 1)\) is a residuated lattice.

It is easily noted that the product operation of a residuated lattice is the algebraic counterpart of a triangular norm, \(\rightarrow\) corresponds to the residual implication based on \(\otimes\), and \(\neg\) corresponds to a fuzzy negation (in general, it is not involutive). However, residuated lattices do not give adequate counterparts of triangular conorms, fuzzy \(S\)–implications and fuzzy involutive negations. For these reasons, so–called extended residuated lattices were proposed ([29], [32], [35]).

**Definition 2.** An extended residuated lattice (ER–lattice, for short) is a system \((L, \wedge, \vee, \otimes, \rightarrow, \sim, 0, 1)\) such that

1. \((L, \wedge, \vee, \otimes, \rightarrow, \sim, 0, 1)\) is a residuated lattice;
2. \(\sim \sim a = a\) for every \( a \in L\);
3. \(\sim (a \vee b) = \sim a \wedge \sim b\) for all \( a, b \in L\).

From the above definition it immediately follows that an ER–lattice is a residuated lattice endowed with a De Morgan negation.

An ER–lattice \((L, \wedge, \vee, \otimes, \rightarrow, \sim, 0, 1)\) is complete iff the underlying lattice \((L, \wedge, \vee, 0, 1)\) is complete.
Given an ER–lattice \((L, \land, \lor, \otimes, \rightarrow, \sim, 0, 1)\), let us define the following binary operations in \(L\): for all \(a, b \in L\),

\[
a \oplus b = \sim(\sim a \otimes \sim b) \tag{4}
\]

\[
a \Rightarrow b = \sim a \oplus b \tag{5}
\]

One can easily note that the operation \(\sim\) of an ER–lattice corresponds to a fuzzy De Morgan negation, \(\oplus\) and \(\Rightarrow\) are algebraic counterparts of a triangular conorm, and a fuzzy \(S\)–implication, respectively. Therefore, the signature of an ER–lattice gives algebraic counterparts of main fuzzy logical connectives: a triangular norm and a triangular conorm, an \(R\)–implication and an \(S\)–implications, and two fuzzy negations (including a De Morgan negation). Properties of ER–lattices can be found, for example, in [35].

Important properties of ER–lattices are listed in the following two propositions.

**Proposition 1.** Let \((L, \land, \lor, \otimes, \rightarrow, \sim, 0, 1)\) be an ER–lattice and let \(\leq\) be the natural lattice ordering. Then for all \(a, b, c \in L\), the following conditions hold:

1. \(\otimes\) and \(\oplus\) are isotone in both arguments
2. \(\rightarrow\) and \(\Rightarrow\) are antitone in the 1st and isotone in the 2nd argument
3. \(\sim\) and \(\neg\) are antitone
4. \(a \otimes b \leq a \land b, a \lor b \leq a \oplus b\)
5. \(a \otimes 0 = 0, a \oplus 1 = 1\)
6. \(a \rightarrow a = a \rightarrow 1 = 0 \rightarrow a = 1, \ 1 \rightarrow a = a\)
7. \(a \Rightarrow 1 = 0 \Rightarrow a = 1\)
8. \(a \leq b\) iff \(a \rightarrow b = 1\)
9. \(a \otimes (a \rightarrow b) \leq b\)
10. \(a \otimes (b \rightarrow c) \leq b \rightarrow (a \otimes c)\)
11. \((a \rightarrow b) \otimes (b \rightarrow c) \leq (a \rightarrow c)\)
12. \((a \Rightarrow c) \leq (a \Rightarrow b) \oplus (b \Rightarrow c)\)
13. \((a \rightarrow b) \leq (c \rightarrow a) \Rightarrow (c \rightarrow b)\)
14. \(a \rightarrow b \leq (a \otimes c) \rightarrow (b \otimes c)\)
15. \(b \leq a \rightarrow (a \otimes b)\)
16. \(a \rightarrow (b \rightarrow c) = (a \otimes b) \rightarrow c\)
17. \(a \Rightarrow (b \Rightarrow c) = (a \otimes b) \Rightarrow c\)
18. \(a \rightarrow \neg b = \neg(a \otimes b)\)
19. \(a \Rightarrow \sim b = \sim(a \otimes b)\)
20. \(a \rightarrow b \leq \neg b \rightarrow \neg a\)
21. \(a \Rightarrow b = \sim b \Rightarrow \sim a\)
22. \(a \Rightarrow b = \sim(a \otimes b)\)
23. \(a \leq \neg \neg a\).

**Proposition 2.** Let \((L, \land, \lor, \otimes, \rightarrow, \sim, 0, 1)\) be an ER–lattice. Then for every \(a \in L\) and for every family \((b_i)_{i \in I}\) of elements of \(L\), if the respective infima and suprema exist, then the following conditions hold:
Throughout the paper we will use the same symbol $L$ to denote the ER–lattice as well as its underlying domain.

Let $L$ be an ER–lattice and let $X \neq \emptyset$. An $L$–fuzzy set in $X$ is a mapping $F : X \to L$. For any $x \in X$, $F(x)$ is the degree to which $x$ belongs to $F$. The family of all $L$–fuzzy sets in $X$ will be denoted by $\mathcal{F}_L(X)$.

Basic operations on $L$–fuzzy sets in $X$ are defined as follows: for all $A, B \in \mathcal{F}_L(X)$ and for every $x \in X$,

\[
\begin{align*}
(A \cap_L B)(x) &= A(x) \wedge B(x) \\
(A \cup_L B)(x) &= A(x) \vee B(x) \\
(A \oplus_L B)(x) &= A(x) \oplus B(x) \\
(\ominus_L A)(x) &= \ominus A(x), \text{ where } \ominus \in \{\neg, \sim\},
\end{align*}
\]

Given a complete ER–lattice $L$ and an indexed family $(A_i)_{i \in I}$ of $L$–fuzzy sets in $X$, we write $\bigcup_{i \in I} A_i$ and $\bigcap_{i \in I} A_i$ to denote $L$–fuzzy sets in $X$ defined respectively as: for every $x \in X$,

\[
\begin{align*}
(\bigcup_{i \in I} A_i)(x) &= \sup_{i \in I} A_i(x) \\
(\bigcap_{i \in I} A_i)(x) &= \inf_{i \in I} A_i(x).
\end{align*}
\]

For two $L$–fuzzy sets $A$ and $B$ in $X$, we will write $A \subseteq_L B$ iff $A(x) \leq B(x)$ for every $x \in X$.

A binary $L$–fuzzy relation in $X$ is a mapping $X \times X \to L$, i.e., an $L$–fuzzy set in $X \times X$. The family of all binary $L$–fuzzy relations in $X$ will be denoted by $\mathcal{R}_L(X)$.

For a binary $L$–fuzzy relation $R$ in $X$ and for $x \in X$, we will write $xR$ to denote the $L$–fuzzy set in $X$ defined as: $(xR)(y) = R(x, y)$ for every $y \in X$.

A binary $L$–fuzzy relation $R$ is called

- **serial** iff $\sup_{y \in X} R(x, y) = 1$ for every $x \in X$
- **reflexive** iff $R(x, x) = 1$ for every $x \in X$
- **symmetric** iff $R(x, y) = R(y, x)$ for all $x, y \in X$
- **$L$–transitive** iff $R(x, y) \otimes R(y, z) \leq R(x, z)$ for all $x, y, z \in X$
- **$L$–Euclidean** iff $R(x, y) \otimes R(x, z) \leq R(y, z)$ for all $x, y, z \in X$.

A binary $L$–fuzzy relation $R$ in $X$ is a **tolerance** relation iff it is reflexive and symmetric, it is called an **$L$–quasi ordering** iff it is reflexive and $L$–transitive, and it is an **$L$–equivalence** relation iff it is reflexive, symmetric and $L$–transitive.
5 L–Fuzzy Rough Sets

Let \( L \) be a complete \( \mathcal{E} \mathcal{R} \)–lattice, \( X \) be a non–empty universe, and let \( R \in \mathcal{R}_L(X) \). A structure \( \mathcal{FAS} = (L, X, R) \) is called an \( L \)–fuzzy approximation space.

Given an \( L \)–fuzzy approximation space \( (L, X, R) \), define the following operations: for every \( A \in \mathcal{F}_L(X) \) and for every \( x \in X \),

\[
\begin{align*}
\mathcal{FAS}_R(A)(x) &= \inf_{y \in X} (R(x, y) \to A(y)), \\
\mathcal{FAS}_S(A)(x) &= \inf_{y \in X} (R(x, y) \Rightarrow A(y)), \\
\mathcal{FAS}(A)(x) &= \sup_{y \in X} (R(x, y) \otimes A(y)).
\end{align*}
\]

The operation (6) (resp. (7)) is called a lower \( L_R \)–fuzzy (resp. lower \( L_S \)–fuzzy) rough approximation of \( A \) in \( \mathcal{FAS} \), and the operation (8) is an upper \( L \)–fuzzy rough approximation of \( A \) in \( \mathcal{FAS} \). In (6) and (7) the subscripts \( R \) and \( S \) indicate that the underlying arrow operations correspond to an \( R \)–implication and an \( S \)–implication, respectively.

Note that a lower \( L_R \)–fuzzy (resp. lower \( L_S \)–fuzzy) rough approximation operation corresponds to a fuzzy box modal operator and an upper \( L \)–fuzzy rough approximation operation corresponds to a fuzzy diamond modal operator (see, e.g., [12], [13]). In consequence, they have similar intuitive interpretation. Namely, \( \mathcal{FAS}_R(A)(x) \) (resp. \( \mathcal{FAS}_S(A)(x) \)) is the degree to which \( x \) certainly belongs to \( A \), while \( \mathcal{FAS}(A)(x) \) is the degree to which \( x \) possibly belongs to \( A \).

**Definition 3.** Let an \( L \)–fuzzy approximation space \( (L, X, R) \) be given. We say that \( F \in \mathcal{F}_L(X) \) is a lower \( L_R \)–fuzzy (resp. \( L_S \)–fuzzy) rough set in \( \mathcal{FAS} \) iff \( F = \mathcal{FAS}_R(A) \) (resp. \( F = \mathcal{FAS}_S(A) \)) for some \( A \in \mathcal{F}_L(X) \), \( F \) is called an upper \( L \)–fuzzy rough set in \( \mathcal{FAS} \) iff \( F = \mathcal{FAS}(A) \) for some \( A \in \mathcal{F}_L(X) \). By an \( L_R \)–fuzzy rough set in \( \mathcal{FAS} \) (resp. \( L_S \)–fuzzy rough set in \( \mathcal{FAS} \)) we mean a pair \( (F_1, F_2) \in \mathcal{F}_L(X) \times \mathcal{F}_L(X) \) such that \( F_1 = \mathcal{FAS}_R(A) \) and \( F_2 = \mathcal{FAS}(A) \) (resp. \( F_1 = \mathcal{FAS}_S(A) \) and \( F_2 = \mathcal{FAS}(A) \)) for some \( A \in \mathcal{F}_L(X) \).

The family of all lower \( L_R \)–fuzzy (resp. \( L_S \)–fuzzy) rough sets in \( \mathcal{FAS} \) will be respectively written \( \mathcal{L}_R(\mathcal{FAS}) \) and \( \mathcal{L}_S(\mathcal{FAS}) \) and the family of all upper \( L \)–fuzzy rough sets in \( \mathcal{FAS} \) will be denoted by \( \mathcal{U}(\mathcal{FAS}) \).

Basic properties of \( L \)–fuzzy rough sets are listed below.

**Theorem 3.** ([30], [34], [35]) For every \( L \)–fuzzy approximation space \( (L, X, R) \), for all \( A, B \in \mathcal{F}_L(X) \), and for every indexed family \( (A_i)_{i \in I} \) of \( L \)–fuzzy sets in \( X \), the following conditions hold:

(i) \( \mathcal{FAS}_R(X) = \mathcal{FAS}_S(X) = X, \mathcal{FAS}(\emptyset) = \emptyset \)

(ii) if \( A \subseteq L B \), then

\[
\begin{align*}
\mathcal{FAS}_R(A) &\subseteq L \mathcal{FAS}_R(B), \quad \mathcal{FAS}_S(A) \subseteq L \mathcal{FAS}_S(B) \\
\mathcal{FAS}(A) &\subseteq L \mathcal{FAS}(B)
\end{align*}
\]

(iii) \( \mathcal{FAS}_R(A) \subseteq L \neg_L \mathcal{FAS}(\neg_L A) \)

\[
\mathcal{FAS}(A) \subseteq L \neg_L \mathcal{FAS}_R(\neg L A)
\]
\[
\text{FAS}_S(A) = \sim_L \text{FAS}(\sim_L A)
\]
\[
\text{FAS}(A) = \sim_L \text{FAS}_S(\sim_L A).
\]

(iv) \( \text{FAS}_S(\cap_{i \in I} A_i) = \cap_{i \in I} \text{FAS}_S(A_i), \quad \text{FAS}_S(\cap_{i \in I} A_i) = \cap_{i \in I} \text{FAS}_S(A_i) \)
\[
\text{FAS}_R(\cap_{i \in I} A_i) \subseteq_R \cap_{i \in I} \text{FAS}_R(A_i), \quad \text{FAS}_R(\cap_{i \in I} A_i) \subseteq_R \cap_{i \in I} \text{FAS}_R(A_i)
\]
\[
\text{FAS}(\cap_{i \in I} A_i) \subseteq_L \cap_{i \in I} \text{FAS}(A_i)
\]
\[
\text{FAS}(\cup_{i \in I} A_i) \subseteq_L \cap_{i \in I} \text{FAS}(A_i)
\]

The property (ii) of the above theorem states the monotonicity of \(L\)-fuzzy approximation operators and (iii) states the duality (resp. weak duality) between \(\text{FAS}_S\) (resp. \(\text{FAS}_R\)) and \(\text{FAS}\). Observe that for \(L_S\)-fuzzy rough sets all properties mentioned above correspond to the ones in the crisp case (see Theorem 1). For \(L_R\)-fuzzy rough sets, however, we have a weaker form of duality only.

Depending on classes of \(L\)-fuzzy relations \(L\)-fuzzy rough sets have more specific properties.

**Theorem 4.** Let \((L, X, R)\) be an \(L\)-fuzzy approximation space.

(i) \(R\) is serial \(\iff\) \(\text{FAS}_R(A) \subseteq L \text{FAS}(A)\)
\(\iff\) \(\text{FAS}_S(A) \subseteq L \text{FAS}(A)\)

(ii) \(R\) is reflexive \(\iff\) \(\text{FAS}_R(A) \subseteq L A\)
\(\iff\) \(\text{FAS}_S(A) \subseteq L A\)
\(\iff\) \(A \subseteq L \text{FAS}(A)\)

(iii) \(R\) is symmetric \(\iff\) \(\text{FAS}(\text{FAS}_R(A)) \subseteq L A\)
\(\iff\) \(A \subseteq L \text{FAS}_R(\text{FAS}(A))\)

(iv) \(R\) is \(L\)-transitive \(\iff\) \(\text{FAS}_R(A) \subseteq L \text{FAS}_R(\text{FAS}_R(A))\)
\(\iff\) \(\text{FAS}_S(A) \subseteq L \text{FAS}_S(\text{FAS}_S(A))\)
\(\iff\) \(\text{FAS}(\text{FAS}(A)) \subseteq L \text{FAS}(A)\)

(v) \(R\) is \(L\)-Euclidean \(\iff\) \(\text{FAS}(A) \subseteq L \text{FAS}_R(\text{FAS}(A))\)
\(\iff\) \(\text{FAS}(\text{FAS}_R(A)) \subseteq L \text{FAS}_R(A)\).

**Proof.** By way of example we show (ii) and (iv).

(ii) \(\Rightarrow\) Assume that \(R\) is reflexive. Then for every \(A \in \mathcal{F}_L(X)\) and for every \(x \in X\),

\[
\text{FAS}_R(A)(x) = \inf_{y \in X} (R(x, y) \rightarrow A(y))
\]
\[
\leq R(x, x) \rightarrow A(x) = 1 \rightarrow A(x) = A(x).
\]

Hence \(\text{FAS}_R(A) \subseteq L A\). Also,

\[
\text{FAS}_S(A)(x) = \inf_{y \in X} (R(x, y) \Rightarrow A(y))
\]
\[
\leq R(x, x) \Rightarrow A(x) = \sim 1 \oplus A(x) = 0 \oplus A(x) = A(x),
\]

so \(\text{FAS}_S(A) \subseteq L A\). Furthermore,

\[
\text{FAS}(A)(x) = \sup_{y \in X} (R(x, y) \otimes A(y))
\]
\[
\geq R(x, x) \otimes A(x) = 1 \otimes A(x) = A(x),
\]

where \(\geq = \leq^{-1}\). Therefore, \(A \subseteq L \text{FAS}(A)\).
Assume that \( R \) is not reflexive, i.e., \( R(x_0, x_0) \neq 1 \) for some \( x_0 \in X \). For \( A = x_0R \), we have:

\[
FAS_R(A)(x_0) = \inf_{y \in X} (R(x_0, y) \rightarrow A(y))
\]

\[
= \inf_{y \in X} (R(x_0, y) \rightarrow R(x_0, y))
\]

\[
= 1 \quad \text{by Proposition} \, \,(\text{vii}).
\]

However, \( A(x_0) = R(x_0, x_0) < 1 \). Hence \( FAS_R(A) \not\subseteq L \, A \).

Next, let us take \( A = \{x_0\} \). Then we have:

\[
FAS_S(A)(x_0) = \inf_{y \in X} (R(x_0, y) \implies A(y))
\]

\[
= \inf_{y \in X} (\sim R(x_0, y) \oplus A(y))
\]

\[
= \inf_{y \in X} (\sim (\sim R(x_0, y) \otimes \sim A(y)))
\]

\[
= \inf_{y \in X} ((R(x_0, y) \otimes \sim A(y))
\]

\[
= \sim R(x_0, x_0)
\]

\[
> 0 = A(x_0)
\]

so \( FAS_S(A) \not\subseteq L \, A \).

Finally, for \( A = \{x_0\} \), we have:

\[
FAS(A)(x_0) = \sup_{y \in X} (R(x_0, y) \otimes A(y))
\]

\[
= R(x_0, x_0) \otimes A(x_0) = R(x_0, x_0) \otimes 1 = R(x_0, x_0) < 1 = A(x_0),
\]

thus \( A \not\subseteq L \, FAS(A) \).

(iv) \( \Rightarrow \) Assume that \( R \) is \( L \)-transitive. Then for every \( A \in \mathcal{F}_L(X) \) and for every \( x \in X \),

\[
FAS_R(FAS_R(A))(x)
\]

\[
= \inf_{y \in X} (R(x, y) \rightarrow (\inf_{z \in X} (R(y, z) \rightarrow A(z))))
\]

\[
= \inf_{y \in X} \inf_{z \in X} (R(x, y) \rightarrow (R(y, z) \rightarrow A(z)))
\]

\[
= \inf_{y \in X} \inf_{z \in X} (R(x, y) \otimes R(y, z) \rightarrow A(z))
\]

\[
\geq \inf_{z \in X} (R(x, z) \rightarrow A(z))
\]

\[
= FAS_R(A)(x).
\]

Hence \( FAS_R(A) \subseteq L \, FAS_R(FAS_R(A)) \).

In the similar way one can show that \( FAS_S(A) \subseteq L \, FAS_S(FAS_S(A)) \). Also,

\[
FAS(FAS(A))(x)
\]

\[
= \sup_{y \in X} (R(x, y) \otimes \sup_{z \in X} (R(y, z) \otimes A(z)))
\]

\[
= \sup_{y \in X} \sup_{z \in X} (R(x, y) \otimes R(y, z) \otimes A(z))
\]

\[
\leq \sup_{z \in X} (R(x, z) \otimes A(z))
\]

\[
= FAS(A)(x),
\]

thus \( FAS(FAS(A)) \subseteq L \, FAS(A) \).

\( \Leftarrow \) Assume that \( R \) is not \( L \)-transitive, that is \( R(x_0, y_0) \otimes R(y_0, z_0) \not\subseteq R(x_0, z_0) \) for some \( x_0, y_0, z_0 \in X \). By Proposition(\text{ix}), \( (R(x_0, y_0) \otimes R(y_0, z_0)) \rightarrow R(x_0, z_0) \not\subseteq 1 \) for
some $x_0, y_0, z_0 \in X$. Let us take $A = x_0R$. By Proposition [Pvii] we get $FAS_R(A)(x_0) = \inf_{y \in X} (R(x_0,y) \rightarrow R(x_0,y)) = 1$. Furthermore,

$$FAS_R(A)(x_0) \rightarrow FAS_R(FAS_R(A))(x_0)$$

$$= 1 \rightarrow FAS_R(FAS_R(A))(x_0) \quad \text{by Proposition [Pvii]}$$

$$= \inf_{y \in X} (R(x_0,y) \rightarrow \inf_{z \in X} (R(y,z) \rightarrow R(x_0,z)))$$

$$= \inf_{y \in X} \inf_{z \in X} (R(x_0,y) \rightarrow (R(y,z) \rightarrow R(x_0,z))) \quad \text{by Proposition [Pii]}$$

$$= \inf_{y \in X} \inf_{z \in X} (R(x_0,y) \otimes R(y,z) \rightarrow R(x_0,z)) \quad \text{by Proposition [Pvi]}$$

$$\leq (R(x_0,y_0) \otimes R(y_0,z_0)) \rightarrow R(x_0,z_0)) \neq 1,$$

so again, by Proposition [Pix], $FAS_R(A)(x_0) \not\subseteq FAS_R(FAS_R(A))(x_0)$. Hence we have $FAS_R(A) \not\subseteq FAS_R(FAS_R(A))$, as required.

Now we show that for $A = X \setminus \{z_0\}$ it holds $FAS_S(A) \not\subseteq FAS_S(FAS_S(A))$. Note that $R(x_0,y_0) \otimes R(y_0,z_0) \not\subseteq R(x_0,z_0)$ implies $\sim R(x_0,z_0) \not\subseteq \sim (R(x_0,y_0) \otimes R(y_0,z_0))$, since $\sim$ is order-reversing. Then, by Proposition [Piv],

$$\sim R(x_0,z_0) \rightarrow \sim (R(x_0,y_0) \otimes R(y_0,z_0)) \neq 1. \quad (9)$$

Also, for every $y \in X$, we have

$$\text{FAS}_S(A)(y) = \inf_{z \in X} (R(y,z) \Rightarrow A(z))$$

$$= \inf_{z \in X} \sim (R(y,z) \otimes \sim A(z)) \quad \text{by Proposition [Pxi]}$$

$$= \sim \sup_{z \in X} (R(y,z) \otimes \sim A(z)) \quad \text{by Proposition [Piv]}$$

$$= \sim R(y,z_0).$$

Now we have:

$$FAS_S(A)(x_0) \rightarrow FAS_S(FAS_S(A))(x_0)$$

$$= \sim R(x_0,z_0) \rightarrow \inf_{y \in X} (R(x_0,y) \Rightarrow FAS_S(A)(y))$$

$$= \sim R(x_0,z_0) \rightarrow \inf_{y \in X} (R(x_0,y) \Rightarrow \sim R(y,z_0))$$

$$= \sim R(x_0,z_0) \rightarrow \inf_{y \in X} \sim (R(x_0,y) \otimes \sim R(y,z_0)) \quad \text{by Proposition [Pxi]}$$

$$= \sim R(x_0,z_0) \rightarrow \sim \sup_{y \in X} (R(x_0,y) \otimes R(y,z_0)) \quad \text{by Proposition [Pvi]}$$

$$\leq \sim R(x_0,z_0) \rightarrow \sim (R(x_0,y_0) \otimes R(y_0,z_0)) \neq 1 \quad \text{by (9)}.$$

Hence, by Proposition [Pix], $FAS_S(A)(x_0) \not\subseteq FAS_S(FAS_S(A))(x_0)$, which immediately implies $FAS_S(A) \not\subseteq FAS_S(FAS_S(A))$, as expected.

Finally, take $A = \{z_0\}$. Observe that $FAS(A)(y) = \sup_{z \in X} (R(y,z) \otimes A(z)) = R(y,z_0)$ for every $y \in X$. Now,

$$\text{FAS}(\text{FAS}(A))(x_0) \rightarrow \text{FAS}(A)(x_0)$$

$$= \sup_{y \in X} (R(x_0,y) \otimes \text{FAS}(A)(y)) \rightarrow \text{FAS}(A)(x_0)$$

$$= \sup_{y \in X} (R(x_0,y) \otimes R(y,z_0)) \rightarrow R(x_0,z_0)$$

$$\leq (R(x_0,y_0) \otimes R(y_0,z_0)) \rightarrow R(x_0,z_0) \quad \text{by Proposition [Pii]}$$

$$\neq 1,$$

so $FAS(\text{FAS}(A))(x_0) \not\subseteq FAS(A)(x_0)$ by Proposition [Pix], thus $\text{FAS}(\text{FAS}(A)) \not\subseteq FAS(A)$, as required. \(\square\)
It is worth emphasizing that all properties mentioned in Theorem 4 hold for lower $L_R$–fuzzy rough approximation operators. As stated in Theorem 2, they characterize rough approximation operators. Then these operations are adequate generalizations of their crisp counterparts. However, in the general case properties (iii) and (v) of the above theorem do not hold for lower $L_S$–fuzzy rough approximations, as the following example shows.

Example 2. Let us take a lattice $(L, \min, \max, t_z, l_z, \eta, 0, 1)$ with $L = [0, 1]$. Then $\oplus$ is given by $s_Z(x, y) = \max(x, y)$. Let $X = \{x_1, x_2, x_3\}$ and let $R$ be defined as: $R(x_1, x_1) = R(x_2, x_2) = 1$ and $R(x_1, x_2) = \frac{1}{2}$ elsewhere. Clearly, $R$ is symmetric. Consider $FAS = (L, X, R)$ and $A = \{x_3\}$. By simple calculations we get $FAS(FAS_S(A)) = FAS_S(FAS(A)) = (x_1 : \frac{1}{2}, x_2 : \frac{1}{2}, x_3 : \frac{1}{2})$. Obviously, neither $FAS(FAS_S(A)) \subseteq A$ nor $A \subseteq FAS_S(FAS(A))$ holds. Hence, for lower $L_S$–fuzzy rough approximation operation, the property (iii) of Theorem 4 does not hold.

Also, it is easy to verify that $R$ is an $L$–Euclidean relation. Since $FAS_S(A) = (x_1 : 0, x_2 : \frac{1}{2}, x_3 : \frac{1}{2})$ and $FAS(A) = (x_1 : \frac{1}{2}, x_2 : \frac{1}{2}, x_3 : 1)$, neither $FAS(A) \subseteq FAS_S(FAS(A))$ nor $FAS(FAS_S(A)) \subseteq FAS_S(A)$.

As mentioned in Theorem 3, lower $L_S$–fuzzy rough approximation operators are dual to upper $L$–fuzzy rough approximation operators, which in general is not the case when lower $L_R$–fuzzy and upper $L$–fuzzy rough approximations are taken. Recall that the Łukasiewicz implication is both an $R$–implication and an $S$–implication. For such operators all properties mentioned in Theorem 3 and 4 are satisfied. Therefore, lattice–based fuzzy rough sets, where such operators are taken, are straightforward generalizations of Pawlak’s rough sets.

From the above theorem the following corollary easily follows.

**Corollary 1.** Let $(L, X, R)$ be an $L$–fuzzy approximation space. Then for every $A \in \mathcal{P}_L(X)$,

(i) if $R$ is an $L$–quasi ordering, then

- $FAS_R(FAS_R(A)) = FAS_R(A)$
- $FAS_S(FAS_S(A)) = FAS_S(A)$
- $FAS(FAS(A)) = FAS(A)$.

(ii) if $R$ is a tolerance relation, then

- $FAS_R(A) \subseteq_L FAS(FAS_R(A)) \subseteq_L A$
- $A \subseteq_L FAS_R(FAS(A)) \subseteq_L FAS(A)$

(iii) if $R$ is an $L$–equivalence relation, then

- $FAS(FAS_R(A)) = FAS_R(A)$
- $FAS_R(FAS(A)) = FAS(A)$.

Therefore, in the case of a tolerance relation $FAS FAS_R$ and $FAS_R FAS$ give a tighter approximation of $A$ than $FAS_R$ and $FAS$, respectively.

We conclude this section by the following observation. In classical settings it is well-known that box and diamond modal operators coincide with interior and
closure topological operators, respectively, provided that the underlying relation is a quasi-ordering, i.e., it is reflexive and transitive. The analogous result holds in the fuzzy case.

First, let us recall (e.g., [2],[34]) the definitions of $L$–fuzzy topological operators of interior and closure. Let $\text{int}_L, \text{cl}_L : \mathcal{P}_L(X) \to \mathcal{P}_L(X)$. We say that $\text{int}_L$ is an $L$–fuzzy interior operator iff

\begin{align*}
(\text{I1}) & \quad \text{int}_L(X) = X \\
(\text{I2}) & \quad \text{int}_L(A) \subseteq_L A \text{ for every } A \in \mathcal{P}_L(X) \\
(\text{I3}) & \quad \text{int}_L(\text{int}_L(A)) = \text{int}_L(A) \text{ for every } A \in \mathcal{P}_L(X) \\
(\text{I4}) & \quad \text{int}_L(A \cap_L B) = \text{int}_L(A) \cap_L \text{int}_L(B).
\end{align*}

We say that $\text{cl}_L$ is an $L$–fuzzy closure operator iff

\begin{align*}
(\text{C1}) & \quad \text{cl}_L(\emptyset) = \emptyset \\
(\text{C2}) & \quad A \subseteq_L \text{cl}_L(A) \text{ for every } A \in \mathcal{P}_L(X) \\
(\text{C3}) & \quad \text{cl}_L(\text{cl}_L(A)) = \text{cl}_L(A) \text{ for every } A \in \mathcal{P}_L(X) \\
(\text{C4}) & \quad \text{cl}_L(A \cup_L B) = \text{cl}_L(A) \cup_L \text{cl}_L(B).
\end{align*}

From Theorem [3] and Corollary [11] we immediately get the following result.

**Theorem 5.** Let $\text{FAS} = (L,X,R)$ be an $L$–fuzzy approximation space such that $R$ is an $L$–quasi ordering. Then $\text{FAS}_R$ and $\text{FAS}_S(A)$ are $L$–fuzzy interior operators and $\text{FAS}$ is an $L$–fuzzy closure operator.

### 6 Lattices of Lower and Upper $L$–Fuzzy Rough Sets

Let an $L$–fuzzy approximation space $(L,X,R)$ be given. In this section we will consider algebraic structures for lower $L_R$–fuzzy rough sets, lower $L_S$–fuzzy rough sets, and upper $L$–fuzzy rough sets.

Let $(F_i)_{i \in I}$ be the family of lower $L_R$–fuzzy rough sets (or lower $L_S$–fuzzy rough sets). By Theorem [3] it follows that $\bigcap_{i \in I} F_i \in \mathcal{L}_R(\text{FAS})$ and $\bigcap_{i \in I} F_i \in \mathcal{L}_S(\text{FAS})$, yet in general neither $\bigcup_{i \in I} F_i \in \mathcal{L}_R(\text{FAS})$ nor $\bigcup_{i \in I} F_i \in \mathcal{L}_S(\text{FAS})$, as the following example shows.

**Example 3.** Let $X = \{x_1, x_2, x_3\}$ and let $R$ be a crisp relation given by: $R = \{(x_1, x_1), (x_1, x_2), (x_2, x_1), (x_2, x_3), (x_3, x_2), (x_3, x_3)\}$. Let $(L, \min, \max, t, i, n, 0, 1)$ be an extended residuated lattice such that where $L = [0, 1]$, $t$ is an arbitrary left–continuous triangular norm, $i$ is the residual implication based on $t$, and $n$ is the standard fuzzy negation $n(x) = 1 - x$. Let $\text{FAS} = (L,X,R)$. Consider two fuzzy sets $A$ and $B$ defined as $A = (x_1 : 0.1, x_2 : 0.8, x_3 : 0.3)$ and $B = (x_1 : 0.9, x_2 : 0.4, x_3 : 0.2)$, respectively. By simple calculations one can easily verify that $\text{FAS}_R(A) = \text{FAS}_S(A) = (x_1 : 0.1, x_2 : 0.1, x_3 : 0.3)$ and $\text{FAS}_R(B) = (x_1 : 0.4, x_2 : 0.2, x_3 : 0.2)$. Put $F = \text{FAS}_R(A) \cup_L \text{FAS}_R(B)$, or equivalently, $F = \text{FAS}_R(A) \cup_L \text{FAS}_R(B)$. Then $F = (x_1 : 0.4, x_2 : 0.2, x_3 : 0.3)$. Assume that $F \in \mathcal{L}(\text{FAS})$. Then there exists $C \in \mathcal{P}(X)$ such that $F = \text{FAS}_R(C)$. However,
(a) \( F(x_1) = \min\{1 \rightarrow C(x_1), 1 \rightarrow C(x_2), 0 \rightarrow C(x_3)\} = \min\{C(x_1), C(x_2)\} = 0.4 \),
(b) \( F(x_2) = \min\{1 \rightarrow C(x_1), 0 \rightarrow C(x_2), 1 \rightarrow C(x_3)\} = \min\{C(x_1), C(x_3)\} = 0.2 \),
(c) \( F(x_3) = \min\{0 \rightarrow C(x_1), 1 \rightarrow C(x_2), 1 \rightarrow C(x_3)\} = \min\{C(x_2), C(x_3)\} = 0.3 \).

From (a), \( C(x_1) \geq 0.4 \), so by (b), \( C(x_3) = 0.2 \), which contradicts (c). Hence, for any \( C \in \mathcal{F}_L(X), F \not\in \mathcal{F}_R^{\text{AS}}(C) \), which means that \( F \not\in \mathcal{L}_R^{\text{AS}}(FAS) \). Using similar arguments one can easily check that for every \( C \in \mathcal{F}_L(X), F \not\in \mathcal{L}_R^{\text{AS}}(FAS) \).

Similarly, for an arbitrary family \( (H_k)_{k \in K} \) of upper \( L \)-fuzzy rough sets it holds, \( \bigcup_{k \in K} H_k \in \Omega(FAS) \), but not necessarily \( \bigcap_{k \in K} H_k \in \Omega(FAS) \).

Let \( (F_i)_{i \in I} \) be the family of lower \( L_R \)-fuzzy rough sets, let \( (G_j)_{j \in J} \) be the family of lower \( L_S \)-fuzzy rough set, and let \( (H_k)_{k \in K} \) be the family of upper \( L \)-fuzzy rough sets. Consider the sets

\[
\begin{align*}
\mathcal{U}_R^{\text{AS}}(FAS, \bigcup_{i \in I} F_i) &= \{ F \in \mathcal{L}_R^{\text{AS}}(FAS) : \bigcup_{i \in I} F_i \subseteq_L F \} \\
\mathcal{U}_S^{\text{AS}}(FAS, \bigcup_{j \in J} G_j) &= \{ G \in \mathcal{L}_S^{\text{AS}}(FAS) : \bigcup_{j \in J} G_j \subseteq_L G \} \\
\mathcal{L}^{\text{AS}}(FAS, \bigcap_{k \in K} H_k) &= \{ H \in \Omega(FAS) : H \subseteq_L \bigcap_{k \in K} H_k \}.
\end{align*}
\]

The set \( \mathcal{U}_R^{\text{AS}}(FAS, \bigcup_{i \in I} F_i) \) is the family of all upper bounds of \( \bigcup_{i \in I} F_i \) in \( \mathcal{L}_R^{\text{AS}}(FAS) \) w.r.t. \( \subseteq_L \), the set \( \mathcal{U}_S^{\text{AS}}(FAS, \bigcup_{j \in J} G_j) \) is the family of all upper bounds of \( \bigcup_{j \in J} G_j \) in \( \mathcal{L}_S^{\text{AS}}(FAS) \) w.r.t. \( \subseteq_L \), and the set \( \mathcal{L}^{\text{AS}}(FAS, \bigcap_{k \in K} H_k) \) is the family of all lower bounds of \( \bigcap_{k \in K} H_k \) in \( \Omega(FAS) \) w.r.t. \( \subseteq_L \). Put

\[
\begin{align*}
F^*(FAS, \bigcup_{i \in I} F_i) &= \bigcap \mathcal{U}_R(FAS, \bigcup_{i \in I} F_i) \\
G^*(FAS, \bigcup_{j \in J} G_j) &= \bigcap \mathcal{U}_S(FAS, \bigcup_{j \in J} G_j) \\
H^*(FAS, \bigcap_{k \in K} H_k) &= \bigcup \mathcal{L}(FAS, \bigcap_{k \in K} H_k).
\end{align*}
\]

By Proposition 3 iv), \( F^*(FAS, \bigcup_{i \in I} F_i) \in \mathcal{L}_R^{\text{AS}}(FAS) \) and \( G^*(FAS, \bigcup_{j \in J} G_j) \in \mathcal{L}_S^{\text{AS}}(FAS) \). Clearly, \( \bigcup_{i \in I} F_i \subseteq_L F^*(FAS, \bigcup_{i \in I} F_i) \) and \( \bigcup_{j \in J} G_j \subseteq_L G^*(FAS, \bigcup_{j \in J} G_j) \). Therefore, \( F^*(FAS, \bigcup_{i \in I} F_i) \in \mathcal{U}_R^{\text{AS}}(FAS, \bigcup_{i \in I} F_i) \) and \( G^*(FAS, \bigcup_{j \in J} G_j) \in \mathcal{U}_S^{\text{AS}}(FAS, \bigcup_{j \in J} G_j) \). Also, \( F^*(FAS, \bigcup_{i \in I} F_i) \subseteq_L F \) for any \( F \in \mathcal{U}_R^{\text{AS}}(FAS, \bigcup_{i \in I} F_i) \) and \( G^*(FAS, \bigcup_{j \in J} G_j) \subseteq_L G \) for any \( G \in \mathcal{U}_S^{\text{AS}}(FAS, \bigcup_{j \in J} G_j) \). Hence, \( F^*(FAS, \bigcup_{i \in I} F_i) \) is the least upper bound of \( \bigcup_{i \in I} F_i \) in \( \mathcal{L}_R^{\text{AS}}(FAS) \) w.r.t. \( \subseteq_L \) and \( G^*(FAS, \bigcup_{j \in J} G_j) \) is the least lower bound of \( \bigcup_{j \in J} G_j \) in \( \mathcal{L}_S^{\text{AS}}(FAS) \) w.r.t. \( \subseteq_L \). Analogously, \( H^*(FAS, \bigcap_{k \in K} H_k) \in \Omega(FAS) \), \( H^*(FAS, \bigcap_{k \in K} H_k) \subseteq_L \bigcap_{k \in K} H_k \), and \( H \in \bigcup H^*(FAS, \bigcap_{k \in K} H_k) \) for every \( H \in \bigcup H(FAS, \bigcap_{k \in K} H_k) \). Therefore, \( H^*(FAS, \bigcap_{k \in K} H_k) \) is the greatest lower bound of \( \bigcap_{k \in K} H_k \) in \( \Omega(FAS) \) w.r.t. \( \subseteq_L \).

For all \( F_1, F_2 \in \mathcal{L}_R^{\text{AS}}(FAS) \), for all \( G_1, G_2 \in \mathcal{L}_S^{\text{AS}}(FAS) \), and for all \( H_1, H_2 \in \Omega(FAS) \), let us denote

\[
\begin{align*}
F_1 \Delta_R F_2 &= F_1 \cap_L F_2 \\
F_1 \Vee_R F_2 &= \bigcap \mathcal{U}_R^{\text{AS}}(FAS, \{ F_1 \cup_L F_2 \}) \\
G_1 \Delta_S G_2 &= G_1 \cap_L G_2 \\
G_1 \Vee_S G_2 &= \bigcap \mathcal{U}_S^{\text{AS}}(FAS, \{ G_1 \cup_L G_2 \}) \\
H_1 \frown H_2 &= \bigcup \mathcal{L}^{\text{AS}}(FAS, \{ H_1 \cap_L H_2 \}) \\
H_1 \frown H_2 &= H_1 \cup_L H_2.
\end{align*}
\]
From the above discussion we get the following

**Theorem 6.** The algebras \((\mathcal{L}_R(\text{FAS}), \Delta_R, \vee_R, \emptyset, X)\), \((\mathcal{L}_S(\text{FAS}), \Delta_S, \vee_S, \emptyset, X)\), and \((\mathcal{U}(\text{FAS}), \wedge, \vee, \emptyset, X)\) are complete lattices.

Let an \(L\)-fuzzy approximation space \(\text{FAS} = (L, X, R)\) be such that \(R\) is an \(L\)-quasi ordering. Assume that \(F_1, F_2 \in \mathcal{L}_R(\text{FAS})\). Then there are two \(L\)-fuzzy sets, say \(A_1\) and \(A_2\), such that \(F_1 = FAS_R(A_1)\) and \(F_2 = FAS_R(A_2)\). So we have:

\[
F_1 \cup_L F_2 = FAS_R(A_1) \cup_L FAS_R(A_2) \\
= FAS_R(FAS_R(A_1)) \cup_L FAS_R(FAS_R(A_2)) \quad \text{by Corollary [H1]}
\]

\[
\subseteq L FAS_R(FAS_R(A_1)) \cup_L FAS_R(FAS_R(A_2)) \quad \text{by Proposition [IV]}
\]

\[
= FAS_R(F_1 \cup_L F_2) \\
\subseteq L F_1 \cup_L F_2 \quad \text{by Proposition [III].}
\]

Hence \(FAS_R(F_1 \cup_L F_2) = F_1 \cup_L F_2\). Also, \(F_1 \cup_L F_2 \in \mathcal{U}L(R(\text{FAS}, F_1 \cup_L F_2))\). Clearly, \(F_1 \cup_L F_2 = \bigcap \mathcal{U}L(R(\text{FAS}, F_1 \cup_L F_2))\), so \(F_1 \vee_R F_2 = F_1 \cup_L F_2\). Since \(F_1 \Delta_R F_2 = F_1 \cap_L F_2\), we get

\[
F_1 \Delta_R (F_2 \vee_R F_3) = (F_1 \Delta_R F_2) \vee_R (F_1 \Delta_R F_3) \\
F_1 \vee_R (F_2 \Delta_R F_3) = (F_1 \vee_R F_2) \Delta_R (F_1 \vee_R F_3)
\]

for all \(F_1, F_2, F_3 \in \mathcal{L}_R(\text{FAS})\). Using similar arguments it is easy to show that for all \(G_1, G_2, G_3 \in \mathcal{L}_S(\text{FAS})\),

\[
G_1 \Delta_S (G_2 \vee_S G_3) = (G_1 \Delta_S G_2) \vee_S (G_1 \Delta_S G_3) \\
G_1 \vee_S (G_2 \Delta_S G_3) = (G_1 \vee_S G_2) \Delta_S (G_1 \vee_S G_3).
\]

Similarly, for all \(H_1, H_2, H_3 \in \mathcal{U}(\text{FAS})\),

\[
H_1 \Delta (H_2 \vee H_3) = (H_1 \Delta H_2) \vee (H_1 \Delta H_3) \\
H_1 \vee (H_2 \Delta H_3) = (H_1 \vee H_2) \Delta (H_1 \vee H_3).
\]

Therefore, we get the following

**Proposition 3.** Let \(\text{FAS} = (L, X, R)\) be an \(L\)-fuzzy approximation space such that \(R\) is an \(L\)-quasi ordering. Then the structures \((\mathcal{L}_R(\text{FAS}), \subseteq_L)\), \((\mathcal{L}_S(\text{FAS}), \subseteq_L)\), and \((\mathcal{U}(\text{FAS}), \subseteq_L)\) are complete, distributive lattices.

We complete this section by showing some link between the lattices of \(\mathcal{L}_S(\text{FAS})\) and \(\mathcal{U}(\text{FAS})\). To this end, let us define the following notion (see [42]). We say that two lattices \((X, \leq_X)\) and \((Y, \leq_Y)\) are **dually isomorphic** iff there is a bijection \(\phi : X \rightarrow Y\) such that for all \(x_2, x_2 \in X, x_1 \leq_X x_2\) iff \(\phi(x_2) \leq_Y \phi(x_1)\).

**Proposition 4.** For every \(L\)-fuzzy approximation space \(\text{FAS} = (L, X, R)\), the lattices \((\mathcal{L}_S(\text{FAS}), \subseteq_L)\) and \((\mathcal{U}(\text{FAS}), \subseteq_L)\) are dually isomorphic.
Proof. Let \( F \in \mathcal{L}_S(\text{FAS}) \). Put \( \Gamma(F) = \{ A \in \mathcal{L}_S(\text{FAS}) : \text{FAS}_S(A) = F \} \) and define a mapping \( \phi : \mathcal{L}_S(\text{FAS}) \rightarrow \mathcal{U}(\text{FAS}) \) by: for every \( F \in \mathcal{L}_S(\text{FAS}) \),

\[
\phi(F) = \overline{\text{FAS}}(\sim_L \cap \Gamma(F)).
\]

Note that for every \( F \in \mathcal{L}_S(\text{FAS}) \) we have:

\[
\phi(F) = \overline{\text{FAS}}(\sim_L \cap \Gamma(F)) \\
= \sim_L \text{FAS}_S(\cap \Gamma(F)) \quad \text{by Theorem 3(iii)} \\
= \sim_L F \quad \text{by Theorem 3(iv)}.
\]

We show that \( \phi \) is a bijection. Let \( F_1, F_2 \in \mathcal{L}_S(\text{FAS}) \) be such that \( F_1 \neq F_2 \). Since \( \sim \) is an involution, \( \sim_L F_1 \neq \sim_L F_2 \). Hence \( \phi(F_1) \neq \phi(F_2) \). Now take an arbitrary \( G \in \mathcal{U}(\text{FAS}) \). Then for some \( A \in \mathcal{P}_L(X) \), \( G = \overline{\text{FAS}}(A) \). Therefore, by Theorem 3(iii),

\[
G = \sim_L \text{FAS}_S(\sim_L A).
\]

Whence there is \( F \in \mathcal{L}_S(\text{FAS}) \), namely \( F = \overline{\text{FAS}}_S(\sim_L A) \), such that \( G = \sim_L F = \phi(F) \).

Finally, let us take \( F_1, F_2 \in \mathcal{L}_S(\text{FAS}) \). Then we have the following equivalences:

\[
F_1 \subseteq_L F_2 \iff \sim_L F_2 \subseteq_L \sim_L F_1 \iff \phi(F_2) \subseteq_L \phi(F_1),
\]

which completes the proof. \( \square \)

7 Conclusions

In this paper we have presented an algebraic extension of fuzzy rough sets. As an underlying algebraic structure a residuated lattice endowed with a De Morgan negation has been taken. The signature of these structures gives algebraic counterparts of main fuzzy logical connectives: triangular norms and triangular conorms, \( R \)–implication and \( S \)–implications, and two fuzzy negations, including a De Morgan negation. Taking either of arrow operations we obtain two classes of lower \( L \)–fuzzy rough sets, namely lower \( L_R \)–fuzzy rough sets and lower \( L_S \)–fuzzy rough sets. Moreover, depending on a choice of a class of binary \( L \)–fuzzy relations, we get further respective classes of \( L \)–fuzzy rough sets. Properties of these structures have been studied. As it turned out, for \( L_R \)–fuzzy rough sets most properties of crisp rough sets hold (except duality). Since the \( \text{Łukasiewicz} \) implication is both an \( R \)–implication and an \( S \)–implication, for such operators all properties of \( L \)–fuzzy rough sets are straightforward generalizations of Pawlak’s rough sets. We have proved that the family of lower (resp. upper) \( L \)–fuzzy rough sets is a lattice w.r.t. \( L \)–fuzzy inclusion \( \subseteq_L \), and, under specific conditions, they are distributive lattices. It was also shown that the family of lower \( L \)–fuzzy (resp. \( L_S \)–fuzzy) rough sets are dually isomorphic to the family of upper \( L \)–fuzzy rough sets. Finally, it has been pointed out that under some assumptions lower and upper \( L \)–fuzzy rough approximation operations are \( L \)–fuzzy interior and \( L \)–fuzzy closure operators, respectively.

It is well–known that the family of all (crisp) rough sets defined on the basis of an equivalence relation, constitute a regular double Stone algebra. The question is what kind of algebraic structure \( L \)–fuzzy rough sets are. Since I deeply hope that the research cooperation with Etienne will be going on, it may be a challenging topic for further joint studies.
References


Graduality, Uncertainty and Typicality in Formal Concept Analysis

Yassine Djouadi, Didier Dubois, and Henri Prade

Abstract. There exist several proposals for extending formal concept analysis (FCA) to fuzzy settings. They focus mainly on mathematical aspects and assume generally a residuated algebra in order to maintain the required algebraic properties for the definition of formal concepts. However, less efforts have been devoted for discussing what are the possible reasons for introducing degrees in the relation linking objects and properties (which defines a formal context in the FCA sense), and thus what are the possible meanings of the degrees and how to handle them in agreement with their intended semantics. The paper investigates three different semantics, namely i) the graduality of the link associating properties to objects, pointing out various interpretations of a fuzzy formal context; ii) the uncertainty pervading this link (in case of binary properties) when only imperfect information is available and represented in the framework of possibility theory; and lastly, iii) the typicality of objects and the importance of definitional properties within a class. Remarkably enough, the uncertainty semantics has been hardly considered in the FCA setting, and the third semantics apparently not. Moreover, we provide an algorithm for building the whole fuzzy concept lattice based on Gödel implication for handling gradual properties in a qualitative manner.

1 Introduction

Formal concept analysis (FCA) was independently introduced by Wille in the 1980’s but its mathematical basis had been pioneered in the setting

Yassine Djouadi
University of Tizi-Ouzou, BP 17, RP, 15000 Tizi-Ouzou, Algeria
e-mail: ydjouadi@mail.ummto.dz

Didier Dubois · Henri Prade
IRIT, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France
e-mail: dubois@irit.fr, prade@irit.fr
of social sciences by Barbut and Monjardet [1] (see also [21]). Since then, it has become increasingly popular among various methods of conceptual data analysis and knowledge representation. It is formulated by means of the notion of a formal context, which is a triple \((\mathcal{O}, \mathcal{P}, \mathcal{R})\), where \(\mathcal{O}\) is a set of objects, \(\mathcal{P}\) is a set of properties and \(\mathcal{R}\) specifies a binary relation between the sets \(\mathcal{O}\) and \(\mathcal{P}\) \((\mathcal{R} \subseteq \mathcal{O} \times \mathcal{P})\). In the classical setting, it is always assumed that for any pair \((o, p)\): i) \(o \mathcal{R} p\) means that it is well-known that the object \(o\) satisfies the property \(p\), ii) \(o \mathcal{R} \bar{p}\) means that it is well-known that the object \(o\) does not satisfy the property \(p\), iii) exclusively one of the two above cases applies.

During the last years, FCA has been applied in many different areas like psychology, sociology, anthropology, medicine, biology, linguistics, etc. In such cases, FCA unavoidably deals with relational information structures (formal contexts) derived from human investigation (judgement, observation, measure, etc.). However, it is widely agreed that such information is often of a gradual nature and possibly tainted with uncertainty, while FCA has been almost exclusively developed in a Boolean setting.

FCA constitutes a particular instance of successful application of the calculus of relations, and can be formally extended to fuzzy relations. As such, it may be paralleled with other applications of fuzzy relational calculus [22], ranging from approximate reasoning [10, 17] to image processing [23].

Several fuzzy extensions of formal concept analysis have nevertheless been addressed in the literature, however to different extents. A central notion in existing approaches is the so-called fuzzy formal context whose entries become now degrees from a totally ordered scale \(L\) (generally \(L=[0,1]\)), whereby property satisfaction becomes a matter of degree. Since the first paper by Burusco and Fuentes-Gonzalez [8], different fuzzy logic settings have been proposed, especially by Bělohlávek (see [2, 3] for a survey, [18] for a fuzzy inference perspective). In a different direction, Burmeister and Holzer [7] have made a proposal that introduces a third value in the relational context that stands for the case where it is not known if a property holds or not for an object. Messai et al. [25] handle many-valued attribute domains based on a fuzzy logic encoding.

The different fuzzy extensions of FCA that have been proposed are mainly justified by the requirement of preserving good mathematical (algebraic) properties (i.e. closure and opening properties), without always sufficiently discussing what the extension is good for from a knowledge representation point of view. In particular, one may introduce degrees in a relational context, with many different intended meanings. For instance, Table 1 illustrates different generalizations of a formal context which may naturally appear. Entries in the first column of Table 1 express a satisfaction degree of the gradual property Young. Most existing approaches deal with this kind of formal context [2, 3, 8, 9, 20, 24, 26]. Entries in the second column give the proficiency level of a given student in the English course. This level may be precisely known (e.g. like for Pierre and Mike) or just approximated using an interval
Table 1 Different formal context generalizations.

<table>
<thead>
<tr>
<th>( R_1 )</th>
<th>Young</th>
<th>English Proficiency</th>
<th>Married</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierre</td>
<td>1</td>
<td>0.9</td>
<td>( \times )</td>
</tr>
<tr>
<td>Sophie</td>
<td>0.7</td>
<td>( (0,1) )</td>
<td>?</td>
</tr>
<tr>
<td>Mike</td>
<td>0.6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nahla</td>
<td>1</td>
<td>( [0.2,0.4] )</td>
<td>( (0.7; 0) )</td>
</tr>
<tr>
<td>Peter</td>
<td>0.2</td>
<td>( [0,1] )</td>
<td>( (0; 0) )</td>
</tr>
</tbody>
</table>

(e.g., like for Sophie, Nahla, and Peter). An interesting feature of the use of intervals is that they provide the ability to encode partial/total ignorance \[12, 13\]. For instance, Peter’s proficiency level in English is totally unknown whereas we just know that it is not true that Sophie doesn’t at all master English.

The third column of Table 1 illustrates incompleteness and uncertainty. For instance, as it is well known that Pierre is married and Mike is not. However, nothing may be stated for Sophie, which is represented using a question mark. In the presence of total ignorance, Burmeister and Holzer \[7\] have proposed to consider the “?” symbol as third value. In the possibility theory setting, both partial and total ignorance may be taken into account. For instance, the pair \( (0.7 ; 0) \) indicates that the certainty of Nahla being married is 0 (no certainty at all), and the certainty that she is not is 0.7, whereas the pair \( (0; 0) \) illustrates total ignorance about Peter and corresponds to the question mark “?”.

In a preliminary version of this paper \[11\], we have distinguished different possible situations leading to the introduction of degrees in fuzzy formal contexts. We have also provided a preliminary discussion of a possibilistic handling of uncertainty, and introduced the idea of taking into account the typicality of objects and the importance of properties in the setting of formal concept analysis. In this paper, we enlarge the discussion about these three motivations and define different graded extensions of formal concept analysis: accounting for gradual properties, handling uncertainty, and acknowledging typicality. The paper is structured as follows. After a background on fuzzy extensions of formal concept analysis in Section 2 and a discussion of suitable operators, Section 3 both provides an introductory discussion of different situations where graduality is encountered, and an algorithm for building the whole fuzzy concept lattice based on Gödel implication for handling gradual properties in a qualitative manner. Section 4 presents the possibilistic view of uncertain formal concepts, while Section 5 illustrates the potential interest of tolerating missing properties that are not compulsory for non-typical objects in a formal concept.
2 Fuzzy Extensions of Formal Concept Analysis

In FCA, a formal context $K = (O, P, R)$ is represented by a Boolean table where rows correspond to objects and columns to properties (or conversely), and a table entry contains “$\times$” or “blank mark”, depending on whether the object has or not the corresponding property. Let $R(o) = \{ p \in P \mid oRp \}$ be the set of properties satisfied by object $o$. Similarly, let $R^{-1}(p) = \{ o \in O \mid oRp \}$ be the set of objects that satisfy the property $p$. We can define powerset operators $(.)^\uparrow : 2^O \to 2^P$ and $(.)^\downarrow : 2^P \to 2^O$ as:

$$O^\uparrow = \{ p \in P \mid \forall o \in O : (o \in O \Rightarrow oRp) \} = \{ p \in P \mid R^{-1}(p) \supseteq O \}$$

$$P^\downarrow = \{ o \in O \mid \forall p \in P : (p \in P \Rightarrow oRp) \} = \{ o \in O \mid R(o) \supseteq P \}$$

The formal concept analysis problem is that of extracting formal concepts from object/property relations. A formal concept is a pair $\langle O, P \rangle$, such that $O = \{ o \in O \mid R(o) \supseteq P \}$ and $P = \{ p \in P \mid R^{-1}(p) \supseteq O \}$. Thus, $P$ is the set of properties shared by all the objects in $O$, and $O$ is the set of objects that possess all the properties in $P$. The set $O$ (resp. $P$) is called extent (resp. intent). It is easy to remark that both extents (resp. intents) are fixed points w.r.t. to the composition $(.)^\uparrow \downarrow$ (resp. $(.)^\downarrow \uparrow$). Following Birkhoff’s result [5], the set of all formal concepts (denoted $L(K)$) is a complete lattice. A formal concept can be also equivalently defined as a maximal pair $(O, P)$ (in the sense of set inclusion) that satisfies the condition $O \times P \subseteq R$.

Fuzzy extensions of FCA deal mainly with a fuzzy relation $R \in L^O \times P$ which is a mapping: $O \times P \to L$. Whereas both operators $(.)^\uparrow : L^O \to L^P$ and $(.)^\downarrow : L^P \to L^O$ become now based on a fuzzy implication $I$ as defined below [8]:

$$O^\uparrow (p) = \bigwedge_{o \in O} \left( I(O(o), R(o, p)) \right) \quad P^\downarrow (o) = \bigwedge_{p \in P} \left( I(P(p), R(o, p)) \right)$$

However, the development of a fuzzy formal concept analysis theory requires an appropriate algebra of fuzzy sets [3] in order to maintain the closure property for the composition of the above defined derivation operators. This property, recalled in the following, turns out to be crucial for the soundness of the theory.

**Definition 1.** Given a universe $U$, a mapping $\Phi : L^U \to L^U$ is a fuzzy closure operator iff $\forall U, V \in L^U$ it satisfies:

(CL1): $U \subseteq V \Rightarrow \Phi(U) \subseteq \Phi(V)$ (isotone)

(CL2): $U \subseteq \Phi(U)$ (extensive)

(CL3): $\Phi(\Phi(U)) = \Phi(U)$ (idempotent)

In order to satisfy the above property, the existing approaches use residuated algebras. Let us recall that a residuated algebra [27], is an algebra $L = (L, \wedge, \vee, *, \to)$ s.t. the pair $(*, \to)$ satisfies the adjointness property
(namely, \( p \leq q \rightarrow r \) iff \( p \ast q \leq r \)). For maintaining the closure property of the compositions \((\cdot)^\downarrow\uparrow\) and \((\cdot)^\uparrow\downarrow\) Theorem 1 proposes a sufficient condition (R1), weaker than the residuation principle. Consequently this theorem allows us to use more general fuzzy algebras \( L = (L, \land, \lor, \rightarrow, \ast, \sim) \), since requirement (R1) does not refer to the conjunction \( \ast \) nor the negation \( \sim \). Let us recall that a fuzzy implication operator (\( \rightarrow \)) is decreasing (in the broad sense) in its first component and increasing (in the broad sense) in its second component and verifies boundaries conditions \( (0 \rightarrow 0 = 0 \rightarrow 1 = 1 \rightarrow 1 = 1 \) and \( 1 \rightarrow 0 = 0 \)). Whereas, a fuzzy conjunction (\( \ast \)) is a binary increasing operator (in the broad sense) which verifies identity condition \( (p \ast 1 = 1 \ast p = p) \) and boundaries conditions \( (0 \ast 0 = 0 \ast 1 = 1 \ast 0 = 0) \). Note that \( \ast \) is not necessarily a t-norm.

**Theorem 1.** (Djouadi and Prade [13]). Given an algebra \( L = (L, \land, \lor, \rightarrow) \) with a fuzzy implication \( \rightarrow \), the composition \((\cdot)^\downarrow\uparrow\) (symmetrically \((\cdot)^\uparrow\downarrow\)) is a fuzzy closure operator if the following property is satisfied \( \forall p, q \in L: \)

\[
(R1) \quad p \leq (p \rightarrow q) \rightarrow q
\]

It is noticeable that Theorem 1 does not restrict the set of admissible implications to the set of residuated implications. The following counter-example shows a tri-valued implication (\( \rightarrow \)) which satisfies the above condition (R1) although it is not residuated.

**Example 1.** Let us consider the tri-valued implication (denoted here \( \rightarrow \)) defined in \( \{0, a, 1\} \) as shown in Table 2:

\[
\begin{array}{c|c|c|c}
\rightarrow & 0 & a & 1 \\
0 & 1 & 1 & 1 \\
a & a & a & 1 \\
1 & 0 & 0 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
(p \rightarrow q) \rightarrow q & 0 & a & 1 \\
0 & 0 & 0 & 1 \\
a & a & a & 1 \\
1 & 1 & 1 & 1 \\
\end{array}
\]

It is easy to check that the implication \( \rightarrow \) verifies the condition (R1). However, this implication is not residuated, as now checked. Indeed, let us try to prove one half of the residuation principle (namely, \( p \ast q \leq r \Rightarrow p \leq q \rightarrow r \)). Since the (fuzzy) conjunction \( \ast \) verifies the boundary condition \( p \ast 1 = p \), by substituting \( p \) by 1 and \( q \) and \( r \) by \( a \) we get:

\[
1 \ast a \leq a \Rightarrow 1 \leq a \rightarrow a
\]

\[
\iff a \leq a \Rightarrow 1 \leq a
\]

Thus, it is obvious that the half \( (\Rightarrow) \) of the residuation principle does not hold. Then, the implication \( \rightarrow \) is not residuated.
Remark 1. Admissible algebras of the form \( L = (L, \wedge, \vee, \rightarrow, *, \sim) \) are not restricted to algebras with pairs \((\rightarrow, *)\) s.t. \(\rightarrow\) is the residuum of \(*\). For instance, a fuzzy algebra equipped with Gödel implication \((p \rightarrow q = 1 \text{ if } p \leq q, q \text{ elsewhere})\) may involve any conjunction \(*\) rather than the min \(t-\)norm.

We give now an example of implication that satisfies the condition (R1), and which as such is suitable for the definition of \((_)\uparrow\downarrow\) (symmetrically \((_)\downarrow\uparrow\)).

Example 2. Let us consider a fuzzy algebra \( L = (L, \wedge, \vee, \rightarrow, *, \sim) \) where the implication \( \rightarrow \) is defined as:

\[
(p \rightarrow q) = \begin{cases} 
1 & \text{if } p \leq q \\
(1 - p) \vee q & \text{elsewhere}
\end{cases}
\]

This implication is remarkable since it can be obtained both as the \(\vee\)-based disjunction of Gödel and its contrapositive, and as \(\sim (p \otimes \sim q)\), where \(\otimes\) is the nilpotent minimum and \(\sim\) is an involutive negation \((p \otimes q = 0 \text{ if } p + q \leq 1, p \wedge q \text{ if } p + q > 1)\). Note that the pair \((\rightarrow, *)\) is a residuated pair when taking the conjunction \(* = \otimes\), namely the nilpotent minimum which is a non-continuous \(t-\)norm. In any case, this implication is admissible according to Theorem 1 since property \(p \leq (p \rightarrow q) \rightarrow q\) is verified.

Remark 2. Note that condition (R1) is satisfied in a residuated structure, when \(*\) is commutative as claimed in the following proposition.

Proposition 1. Condition (R1) is satisfied by any residuated algebra if the conjunction \(*\) is commutative.

Proof. We have to prove that: \(\langle p \leq q \rightarrow r \iff p \ast q \leq r, * \text{ commutative } \rangle \Rightarrow p \leq (p \rightarrow q) \rightarrow q\).

Since \(*\) is commutative, we have: \(p \leq q \rightarrow r \iff p \ast q \leq r \iff q \ast p \leq r \iff q \leq p \rightarrow r\).

Since the previous equivalences hold \(\forall p, q, r \in L\), assuming \(q = p \rightarrow q\) and \(r = q\), we get condition (R1):

\[
p \leq (p \rightarrow q) \rightarrow q \iff p \rightarrow q \leq p \rightarrow q
\]

The following example illustrates the case where a particular implication, residuated w.r.t. a non commutative conjunction, does not satisfy the condition (R1).

Example 3. The Dienes implication (i.e. \(p \leftrightarrow q = (1 - p) \vee q\)) is a remarkable implication since it forms a residuated pair with the non commutative conjunction \(*\) (defined as: \(p \ast q = p \text{ if } p + q > 1; 0 \text{ elsewhere}\)) as we shall prove in the following. Let us first determine the left member of the equivalence \(p \leq q \leftrightarrow r \iff p \ast q \leq r:\)
Graduality, Uncertainty and Typicality in Formal Concept Analysis

\[ p \leq q \rightarrow r \]
\[ \iff p \leq (1 - q) \lor r \]
\[ \iff \begin{cases} p \leq r & \text{if } p + q > 1 \\ \top & \text{if } p + q \leq 1 \end{cases} \tag{4} \]

Let us now determine the right member:
\[ p \ast q \leq r \]
\[ \iff \begin{cases} p \leq r & \text{if } p + q > 1 \\ 0 \leq r & \text{if } p + q \leq 1 \end{cases} \tag{5} \]

Thus, it is easy to state that Expression 4 is equivalent to Expression 5. However, this fuzzy implication does not satisfy the condition (R1). Indeed, taking \( p = 0.8 \) and \( q = 0.7 \), we have \((p \rightarrow q) \leftrightarrow q = ((1 - 0.8) \lor 0.7) \rightarrow 0.7 \neq p = 0.8\). However, as shown by Example 2, it can be mixed with the requirement that \( p \rightarrow q = 1 \) when \( p \leq q \) into a suitable implication.

**Remark 3.** Gödel, Goguen and Lukasiewicz implications are admissible implications since they are residuated with respect to a commutative continuous triangular norm.

Assuming a sound algebra w.r.t. the closure property, fuzzy FCA consists then of extracting pairs of \( L \)-sets \( \langle O, P \rangle \) from a fuzzy context \( K = (L, O, P, R) \), where \( L \)-sets \( O \) and \( P \) determine each other pointwisely, being s.t. \( O(o) = P^\downarrow(o), \forall o \in O \) and \( P(p) = O^\uparrow(p), \forall p \in P \). The set of all formal \( L \)-concepts is also a complete lattice.

### 3 Gradual Link between a Property and an Object

This section highlights first different gradual interpretations of a fuzzy formal context. We emphasize the representation aspect and the underlying issues of achieving a well-suited representation for many valued contexts (conceptual scaling vs. fuzzy properties). Lastly, we take advantage of both semantical and computational aspects of Gödel implication and provide an algorithm for building the whole concept lattice.

#### 3.1 Unipolar vs. Bipolar Scale Interpretation

A “fuzzy” or graded extension of binary formal contexts may convey different semantics. In a first interpretation, the values in the table (which are scalars in \( L \)) may be understood as providing a refinement of the cross marks. Namely, they represent to which extent an object has a property, while in the classical model, this relationship was not a matter of degree. It is important to remark that in this view, we do not refine the absence of a property for an object (the blank mark is always replaced by the bottom element \( 0 \) of \( L \)). This view
Table 3 Many valued relation

<table>
<thead>
<tr>
<th>( \mathcal{R}_2 )</th>
<th>Pierre</th>
<th>Sophie</th>
<th>Mike</th>
<th>Nahla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>22</td>
<td>28</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>Salary</td>
<td>1100</td>
<td>1300</td>
<td>1500</td>
<td>1500</td>
</tr>
</tbody>
</table>

Table 4 Context subsumption

<table>
<thead>
<tr>
<th>( \mathcal{R}_3 )</th>
<th>Pierre</th>
<th>Sophie</th>
<th>Mike</th>
<th>Nahla</th>
</tr>
</thead>
<tbody>
<tr>
<td>age ( \geq 20 )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
</tr>
<tr>
<td>age ( \geq 25 )</td>
<td>( \times )</td>
<td>( \times )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>age ( \geq 30 )</td>
<td></td>
<td></td>
<td>( \times )</td>
<td></td>
</tr>
<tr>
<td>salaire ( \geq 1000 )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
</tr>
<tr>
<td>salaire ( \geq 1200 )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td></td>
</tr>
<tr>
<td>salaire ( \geq 1400 )</td>
<td></td>
<td>( \times )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

will be referred to as the positive unipolar interpretation. In this interpretation, \( R^{-1}(p) \) (resp. \( R(o) \)) is considered as the support of the fuzzy set of objects (resp. properties) satisfying the property \( p \) (resp. the object \( o \)). One could also consider the opposite convention namely the negative unipolar interpretation where degrees would represent to which extent an object does not have a property and equivalently provide a refinement of the blank marks.

The most commonly used interpretations, through existing FCA proposals, are implicitly based on the positive unipolar interpretation that allows to map a formal context with quantitative attributes into a fuzzy formal context. In this spirit, conceptual scale theory \[29\] may be used to achieve a suited (Boolean) representation by successive subsumptions. For instance, the formal context illustrated in Table 4 is obtained from Table 3 by a conceptual scaling of both many-valued attributes “Age” and “Salary”. As can be seen, we have two sets of properties with obvious subsumption relations between them. Pairs (\{Peter, Sophie, Mike, Joe\}, \{age \( \geq 20 \), salaire \( \geq 1000 \}\)), (\{Sophie, Mike\}, \{age \( \geq 20 \), age \( \geq 25 \), salaire \( \geq 1000 \), salaire \( \geq 1200 \}\)), or (\{Mike\}, \{age \( \geq 20 \), age \( \geq 25 \), age \( \geq 30 \), salaire \( \geq 1000 \), salaire \( \geq 1200 \), salaire \( \geq 1400 \}\}) are formal concepts.

Knowing the ages and the salaries, the formal context \( \mathcal{R}_3 \) can be re-encoded in a more compact way, using two fuzzy sets ‘young’ and ‘small’ with decreasing membership functions, as illustrated in Table 5. Observe also that \( \mathcal{R}_4 \) offers a more precise representation of initial data than Table 4. The context in Table 5, event tough more compact than Table 4 highlights the fact that Mike, and to a lesser extent Sophie are not very young and have a salary that is not really low. It constitutes in some sense the negative of the picture shown on Table 3. Note that the type of representation on Table 5 can be obtained even without providing interpretable fuzzy sets.
Table 5 Context summarization

<table>
<thead>
<tr>
<th>( \mathcal{R}_4 )</th>
<th>Pierre</th>
<th>Sophie</th>
<th>Mike</th>
<th>Nahla</th>
</tr>
</thead>
<tbody>
<tr>
<td>age ‘young’</td>
<td>1</td>
<td>0.7</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>salary ‘low’</td>
<td>1</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

and thus, by normalizing in \( L \) the domain of attribute values. This approach is used in \([25]\).

Another interpretation of the degrees, maybe more in the standard spirit of fuzzy logic would be to replace both the cross marks and the blank marks by values in the scale \( L (L = [0, 1]) \). Then \( L \) possesses a mid-point acting as a pivoting value between the situations where the object possesses the property to some extent and the converse situation where the object possesses the opposite property to some extent. Under this view, a fuzzy formal concept should be learnt together with its negation. This view corresponds to a bipolar scale interpretation.

3.2 Concept Lattice Based on Gödel Implication

From a semantic point of view, the use of Gödel implication, which is at the basis of the expression of a graduality in fuzzy FCA, can also refine fuzzy contexts into sub-contexts for which a gradual rule \([16]\) is satisfied. Thus, in the above example the pair: \( (\{\text{Pierre, Sophie, Mike, Nahla}\}, \{\text{young}_{0.6}, \text{low}_{0.6}\}) \) is a formal concept. Note however that Nahla violates the rule “the more the person is young the more the salary is low”, which is however completely satisfied for Pierre, Sophie and Mike: indeed, \( \mathcal{I}_G(\text{young}, \text{low}) = 1 \) (where \( \mathcal{I}_G \) denotes Gödel implication) if the extension of this concept is limited to Pierre, Sophie and Mike. The search for fuzzy concepts where a gradual rule is satisfied can be a topic of interest.

Computing the whole concept lattice (i.e. the set of all \( L \)-closed elements) remains the main objective of FCA. However, it is noticeable that the set of all fixed points is infinite under an arbitrary implication and computing this whole set is a non envisageable task. Some approaches \([4], [30]\) compute only the closure of a given \( L \)-set \( O \in L^O \) (or \( P \in L^P \)). Another approach \([9]\) assumes a finite scale \( L \) with cardinality \( k \) in order to compute the whole (finite) set of closed elements of \( L^O \). The complexity is thus reduced to \( k^{|O|} \) where \( |O| \) is the cardinality of the set \( O \).

Beyond the “gradual” semantics of the Gödel implication, an important feature is that the set of all fixed points is finite under the use of this implication as given in the following proposition. Due to this important property, the computation of the whole concept lattice then becomes feasible.

**Proposition 2.** Let \( \mathcal{K} = (L, O, P, R) \) a \( L \)-context where \( L \) is an arbitrary scale (not necessarily finite). Then, the \( L \)-concept lattice \( \mathcal{L}(\mathcal{K}) \), is a finite set under Gödel implication.
Proof. It is a well known result that the sets of all intents and the set of all extents are isomorphic complete lattices \[2\]. Thus, it is sufficient to prove that one of them is a finite set, namely the set of all extents. Thus, we have:

\[
O_{\downarrow\uparrow}(o) = \bigwedge_{p \in P} \mathcal{I}_G(O_{\downarrow}(p), \mathcal{R}(o, p))
\]

\[
= \bigwedge_{p \in P} \left( \begin{cases} 1 & \text{if } O_{\downarrow}(p) \leq \mathcal{R}(o, p) \\ \mathcal{R}(o, p) & \text{elsewhere} \end{cases} \right)
\]

Thus, necessarily \(O_{\downarrow\uparrow}(o) \in \{\mathcal{R}(o_i, p_j)\} \cup \{1\}\) where \(i, j\) are integers bounded by the numbers of lines and columns of \(\mathcal{R}\). This implies that the set of all extents is finite. \(\square\)

It has been shown through the proof of Proposition \[2\] that \(O_{\downarrow\uparrow}(o) \in \{\mathcal{R}(o_i, p_j)\} \cup \{1\}\) (symmetrically \(P_{\downarrow\uparrow}(p) \in \{\mathcal{R}(o_i, p_j)\} \cup \{1\}\)). This means that each closed element is valued on the set of entries of the relation \(\mathcal{R}\). Thus, in order to determine all closed elements, it is then sufficient to enumerate \(L\)-sets valued in \(\{\mathcal{R}(o, p), o \in O, p \in P\}\) from the bottom element of the concept lattice to the top element (or conversely). The following proposition characterizes these elements. In the following, we assume that \(1_O\) and \(0_O\) (resp. \(1_P\) and \(0_P\)) stands for the extreme membership values to the corresponding \(L\)-set s.t. \(1_O(o) = 1\) and \(0_O(o) = 0\ \forall o \in O\) (resp. \(1_P(p) = 1\) and \(0_P(p) = 0\ \forall p \in P\)). We also denote by \(O^{\wedge}\) (resp. \(O^{\vee}\)) the particular \(L\)-sets s.t. \(O^{\wedge}(o) = \bigwedge_{p \in P} \mathcal{R}(o, p)\) (resp. \(O^{\vee}(o) = \bigvee_{p \in P} \mathcal{R}(o, p)\)) \(\forall o \in O\), and by \(P^{\wedge}\) (resp. \(P^{\vee}\)) the particular \(L\)-sets s.t. \(P^{\wedge}(p) = \bigwedge_{o \in O} \mathcal{R}(o, p)\) (resp. \(P^{\vee}(p) = \bigvee_{o \in O} \mathcal{R}(o, p)\) \(\forall p \in P\)).

**Proposition 3.** Under Gödel implication, the lower and upper bound of the \(L\)-concept lattice \(\mathcal{L}(\mathcal{K})\) are given by:

\[
\bigwedge_{(O, P)} \mathcal{L}(\mathcal{K}) = \langle O^{\wedge}, 1_P \rangle \quad \text{and} \quad \bigvee_{(O, P)} \mathcal{L}(\mathcal{K}) = \langle 1_O, P^{\wedge} \rangle
\]

Proof. \((.)^{\downarrow\uparrow}\) is isotone, so \(\forall o \in O : 0 \leq O(o) \leq 1 \Rightarrow 0^{\downarrow\uparrow}_O(o) \leq O^{\downarrow\uparrow}(o) \leq 1^{\downarrow\uparrow}_O(o)\).

i) let us determine the \(L\)-set \(0^{\downarrow\uparrow}_O\):

\[
0^{\downarrow\uparrow}_O(o) = \bigwedge_{p \in P} \mathcal{I}_G \left( \bigwedge_{o \in O} \mathcal{I}_G(0, \mathcal{R}(o, p)), \mathcal{R}(o, p) \right)
\]

\[
= \bigwedge_{p \in P} \mathcal{I}_G(1, \mathcal{R}(o, p)) \quad \text{since } \mathcal{I}_G(0, v) = 1 \ \forall v \in L
\]

\[
= \bigwedge_{p \in P} \mathcal{R}(o, p) \quad \text{since } \mathcal{I}_G(1, v) = v \ \forall v \in L
\]
ii) let us determine the $L$-set $1^\uparrow \downarrow$: 

$$1^\uparrow \downarrow (o) = \bigwedge_{p \in P} \mathcal{I}_G \left( \bigwedge_{o \in O} \mathcal{I}_G (1, \mathcal{R}(o, p)), \mathcal{R}(o, p) \right)$$

$$= \bigwedge_{p \in P} \mathcal{I}_G \left( \bigwedge_{o \in O} \mathcal{R}(o, p), \mathcal{R}(o, p) \right)$$

$$= \bigwedge_{p \in P} 1 = 1$$

Top and bottom intents are similarly determined. □

The following proposition will be useful for pruning useless potential fixed points in the algorithm.

**Proposition 4.** (Djouadi and Prade [13]) Let $O_1, O_2 \in L^O$ (resp. $P_1, P_2 \in L^P$) two $L$-sets such that $O_1 \subseteq O_2$ (resp. $P_1 \subseteq P_2$). If $O_2 \subseteq O_1^\uparrow$ (resp. $P_2 \subseteq P_1^\uparrow$) then $O_2^\downarrow = O_1^\downarrow$ (resp. $P_2^\downarrow = P_1^\downarrow$).

### 3.2.1 Construction Algorithm

The proposed algorithm is intended to build the lattice $\mathcal{L}(K)$ of all $L$-formal concepts (i.e. pairs $(O, P)$ of fixed points s.t. $O = P^\uparrow$ and $P = O^\downarrow$). Since the set of extents and the set of intents are isomorphic complete lattices, we may just build one of them, for instance the intent-concept lattice. The lattice construction process is organized through two procedures. The main procedure CLOSED\_LATTICE computes top and bottom intent-concept and calls the recursive procedure RECURSIVE\_CLOSURE which calls itself in order to generate all potential closed elements. The following notations and functions are used by the procedures.

- $m$: The cardinality of the set $P$ ($m = |P|$)
- $\text{INTENT\_SET}$: Set containing all intent concepts.
- $\text{Closed}$: Corresponds to a closed $L$-set.
- $\text{DIR\_SUCC}(\text{Closed}, p_j)$: Returns the direct successor of the $L$-set $\text{Closed}$ w.r.t the property $p_j$. The result (denoted $V$) is defined s.t.
  - $V(p_k) = \text{Closed}(p_k)$ when $k \neq j$,
  - $V(p_j) = \text{Closed}(p_j)$ when $\text{Closed}(p_j) = P^\uparrow(p_j)$,
  - $V(p_j) = \beta$ s.t. $\beta > \text{Closed}(p_j)$ and $\not\exists \beta' : \beta > \beta' > \text{Closed}(p_j)$, elsewhere.
- $\text{Successor}$: This $L$-set contains a direct successor of the last computed closed element.
- $\text{Closure}$: This $L$-set contains the closure of $\text{Successor}$.
- $\text{Next}$: Contains the direct successor of the last computed closure w.r.t the attribute $p_j$.
- $\text{FLAG}$: This Boolean is used to prune already computed closed elements.
Algorithm. CLOSED_LATTICE(\(\mathcal{K}\))

Input: A fuzzy context \(\mathcal{K} = (L, X, A, R)\).
Output: The set \(\text{INTENT_SET}\) of all fuzzy intent concepts.

Begin
1: \(\text{Top}(p_j) \leftarrow 1_p(p_j) \forall p_j \in P\); 
2: \(\text{Bottom}(p_j) \leftarrow P^\lor(p_j) \forall p_j \in P\); 
3: \(\text{INTENT SET} \leftarrow \{\text{Top}\} \cup \{\text{Bottom}\}\); 
4: \(\text{RECURSIVE_CLOSURE}(\text{Bottom})\); 
End

Proposition 4 is used by the procedure \(\text{RECURSIVE_CLOSURE}\) for pruning potential closed elements for which the closure is already computed. It allows also to fix the next potential closed element. This avoids to have to enumerate all potential closed elements and significantly reduces complexity, as shown through Example 4. The Boolean \(\text{FLAG}\) prevents the procedure \(\text{RECURSIVE_CLOSURE}\) from calling itself for already computed closed elements. Note that the use of a recursive procedure provides the ability to build the whole lattice structure while computing the closed elements.

Algorithm. \(\text{RECURSIVE_CLOSURE}(\text{Closed})\)

Begin
1: \(j \leftarrow 1\); /* index \(j\) allows to enumerate the attributes */ 
2: \(\text{FLAG} \leftarrow \text{TRUE}\); 
3: While \((j \leq m)\) and \(\text{FLAG}\) Do 
4: Begin 
5: \(\text{Next} \leftarrow \text{DIR_SUCC}(\text{Closed}, p_j)\); 
6: While \((\text{Next} \neq P \lor a_j)\) and \(\text{FLAG}\) Do 
7: Begin /* there exists a direct successor */ 
8: \(\text{Successor} \leftarrow \text{Closed}\); 
9: \(\text{Successor}(p_j) \leftarrow \text{DIR_SUCC}(\text{Closed}, p_j)\); 
10: \(\text{Closure} \leftarrow (\text{Successor})^{\uparrow}\); 
11: If \(\neg(\text{Closure} \in \text{INTENT_SET})\) 
12: Then 
13: \(\text{INTENT_SET} \leftarrow \text{INTENT_SET} \cup \{\text{Closure}\}\); 
14: \(\text{RECURSIVE_CLOSURE}(\text{Closure})\); 
15: \(\text{Next} \leftarrow \text{DIR_SUCC}(\text{Closure}, p_j)\); 
16: Else 
17: \(\text{FLAG} \leftarrow \text{FALSE}\); 
18: End If 
19: End While 
20: \(j++\); 
21: End While 
End
Graduality, Uncertainty and Typicality in Formal Concept Analysis

**Proposition 5.** The procedure CLOSED_LATTICE is complete i.e. it finds all closed sets \( P \in L^P \) s.t. \( P = P_{\text{cl}} \).

**Proof.** The proof follows from Proposition 2, Proposition 4 and the structure of the procedure RECURSIVE_CLOSURE which enumerates recursively all non useless potential closed elements. \( \Box \)

### 3.2.2 Illustrative Example

An example is given to illustrate the proposed algorithm. It is based on a fuzzy formal context that expresses relationships about meteorological observations. A special notation \( P = p_1^{\alpha_1}p_2^{\alpha_2} \ldots p_m^{\alpha_m} \) where \( \alpha_j = P(p_j) \), (resp. \( O = o_1^{\beta_1}o_2^{\beta_2} \ldots o_n^{\beta_n} \) where \( \beta_i = O(o_i) \)) will be used to represent \( L \)-sets \( P \in L^P \) (resp. \( O \in L^O \)).

<table>
<thead>
<tr>
<th>( R )</th>
<th>Low Temperature</th>
<th>Low Pressure</th>
<th>High Humidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm Day</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Cold Day</td>
<td>1</td>
<td>0.9</td>
<td>0.6</td>
</tr>
<tr>
<td>Rather Dry Day</td>
<td>0.0</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>Violent Wind Day</td>
<td>0.7</td>
<td>0.9</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 6 Fuzzy formal context

![Diagram of L-sets](image)

**Fig. 1** Pruning useless \( L \)-sets.
Fig. 2 Whole lattice of fuzzy formal concepts.
Example 4. Let us consider a fuzzy context which consists of:

- A set of properties $P = \{\text{Low Temperature, Low Pressure, High Humidity}\}$ respectively abbreviated as LT, LP, HH.
- A set of objects $O = \{\text{Warm Day, Cold Day, Rather Dry Day, Violent Wind Day}\}$ respectively abbreviated as WD, CD, RDD, VWD.
- A fuzzy relation $R$ illustrated in Table 6.
- A scale $L = [0, 1]$.

The proposed algorithm proceeds as follows (the variable $m$ is equal to 3):

Step 1: “Top” and “Bottom” on the intent lattice are determined. $Top = LT^1 LP^1 HH^1$. $Bottom = LT^{0.0} LP^{0.2} HH^{0.1}$.

Step 2: The procedure RECURSIVE CLOSURE($LT^{0.0} LP^{0.2} HH^{0.1}$) is called. The variable $Next$ takes the value 0.7 that corresponds to the direct successor of the degree 0.0 w.r.t the attribute LT in the relation $R$. The variable $Successor$ is first assigned the L-set $LT^{0.0} LP^{0.2} HH^{0.1}$ in line 8. Whereas in line 9, $Successor$ is assigned the degree $Next$ and becomes $LT^{0.7} LP^{0.2} HH^{0.1}$. In line 10, the closure of the L-set $Successor$ is computed and the variable $Closure$ is assigned the L-set $LT^{0.7} LP^{0.9} HH^{0.5}$.

Step 3: Since the intent-concept $LT^{0.7} LP^{0.9} HH^{0.5}$ has not been already generated, it will be added in the set INTENT SET and the Boolean FLAG remains true. At this level the Proposition 4 will be used to prune all useless potential closed elements (i.e. which will not generate new closed elements). In this precise case, Figure 1 illustrates a lattice of useless L-sets which will be pruned.

Step 4: The procedure RECURSIVE CLOSURE is again called with $(LT^{0.7} LP^{0.8} HH^{0.5})$ as parameter since the next L-set which may generate (or may be) a closed element is inevitably a direct successor of the L-set $LT^{0.7} LP^{0.8} HH^{0.5}$.

Step n: The procedure ends when all potentially closed elements are generated.

The whole lattice structure is depicted in Figure 2.

4 Uncertainty

Neither the standard FCA approach nor its fuzzy extension are equipped for representing situations of partial or complete ignorance. To this end, in the Boolean case, we need to introduce a proper representation of partial uncertainty including ignorance in the relational table of the formal context. One may think of introducing gradations of uncertainty by changing crosses
and blanks in the table into probability degrees, or by possibility or necessity degrees. In the probabilistic case, one number shall assess the probability that a considered property holds for a given object (its complement to 1 corresponding to the probability it does not hold). However, this is assuming precise knowledge on the probability values, which is not really appropriate if we have to model the state of complete ignorance. It is why we investigate the use of the possibilistic setting in the following.

In the possibilistic setting, crosses may be replaced by positive degrees of necessity for expressing some certainty that an object satisfies a property. The blanks could be refined by possibility degrees less than 1, expressing that it is little possible that an object satisfies a property. However, this convention using a single number in the unit interval for each entry in the context may be misleading as according to whether the number replaces a blank or a cross the meaning of the number is not the same.

In the possibilistic setting, possibility and necessity functions are related by the duality relation \( N(A) = 1 - \Pi(A) \), that holds for any event \( A \), where \( \overline{A} \) denotes the opposite event \[1\]. Then, for entries \((o, p)\) in the table, we use a representation as a pair of necessity degrees \((\alpha, 1 - \beta)\) where \( \alpha = N(o\overline{p}) \) (resp. \( 1 - \beta = N(o\overline{p}) \)) corresponds to the necessity (certainty) that object \( o \) has (resp. does not have) property \( p \). Moreover, we should respect the property \( \min(\alpha, 1 - \beta) = 0 \), since \( \min(N(A), N(\overline{A})) = 0 \) in agreement with complete ignorance, in which case nothing (i.e., neither \( A \) nor \( \overline{A} \)) is even somewhat certain. Pairs \((1,0)\) and \((0,1)\) correspond to completely informed situations where it is known that object \( o \) has, respectively does not have, property \( p \). The pair \((0,0)\) reflects total ignorance, whereas pairs \((\alpha, 1 - \beta)\) s.t. \( 1 > \max(\alpha, 1 - \beta) > 0 \) correspond to partial ignorance.

An uncertain formal context is thus represented by

\[
\mathcal{R}^U = \{ (\alpha(o,p), 1 - \beta(o,p)) \mid o \in \mathcal{O}, p \in \mathcal{P} \}
\]

where \( \alpha(o,p) \in [0,1] \), \( \beta(o,p) \in [0,1] \). A relational database with fuzzily-known attribute values is theoretically equivalent to the fuzzy set of all ordinary databases corresponding to the different possible ways of completing the information consistently with the fuzzy restrictions on the attribute values \[6\]. In the same way, an uncertain formal context may be viewed as a weighted family of all standard formal contexts obtained by changing uncertain entries into sure ones. More precisely, one may consider all the completions of an uncertain formal context. This is done by substituting entries \((o,p)\) that are uncertain, i.e., such that \( 1 > \max(\alpha(o,p), 1 - \beta(o,p)) \) by a pair \((1,0)\), or a pair \((0,1)\). Replacing \((\alpha(o,p), 1 - \beta(o,p))\) by \((1,0)\) is possible at degree \( \beta(o,p) \), the possibility that \( o \) has property \( p \). Similarly, replacing \((\alpha(o,p), 1 - \beta(o,p))\) by \((0,1)\) is possible at degree \( 1 - \alpha(o,p) \), the possibility that \( o \) does not have the property \( p \). In this way, one may determine to what extent a particular completion (a context \( \mathcal{C} \)) is possible, by aggregating the possibility degrees
associated with each completed entry (using min operator). Formally, one can write
\[ \pi(C) = \min(\min_{o,p : o \in C \land p \in C} \beta(o, p), \min_{o,p : o, p \in C} 1 - \alpha(o, p)) . \]
Likewise the degree of possibility that \((X, Y)\) is a formal context of \(R^U\) is
\[ \pi(X, Y) = \sup\{\pi(C) : C\text{ such that } (X, Y)\text{ is a formal context of } C\} . \]
Useful completions are those where partial certainty becomes full certainty. Indeed, given an uncertain formal context and a threshold pair \((u, v)\), let us replace all entries of the form \((\alpha, 0)\) such that \(\alpha \geq u\) with \((1, 0)\) and entries of the form \((0, 1 - \beta)\) such that \(1 - \beta \geq v\) with \((0, 1)\). All such replacements have possibility 1 according to the above formula. Remaining entries, which are more uncertain, can be systematically substituted either by \((1,0)\), or by \((0,1)\). Considering, the two extreme cases where all such entries are changed into \((1,0)\) and the case where all such entries are changed into \((0,1)\) gives birth to upper and lower completions, respectively. In this way, two classical (Boolean) formal contexts, denoted \(R^*_\{(u,v)\}\) and \(R^{\ast\ast}_{\{(u,v)\}}\) are obtained as respective results of the two completions. They allow to determine, for a given threshold \((u, v)\), maximal extensions (resp. minimal intensions) and minimal extensions (resp. maximal intensions) of uncertain formal concepts. It is clear that \(R^*_\{(u,v)\} \subseteq R^{\ast\ast}_{\{(u,v)\}}\).

**Example 5.** Table 7 exhibits a formal context where some entries are pervaded with uncertainty. Let us examine the situation regarding formal concepts. Take \(u = 0.7, v = 0.5\) for instance. In context \(R^\ast_{\{(0.7,0.5)\}}\), examples of formal concepts are pairs \(\{(6,7,8), \{c,d,e\}\}\), or \(\{(5,6,7,8), \{d,e\}\}\), or \(\{(2,3,4), \{g,h\}\}\), although with \(u = 0.9\), the last formal concept would reduce to \(\{(2,3), \{g,h\}\}\), i.e. the extent of the concept is smaller.

Now consider \(R^\ast_{\{(0.7,0.5)\}}\), where the entries with low certainty levels (either in favor or against the existence of the link between \(o\) and \(p\)) are turned into positive links. Then, \(\{(2,3,4), \{g,h\}\}\) remains unchanged as a formal concept, while a larger concept now emerges, namely \(\{(5,6,7,8), \{c,d,e\}\}\). However, one may prefer to consider the results obtained from \(R^\ast_{\{(0.7,0.5)\}}\), where only the almost certain information is changed into positive links. In the example, if we move down \(u\) to 0.5, and use \(R^\ast_{\{(0.5,0.5)\}}\) we still validate the larger former concept \(\{(5,6,7,8), \{c,d,e\}\}\). This illustrates the fact that becoming less and less demanding on the level of certainty, may enable the fusion of close concepts (here \(\{(6,7,8), \{c,d,e\}\}\), and \(\{(5,6,7,8), \{d,e\}\}\), providing a more synthetic view of the formal context.

This small example is intended to illustrate several points. First of all, it should be clear that being uncertain about the existence of a link between an object and a property is not the same as being certain about a gradual link. Second, under uncertainty, there are formal concepts whose boundaries
are not affected by uncertainty, while others are. Lastly, regarding certain enough pieces of information as certain may help simplifying the analysis of the formal context. Besides, the proposed setting may also handle inconsistent information by relaxing the constraint \( \min(\alpha, 1 - \beta) = 0 \). This would amount to introducing paraconsistent links between objects and properties.

## 5 Typical Objects and Less Important Properties

Beyond the formal concepts which are strongly related to the mathematical notion of closure, one may consider that a formal context conveys an implicit semantics related to the typicality of objects and to the importance of properties. The notions of typicality and importance are dually stated according to the following principles:

A) an object \( o \) is all the more typical with respect to a set of properties \( P \) as it has all the properties \( p \in P \) that are sufficiently important;  
B) a property \( p \) is all the more important with respect to a set of objects \( O \) as all the objects \( o \in O \) that are sufficiently typical possess it.

Let us illustrate these notions through an example. We consider the set of animals \( O = \{ \text{albatross, parrot, penguin, kiwi, turtle} \} \), and the set of properties \( P = \{ \text{laying eggs, having two legs, flying, having feathers} \} \). Table 8 describes these animals with respect to relevant properties. It can be easily seen from a FCA point of view w. r. t. \( \mathcal{R}_5 \), there are a set of “regular birds” (here ‘albatross’, ‘parrot’), a set of “more or less regular birds” (here ‘albatross’, ‘parrot’, ‘penguin’), a set of “less regular birds” including ‘kiwis’ and, a set of animals which are not at all categorized as “birds”. Thus, the idea of a “bird” (as in Table 8) may appear as a vague concept if on the one hand one considers that some birds (e.g. ‘albatross’, ‘parrot’) are more typical than others (e.g. ‘penguin’, ‘kiwi’), and on the other hand that the satisfaction of some properties (e.g. ‘laying eggs’, ‘having two legs’) is more important than others (e.g. ‘flying’, or even ‘having feathers’). More formally, given a

### Table 7 Uncertain formal concepts

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \times )</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \times )</td>
<td>( \times )</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.5,0)</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td></td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td>( \times )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>(0, 0.8)</td>
<td>( \times )</td>
<td>(0, 0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>( \times )</td>
<td>( \times )</td>
<td>(0.8, 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>( \times )</td>
<td>( \times )</td>
<td>(0.8, 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>( \times )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
formal context $K = (O, P, R)$, the definitions of typicality and importance of a concept $C$ are given as:

$$
\text{Table 8  Birds example}
$$

<table>
<thead>
<tr>
<th>$R_5$</th>
<th>eggs</th>
<th>2 legs</th>
<th>feather</th>
<th>fly</th>
</tr>
</thead>
<tbody>
<tr>
<td>albatross</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>parrot</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>penguin</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>kiwi</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>turtle</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Definition 2. Given a formal context $K$, the typicality $\tau_C$ (resp. the importance $\iota_C$) of a concept $C$ is a mapping $\tau_C : O \to L$ (resp. $\iota_C : P \to L$). Typicality and importance define themselves mutually as:

$$
\iota_C(p) = \bigwedge_{o \in O} \tau_C(o) \to R(o, p) \quad \text{or} \quad \tau_C(o) = \bigwedge_{p \in P} \iota_C(p) \to R(o, p)
$$

It is interesting to remark that under some conditions, the pair $(\tau_C, \iota_C)$ forms a Galois connection. Indeed, let us assume that the typical extension of a concept $C$ is given by $\tau_C(\text{albatross}) = \tau_C(\text{parrot}) = 1$, $\tau_C(\text{penguin}) = \alpha$, $\tau_C(\text{kiwi}) = \beta$, and $\tau_C(\text{turtle}) = 0$ with $1 > \alpha > \beta > 0$. Let us compute the fuzzy set of important properties according to Definition 2, i.e. $\iota_C(p) = \bigwedge_{o \in O} \tau_C(o)$, with the condition $a \to 1 = 1$ and $a \to 0 = 1 - a$. This choice expresses the idea that a bird not having property $p$ makes a property all the less important in the definition of the concept `bird` as this bird is considered more typical. Then, we obtain $\iota_C(\text{`laying eggs'}) = \iota_C(\text{`having two legs'}) = 1$; $\iota_C(\text{`flying'}) = 1 - \alpha$; $\iota_C(\text{`having feathers'}) = 1 - \beta$. Thus, `having feathers` is more important than `flying`, since $1 - \alpha < 1 - \beta$.

Let us now compute the fuzzy set of typical objects w.r.t. the obtained fuzzy concept intension $\iota_C(p)$. We get $\tau_C(\text{albatross}) = \tau_C(\text{parrot}) = 1$; $\tau_C(\text{penguin}) = \alpha$; $\tau_C(\text{kiwi}) = \beta$; $\tau_C(\text{turtle}) = 0$ (since $(1 - a) \to 0 = a$). We thus recover $\left( \bigwedge_{p \in P} \left( \bigwedge_{o \in O} \tau_C(o) \to R(o, p) \right) \to R(o, p) \right) = \tau_C(o) \forall o$, and we recognize a fuzzy Galois connection.

This example illustrates the idea that, viewing a formal concept as a maximal rectangle included in the formal context, one may tolerate some missing object-property links in such a rectangle, provided that they pertain to non-fully typical objects that may miss non-fully compulsory properties. This may facilitate the emergence of larger formal concepts, with wider range of applicability, but having exceptions.
6 Conclusion

We claim that the semantics of the grades that may be introduced in FCA is crucial and influences the way these grades should be handled. Starting with the view of a formal concept in a formal context as a maximal subset where the relation holds everywhere, we have suggested how the ideas of gradualness, of uncertainty, of typicality or of importance may be used for refining formal concepts in a graded way.

References

Choice Functions in Fuzzy Environment: An Overview*

Xuzhu Wang, Caiping Wu, and Xia Wu

Abstract. In this chapter, we present an overview of choice functions. Firstly, we introduce some important research topics on classic choices which serve as a guideline for the fuzzification research of choice functions. Then we begin with the fuzzy choice function defined by Banerjee. After a brief introduction to Banerjee’s work, various preferences derived from a choice function are investigated. Main rationality conditions are presented and the relationships between them are presented. Some necessary and sufficient conditions for $T$-transitive rationality are summarized. Afterwards, we turn to a special family of choice functions which are associated with fuzzy preferences. In this case, the final choices can be exact or fuzzy. For the former, various characterizations of the Orlovsky choice function are primary topics. For the latter, Roubens’s work is mainly introduced. Finally, the research around Georgescu fuzzy choice functions is surveyed. The investigation includes three parts: rationality characterization, rationality conditions and rationality indicators. In the first part, $G$-rationality, $G$-normality, $M$-rationality and $M$-normality

Wang Xuzhu
Department of Mathematics, Taiyuan University of Technology, Taiyuan, Shanxi, P.R. China
Tel.: 0086-351-6014819
e-mail: wangxuzhu2006@126.com

Wu Caiping
Department of Mathematics, Taiyuan University of Technology, Taiyuan, Shanxi, P.R. China
Tel.: 0086-351-6699148
e-mail: tywcp@sohu.com

Wu Xia
Department of Mathematics, Taiyuan University of Technology, Taiyuan, Shanxi, P.R. China Tel.: 0086-351-6373577
e-mail: zhaojiangxin@sohu.com

* The work is supported by Shanxi Natural Science Foundation 200901106.
are discussed. In the second part, some important fuzzy versions of rationality conditions are displayed and their interrelationships are presented. The third part is devoted to the introduction of various rationality and rationality condition indicators.

**Keywords:** Fuzzy choice function, Revealed preference, Rationality condition, Rationalization.

# 1 Classic Choice Functions

In the real world, we are frequently confronted with the problem of how to choose one or more alternatives from a set $S$ of alternatives. This problem is called a choice problem and the set of the chosen alternatives from $S$ is normally called a choice set which is denoted by $C(S)$. A choice function is a mapping (function) that associates a non-empty choice set $C(S)$ to any $S$ in a certain domain $B$ which is a subset of the set of all alternatives. The concept of a choice function is initiated for the research on Economics, both a consumer [3] and a demand function [1] in Economics are choice functions in essence. In consumer theory, the involved domain $B$ is a class of convex polyhedra representing “budget sets” which are obtained by taking the income and prices into account. It is generally assumed [1] that $B$ contains all finite subsets of $X$, where $X$ is the set of all alternatives. For a close examination of the domain $B$, the reader may refer to [4]. For simplicity, we assume that the set $X$ of all alternatives is finite and $B$ is the set of all non-empty subsets of $X$, which will be denoted by $P(X)$ in the sequel.

Given a choice function, there are many ways to derive preferences among which we mention the following important ones:

1. **Revealed preference** $R$: $xRy \iff \exists S \in P(X), x, y \in S$ and $x \in C(S)$.
2. **Generated preference** $\bar{R}$: $x \bar{R} y \iff x \in C\{\{x, y\}\}$.
3. **Strict revealed preference** $\tilde{R}$: $x \tilde{R} y \iff \exists S \in P(X), x \in C(S)$ and $y \notin C(S)$.

Those preferences were introduced in different names in the literature. For example, generated preference is called “relation generated by $C$” by Arrow [1] and “base relation” by Herzberger [15]. Since each of $R$, $\bar{R}$ and $\tilde{R}$ reveals the consumer’s (or the decision-maker’s) preference from a different angle, all of them are commonly called revealed preferences by many authors [3, 10, 11] in different senses. We give them the current names just to distinguish them for the sake of convenience. In the literature, other preferences derived from a choice function can be found, e.g., the transitive closure $t(R)$ and $t(\tilde{R})$ of $R$ and $\tilde{R}$ in [3, 4], $R_*$ in [10], $\tilde{P}$ in [17] etc.

It is easily checked that both $R$ and $\bar{R}$ are reflexive and complete since all singletons and all the sets of two elements are included in the domain $P(X)$. In addition, $R$ is acyclic [22]. An important role of $R$ is to describe normality of choice functions. A choice function $C$ is called normal
if \( \forall S \in P(X), C(S) = \{ x \in S | \forall y \in S, xRy \} \). The normality indicates that the choice function can be generated by its revealed preference \( R \).

In the study of choice functions, the most important issue is their rationality. To deal with this issue, various rationality conditions have been presented by a number of authors. In the following, we introduce some important ones.

(1) Condition \( \alpha \): \( \forall S, T \in P(X), x \in C(T) \) and \( S \subseteq T \) imply \( x \in C(S) \).

Its interpretation is: If an alternative is chosen from a set, then it is chosen from any subset of it.

(2) Condition \( \beta \): If \( S \subseteq T \) and \( x, y \in C(S) \), then \( x \in C(T) \) iff \( y \in C(T) \).

Its interpretation is: If both \( x \) and \( y \) are chosen from \( S \), a subset of \( T \), then \( x \) is chosen from \( T \) iff \( y \) is chosen from \( T \).

(3) Condition \( \gamma \): If \( S = \bigcup_{i \in I} S_i \) and \( x \in C(S_i) \) for all \( i \in I \), then \( x \in C(S) \).

Its interpretation is: If \( x \) is chosen from every subset \( S_i \) of \( S \) and the union of all \( S_i \) is \( S \) itself, then \( x \) is chosen from \( S \).

(4) Condition \( \delta \): For any \( S, T \in P(X) \) and \( x, y \in C(S) \), if \( S \subseteq T \), then \( C(T) \neq \{ x \} \).

Its interpretation is: If both \( x \) and \( y \) are chosen from \( S \), a subset of \( T \), then neither of them can be uniquely chosen from \( T \).

Condition \( \alpha \) was firstly proposed by Chernoff [2], Condition \( \beta, \gamma \) and \( \delta \) were due to Sen [4, 5]. These rationality conditions formulate the behavior of a choice function when a choice set is expanded or contracted, and thus they are frequently called expansion-contraction conditions. In [4], Arrow also put forward some similar conditions. For example, Arrow’s C2 or C3 is just Condition \( \alpha \) in a different form and the condition C4 combines \( \alpha \) and \( \gamma \).

It is shown in [4] that Condition \( \alpha \) and Condition \( \gamma \) together are equivalent to the normality of the choice function.

In [6], a weakened version \( \alpha_2 \) of \( \alpha \), a weakened version \( \gamma_2 \) of \( \gamma \) and two strengthened versions \( \beta^+ \) and \( \beta' \) of \( \beta \) can be found. For other expansion-contraction conditions, the reader may refer to [17] for WD, ND, WND, SS and MR, [24] for W1-W5, [16] for PI etc.

Besides expansion-contraction conditions, there exist other type of rationality conditions which are related to preferences derived from a choice function. The most famous of these are the Weak Congruence Axiom (WCA), the Weak Axiom of Revealed Preference (WARP) and their strengthened versions: the Strong Congruence Axiom (SCA) and the Strong Axiom of Revealed Preference (SARP), which are defined as follows.

(1) WCA: If \( y \in C(S), x \in S \) and \( xRy \), then \( x \in C(S) \).

(2) WARP: If \( x \tilde{P} y \), then \( yRc\tilde{P}x \).

(3) SCA: If \( y \in C(S), x \in S \) and \( xt(R)y \), then \( x \in C(S) \), where \( t(R) \) stands for the transitive closure of \( R \).

(4) SARP: If \( xt(\tilde{P})y \), then \( yRc\tilde{P}x \), where \( t(\tilde{P}) \) stands for the transitive closure of \( \tilde{P} \).

SCA was introduced by Richter [3] and its weakened form WCA by Sen [4]. WARP and SARP were introduced in the classic consumer theory by
An amazing conclusion is that all the four conditions are equivalent. They are also equivalent to the combination of Condition $\alpha$ and $\beta$.

There are a number of other revealed-preference-based conditions in the literature in addition to the preceding conditions. The reader may refer to [1] for C1, [10] for HARP, [17] for QRPA and [18] for Axiom 1 – Axiom 4 etc.

Given a choice function $C$, one main concern by researchers is whether or not there exists a binary relation $Q$ such that $C(S) = \{x \mid \forall y \in S, xQy\}$ for all $S$ in $P(X)$. If such relation does exist, then $C$ is called rational. Since we define choice functions on the domain $P(X)$, it is easy to show that if a choice function is rational then $Q = R$, and thus rationality and normality are equivalent in our context. If additional conditions are imposed on $R$, then we have some special rationality characterizations. A choice function is called transitive rational if the choice function is rational and $R$ is transitive. Considering the equivalency of some rationality conditions [4, 9], each of the following conditions is a necessary and sufficient condition of transitive rationality: WCA, SCA, WARP, SARP, $\alpha$ and $\beta$ together, $R = \tilde{R}$, $R = R_*$ ($R_* = \{(x, y) \mid x \notin S \text{ or } x \in C(S) \text{ or } y \notin C(S)\}$), the transitivity of $\tilde{R}$ together with the normality of the choice function, $\tilde{R} = \tilde{R}$ together with the normality of the choice function etc.

Besides transitive rationality, the investigation of other rationalities are also visible in the literature among which we mention mainly quasi-transitive rationality, semi-transitive rationality, pseudo-transitive rationality and semi-order rationality. A fuzzy choice function is called quasi-transitive rational if the fuzzy choice function is rational and $P_R$ is transitive, where $P_R$ stands for the strict relation of $R$, defined by $P_R = R \cap R^d$ ($R^d = \{(x, y) \mid yR^c x\}$). Concerning quasi-transitive rationality, a main conclusion is: for a normal choice function $C$, $R$ is quasi-transitive iff Condition $\delta$ is satisfied [4]. Other necessary and sufficient conditions can be found in [18, 19, 21, 24]. A fuzzy choice function is called semi-transitive rational if the fuzzy choice function is rational and $R$ is semi-transitive, i.e., $\forall a, b, c, d \in X, aRb$ and $bRc$ imply $aRd$ or $dRc$. For the characterization of semi-transitive rationality, the reader may refer to [19]. As for pseudo-transitive rationality and semi-order rationality, see [18, 19, 21, 24] for details.

Among choice functions, an important special family is that whose choice sets are based on a preference relation on $X$. Two famous preference-based choice sets are the set of best elements and the set of maximal elements [5]. Given a relation $R$ on $X$ and $S \in P(X)$, an element $x$ in $S$ is called a “best” (or “greatest”) element of $S$ w.r.t. $R$ if $\forall y \in S, xRy$, and the set of all best elements in $S$ is denoted by $G_R(S)$, i.e.,

$$G_R(S) = \{x \mid x \in S \text{ and } \forall y \in S, xRy\}.$$ 

An element $x$ is called a maximal element of $S$ w.r.t. $R$ if $\forall y \in S, yP^c_R x$, where $P_R$ is the strict relation of $R$
and the set of all maximal elements in $S$ is denoted by $M_R(S)$, i.e.,
\[ M_R(S) = \{ x | x \in S \text{ and } \forall y \in S, y P^c x \}. \]

Clearly, $G_R(S) \subseteq M_R(S)$ for any $S \in P(X)$. Unfortunately, $G_R(S)$ or $M_R(S)$ may be empty, and either $G_R$ or $M_R$ is not necessarily a choice function as a result. It is shown in [5] that if $R$ is reflexive and complete then $G_R(S) \neq \emptyset$ iff $R$ is acyclic on $X$, i.e.,
\[ \forall x_1, x_2, \ldots, x_n \in X, x_1 P x_2 P \ldots P x_n \text{ implies } x_1 R x_n. \]

It is pointed out in [23] that $\forall S \in P(X), G_R(S) = M_R(S)$ if $R$ is reflexive and complete. As a consequence, the acyclicity of $R$ is also a necessary and sufficient condition for $M_R(S) \neq \emptyset$ provided that $R$ is reflexive and complete.

In [5], the properties of $G_R$ are investigated. It is verified that (1) $G_R$ satisfies Condition $\alpha$ and $\gamma$ but not necessarily Condition $\beta$; (2) Under the assumption that $G_R$ is a choice function, $G_R$ satisfies Condition $\beta$ iff $R$ satisfies $P \circ I \subseteq P$, where $I$ is the indifference part of $R$ defined by $x I y \iff x R y$ and $y R x$.

## 2 Banerjee Fuzzy Choice Functions

In 1995, Banerjee [44] presented the definition of a fuzzy choice function, which is called the Banerjee fuzzy choice function [31]. Let $F(X)$ denote the set of all non-empty fuzzy subsets of $X$. A Banerjee fuzzy choice function is a mapping $C : P(X) \rightarrow F(X)$ such that $\forall S \in P(X)$, supp$C(S) \subseteq S$, where supp$C(S)$ denotes the support of the fuzzy choice set $C(S)$. It is easily checked that supp$C(S) \subseteq S$ is equivalent to $C(S) \subseteq S$. So supp$C(S) \subseteq S$ can be replaced by $C(S) \subseteq S$ which is in the same form as in the crisp case.

In this section, we assume that every choice set is normal, i.e., $\forall S \in P(X)$, $\exists x \in S$ such that $C(S)(x) = 1$.

Then, he carried out an investigation into the characterization of rationality of fuzzy choice functions in a general sense. To examine the rationality of a choice function $C$, Banerjee introduced some concepts related to a fuzzy choice function $C$, among which the most important ones include fuzzy revealed preference relation $R$ defined by
\[ R(x, y) = \bigvee_{\{S | x, y \in S\}} C(S)(x), \]
and normality:
\[ \forall S \in P(X), \forall x \in S, C(S)(x) = \bigwedge_{y \in S} R(x, y). \]
Clearly, (1) \( \forall x, y \in X, C(S)(x) \leq R(x, y) \); (2) If every choice set is normal, i.e., \( \forall S \in P(X), \exists x \in S \) such that \( C(S)(x) = 1 \), then \( R \) is reflexive and complete. Afterwards, Banerjee proposed three fuzzy congruence conditions FC1, FC2, FC3.

1. FC1: \( \forall S \in P(X), y \) is dominant in \( S \), i.e., \( C(S)(y) \geq C(S)(x) \) for all \( x \in S \), implies \( C(S)(x) = R(x, y) \) is valid for all \( x \in S \).
2. FC2: \( \forall S \in P(X) \) and \( \forall x, y \in S, y \) is dominant in \( S \) and \( R(x, y) \geq R(y, x) \) imply that \( x \) is dominant in \( S \).
3. FC3: \( \forall S \in P(X), \forall \alpha \in (0, 1] \) and \( \forall x, y \in S, C(S)(y) \geq \alpha \) and \( R(x, y) \geq \alpha \) imply \( C(S)(x) \geq \alpha \).

It is easily checked that all the three conditions are fuzzy versions of WCA in the crisp case. Banerjee showed that FC1, FC2 and FC3 are necessary and sufficient conditions for characterizing the rationality of a fuzzy choice function. Meanwhile, Banerjee pointed out that the three congruence conditions are independent. This is an incorrect assertion. As a matter of fact, the following dependencies are verified in [45]:

1. FC1 implies FC2;
2. FC3 implies FC2.

Hence FC1 and FC3 are sufficient to characterize the rationality of a Banerjee fuzzy choice function.

As in the crisp case, various preference relations can be derived from a fuzzy choice function. Besides the fuzzy revealed preference relation \( R \), some other preferences \( \bar{R}, \bar{R}_n, R_* \) derived from a fuzzy choice function \( C \) are also defined in different frameworks [34, 35, 51]. Here we assume that \( (T, S, n) \) is a De Morgan triple.

1. \( \bar{R}(x, y) = C(\{x, y\})(x) \);
2. \( \bar{R}_n(x, y) = n(\bar{P}_n(y, x)), \) where \( \bar{P}_n(x, y) = \bigvee_{\{S|x, y \in S\}} T(C(S)(x), n(C(S)(y))) \);
3. \( R_*(x, y) = \bigwedge_{\{S|x, y \in S\}} I_T(C(S)(y), C(S)(x)), \) where \( T \) is a \( t \)-norm,

\[ I_T(a, b) = \text{sup}\{c|T(a, c) \leq b\} \]

is the \( R \)-implication associated with \( T, n \) a negation.

In [51], under the condition that \( n \) is a strong negation, the relationships between these preferences are investigated in detail and the main results include:

1. \( \bar{R}_n \subseteq \bar{R} \subseteq R \);
2. \( R_* \subseteq \bar{R} \);
3. \( R = \bar{R}_n \) iff the choice function is normal and \( \bar{R} = \bar{R}_n \);
4. \( R_* = \bar{R}_n \) under a strong De Morgan triple;
5. If \( T \) is a continuous \( t \)-norm, then \( R = R_* \) iff \( \bar{R} = R_* \) and the choice function is normal.
Remark 1. Georgescu proved in [31] that the normality of the choice function and $\bar{R} = \bar{R}_n$ imply $R = \bar{R}_n$ in a different framework.

In the study of fuzzy choice functions, some important rationality conditions are fuzzified. As in the crisp case, they can be divided into revealed-preference-based conditions and expansion-contraction conditions. The former includes Weak Fuzzy Congruence Axiom (WFCA), Weak Axiom of Fuzzy Revealed Preference (WAFRP) and their strengthened versions: Strong Fuzzy Congruence Axiom (SFCA) and Strong Axiom of Fuzzy Revealed Preference (SAFRP). Let $(T, S, n)$ be a De Morgan triple. Then these rationality conditions are defined as follows.

1. WFCA: $\forall S \in P(X), \forall x, y \in S, T(R(x, y), C(S)(y)) \leq C(S)(x)$;
2. WAFRP: $\forall x, y \in X, \bar{P}_n(x, y) \leq n(R(y, x))$;
3. SFCA: $\forall S \in P(X), \forall x, y \in S, T(t(R)(x, y), C(S)(y)) \leq C(S)(x)$;
4. SAFRP: $\forall x, y \in X, t(\bar{P}_n)(x, y) \leq n(R(y, x))$.

It is easily checked that WFCA is equivalent to FC3 if $T = \min$. WFCA and WAFRP are not generally equivalent although their crisp counterparts WCA and WARP are equivalent. However, the equivalency can be obtained again by imposing the condition that $(T, S, n)$ is a strong De Morgan triple [52].

The most important expansion-contraction conditions include $F\alpha, F\beta, F\gamma, F\delta$ [48, 53].

1. $F\alpha$: If $S_1 \subseteq S_2$ then $C(S_2)(x) \leq C(S_1)(x)$ for any $x \in S_1$.
2. $F\beta$: If $S_1 \subseteq S_2$ and $x, y \in S_1$, then
   
   $\begin{array}{l}
   T(C(S_1)(x), C(S_1)(y), C(S_2)(x)) \leq C(S_2)(y).
   \end{array}$

3. $F\gamma$: For any $x \in S_1 \cap S_2$,
   
   $\begin{array}{l}
   T(C(S_1)(x), C(S_2)(x)) \leq C(S_1 \cup S_2)(x).
   \end{array}$

4. $F\delta$: If $S_1 \subseteq S_2$, $x, y \in S_1$ and $x \neq y$, then
   
   $\begin{array}{l}
   T(C(S_1)(x), C(S_1)(y)) \leq n(T(C(S_2)(x), T(n(C(S_2)(t))))).
   \end{array}$

It can be checked that they are indeed fuzzy versions of the conditions $\alpha, \beta, \gamma, \delta$ respectively. Concerning the connections between the above mentioned conditions, it is shown that

1. If $n$ is a strong negation, then $F\beta$ implies $F\delta$ [53].
2. If $T = \min$, then $F\alpha$ and $F\gamma$ are equivalent to the normality of the choice function [48].
3. If $T = \min$, then $F\alpha$ and $F\beta$ are equivalent to WFCA [33].
4. The normality of the choice function and $\bar{R} = \bar{R}_n$ $\iff$ WAFRP $\iff$ SAFRP $\iff$ $R = \bar{R}_n$ if the involved negation is a strong negation [51].
(5) \( T \)-regularity (reflexivity, completeness and \( T \)-transitivity) of \( R \) and the normality of the choice function \( \iff \) \( T \)-regularity of \( \bar{R} \) and the normality of the choice function \( \iff \) \( WFCA \iff SFCA \) \cite{34, 51}.

In addition to these popular rationality conditions, many other crisp rationality conditions employed by Suzumura \cite{9}, Schwartz \cite{24}, Arrow \cite{1} and Bandyopadhyay \cite{17} are also extended to the fuzzy case and the relationships between the fuzzified conditions are investigated in \cite{49, 53, 56}.

Another important research topic is the rationalization of fuzzy choice functions. If there exists a reflexive and complete fuzzy relation \( Q \) on \( X \) such that \( \forall S \in P(X), \forall x \in S, C(S)(x) = \bigwedge_{y \in S} Q(x, y) \), then we call the choice function \( C \) rationalized by \( Q \), or simply the choice function \( C \) rational. It is interesting that rationality and normality are equivalent if every choice set is normal. Firstly, it is clear that normality implies rationality. On the other hand, if \( C \) is rationalized by \( Q \), then for any \( x, y \in X \),

\[
R(x, y) = \bigvee_{\{S \mid x, y \in S\}} C(S)(x) = \bigvee_{\{S \mid x, y \in S\}} (\bigwedge_{z \in S} Q(x, z)) \leq Q(x, y).
\]

Thus \( C(S)(x) = \bigwedge_{y \in S} Q(x, y) \geq \bigwedge_{y \in S} R(x, y) \). On the other hand, \( C(S)(x) \leq R(x, y) \) for all \( x, y \in X \), which means \( C(S)(x) \leq \bigwedge_{y \in S} R(x, y) \). Hence, \( C(S)(x) = \bigwedge_{y \in S} R(x, y) \) for all \( x, y \in X \), i.e., \( C \) is normal.

There are various rationality characterizations dependent on the properties of the binary relations involved, of which \( T \)-transitive rationality is the most extensively studied one. A fuzzy choice function is called \( T \)-transitive rational if the fuzzy choice function is rational and \( R \) is \( T \)-regular. Each of the following statements is a necessary and sufficient condition of \( T \)-transitive rationality provided that \( T \) is continuous \cite{34, 51}.

(1) \( \bar{R} \) is \( T \)-regular and the choice function is normal;
(2) The choice function satisfies \( WFCA \);
(3) The choice function satisfies \( SFCA \);
(4) The choice function satisfies \( F\alpha \) and \( F\beta^+ \), where \( F\beta^+ \) means that \( S_1 \subseteq S_2 \) and \( x, y \in S_1 \), imply \( T(C(S_1)(x), C(S_2)(y)) \leq C(S_2)(x) \).
(5) \( R = R_\ast \);
(6) \( \bar{R} = R_\ast \) and the choice function is normal.

Furthermore, if \( (T, S, n) \) is a strong De Morgan triple, then the following statements are equivalent \cite{51} and thus every one can serve as a necessary and sufficient condition of \( T \)-transitive rationality.

(1) \( R = \bar{R}_n \);
(2) The choice function satisfies \( WAFRP \);
(3) The choice function satisfies \( SAFRP \);
(4) The choice function is normal and \( \bar{R} = \bar{R}_n \).
In addition, if $T = \min$, $T$-transitive rationality is equivalent to $F\alpha$ and $F\beta$.

Besides $T$-transitive rationality, the investigation of other rationalities are also visible in the literature among which we mention mainly $T$-quasi-transitive rationality and acyclic rationality. A fuzzy choice function is called $T$-quasi-transitive rational if the fuzzy choice function is rationalized by a reflexive, complete fuzzy relation $Q$ such that $P_Q$ is $T$-transitive, where $P_Q$ stands for the strict relation of $Q$. It was shown in [34] if a fuzzy choice function satisfies $F\alpha$ and $F\delta$, then $R$ is $T$-quasi-transitive rational. In [49], two necessary conditions of $T$-quasi-transitive rationality are presented by the fuzzification of the Schwartz’s crisp rationality conditions [24]. In [57], by using fuzzy counterparts of Aizerman’s [14] and Nehring’s [13] rationality conditions, several necessary conditions of $T$-quasi-transitive rationality are also proposed. Unfortunately, these necessary conditions are not sufficient to assure $T$-quasi-transitive rationality although all of them are sufficient in the crisp case.

If a fuzzy choice function $C$ is rational and rationalized by an acyclic relation $Q$ then $C$ is called acyclic rational. A fuzzy relation $Q$ on $X$ is acyclic if, for any $x_1, x_2, \ldots, x_n \in X$, $Q(x_1, x_2) > Q(x_2, x_1)$, $Q(x_2, x_3) > Q(x_3, x_2)$, \ldots, $Q(x_{n-1}, x_n) > Q(x_n, x_{n-1})$ imply $Q(x_1, x_n) \geq Q(x_n, x_1)$. As we know, a choice function is normal if and only if it is rational. Moreover, it is shown that if every choice set of a fuzzy choice function is normal, then the revealed preference is acyclic [49]. Hence we can draw the conclusion that every rational fuzzy choice function is acyclic rational provided that every choice set is normal. In [56], a necessary and sufficient condition is presented to characterize acyclic rationality by fuzzifying the Bandyopadhyay’s rationality conditions in [17].

As far as $T$-pseudo-transitive rationality and $T$-semi-transitive rationality are concerned, some elementary results can be found in [50].

Remark 2. It is seen that relations and their properties are extensively employed in the study of choice functions, particularly in the characterization of choice functions. They are deeply investigated in preference modeling theory [58]. In this theory, relations are employed to model pairwise comparison results between alternatives including strict preference, indifference, incomparability and large preference. Some properties of relations, e.g., reflexivity, completeness, transitivity, semi-transitivity, Ferrers property are among the most important concepts for modeling particular preference structures such as weak order, partial order, interval order, semi-order etc. As relations are fuzzified, their properties are also fuzzified, and as a result fuzzy preference structures are correspondingly investigated. For the literature on fuzzy preference structures, we recommend the references [29, 30, 55, 60, 61, 62].
Finally, we point out that the definition of rationality is adopted from [31]. There is an alternative definition in [44].

3 Choices Based on Fuzzy Preferences

3.1 Exact Choices Based on Fuzzy Preferences

In fuzzy environment, the study of choice functions was essentially initiated by Orlovsky when modeling fuzzy preferences in 1978 [37]. Orlovsky argued that a decision-maker may have a vague idea of preferences between alternatives, i.e. he find it difficult to state definitely that an alternative is preferred to another. In such cases, a number in [0, 1] may be readily employed to reflect the degree of preference. From then on, fuzzy preference relations have been extensively utilized in decision-making analysis including the research of choice functions.

Let $\mathcal{R}(X)$ the set of some fuzzy relations on $X$. The following definition is adopted from [39]. An $\mathcal{R}$-preference-based choice function (PCF) is a mapping $C : P(X) \times \mathcal{R}(X) \to P(X)$ such that

$$\forall S \in P(X), \forall R \in \mathcal{R}(X), C(S, R) \subseteq S.$$ 

For any $S \in P(X)$, $R \in \mathcal{R}(X)$ and $x \in X$, let

$$OV(x, S, R) = \min_{y \in S} \min \{1 - R(y, x) + R(x, y), 1\}.$$

Then a PCF is defined by

$$C_{OV}(S, R) = \{x \in S | \forall y \in S, OV(x, S, R) \geq OV(y, S, R)\},$$

which is called the Orlovsky choice function [38].

Remark 3. Take the notation

$$PD(S, R) = \{x \in S | \forall y \in S, R(x, y) \geq R(y, x)\}.$$

It is shown that $C_{OV} = PD$ provided $R$ satisfies acyclicity. Acyclicity is also a necessary and sufficient condition for $PD(S, R) \neq \emptyset$ (or equivalently, $C_{OV}(S, R) \neq \emptyset$) [23].

Besides the Orlovsky choice function, some other crisp choice functions based on fuzzy preferences can be found in the literature. Dutta et al. suggested four choice functions in [25] and Barrett et al. proposed nine unambiguous choice functions in [42]. In addition, Basu [26], Switalski [27] and Roubens [43] also suggested some ways to choose alternatives when confronted with fuzzy preferences. Certainly, the use of different choice function reflects the decision-maker’s different angle of viewing alternatives and may lead to different choice results.
The Orlovsky choice function attracted much attention among which the most important and fruitful research is the characterization of the Orlovsky choice function. In 1993, Banerjee firstly presented a characterization of the Orlovsky choice function for the so-called Banerjee transitive class of fuzzy relations [38]. By a Banerjee transitive fuzzy relation, we mean the fuzzy relation $R$ on $X$ such that $\forall x, y, z \in X$,

$$R(x, y) \geq R(y, x) \text{ and } R(y, z) \geq R(z, y) \text{ imply } R(x, z) \geq R(z, x).$$

This property was firstly proposed by Ponsard [59]. We adopted the name from [39]. Bouyssou pointed out that the contraction condition $\alpha$ is mistakenly taken for granted [39], and the characterization is incorrect as a result. Then he reformulated the characterization for the same class of fuzzy relations. Meanwhile, Bouyssou proposed a characterization of the Orlovsky choice function for the acyclic class of fuzzy relations. Afterwards, Sengupta characterized the Orlovsky choice function for the weak transitive class of fuzzy relations [41]. A fuzzy relation $R$ on $X$ is called weakly transitive if $\forall x, y, z \in X$,

$$R(x, y) > R(y, x) \text{ and } R(y, z) > R(z, y) \text{ imply } R(x, z) > R(z, x).$$

In addition to characterizing the Orlovsky choice function, Bouyssou characterized the “sum of differences” choice function by using Neutrality, Monotonicity and Independence of circuits [40].

Another research topic is the rationality properties of crisp preference-based choice functions. In [42], Barrett et al. proposed two classes of rationality conditions: conditions for choosing an alternative and conditions for rejecting an alternative. The first class of conditions includes:

1. **reward for pairwise weak dominance:**
   $$\forall y \in S, R(x, y) \geq R(y, x) \Rightarrow x \in C(S, R);$$

2. **reward for pairwise strict dominance:**
   $$\forall y \in S \setminus \{x\}, R(x, y) > R(y, x) \Rightarrow x \in C(S, R).$$

The second class of conditions includes:

1. **strict rejection condition:**
   $$\forall x \in S, R(y, x) \geq R(x, y) \text{ and } \exists y \in S, R(y, x) > R(x, y) \text{ implies } x \notin C(S, R);$$

2. **weak rejection condition:**
   $$\forall y \in S \setminus \{x\}, R(y, x) > R(x, y) \text{ implies } x \notin C(S, R).$$
Clearly, these conditions are all based on pairwise comparisons. In addition, they presented faithfulness (upper faithfulness, lower faithfulness) as a property to assess the rationality of preference-based choice functions. Afterwards, nine choice functions are examined based on these conditions. The results show that \( \max -mD \) satisfies most rationality properties, where \( \max -mD \) is defined by

\[
C_{mD}(S, R) = \{x | \forall y \in S, mD(x, S, R) \geq mD(y, S, R)\}
\]

and for any \( S \in P(X) \), \( x \in S \),

\[
mD(x, S, R) = \min_{y \in S \setminus \{x\}} (R(x, y) - R(y, x)).
\]

### 3.2 Fuzzy Choices Based on Fuzzy Preferences

In 1993, Roubens defined several fuzzy choice set based on a fuzzy relation \( R \) on \( X \), among which the following are typical ones [43]. Let \( P \) be the strict part of \( R \) defined by \( P(x, y) = T(R(x, y), 1 - R(y, x)) \) with \( T \) a \( t \)-norm. Let \( S \in P(X) \) and \( y \in S \).

1. \( D^R(y) = \bigwedge_{x \in S \setminus \{y\}} R(y, x) \) is the degree to which \( y \) is preferred to any element of \( S \).
2. \( N^R(y) = 1 - \bigvee_{x \in S \setminus \{y\}} R(x, y) \) is the degree to which any element of \( S \) is not preferred to \( y \).
3. \( SD^R(y) = \bigwedge_{x \in S \setminus \{y\}} P(y, x) \) is the degree to which \( y \) is strictly preferred to any element of \( S \).
4. \( SN^R(y) = 1 - \bigvee_{x \in S \setminus \{y\}} P(x, y) \), is the degree to which any element of \( S \) is not strictly preferred to \( y \).

Then he proposed some rationality properties to examine the rationality of these choice actions which include heritage property, condordance property, independence of irrelevant alternatives and reward for strict dominance. It can be easily checked heritage property and condordance property are just \( F\alpha \) and \( F\gamma \) respectively while reward for strict dominance is similar to reward for pairwise strict dominance proposed by Barrett et al [42]. Independence condition means that if \( x \) is better than \( y \) in a set, then the same is true for a larger set of alternatives. Afterwards, Roubens showed the fulfilment of these properties for the above mentioned preference-based fuzzy choices under some conditions.

As a matter of fact, exact choices and fuzzy choices are closely related to each other if a fuzzy preference is given. For instance, Barrett et al. [42] proposed nine choice functions in order to make unambiguous choices. However, we can see a fuzzy choice function behind every exact choice function.
Choice Functions in Fuzzy Environment: An Overview

Take their \( C_{mD} \) as an example. In this case, \( C_{mD} \) is the final exact choice function. In forming this choice function, \( mD(x, S, R) \) is employed, which may be regarded as a fuzzy choice set and thus a fuzzy choice function when \( S \) varies in \( P(X) \). Conversely, a fuzzy choice function can be also employed to make crisp choices. More specifically, assume that \( C(S, R) (S \in P(X)) \) is a fuzzy choice related to a fuzzy relation \( R \). Then by simply letting \( C'(S, R) = \{x|\forall y \in S, C(S, R)(x) \geq C(S, R)(y)\} \), we have the exact choice function \( C' \). Particularly, if \( C(S, R) \) is normal for every \( S \in P(X) \), then we may make exact choice \( C'(S, R) = \{x|C(S, R)(x) = 1\} \) for any \( S \), which is called the set of non-fuzzy non-dominated alternatives in \([37]\). Similar practice can be found in \([43]\) for \( DR, NR, SDR \) and \( SNR \).

For a survey of preference-based choice functions, the reader may refer to \([63]\).

4 Georgescu Fuzzy Choice Functions

In the definition of Banerjee fuzzy choice function, the available set of alternatives are crisp sets, i.e, we make the fuzzy choice \( C(S) \) from the crisp set \( S \). In other words, the domain of a Banerjee choice function is crisp although its range is fuzzy. Georgescu argued that “If a fuzzy subset \( S \) of \( X \) will represent a vague criterion and \( x \) is an alternative, then the real number \( S(x) \) means the degree to which the alternative \( x \) verifies the criterion \( S \). \( S(x) \) will be called the availability degree of \( x \)” \([31]\), Page 87). Hence, she presented a new definition of fuzzy choice function by extending the domain from crisp sets to fuzzy sets. A Georgescu fuzzy choice function is a mapping \( C: F(X) \rightarrow F(X) \) such that \( C(S) \subseteq S \) is valid for any \( S \in F(X) \). This is the most general definition of a fuzzy choice function in the literature. In this section, a fuzzy choice function means a Georgescu fuzzy choice function and the involved work is mainly done by Georgescu \([31, 32, 33, 34, 35]\) in the framework \((*, \rightarrow, \neg)\), where \(*\) is a continuous \( t \)-norm, \( x \rightarrow y = \sup\{z|x*z \leq y\} \) and \( \neg x = x \rightarrow 0 \). We assume that every choice set is normal, i.e., \( \forall S \in F(X), \exists x \in [0, 1], C(S)(x) = 1 \).

Remark 4. Fuzzy logic connectives are fundamental tools for the research on Georgescu fuzzy choice functions. For their detailed discussion, the reader may refer to \([28, 46, 47, 65, 66, 67]\).

For a Georgescu fuzzy choice function \( C \), some main fuzzy preferences derived from \( R \) are defined as follows.

1. \( R(x, y) = \bigvee_{S \in F(X)} (C(S)(x) * S(y)). \)
2. \( \bar{R}(x, y) = C(\{x, y\})(x) \) and \( \bar{P}(x, y) = \bar{R}(x, y) * \neg \bar{R}(y, x). \)
3. \( \tilde{R}(x, y) = \neg \tilde{P}(y, x) \), where \( \tilde{P}(x, y) = \bigvee_{S \in F(X)} (C(S)(x) * S(y) * \neg C(S)(y)). \)
4.1 Rationality Characterization

Let $Q$ be a fuzzy relation on $X$. Define $M_Q$ and $G_Q$ by:

\[ M_Q(x) = S(x) \ast \bigwedge_{y \in X} [(S(y) \ast Q(y, x)) \rightarrow Q(x, y)], \]

\[ G_Q(x) = S(x) \ast \bigwedge_{y \in X} ((S(y) \rightarrow Q(x, y))). \]

Clearly, $G_Q(S) \subseteq M_Q(S)$ for any $S \in F(X)$. It is verified [31] that $G_Q(S) = M_Q(S)$ if $Q$ is reflexive and complete. A fuzzy choice function is called $G$-rational ($M$-rational respectively) if there exists a fuzzy relation $Q$ on $X$ such that $C = G_Q$ ($C = M_Q$ respectively). Particularly, $G$-rational is called $G$-normal and $M$-rational is called $M$-normal when $Q = R$. Concerning normality and rationality, there are the following conclusions [31]:

1. $M$-rationality implies $G$-rationality;
2. If $\ast = \text{min}$, $G$-normality and $G$-rationality are equivalent;
3. If $\ast = \text{min}$, $M$-normality implies $G$-normality.

If additional conditions are imposed, then we have special rationalities. A fuzzy choice function $C$ is called full rational if there exists a fuzzy relation $Q$ on $X$ which is reflexive, $\ast$-transitive and strongly total ($\forall x, y \in X, x \neq y, Q(x, y) = 1$ or $Q(y, x) = 1$) such that $C = G_Q$. A fuzzy choice function $C$ is called strongly complete rational if there exists a fuzzy relation $Q$ on $X$ which is strongly complete (reflexive, strongly total) such that $C = G_Q$. It is verified that any $M$-normal fuzzy choice function $C$ is strongly complete rational under the $t$-norm min.

Compared with the rationality characterization of crisp choice functions, the results related to the fuzzy case in the sense of Georgescu are much fewer.

4.2 Rationality Conditions

Some important rationality conditions including WCA, SCA, WARP, SARP, etc. have been fuzzified by Georgescu. To distinguish them from the conditions in the sense of Banerjee, we take the notations WFCAG, SFCA(G), WAFRG(G), SAFRG(G) etc.

1. $\text{WAFRG}(G)$: $\tilde{P}(x, y) \leq \neg R(y, x)$ for all $x, y \in X$.
2. $\text{SAFRG}(G)$: $t(\tilde{P})(x, y) \leq \neg R(y, x)$ for all $x, y \in X$, where $t(\tilde{P})$ is the $\ast$-transitive closure of $\tilde{P}$.
3. $\text{WFCA}(G)$: $R(x, y) \ast C(S)(y) \ast S(x) \leq C(S)(x)$ for all $S \in F(X)$ and $x, y \in X$.
4. $\text{SFCA}(G)$: $t(R)(x, y) \ast C(S)(y) \ast S(x) \leq C(S)(x)$ for all $S \in F(X)$ and $x, y \in X$, where $t(R)$ is the $\ast$-transitive closure of $R$. 
Generally, WFCA(G) implies WAFRP(G), and SFCA(G) implies SAFRP(G) \[31\]. In addition, if we consider the following statements:

(i) \( R \) is regular and \( C \) is \( G \)-normal;
(ii) \( \tilde{R} \) is regular and \( C \) is \( G \)-normal;
(iii) \( C \) verifies WFCA(G);
(iv) \( C \) verifies SFCA(G);
(v) \( C \) verifies WAFRP(G);
(vi) \( C \) verifies SAFRP(G);
(vii) \( R = \tilde{R} \);
(viii) \( \tilde{R} = \tilde{R} \) and \( C \) is \( G \)-normal,

then the following assertions are true \[31\]:

(1) (i) and (ii) are equivalent.
(2) (i) implies (iii); if \( * = \min \), then (iii) implies (i).
(3) (iii) and (iv) are equivalent.
(4) If \( * \) is the Lukasiewicz \( t \)-norm, then (iii), (v), (vi), (vii) are equivalent.
(5) (viii) implies (vii).

These assertions extend the results obtained by Sen in \[4\], where all the assertions are equivalent in the crisp case.

Another fuzzy version WAFRP\(^0\) of WARP also plays an important role in the study of Georgescu fuzzy choice function. Let \( A, B \in F(X) \). Take the notations:

\[
\begin{align*}
I(A, B) &= \bigwedge_{x \in X} (A(x) \rightarrow B(x)); \\
E(A, B) &= \bigwedge_{x \in X} (A(x) \leftrightarrow B(x)).
\end{align*}
\]

Then we have WAFRP\(^0\): For any \( x, y \in X \) and \( S_1, S_2 \in F(X) \),

\[
(S_1(x) * C(S_2)(x)) * (S_2(y) * C(S_1)(y)) \leq E(S_1 \cap C(S_2), S_2 \cap C(S_1)).
\]

WAFRP\(^0\) implies \( M \)-normality and thus \( G \)-rationality and \( M \)-rationality as well. In addition, WAFRP\(^0\) is equivalent to WFCA(G) under the \( t \)-norm \( \min \).

For a Georgescu fuzzy choice function, some expansion-contraction conditions have the following fuzzy versions:

(1) \( F\alpha(G) : \forall x \in X, \forall S, T \in F(X), I(S, T) * S(x) * C(T)(x) \leq C(S)(x) \).
(2) \( F\beta(G) : \forall x, y \in X, \forall S, T \in F(X), I(S, T) * S(x) * C(T)(y) \leq C(T)(x) \leftrightarrow C(T)(y) \).
(3) \( FAA : \forall x \in X, \forall S, T \in F(X), I(S, T) * S(x) * C(T)(x) \leq E(S \cap C(T), C(S)) \).

Clearly, normality implies \( F\alpha(G) \). Concerning the relationships between these conditions, it is worth mentioning that, if the involved \( t \)-norm \( * \) is \( \min \), then \( (1) WFCA(G) \) is equivalent to the combination of \( F\alpha(G) \) and \( F\beta(G) \); \( (2) FAA \) is equivalent to WFCA(G) and can be employed to characterize the full rationality of the choice function.
4.3 Rationality Indicators

In the study of Georgescu fuzzy choice functions, a unique aspect is the measure of the degree of rationality and the degree of the satisfaction of rationality conditions. Instead of examining whether a fuzzy choice function is rational or not, a rationality degree is defined to measure the rationality of the choice function. Similarly, instead of checking whether or not a rationality property is satisfied or not, an indicator is defined to indicate the degree to which the property is satisfied. In order to achieve this aim, we firstly introduce a basic notion: the similarity of two fuzzy choice functions $C_1$ and $C_2$ defined by

$$E(C_1, C_2) = \bigwedge_{x \in X} \bigwedge_{S \in F(X)} (C_1(S)(x) \leftrightarrow C_2(S)(x)).$$

Let $R = F(X \times X)$ and $R$ be the revealed preference. Firstly, we present the rationality and normality degrees:

- **G-rationality degree:** $\text{Rat}_G(C) = \bigvee_{Q \in R} E(C, G_R)$;
- **M-rationality degree:** $\text{Rat}_M(C) = \bigvee_{Q \in R} E(C, M_R)$;
- **G-normality degree:** $\text{Norm}_G(C) = E(C, G_R)$;
- **M-normality degree:** $\text{Norm}_M(C) = E(C, M_R)$.

Concerning the connections between these rationality indicators, the following are valid:

$$\text{Rat}_M(C) \leq \text{Rat}_G(C), \text{Norm}_M(C) \leq \text{Norm}_G(C) \text{ and } \text{Rat}_G(C) = \text{Norm}_G(C),$$

which extend the results related to the relationships between $M$-rationality, $M$-normality, $G$-rationality and $G$-normality. Moreover, Georgescu defined and discussed strongly complete rationality indicator.

As for rationality conditions, the following indicators are typical:

1. $\text{WAFRP}(C) = \bigwedge_{x,y \in X} (\bar{P}(x,y) \rightarrow \neg R(y,x))$;
2. $\text{SAFRP}(C) = \bigwedge_{x,y \in X} (t(\bar{P})(x,y) \rightarrow \neg R(y,x))$;
3. $\text{WFCA}(C) = \bigwedge_{x,y \in X} \bigwedge_{S \in F(X)} ((S(x) \ast C(S)(y) \ast R(x,y)) \rightarrow C(S)(x))$;
4. $\text{SFCA}(C) = \bigwedge_{x,y \in X} \bigwedge_{S \in F(X)} ((S(x) \ast C(S)(y) \ast t(R)(x,y)) \rightarrow C(S)(x))$.

If the involved $t$-norm is min, the equalities $\text{WFCA}(C) = \text{SFCA}(C) = E(C, G_R) \land \text{Trans}(R) = E(C, G_R) \land \text{Trans}(\bar{R})$ can be found in [31], which extend the conclusions in [4] to degree descriptions. If $\ast$ is the Lukasiewicz $t$-norm, then $\text{WAFRP}(C) = \text{WFCA}(C) = E(R, \bar{R})$. 

The rationality condition \( WAFRP^0 \) indicator was also investigated. Its definition is
\[
WAFRP^0(C) = \bigwedge_{x,y \in X} \bigvee_{S,T \in F(X)} [(S(x) \ast C(T)(x) \ast (T(y) \ast C(S)(y))) \rightarrow E(S \cap C(T), T \cap C(S))].
\]

It is verified in [64] that
\[
WFCA(C) = WAFRP^0(C) \leq \min\{WAFRP(C), Norm_M(C), E(G_R, M_R)\}
\]
if \(* = \min\).

Important expansion-contraction condition indicators include \( F\alpha(C), F\beta(C), FAA(C), F\alpha_2(C), F\beta_2(C) \), etc.

1. \( F\alpha(C) = \bigwedge_{x \in X} \bigwedge_{S,T \in F(X)} [(I(S, T) \ast S(x) \ast C(T)(x)) \rightarrow C(S)(x)]; \)
2. \( F\beta(C) = \bigwedge_{x,y \in X} \bigwedge_{S,T \in F(X)} [(I(S, T) \ast C(S)(x) \ast C(S)(y)) \rightarrow (C(T)(x) \leftrightarrow C(T)(y))]; \)
3. \( F\alpha_2(C) = \bigwedge_{x \in X} \bigwedge_{S \in F(X)} [(I(\{x\}, S) \ast C(S)(x)) \rightarrow C(\{x\})(x)]; \)
4. \( F\beta_2(C) = \bigwedge_{x,y \in X} \bigwedge_{S \in F(X)} [(I(S\{x,y\}, S) \ast C(\{x,y\})(x) \ast C(\{x,y\})(y)) \rightarrow (C(S)(x) \leftrightarrow C(S)(y))]; \)
5. \( FAA(C) = \bigwedge_{x \in X} \bigwedge_{S,T \in F(X)} [(I(S,T) \ast S(x) \ast C(T)(x)) \rightarrow E(S \cap C(T), C(S))].\)

If the involved \( t\)-norm is min, then \( WFCA(C) = F\alpha(C) \land F\beta(C). \)

**Remark 5.** In [31], it was claimed that \( WFCA(C) = F\alpha(C) \land F\beta(C) \) without the requirement \(* = \min\). However, the established proof was based on this requirement.

Meanwhile, two transitive rationality indicators \( TrRat_G(C) \) and \( TrRat_M(C) \) were defined in [64].

\[
TrRat_G(C) = \bigvee_{Q \in \mathcal{R}} E(C, G_Q) \land Trans(Q);
\]
\[
TrRat_M(C) = \bigvee_{Q \in \mathcal{R}} E(C, M_Q) \land Trans(Q).
\]

Main conclusions include \(* = \min\)

1. \( F\alpha_2(C) \land F\beta_2(C) = E(C, G_R) \land Trans(R) \leq TrRat_G(C); \)
2. \( F\alpha_2(C) \land F\beta_2(C) = F\alpha(C) \land F\beta_2(C) = F\alpha_2(C) \land F\beta(C); \)
3. \( FAA(C) = E(C, G_R) \land Trans(R). \)

Recently, two acyclic rationality indicators \( AcRat_G(C) \) and \( AcRat_M(C) \) are put forward by Georgescu in [36] and the equality \( Rat_G(C) = AcRat_G(C) \) is verified.
5 Concluding Remarks

This state-of-the-art survey presents a brief introduction to various types of choice functions in fuzzy environment. The involved topics mainly focus on revealed preferences, rationality properties and rationalization. It should be noted that all the results are heavily dependent on the context. In our context, the domain of a crisp choice function is assumed to be $P(X)$. This assumption largely simplifies our discussion. In this case, normality and rationality are equivalent, the revealed preference become reflexive, complete and acyclic. Meanwhile, some notions in the literature become indifferent. For example, rationality and acyclic rationality become identical. $M$-rationality and $G$-rationality defined by Suzumura [10] become the same notion etc. In addition, in the discussion of Banerjee fuzzy choice functions and Georgescu fuzzy choice functions, different frameworks are adopted in considering the current research state. For the former, the framework is a De Morgan triple and it is a residuated lattice $([0, 1], \lor, \land, *, \rightarrow, 0, 1)$ for the latter. Hence, different fuzzy logic connectives may be employed in formulating rationalities although their intuitive interpretations are the same. This difference of framework may make the difference of the research results extremely subtle although some conclusions appear in almost the same or similar form. A Georgescu fuzzy choice function is the most general choice function since both the domain and range of the choice function are fuzzy. However, this generality often makes the investigation difficult. As a consequence, it is observed that many results related to Georgescu fuzzy choice functions are obtained with the condition $* = \min$. In this case, as we know, the corresponding negation should be the intuitionistic negation $[29]$ with the definition $\neg a = 1$ if $a = 0$ and $\neg a = 0$ elsewhere, which makes many rationality properties associated with the negation become special.

Acknowledgements. One of the authors, Wang Xuzhu, wish to express his sincere gratitude to Prof. Kerre for the patient and professional guidance during Wang's Ph. D. work and for the successful cooperation after Wang left Gent University from which Wang benefited greatly.

References

64. Wu, X.: A further study on rationality indicators of fuzzy choice functions, MS degree thesis, Taiyuan University of Technology (2010) (in chinese)
A Fuzzy Extension to Compact and Accurate Associative Classification

Guoqing Chen, Yixi Xiong, and Qiang Wei

Abstract. Classification has been one of the focal points in data mining research and applications. With an effective approach to building compact and accurate associative classification (namely GARC – Gain-based Association Rule Classification (Chen, Liu, Yu, Wei, & Zhang, 2006)) in forms of association rules, this chapter explores a way of fuzzy extension to GARC in dealing with the problem caused by crisp partitions for continuous attribute domains in data. Concretely, the sharp boundaries of the partitioned intervals are smoothened using fuzzy sets (or often conveniently labeled in linguistic terms) so as to reflect a variety of fuzziness on the domains (parameterized in $f_2$), giving rise to a fuzzy associative classifier (i.e., GARC$_{f_2}$). Furthermore, due to the fuzziness involved, the notions of information gain, rule redundancy and conflicts are extended, aimed at providing the desirable features of GARC in the fuzzy extension context for accuracy and compactness. Moreover, data experiments on benchmarking datasets as well as a real-world application illustrate the effectiveness of the proposed fuzzy associative classifier.

Keywords: Associative Classification, Information Gain, Fuzzy Partitions, Credit Rating.

In recent decades, there have been numerous research efforts as well as a wide range of applications in both classification and association rule mining (Fayyad & Utthurasamy, 1994; Han, Kamber, & Pei, 2005; Weiss & Kulikowski, 1991). In the context of massive data, association rule mining is to discover associative patterns between data items in a manner they occur together frequently, while classification is to build a classifier usually based on training/testing datasets, which enables to

Guoqing Chen · Yixi Xiong · Qiang Wei
School of Economics and Management, Tsinghua University
Beijing 100084 China
classify the newly coming data to a specific group (class). Though one may find a number of existing approaches that could build classifiers effectively, such as Decision Tree, Bayesian networks, Support Vectors Machine, Neural Network, etc. (Berzal, Cubero, Sanchez, & Serrano, 2004; Friedman, Geiger, & Goldszmidt, 1997; Quinlan, 1993; Shafer, Agrawal, & Mehta, 1996; Sharpe & Glover, 1999; Shie & Chen, 2008), associative classification, which results from a combined perspective of classification and association rule mining, has attracted an increasing amount of attention of academia and practitioners due to its accuracy and understandability (Ali, Manganaris, & Srikant, 1997; Berzal et al., 2004; Chen, Liu, Yu, Wei, & Zhang, 2006; Chen, Zhang, & Yu, 2006; Dong, Zhang, Wong, & Li, 1999; Li, Han, & Pei, 2001; Liu, Hsu, & Ma, 1998; Liu, Ma, & Wong, 2000; Wang, Zhou, & He, 2000; Yin & Han, 2003; Zhang, Chen, & Wei, 2009).

Concretely, an association rule is a link between data items in forms of $X \Rightarrow Y$, where $X$ and $Y$ are itemsets (i.e., set of items/attributes) with $X \cap Y = \emptyset$, expressing that the event for $X$ is associated with the event for $Y$. For example, a rule like “Milk $\Rightarrow$ Bread” means that the customers who buy milk may also tend to buy bread at the same time. There are two basic measures to evaluate a rule, namely, Degree of Support (Dsupp) and Degree of Confidence (Dconf). $\text{Dsupp}(X \Rightarrow Y) = \frac{||X \cup Y||}{|D|}$ is the percentage of transactions containing both $X$ and $Y$ with respect to the whole transaction set (dataset), while $\text{Dconf}(X \Rightarrow Y)$ is the percentage of transactions containing $Y$ among those transactions (i.e., tuple/record) containing $X$. That is, for itemsets $X$ and $Y$, $\text{Dsupp}(X \Rightarrow Y) = \frac{||X \cup Y||}{|D|}$, and $\text{Dconf}(X \Rightarrow Y) = \frac{||X \cup Y||}{||X||}$, where $|D|$ is the number of transactions in dataset $D$, and $||X||$ and $||X \cup Y||$ are the numbers of transactions that contain, respectively, $X$ and $X \cup Y$. Given two thresholds for degrees of support and confidence (i.e., minimal support $\alpha$ and minimal confidence $\beta$, with $\alpha, \beta \in [0,1]$), an itemset $X$ is called frequent, if $\text{Dsupp}(X) = \frac{||X||}{|D|} \geq \alpha$. Furthermore, for any rule $X \Rightarrow Y$, it is called a valid association rule if $\text{Dsupp}(X \Rightarrow Y) \geq \alpha$ and $\text{Dconf}(X \Rightarrow Y) \geq \beta$. To find such valid rules, many approaches have been proposed (Agrawal, Imielinski, & Swami, 1993; Agrawal & Srikant, 1994; Chen, Wei, Liu, & Wets, 2002; Mannila, Toivonen, & Verkamo, 1994; Srikant & Agrawal, 1996; Srikant, Vu & Agrawal, 1997), among which the Apriori method is usually deemed as a classical algorithm (Agrawal & Srikant, 1994).

In the spirit of association rules, associative classification can be viewed as a special case of association rule mining, resulting in a classifier composed of a number of class association rules (CARs) such as CBA (Liu et al., 1998; Liu et al., 2000). A CAR is of the form $X \Rightarrow C_i$, where $X$ is a set of items and $C_i$ is a predetermined class label. Let $t$ be a transaction in dataset $D$, $t$ is called to satisfy $X \Rightarrow C_i$ if $X \subseteq t$ and $C_i$ is the predetermined class for $t$. Associative classification usually searches for all the CARs that satisfy given minimum support and minimum confidence thresholds. Notably, associative classification has been considered effective in many cases (Berzal et al., 2004; Chen et al., 2006; Chen et al., 2006; Dong et al., 1999; Li et al., 2001; Liu et al., 1998; Liu et al., 2000; Wang et al, 2000; Yin & Han, 2003; Zhang et al., 2009).
While classifiers in forms of association rules are usually appealing for use and explanation by human users, one may however suffer from the fact that the number of discovered CARs is huge, thus hard to manage in understanding the rules as well as in resolving possible rule redundancy and conflicts. Therefore, the construction of an effective classifier to cope with the size of rules and redundant/conflict rules is desirable. In this regard, an approach, namely GARC (i.e., Gain-based Association Rule Classification) (Chen et al., 2006), has been introduced that possesses satisfied accuracy for classification as well as compactness in size of the resultant classifier.

In GARC, only those “informative” rules are generated. The characteristic of this approach is threefold (Chen et al., 2006): First, GARC employs information gain measure to select the best-split attribute for 1-itemsets (i.e., the itemsets each containing a single item), which is to be included in the generation of candidate k-itemsets (i.e., the itemsets each containing k items, k ≥ 2). Second, both frequent itemsets and excluded itemsets (i.e., infrequent itemsets) are used in the generation of candidate itemsets. Third, GARC defines rule redundancy and rule conflicts as follows: in a set of classification rules φ, if rule X⇒Ci exists, then (1) any rule of the form Z⇒Ci (Z⊃X) is called redundant, that is, Z⇒Ci would not be regarded semantically necessary from the perspective of classification; and (2) for i≠j, any rule of the form Z⇒Cj (Z⊇X) is called conflicting if Z⇒Ci does not precedes X⇒Ci (i.e., (i) Dconf(X⇒Ci) > Dconf(Z⇒Cj), or (ii) Dconf(X⇒Ci) = Dconf(Z⇒Cj) while Dsupp(X⇒Ci) > Dsupp(Z⇒Cj)), which may lead to misclassification in the decision making process. A nonempty classification rule set is called a compact set of classification rules if neither rule redundancy nor rule conflicts exists. Furthermore, GARC applies two strategies in avoiding rule conflicts and redundancy: (1) If β>50%, then rules X⇒Ci and X⇒Cj will not hold simultaneously. (2) For a classification rule X⇒Ci, if 1−Dconf(X⇒Ci)<α or Dsupp(X)<2α, then any itemset like Z⇒Cj (Z⊇X, i≠j) is an excluded itemset. As a result, GARC constructs a more condensed and understandable (in terms of much fewer rules than CBA) classifier while the classification accuracy is satisfactory compared with other major approaches (Chen et al., 2006).

It is worth mentioning that classical association rule mining methods and related associative classification methods were originally designed for binary databases in that the values of attributes (items) are either 1 (true) or 0 (false), reflecting whether or not a corresponding event takes place. In other words, for itemsets X and Y, the semantics of rule X⇒Y could be explained in a way as whether the events for X and X∪Y occur or not. However, in many real world applications, not only shall the occurrence of events be considered important, but also the quantity involved in the events. That is, we often encounter the cases where the values of attributes (such as Age, Income, Price, etc.) are continuous in the domains. These cases are usually dealt with in light of quantitative association rules (QAR) (such as “Milk[6,12] ⇒Bread[2,6]”), meaning that the customers who buy 6 to 12 bottles of milk may also tend to buy 2 to 6 pieces of bread at the same time) (Aumann & Lindell, 1999; Dougherty, Kohavi, & Sahami, 1995; Srikant & Agrawal, 1995). In doing so, typical approaches use sharp partitioning for the respective domains to generate several intervals, each then treated as a new attribute. Then, the original dataset with continuous attribute values could be transformed into a new dataset with binary
attribute values, on which the classical Apriori method can be used in finding such QARs. Another example of the sharp partitioning is on an Age domain (0,100) to form such new attributes as Age(0, 30], Age(30, 60], Age(60, 100).

Notably, though sharp partitioning is considered effective, it may suffer from the “boundary problem”, leading to misclassification (Bosc & Pivert, 2001; Chen, Wei, & Kerre, 2006; Chen, Yan, & Kerre, 2004; Chen, Yan, & Wei, 2009). That is to say, the intervals with sharp boundaries may result in over-sensitive grouping of attribute values near the boundaries, which then affect the classification outcomes. For instance, a transaction \( t \) with age of 31 will have the values 0, 1, 0 for attributes Age(0, 30], Age(30, 60], Age(60, 100), respectively, meaning that the event relating to Age(30,60] is relevant to transaction \( t \). On the other hand, age of 30 will be grouped to another attribute Age(0,30), though semantically age of 30 or 31 may not look like so different. In other words, the intervals resulting from sharp partitioning may amply (or over-emphasize) the difference between the values at boundaries.

In this regard, fuzzy logic (Zadeh, 1965) is considered useful to smooth the differences with gradual grades for the values around boundaries. This will extend the belongings of these values to attributes from \{0,1\} to [0,1]. An approach that is considered effective is to define fuzzy sets upon the sharply partitioned intervals (e.g., the crisp intervals used in classical classification methods such as CBA and GARC). Whereas there are different ways to define the fuzzy sets (e.g., parameterized by \( f_1 \) as the degree of fuzziness in (Xiong, 2010)), this chapter presents a fuzzy extension to GARC in that fuzzy sets are defined on the boundaries of the crisp intervals in consideration of a boundary value possessing the greatest degree of fuzziness. With the degree of fuzziness as parameterized by \( f_2 \), the fuzzy extension is then referred to as GARC\(_{f2}\).

The chapter is organized as follows. Section 1 introduces some preliminary notions on fuzzy sets. Section 2 elaborates on the steps of GARC\(_{f2}\) in constructing classifiers including the respective algorithm. Section 3 shows the experimental results and analysis on benchmarking datasets. Finally, section 4 illustrates the advantages of GARC\(_{f2}\) through a real-world application.

1 Preliminary Notions

Fuzzy sets theory provides a means to deal with imprecision and uncertainty (Zadeh, 1965). Let \( U \) be a domain, a fuzzy set \( A \) (sometimes labeled in a linguistic term) on \( U \) is characterized by a membership function \( \mu_A(u): U \to [0,1] \), where, for any \( u \) in \( U \), \( \mu_A(u) \) is the grade of membership of \( u \) for \( A \), representing the degree of \( u \) belonging to \( A \). For example, Young-Age, Middle-Age, Old-Age could be defined on Domain(Age) with the corresponding membership functions being trapezoid as shown in Figure 1.

Furthermore, fuzzy partitioning (Bezdek, 1981; Dunn, 1974) is used for transforming quantitative datasets to fuzzy datasets. Generally, suppose there are \( n \) transactions \{\( t_1, t_2, \ldots, t_m \)\} in an original dataset \( D \) with schema \( \mathbf{R}(A) \), where \( A \) is a set of \( m \) attributes, i.e., \( A = \{A_1, A_2, \ldots, A_m\} \), each \( A_k \) \((1 \leq k \leq m)\) can be associated
with \( q_k \) sets (i.e., intervals) defined on \( \text{domain}(A_k) \), denoted as \( \{ A_1^k, \ldots, A_{q_k}^k \} \). Traditionally, for each value \( t_i(A_k) (1 \leq i \leq n) \), i.e., value of \( t_i \) for attribute \( A_k \), it will totally belong to one of the sets \( \{ A_1^k, \ldots, A_{q_k}^k \} \). The result can be described by an \( n \times q_k \) matrix \( X_k \), where each element \( x_{ij} \) values 1 if \( x_{ij} \in A_j^k (1 \leq j \leq q_k) \) or 0 otherwise. In other words, the matrix \( X_k \) has the following properties: (1) \( x_{ij} \in \{0,1\} \), (2) \( \sum_{j=1}^{q_k} x_{ij} = 1, \forall i \), (3) \( 0 < \sum_{i=1}^{n} x_{ij} < n, \forall j \), then \( X_k \) is called to be a sharp partition of attribute \( A_k \). If the value of \( x_{ij} \) is not constrained in Boolean value \( \{0,1\} \) but in \([0,1]\), which means that \( t_i(A_k) \) can partially belong to one of the sets \( \{ A_1^k, \ldots, A_{q_k}^k \} \), then \( X_k \) is called to be a fuzzy partition of attribute \( A_k \) if (1) \( x_{ij} \in [0,1] \), (2) \( \sum_{j=1}^{q_k} x_{ij} = 1, \forall i \), (3) \( 0 < \sum_{i=1}^{n} x_{ij} < n, \forall j \).

For instance, given a dataset \( D \) with continuous attributes as shown in Table 1, the resultant datasets based on sharp partitioning and fuzzy partitioning are shown as Table 2 and Table 3.

**Table 1** Original Dataset \( D \)

<table>
<thead>
<tr>
<th>( D )</th>
<th>Age</th>
<th>Income</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>33</td>
<td>64000</td>
<td>( C_1 )</td>
</tr>
<tr>
<td>#2</td>
<td>25</td>
<td>26000</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#3</td>
<td>46</td>
<td>39000</td>
<td>( C_3 )</td>
</tr>
<tr>
<td>#4</td>
<td>62</td>
<td>22000</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#5</td>
<td>36</td>
<td>52000</td>
<td>( C_1 )</td>
</tr>
</tbody>
</table>
Table 2 Sharp-Partitioned Dataset \( D_s \)

<table>
<thead>
<tr>
<th>( D_s )</th>
<th>Age (0,35]</th>
<th>Age (35,65]</th>
<th>Age (65,100]</th>
<th>Income [0,25000]</th>
<th>Income (25000,60000]</th>
<th>Income (60000,( \infty ))</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>( C_1 )</td>
</tr>
<tr>
<td>#2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>( C_3 )</td>
</tr>
<tr>
<td>#4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>( C_1 )</td>
</tr>
</tbody>
</table>

Table 3 Fuzzy-Partitioned Dataset \( D_f \)

<table>
<thead>
<tr>
<th>( D_f )</th>
<th>Young-Age</th>
<th>Middle-Age</th>
<th>Old-Age</th>
<th>Low-Income</th>
<th>Middle-Income</th>
<th>High-Income</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.7</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.7</td>
<td>( C_1 )</td>
</tr>
<tr>
<td>#2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.6</td>
<td>0</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>( C_3 )</td>
</tr>
<tr>
<td>#4</td>
<td>0</td>
<td>0.2</td>
<td>0.8</td>
<td>0.8</td>
<td>0.2</td>
<td>0</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>#5</td>
<td>0.4</td>
<td>0.6</td>
<td>0</td>
<td>0</td>
<td>0.9</td>
<td>0.1</td>
<td>( C_1 )</td>
</tr>
</tbody>
</table>

Note that the fuzzy sets on the domain of \( Income \) for Table 3 are shown as defined in Figure 2.

![Fig. 2 Membership functions for Low-Income, Middle-Income and High-Income](image)

2 Gain Based Associative Classification with Fuzzy Partitioning

This section will introduce \( \text{GARC}_{f2} \) (Gain based Associative Classification with Fuzzy Partitioning in \( f_2 \)) in three steps: (1) transforming the traditional datasets with continuous attributes to fuzzy-partitioned datasets, (2) mining a compact set of fuzzy class association rules, and (3) building a classifier based on the discovered rules.
2.1 Fuzzy Partitioning with Trapezoidal Membership Functions

As mentioned in Section 2, fuzzy partitioning is to define a number of fuzzy sets, usually labeled by linguistic terms, upon each attribute domain. In this chapter, trapezoidal membership functions will be adopted, since they are not only easy to construct, but also convenient to illustrate the values close to the boundaries (Chen, 1998; Klir & Yuan, 1995). In addition, though represented in linear segments, they are sufficient for use in reflecting a wide variety of real-world semantics.

Concretely, a membership function concerned is of the form: \( \text{trapezia}(a, b, c, d) \) (denoted as \( \pi \) in brief), where \( a, b, c \) and \( d \) are four parameters. Furthermore, the membership functions are assumed to be normalized, i.e., for any \( \mu: \pi(a, b, c, d), \mu(b) = \mu(c) = 1 \). Moreover, given an attribute \( A \) in \( D \), suppose Domain(\( A \)) is fuzzy-partitioned into a series of \( v \) attributes \( A_1, A_2, \ldots, A_v \) in \( D_f \), where \( A_q (q = 1, 2, \ldots, v) \) is a fuzzy set with a trapezoidal membership function \( \mu_{A_q}: \pi(a_q, b_q, c_q, d_q) \), then \( A_1, A_2, \ldots, A_v \) could be semantically viewed as in an increasing order across the value range in Domain(\( A \)), e.g., denoted as \( A_q ≺ A_{q+1} \) (where \( c_q ≤ b_{q+1} \) and \( \mu_{A_q}(c_q) = \mu_{A_{q+1}}(b_{q+1}) = 1 \). For instance, \( \text{Low-Income} ≺ \text{High-Income} \), and \( \text{Young-Age} ≺ \text{Middle-Age} \). This represents a gradual smoothing around the juncture of two adjacent fuzzy sets (attributes/labels). To make \( \{A_1, A_2, \ldots, A_v\} \) a fuzzy partition on Domain(\( A \)), we shall also have \( a_q = c_{q-1}, b_q = d_{q-1}, (q = 2, 3, \ldots, v) \).

Thus, importantly, for any \( u \) in Domain(\( A \)), there exist two membership functions \( \mu_{A_q} \) and \( \mu_{A_{q+1}} \) defined on Domain(\( A \)) such that \( \mu_{A_q} ≺ \mu_{A_{q+1}} \) and \( \mu_{A_q}(u) + \mu_{A_{q+1}}(u) = 1 \).

In particular, given a domain \([p_0, p_v]\) with a crisp partitioning of \( v \) intervals \([p_0, p_1], [p_1, p_2], \ldots, [p_{v-1}, p_v]\), if a fuzzy partitioning with \( v \) attributes is still desired, then a trapezoidal membership function \( \pi(a_q, b_q, c_q, d_q) \) could be defined for each attribute \( A_q (q = 1, 2, \ldots, v) \) in \( D_f \) with respect to interval \([p_{q-1}, p_q]\). Moreover, a single parameter \( f_2 \) could be defined, as follows, to represent the degree of smoothing around the boundary of each crisp interval. In other words, \( f_2 \) could be viewed as the degree of fuzziness that \( A_q \) reflects for \([p_{q-1}, p_q]\) (See Figure 3).

\[
f_2 = \frac{p_{q-1} - a_q}{\min\left(p_{q-1} - p_{q-2}, p_q - p_{q-1}\right)} = \frac{b_q - p_{q-1}}{\min\left(p_{q-1} - p_{q-2}, p_q - p_{q-1}\right)} = \frac{p_q - c_q}{\min\left(p_q - p_{q-1}, p_{q+1} - p_q\right)} = \frac{d_q - p_q}{\min\left(p_q - p_{q-1}, p_{q+1} - p_q\right)}
\]

where \( q = 2, 3, \ldots, v-1 \). It is easy to see that \((a_q+b_q)/2=p_{q-1}, (c_q+d_q)/2=p_q\).
For example, for the fuzzy set $A_q$ defined on the domain of $A$, of which trapezoidal membership function is showed as $\mu_{A_q}: \pi(a_q, b_q, c_q, d_q)$, $x$ is the distance between $a_q$ and $p_{q-1}$, while $y$ and $z$ are the length of intervals $[p_{q-2}, p_{q-1}]$ and $[p_{q-1}, p_q]$, respectively (See Figure 3). Given parameter $f_2 = 0.25$, then we have $x / \min(y, z) = x / y = 0.25$.

It is worth mentioning that crisp intervals is a special case of trapezoidal membership functions, i.e., $\pi(b, b, c, c)$ with $b \neq c$. Apparently, the higher $f_2$ is, the fuzzier $A_q$ is. The range of $f_2$ is from 0 to 0.5 expressing an increasing degree of fuzziness, where $f_2 = 0$ represents a crisp interval (i.e., $a = b$ and $c = d$). An advantage of using $f_2$ is to only set a single parameter instead of setting many ones (e.g., $4 \times 4$), which is practically appealing.

### 2.2 Generating a Fuzzy Associative Classifier (GARC$_{f_2}$)

#### 2.2.1 Fuzzy Class Association Rules (Fuzzy CARs)

The fuzzy partitioned dataset gives rise to fuzzy association rules. In classification, it results in fuzzy class association rules. For example, compared with a classical CAR: $(\text{Age}(0,35], \text{Income}(25000, 60000]) \Rightarrow \text{Class 1}$ (meaning that the customers at ages of (0, 35] and with income of (25000, 60000] will be classified into Class 1), a fuzzy CAR may look like $(\text{Young-Age}, \text{Middle-Income}) \Rightarrow \text{Class 1}$ (meaning that the customers who are young and have middle income will be classified into Class 1), which is more general and is of natural-language nature.

Generally, for the original set of items/attributes $I = \{I_1, I_2, \ldots, I_m\}$ and database $D$ with schema $R(I)$, each $I_k (1 \leq k \leq m)$ can be associated with $q_k$ fuzzy sets defined on domain($I_k$), and usually labeled as $q_k$ new attributes in $D_f$. That is, the new database $D_f$ is with respect to schema $R(I_f)$ where $I_f = \{I_1^1, \ldots, I_1^{q_1}, \ldots, I_k^1, \ldots, I_k^{q_k}, \ldots, I_m^1, \ldots, I_m^{q_m}\}$. Let $G = \{C_1, C_2, \ldots, C_y\}$ be a set of class labels, then a fuzzy rule item is of the form $X \Rightarrow C$, where $X \subseteq I_f$, and $C \in G$. 

![Fig. 3 Sharp and Fuzzy Partitions with $f_2 = 0.25$](image-url)
In fuzzy context, the notions of degrees of support and confidence can be extended as follows (Chen & Wei, 2002; Chen et al., 2006). For any \( X = \{x_1, x_2, \ldots, x_p\} \subseteq I_f \cup \{C_j\} (j = 1, 2, \ldots, g) \) and transaction \( t \) in \( D_f \),

\[
Dsuppt(X) = \mu_X(t) = \min(t(x_1), t(x_2), \ldots, t(x_p)),
\]

\[
Dsupp(X) = \frac{\|X\|}{|D_f|} = \frac{\sum_{t} \text{count}(\mu_X(t))}{|D_f|},
\]

\[
Dsupp(X \Rightarrow C_i) = Dsupp(XC_i) = \frac{\|XC_i\|}{|D_f|} = \frac{\sum_{t : \text{class}=C_i} \text{count}(\mu_X(t))}{|D_f|},
\]

\[
Dconf(X \Rightarrow C_i) = \frac{\|XC_i\|}{\|X\|} = \frac{\sum_{t : \text{class}=C_i} \text{count}(\mu_X(t))}{\sum_{t} \text{count}(\mu_X(t))},
\]

where \( t(x_i) (i = 1, 2, \ldots, p) \) is the corresponding membership degree of \( t \) for attribute \( x_i \), \( |D_f| \) is the cardinality of \( D_f \) (i.e., the number of transactions in \( D_f \)), and \( \|\|X\| \) is the fuzzy cardinality of \( X \) calculated using \( \Sigma \text{count} \) operator. Note that t-norm minimum is used in the definition of \( Dsuppt(X) \) due to its semantics. Other t-norms may also be considered under specific application contexts.

Take Table 3 for example, \( Dsupp_{\text{th}}(\text{Young-Age, Middle-Income}) = \min(0.7, 0.3) = 0.3 \). Moreover, \( Dsupp(\text{Young-Age, Middle-Income}) = (\min(0.7, 0.3) + \min(1, 0.6) + \min(0, 1) + \min(0.2) + \min(0.4, 0.9))/5 = 1.3/5 = 0.26 \), \( Dsupp(\text{Young-Age, Middle-Income} \Rightarrow C_1) = (\min(0.7, 0.3) + \min(0.4, 0.9))/5 = 0.7/5 = 0.14 \), then \( Dconf((\text{Young-Age, Middle-Income}) \Rightarrow C_1) = 0.14/0.26 = 0.54 \).

Then, given minimal support \( \alpha \) and minimal confidence \( \beta \), with \( \alpha, \beta \in [0, 1] \), a fuzzy rule item \( X \Rightarrow C_i \) is called frequent if \( Dsupp(X \Rightarrow C_i) \geq \alpha \). Importantly, if fuzzy rule item \( X \Rightarrow C_i \) is not frequent, then any fuzzy rule item of the form \( Z \Rightarrow C_i (Z \supset X) \) is also infrequent, this is because \( \mu_Z(t) \leq \mu_X(t) \) for each \( t \), and consequently we have:

\[
Dsupp(Z \Rightarrow C_i) = \frac{\sum_{t : \text{class}=C_i} \text{count}(\mu_Z(t))}{|D_f|} \leq \frac{\sum_{t : \text{class}=C_i} \text{count}(\mu_X(t))}{|D_f|} = Dsupp(X \Rightarrow C_i)
\]

A fuzzy rule item \( X \Rightarrow C_i \) is called a valid fuzzy class association rule (fuzzy CAR) if \( Dsupp(X \Rightarrow C_i) \geq \alpha \) and \( Dconf(X \Rightarrow C_i) \geq \beta \). It is worthwhile to indicate that, since crisp intervals are special cases of fuzzy sets (e.g., an interval \([a, b]\) can be represented by a fuzzy set \( \int_{u=a}^{u=b} 1.0 / u \)), quantitative association rules are special cases of fuzzy association rules (with fuzzy partitioning), and therefore, CARs are special cases of fuzzy CARs.
2.2.2 Fuzzy Information Gain

Information gain is one of the measures used in decision tree classification methods for selecting best split attributes (Breiman, 1984; Weiss & Kulikowski, 1991). It is also used as a measure to reduce the number of itemsets in some associative classification (e.g. GARC).

In GARC, information gain measure is used to reduce the search space in generating candidate itemsets as well as rules. That is to say, only those candidate itemsets containing the so-called best split attribute value will be generated. This will significantly help improve the rule set without loss of accuracy (Chen et al., 2006). In the fuzzy context with $f_2$, we extend the fuzzy information entropy of a fuzzy sub-dataset $S$ in $D_f$ with class labels as follows (De Luca & Termini, 1972):

$$\text{info}(S) = \sum_{p=1}^{g} \frac{\sum_{\text{count}(C_p, S)}}{|S|} \times \log \left( \frac{\sum_{\text{count}(C_p, S)}}{|S|} \right)$$

where $\sum_{\text{count}(C_p, S)}$ represents the fuzzy cardinality of transactions in $S$ with class label $C_p$. In addition, given an attribute $A$ in $D$, without loss of generality, suppose $\text{Domain}(A)$ is partitioned into $v$ attributes each being a fuzzy set on $\text{Domain}(A)$, e.g., $A_1, A_2, \ldots, A_v$, and $D_{f_i}^{A_q}$ represents the fuzzy sub-dataset in $D_f$ in which each transaction $t$ has $\mu_{A_q}(t) > 0$, for $q = 1, 2, \ldots, v$, then the fuzzy information entropy of $A$ could be calculated below:

$$\text{info}_A(D_f) = \sum_{q=1}^{v} \frac{|D_{f_i}^{A_q}|}{|D_f|} \times \text{info}(D_{f_i}^{A_q})$$

$$= -\sum_{q=1}^{v} \| A_q \| \times \left( \sum_{p=1}^{g} \frac{\sum_{\text{count}(C_p, D_{f_i}^{A_q})}}{|D_{f_i}^{A_q}|} \times \log \left( \frac{\sum_{\text{count}(C_p, D_{f_i}^{A_q})}}{|D_{f_i}^{A_q}|} \right) \right)$$

$$= -\sum_{q=1}^{v} \text{Dsupp}(A_q) \times \left( \sum_{k=1}^{g} \text{Dconf}(A_q \Rightarrow C_k) \times \log \left( \text{Dconf}(A_q \Rightarrow C_k) \right) \right)$$

Hence, the fuzzy information gain for attribute $A$ is obtained as follows:

$$\text{gain}(A) = \text{info}(D_f) - \text{info}_A(D_f)$$

2.2.3 Pruning Strategies

As indicated previously, building an associative classifier needs to deal with the situation where there exist redundant and conflicting rules. This is also the case in generating fuzzy CARs. For example, suppose we already have High_Income $\Rightarrow C_2$, then (Young_Age, High_Income) $\Rightarrow C_2$ might be redundant; and High_Income $\Rightarrow C_3$ might be conflicting. Rather than filtering these redundant and conflicting
rules after generating all fuzzy CARs, a more effective and desirable way is to incorporate some pruning strategies in the process of generating itemsets as well as CARs, resulting in fewer itemsets (therefore rules) to be generated.

First of all, as in GARC, the strategy could be applied that if $X \Rightarrow C_i$ holds, then any candidate itemsets containing $XC_i$ need not to be generated and tested. In this way, the number of candidate itemsets to be generated is reduced. Moreover, the following strategies would be helpful in pruning candidate rule items.

**Strategy 1.** Given a rule item $X \Rightarrow C_i$ in $D_f$, if the condition

$$\text{Dsupp}(X) \times (1 - \text{Dconf}(X \Rightarrow C_i)) < \alpha$$

is satisfied, then rule items of the form $Z \Rightarrow C_j$ ($Z \supseteq X, j \neq i$) will not hold (i.e., not valid) in $D_f$.

**Proof.** Since $\text{Dsupp}(Z \Rightarrow C_j) \leq \text{Dsupp}(X \Rightarrow C_i)$, we only need to prove that $X \Rightarrow C_j$ ($j \neq i$) is infrequent.

As each $t$ in $D_f$ will be assigned with a unique class label $C_i$, $i = 1, 2, \ldots, g$, we have

$$\|X\| = \sum_t \text{count}(\mu_X(t)) = \sum_{rt.\text{class} = C_i} \text{count}(\mu_X(t)) + \sum_{rt.\text{class} \neq C_i} \text{count}(\mu_X(t)), \forall i,$$

which means:

$$\|X\| = \sum_{rt.\text{class} = C_i} \text{count}(\mu_X(t)) + \sum_{j \neq i} \left( \sum_{rt.\text{class} = C_j} \text{count}(\mu_X(t)) \right) = \|XC_i\| + \sum_{j \neq i} \left( \|XC_j\| \right).$$

Further, from

$$\|X\| = \frac{\|XC_i\|}{|D_f|} \sum_{j \neq i} \frac{\|XC_j\|}{|D_f|},$$

we have:

$$\sum_{j \neq i} \text{Dconf}(X \Rightarrow C_j) = 1 - \text{Dconf}(X \Rightarrow C_i).$$

Thus, $\text{Dconf}(X \Rightarrow C_j) \leq 1 - \text{Dconf}(X \Rightarrow C_i), \forall j \neq i$.

Since

$$\text{Dsupp}(X \Rightarrow C_j) = \frac{\|XC_j\|}{|D_f|} \times \frac{\|XC_j\|}{\|X\|} = \frac{\text{Dsupp}(X)}{|D_f|} \times \text{Dconf}(X \Rightarrow C_j),$$

we have $\text{Dsupp}(X \Rightarrow C_j) \leq \text{Dsupp}(X) \times (1 - \text{Dconf}(X \Rightarrow C_i)) \times \alpha$, which means $X \Rightarrow C_j$ is infrequent, and therefore will not hold in $D_f$. \qed

**Strategy 2.** Suppose rule $X \Rightarrow C_i$ holds in $D_f$, if the condition

$$\text{Dsupp}(X) < \max \left( \frac{1}{1 - \beta}, 2 \right) \times \alpha$$

is satisfied, then rule items of the form $Z \Rightarrow C_j$ ($Z \supseteq X, j \neq i$) will not hold in $D_f$. 

Proof. Similarly, we only need to prove that \( X \Rightarrow C_j (j \neq i) \) is infrequent. The following proof consists of two situations:

1. If \( \beta \leq 0.5 \), then \( \frac{1}{1-\beta} \leq 2 \). In this situation, the condition is equivalent to

\[
\text{Dsupp}(X) < 2\alpha.
\]

Since
\[
\|X\| = \sum_{\text{tr.class} = \text{C}_i} \text{count}(\mu_X(t)) + \sum_{j \neq i} \left( \sum_{\text{tr.class} = \text{C}_j} \text{count}(\mu_X(t)) \right) = \|X\| + \sum_{j \neq i} \left( \|X\| \right) ,
\]

then
\[
\|X\| = \|X\| + \sum_{j=1}^{g} \|X\| = \sum_{j=1}^{g} \|X\|.
\]

Thus, we have
\[
\sum_{j=1}^{g} \text{Dsupp}(X \Rightarrow C_j) = \text{Dsupp}(X) - \text{Dsupp}(X \Rightarrow C_i) .
\]

Since \( X \Rightarrow C_i \) holds in \( D_f \), then \( \text{Dsupp}(X \Rightarrow C_i) \geq \alpha \). Further, we have
\[
\text{Dsupp}(X \Rightarrow C_j) \leq \text{Dsupp}(X) - \text{Dsupp}(X \Rightarrow C_i) < 2\alpha - \alpha = \alpha, \forall j \neq i .
\]

So \( X \Rightarrow C_j \) is infrequent, thus will not hold in \( D_f \).

2. If \( \beta > 0.5 \), then \( \frac{1}{1-\beta} \leq 2 \). In this situation, the condition is equivalent to

\[
(1-\beta) \text{Dsupp}(X) < \alpha .
\]

Since \( X \Rightarrow C_i \) holds in \( D_f \), then \( \text{Dconf}(X \Rightarrow C_i) \geq \beta \). Consequently, we have
\[
\text{Dsupp}(X) \times (1 - \text{Dconf}(X \Rightarrow C_i)) \leq (1-\beta) \text{Dsupp}(X) < \alpha ,
\]

which means the condition in Strategy 1 is satisfied. Therefore, \( X \Rightarrow C_j \) is infrequent, and will not hold in \( D_f \).

Importantly, these strategies will be incorporated into the mining process as pruning strategies (rather than post-process filtering strategies) to reduce the number of candidate itemsets (rule items) generated.

Finally, with the built classifier, a newly coming transaction \( t \) could be classified using a particular fuzzy CAR: \( X \Rightarrow C_i \), chosen in the classifier. That requires a match between \( t \) and \( X \), which can be determined with a measure called weighted confidence, denoted as \( \text{Wconf}_f(X \Rightarrow C_i) \) (Chen & Chen, 2008):
\[
\text{Wconf}_f(X \Rightarrow C_i) = \text{Dsupp}(X) \times \text{Dconf}(X \Rightarrow C_i) .
\]

For each fuzzy CAR, there is a \( \text{Wconf} \) value for \( t \), then the fuzzy CAR with the highest \( \text{Wconf} \) value will be used to classify \( t \).

2.3 The GARC\textsubscript{f2} Algorithm

The algorithmic detail of the fuzzy extension to GARC with \( f_2 \) (i.e., GARC\textsubscript{f2}) is provided in its general form as shown in Figure 4.
1. \( D_1 = \text{Trapezia\_Fuzzy}(D) \); // Fuzzy partitioning with \( f_2 \)
2. \( f\text{CAR}_1 = \{ r \mid r \text{ is an 1-itemset, } D\text{supp}(r) \geq \alpha \text{ and } D\text{conf}(r) \geq \beta \} \); // Initial 1-itemset rule set
3. \( \text{prfCAR}_1 = \text{prunRules}(f\text{CAR}_1) \); // Deleting redundant rules and conflicting rules
4. \( F_1 = \{ p \mid p \text{ is an 1-itemset, } D\text{supp}(p) \geq \alpha \text{ and } p \text{ is not in } \text{prfCAR}_1 \text{ set} \} \) // Strategy 1
5. \( \text{bestattr} = \text{gain}() \); // function \( \text{gain}() \) returns best attribute with maximal gain
6. for \( k \) from 2 to \( m \) do
7. \( C_k = \text{CandidateGen}(F_{k-1}, \text{bestattr}) \); // Strategy 2
8. \( f\text{CAR}_k = \{ r \mid r \in C_k, D\text{supp}(r) \geq \alpha, D\text{conf}(r) \geq \beta \} \);
9. \( \text{prfCAR}_k = \text{prunRules}(f\text{CAR}_k) \); // Deleting redundant rules and conflicting rules
10. \( F_k = \{ p \mid p \text{ is a } k\text{-itemset, } D\text{supp}(p) \geq \alpha \text{ and } p \text{ is not in } \text{prfCAR}_k \text{ set} \} \) // Strategy 1
11. end for;
12. \( \text{CompactfCARs} = \bigcup_k(\text{prfCAR}_k) \); // Generating a Compact Set of Fuzzy CARs
13. Sort(\text{CompactfCARs});

Fig. 4 The \( \text{GARC}_{12} \) algorithm

In Figure 4, line 1 performs the fuzzy partitioning operation based on a given parameter \( f_2 \), so that the original dataset \( D \) is then transformed to the fuzzy-partitioned dataset \( D_1 \). Lines 2-5 scan the dataset to measure all 1-itemsets rule items and then to get the best split attribute by fuzzy information gain calculation. Rule items that satisfy the minimum support and minimum confidence criteria are written into the set of fuzzy CARs (line 2). The generated 1-itemset rule set (called \( f\text{CAR}_1 \)) is subject to a pruning operation (line 3), which deletes the redundant rules and conflicting rules. The pruning operation is also done in each subsequent pass of \( f\text{CAR}_k \) (line 9). Lines 6-11 consecutively run operations of generation, scan and filtering to discover fuzzy CARs without rule redundancy and conflicts. For each subsequent pass, say pass \( k \), frequent \((k-1)\)-itemsets that are not in the pruned rule set \( F_{k-1} \) (line 4 and line 10) are used to generate candidate rule items in the next loop, according to Strategy 1, to avoid generating redundant rules. The \text{CandidateGen} function generates candidate rule items that include the best split attribute (line 7). Strategy 2 is applied here for avoid conflicting rules so as to reduce the number of candidate rules. The generated fuzzy CARs of each pass, with rule redundancy and rule conflicts eliminated, construct the compact set of fuzzy CARs (line 12). Finally, all the rules in the compact set will be sorted with precedence (line 13). In addition, the algorithm will terminate when \( F_{k-1} = \emptyset \) or otherwise, in at most \( m \) passes, which should be less than the number of attributes in \( D \).
3 Data Experiments

This section shows the experimental results for the proposed GARC\textsubscript{f2} algorithm, along with some comparisons with GARC. Briefly, the results show that GARC\textsubscript{f2} is at the same level of accuracy as GARC statistically, while the number of rules is significantly fewer than that of GARC.

The experimental environment is with Window XP, Intel Core Duo 3GHz, 2G RAM and MatLab 2009a. A commonly used benchmarking database in this field, namely the UCI Machine Learning Repository (Merz & Murphy, 1996) is used, including the 20 datasets that GARC selected with continuous attributes. Table 4 provides the basic information about the datasets, each containing two parts: a training data part and a testing data part. The former was used for generating the classifiers while the latter was used for testing purposes.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Attributes</th>
<th>Number of attributes</th>
<th>Null Value (Y/N)</th>
<th>Number of Classes</th>
<th>Number of training data</th>
<th>Number of testing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anneal</td>
<td>Discrete, continuous</td>
<td>38</td>
<td>Y</td>
<td>6</td>
<td>598</td>
<td>300</td>
</tr>
<tr>
<td>2. Australian</td>
<td>Discrete, continuous</td>
<td>14</td>
<td>N</td>
<td>2</td>
<td>460</td>
<td>230</td>
</tr>
<tr>
<td>3. Balance</td>
<td>Continuous</td>
<td>4</td>
<td>N</td>
<td>3</td>
<td>416</td>
<td>209</td>
</tr>
<tr>
<td>4. Breast</td>
<td>Continuous</td>
<td>10</td>
<td>Y</td>
<td>2</td>
<td>466</td>
<td>233</td>
</tr>
<tr>
<td>5. Cleve</td>
<td>Discrete, continuous</td>
<td>13</td>
<td>Y</td>
<td>2</td>
<td>202</td>
<td>101</td>
</tr>
<tr>
<td>6. Crx</td>
<td>Discrete, continuous</td>
<td>15</td>
<td>Y</td>
<td>2</td>
<td>460</td>
<td>230</td>
</tr>
<tr>
<td>7. Diabetes</td>
<td>Continuous</td>
<td>8</td>
<td>N</td>
<td>2</td>
<td>512</td>
<td>256</td>
</tr>
<tr>
<td>8. Glass</td>
<td>Continuous</td>
<td>9</td>
<td>N</td>
<td>7</td>
<td>142</td>
<td>72</td>
</tr>
<tr>
<td>9. Heart</td>
<td>Discrete, continuous</td>
<td>13</td>
<td>N</td>
<td>2</td>
<td>180</td>
<td>90</td>
</tr>
<tr>
<td>10. Hepatitis</td>
<td>Discrete, continuous</td>
<td>19</td>
<td>Y</td>
<td>2</td>
<td>103</td>
<td>52</td>
</tr>
<tr>
<td>11. Hypothyroid</td>
<td>Discrete, continuous</td>
<td>25</td>
<td>Y</td>
<td>2</td>
<td>2108</td>
<td>1055</td>
</tr>
<tr>
<td>12. Ionosphere</td>
<td>Continuous</td>
<td>34</td>
<td>Y</td>
<td>2</td>
<td>234</td>
<td>117</td>
</tr>
<tr>
<td>13. Iris</td>
<td>Continuous</td>
<td>4</td>
<td>N</td>
<td>3</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>14. Labor</td>
<td>Discrete, continuous</td>
<td>16</td>
<td>Y</td>
<td>2</td>
<td>38</td>
<td>19</td>
</tr>
<tr>
<td>15. Pima</td>
<td>Continuous</td>
<td>8</td>
<td>N</td>
<td>2</td>
<td>512</td>
<td>256</td>
</tr>
<tr>
<td>16. Sick</td>
<td>Discrete, continuous</td>
<td>29</td>
<td>Y</td>
<td>2</td>
<td>2515</td>
<td>1257</td>
</tr>
<tr>
<td>17. Sonar</td>
<td>Continuous</td>
<td>60</td>
<td>N</td>
<td>2</td>
<td>138</td>
<td>70</td>
</tr>
<tr>
<td>18. Vehicle</td>
<td>Continuous</td>
<td>18</td>
<td>N</td>
<td>4</td>
<td>564</td>
<td>282</td>
</tr>
<tr>
<td>19. Waveform</td>
<td>Continuous</td>
<td>40</td>
<td>N</td>
<td>3</td>
<td>3333</td>
<td>1667</td>
</tr>
<tr>
<td>20. Wine</td>
<td>Continuous</td>
<td>13</td>
<td>N</td>
<td>3</td>
<td>118</td>
<td>60</td>
</tr>
</tbody>
</table>

3.1 Accuracy

Accuracy, as one of the basic and important measures for classifiers, is the ratio of the number of cases truly predicted by the classifier over the total number of cases
in the test dataset. In this experiment, GARC$_{f_2}$ with different $f_2$ values is compared to GARC (at $\alpha = 0.01$ and $\beta = 0.7$ as in GARC (Chen et al., 2006)). The results are tabulated in Table 5.

Table 5 Accuracies of GARC and GARC$_{f_2}$ ($f_2 = 0.1, 0.2, 0.3, 0.4, 0.5$)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>GARC</th>
<th>GARC$_{f_2}$ ($f_2=0.1$)</th>
<th>GARC$_{f_2}$ ($f_2=0.2$)</th>
<th>GARC$_{f_2}$ ($f_2=0.3$)</th>
<th>GARC$_{f_2}$ ($f_2=0.4$)</th>
<th>GARC$_{f_2}$ ($f_2=0.5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anneal</td>
<td>87.00</td>
<td>88.33</td>
<td>87.33</td>
<td>87.33</td>
<td>87.33</td>
<td>87.33</td>
</tr>
<tr>
<td>2. Australian</td>
<td>87.39</td>
<td>87.39</td>
<td>87.39</td>
<td>87.39</td>
<td>87.39</td>
<td>87.39</td>
</tr>
<tr>
<td>3. Balance</td>
<td>71.29</td>
<td>71.29</td>
<td>71.29</td>
<td>71.29</td>
<td>71.29</td>
<td>71.29</td>
</tr>
<tr>
<td>4. Breast</td>
<td>94.42</td>
<td>94.42</td>
<td>94.42</td>
<td>93.99</td>
<td>94.42</td>
<td>94.42</td>
</tr>
<tr>
<td>5. Cleve</td>
<td>66.34</td>
<td>69.31</td>
<td>69.31</td>
<td>69.31</td>
<td>69.31</td>
<td>69.31</td>
</tr>
<tr>
<td>6. Crx</td>
<td>85.65</td>
<td>85.65</td>
<td>85.65</td>
<td>85.22</td>
<td>85.22</td>
<td>85.22</td>
</tr>
<tr>
<td>7. Diabetes</td>
<td>67.97</td>
<td>68.36</td>
<td>68.75</td>
<td>67.58</td>
<td>63.67</td>
<td>64.06</td>
</tr>
<tr>
<td>8. Glass</td>
<td>62.50</td>
<td>62.50</td>
<td>62.50</td>
<td>61.11</td>
<td>51.39</td>
<td>55.56</td>
</tr>
<tr>
<td>9. Heart</td>
<td>78.89</td>
<td>77.78</td>
<td>77.78</td>
<td>77.78</td>
<td>77.78</td>
<td>77.78</td>
</tr>
<tr>
<td>10. Hepatitis</td>
<td>86.54</td>
<td>86.54</td>
<td>86.54</td>
<td>86.54</td>
<td>86.54</td>
<td>86.54</td>
</tr>
<tr>
<td>11. Hypothyroid</td>
<td>94.79</td>
<td>94.79</td>
<td>94.79</td>
<td>94.79</td>
<td>94.79</td>
<td>94.79</td>
</tr>
<tr>
<td>12. Ionosphere</td>
<td>91.45</td>
<td>93.16</td>
<td>93.16</td>
<td>92.31</td>
<td>92.31</td>
<td>90.60</td>
</tr>
<tr>
<td>13. Iris</td>
<td>94.00</td>
<td>94.00</td>
<td>94.00</td>
<td>94.00</td>
<td>94.00</td>
<td>94.00</td>
</tr>
<tr>
<td>14. Labor</td>
<td>84.21</td>
<td>84.21</td>
<td>84.21</td>
<td>84.21</td>
<td>84.21</td>
<td>84.21</td>
</tr>
<tr>
<td>15. Pima</td>
<td>76.17</td>
<td>75.39</td>
<td>75.39</td>
<td>74.22</td>
<td>74.61</td>
<td>75.39</td>
</tr>
<tr>
<td>16. Sick</td>
<td>93.79</td>
<td>93.79</td>
<td>93.79</td>
<td>93.79</td>
<td>93.79</td>
<td>93.79</td>
</tr>
<tr>
<td>17. Sonar</td>
<td>74.29</td>
<td>70.00</td>
<td>71.43</td>
<td>71.43</td>
<td>71.43</td>
<td>74.29</td>
</tr>
<tr>
<td>18. Vehicle</td>
<td>60.99</td>
<td>61.35</td>
<td>61.70</td>
<td>58.87</td>
<td>57.80</td>
<td>58.87</td>
</tr>
<tr>
<td>19. Waveform</td>
<td>73.73</td>
<td>73.67</td>
<td>71.51</td>
<td>71.51</td>
<td>71.63</td>
<td>71.57</td>
</tr>
<tr>
<td>20. Wine</td>
<td>83.33</td>
<td>91.67</td>
<td>93.33</td>
<td>90.00</td>
<td>90.00</td>
<td>91.67</td>
</tr>
<tr>
<td>Mean</td>
<td>80.74</td>
<td>81.18</td>
<td>81.21</td>
<td>80.80</td>
<td>79.95</td>
<td>80.40</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>11.01</td>
<td>11.37</td>
<td>11.39</td>
<td>11.80</td>
<td>12.99</td>
<td>12.29</td>
</tr>
</tbody>
</table>

For illustrative purposes, Figure 5 depicts the mean accuracy with respect to different $f_2$ values. It is worth noticing that, at $f_2 = 0$, GARC$_{f_2}$ degenerates to GARC. Furthermore, Table 6 lists the results of significance tests on pairwise mean difference for GARC and GARC$_{f_2}$ ($f_2 = 0.1, 0.2, 0.3, 0.4, 0.5$) with t-test, which indicates that on average the accuracy of GARC$_{f_2}$ (though varying with $f_2$) is not significantly different from that of GARC.
3.2 Number of Rules

As stated in previous discussions, the number of rules reflects the understandability/compactness of a classifier. Briefly, the smaller the number of rules is, the better understandability a classifier obtains (e.g., at the same accuracy level). The introduction of fuzzy partitioning could express appropriate semantics of data, which may lead to a decrease in the number of rules and therefore an increase in the understandability. The experiments on the number of rules as listed in Table 7 confirmed this statement.
Table 7 Number of Rules of GARC and GARC_{f_2} (f_2 = 0.1, 0.2, 0.3, 0.4, 0.5)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>GARC</th>
<th>GARC_{f_2} (f_2=0.1)</th>
<th>GARC_{f_2} (f_2=0.2)</th>
<th>GARC_{f_2} (f_2=0.3)</th>
<th>GARC_{f_2} (f_2=0.4)</th>
<th>GARC_{f_2} (f_2=0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Anneal</td>
<td>51</td>
<td>48</td>
<td>47</td>
<td>47</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>2. Australian</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>3. Balance</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4. Breast</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>5. Cleve</td>
<td>15</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>6. Crx</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>7. Diabetes</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>8. Glass</td>
<td>21</td>
<td>21</td>
<td>17</td>
<td>14</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>9. Heart</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>10. Hepatitis</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>11. Hypothyroid</td>
<td>43</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>12. Ionosphere</td>
<td>67</td>
<td>67</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>13. Iris</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>14. Labor</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>15. Pima</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>16. Sick</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>17. Sonar</td>
<td>16</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>18. Vehicle</td>
<td>122</td>
<td>102</td>
<td>86</td>
<td>82</td>
<td>68</td>
<td>78</td>
</tr>
<tr>
<td>19. Waveform</td>
<td>33</td>
<td>33</td>
<td>27</td>
<td>27</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>20. Wine</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

Mean: 28.35, 27.15, 25.45, 24.95, 23.90, 24.20

Similarly, Figure 6 represents the mean number of rules with respect to different \( f_2 \) values. Again, at \( f_2 = 0 \), GARC_{f_2} degenerates to GARC.

Next, statistical tests had been performed to examine the significance on the difference of number of rules of GARC and GARC_{f_2} (\( f_2 = 0.1, 0.2, 0.3, 0.4, 0.5 \)) using the non-parametric Friedman test (Conover, 1999; Vapnik, 1998), with the results shown in Table 8, revealing the advantage of GARC_{f_2} over GARC in number of rules. That is, in general, the number of rules in the GARC_{f_2} classifier was fewer than that of GARC, which appeared to be significant in the case of \( f_2 > 0.1 \) (while insignificant at \( f_2 = 0.1 \)).
Fig. 6 Mean Number of Rules of GARC, GARC_f (f = 0.1, 0.2, 0.3, 0.4, 0.5)

Table 8 Significance Test on Pairwise Difference of Number of Rules of GARC and GARC_f

<table>
<thead>
<tr>
<th>Friedman Test</th>
<th>95%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARC_f (f = 0.1) – GARC</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>GARC_f (f = 0.2) – GARC</td>
<td>Yes (-)*</td>
<td>Yes (-)</td>
</tr>
<tr>
<td>GARC_f (f = 0.3) – GARC</td>
<td>Yes (-)</td>
<td>Yes (-)</td>
</tr>
<tr>
<td>GARC_f (f = 0.4) – GARC</td>
<td>Yes (-)</td>
<td>Yes (-)</td>
</tr>
<tr>
<td>GARC_f (f = 0.5) – GARC</td>
<td>Yes (-)</td>
<td>Yes (-)</td>
</tr>
</tbody>
</table>

*Note. Yes (-) for A – B represents that number of rules of A is significantly smaller than that of B.

To summarize, from the data experiments, in a statistically significant sense, the accuracy performance of GARC_f is as good as that of GARC, and the compactness/understandability performance of GARC_f is generally better than that of GARC. Overall, these are deemed desirable.

4 A Case Study: Bank Credit Ratings

This section examines the effectiveness of GARC_f in a real-world classification application, namely, bank credit rating.
The data were collected from a rating institution in China, composed of two sets: one is 140 credit rating records of bank G in years 2000-2002 as the training sample, the other is 104 credit rating records of banks Z, J, X and R in years 1999-2002 as the testing sample. The ratings consist of three categories with a total of nine levels, i.e., AAA, AA, A, BBB, BB, B, CCC, CC, C. Further, in usual practice, they were grouped into two classes: Investment Grade (for the ratings of BBB or higher) and Speculative Grade (for the ratings of BB or lower). Tables 9 and 10 show the data and their attributes.

Table 9 Class distribution in the credit rating dataset

<table>
<thead>
<tr>
<th>Class</th>
<th>No. of records in Training Dataset</th>
<th>No. of records in Testing Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment Grade</td>
<td>88</td>
<td>72</td>
</tr>
<tr>
<td>Speculative Grade</td>
<td>52</td>
<td>32</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>104</td>
</tr>
</tbody>
</table>

Table 10 Attributes of the credit rating dataset

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Crisp Boundaries*</th>
<th>Fuzzy Partitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1: Net asset to loan for the period ending at Dec 31</td>
<td>65.985</td>
<td>Low, High</td>
</tr>
<tr>
<td>A2: Asset to liability ratio</td>
<td>69.72</td>
<td>Low, High</td>
</tr>
<tr>
<td>A3: Capital expenditure to equity ratio</td>
<td>130.275, 203.85</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A4: Current ratio</td>
<td>95.965, 139.385</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A5: Acid ratio</td>
<td>62.845, 106.96</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A6: Net non-financing cash inflow to liquidity liability ratio</td>
<td>-12.755, 9.785</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A7: Operating cash inflow to liquidity liability ratio</td>
<td>-11.97, 7.305</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A8: Account receivable turnover ratio</td>
<td>5.03</td>
<td>Low, High</td>
</tr>
<tr>
<td>A9: Inventory turnover ratio</td>
<td>1.34, 3.555</td>
<td>Low, Medium, High</td>
</tr>
<tr>
<td>A10: Fixed asset turnover ratio</td>
<td>2.025</td>
<td>Low, High</td>
</tr>
<tr>
<td>A11: Gross profit margin</td>
<td>23.035</td>
<td>Low, High</td>
</tr>
<tr>
<td>A12: Operating income margin</td>
<td>0.305</td>
<td>Low, High</td>
</tr>
<tr>
<td>A13: ROE</td>
<td>0.745</td>
<td>Low, High</td>
</tr>
<tr>
<td>A14: Return on total assets</td>
<td>3.975</td>
<td>Low, High</td>
</tr>
</tbody>
</table>

*Note. Crisp boundaries are the results of the discretization of continuous attributes, done by the Entropy method in (Fayyad & Irani, 1993), which is used in CBA/GARC. Crisp boundaries for each attribute correspond to a group of crisp partitions. For example, the crisp boundaries of attribute A4 correspond to crisp partitions: A4<95.965, 95.965< A4<139.385, A4>139.385.
In Table 10, fuzzy partitions were defined upon the intervals with crisp boundaries determined in CBA/GARC, resulting in respective fuzzy sets such as high, medium or low using the trapezoidal membership functions discussed in previous sections. Subsequently, GARC and GARC_{f2} were applied to the data with the settings of $\alpha$ and $\beta$ in order to get a classifier at the highest level of accuracy then smallest number of rules for each dataset (see Figure 7).

It can be seen from Figure 6 that at the best cases GARC_{f2} was advantageous over GARC in accuracy. Furthermore, looking at the case with $f_2 = 0.5$, we could reach a situation where GARC_{f2} performed better than GARC in accuracy and the number of rules (Table 11).

**Table 11 Performances of GARC and GARC_{f2}**

<table>
<thead>
<tr>
<th></th>
<th>Highest accuracy</th>
<th>No. of rules</th>
<th>Parameter Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARC</td>
<td>85.58%</td>
<td>15</td>
<td>$\alpha$ $\beta$</td>
</tr>
<tr>
<td>GARC_{f2} ($f_2 = 0.5$)</td>
<td>95.19%</td>
<td>14</td>
<td>0.01 0.90</td>
</tr>
</tbody>
</table>

Finally, for illustrative purposes, the fuzzy CARs generated by CARC_{f2} in this application are exemplified in Table 12 (with $f_2=0.5$).
Table 12 Rules in the classifier built by GARC$_{f2}$ ($f_2 = 0.5$)

<table>
<thead>
<tr>
<th>Rule</th>
<th>Condition</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Capital expenditure to equity ratio is low) $\land$ (Account receivable turnover ratio is high)</td>
<td>Investment Grade</td>
</tr>
<tr>
<td>2</td>
<td>(Current ratio is high)</td>
<td>Investment Grade</td>
</tr>
<tr>
<td>4</td>
<td>(Acid ratio is high)</td>
<td>Investment Grade</td>
</tr>
<tr>
<td>6</td>
<td>(Inventory turnover ratio is low)</td>
<td>Speculative Grade</td>
</tr>
</tbody>
</table>

5 Conclusion

This chapter has presented a fuzzy extension to a compact and accurate associative classifier (namely GARC) by introducing fuzzy sets on the attribute domains in order to smooth the belongings of the data values around the boundaries of crisp intervals. This has been done via fuzzy partitioning using trapezoidal membership functions, designed with a single parameter $f_2$ in reflecting the degree of fuzziness around the boundaries, giving rise to the GARC$_{f2}$ approach. Moreover, the notions of support, confidence, information gain, rule redundancy and conflicts have also been extended in the fuzzy context, including the incorporation of pruning strategies in the rule generation process. This incorporation as well as the extended information gain have been proven to be important for compactness (i.e., with a reduced number of fuzzy CARs). Benchmarking datasets showed that on average GARC$_{f2}$ had accuracy similar to that of GARC, and was significantly advantageous over GARC in compactness. Finally, a real application on bank credit ratings also demonstrated the effectiveness of the extended approach.

Acknowledgements. The work was partly supported by the National Natural Science Foundation of China (70890080), the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities of China (07JJD63005) and Tsinghua University’s Research Center for Contemporary Management.

References


Fayyad, U., Uthurusamy, R. (eds.): AAAI Workshop on Knowledge Discovery in Databases, Seattle, Washington, DC, USA (1994)


Abstract. Fuzzy set theory and fuzzy logic provide a mathematical model for approximate reasoning, which mimics human reasoning and can therefore be very powerful. Also in the field of image processing many applications can be found. It was in the late nineties that Etienne’s research group started to work on image processing, and took off on a journey that lasts until today. This chapter is a “crew member’s log” of this journey, with Etienne as our captain, and provides an overview of our work.

1 Getting Ready to Sail Off

The first papers on fuzzy techniques in image processing, originating from our research group at Ghent University, were published in the period 1995-1997. They were co-authored by Bernard De Baets, Etienne Kerre, and foreign colleagues, and dealt with the construction of fuzzy mathematical morphology [1, 2, 3, 4, 5]. Inspired by the potential that fuzzy set theory had to offer in this field, Etienne submitted and obtained a large university research project. This marked the beginning of a 12-year period, starting in 1998 until today, in which fuzzy set theory and fuzzy logic were used for and applied in image processing problems.

Looking back at these years, one can distinguish 3 different research “axes”: fuzzy mathematical morphology (which is mainly developed from a theoretical point of view), filters for noise reduction in images (which is very practical...
oriented), and similarity measures for image comparison (which holds the balance
between theory and practical applicability). It is remarkable to see how each of these
fields has evolved in this 12-year period.

We invite the reader to embark on the summary of this 12-year journey. It is
our tribute to Etienne Kerre (who was promotor of our PhD’s, which all dealt with
image processing), and our way to say “thank you” for the wonderful years we had
in his research group. As such, this chapter is intended to serve as a kind of overview
chapter, in which we focus on the work w.r.t. image processing that has been done
in Etienne’s research group.

2 Crew Member’s Logs of a 12-Year Journey

2.1 Journey 1: Fuzzy Mathematical Morphology

In the nineties several models for mathematical morphology based on fuzzy set
theory were developed. In general, three types of models could be distinguished:
models based on a fuzzification of the underlying logical operators (e.g., the model
of Bloch & Maître and the model of De Baets), models based on a fuzzification of
the concept of inclusion (e.g., the models of Zadeh, Sinha & Dougherty, Kitainik,
Bandler & Kohout), and other models (e.g., the model based on the Minkowski
addition, the model based on fuzzy integrals, and the model based on the erosions
of Di Gesu). This wide variety of new models illustrated the fact that fuzzy set
theory indeed had potential in this field, but on the other hand this diversity also
made it difficult to keep the general overview and raised the question on how to
choose which model is best fitted for a specific problem.

One of our first contributions to fuzzy mathematical morphology was an exten-
sive study of all the different existing models. This was a very interesting work, as
it turned out that several connections between the different models could be estab-
lished [6,7,8]. The model developed by De Baets, based on morphological dilations
and erosions that are defined using general conjunctors and implicators, is the most
general one and includes every of the other mentioned models. The results of our
study made it possible to gain insight in the variety of fuzzy morphological mod-
els, and also implied that future research on fuzzy mathematical morphology should
focus on the most general model, since properties of other models could easily be
derived from the properties of the general model.

Just as an example to familiarize the reader with some typical expressions, we
recall the definition of the two basic binary morphological operations, namely the
dilation and erosion:

\[
D(A, B) = \{y \in X | T_y(B) \cap A \neq \emptyset\},
\]

\[
E(A, B) = \{y \in X | T_y(B) \subseteq A\},
\]

where \(A, B\) are crisp subsets of a universe \(X\) (\(A\) represents the binary image and \(B\) is
the so-called structuring element which can be chosen by the user), and \(T_y(B)\) is the
Fuzzy Techniques in Image Processing at Ghent University 199

translation of $B$ over the vector $y$. In the most general fuzzy morphological model, these definitions can be fuzzified as follows:

$$D_C(A, B)(y) = \sup_{x \in T_Y(d_B) \cap d_A} C(B(x - y), A(x)),$$

$$E_I(A, B)(y) = \inf_{x \in T_Y(d_B)} I(B(x - y), A(x)),$$

where $A, B$ are fuzzy sets in a universe $X$ ($A$ represents a grayscale image and $B$ a grayscale structuring element), $C$ is a conjunctor on $[0,1]$ (to model the intersection) and $I$ is an implicator on $[0,1]$ (to model the concept of subsetness). Conjunctors and implicators are fuzzifications of the Boolean conjunction and the binary implication, respectively, and are defined as follows: a binary operator $C$ on $[0,1]$ is a conjunctor if it is an increasing mapping (i.e. it has increasing partial mappings) that satisfies $C(0,0) = C(0,1) = C(1,0) = 0$ and $C(1,1) = 1$, and a binary operator $I$ on $[0,1]$ is an implicator on $[0,1]$ if it is a hybrid monotonous mapping (i.e. it has decreasing first and increasing second partial mappings) that satisfies $I(0,0) = I(0,1) = I(1,1) = 1$ and $I(1,0) = 0$. $d_A$ and $d_B$ in the above formulas represents the support of the fuzzy sets $A$ and $B$, respectively, e.g. $d_A = \{x \in X | A(x) > 0\}$.

Regarding the investigation of properties, we explicitly paid attention to the so-called decomposition and construction properties [9, 10]. Decomposition properties reveal the relation between fuzzy morphological operations and their $\alpha$-cuts, while construction properties show which fuzzy morphological operations can be constructed from their binary counterparts. For a brief period we also looked into the fuzzification of the concept of adjunction, which plays a crucial role in the lattice based approach to mathematical morphology [11]. We vividly remember our nice cooperation with Henk Heijmans (which included a short stay of Mike at Henk’s institution), one of the most well-known researchers in this field. Unfortunately, Henk got very ill and until today is no longer able to perform academic work. Several years later, we published a second paper on fuzzy adjunctions, inspired by the work of our colleagues Yun Shi and Da Ruan on fuzzy implications [12].

In the same period that we were working on fuzzy mathematical morphology, our colleague Martine De Cock was working on linguistic modifiers and our Polish colleague Anna Radzikowska was working on fuzzy rough set theory. Looking at each other’s work we noticed that there was a common ground: the operators that constitute the basis of each of the three different theories (mathematical morphology, linguistic modifiers, fuzzy rough sets) showed a lot of similarity. A deeper investigation of this similarity led to the conclusion that the three theories are actually special cases of a more general theoretical framework based on images under fuzzy relations. This theoretical generality was further investigated, which lead to several publications [13, 14, 15, 16, 17]. The practical result was that we could investigate the general framework, and translate the properties of this general framework to each specific application area.

A few years later a completely different contribution to the field of mathematical morphology was realized by constructing morphological models for colour images. Indeed, it must be noted that all of the above was developed in the context of
grayscale images. The extension to colour images, whether this extension is realized for morphological operators such as in this case or for noise reduction filters (see further in this chapter), certainly is not straightforward. From an abstract point of view a colour image can be regarded as a triplet of grayscale images (e.g., the well-known RGB colour model), but it is not sufficient to treat these three components separately to obtain a good extension. On the contrary, such a simplistic approach does not take into account any correlations between the different components and will most likely introduce artefacts. Therefore, we looked at several alternative extensions to colour morphology [18, 19]. We also applied the developed colour morphology in the context of image magnification [20].

More recently, we shifted our attention to a third track in the world of (fuzzy) mathematical morphology. Until now, we only used fuzzy set theory and fuzzy logic as a tool. The application of fuzzy set theory in mathematical morphology was inspired by the fact that a grayscale image and a fuzzy set are mathematically modelled in exactly the same way, namely as mappings from some universe to the unit interval \([0,1] \). The interpretation of course might be completely different, but techniques from one field could easily be applied to the other field. In this case, this is best illustrated by the fact that fuzzy logical operators were tools to construct alternative morphological operators. The reader familiar with the field will know that this has nothing to do with modeling uncertainty or imprecision. However, since 2008 we established a very nice and fruitful cooperation with our colleague Peter Sussner. The shift in the work is quite clear: we do not only apply fuzzy techniques as a tool, but we also apply concepts related to fuzzy set theory to actually model uncertainty that comes along with image capture [21, 22, 23, 24]. In that way the picture gets more complete: fuzzy set theory not only serves as a tool, but also as a model in mathematical morphology.

The basic idea is to model grayscale images no longer as mappings from a universe to the unit interval \([0,1] \), where the value between 0 and 1 represents the measured gray value, but as mappings from a universe to the set of closed subintervals of the unit interval. This means that a pixel is no longer mapped onto a single value, but to a closed interval of likely values. This actually corresponds to the real-life situation, in which one is not always sure of the measured gray value: such values are not only rounded up or down (for computational storage and processing), but might also slightly shift in value depending on the recording circumstances (e.g., a take of a scene with a sunny sky versus the take of the same scene with a cloudy sky). The uncertainty that comes along with this image capture can be modelled by closed intervals, and in a natural way coincides with the already developed framework of interval-valued fuzzy sets. An illustration of this interval-valued approach is shown in Figures 1 and 2. We can also make the link with intuitionistic fuzzy set theory (since interval-valued fuzzy set theory and intuitionistic fuzzy set theory are equivalent), and even attach a related interpretation to the corresponding image model [25, 26].

Once the new image model was established, the construction and deeper investigation of the corresponding morphological model was the next challenge. We constructed a general interval-valued framework [27], but also looked at very specific
Fig. 1 Different captures of the cameraman image: top = take with cloudy sky, middle = take with sunny sky, bottom = take with distortion. This example illustrates that the capture circumstances can cause uncertainty regarding the real pixel values. Also, all recorded values and positions are approximations of the real situation due to technical limitations.
models (such as in [28, 29]) or at further generalizations based on \( L \)-fuzzy set theory [30], where \( L \) is a complete lattice. Just as we did for the “classical” fuzzy morphological operators, we also investigated decomposition and construction properties for interval-valued morphological operators [31, 32, 33].

The above overview of our activities in the field of mathematical morphology clearly illustrates the different “waves” that we experienced: starting from investigating the wide variety of fuzzy-inspired morphological models, we moved forward to colour extensions, and are currently on the track of interval-valued fuzzy morphology to also take into account the uncertainty that is involved with image capturing and modeling. Our future work will focus on theoretical items on the one hand (e.g., what is the relation between interval-valued morphological operators followed by defuzzification and defuzzification of the interval-valued image followed by classical morphological operators?), and on practical issues on the other hand (e.g., how can the interval-valued approach contribute to real-world edge detection applications?). One thing is sure: this journey is far from over . . .

2.2 Journey 2: Similarity Measures

There is a big need for objective similarity measures in image processing. Two different categories of applications demonstrate this. First, similarity measures are crucial in image retrieval applications: given a source image, similarity measures are required to retrieve those images from a database that are most similar to the source image. This also can have internet-based applications, where you feed a source image and try to retrieve similar images from the internet in order to obtain information regarding the content of the source image. Second, similarity measures are also important for more theoretical purposes. For example, if one develops a filter for noise reduction (such as our research group did in the same period as we were looking at
similarity measures) it is very important that this filter is compared to other existing filters; otherwise one cannot demonstrate the usefulness and/or added value of that new filter. Popular similarity measures in that regard include the MSE (Mean Square Error) and the related PSNR (Peak-Signal-to-Noise-Ratio). However, experience learns that the numerical results of these measures do not always correspond to the visual evaluation of the images by humans. In other words: two images that have the lowest MSE (or highest PSNR) are not necessarily the two most similar images from a visual point of view. Of course, this observation is quite important when reading publications in which algorithms are compared to each other only using such measures. Either a visual evaluation should be included, or either new similarity measures that are more closely related to human evaluation have to be constructed.

It is the latter which inspired our research group to start working on similarity measures. We were familiar with similarity measures in the context of fuzzy set theory, which were developed to quantify the ressemblance between different fuzzy sets. Since fuzzy sets and grayscale images are modelled in the same way (namely as mappings from a universe to the unit interval [0,1]) the first step in our investigation was quite straightforward: to what extent can we apply existing similarity measures for fuzzy sets in image processing applications? For that purpose, more than 40 different measures were evaluated w.r.t. relevant image processing properties such as reflexivity (an image should be similar to itself to the degree 1), symmetry (the order in which the images are processed may not influence the result of the similarity measure), reaction to noise (a noisy image should have a large similarity w.r.t. the original image), reaction to enlightening or darkening (an enlightened or darkened image should have a large similarity w.r.t. the original image) and reaction to binary images (we expect similarity values between 0 and 1, and not only 0 or 1). From this investigation it followed that just a handful of the existing measures are appropriate for comparison of images \[34, 35\]. One of these measures is the following, which we refer to as $M_6$:

$$M_6(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{\sum_{(i,j) \in X} \min(A(i,j), B(i,j))}{\sum_{(i,j) \in X} \max(A(i,j), B(i,j))},$$

where $A$ and $B$ are grayscale images (fuzzy sets in $X$) and $A(i,j)$ and $B(i,j)$ represent the gray value of the pixel at position $(i, j)$ in the respective images.

Because the perceptual behaviour of the remaining similarity measures was not always convincing (mainly due to the fact that the measures are pixel-based, and do not take into account any other relevant information), we constructed new similarity measures ourselves. First of all, we applied the similarity measures to partitioned images in order to construct neighbourhood-based similarity measures, with a more robust behaviour. However, simply applying the similarity measures to corresponding image parts did not yield satisfactory results either, so we needed to look further to other techniques in order to improve the behaviour of the similarity measures. In that way, we constructed neighbourhood-based similarity measures which also
incorporate some characteristics of the human visual system, such as contrast sensitivity, by taking into account the homogeneity in the considered neighbourhoods \[36, 37, 38, 39, 40\]. The results were much better compared to the results of the pixel-based similarity measures.

In a next step we investigated whether similarity measures from fuzzy set theory could be applied in a meaningful way to the histograms of images \[41, 42, 43, 44\]. It is indeed meaningful to compare two histograms in the framework of fuzzy set theory, because the histogram of an image can be transformed to a fuzzy set in the universe of gray levels by dividing the values of the histogram in every gray level by the maximum amount of pixels with the same gray value (resulting in so-called normalized histograms). In this way the most typical gray value gets membership degree 1 in the fuzzy set associated with the histogram and every other less typical gray value gets a smaller membership degree. Consequently, a normalized histogram is in accordance with the intuitive idea behind a fuzzy set: the most typical element in the universe gets membership degree 1 and all other less typical elements belong to the fuzzy set to a less extent. We found 15 similarity measures to be useful for direct application to (normalized) histograms. Furthermore, 22 similarity measures turned out to be appropriate for application to (normalized) ordered histograms; ordered histograms are obtained by placing the least occurring gray value in the first position of the histogram and by ordering the remaining frequencies in increasing order. Also in this approach, experimental results showed a better perceptual performance than the pixel-based similarity measures and the classical MSE or PSNR.

In order to confirm the applicability of neighbourhood-based and histogram-based similarity measures, a large psycho-visual experiment was conducted using a Multi-Dimensional Scaling approach (MDS) towards analyzing and modeling image quality variations \[45\]. This work was done in close cooperation with our colleagues from the Telecommunications and Information Department of the Faculty of Engineering Sciences. In total, 35 individuals evaluated a whole range of images (original images and slightly distorted ones) and these psycho-visual results were confronted with the output of our best neighbourhood- and histogram-based similarity measures. The conclusion of this large experiment was that the fuzzy similarity measures outperform the classical MSE and PSNR, i.e., they are better in accordance with the human visual evaluation. Another conclusion was that the neighborhood-based similarity measures perform better than the histogram-based similarity measures, which is probably due to the fact that histogram similarity measures do not incorporate the spatial properties of the different gray values, but only consider the frequency of occurrence.

Since both neighbourhood-based and histogram-based similarity measures take into account different additional information from the images, we also constructed similarity measures which combine both approaches. This resulted in a new series of similarity measures with a strong global perceptual performance \[46, 47\].

Next, our research shifted to the construction of similarity measures for colour images. In a first approach, good similarity measures for grayscale images were applied to the different colour components of colour images (either in the RGB, HSV or Lab colour space) and then the results for each pixel were averaged. The
results were promising, but we anticipated that a better approach would be to treat the colours as vectors, instead of just applying grayscale measures to each colour component. To achieve this, we introduced a new vector ordering in the RGB colour space which allowed us to extend the similarity measures that were constructed for grayscale images to similarity measures for colour images (such an ordering is required to be able to select minima and maxima). Experiments confirmed that the results were better than in the component-based approach [48,49,50,51,52].

Finally, we also constructed a colour image retrieval application. Colour image retrieval is becoming more and more important (also in the context of video), and so is the quest for automated and reliable retrieval systems. The novelty of our approach was the use of a fuzzy partition of the HSI colour space and the use of one of our similarity measures for histogram comparison (namely the one based on the above mentioned similarity measure $M_6$). The resulting retrieval system has the advantage that the images do not have to be characterized in advance using several features, and it is quite flexible since the database images are not required to have the same dimensions [53,54]. We also situated our approach to image retrieval in a more larger framework in [55]. An example can be found in Figure 3 which shows an input image and the 9 most similar retrieved results from a large database of over 500 natural images of animals, flowers, buildings, cars, texture images, and so on. The results are quite good: the three most similar retrieved images are flowers in the same colour as the one in the query image; the other retrieved images do not contain flowers but have a very similar layout.

2.3 Journey 3: Fuzzy Filters for Noise Removal

Noise reduction filters are without doubt the most practical results from our research in image processing. During the past 12 years, a nice trajectory can be detected: starting with filters for grayscale images, the next step was filters for colour images, and finally filters for video sequences (both grayscale and colour). We focussed on two very common noise types: impulse noise (where a fraction of the pixel values is replaced by either fixed noise values or random noise values) and gaussian noise (additve noise). Table 1 presents a bird’s eye view on the filters that were developed in our group.

It all started with the so-called GOA-filter, named after the university project that enabled us to do this research and which resulted from a very nice cooperation with our colleagues from the Faculty of Engineering Sciences, in particular Dimitri Van De Ville, Wilfried Philips and Ignace Lemahieu [56,57]. The filter is designed for the removal of gaussian noise in grayscale images, and uses fuzzy rules to detect the degree to which the gradient in a certain direction is small (the idea is that a small gradient is caused by noise, while a large gradient is caused by image structure). Fuzzy rules are also applied to calculate the correction term that is used for the denoising; the contribution of neighbouring pixels depends on their gradient values. The results of this filter were very good, and confirmed the usefulness of fuzzy logic for the construction of noise reduction filters. The main advantage of fuzzy filters is
that they allow us to work and to reason with linguistic information, just as experts do (approximate reasoning); see the scheme in Figure 4. In order to confirm these good results, we carried out extensive comparative studies of existing classical and fuzzy filters, both for impulse noise and gaussian noise [58, 59, 60, 61, 62, 63].

A second filter for the reduction of gaussian noise from grayscale images was presented a few years later [64]. This FuzzyShrink-filter can be seen as a fuzzy variant of an existing probabilistic shrinkage method, and was developed in the wavelet domain. The filter outperformed fuzzy non-wavelet methods and was comparable with other recent but more complex wavelet methods.

After the succesful GOA-filter for gaussian noise, we developed a filter for the removal of fixed impulse noise in grayscale images. This filter was called the Fuzzy Impulse noise Detection and Reduction Method, or FIDRM for short [65]. The filter
Table 1 A summary of the different fuzzy filters for noise reduction that were developed in our research group.

<table>
<thead>
<tr>
<th></th>
<th>Still gray</th>
<th>Still colour</th>
<th>Video gray</th>
<th>Video colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>fixed impulse</td>
<td>FIDRM</td>
<td>FIDRMC</td>
<td>HFMRC</td>
<td>HFC OWA</td>
</tr>
<tr>
<td>random impulse</td>
<td>FRINR</td>
<td>HFC</td>
<td>FRINV-G</td>
<td>FRINV-C</td>
</tr>
<tr>
<td>gaussian</td>
<td>GOA</td>
<td>FCG</td>
<td>FMDAF</td>
<td>FMDAF-RGB</td>
</tr>
</tbody>
</table>

Fig. 4 Fuzzy filters not only use numerical to filter out the noise in images, but can also work with linguistic information. Furthermore, fuzzy logic allows us to reason with this linguistic information and enables us to better approximate human reasoning.

followed a similar approach as the GOA filter, as it used gradient values to detect and remove the noise. Again, extensive experiments confirmed the state-of-the-art results of the filter [66, 67]. The filter could easily be extended to colour images by applying the filter on each of the corrupted colour bands separately. The results for colour images were relatively good [68], but the disadvantage of this approach is of course that correlations between colour band were neglected and small colour artefacts were introduced.

The latter observation inspired us to construct other filters, specifically to remove impulse noise from colour images, and led to the FIDRMC and HFMRC filters. The FIDRMC filter consists of two separated steps: the detection phase and the filtering phase. The detection phase is applied separately to each colour component, where fuzzy rules are used to determine whether a pixel pigment is corrupted with impulse noise or not. After the detection phase the filter only focuses on those pixel pigments which have a non-zero membership degree in the fuzzy set “impulse noise”. In the filtering phase we also take into account the colour information of a certain neighbourhood around a given central pixel [69, 70]. The HFMRC filter follows a different approach and uses the histograms of the colour component differences to detect and filter the fixed impulse noise [71]. The HFMRC filter was later upgraded to the more complex HFC filter [72] that could also tackle randomly valued impulse noise in colour images. Previously, our FRINR filter already achieved the goal of removing randomly valued impulse noise in grayscale images [73, 74]. The detection
phase of the FRINR filter consists out of two units that are both used to define corrupted impulse noise pixels. The first unit investigates the neighbourhood around a pixel to conclude whether the pixel can be considered as impulse noise or not, while the second unit uses fuzzy gradient values to determine the degree to which a pixel can be considered as impulse noise and the degree to which a pixel can be considered as noise free.

Regarding the removal of gaussian noise from colour images, we developed the FCG filter \[75\]. In contrast to most of the other existing methods, the first subfilter of the FCG filter distinguishes between local variations due to noise and local variations due to image structures (such as edges) by using the colour component distances instead of component differences. The second subfilter is used as a complementary filter which especially preserves differences between the colour components. A few years later, resulting from an idea of and our cooperation with Arya Basu, the OWA filter was developed for noise reduction in colour images corrupted by either impulse noise or a combination of salt & pepper noise and gaussian noise \[76\]. This OWA filter is an improvement of classical bilateral filtering, achieved by using Ordered Weighted Averaging operators.

With all of the above, we delivered fuzzy filters for fixed impulse noise, random impulse noise and gaussian noise, both for grayscale and colour images. The next step was the design of noise reduction filters for video. It is quite clear that video makes things much more complicated: the introduction of a temporal component implies that good filtering can only be achieved if the temporal aspect is taken well into account, i.e., if motion information is incorporated in the filter design. Filters for colour video are even more complex, because the temporal information has to be combined with the colour information. Nevertheless, a nice series of video filters were constructed.

The first achievement was the FMDAF or Fuzzy Motion and Detail Adaptive Filter, which was constructed in close cooperation with our colleagues from the Faculty of Engineering Sciences. We also developed a recursive version of the filter (RFMDAF), and a non-recursive and recursive version in the wavelet domain (WFDMAF and WRFMDAF). The method can be seen as a fuzzy variant of a multiple class video denoising method that automatically adapts to detail and motion. Experimental results show that the FMDAF filter efficiently removes gaussian noise from grayscale image sequences and outperforms other state-of-the-art filters of comparable complexity \[77\] \[78\].

Afterwards we extended this filter to colour video sequences in three different ways. First the FMDAF-RGB filter was an extension of the FMDAF filter in such a way that colours in the RGB colour space were treated as vectors; no separate filtering of colour bands occurred \[79\]. Next we developed the FMDAF-CR filter \[80\]. Although the filtering here is performed in each colour band separately in the RGB colour space, the fuzzy rules also require information from the other colour bands such that correlations between colour bands are taken into account. We also constructed a third extension, called the FMDAF-YUV filter \[80\], in which the WFMDF filter (in the wavelet domain) is applied to the Y-component in the YUV colour space and where the chrominance components U and V are averaged.
over a small window. All these filters have a very competitive numerical and visual performance.

It is quite remarkable that most video filters that can be found in literature are designed for sequences corrupted by gaussian noise, and much less video filters exist for the removal of impulse noise. Existing 2D filters for impulse noise can of course be extended to video by applying the 2D filter to each frame of the video sequence, but temporal inconsistencies will arise when motion is neglected in the filtering

Fig. 5 20-th frame of the “Deadline” video sequence: top = original frame, middle = noisy frame (25% random impulse noise), bottom = result of our FRINV-G filter.
process. Existing 3D filters on the other hand often suffer from detail loss because too many pixels are filtered or, vice versa, too many pixels are left unfiltered in order to preserve the image details. Quite recently, this observation encouraged us to develop a fuzzy filter for the removal of random impulse noise in both grayscale and colour video. The grayscale version of our filter (which we here refer to as the FRINV-G filter [81]) consists out of three different successive filtering steps and a fourth refinement step. In each filtering step, only the pixels that are detected as being noisy (a detection that is achieved by fuzzy rules) are filtered and to exploit the temporal information detected pixels are filtered in a motion compensated way. Experimental results clearly demonstrate that the FRINV-G filter outperforms other state-of-the-art filters, both numerically and visually. An illustrative example is shown in Figure 5. Finally, also a vector-based colour extension was constructed; here we refer to this colour video filter as the FRINV-C filter [82].

3 The Ship Has Sailed

Mathematical morphology, similarity measures, noise reduction filters: these were the three main tracks of the research on image processing in Etienne Kerre’s research group. The short overviews above, together with the extensive literature list, show that each of these tracks was a very intensive and interesting journey. During all of these journeys we also had the opportunity to put some effort in the international promotion of soft computing in image processing in general. In 2002 we established the so-called “SCIP Working Group”, where SCIP is an acronym for Soft Computing in Image Processing. The goal of this informal working group, of which the organisation is mainly internet-based, was and is to promote communication and cooperation between colleagues that are active in this exciting field. We firmly believe that it is very important that all efforts are undertaken to intensify this communication and cooperation: it leads to a cross-fertilization of ideas and new cooperations, and avoids that several groups are working on the same problems without knowing this from each other. The SCIP Working Group, which has around 160 members from over 40 countries, mainly tries to achieve these goals by regularly organizing special sessions on soft computing in image processing at international conferences. If the opportunity is there, we will also take the initiative to edit a book or journal issue that brings together interesting contributions from colleagues from all over the world. In that regard we certainly mention the three books in the series “Studies in Fuzziness and Soft Computing” [83] [84] [85] and the special issue of the international journal “Soft Computing - A Fusion of Foundations, Methodologies and Applications” [86] and of the “International Journal of Approximate Reasoning” [87].

Although each of the five authors of this chapter – which all five obtained their PhD under Etienne’s guidance – were involved in one or more of the three discussed research tracks and in the SCIP Working Group, there always was one unifying factor: our promotor Etienne was there to support and guide us! Despite the fact that most of the authors have left the academic world, the work is still not done.
Currently, we are involved in a European Marie-Curie project on soft computing techniques in medical image processing. Our goal in this project is to use our expertise on similarity measures and noise reduction filters in order to apply these techniques on medical images. Two new Phd students are already working on the project! Etienne will not be around full-time, but we are all convinced that he will be part of this new journey . . .

References


The Need to Use Fuzzy Extensions in Fuzzy Thresholding Algorithms

Humberto Bustince, Miguel Pagola, Edurne Barrenechea, and Javier Fernández

Abstract. In this chapter we present some recent applications of fuzzy extensions in image segmentation. First we review some basic concepts of Interval-valued fuzzy sets, which is the extension that is mainly used. Next we present the fuzzy thresholding algorithm and we discuss its main problem that leads to use the extensions of fuzzy sets. In section 3 we review some methods recently published that use extensions of fuzzy sets in image thresholding. Finally we show some experimental results comparing the classical fuzzy thresholding algorithm against the algorithms based on extensions of fuzzy sets.

1 Introduction to Extensions of Fuzzy Sets: Interval Valued Fuzzy Sets

From the beginning it was clear that fuzzy set theory [29] was an extraordinary tool for representing human knowledge. Nevertheless, Zadeh himself established (see [30]) that sometimes, in decision-making processes, knowledge is better represented by means of some generalizations of fuzzy sets. A key problem of representing the knowledge by means of fuzzy sets is to choose the membership function which best represents such knowledge.

Sometimes, it is appropriate to represent the membership degree of each element to the fuzzy set by means of an interval. From these considerations arises the extension of fuzzy sets called theory of interval-valued fuzzy sets (IVFSs), that is, fuzzy sets such that the membership degree of each element of the fuzzy set is given by a closed subinterval of the interval \([0, 1]\). Hence, not only vagueness (lack of sharp class boundaries), but also a feature of uncertainty (lack of information) can be addressed intuitively.

These sets were first introduced in the 1970s. In May 1975 Sambuc (see [24]) presented in his doctoral thesis the concept of an interval-valued fuzzy set named...

The concept of a type 2 fuzzy set was introduced by Zadeh [30] as a generalization of an ordinary fuzzy set. Type 2 fuzzy sets are characterized by a fuzzy membership function, that is, the membership value for each element of the set is itself a fuzzy set in $[0, 1]$.

Formally, given the referential set $U$, a type 2 fuzzy set is defined as an object $\overline{A}$ which has the following form:

$$\overline{A} = \{(u, x, \mu_u(x)) | u \in U, x \in [0, 1]\},$$

where $x \in [0, 1]$ is the primary membership degree of $u$ and $\mu_u(x)$ is the secondary membership level, specific to a given pair $(u, x)$.

One year later, Grattan-Guinness [17] established a definition of an interval-valued membership function. In that decade interval-valued fuzzy sets appeared in the literature in various guises and it was not until the 1980s, that the importance of these sets, as well as their name, was definitely established.

A particular case of a type 2 fuzzy set is an interval type 2 fuzzy set (see [20]–[21]). An interval type 2 fuzzy set $\overline{A}$ in $U$ is defined by

$$\overline{A} = \{(u, A(u), \mu_u(x)) | u \in U, A(u) \in L([0, 1])\},$$

**Fig. 1** Fuzzy membership function.

**Fig. 2** Interval valued fuzzy membership function. Lower bound and upper bound
where \( A(u) \) is a closed subinterval of \([0,1]\), and the function \( \mu_u(x) \) represents the fuzzy set associated with the element \( u \in U \) obtained when \( x \) covers the interval \([0,1]\); \( \mu_u(x) \) is given in the following way:

\[
\mu_u(x) = \begin{cases} 
  a & \text{if } A(u) \leq x \leq \overline{A}(u) \\
  0 & \text{otherwise}
\end{cases},
\]

where \( 0 \leq a \leq 1 \). It turns out that an interval type 2 fuzzy set is the same as an IVFS if we take \( a = 1 \).

Another important extension of fuzzy set theory is the theory of Atanassov’s intuitionistic fuzzy sets (A-IFSs). Atanassov’s intuitionistic fuzzy sets assign to each element of the universe not only a membership degree, but also a nonmembership degree, which is less than or equal to 1 minus the membership degree.

An Atanassov’s intuitionistic fuzzy set (A-IFS) on \( U \) is a set

\[
\hat{A} = \{(u, \mu_{\hat{A}}(u), \nu_{\hat{A}}(u)) | u \in U\},
\]

where \( \mu_{\hat{A}}(u) \in [0,1] \) denotes the membership degree and \( \nu_{\hat{A}}(u) \in [0,1] \) the nonmembership degree of \( u \) in \( \hat{A} \) and where, for all \( u \in U \), \( \mu_{\hat{A}}(u) + \nu_{\hat{A}}(u) \leq 1 \).

In \([1]\) Atanassov established that every Atanassov intuitionistic fuzzy set \( \hat{A} \) on \( U \) can be represented by an interval-valued fuzzy set \( A \) given by

\[
A : U \rightarrow L([0,1]) \quad u \rightarrow [\mu_{\hat{A}}(u), 1 - \nu_{\hat{A}}(u)], \quad \text{for all } u \in U.
\]

Using this representation, Atanassov proposed in 1983 that Atanassov’s intuitionistic fuzzy set theory was equivalent to the theory of interval-valued fuzzy sets. This equivalence was proven in 2003 by Deschrijver and Kerre \([12]\). Therefore, from a mathematical point of view, the results that we obtain for IVFSs are easily adaptable to A-IFSs and vice versa. Nevertheless, we need to point out that, conceptually, the two types of sets are totally different. This is made clear when applications of these sets are constructed (see \([28]\)).

In 1993, Gau and Buehrer introduced the concept of vague sets \([16]\). Later, in 1996, it was proven that vague sets are in fact A-IFSs \([5]\).

A compilation of the sets that are equivalent (from a mathematical point of view) to interval-valued fuzzy sets can be found in \([13]\). Two conclusions are drawn from this study:

1.- Interval-valued fuzzy sets are equivalent to A-IFSs (and therefore vague sets), to grey sets (see \([11]\)) and to \( L \)-fuzzy set in Goguen’s sense with respect to a special lattice \( L([0,1]) \).

2.- IVFSs are a particular case of probabilistic sets (see \([14]\)), of soft sets (see \([3]\)), of Atanassov’s interval-valued intuitionistic fuzzy sets and evidently of Type 2 fuzzy sets.
2 Interval Valued Fuzzy Sets and Image Segmentation

Image segmentation is the first step in image analysis and pattern recognition. It is a critical and essential component of image analysis and/or pattern recognition system and is one of the most difficult tasks in image processing, that can determine the quality of the final result of the system.

The goal of image segmentation is the partition of an image in different areas or regions.

Definition 1. Segmentation is grouping pixels into regions such that
1. \( \bigcup_{i=1}^{k} P_i = \text{Entire image} \) \( \{P_i\} \) is an exhaustive partitioning).
2. \( P_i \cap P_j = 0, i \neq j \) \( \{P_i\} \) is an exclusive partitioning).
3. Each region \( P_i \) satisfies a predicate; that is, all points of the partition have some common property.
4. Pixels belonging to adjacent regions, when taken jointly, do not satisfy the predicate.

There exist three different approaches using fuzzy methods:

- Histogram thresholding.
- Feature space clustering.
- Rule based systems.

2.1 Thresholding

One of the earlier papers of fuzzy thresholding is [23] in 1983. The idea behind the fuzzy thresholding is to first transfer the selected image feature into a fuzzy subset by means of a proper membership function and then select and optimize a global or local fuzzy measure to attain the goal of image segmentation.

In [23] the authors used the S-function to fuzzify the image. They minimize the entropy in such a way that the final segmented image is the one which has less doubtful pixels.

Said membership function represents the brightness set within the image. The basic idea of using this membership function is that, if we take the value of a parameter as the threshold value, the dark pixels should have low membership degrees, and on the contrary, brighter pixels should have high membership degrees. The pixels with membership function near 0.5 should be the ones that are not clearly classified. Therefore the set with less entropy is the set with less amount of pixels with uncertain membership (around 0.5).

Since the first works that used S-functions up to now the problem is: finding the membership function that will accurately represent the membership degree of each pixel to each region of the image (in the case of two regions: the background and the object).

If the expert knows the exact membership degree of each pixel to each region of the image, the problem is solved. This does not normally happen, that is, there are
The Need to Use Fuzzy Extensions in Fuzzy Thresholding Algorithms

pixels which the expert does not know whether they belong to the background or to the object.

Depending on the type of image we are working with, it is necessary to use one membership function or another. That is, in order to obtain a good segmentation the user or expert must choose the membership function that will best represent that image.

Therefore, in any segmentation algorithm there is always a degree of ignorance in the choice of the membership function made by the user. It is established (see [15, 18]) that the membership function must always try to represent:

- The similarity of the pixels within the regions.
- The connectivity of the pixels within the regions.
- The difference between the pixels of different regions.

We do not know how to correctly express these three properties in the membership function. For this reason there is always an uncertainty associated with thresholding algorithms. According to Klir and Wierman [19] “Uncertainty has a pivotal role in any efforts to maximize the usefulness of systems models ... uncertainty becomes very valuable when considered in connection to the other characteristics of systems models: a slight increase in uncertainty may often significantly reduce the complexity and, at the same time, increase the credibility of the model. Uncertainty is thus an important commodity in the modeling business, a commodity which can be traded for gains in the other essential characteristics of models”.

In this sense Zadeh [30], Mendel [21], and other authors establish that the main characteristic of the extensions of fuzzy sets is their capacity to model uncertainty in the membership function selection.

In this work we are going to present some recent methods to add uncertainty by means of extensions of fuzzy sets into the thresholding process. All of these techniques are based on the classical fuzzy thresholding algorithm.

3 Fuzzy Algorithm of Image Thresholding and Its Extensions

The basic structure of the algorithms for the calculation of the threshold of an image that use fuzzy techniques (see [7, 15, 18, 22]) is composed of the following steps:

(a) Assign $L$ fuzzy sets $\tilde{Q}_t$ to each image $Q$. Each one associated to a level of intensity $t$, ($t = 0, 1, \cdots, L - 1$), of the grayscale $L$ used.

(b) Calculate the entropy of each one of the $L$ fuzzy sets $\tilde{Q}_t$ associated to $Q$.

(c) Take as best threshold $t$ the level of gray associated to the fuzzy set with lowest entropy.

Alg. Fuzzy

Membership functions are normally defined by the expert on the basis of his/her knowledge. The membership functions used in [15, 18] represent how similar the
intensity of each pixel is to the mean of the intensities of the object or to the mean of the intensities of the background. This type of membership functions are the ones that provide the best results (see [22, 25]). By defining the membership functions in this way, the set with lowest entropy is the set that contains the greatest number of pixels around the mean of the intensities of the background and the mean of the intensities of the object.

In these conditions given an image \( Q \) and an intensity threshold \( t \) set, we have proposed a general method [7] of constructing the membership function of each intensity to the set \( \tilde{Q}_t \) in the following way:

\[
\mu_{\tilde{Q}_t}(q) = \begin{cases} 
F(REF(q, m_b(t))) & \text{if } q \leq t \\
F(REF(q, m_o(t))) & \text{if } q > t 
\end{cases}
\]  

(1)

Where \( m_b(t) \) and \( m_o(t) \) are given by the following expressions:

\[
m_b(t) = \frac{\sum_{q=0}^{t}qh(q)}{\sum_{q=0}^{t}h(q)}
\]  

(2)

\[
m_o(t) = \frac{\sum_{q=t+1}^{L-1}qh(q)}{\sum_{q=t+1}^{L-1}h(q)}.
\]  

(3)

\( h(q) \) being the number of pixels of the image with intensity \( q \).

Depending on the functions \( F \) and \( \text{REF} \) chosen we can construct most of the expressions of membership functions used in the fuzzy thresholding in the literature [18, 25].

Once the manner of constructing the membership function has been chosen, we calculate the \( L \) fuzzy sets associated with each \( t = 0, 1, \ldots, 255 \). In the following step of the algorithm we calculate the entropy of each one of those sets, which is why we now present the definition of entropy and some expressions.

### 3.1 Examples of the Fuzzy Algorithm Using Different Membership Functions

Step (c) of the fuzzy algorithm establishes that once the entropy is calculated for all the fuzzy sets, we must take as threshold the value of \( t \) associated with the fuzzy set \( \tilde{Q}_t \) with lowest entropy. Taking into account that we have constructed the membership function using \( \text{REFs} \), then the membership value equal to 1 indicates that the value of the pixel is very close to the value of the mean of the intensities of the background or to the value of the mean of the intensities of the object. As a result the expert is sure of the membership, either to the background or to the object, of most pixels.

In the following example we prove experimentally that the solution to the algorithm depends on the membership function selected.
Table 1 Different membership functions

<table>
<thead>
<tr>
<th></th>
<th>1) $F(x) = 0.5(1+x)$</th>
<th>2) $F(x) = 0.5(1+x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$REF(x,y) = 1 -</td>
<td>x - y</td>
</tr>
<tr>
<td>3)</td>
<td>$F(x) = 0.5(1+x)$</td>
<td>$F(x) = 0.5(1+x)$</td>
</tr>
<tr>
<td></td>
<td>$REF(x,y) = 1 -</td>
<td>x - y</td>
</tr>
<tr>
<td>5)</td>
<td>$F(x) = \frac{1}{(2-x)}$</td>
<td>$F(x) = \frac{1}{(2-x)}$</td>
</tr>
<tr>
<td></td>
<td>$REF(x,y) = 1 -</td>
<td>x - y</td>
</tr>
<tr>
<td>7)</td>
<td>$F(x) = \frac{1}{(2-x)}$</td>
<td>$F(x) = \frac{1}{(2-x)}$</td>
</tr>
<tr>
<td></td>
<td>$REF(x,y) = 1 -</td>
<td>x - y</td>
</tr>
</tbody>
</table>

Example 1. The membership functions are constructed from the eq. (1). On Table 1 we show the different functions $F$ and $REF$ taken. To calculate the entropy of fuzzy sets we use the eq. (4) with

$$E(\tilde{A}) = \sum_{x \in X} 1 - |1 - 2\mu_{\tilde{A}}(x)|^2$$

(4)

Fig. 3 Different solutions to the fuzzy algorithm with different membership functions
In figure 3 we show the binary images obtained. From the visual analysis of these images we deduce that the results of the fuzzy algorithm depend on the membership function chosen (in this case functions (5) and (6) would be the most suitable solutions).

3.2 Generalization of the Fuzzy Algorithm Using IVFS

In 2005 Tizhoosh [26] presented a paper that uses Interval-valued fuzzy sets in image thresholding (we must point out that he tries to use type 2 fuzzy sets, however in the paper he only uses Interval-valued fuzzy sets [9]). His study is based on the modification of the classical fuzzy algorithm of Huang and Wang [18], so that he applies an $\alpha$ factor as interval generator to the membership function. Starting from a membership function, Tizhoosh obtains an interval valued fuzzy set that “contains” different membership functions and is useful for finding the threshold of an image. In this manner we can handle the inaccuracy of the selection of the membership function and obtain good results even though the best selection has not been made. That is why to each pixel we are going to assign two values:

- We will obtain the first one by changing the membership function chosen (constructed from the restricted equivalence functions), which we interpret as a pessimistic measure of the membership degree of said pixel to the set that represents the image.
- The second value, also obtained from the membership function chosen, we interpret as an optimistic measure of the membership degree of said pixel to the set that represents the image.

The structure of the algorithm is the following:

(a) Assign $L$ fuzzy sets $\tilde{Q}_t$ to each image $Q$. Each one associated to a level of intensity $t$, ($t = 0, 1, \cdots, L - 1$), of the grayscale $L$ used.

(b) Construct an Interval valued fuzzy set $Q_t$ with the parameters $\alpha$ and $\beta$ from each one of the $L$ fuzzy sets $\tilde{Q}_t$ associated to $Q$. It must reflect the uncertainty in the choice of membership function.

(c) Calculate the Interval valued entropy of each one of the $L$ interval valued fuzzy sets $Q_t$ associated to $Q$.

(d) Take as best threshold $t$ the level of gray associated to the fuzzy set with lowest IV entropy.

Alg. IVFS

So in this version the main fact is to construct the IVFSs in such a way that it represents the original shape of the membership function given by the expert and
also brings the uncertainty of the election. The other different aspect is that we choose as threshold the $t$ associated with the set with minimum IV-entropy.

A generalization of the Tizhoosh algorithm can be made using two different parameters to estimate the uncertainty of the membership function, so we can construct an interval-valued fuzzy set in the following way:

$$A_{\alpha, \beta} = \{(x, [\mu^\alpha_A(x) \mu^\beta_A(x)]) | x \in X \} \in IVFSs(X).$$

(5)

The verification that $A_{\alpha, \beta} \in IVFSs(X)$ is evident: $0 \leq \mu^\alpha_A(x) \leq \mu^\beta_A(x) \leq 1$. The parameters $\alpha$ and $\beta$ must be related with the ignorance of the expert in the membership function selection. Tizhoosh used only one parameter $\alpha$, so that the IVFS set he constructed is:

$$A_{\alpha} = \{(x, [\mu^\alpha_A(x) \mu^\frac{1}{\alpha}_A(x)]) | x \in X \} \in IVFSs(X).$$

(6)

In [4] Bustince presents the general way of constructing $E_F$. From all of the expressions presented we are going to use the simplest, Sambuc’s indetermination index [24]. That is, for the set constructed in the eq. (5) the entropy is given by:

$$E_F(Q_t) = \frac{1}{N \times M} \sum_{q=0}^{L-1} h(q)(\mu^{1/\beta}_{Q_t}(q) - \mu^\alpha_{Q_t}(q)).$$

(7)

The justification for choosing the set with lowest $E_F$ is the following:

if $E_F$ is close to zero, then we can consider that generally, for all $q$ this holds:

$$\mu^{1/\beta}_{Q_t}(q) \approx \mu^\alpha_{Q_t}(q).$$

By the way of constructing the membership functions we know that:

$$0.5 \leq \mu_{Q_t}(q) \leq 1$$

for each $q \in \{0, \cdots, L - 1\};$

therefore $\mu^{1/\beta}_{Q_t}(q) \approx 1$, (as long as we take $\alpha \neq 1$ and $\beta \neq 1$). In these conditions, by the construction of $Q_t$ we have

$$\mu_{Q_t}(q) = \begin{cases} F(REF(q,m_b(t))) & \text{if } q \leq t \\ F(REF(q,m_o(t))) & \text{if } q > t \end{cases} \approx 1$$

By the condition: $F(x) = 1$ if and only if $x = 1$, we have that $\mu_{Q_t}(q)$ is close to one if $REF(q,m_b(t))$ when $q \leq t$ (or $REF(q,m_o(t))$ when $q > t$) is close to one; Bearing in mind the definition of function of $REF$ we have that this happens when $q$ is very close to $m_b(t)$ or to $m_o(t)$ depending on the case. Furthermore we know that (see [15, 18]) that if $q$ is very close to $m_b(t)$ or to $m_o(t)$, then we are in the best situation for taking the threshold. Therefore if we take the set with lowest value of $E_F$, by the reasoning above, then that IVFS is the one that best represents the image and the one with the minimum in that representation [9].
3.3 Modification of the Fuzzy Algorithm by Means of A-IFSs

Vlachos and Sergiadis [28] also propose a modification of Tizhoosh’s algorithm, using Atanassov’s intuitionistic fuzzy sets (see [1]). Their basis are membership functions similar to Huang’s but, instead of minimizing the entropy, the algorithm minimizes the divergence with set $\hat{I}$ (see [10]). The structure of the intuitionistic algorithm is the same as Tizhoosh’s. The construction of the intuitionistic fuzzy sets is done in the following way:

$$
\mu_{A}(g,t) = \lambda \mu_{A}(g,t)
$$

$$
\nu_{A}(g,t) = (1 - \lambda \mu_{A}(g,t))^\lambda
$$

With $\lambda \in [0,1]$, being $A$ an intuitionistic fuzzy set, and the divergence:

$$
D_{IFS}(A, \hat{I}, t) = \sum_{g=0}^{L-1} h_{A}(g) \left( \frac{2 \mu_{A}(g,t)}{1 + \mu_{A}(g,t)} + v_{A}(g,t) \ln 2 + \frac{2}{1 + \mu_{A}(g,t)} \right)
$$

In this case the uncertainty is represented by the hesitation index:

$$
\pi_{A}(g,t) = 1 - \mu_{A}(g,t) - \nu_{A}(g,t)
$$

3.4 Construction of the IVFSs Using Ignorance Functions

In this case the classical fuzzy thresholding algorithm is modified due to the user should pick two functions, one to represent the background and another one to represent the object. We have chosen this representation since, in this way, the expert is able to get a better representation of the pixels for which he is not sure of their membership to the object or the background. In figure 4 we show two membership functions, one to represent the background and the other to represent the object.

As we have already said in the previous paragraph, we are going to represent the images by means of two different fuzzy sets. For this reason, in our proposed algorithm we introduce the concept of ignorance function $G_{u}$. Such functions are a way to represent the user’s ignorance for choosing the two membership functions.

Fig. 4 Two different membership functions to represent the background and the object with t=150
used to represent the image (object and background). Therefore, in our algorithm we will associate to each pixel three numerical values:

- A value for representing its membership to the background, which we will interpret as the expert’s knowledge of the membership of the pixel to the background.
- A value for representing its belongingness to the object, which we will interpret as the expert’s knowledge of the membership of the pixel to the object.
- A value for representing the expert’s ignorance of the membership of the pixels to the background or to the object. This ignorance hinders the expert from making an exact construction of the membership functions described in the first two items and therefore it also hinders the proper construction of step (a) of the fuzzy algorithm. The lower the value of ignorance is, the better the membership function chosen to represent the membership of that pixel to the background and the one chosen to represent the membership to the object will be. Evidently, there will be pixels of the image for which the expert will know exactly their membership to the background or to the object but there will also be pixels for which the expert is not able to determine if they belong to the background or to the object.

Under these conditions, if the value of the function of ignorance ($G_u$) for a certain pixel is zero, it means that the expert is positively sure about the belongingness of the pixel to the background or to the object. However, if the expert does not know at all whether the pixel belongs to the background or to the object he must represent its membership to both with the value 0.5, and under these conditions we can say that the expert has total ignorance regarding the membership of the pixel to the background and the membership of the same pixel to the object.

In [8] a methodology is proposed to construct ignorance functions. In the following example we show an ignorance function that can be constructed by means of the t-norm minimum.

**Example 2.** Using the t-norm $T = T_M$

$$G_u(x, y) = \begin{cases} 
2 \cdot \min(1 - x, 1 - y) & \text{if } \min(1 - x, 1 - y) \leq 0.5 \\
\frac{1}{2 \min(1 - x, 1 - y)} & \text{otherwise}
\end{cases}$$

is a continuous ignorance function.

Also Ignorance functions can be constructed from functions like the geometric mean that are not t-norms (please see [8]).

**Example 3.** If we take $\phi(x) = \sqrt{x}$ for all $x \in [0, 1]$ we recover the following ignorance function:

$$G_u(x, y) = \begin{cases} 
2 \sqrt{(1 - x) \cdot (1 - y)} & \text{if } (1 - x) \cdot (1 - y) \leq 0.25 \\
\frac{1}{2 \sqrt{(1 - x) \cdot (1 - y)}} & \text{otherwise}
\end{cases}$$

In the following proposition we can see how we can construct IVFS by means of the ignorance function.
Proposition 1. Let $\tilde{Q}_B$ and $\tilde{Q}_O$ be the fuzzy sets associated to the background and the object of the image built by an expert. Let $G_u$ be the ignorance function associated to the construction of the said fuzzy sets. Under these conditions, if we define

$$\Phi : \mathcal{F}_{s}(U) \times \mathcal{F}_{s}(U) \longrightarrow \mathcal{F}_{s}(U)$$

given by

$$\Phi(\tilde{Q}_B, \tilde{Q}_O) = \{(u, [M_L(u), M_U(u)]) | u \in U\}$$

such that

$$[M_L(u), M_U(u)] = [G_u(0.5, 0.5) - G_u(\mu_{\tilde{Q}_B}(u), \mu_{\tilde{Q}_O}(u)), G_u(0.5, 0.5)],$$

then for each membership $[M_L(u), M_U(u)]$ defined by $\Phi$ the following equation holds:

$$W([M_L(u), M_U(u)]) = G_u(\mu_{\tilde{Q}_B}(u), \mu_{\tilde{Q}_O}(u))$$

If the IV entropy $E_F(Q)$ that represents the total influence of the ignorance in the construction of fuzzy sets $Q_{Bt}$ and $Q_{Ot}$ is such that it verifies:

$$E_F \rightarrow 0$$

then bearing in mind the properties required to $\mathcal{M}$ we have

$$G_u(\mu_{Q_{Bt}}(q), \mu_{Q_{Ot}}(q)) \rightarrow 0 \text{ for all } q \in \{0, \cdots, L - 1\}.$$ 

By the property "$G_u(x, y) = 0$ if and only if $x = 1$ or $y = 1$" that is required to the functions $G_u$, we have $\mu_{Q_{Bt}}(q) \rightarrow 1$ or $\mu_{Q_{Ot}}(q) \rightarrow 1$, therefore $q$ is very close to $m_B(t)$ ($m_O(t)$), that is, we are certain that the pixels with intensity $q$ belong to the background (object). This is due to the following reasoning:

1. If $\mu_{Q_{Bt}}(q) \rightarrow 1$, then $REF(\frac{q}{L-1}, \frac{m_B(t)}{L-1}) \rightarrow 1$, therefore $q \approx m_B(t)$. In this case the pixels with intensity $q$ are such that their intensity is very close to the average intensity of the pixels that represent the background. This fact enables us to assure that the pixel in question belongs to the background.

2. If $\mu_{Q_{Ot}}(q) \rightarrow 1$, then $REF(\frac{q}{L-1}, \frac{m_O(t)}{L-1}) \rightarrow 1$, therefore $q \approx m_O(t)$. In this case the pixels with intensity $q$ are such that their intensity is very close to the average intensity of the pixels that represent the object. This fact enables us to assure that the pixel in question belongs to the object.

4 Examples and Experimental Results

In this section we are going to present some results obtained for the different algorithms presented before. Recalling the objective of these algorithms is that the IVFS algorithm produces solutions comparable in quality to those of the fuzzy algorithm when an adequate membership function has been used and surpasses the fuzzy algorithm when the membership function chosen does not represent the image properly. In this last case we deduce that there exists great uncertainty in the election of the membership function. In the first experiment we calculate the threshold using the fuzzy algorithm using 8 different membership functions (table 1) and then we calculate the threshold with the IVFS algorithm with $\alpha = 2$ and $\beta = 2$ for the set
of natural images depicted in figure 5. On Table 2 we present the percentage of wrongly classified pixels in each case. We observe again that for a single image the fuzzy algorithm returns different solutions depending on the membership function we select. The lower the number of wrongly classified pixels, the better the solution and therefore the better is the membership function for that image. In the last row the mean of the values of the ten images is shown, this way we can choose the best membership function for this set of images. In this case the best membership function is $\mu_7$.

On the next table, Table 3 we show the values obtained with the IVFS algorithm using in every case the same 8 membership functions and the parameters $\alpha = 2$ and $\beta = 2$ in order to construct IVFSs.

The values in which the mean of error of the IVFS algorithm is less than the fuzzy case are highlighted in bold type. We observe in this experiment that the IVFS algorithm, with the exception of a membership function $\mu_7$ (which is the best), is

### Table 2 Results of the Fuzzy algorithm

<table>
<thead>
<tr>
<th>Image</th>
<th>$\mu_1$</th>
<th>$\mu_2$</th>
<th>$\mu_3$</th>
<th>$\mu_4$</th>
<th>$\mu_5$</th>
<th>$\mu_6$</th>
<th>$\mu_7$</th>
<th>$\mu_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.5763</td>
<td>2.8441</td>
<td>5.7520</td>
<td>2.6247</td>
<td>2.8441</td>
<td>3.3262</td>
<td>3.9639</td>
<td>2.5916</td>
</tr>
<tr>
<td>1</td>
<td>7.2416</td>
<td>7.2135</td>
<td>7.7773</td>
<td>7.2569</td>
<td>7.2365</td>
<td>7.5502</td>
<td>7.7543</td>
<td>7.2238</td>
</tr>
<tr>
<td>2</td>
<td>1.6334</td>
<td>1.6234</td>
<td>1.9852</td>
<td>1.6611</td>
<td>1.6334</td>
<td>1.7314</td>
<td>2.0028</td>
<td>1.6234</td>
</tr>
<tr>
<td>8</td>
<td>39.8728</td>
<td>42.2905</td>
<td>30.8456</td>
<td>38.6832</td>
<td>42.2905</td>
<td>44.4647</td>
<td>30.8456</td>
<td>41.0650</td>
</tr>
</tbody>
</table>
Table 3 Results of the IVFS algorithm (a)

<table>
<thead>
<tr>
<th>Image</th>
<th>$\mu_1$</th>
<th>$\mu_2$</th>
<th>$\mu_3$</th>
<th>$\mu_4$</th>
<th>$\mu_5$</th>
<th>$\mu_6$</th>
<th>$\mu_7$</th>
<th>$\mu_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.5661</td>
<td>2.7217</td>
<td>5.7520</td>
<td>2.6247</td>
<td>3.1961</td>
<td>3.9639</td>
<td>2.5763</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.2391</td>
<td>7.2340</td>
<td>7.7773</td>
<td>7.2620</td>
<td>7.1906</td>
<td>7.3666</td>
<td>7.7543</td>
<td>7.2416</td>
</tr>
<tr>
<td>2</td>
<td>1.6284</td>
<td>1.6108</td>
<td>1.9852</td>
<td>1.6962</td>
<td>1.6058</td>
<td>1.6636</td>
<td>2.0028</td>
<td>1.6334</td>
</tr>
<tr>
<td>3</td>
<td>4.2722</td>
<td>4.4088</td>
<td>4.1717</td>
<td>4.2722</td>
<td>4.4088</td>
<td>4.1717</td>
<td>4.2722</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>39.8728</td>
<td>42.2905</td>
<td>29.9969</td>
<td>38.6832</td>
<td>42.2905</td>
<td>43.4468</td>
<td>30.8456</td>
<td>39.8728</td>
</tr>
</tbody>
</table>

better than the fuzzy case. That is, if we choose a wrong membership function the IVFS algorithm provides a better result.

Similar results can be obtained for the A-IFS algorithm. Please see [28] for other results in pattern recognition.

Case of IVFS constructed by ignorance functions

Finally we evaluate the performance of the algorithm that uses ignorance functions in prostate ultrasound images. For each image the expert has provided a point at the center of the prostate making object detection/extraction easier. The image is then filtered via median filter in $5 \times 5$ neighborhoods. Further, selective contrast enhancement as described in [27] has been applied to increase the image quality.

We must take into account that, since ultrasound images depend on the particular settings of the machine is very important that our algorithm gives good solutions even if some membership functions that do not represent accurately the background and the prostate are chosen.

In this experiment we take 8 different membership functions. $\mu(q) = 1 - |q - m_z(t)|^\lambda$ with $\lambda = 0.1, 0.3, 0.8, 1, 1.3, 2, 6, 15$, where $z$ could denote the object or the background. For all the images of each set, we execute the Generalized fuzzy

![Fig. 6 Set of Ultrasound prostate images with their ideal segmentation made by an expert radiologist.](image-url)
algorithm with two membership functions using the classical fuzzy algorithm with all the possible combinations of the membership functions; i.e., $8^2 = 64$. Then we execute the ignorance based algorithm using the geometric mean based ignorance function for all the cases, so we’ll obtain 64 different solutions.

To interpret the results of this experiment we study the graphic in Fig. 7. This graphic is obtained in the following way:

1. We arrange all the cases from the smallest to the biggest percentage of badly classified pixels (error) in the solution of the fuzzy algorithm.
2. For each pair of membership functions, we calculate the error obtained with the ignorance based algorithm.

The crosses represents the error we get with the ignorance based algorithm, and the dotted line, the error we get with the fuzzy algorithm. (If the dots are under the cross, it means that for that pair of membership functions the error of the ignorance based algorithm is smaller than the error of the fuzzy algorithm). Observe that:

1. For the pairs of membership functions such that the fuzzy algorithm solution is good (small error), the ignorance based algorithm does not provide better results.
2. If the error we get with the fuzzy algorithm begins to be high (i.e., if we have used bad-chosen membership functions), then the result of the ignorance based algorithm improves the other algorithm’s result.

We observe that in the fuzzy algorithm case there are around 40% of the cases on which the membership functions do not represent correctly areas of the image and a very high error is obtained. However, the IVFS algorithm gets a good solution for almost all the cases.

![Fig. 7](image-url) Percentage of error in all cases of first prostate ultrasound image in Fig. 7 with ignorance function using the geometric mean.

If we analyze the table of mean errors for all the cases, we see that in the case of real ultrasound images ignorance based algorithm has a total mean error less than the fuzzy algorithm.
5 Conclusions and Future Research

A key problem of the fuzzy thresholding algorithm is the accurate election of the membership function. In this chapter we have presented methods which use extensions fuzzy sets as a tool to represent the uncertainty presented in the election of the correct membership function. The experimental results shown in section 4 allow us to conclude that algorithms that adds the uncertainty presented in the problem by means of extensions of fuzzy sets provides similar or better results that the ones obtained by fuzzy classical techniques.

As a future research, form the extensions point of view, must focus on the use of general type 2 fuzzy sets, but the computation time of general type usually is not tractable. Also, finding applications or image representations in which membership and non-membership can be generated independently in order to use all of the power of A-IFSs, should be researched.

References

24. Sambuc, R.: Function ϕ-Flous, Application a l’aide au Diagnostic en Pathologie Thyroidienne, These de Doctorat en Medicine, University of Marseille (1975)
Author Index

Atanassov, Krassimir T. 19
Barrenechea, Edurne 219
Bustince, Humberto 219

Chen, Guoqing 171
Cornelis, Chris 57

De Cock, Martine 237
Deschrijver, Glad 57
De Witte, Valérie 197
Djouadi, Yassine 127
Dubois, Didier 127

Fernández, Javier 219
Kacprzyk, Janusz 267
Komorníková, Magda 3

Makarytska, Nataliya 237
Mélange, Tom 197
Mesiar, Radko 3

Nachtegael, Mike 197
Novák, Vilém 39
Nowacka, Katarzyna 267
Pagola, Miguel 219
Perfilieva, Irina 39
Prade, Henri 127
Radzikowska, Anna Maria 107
Ruan, Da 83
Schockaert, Steven 237
Schulte, Stefan 197
Shi, Yun 83
Van der Weken, Dietrich 197
Van Gasse, Bart 57, 83
Wang, Xuzhu 149
Wei, Qiang 171
Wu, Caiping 149
Wu, Xia 149
Xiong, Yixi 171
Zadrożny, Sławomir 267